A Multiresolution Fuzzy Clustering of Images

Kalyani Mali, Sushmita Mitra and Tinku Acharya

Abstract—**A wavelet based multiresolution fuzzy clustering of digital images is described. We have applied wavelet transforms to decompose the images into different subbands in multiple resolutions, thereby eliminating noise in each low frequency subband. Then a partitive fuzzy clustering technique is proposed to be applied in multiple resolutions of the transformed image, in order to find coarser to finer clustering of the images in the image database. The relevance of the proposed multiresolution technique to clustering and retrieval of images in the compressed domain is also highlighted.**

A combination of texture, shape, topology and fuzzy geometric features, that is invariant to orientation, scale and object deformation, are extracted from the low frequency subbands of the transformed image in coarser to finer resolutions. Partitive fuzzy clustering is performed, to group these images according to similarity at different levels of resolution. We use clustering validity indices to determine the optimal number of image categories. The images in each cluster are graded based on their distances from the corresponding centroid. The use of wavelet transform eliminates the need for any other preprocessing for noise removal, to make it amenable for feature extraction. The extracted features serve as the signature of the images, in terms of their content. Their use in content based image retrieval is also demonstrated.

Index Terms— **Image mining, wavelets, soft computing, content based image retrieval, fuzzy clustering, image compression.**

I. Introduction

A CLUSTER is a collection of data objects which are similar to one another within the same cluster but dissimilar to the objects in other clusters. The problem is to group N patterns into c desired clusters with high *intra-class* similarity and low *inter-class* similarity by optimizing an objective function. In the c-means algorithm [1], each cluster is represented by the center of gravity of the cluster. This need not essentially correspond to an object of the given pattern set. In the c-medoids algorithm [2], on the other hand, each cluster is represented by one of the representative objects in the cluster located near the center. Clustering validity indices, like Davies-Bouldin and Dunn's, may be used to determine the optimal number of clusters.

Soft computing methodologies, involving fuzzy sets, neural networks, genetic algorithms, rough sets, wavelets, and their hybridizations, have recently been used to solve data mining problems [3]. They strive to provide approximate solutions at low cost, thereby speeding up the process. Fuzzy sets, which constitute the oldest component of soft computing, are suitable for handling the issues related to understandability of patterns, incomplete or noisy data, mixed media information, and human interaction, and can provide approximate solutions faster. Incorporation of the fuzzy membership concept, in fuzzy c-means and fuzzy c-medoids clustering, enables appropriate modeling of real life overlapping data.

Clustering is often required at hierarchical levels of coarseness, grouping spatial objects at different levels of accuracy. This gives rise to the concept of multiresolution representation of an image, as depicted in Fig. 1. Wavelets [4] are found to be very useful in appropriately modeling such situations because of the nonstationary property of the image signals formed around the edges and correlation amongst the image pixels. The role of wavelets in different aspects of data mining is gaining significant importance, and it has become a very powerful signal processing tool in different application areas such as image processing, compression, image indexing and retrieval, digital libraries, image clustering and databases [5], [6].

Most of the activities in mining image data have been in the search and retrieval of images based on the analysis of similarity of a query image or its feature(s) with the entries in the image database. In *Content Based Image Retrieval* (CBIR) systems, the images are searched and retrieved by extracting suitable features based on the visual content of the images [3], [7]. CBIR has increasingly become a growing area of study towards the successful development of image mining techniques [3].

In this article we describe a wavelet based multiresolution fuzzy clustering scheme, with potential for CBIR on real life digital images. Wavelet transforms are applied to decompose the images into different subbands in multiple resolutions, thereby eliminating noise in each low frequency subband. The clustering now needs to be done on low-frequency subbands of the images of smaller sizes, from coarser to finer resolutions. This also enables us to eliminate the need for any other preprocessing of the raw images. A combination of texture, shape, topology and fuzzy geometric features are extracted directly from this compressed image. Partitive fuzzy clustering is performed at different levels of resolution and the performance compared with that of the nonfuzzy version. We use clustering validity indices to determine the optimal number of image categories. The images in each cluster are graded based on their distances from the corresponding centroid. The features are next used for CBIR in the wavelet transformed domain. The images retrieved, based on content, are always found to lie in the same partition as the query image.

The rest of the paper is organized as follows. Section II

introduces concepts from wavelet transform. This is followed, in Section III, by a brief description of the different features extracted, involving textural, shape, topology, anf fuzzy geometric properties. The clustering methodologies and validity indices are presented in Section IV. The actual implementation of clustering and CBIR on digital images are provided in Section V. The relevance of the proposed multiresolution technique to clustering and retrieval of images in the compressed domain is highlighted in Section VI. Finally, Section VII concludes the article.

II. Wavelet Transform

Wavelet transform is a signal processing technique that decomposes a signal or image into different frequency subbands at number of levels and multiple resolutions. In every level of decomposition, the high-frequency subband captures the discontinuities in the signals – for example, the edge information in an image. The low-frequency subband is a subsampled version of the original image, with similar statistical and spatial properties as the original signal. As a result, the low-frequency subband can be further decomposed into higher levels of resolution, and it helps in representing spatial objects in different coarser levels of accuracy in multiresolution subbands. The wavelet transform is typically represented as a pair of high-pass and low-pass filters, with many wavelet basis functions being available in literature [4].

Wavelets are functions generated from one single function called the *mother wavelet* by dilations (scalings) and translations (shifts) in time (frequency) domain. If the mother wavelet is denoted by $\psi(t)$, the other wavelets $\psi^{a,b}(t)$ for $a > 0$ and a real number b can be represented as

$$\psi^{a,b}(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), \tag{1}$$

where a and b represent the parameters for dilations and translations in the time domain. The parameter a causes contraction in time domain when $a < 1$ and expansion when $a > 1$.

In this article we focus on the Haar and Daubechies wavelet functions [8]. Haar wavelets use the unit-height, unit-width scaling function

$$\Phi(t) = \begin{cases} 1 & 0 \le t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{2}$$

The Haar wavelet function is expressed as

$$\Psi(t) = \begin{cases} 1 & 0 \le t < 0.5, \\ -1 & 0.5 \le t < 1, \\ 0 & \text{otherwise.} \end{cases} \tag{3}$$

In case of Daubechies wavelets, we have

$$\Phi(t) = 0, \quad \text{if} \quad t \le 0 \quad \text{or} \quad t \ge 3.$$

Therefore, $\Phi(0) = \Phi(3) = 0$, $\Phi(1) = \frac{1+\sqrt{3}}{2}$, and $\Phi(2) = \frac{1-\sqrt{3}}{2}$. The scaling function Φ satisfies the recurrence relation

$$\Phi(t) = h_0\Phi(2t) + h_1\Phi(2t-1) + h_2\Phi(2t-2) + h_3\Phi(2t-3), \tag{4}$$

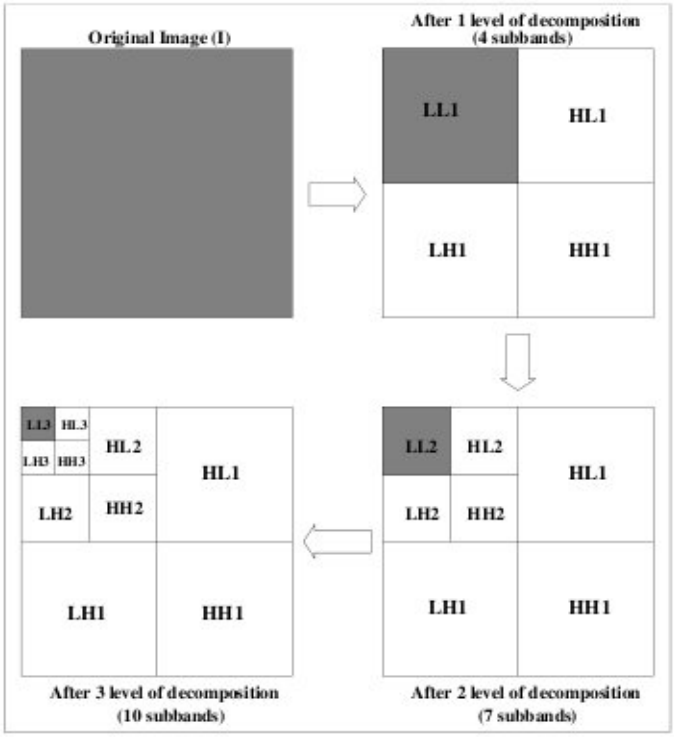

Fig. 1. Three level multiresolution wavelet decomposition of an image.

for $0 < t < 3$, where $h_0 = \frac{1+\sqrt{3}}{4}$, $h_1 = \frac{3+\sqrt{3}}{4}$, $h_2 = \frac{3-\sqrt{3}}{4}$, and $h_3 = \frac{1-\sqrt{3}}{4}$. The associated wavelet function is denoted as

$$\Psi(t) = -h_0\Phi(2t-1) + h_1\Phi(2t) - h_2\Phi(2t+1) + h_3\Phi(2t+2), \tag{5}$$

where $-1 < t < 2$.

In Fig. 1, we show an example of hierarchical wavelet decomposition of an image into ten subbands after three levels of decomposition [3]. After the first level of decomposition, the original image is decomposed into four subbands $LL1$, $HL1$, $LH1$, and $HH1$. The $LL1$ subband is the low-frequency subband which can be considered as a 2:1 subsampled (horizontally and vertically) version of the original image I, and its statistical characteristic is similar to the original image as shown by the shaded regions in Fig. 1. Here $HL1$, $LH1$, and $HH1$ are called the high-frequency subbands, where $HL1$ and $LH1$ correspond to the horizontal and vertical high frequencies, respectively. $HH1$ constitutes the high frequencies that are not in either horizontal or vertical orientations. Each of these spatially oriented (horizontal, vertical, or diagonal) subbands mostly contain information of local discontinuities in the image.

Since the low-frequency subband $LL1$ has similar spatial and statistical characteristics as the original image, it can be further decomposed into four subbands $LL2$, $HL2$, $LH2$, and $HH2$. Continuing the same method for decomposition in $LL2$, the original image is decomposed into 10 subbands $LL3$, $HL3$, $LH3$, $HH3$, $HL2$, $LH2$, $HH2$, $HL1$, $LH1$, and $HH1$ after three levels of pyramidal multiresolution subband decomposition, as shown in Fig. 1. The same procedure can continue to further decompose $LL3$ into higher levels.

The low-pass filter, inherent in the transform, helps to remove the noise. It is possible to identify clusters at different levels of accuracy (i.e., fine or coarse) by using multiresolution partitioning. This implies that there are less clusters at a coarser resolution.

III. Feature Extraction

In this section we describe the input features used for clustering the digital image. These include texture, fuzzy geometry, moment invariants and Euler vector. They are useful in characterizing images, and can be used as a signature of image content. Hence these features have promising application in CBIR.

A. Texture

Texture is one of the important features used in identifying objects or regions of interest in an image [9]. It is often described as a set of statistical measures of the spatial distribution of gray levels in an image. This scheme has been found to provide a powerful input feature representation for various recognition problems.

The textural features are computed from a set of angular nearest neighbor gray-tone spatially dependent matrices. The contextual texture information is specified by the matrix of relative frequencies $P(i,j)$ with which two neighboring resolution cells, having gray levels i and j and separated by a distance δ, occur in the image.

The unnormalized frequencies are defined by the elements $P(i,j,\delta;\theta)$ of a set of cooccurrence matrices, where θ is 0^o, 45^o, 90^o and 135^o for horizontal, right-diagonal, vertical, and left-diagonal neighbor pairs, respectively. For nearest neighbor pairs, we have $\delta = 1$. Then the number of neighboring resolution cell pairs R is given by

$$R = \begin{cases} 2N_y(N_x-1) & \text{for } \theta = 0^o \\ 2N_x(N_y-1) & \text{for } \theta = 90^o \\ 2(N_x-1)(N_y-1) & \text{otherwise.} \end{cases} \quad (6)$$

Here N_x and N_y refer to the number of pixels along the horizontal and vertical directions of the digital image.

The angular second moment (A) gives a measure of the homogeneity of the texture and is defined as

$$A = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\left(\frac{P(i,j)}{R}\right)^2. \quad (7)$$

Note that R, from eqn. (6), is used as the normalizing constant. The parameter N_g indicates the maximum gray-tone value present in the image.

The measure H is the inverse difference moment, and also provides an indication of the amount of homogeneity in the texture. It is expressed as

$$H = \sum_{n=0}^{N_g-1}\frac{1}{1+n^2}\left\{\sum_{|i-j|=n}\frac{P(i,j)}{R}\right\}. \quad (8)$$

Entropy E is a measure of variability in the image, and is zero for a constant image. It is defined as

$$E = -\sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\frac{P(i,j)}{R}\log\frac{P(i,j)}{R}. \quad (9)$$

Contrast C is a difference moment of the matrix P, measuring the amount of local variation present in the image. It is expressed as

$$Contrast = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-j)^2\frac{P(i,j)}{R}. \quad (10)$$

Note that the notation θ was omitted in eqns. (7)–(10) to avoid clutter. Each measure is calculated four times, corresponding to each of the four directional cooccurrence matrices.

B. Shape

Shape can roughly be defined as the description of an object minus its position, orientation and size. Therefore, shape features should be invariant to *translation*, *rotation*, and *scale*, when the arrangement of the objects in the image are not known in advance.

Let $I(x,y)$ denote an image in the two-dimensional spatial domain. *Geometric Moment* [10] of order $p+q$ is denoted as

$$m_{p,q} = \sum_x\sum_y x^p y^q I(x,y), \quad (11)$$

for $p,\ q = 0, 1, 2, \ldots$. The central moments are expressed as

$$\mu_{pq} = \sum_x\sum_y (x-x_c)^p (y-y_c)^q I(x,y), \quad (12)$$

where $x_c = \frac{m_{1,0}}{m_{0,0}}$, $y_c = \frac{m_{0,1}}{m_{0,0}}$, and (x_c, y_c) is called the center of the region or object. Hence the *Central Moments* can be computed as

$$\left.\begin{array}{rcl} \mu_{0,0} &=& m_{0,0} \\ \mu_{1,0} &=& 0 \\ \mu_{0,1} &=& 0 \\ \mu_{2,0} &=& m_{2,0} - x_c m_{1,0} \\ \mu_{0,2} &=& m_{0,2} - y_c m_{0,1} \\ \mu_{1,1} &=& m_{1,1} - y_c m_{1,0}. \end{array}\right\}$$

The *normalized central moments*, denoted by $\eta_{p,q}$, are defined as $\eta_{p,q} = \frac{\mu_{p,q}}{\mu_{0,0}^{\gamma}}$, where $\gamma = \frac{p+q+2}{2}$ for $p+q = 2, 3, \ldots$. Two of the seven *transformation invariant moments*, used here, are expressed as follows.

$$\left.\begin{array}{rcl} \phi_1 &=& (\eta_{2,0}+\eta_{0,2}) \\ \phi_2 &=& (\eta_{2,0}-\eta_{0,2})^2 + 4\eta_{1,1}^2. \end{array}\right\}$$

It is to be noted, however, that the shape features need accurate segmentation of images to detect the object or region boundaries.

C. Topology

A digital image can be represented by one or more topological properties [8], which typically represent the geometric shape of an image. These are also invariant to stretching, deformation, rotation, scaling or translation.

One topological property of a digital image is known as *Euler number*. Although typically computed in a binary image, it can be extended to characterize gray-tone images by defining a vector of Euler numbers of the binary planes of the gray-tone image. This has been called the *Euler Vector* [11]. The *Euler number* is defined as the difference between number of *connected components* and number of *holes* in a binary image. Hence if an image has C connected components and H number of holes, the *Euler number* E of the image can be defined as

$$E = C - H. \tag{13}$$

Intensity value of each pixel in an 8-bit gray-tone image can be represented by an 8-bit binary Euler vector b_i, $i = 0, 1, \cdots, 7$, that is, $(b_7, b_6, b_5, b_4, b_3, b_2, b_1, b_0)$, where $b_i \in \{0, 1\}$. The ith bit plane is formed with b_i's from all the pixels in the gray-tone image. We retain the first two most significant bit planes corresponding to (b_7, b_6), because they contain most of the information of the image.

D. Fuzzy geometry

A fuzzy subset of a set S is a mapping μ from S into $[0, 1]$. For any $x \in S$, $\mu(x)$ is known as the degree of membership of x in μ. Let $\mu(I)$ denote a fuzzy representation of an $N_x \times N_y$ gray level image I, *i.e.*, a mapping μ from $I \in \{1, \ldots, N_g\}$ into $[0, 1]$ representing a fuzzy subset of I, where N_g is the maximum gray value. For convenience, we shall use μ only to denote $\mu(I)$ in this section. The *compactness* ($Comp$) [12] of a fuzzy set μ having area $a(\mu)$ and perimeter $p(\mu)$ is defined as

$$Comp(\mu) = \frac{a(\mu)}{p^2(\mu)}. \tag{14}$$

Physically, compactness means the fraction of maximum area (that can be encircled by the perimeter) actually occupied by the fuzzy region/concept represented by μ. Here the *area* is defined as $a(\mu) = \sum \mu$, the summation being considered over a region outside which $\mu = 0$. The *perimeter* of an image is expressed as $p(\mu) = \sum_{i,j} |\mu(i) - \mu(j)|$, where $\mu(i)$ and $\mu(j)$ are the membership values of two adjacent pixels.

IV. Clustering

In this section we describe the partitive algorithms used for clustering [3], *viz.*, the c-means, c-medoids (PAM), fuzzy c-means (FCM) and fuzzy c-medoids. Two clustering validity indices, employed to determine the optimal number of clusters, are also mentioned.

A. c-means algorithm

The algorithm proceeds by partitioning N objects into c nonempty subsets. During each partition, the centroids or means of the clusters are computed. The main steps of the c-means algorithm [1] are as follows:

- Assign initial means $\mathbf{m}_i$ (also called centroids).
- Assign each data object (pattern point) $\mathbf{X}_k$ to the cluster U_i for the closest mean.
- Compute new mean for each cluster using

$$\mathbf{m}_i = \frac{\sum_{\mathbf{X}_k \in U_i} \mathbf{X}_k}{|c_i|}, \tag{15}$$

  where $|c_i|$ is the number of objects in cluster U_i.
- Iterate until criterion function converges, *i.e.*, there are no more new assignments.

B. Partitioning Around Medoids (PAM)

The algorithm uses the most centrally located object in a cluster, the medoid, instead of the mean. Note that a medoid, unlike a mean, is essentially an existing data object from the cluster. It is closest to the corresponding mean. The basic steps are outlined as follows:

- Arbitrarily choose c objects as the initial medoids or seed points.
- Assign each remaining data object (pattern) to the cluster for the closest medoid.
- Replace each of the medoids by one of all the non-medoids (causing the greatest reduction in square error), as long as the quality of clustering improves.
- Iterate until the criterion function converges.

For large N and c, the c-medoids [2] algorithm is computationally more costly than the conventional c-means.

C. Fuzzy c-means (FCM)

This is a fuzzification of the c-means algorithm. It partitions a set of N patterns $\{\mathbf{X}_k\}$ into c clusters by minimizing the objective function

$$J = \sum_{k=1}^{N} \sum_{i=1}^{c} (\mu_{ik})^{m'} ||\mathbf{X}_k - \mathbf{m}_i||^2, \tag{16}$$

where $1 \leq m' < \infty$ is the fuzzifier, $\mathbf{m}_i$ is the ith cluster center, $\mu_{ik} \in [0, 1]$ is the membership of the kth pattern to it, and $||.||$ is the distance norm, such that

$$\mathbf{m}_i = \frac{\sum_{k=1}^{N} (\mu_{ik})^{m'} \mathbf{X}_k}{\sum_{k=1}^{N} (\mu_{ik})^{m'}} \tag{17}$$

and

$$\mu_{ik} = \frac{1}{\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}}\right)^{\frac{2}{m'-1}}}, \tag{18}$$

$\forall i$, with $d_{ik} = ||\mathbf{X}_k - \mathbf{m}_i||^2$, subject to $\sum_{i=1}^{c} \mu_{ik} = 1$, $\forall k$, and $0 < \sum_{k=1}^{N} \mu_{ik} < N$, $\forall i$. The algorithm proceeds as follows.

1) Pick the initial means $\mathbf{m}_i$, $i = 1, \ldots, c$. Choose values for fuzzifier m' and threshold ϵ. Set the iteration counter $t = 1$.
2) **Repeat** Steps 3-4, by incrementing t, **until** $|\mu_{ik}(t) - \mu_{ik}(t-1)| > \epsilon$.
3) Compute μ_{ik} by eqn. (18) for c clusters and N data objects.
4) Update means $\mathbf{m}_i$ by eqn. (17).

Note that for $\mu_{ik} \in [0, 1]$ the objective function of eqn. (16) boils down to the hard c-means case, whereby a *winner-take-all* strategy is applied in place of membership values in eqn. (17).

D. Fuzzy c-medoids

This is a fuzzification of the c-medoids algorithm and is outlined as follows:

1) Pick the initial medoids $\mathbf{m}_i$, $i = 1, \ldots, c$.
2) **Repeat** Steps 3-4 **until** convergence.
3) Compute μ_{ik} for $i = 1, \ldots, c$ and $k = 1, \ldots, N$.
4) Compute new medoids

$$\mathbf{m}_i = \mathbf{X}_q,$$

where

$$q = arg \min_{1 \le j \le N} \sum_{k=1}^{N} (\mu_{ik})^{m'} ||\mathbf{X}_j - \mathbf{X}_k||^2 \qquad (19)$$

refers to that j for which the minimum value of the expression is obtained.

Note that this boils down to the hard c-medoids (PAM) with $\mu_{ik} = 1$, if $i = q$, and to $\mu_{ik} = 0$ otherwise.

E. Clustering Validity Index

The clustering algorithms described above are partitive, requiring prespecification of the number of clusters. The results are obviously dependent on the choice of c. There exist validity indices to evaluate the goodness of clustering, corresponding to a given value of c. In this article we compute the optimal number of clusters c_0 in terms of the Davies-Bouldin and Dunn cluster validity indices [13].

The Davies-Bouldin index is a function of the ratio of the sum of within-cluster distance to between-cluster separation. The optimal clustering, for $c = c_0$, minimizes

$$\frac{1}{c} \sum_{k=1}^{c} \max_{l \ne k} \left\{ \frac{S(U_k) + S(U_l)}{d(U_k, U_l)} \right\}, \qquad (20)$$

for $1 \le k, l \le c$. In this process, the within-cluster distance $S(U_k)$ is minimized and the between-cluster separation $d(U_k, U_l)$ is maximized. The distance can be chosen as the traditional Euclidean metric for numeric features.

Dunn's index is also designed to identify sets of clusters that are compact and well separated. Here we maximize

$$\min_k \left\{ \min_{l \ne k} \left\{ \frac{d_a(U_k, U_l)}{\max_j S_a(U_j)} \right\} \right\}, \qquad (21)$$

for $1 \le j, k, l \le c$.

V. Implementation and Results

Multiresolution clustering and content based retrieval were performed on 44 images, consisting of 27 flowers, 7 aeroplanes and 10 cars. Dimension of the images range from 128 $\times$ 128 to 256 $\times$ 256. We performed three levels of wavelet decomposition of the image, and we considered the $LL1$, $LL2$ and $LL3$ subbands of Fig. 1 for clustering. The 21 extracted features, employed for clustering and retrieval, were (i) four texture-based [eqns. (7)-(10)], each along the four directions, (ii) two shape-based [eqn. (13)], (iii) first two Euler vector components for topology-based, and (iv) one fuzzy geometric feature [eqn. (14)]. Davies-Bouldin (DB) and Dunn's validity indices helped determine the optimal number of clusters. All distances were computed, using Euclidean norm [with or without fuzzy membership, and $m' = 1.2$ in eqn. (16)], in the 21-features space.

A. Clustering

The basic steps of the algorithm are as follows.

1) Apply wavelet transform on the images using eqns. (2)-(5).
2) Extract features in each level of decomposition (*i.e.*, original, $LL1$, $LL2$, $LL3$) using eqns. (6)-(14).
3) Perform multiresolution clustering by c-means, PAM, fuzzy c-means, fuzzy c-medoids algorithms, using eqns. (15)-(19).
4) Apply clustering validity indices to determine optimal partitioning.

Table I demonstrates the optimal clustering, as determined by different clustering validity indices, using the Haar and Daubechies wavelet decomposition. Fig. 2 depicts representative images from two clusters generated by the fuzzy c-medoids algorithm in the $LL2$ subband, with Daubechies function. Partitioning into two clusters was found to be optimal by both the Davies-Bouldin and Dunn's validity indices for both fuzzy algorithms, as observed from Table I. This demonstrates the utility of fuzzy membership in appropriately modeling real life overlapping or ambiguous data.

The first two and last two rows, in Fig. 2, correspond to images from the two clusters respectively. The first element of row 1 and row 3 represent the two cluster means. The remaining elements are sorted in increasing order of their distance from the respective means. It is found that the first cluster consists of different types of flowers, while the second cluster is made up of various means of transport (*i.e.*, aeroplanes and cars). There were 30 and 14 elements in the two clusters, with three misclassifications.

It is to be noted that the system was able to distinguish between categories "flowers" and "means of transport" in the $LL2$ subspace of Fig. 2, irrespective of the orientation, scale and deformation of the object, as well as the intensity of the background. This is mainly due to the wavelet decomposition and the subsequent extraction of good features.

Clustering was found to demonstrate best results when applied in the $LL2$ subbands whose dimension is one-fourth of the original image. The effect of multiresolution, from the coarser original image to the finer resolution $LL1$ through $LL2$, is particularly evident in case of fuzzy clustering. However, the $LL3$ subband was too small (dimension ranging from 16 $\times$ 16 to 32 $\times$ 32) to contain useful information, thereby resulting in degraded clustering performance. However, with images with higher dimension the $LL3$ subband will contain meaningful information for finer clustering results.

TABLE I

OPTIMAL NUMBER OF CLUSTERS FOR DIFFERENT SUBBANDS IN WAVELET DECOMPOSITION

Wavelet function	Subband	Original		$LL1$		$LL2$		$LL3$	
	Clustering	DB	Dunn	DB	Dunn	DB	Dunn	DB	Dunn
Haar	c-means	4	4	2	2	2	2	5	2
	PAM	3	4	2	2	3	2	2	2
	FCM	4	4	3	3	2	2	2	3
	Fuz c-medoids	4	4	3	3	2	2	3	2
Daubechies	c-means	4	4	2	3	2	2	2	2
	PAM	3	4	2	2	2	2	2	2
	FCM	4	4	3	3	2	2	2	3
	Fuz c-medoids	4	4	3	3	2	2	7	2

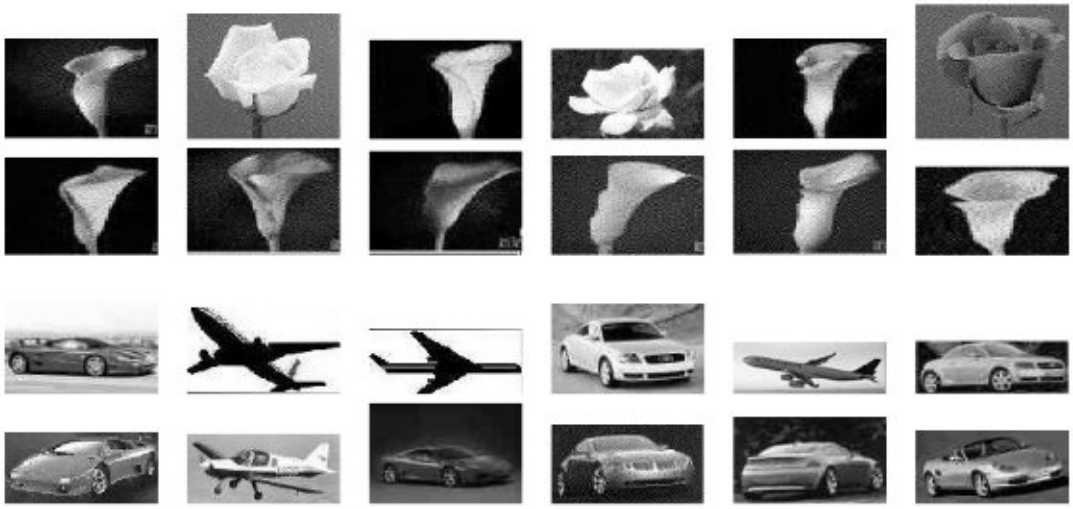

Fig. 2. Images ordered by distance from centroid with fuzzy c-medoids clustering in compressed $LL2$ domain, using Daubechies wavelets for two clusters

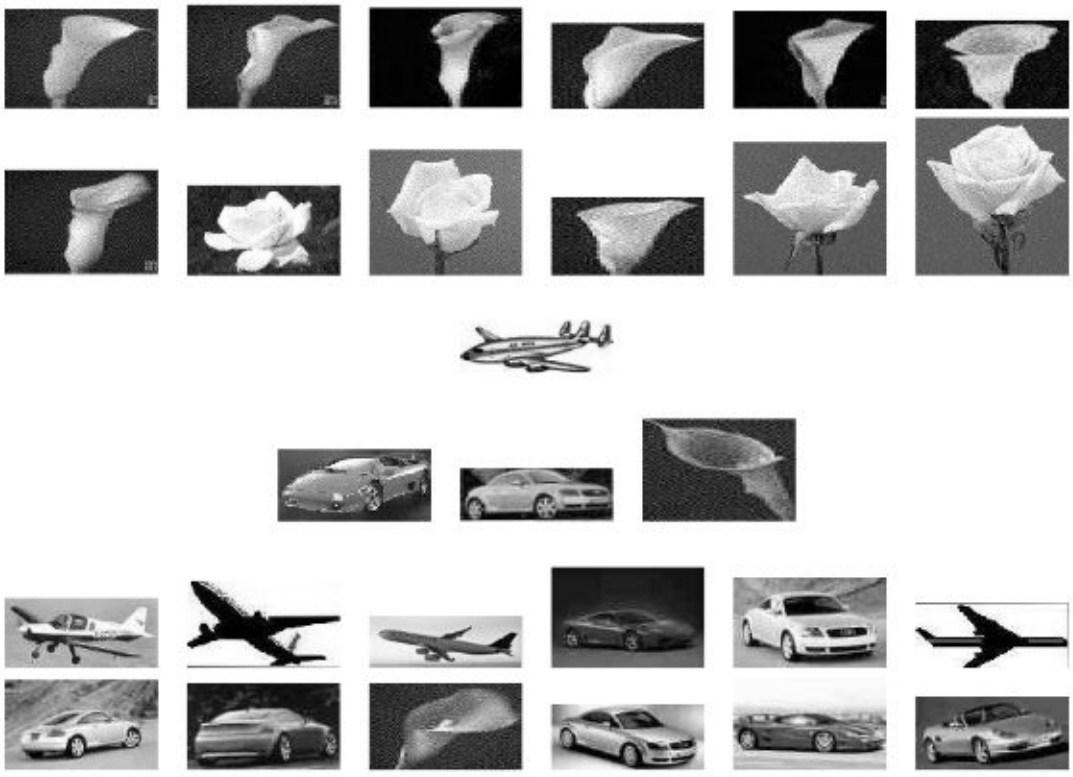

Fig. 3. Images ordered by distance from centroid with fuzzy c-medoids clustering in original space for four clusters

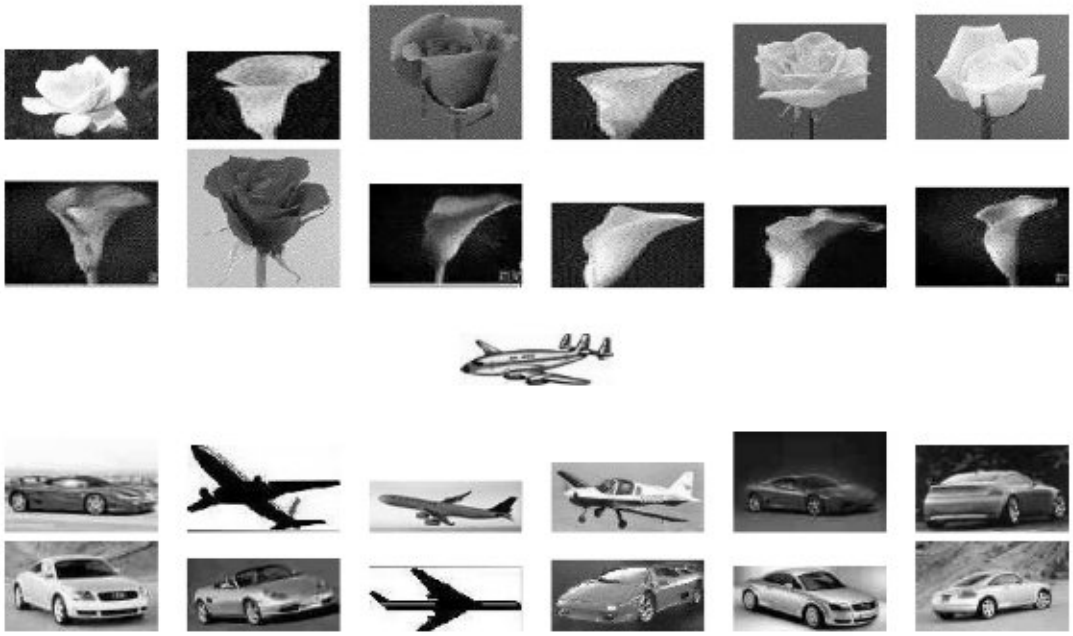

Fig. 4. Images ordered by distance from centroid with fuzzy c-medoids clustering in compressed $LL1$ domain, using Daubechies wavelets for three clusters

As we observe from the table, both the wavelet functions generated 4, 3 and 2 optimal clusters for the two fuzzy clustering algorithms using two validity indices. This is a promising outcome, with the source images being subsampled to four times reduction in size. Its implication to analyzing, manipulating, querying and mining of large image datasets is very obvious. By using this multiresolution wavelet decomposition technique, we could automatically reduce the effect of noise at different levels subbands of the images.

Clustering the original image generated four partitions using fuzzy clustering. Results are depicted in Fig. 3 for fuzzy c-medoids algorithm. Two of these clusters contained only one (row 3) and three (row 4) objects, partly signifying noise. Upon subband decomposition using Daubechies function, the effect of such noise could be minimized, thereby leading to three partitions in the $LL1$ band. This is demonstrated for fuzzy c-medoids clustering in Fig. 4, generating one cluster (row 3) containing only one outlier. Finally, two meaningful partitions were obtained in the $LL2$ subspace, as evidenced from Fig. 2.

B. Retrieval

Here a sample image is presented to the system for content based retrieval of the closest match. Wavelet transform is applied, followed by feature extraction. Results are provided in the $LL2$ subband.

Fig. 5 demonstrates sample CBIR results, from the $LL2$ subband of the wavelet decomposition, using Daubechies function. The first column refers to the query image. The rest of the columns, along each row, depict the content-based retrieved images graded according to increasing distance from the corresponding query image. It is observed, by comparing with Fig. 2, that each query image successfully retrieves images from its own partition.

It is to be noted that the partition "flowers" consists of two varieties of flowers, as illustrated in the first two rows of Fig. 5. The retrieved images lie in the same subcategory as the query image. The partition "means of transport" consists of two subcategories, *viz.*, cars and aeroplanes. Although there is some mismatch in retrieval (between subcategories) in the last case, yet the resultant image is always within the correct partition (means of transport) as observed from the last two rows of the figure. This is perhaps due to the less number of images, available from this category, during system design.

On the whole, the results serve to highlight the utility of the extracted features, in the multiresolution wavelet subband domain, for CBIR and demonstrate the clustering performance from coarser to finer scales.

VI. Utility of Multiresolution Clustering Technique in Compressed Domain

Discrete wavelet transform has been effectively used as an image compression tool. A fundamental shift in the image compression approach came after the Discrete Wavelet Transform (DWT) became popular [14], [15], [16], [17], [18]. The new image compression standard JPEG2000 [19] is based on the discrete wavelet transform and has several advantages over the traditional DCT based image compression standard such as the baseline JPEG. The images are first decomposed into a number of subbands at different levels of multiresolution decomposition. Each subband is then divided into number of smaller size codeblocks (usually of size 32×32) and the codeblocks are independently encoded using entropy encoding scheme. During the entropy encoding scheme, the codeblocks are first processed by a bit-plane coder (BPC) [20] and the output of the BPC is encoded by a binary arithmetic coding scheme (MQ-coder) in order to generate the compressed image. In JPEG2000 image compression standard, the discrete wavelet transform is the computationally most intensive part of the compression system.

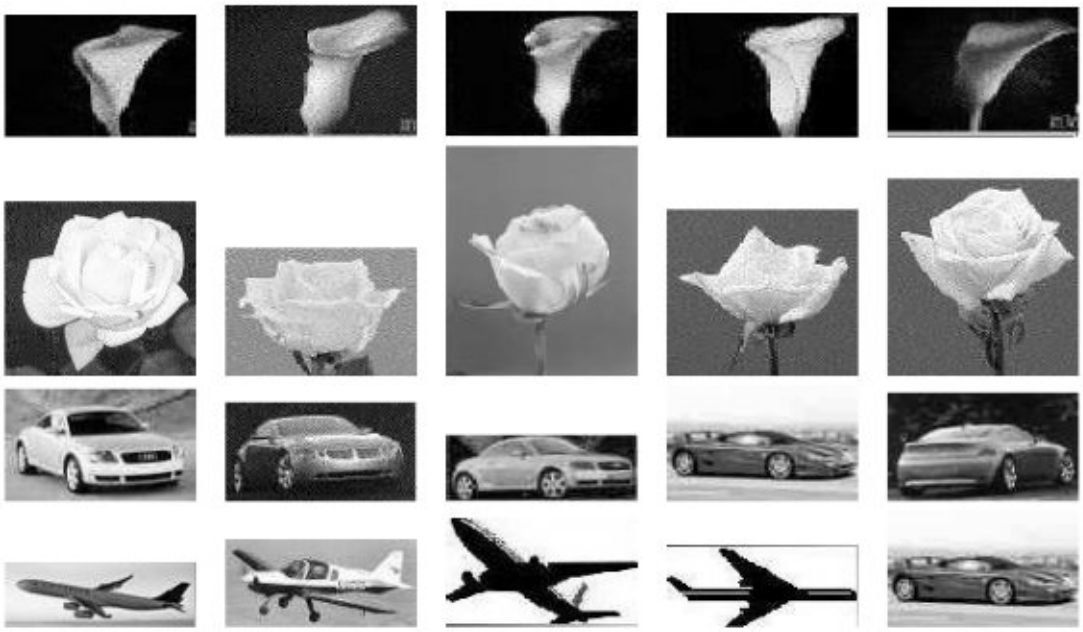

Fig. 5. CBIR for sample images, from both partitions, in $LL2$ subband using Daubechies function

Because of the large data storage requirements, the images are practically stored in compressed form in the image databases. The proposed multiresolution fuzzy clustering scheme is very much suitable to retrieve and cluster the images in compressed domain. In this process, the compressed codes of the images are not completely decompressed to generate the original image. Instead the compressed image is decoded to generate the wavelet subbands only. The proposed multiresolution fuzzy clustering scheme can then be applied to these decoded wavelet subbands. As a result, we do not need to apply the discrete wavelet transform while clustering and retrieving the images.

VII. Conclusions

With the advent of multimedia data mining [3], the need for intelligent storage, search, processing and retrieval of information from large, heterogeneous databases through the application of user-friendly interfaces is assuming utmost importance. This promises wide-ranging applications in fields from photojournalism through medical technology to biometrics. Due to the increasing involvement of pictorial information, the need for image compression and subsequent clustering becomes evident. This is followed by querying of online image databases for CBIR applications. Subsequently these clusters of images may be classified, or associations may be mined between them.

We have described a multiresolution fuzzy clustering of compressed images, with potential for CBIR applications. Wavelet transforms are employed to compress the image, and eliminate noise. This allows us to avoid the need for any other preprocessing of the raw image. The signature of the compressed images, in terms of their visual content, is extracted as features involving texture, shape, topology and fuzzy geometry. The features are invariant to orientation, scale and deformation of object, as well as the background intensity. Partitive clustering, using validity indices, helps us determine the optimal number of clusters. It is observed that fuzzy clustering provides overall better performance, by demonstrating multiresolution partitioning at hierarchical levels of coarseness.

Use of these features for CBIR, in the compressed domain, is also demonstrated. The retrieved images are always found to lie in the same partition as that of the corresponding query image. Since the processing is done on subsampled (compressed) images of smaller size, this strategy now opens up interesting propositions for large scale image mining. The relevance of the proposed multiresolution technique poses tremendous possibilities in clustering and retrieval of images in the compressed domain.

References

[1] J. T. Tou and R. C. Gonzalez, *Pattern Recognition Principles*. London: Addison-Wesley, 1974.

[2] L. Kaufman and P. J. Rousseeuw, *Finding Groups in Data: An Introduction to Cluster Analysis*. New York: John Wiley & Sons, 1990.

[3] S. Mitra and T. Acharya, *Data Mining: Multimedia, Soft Computing, and Bioinformatics*. New York: John Wiley, 2003.

[4] I. Daubechies, *Ten Lectures on Wavelets*. CBMS, Philadelphia: Society for Industrial and Applied Mathematics, 1992.

[5] T. Li, Q. Li, S. Zhu, and M. Ogihara, "A survey on wavelet applications in data mining," *ACM SIGKDD Explorations Newsletter*, vol. 4, December 2002.

[6] G. Van de Wouwer, P. Scheunders, and D. Van Dyck, "Statistical texture characterization from discrete wavelet representations," *IEEE Transactions on Image Processing*, vol. 8, pp. 592–598, 1999.

[7] V. Castelli and L. D. Bergman, *Image Databases: Search and Retrieval of Digital Imagery*. New York: Wiley-Interscience, 2001.

[8] R. C. Gonzalez and R. E. Woods, *Digital Image Processing*. Reading, Massachussetts, USA: Addison-Wesley, 1993.

[9] R. M. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 3, pp. 610–621, 1973.

[10] W. K. Pratt, *Digital Image Processing*. New York, USA: Wiley-Interscience, 1991.

[11] A. Bishnu, B. B. Bhattacharya, M. K. Kundu, C. A. Murthy, and T. Acharya, "Euler vector: A combinatorial signature of gray-tone images," in *Proc. of the IEEE Int. Conf. on Information Technology: Coding & Computing*, (Las Vegas, Nevada, USA), April, 2002.
[12] S. K. Pal and S. Mitra, *Neuro-fuzzy Pattern Recognition: Methods in Soft Computing*. New York: John Wiley, 1999.
[13] J. C. Bezdek and N. R. Pal, "Some new indexes for cluster validity," *IEEE Transactions on Systems, Man, and Cybernetics, Part-B*, vol. 28, pp. 301–315, 1998.
[14] S. G. Mallat, "A theory for multiresolution signal decomposition: The wavelet representation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674–693, 1989.
[15] I. Daubechies, "The wavelet transform, time-frequency localization and signal analysis," *IEEE Transactions on Information Theory*, vol. 36, pp. 961–1005, 1990.
[16] Y. Meyers, *Wavelet: Algorithms and Applications (Translated by R. D. Ryan)*. Philadelphia: SIAM, 1993.
[17] W. Sweldens, "The lifting scheme: A custom-design construction of biorthogonal wavelets," *Applied and Computational Harmonic Analysis*, vol. 3, pp. 186–200, 1996.
[18] I. Daubechies and W. Sweldens, "Factoring wavelet transforms into lifting schemes," *Journal of Fourier Analysis and Applications*, vol. 4, pp. 247–269, 1998.
[19] Information Technology—JPEG2000 Image Coding System—Part 1: Core Coding System ISO/IEC 15444-1, 2000.
[20] D. S. Taubman, "High performance scalable image compression with EBCOT," *IEEE Transactions on Image Processing*, vol. 9, pp. 1158–1170, 2000.