CIREC : Cluster Correlogram Image Retrieval and Categorization using MPEG-7 Descriptors

A. Abdullah and Marco A.Wiering
Intelligent Systems Group
Department of Information and Computing Sciences, University Utrecht
Padualaan 14, 3508TB Utrecht, tel: +31 302539209, fax: +31 302513791
email: {azizi,marco}@cs.uu.nl

***Abstract*— Content-based image retrieval is generally about understanding of information in the images concerned. The more the system is able to understand the content of images the more effective it will be in retrieving desired images. In this paper, we have developed a method that combines a clustering technique with the color correlation histogram. According to this method, an image is divided into several blocks using a fixed partitioning scheme. Then, selected lower-level MPEG-7 features are used to represent the partitions and a clustering technique is applied to group similar patterns into a cluster. The cluster indices will then be used to construct a corellogram which is not based on the lower-level primitives, but on regions consisting of a distribution of primitives. The result shows that the proposed method significantly outperforms other methods on image retrieval and categorization.**

I. Introduction

The need for efficient content-based image retrieval (CBIR) is crucial due to its wide application potential such as for biomedical, commerce and web image classification and searching. One of the main issues in CBIR methods is the result of their searches given a query image. It is common experience for the user to get meaningless information from the query of digital images, although the results seem to be getting better as a database of images grows. Therefore, effective indexing and searching in the large image database is needed and remains a challenge for CBIR.

In this paper, we describe a new method that combines a clustering technique with the color correlation histogram [6]. According to this method, an image is divided into several blocks using a fixed partitioning scheme. Then, selected features are used to represent the partitions and a clustering technique is applied to group similar patterns into a cluster. The cluster indices will then be used to construct a corellogram which is not merely based on the lower-level primitives, but rather on regions consisting of a distribution of primitives. We have implemented this method in the CBIR system called CIREC (Cluster correlogram Image REtrieval and Categorization).

The rest of the paper is organized as follows. Next, we briefly describe related work in CBIR. In Section III, we describe how we use a clustering technique to cluster distributions of lower-level primitives within regions of a fixed partitioning scheme. This technique is used together with the correlogram in CIREC. In Section IV, all indexing features defined in CIREC are introduced with a focus on the MPEG-7 visual descriptors. In Section V the retrieval and categorization effectiveness of CIREC is evaluated by comparing it with other systems on 1000 images of the Corel database. Section VI concludes this paper.

II. Related Work

In this section, we briefly describe some CBIR systems and their properties, but for a much more elaborated survey we refer to [18]. Then we describe basic image features which have been used in CIREC such as the color histogram and the color correlation histogram for image indexing.

A. CBIR Systems

The most frequently cited image features found in literature are color, texture and shape [19], [2], [15], [6], and the most commonly used feature to represent images is color. The color histogram is the best known and most popularly used color feature in CBIR systems and is used in systems such as QBIC [4] and PhotoBook [12]. The color histogram is invariant to rotation, translation and scaling. However, a color histogram does not take into account the spatial information contained in an image, resulting in information loss and coarse indexing. Such indexing can potentially give false results on image queries. Sometimes, two images with dramatically different semantics can give rise to similar histograms. To reduce the problem, Pass and Zabih [6] proposed a split histogram called color coherence vector (CCV). The results produced by this method are quite promising compared to a color histogram. Besides that, Huang et al. [6] proposed another kind of feature called the color correlogram which enables to compute the correlation between colors using spatial information in an image. However, these methods can still not fully solve the problem of fuzziness of the color features inherently exhibited in the color histogram.

The color layout feature was also introduced to overcome the drawbacks of a color histogram. In this method images are partitioned into several blocks and the average color of each block is calculated [19]. However, the color layout is sensitive to shifting, cropping, scaling, and rotation because images are described by a set of local properties [20].

It is quite difficult to construct a good retrieval system. Most of the systems are based on features that are semantically too primitive. Following this, region-based approaches have been introduced to reduce the semantic gap between low-level features and high-level features. A region based approach tries to apply an image segmentation technique to extract regions from images [2]. Thus, similarity between images are measured by calculating the correspondance between their regions. Typical examples of region-based retrieval systems include Blobworld [2], IRM [19], VisualSEEK [15], and SIMPLIcity [19]. However, it is quite difficult to achieve accurate segmentation in an image [15] especially for images with less distinctive objects.

In literature not much attention has been focused to developing a CBIR system that combines information from regions and their spatial correlation. Moghaddan et. al. [1] have proposed a method called the wavelet correlogram as an effort to this direction. Their system is based on a combination of multi resolution image decomposition and the color correlation histogram, where the Gabor wavelet transform is used to compute wavelet coefficients.

B. The Histogram

A color histogram (henceforth referred to as histogram) is frequently used to represent an image's features. In literature, major color indexing methods are based on color histograms [17], [15], [14]. The histogram expresses the frequency distribution of color bins in an image. For simplicity, let's $\mathcal{I}$ denote a digital image and $|\mathcal{I}|$ denote the size of the image. Then, we discretize its color space into *m* distinct colors $c_1, ..., c_m$. Finally, a normalized histogram is computed by dividing the frequency of each color bin by the size of the image. Therefore, the normalized histogram of image $H_{\mathcal{I}}$ can be defined as:

$$H_{\mathcal{I}}(c_i) = \frac{\sum_x \sum_y color(\mathcal{I}_{x,y}) = c_i}{\mid \mathcal{I} \mid} \tag{1}$$

A main advantage of using a histogram is its robustness with respect to the projection of the image. Color histograms are invariant to translation, rotation around the viewing axis, and change slowly with distance to the object and partial occlusion. However, the histogram captures only the color distribution in an image and does not include any spatial correlation between individual pixels. Therefore, it has quite limited discriminative power. The enhanced version of a histogram which is called the color correlogram (henceforth referred to as correlogram) is introduced to reduce this problem. Besides the color distribution, the method also takes into account spatial correlation for individual pixels. It is claimed to be more effective than the histogram [6].

C. The Correlogram

The gray-level co-occurrence matrix [5] is the basis of the correlogram. The co-occurrence matrix is one standard tool for statistical texture analysis and keeps track of the number of pairs of certain intensity pixels that occur at a certain distance and direction in an image. This means that it takes into account the joint histogram probabilities of several neighbor pixels depending on the distance between two pixels.

Let $\mathcal{I}$ be an image, quantized into *m* colors $c_1, ..., c_m$. Let p be a pixel $p = (x, y) \in \mathcal{I}$, and let $p_1 \in \mathcal{I}_{c_i}$ mean that pixel p_1 is of color c_i and $p_2 \in \mathcal{I}_{c_j}$ means that p_2 is of color c_j. The color correlogram [6] matrix C of $\mathcal{I}$ is defined by the joint empirical probability on the image that a color c_i co-occurs with a color c_j at given distance δ and angle φ as:

$$C^{\delta\varphi}(c_i, c_j) = \mathbf{P}(p_1 \in \mathcal{I}_{c_i} \wedge p_2 \in \mathcal{I}_{c_j} \wedge D(p_1, p_2) = (\delta, \varphi)) \tag{2}$$

Where $\mathbf{P}$ means probability, and D(x, y) denotes a distance function using polar coordinates, where $\delta > 0$ and $\varphi \in [0, 2\pi]$. Usually we take a small value for δ since correlation between pixels is more relevant on small distance [13]. The co-occurrence matrix above can be used to construct an asymmetric matrix. The individual elements of an asymmetric matrix are rarely used for texture analysis. Therefore, we construct a symmetric matrix that can be used to analyze overall structure of texture. The symmetric matrix can be constructed by specifying an angle $\varphi \in [0, 2\pi]$ of neighbor pixels. The most common approach is to use the angles 0°,45°,90°,135° and use 1 for the distance. To compute the symmetric matrix the angles are used to compute two values between neigboring points and the results of each angle is combined with others. As a result, one general symmetric matrix can be produced by averaging the results from different angles.

Thus, to compute the correlogram's matrix, the image has to be discretized into *m* distinct colors $c_1, ..., c_m$ and its color co-occurrence matrix is computed to represent overall color texture in the image. The advantages of the correlogram are its easiness to compute and the inclusion of spatial correlations of colors. In their study [7] it is found that a small value of δ is sufficient to represent overall spatial correlation in an image because local correlations are more significant than global correlations.

D. Distance Metric

In content-based image retrieval, a distance metric is usually used to check similarity or dissimilarity between two images. The distance metric tries to capture the strength of relationships between features during comparisons of images in a database. In literature, many similarity measures have been suggested to compare images. In CIREC, we have used the Manhattan distance (also known as the L_1 or city block metric) to measure the similarity between two images. Given a pair of normalized features (f_1,...,f_n) and (g_1,...,g_n) of image $\mathcal{I}$ and $\mathcal{I}'$ respectively, the Manhattan distance $\mathbf{M}$ can be mathematically described as:

$$\mathbf{M}(\mathcal{I}, \mathcal{I}') = \sum_{i=1}^{n} |\mathcal{I}(f_i) - \mathcal{I}'(g_i)| \tag{3}$$

The two images can be said to be similar if the distance $\boldsymbol{M} \approx 0$. We choose the Manhattan distance because it gave the best performance in our experiments.

III. Clustering

The color correlogram works on pixel values such as pixel colors. It cannot be directly used to work on more middle or higher-level features, such as edges. Therefore, CIREC first divides an image into regions, and then uses clustering in order to work with higher-level features based on a group of pixels. The features corresponding to a region are discretized due to the clustering method, and therefore can be directly used to construct a cluster correlogram.

A. K-means Clustering

K-means clustering attempts to subdivide samples consisting of feature values into a set of clusters based on the distances between the samples. Features that are close to each other will be grouped together [8]. The method is quite fast, simple and has been applied and shown to be useful in many applications. To briefly explain the concept of the k-means algorithm suppose the observations are $\{x_i : i = 1, ..., L\}$. The goal of the k-means algorithm is to partition the observations into k groups with mean $\hat{x}_1, \hat{x}_2, ..., \hat{x}_k$ such that

$$KCL(k) = \sum_{i=1}^{L} \min_{1 \leq j \leq k} (x_i - \hat{x}_j)^2 \tag{4}$$

is minimized. K-means clustering works by iterating the following two steps until convergence: (1) assign each observation to the closest cluster-mean, and (2) update the cluster-mean to the centroid of all observations assigned to it in the previous step. The main problem with k-means clustering is that it does not specify how many clusters to choose. In order to determine the appropriate number of clusters k, the clustering technique is executed frequently. We start the experiment with k=2 and increase it by multiplying it with 2 each time. We stop searching when the first derivative of distortion with respect to k, $KCL(k)$ - $KCL(\frac{k}{2})$ is below some small threshold.

B. Cluster Correlogram Approach

Indexing of image files in a feature space is widely used for efficient image retrieval. However, the efficiency of features often decreases with an increase of the size of the feature vector. Recently, region-based feature extraction are developed to enhance the CBIR systems [19], [2]. The advantage of region-based CBIR approaches is that more higher-level features over the regions can be computed, such as shape or texture information, instead of using individual pixel features. Besides that, we also believe a spatial arrangement or distribution for each region plays an important role for efficient retrieval. Based on this consideration, we propose a method that works with any type of feature representation of images to represent regions in an image.

CIREC relies on a fixed partitioning scheme to capture the distribution of image features. This is in contrast with several proposals in literature suggesting methods such as color-based segmentation to characterize the spatial distribution of color information [16]. Although this latter approach provides a more flexible representation and may give more powerful results, we also believe that these advantages are outweighed by the simplicity of the fixed partitioning approach. In the fixed partitioning scheme, each image is divided into B x B blocks as shown in Fig. (1).

Fig. 1. *Left: Original image Right: Partitioned image.*

Afterwards, certain features of all blocks are extracted and clustered into several groups. When clustering has converged, its centroid is used to create cluster indices for each region in each image. Finally, the correlogram is used to construct a spatial arrangement in the image. The cluster correlogram is now computed based on the cluster indices in the regions. Thus, the cluster correlogram uses cluster indices of regions instead of pixel colors, and the rest is similar as the color correlogram described by Equation 2. Clusters which have a (regional) distance smaller than 2 are used for computing the cluster correlogram (which means that every region is correlated with its 8 neighbors). Note that this technique of first clustering particular features and then constructing a cluster correlogram can work with arbitrary features.

IV. Primitive Features Used in CIREC

In CBIR, image features for visual content description is crucial. Good features help to discover meaningful patterns in the image. Until now, there is no agreement what type of features should be used to produce an optimal query result for all images. For instance, a color histogram is quite good to capture the color distribution, but suffers from lack of spatial correlation information i.e. images with different appearances could have similar histograms. Recently, there is one standard called MPEG-7 that has been proposed to provide a standard for automatic indexing for multimedia content [10]. We will use these features as primitives for computing different cluster correlograms in CIREC. When computing the distance between two images, the cluster correlograms of different features are independently computed and finally the Manhattan distances between two images are summed to compute the overall distance.

A. MPEG-7 Content Descriptors

The MPEG-7 [11], [9] standard defines a comprehensive, standardized set for effective searching, identifying, filtering, and browsing on multimedia contents such as images, videos, audios, and other digital or even analog materials [9]. To support various types of descriptors, MPEG-7 is organized into several groups. But in CIREC, we have chosen only primitive MPEG-7 visual descriptors. The reasons why we have chosen to use these visual descriptors is that they are standardized. Besides that, we want to test the effectiveness of using MPEG-7's features in the cluster correlogram. Finally, it gives an easy way to compare our algorithm with other CBIR systems that are based on the same standard. In CIREC, we have used two types of MPEG-7 visual descriptors namely color descriptors and texture descriptors.

1) Color Descriptors: Color is the most instantaneous method of conveying message and meanings in an image. CIREC has used the following color descriptors:

Scalable color - Is a color histogram. The histogram is composed of 256 bins and quantized in HSV color space with 16, 4, 4 values respectively. Then the histogram is encoded by a Haar transform to produce a descriptor.

Color Layout - The main purpose is to represent spatial distribution of colors in an image. It is formed by dividing an image into 8 x 8 non-overlapping blocks and then the representative of the YCbCr color system for each block is obtained. We have chosen 6, 3, 3 for the Y, Cb, Cr coefficients respectively and each color channel is quantized into 64 bins. A Discrete Cosine Transform (DCT) is performed to each block and the coefficients of DCT are used as a descriptor.

Color Structure - The main purpose is to represent local color features in an image. The image is quantized into 128 bins using the HMMD (Hue, Max, Min, Diff) color space. A 8 x 8 sized structuring block is slided over the image. With each shift of the structuring element, the number of times a particular quantized color is contained in the structure element is counted, and a histogram is constructed.

2) Texture Descriptors: Texture is quite important to check homogeneity and non-homogeneity between images. CIREC has used the following texture descriptor:

Edge Histogram - The edge histogram describes a non-homogeneous texture and captures a spatial distribution of edges. First, an image is divided into 4 x 4 non-overlapping blocks. Then, using an edge detection algorithm, six different edge types ($0^o, 45^o, 90^o, 135^o, non-directional, no-edge$) are formed into an eighty-bin histogram.

In literature, texture distribution is quite efficient to represent semantics of an image in uniform background. But sometimes, the semantics might be distracted with other meaningless information. In that sense, color features would be helpful to discriminate homogeneity or non-homogeneity against other information. Color is robust to scale change with the assumption that the lighting condition is constant. For more explanation on the MPEG-7 features, we refer to references cited above. Note that in a normal CBIR system the above features are computed on the whole image, but in CIREC these features are computed in each region separately.

V. Experimental Results

For evaluation of CIREC, we used the first 10 categories and 1000 images of the Corel database. These images have different sizes and were categorized into 10 different groups namely Africans, beaches, buildings, buses, dinosaurs, elephants, flowers, horses, mountains and foods. Each class contains 100 images with different semantics. These images are all in JPEG format with size 384 x 256 or 256 x 384. We will do experiments with retrieval of images and categorization of images. For retrieval we want to have N images returned having the same category as the query image. For categorization we will compute whether a k-nearest neighbor method using majority voting of the retrieved images gives the correct category of the query image.

As mentioned, CIREC is a cluster-based image retrieval and categorization system. The number of clusters (k) is computed manually with a certain threshold value. CIREC is using MPEG-7 visual descriptors as primitives with the following clusters: 32 clusters for scalable color, 32 clusters for color layout, 16 clusters for color structure and 16 clusters for the edge histogram. We have set B for partitioning the image into blocks to 8. CIREC is developed using Java (JDK 1.6.0-beta2) in the Windows 2000 platform. Before the experiment is conducted, the indexing process needs to be performed on each image in the database. The indexing process takes some time and it depends on the number of images, number of features used, and system configuration. We have indexed CIREC on a pentium IV 2.4GHz CPU with 522MB memory. The indexing process has two stages :

- First we cluster the features computed from the regions of all images into a set of clusters. For the 1000 images database it takes 8 hours to complete.
- Second, we use the clusters and the region topology to construct the cluster correlograms for all features we used. This takes 10-15 minutes for 1000 images in the database.

Before evaluating the retrieval effectiveness of CIREC, we first show the results of six query images that were selected from the database. Fig. (2) shows the retrieval results of the queries where only the top 10 images are shown.

For evaluating CIREC and comparing it to using a color correlogram and primitive MPEG-7 features, we compute the precision of the retrieved images on the queries. In our comparison all images will be used one time as a query image. The precision is then computed as follows. Let **Rank**(Q, $\mathcal{I}_i$) = [1,1000] be the rank of retrieved image $\mathcal{I}_i$ from the database, where Q is a query image. The images having a rank below some number N may contain relevant and irrelevant images. Next, let $C(Q, \mathcal{I}_i)$ denote that the retrieved image $\mathcal{I}_i$ has the same category as the query image Q. The precision (P) of the first N retrieved images for a query Q is defined as:

$$P(Q,N) = \frac{\mid \{\mathcal{I}_i \mid \mathbf{Rank}(Q,\mathcal{I}_i) \leq N \bigwedge C(Q,\mathcal{I}_i)\} \mid}{N} \qquad (5)$$

Table I shows the average precision of CIREC for different numbers of retrieved images for each group. Table II and Table III show the average precision of MPEG-7 visual descriptors and the color correlogram respectively. The results clearly show that CIREC outperforms the other methods. Also the

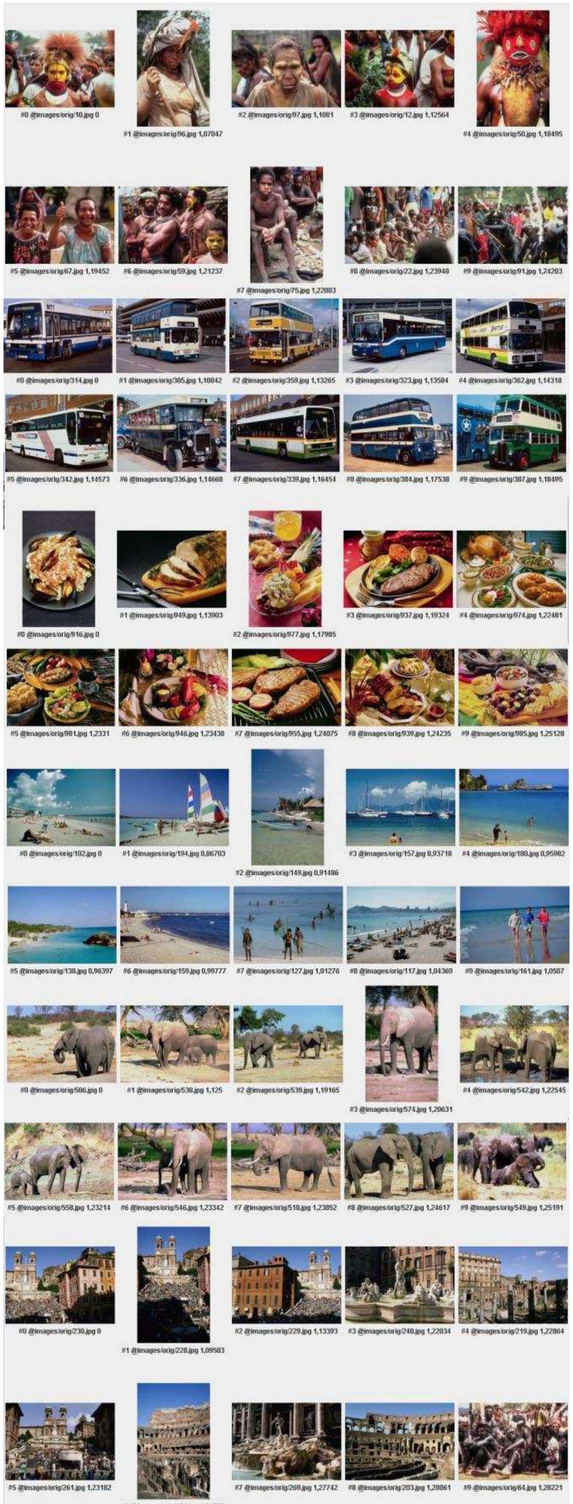

Fig. 2. *Some retrieval results of CIREC on six groups namely Africans, buses, foods, beaches, elephants, buildings respectively; The top-left corner is the query image; The results are arranged first to the right and then downwards.*

TABLE I

AVERAGE PRECISION OF CIREC.

	The number of retrieved images				
	10	20	30	40	50
Africans	0.75	0.70	0.65	0.63	0.60
Beaches	0.69	0.63	0.58	0.55	0.52
Buildings	0.66	0.56	0.51	0.46	0.43
Buses	0.94	0.91	0.89	0.87	0.84
Dinosaurs	1.00	1.00	1.00	1.00	1.00
Elephants	0.72	0.60	0.54	0.49	0.45
Flowers	0.96	0.93	0.91	0.89	0.87
Horses	0.96	0.92	0.90	0.88	0.84
Mountains	0.72	0.66	0.62	0.59	0.57
Foods	0.84	0.78	0.75	0.71	0.66
Total	0.82	0.77	0.74	0.71	0.68

TABLE II

AVERAGE PRECISION OF MPEG-7 FEATURES.

	The number of retrieved images				
	10	20	30	40	50
Africans	0.68	0.57	0.51	0.46	0.42
Beaches	0.52	0.44	0.40	0.37	0.35
Buildings	0.49	0.41	0.37	0.35	0.33
Buses	0.68	0.62	0.57	0.53	0.50
Dinosaurs	1.00	1.00	0.99	0.99	0.91
Elephants	0.58	0.44	0.37	0.34	0.30
Flowers	0.88	0.81	0.78	0.74	0.70
Horses	0.86	0.78	0.72	0.67	0.63
Mountains	0.47	0.40	0.34	0.32	0.29
Foods	0.58	0.49	0.44	0.40	0.38
Total	0.67	0.60	0.55	0.52	0.48

results decrease less fast with an increase of the number of retrieved images. Another CBIR system based on the wavelet correlogram [1] scores a precision of 71% on the Corel database with the same ten categories and $N = 10$.

To measure CIREC's performance for image categorization, we have tested our cluster correlogram technique in combination with the k-nearest neighbor method (k-NN). The k-nearest neighbor method first computes the k closest stored images and then uses majority voting of these k images to decide on the class label. Table IV shows the overall image categorization performance of CIREC, MPEG-7 and Correlogram (Correl.) using k-nearest neighbors. We have tested with various values of k namely k = 1, 3, 5, 7, 9, and 19. When multiple categories

TABLE III

AVERAGE PRECISION OF THE COLOR CORRELOGRAM.

	The number of retrieved images				
	10	20	30	40	50
Africans	0.78	0.69	0.65	0.62	0.59
Beaches	0.53	0.44	0.41	0.38	0.36
Buildings	0.63	0.55	0.49	0.45	0.42
Buses	0.69	0.61	0.57	0.55	0.52
Dinosaurs	1.00	1.00	1.00	1.00	1.00
Elephants	0.68	0.55	0.49	0.44	0.41
Flowers	0.89	0.79	0.74	0.70	0.66
Horses	0.97	0.92	0.88	0.84	0.79
Mountains	0.47	0.41	0.38	0.36	0.35
Foods	0.77	0.70	0.65	0.61	0.57
Total	0.74	0.67	0.63	0.60	0.57

TABLE IV

THE AVERAGE CATEGORIZATION PRECISION RESULTS USING K-NEAREST NEIGHBORS. THE BEST RESULT IS MARKED IN BOLDFACE.

k	1	3	5	7	9	19
$CIREC$	87.0	88.1	88.2	**89.4**	89.0	87.7
$MPEG-7$	75.7	74.8	**76.9**	75.9	76.5	74.0
$CORREL.$	80.7	81.2	80.4	80.7	**81.5**	80.0

TABLE V

THE CONFUSION MATRIX OF IMAGE CATEGORIZATION USING CIREC WITH k=7. A=AFRICANS, B=BEACHES, C=BUILDINGS, D=BUSES, E=DINOSAURS, F=ELEPHANTS, G=FLOWERS, H=HORSES, I=MOUNTAINS, AND J=FOODS.

	A	B	C	D	E	F	G	H	I	J
A	81	1	4	0	1	6	2	2	1	2
B	4	82	1	0	0	0	0	1	12	0
C	6	4	81	0	0	3	4	0	1	1
D	0	1	0	98	0	0	0	1	0	0
E	0	0	0	0	100	0	0	0	0	0
F	6	0	1	0	1	80	0	7	4	1
G	0	0	0	0	0	1	99	0	0	0
H	0	0	0	0	0	0	0	99	1	0
I	1	7	2	0	1	3	1	1	84	0
J	2	0	1	0	4	1	0	1	1	90

Fig. 3. *Some sample images are misclassified. The first row is misclassified as "Beaches" and the second row as "Mountains (with glaciers)". The first and second rows should be classified as "Mountains (with glaciers)" and "Beaches" respectively.*

TABLE VI

THE AVERAGE PRECISION WITH DIFFERENT NUMBERS OF BLOCKS.

	The number of retrieved images				
	10	20	30	40	50
4x4	0.77	0.72	0.68	0.65	0.62
16x16	0.84	0.78	0.74	0.71	0.68
24x24	0.82	0.77	0.72	0.69	0.66
32x32	0.82	0.76	0.72	0.68	0.65

have the same number of votes with a particular $k > 1$, the query image is assigned to the category with the lowest index. CIREC gives the best performance with k = 7 and yields 89.4% correctly classified images. Another categorization system that uses a set of features and support vector machines (SVMs) scored 81.5% on the same dataset [3]. This indicates that the cluster correlogram features are quite expressive and work well in combination with a k-NN classifier.

We also show the results of categorization with CIREC with $k = 7$ using a confusion matrix. The confusion matrix is a square matrix that shows the various classifications and misclassifications of the classifier. In the confusion matrix, numbers on the diagonal are correct classifications and off-diagonal numbers correspond to misclassifications. The confusion matrix is reported in Table V. A detailed examination of the confusion matrix shows that there is one largest number of misclassifications (the underlined number in Table V). The model is slightly confused to make distinctions between "Beaches" and "Mountains (with glaciers)". Fig. 3 shows misclassified images from both categories.

In addition, we have also run some tests to investigate the effect on retrieval and categorization performance of indexing with using different numbers of blocks in CIREC. In this experiment, each image is divided into 4x4, 16x16, 24x24 and 32x32 blocks respectively and the same features and computed cluster centroids are used as for the 8x8 partitioning. Table VI shows the average precision results with different numbers of blocks in CIREC. According to the table, it is clearly seen that the precision varies considerably across the number of retrieved images and gives different results with different blocks. CIREC gives the best overall performance with 16x16 with number of retrieved images = 10 and yields 84%. Table VII also shows that with the same number of blocks, CIREC gives the best overall performance in the average categorization precision. It gives the best performance with k = 5 and yield 90.1% correctly classified images.

In conclusion, we think that with a 16x16 partitioning of an image, CIREC gives a significant resolution to represent regions in an image of visual patterns. The resolution gives an effective representation of spatial distributions of visual clutter for image retrieval. Nonetheless, the disadvantage of using the large number of blocks is that it is computationally more expensive. Unfortunately, CIREC does not have a special algorithm to reduce the size of a large regions while maintaining the resolution characteristics of the original data. However, there are many data reduction techniques which can be used in CIREC. Currently, we are exploring a multiresolution approach as a possible technique to reduce the original size of the data.

VI. CONCLUSION

A new image retrieval and categorization method based on a combination of region-based partitioning, clustering and the color correlogram is proposed. The system called CIREC

TABLE VII

THE AVERAGE CATEGORIZATION PRECISION RESULTS WITH DIFFERENT NUMBERS OF BLOCKS WITH K-NEAREST NEIGHBORS. THE BEST RESULT IS MARKED IN BOLDFACE.

k	1	3	5	7	9	19
4x4	82.1	83.4	84.0	84.3	**85.1**	83.2
16x16	88.7	89.4	**90.1**	89.2	89.1	89.2
24x24	87.9	88.1	**89.3**	88.9	88.8	87.7
32x32	86.7	87.2	**87.7**	87.6	**87.7**	87.1

can work with arbitrary low-level or higher level features, and in the current paper we combined it with the primitive MPEG-7 visual descriptors. On the Corel database, CIREC outperforms a number of other systems on retrieval and categorization. In future work we want to add more features to CIREC. Furthermore, we want to focus on other segmentation techniques that do not need a fixed partitioning scheme.

References

[1] H. Abrishami, A.H. Roohi, and T. Taghizadeh. Wavelet correlogram: A new approach for image indexing and retrieval. *The Journal of the Pattern Recognition Society*, pages 2506–2518, 2005.

[2] Chad Carson, Megan Thomas, Serge Belongie, Joseph M. Hellerstein, and Jitendra Malik. Blobworld: A system for region-based image indexing and retrieval. *Third International Conference on Visual Information Systems*, pages 509–516, 1999.

[3] Y. Chen and J.Z. Wang. Image categorization by learning and reasoning with regions. *Journal of Machine Learning Research*, pages 913–939, 2004.

[4] M. Flickner, H. Sawhney, W. Niblack, J. Ashley, Q. Huang, and B. Dom. Query by image and video content: The QBIC system. In *IEEE Computer Magazine*, pages 23–30, 1995.

[5] R.M. Haralick. Statistical and structural approaches to texture. In *Proceedings of IEEE*, volume 67(5), pages 786–804, 1979.

[6] J. Huang, S. Kumar, M. Mitra, W. Zhu, and R. Zabih. Image indexing using color correlograms. In *Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97).*, pages 762–768, 1997.

[7] J. Huang and R. Zabih. Combining color and spatial information for content-based image retrieval. *In Proceedings of the DARPA Image Understanding Workshop*, pages 687–691, 1997.

[8] A. K. Jain, M. N. Murty, and P. J. Flynn. Data clustering: a review. *ACM Computing Surveys*, 31(3):264–323, 1999.

[9] B.S. Manjunath, P. Salembier, and T. Sikora. Introduction to MPEG 7: Multimedia content description language. In *Wiley, New York*, 2001.

[10] MPEG-7. MPEG-7 overview of the MPEG-7 standard (version 5.0). In *ISO/IEC JTC1/SC29/WG11 N4031*, 2001.

[11] MPEG-7. MPEG-7 visual part of the experimentation model (version 9.0). In *ISO/IEC JTC1/SC29/WG11 N3914*, 2001.

[12] Alex Pentland, Rosalind Picard, and Stan Sclaroff. Photobook: Content-based manipulation of image databases. In *International Journal of Computer Vision*, pages 233–254, 1996.

[13] T.R. Reed and J.M.H. du Buf. A review of recent texture segmentation and feature extraction techniques. *Computer Vision Graphics and Image Processing: Image Understanding*, 57(3):359–372, 1993.

[14] A. W. M. Smeulders, Marcel Worring, Simone Santini, Amarnath Gupta, and Ramesh Jain. Content-based image retrieval at the end of the early years. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22:1349 – 1380, 2000.

[15] John R. Smith and Shih-Fu Chang. Visualseek: A fully automated content-based image query system. In *ACM Multimedia*, pages 87–98, 1996.

[16] J.R. Smith and S-F. Chang. Tools and techniques for color image retrieval. In *Storage and Retrieval for Image and Video Databases (SPIE)*, pages 426–437, 1996.

[17] M. Swain and D. Ballard. Color indexing. *International Journal of Computer Vision*, 22:11–32, 1991.

[18] R.C. Veltkamp, M. Tanase, and D. Sent. Features in content-based image retrieval systems: A survey. In R.C. Veltkamp, H. Burkhardt, and H-P. Kriegel, editors, *State-of-the-Art in Content-Based Image and Video Retrieval*, pages 97–124. Kluwer, 2001.

[19] James Ze Wang, Jia Li, and Gio Wiederhold. SIMPLIcity: Semantics-sensitive integrated matching for picture LIbraries. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(9):947–963, 2001.

[20] James Ze Wang, Gio Wiederhold, Oscar Firschein, and Sha Xin Wei. Content-based image indexing and searching using Daubechies' wavelets. *International Journal on Digital Libraries*, 1(4):311–328, 1997.