Object-Based Image Indexing and Retrieval in DCT Domain using Clustering Techniques

Hossein Nezamabadi-pour   and   Saeid Saryazdi

Abstract- In this paper, we present a new and effective image indexing technique that extracts features directly from DCT domain. Our proposed approach is an object-based image indexing. For each block of size 8*8 in DCT domain a feature vector is extracted. Then, feature vectors of all blocks of image using a k-means algorithm is clustered into groups. Each cluster represents a special object of the image. Then we select some clusters that have largest members after clustering. The centroids of the selected clusters are taken as image feature vectors and indexed into the database.

Also, we propose an approach for using of proposed image indexing method in automatic image classification. Experimental results on a database of 800 images from 8 semantic groups in automatic image classification are reported.

Keywords- Object-based image retrieval, DCT domain, Image indexing, Image classification.

I. INTRODUCTION

In recent years, there has been an intense activity in development image retrieval technique based on image content. Various systems have been introduced for content-based image retrieval (CBIR). CBIR systems operate in two phases: indexing and searching. In the indexing phase, each image of the database is represented using a set of image attribute, such as color[1][2], shape[3][4], texture[5] and layout[6]. Extracted features are stored in a visual feature database. In the searching phase, when a user makes a query, a feature vector for the query is computed. Using a similarity criterion, this vector is compared to the vectors in the feature database. The images most similar to the query are returned to the user.

Rapid advances in hardware technology and growth of computer power, make facilities for spread use of WWW. This causes that digital libraries manipulate huge amounts of image data. Due to the limitations of space and time, the images are represented in compressed formats. As results, techniques used for segmentation and indexing images directly in the compressed domain have become one of the most important topics in digital libraries[7]. Therefore, new waves of research efforts are directed to feature extraction in compressed domain[7][8][9][10][11]. Among compressed domain, JPEG format has been used more than others. As an example more than %95 images on the web are in JPEG compressed format[8].

Region-based image retrieval systems, use local properties of regions(ideally objects) as opposed to the global properties of the entire image. Example of regions-based systems includes SIMPLIcity[12]. If objects within

H. Nezamabadi-pour is with the Electrical Engineering Department, Shahid Bahonar University of Kerman, Kerman, Iran (phone: (98) 341-3335711, fax: (98) 341-3335711, e-mail: nezam@mail.uk.ac.ir).
S. Saryazdi is with the Electrical Engineering Department of Shahid Bahonar University of Kerman, Kerman, Iran (e-mail: saryazdi@mail.uk.ac.ir).

the image are segmented and each object feature is extracted automatically, these features make possible object-based image retrieval[13]. However, representing of the images by adequate number of clusters (objects in the image) can better present the content of them but this approach is time consuming.

In this paper, we investigate the use of DCT coefficient, which are the major components of JPEG, in CBIR. We present a technique to develop and index for an image using DCT coefficients of compressed image. For each block of size 8*8 in DCT domain a feature vector is extracted. Then, feature vectors of all blocks of image using a k-means algorithm is clustered into groups. Each cluster represents a special object of the image. Then we select some clusters that have largest members after clustering. The centroids of the selected clusters are taken as image feature vectors and indexed into the database. We apply proposed image indexing method in automatic image indexing. Experimental results on a database of 800 images from 8 semantic groups in automatic image classification are reported.

In the next section, we present the problem background and review the literature. In section III, we describe the basics of JPEG compression. We illustrate the proposed feature indexing method in section IV. In section V, we present our methodology for object-based image retrieval. Automatic image classification using the propose image indexing approach is introduced in section VI. In section VII, we show the experimental set up and results obtained. Finally, we conclude the paper by presenting a summery.

II. BACKGROUND

A. Image indexing in DCT domain

Ngo et al. developed an image indexing algorithm via reorganization of DCT coefficients in Mandala domain, and representation of color, shape and texture features in compressed domain[7]. Feng et al. proposed an indexing approach by direct extraction of statistic parameters in DCT domain to combine the nature of texture and shape into an integrated feature[9]. The statistical feature extraction is mainly through computing a set of moments directly from DCT coefficients without involving full decompression of inverse DCT.

Jiang et al. proposed an image indexing algorithm that able to extract content from JPEG compressed domain[8]. By analyzing the relationship between DCT coefficients of one block of 8*8 pixels and its four sub-blocks of 4*4 pixels, their proposed algorithm extract an approximated image with smaller size for indexing and content browsing without incurring full decomposition. Ladert and Guerin-Dugue proposed and algorithm for extracting the global distribution of local dominant orientation from DCT domain[11]. They applied the proposed algorithm for categorization of four different semantic groups.

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:3, 2007

B. Region-based image indexing

SIMPLIcity is a region-based system that partitions images into predetermined semantic classes prior to extracting the signature. Signature construction and distance formulations are varied according to semantic class. The k-means algorithm and Haar wavelet and color information of 4*4 blocks are used to segment the image into regions[12].

An image representation method using vector quantization(VQ) on color and texture is proposed by Yoo et al.[13]. Their basic idea is a transformation from raw pixel data to a small set of images region, which are coherent in color and texture space. Features for image indexing are three color components from the HSV color space and five textural features from gray level co-occurrence matrices. Once the features are extracted from an image, 8-D feature vectors represent each pixel in the image. The VQ algorithm is used to cluster those feature vectors into groups. The centroids of clusters are taken as image feature vectors and indexed into the database.

III. JPEG COMPRESSEDOF STILL IMAGES

In this section, we describe the minimal subset of JPEG compression standard, known as the baseline JPEG that is based on DCT. To apply DCT, each pixel in the image is level shifted by 128 by subtracting 128 from each value. Then, the image is divided into fixed size blocks and a DCT is applied to each block, yielding DCT coefficients for the block[10][14]. These coefficients are quantized using weighting functions optimized for the human eye. The resulting coefficients are encoded using Haffman variable word-length algorithm to remove redundancies.

The 2-D DCT for a block of 8*8 pixels is given by(1).

$$C(u,v) = \frac{1}{4}k(u)k(v)\sum_{i=0}^{7}\sum_{j=0}^{7}x(i,j)\cos\left(\frac{(2i+1)u\pi}{16}\right)\cos\left(\frac{(2j+1)v\pi}{16}\right) , \quad 0 \le u,v \le 7$$

$$k(u),k(v)=\begin{cases} \frac{1}{\sqrt{2}} & if\ u\ and\ v=0 \\ 1 & otherwise \end{cases}$$
(1)

Where x(i,j) is the intensity value of pixel in location of (i,j). So, we have

$$C(0,0) = \frac{1}{8}\sum_{i=0}^{7}\sum_{j=0}^{7}x(i,j)$$
(2)

As the formula 2 shows, if C(0,0) is divided by 8 then the average intensity of block is yielding. In an 8*8 block in DCT domain, C(0,0) is DC coefficient and the others are AC.

IV. FEATURE EXTRACTION AND IMAGE INDEXING

in the proposed method, each block of 8*8 pixels in DCT domain is divided into 10 sub-bands(Fig.1).

As we known, in the JPEG compression of color images, the YCbCr color space is used more than other color spaces. For compression of color images, in this space each sub-image(Y, Cb and Cr) is coded separately. In the proposed approach for each color block of size 8*8 a 12-D feature vector is extracted.

The components of this vector are DC components of sub-image of Y, Cb and Cr in DCT domain ($C_Y(0,0)/8$, $C_{Cb}(0,0)/8$, $C_{Cr}(0,0)/8$), components of B1, B2 and B3 of Y sub-image ($C_Y(0,1)$, $C_Y(1,0)$, $C_Y(1,1)$) and the standard deviation of B4, B5, …, B9 blocks of sub-image Y ($std(B_Y4)$, $std(B_Y5)$, …, $std(B_Y9)$). So the feature vector of each color block of 8*8 in DCT domain is as follows.
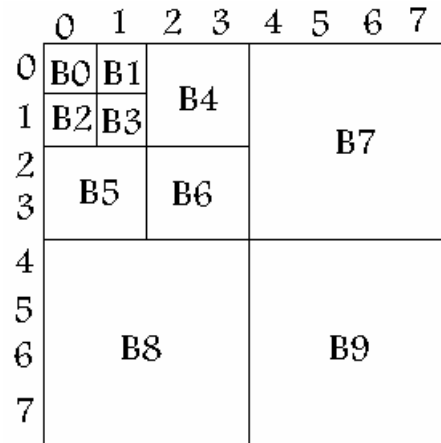

Fig.1. The coefficients in a DCT block of size 8*8 which is divided into 10 sub-bands.

$$F=[C_Y(0,0)/8, \quad C_{Cb}(0,0)/8, \quad C_{Cr}(0,0)/8, \quad C_Y(0,1), \quad C_Y(1,0), \quad C_Y(1,1), std(B_Y4), std(B_Y5), …, std(B_Y9)]$$
(3)

We extract feature vectors of all blocks of image. Then, using a clustering algorithm such as k-means or SOM neural network, these feature vectors cluster into groups. In this work we use k-means algorithm to categorize the image blocks into k clusters. Each cluster represents a special object of the image. Then we select the m clusters that have largest members after clustering. The centroids of selected clusters are taken as image feature vectors (12-D vector for each centroid) and indexed into the database. In our experiments we sets k=10 and m=5. A sample image is depicted in Fig.2 and five representative features of the major objects are listed in Table 1.

Fig.2. A sample image

Table 1. Representative features of major objects

Cluster		1	2	3	4	5
The number of cluster member		631	486	212	154	53
centroid	F(1)	0.2557	0.6198	0.4426	0.4111	0.5038
	F(2)	0.4882	0.4650	0.4610	0.4798	0.4711
	F(3)	0.4981	0.5611	0.5253	0.5397	0.5423
	F(4)	0.1366	0.1235	0.2679	1.0415	0.2320
	F(5)	0.1104	0.1119	0.3934	0.2224	0.9544

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:3, 2007

F(6)	0.0859	0.0786	0.1665	0.1419	0.1693
F(7)	0.0418	0.0404	0.0711	0.1602	0.0569
F(8)	0.0361	0.0440	0.0684	0.0339	0.1567
F(9)	0.0207	0.0261	0.0376	0.0319	0.0518
F(10)	0.0131	0.0208	0.0268	0.0448	0.0262
F(11)	0.0099	0.0224	0.0241	0.0146	0.0497
F(12)	0.0018	0.0131	0.0132	0.0100	0.0240

V. IMAGE RETRIEVAL USING OBJECT FEATURES

As mentioned above, representative features extracted from images are stored in feature database and used for object-based image retrieval. With current computer technology, it is impossible to exactly extract objects in the image and index the object feature. In this paper, we consider five major image's content as representative object features, which described by color and textural information extracted from DCT domain.

In our experiments, we used "query by example" method, QBE, where the user specifies an image, and the system tries to retrieve the most similar image from the database. It should be denoted that the proposed indexing method is based on objects of image. Retrieval procedures are described as proposed method by Yoo et al. as follows[13].

Top-five largest clustering is chosen from the query image. Each object feature contains 12-D vector. We design a graphic user interface that represents each object of query image by its color information only. Therefore, for each query image, color table describes five major objects. For simplicity, only color information is used for describing, but actually each color box includes color textural information.

Choosing one from five boxes starts query. Image similarity is computed using the chi-square distance measure (4) between the query object features and five major object features of first image in the database. We choose closest one from five similarity results as similarity of target object of first image in the database. Repeat above computing procedures across all database images. Retrieve the database images in high similarity order.

$$dissimilarity = \sum_{i=1}^{12} \left(\frac{F^Q(i) - F^T(i)}{F^Q(i) + F^T(i)} \right)^2 \qquad (4)$$

Where Q and T denote query index and target index respectively.

VI. AUTOMATIC IMAGE CLASSIFICATION USING OBJECT FEATURES

For large database with over tens of thousands of image, effective indexing becomes an important issue in CBIR. This problem has not been solved very successfully in current image database systems. A successful categorization of images will greatly enhance the performance of CBIR systems by filtering out images from irrelevant classes during matching [15].

It seems that automatic image classification using the proposed image indexing method is very difficult. As mentioned above, in our method five representative features of the major objects present each image. Now, how we can compute the similarity objects between two images based on their representative features. If we solve this problem, we can use our proposed method in automatic image classification.

We propose an approach for this as follows. The distance of first object of image 1 from five objects of image 2, is computed using the chi-square distance measure(4). This computing procedure is repeated for other four objects of image 1. Then, two objects of image 1 that have the minimum distances from two objects of image 2 are chosen for computing the dissimilarity between two images. Sum of distances of these objects make the dissimilarity between two images. An example of distances between objects of two images is listed in Table 2. The dissimilarity between these images is 0.4788+0.5069=0.9857.

Table 2.
An example of distances between objects of two images

		Clusters of image 2				
		1	2	3	4	5
Clusters of image 1	1	1.689	0.8145	0.5069	1.623	0.8690
	2	0.4788	0.8768	1.5693	1.002	1.1135
	3	0.5690	0.9835	1.238	1.6147	2.5778
	4	1.4116	1.1135	0.5739	2.5183	1.9345
	5	1.5900	1.2018	2.9515	2.0040	0.7563

VII. EXPERIMENTAL RESULTS

A. Image database

We implemented the proposed system on a database of 800 images taken from the Corel collection [12]. These images were arranged in 8 semantic groups: lions, elephants, horses, flowers, foods, mountains, interior design and buses. It includes 100 images from each semantic group. The images are in JPEG format with two sizes 256*384 and 384*256.

B. Evaluation method

There are two methods for evaluation of an image indexing method. One is evaluation of it in an image retrieval system and other is use of it in an automatic image classification. Since the proposed approach is based on objects of images, evaluation of it in an image retrieval system is a subjective matter. To avoid this, we evaluate proposed indexing method in automatic feature classification.

C. Classification

The k-Nearest Neighbor classifier with leave-one-out option is used to classify an image into one of the predefined classes[15][16][17]. Dissimilarity is based on section VI. In our experiments, we set k to 5. Table 3 shows the classification results for k-NN classifier.

Table 3.
Classifier performance for image indexing method

Classification Problem	Accuracy Rate(%)	Misclassification Rate(%)
Lions	95	5
Buses	94	6
Interior design	97	3
Elephants	85	15
Flowers	96	4

World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:1, No:3, 2007

Horses	93	7
Mountains	58	42
Foods	89	11
Total	88.38	11.62

The average accuracy of image classifier is %88.38. Of the 93 misclassified images, 42 were Mountains. If we exclude "Mountains" class, the classifier accuracy increases up to about %95.

VIII. CONCLUSION

In this paper, a novel and effective image indexing technique is presented that extracts features directly from DCT domain. Our proposed approach is an object-based image indexing. For each color image block of size 8*8 in DCT domain a feature vector is extracted. Then, feature vectors of all blocks of an image using the k-means algorithm is clustered into groups. Each cluster represents a special object of the image. Then we select some clusters that have largest members after clustering. The centroids of the selected clusters are taken as image feature vectors and indexed into the database.

Also, we proposed an approach for using of the proposed image indexing method in automatic image classification. Experimental results on a database of 800 images from 8 semantic groups in automatic image classification were reported. The average accuracy of image classifier was %88.38.

REFERENCES

[1] M.J.Swain and D.H.Ballard, "Color indexing", International Journal of Computer Vision, 1991, vol.7, no.1, pp.11-32.
[2] H.Nezamabadi-pour, and E.Kabir,"Image retrieval using histograms of unicolor and bicolor blocks and directional changes in intensity gradient", 2004, Pattern Recognition Letters, vol. 25, no.14, pp. 1547-1557.
[3] F.Mokhtarian and S.Abbasi, "Shape similarity retrieval under affine transforms", Pattern Recognition, 2002, vol. 35, pp. 31-41.
[4] A.K.Jain and A.Vailaya, "Image retrieval using color and shape", Pattern Recognition, 1996, vol.29, no.8, pp.1233-1244.
[5] B.S.Manjunath and W.Y.Ma, "Texture feature for browsing and retrieval of image data", IEEE PAMI, 1996, no. 18, vol. 8, pp. 837-842.
[6] J.R.Smith and C.S.Li, "Image classification and quering using composite region templates", Academic Press, Computer Vision and Understanding, 1999, vol.75, pp.165-174.
[7] C.W.Ngo, T.C.Pong and R.T.Chin, "Exploiting image indexing techniques in DCT domain", pattern Recognition, 2001, vol. 34, pp. 1841-1851.
[8] J.Jiang, A.Armstrong and G.C.Feng, "Direct content access and extraction from JPEG compressed images", Pattern Recognition, 2002, vol. 35, pp. 2511-2519.
[9] G.Feng and J.Jiang, "JPEG compressed image retrieval via statistical features", Pattern Recognition, 2003, vol. 36, pp. 977-985.
[10] S.Climer and S.K.Bhatia, "Image database indexing using JPEG coefficients", Pattern Recognition, 2002, vol. 35, pp. 2479-2488.
[11] P.Ladret and A.G.Dugue, "Categorization and retrieval of scene photographs from a JPEG compressed database", Pattern Analysis and Applications, 2001, no. 4, pp. 185-199.
[12] J.Z.Wang, J.Li and G.Wiederhold, "SIMPLIcity: semantic sensitive integrated matching for picture libraries", IEEE Trans. on Pattern Analysis and Machine Intelligence, 2001, vol.23, no.9, pp.947-963.
[13] H.W.Yoo, S.H.Jung, D.H.Jang and Y.K.Na, "Extraction of major object features using VQ clustering for content-based image retrieval", Pattern Recognition, 2002, vol. 35, pp. 1115-1126.
[14] W.Pennebaker and J.Mitchell, "JPEG still image data compression standard" , 1993, New York: Vann strand.
[15] A.Vailaya, A.K.Jain and H.J.Zhang, "On image classification: city vs. landscape", Pattern Recognition, 1998, vol. 31, pp. 1921-1935.
[16] M.Szummer and R.W.Picard, "Indoor-outdoor image classification", IEEE International Workshop on Content-Based Access of Image and Video Database, in conj. With ICCV'98, Bombay, 1998.
[17] Y.Rubner, J.Puzicha, C.Tomasi and J.M.Buhmann, "Empirical evaluation of dissimilarity measures for color and texture", Computer Vision and Image Understanding, 2001, vol. 84, pp. 25-43.