

## Technical Note

# Content-Based Image Retrieval for Tourism Application Using Handheld Devices

Zobeir Raisi

Faculty of Electrical & Computer Engineering  
University of Sistan & Baluchestan,  
Zahedan, Iran  
[Zobeir.raisi@gmail.com](mailto:Zobeir.raisi@gmail.com)

Farahnaz Mohanna

Faculty of Electrical & Computer Engineering  
University of Sistan and Baluchestan  
Zahedan, Iran  
[f\\_mohanna@ece.usb.ac.ir](mailto:f_mohanna@ece.usb.ac.ir)

Mehdi Rezaei

Faculty of Electrical & Computer Engineering  
University of Sistan and Baluchestan  
Zahedan, Iran  
[Mehdi.rezaei@ece.usb.ac.ir](mailto:Mehdi.rezaei@ece.usb.ac.ir)

Received: September 9, 2011- Accepted: November 29, 2011

**Abstract**—This paper proposed a scenario for using a CBIR (Content-Based Image Retrieval) system in tourism application. Several CBIR algorithms are studied and applied for the proposed scenario. An image database specialized for this application is made to be used for the study purpose. The performance of applied methods were evaluated and compared based on known measures. Among the studied CBIR methods, two algorithms perform better for this application.

**Keywords**- ANMRR; Content-Based Image Retrieval; Handheld; MPEG-7; Mobile; Tourism

## I. INTRODUCTION

In the last decade, image retrieval has become one of the most active research directions in multimedia information processing field because of the rapidly increasing requirements in many practical applications such as architectural management, museum management and education [1, 2]. Text annotation to all images manually is impractical because of large labelling cost and the subjective of human perception [3, 4]. In recent years, Content-Based Image Retrieval (CBIR), which is based on automatically extracted visual features, has attracted more and more attention. In a typical CBIR system, features related to visual content such as shape, color, and texture are initially extracted from a query image [5, 6]. Those of each target image in a database are already computed, and target images which are most similar to query image

are retrieved. In CBIR extraction characterizing of good features is one of the most important tasks [4].

Color is one of the most widely used visual features in CBIR and is invariant to image size and orientation [7]. Colors are defined on a selected color space. Variety of color spaces is available. However, they often serve for different applications. Color histogram is a conventional color feature that is used in CBIR but it does not include any spatial information [8]. Texture is also a visual feature that refers to innate surface properties of an object and their relationship to the surrounding environment [7, 9]. The conventional texture features, used for CBIR, are co-occurrence matrix and edge histogram descriptor (EHD), which are the MPEG-7 standard texture descriptors [10, 11]. Shape features are important image features though they have not been widely used in CBIR as color and

texture features. Shape features have shown to be useful in many domain specific images such as man-made objects. The classic methods of describing shape features are moment invariants, Fourier transform coefficients, edge curvatures, and arc length [12].

The mobile phone industry is going to change phenomenally over the past few years with significant advances in areas of communications and multimedia. Currently, the state-of-the-art multimedia compliant handheld devices such as mobile phones, equipped with digital camera and wireless network connection, enable accessing to large amount of digital media [13, 14]. Moreover, such a powerful device enables new applications. In [15], a client-server content-based image retrieval framework for mobile platforms is developed, which provides the capability of content-based query and browsing from mobile devices. The proposed framework provides an adaptive user interface and a generic structure, which supports a wide range of mobile devices.

Regarding the proposed framework in [15], as a new application, we propose a scenario in which handheld devices and CBIR are used for tourism application. In the proposed scenario a CBIR algorithm is run by a server on a database of images taken from places or other subjects that may be interesting for tourists. Several images of each place or subject exist in the database while useful information is provided about the subject on the server. When a tourist faces an interesting subject, he or she can take an image of subject by a handheld device and send it to the server as query image of CBIR system. In the server, images similar to the query are retrieved and results are returned to the handheld device to be shown on a web browser. Then, the tourist can access the useful information about the subject by clicking on one of the right retrieved images.

In this research, a study was carried out on several CBIR methods, currently proposed by researchers. The methods employed color, texture, and edge features to characterize a query image. The methods were then compared based on their accuracy, running time, and suitability for our application. In this regard, a database was prepared including 1000 images taken from attractions of Zahedan city and University of Sistan and Baluchestan. The photographs were taken in different times of a day and from different angles and distances.

The rest of the paper is organized as follow. Section II introduces a brief overview and comprehensive survey on several recently published image retrieval methods. Experimental results of applying these CBIR methods to our database are shown in Section III and the paper is concluded in Section IV.

II. STUDIED METHODS

A. Color Layout Descriptor

The MPEG-7 color layout descriptor (CLD) is designed to represent the spatial distribution of the color features in an image [16]. During the CLD extraction process, it is assumed that the image consists of three color channels R, G and B. The

steps of CLD descriptor is demonstrated in Fig. 1 and Fig. 2.

The feature extraction process consists of two steps: first the input image is divided into 8x8 non-overlapping blocks and a representative color for each block is determined. Then, the representative color of 8x8 blocks are transformed to YCbCr color space to obtain a down-sampled version of the image. Average color of each block is computed and used as representative color for each block. The CLD is obtained by applying 2-D discrete cosine transform (DCT) on the obtained image. In the next step, a set of low level frequency DCT components of each YCbCr plane are selected using zigzag scanning and quantized to form a CLD [17,18].

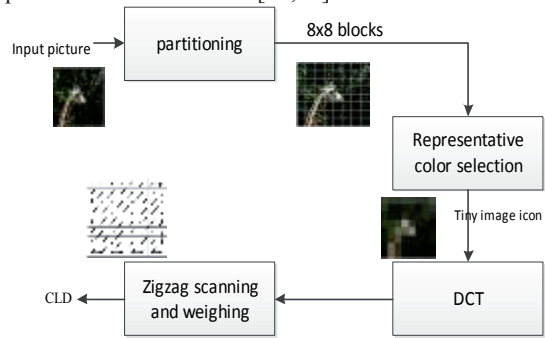


Fig. 1: The CLD extraction process block diagram

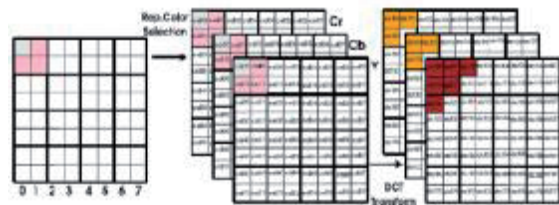


Fig. 2: Zigzag scanning in the CLD extraction process

In this case, the CLD descriptor was formed by reading in zigzag order six coefficients from the Y-DCT matrix and three coefficients from each DCT matrix of the two chrominance components. The descriptor is saved as an array of 12 values (Fig. 2).

For matching two CLDs, {DY,DCr,DCb }, and {DY',DCr',DCb' }, the following distance measure is used [18]:

$$D = \sqrt{\sum_i w_{yi}(DY_i - DY'_i)^2} + \sqrt{\sum_i w_{bi}(DCb_i - DCb'_i)^2} + \sqrt{\sum_i w_{ri}(DCr_i - DCr'_i)^2} \tag{1}$$

where ( w<sub>yi</sub> , w<sub>bi</sub> , w<sub>ri</sub> ) represent the i<sup>th</sup> DCT coefficients of the representative color components. The distances are weighted appropriately, with larger weights given to lower frequency components.

B. Edge Histogram Descriptor

The MPEG-7 edge histogram descriptor (EHD) is an efficient texture descriptor for images with textual



purchase and it can also work as a shape descriptor as long as the edge field contains the true object boundaries. The EHD method represents the distribution of 5 edge types in each local area called sub-image which are: vertical, horizontal, 45-degree diagonal, 135-degree diagonal, and non-edge (Fig. 2) [19, 20]. An example image is shown in Fig. 3. In this example the sub-images are defined by dividing the image into 16 non-overlapping blocks. To characterize the EHD, each sub-image serves a basic region to generate an edge histogram which consists of 5 bins with vertical, horizontal, 45-degree diagonal, 135-degree diagonal, and non-directional edge types. Thus, the histogram for each sub-image represents the relative frequency of occurrence types of edge in the corresponding sub-image [21,22]. For example, an image with 16 sub-images yields a local edge histogram with a total of 80 bins. By scanning sub-images according to the order shown in Fig. 2, the semantics of the bins are

Table I. Semantics of Local Edge Bins

Histogram bins	Semantics
Local_Edge [0]	Vertical edge of sub-image at (0,0)
Local_Edge [1]	Horizontal edge of sub-image at (0,0)
Local_Edge [2]	45 degree edge of sub-image at (0,0)
Local_Edge [3]	135 degree edge of sub-image at (0,0)
Local_Edge [4]	Non-directional edge of sub-image at (0,0)
Local_Edge [5]	Vertical edge of sub-image at (0,1)
⋮	⋮
Local_Edge [74]	Non-directional edge of sub-image at (3,2)
Local_Edge [75]	Vertical edge of sub-image at (3,3)
Local_Edge [76]	Horizontal edge of sub-image at (3,3)
Local_Edge [77]	45degree edge of sub-image at (3,3)
Local_Edge [78]	135 degree edge of sub-image at (3,3)
Local_Edge [79]	Non-directional edge of sub-image at (3,3)



Fig. 3: Five types of edges

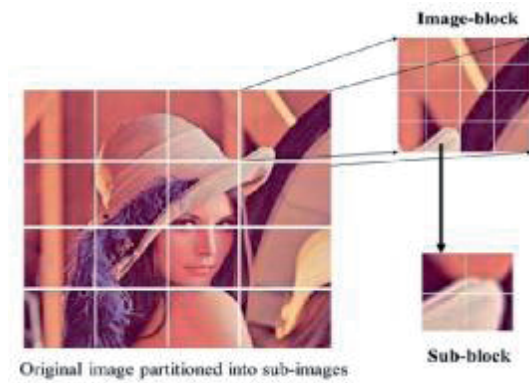


Fig. 4: Definition of sub-image and image-blocks

defined as in Table I. These 80 normalized and quantized bins constitute the standardized EHD of MPEG-7.

To extract edge features, each sub-image or image-block is further divided into four sub-blocks (Fig. 5). Then, the luminance values of the four sub-blocks are used for the edge detection. The mean values of four sub-blocks are convolved with filter coefficients to obtain edge magnitude (Fig. 6). More specifically, the sub-blocks are labelled from 0 to 3 as shown in Fig. 5. For the  $k^{th}$  ( $k=0, 1, 2, 3$ ) sub-block of the  $(i, j)^{th}$  image-block, the average grey level  $a_k(i, j)$  is

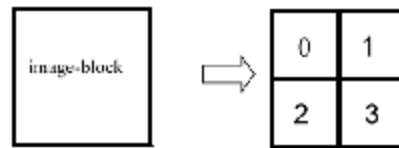


Fig. 5: Sub-blocks and their labelling



Fig.6: Filters for edge detection

calculated. The filter coefficients for vertical, horizontal, 45-degree diagonal, 135-degree diagonal, and non-directional edges are shown as  $f_v(k)$ ,  $f_h(k)$ ,  $f_{d-45}(k)$ ,  $f_{d-135}(k)$  and  $f_{nd}(k)$  respectively. Now the respective edge magnitudes  $m_v(i, j)$ ,  $m_h(i, j)$ ,  $m_{d-45}(i, j)$ ,  $m_{d-135}(i, j)$  and  $m_{nd}(i, j)$  for the  $(i, j)^{th}$  image-block can be obtained as follows:

$$m_v(i, j) = \left| \sum_{k=0}^3 a_k(i, j) \times f_v(k) \right| \tag{2}$$

$$m_h(i, j) = \left| \sum_{k=0}^3 a_k(i, j) \times f_h(k) \right| \tag{3}$$

$$m_{d-45}(i, j) = \left| \sum_{k=0}^3 a_k(i, j) \times f_{d-45}(k) \right| \tag{4}$$



$$m_{d-135}(i, j) = \left| \sum_{k=0}^3 a_k(i, j) \times f_{d-135}(k) \right| \quad (5)$$

$$m_{nd}(i, j) = \left| \sum_{k=0}^3 a_k(i, j) \times f_{nd}(k) \right| \quad (6)$$

$$\max \{m_v(i, j), m_h(i, j), m_{d-45}(i, j), m_{d-135}(i, j), m_{nd}(i, j)\} > Th_{edge} \quad (7)$$

If the maximum value among 5 edges magnitude obtained from (2) to (6) is greater than a threshold ( $Th_{edge}$ ), then the image-block is considered to have the corresponding edge (7). Otherwise, the image-block has no edge. In this paper, the  $Th_{edge}$ , was set at 11 according to the experiments.

C. Co-occurrence Matrix

Co-occurrence matrix is a texture analysis technique for CBIR. The co-occurrence matrix extracts texture information relevant to higher frequency components accurately [10]. Co-occurrence matrices of a binary image  $I$  for a certain  $\Delta x$  and  $\Delta y$  are obtained by the following equation:

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{otherwise} \end{cases} \quad (8)$$

Co-occurrence matrices are computed on the binary image resulted from applying canny edge detector on the gray scale images. Using 12 pair of ( $\Delta x, \Delta y$ ), including  $(0, d)$ ,  $(-d, d)$ ,  $(d, 0)$  and  $(-d, -d)$  where  $d=1, 2, 3$ , will result 12 matrices with the size 2x2 and totally 48 real numbers [23].

D. Scale Color Descriptor

Scale Color Descriptor (SCD) is the major part of MPEG-7 visual standard consisting of color space, color quantization and histogram descriptors [18]. This would allow specification of color histogram with varying number of bins and non-uniform quantization of different color spaces. The SCD uses the HSV color space that is uniformly quantized into a total of 256 bins. This includes 16 levels in H, 4 levels in S and 4 levels in V. The histogram values are truncated into an 11-bit integer representation. Histograms of images with different sizes have different sizes as well. Therefore, different sized histograms can be compared using size conversion provided by Haar transform.

E. Dominant Color Descriptor

Dominant Color Descriptor (DCD) provides an effective, compact and intuitive representation of colors present in an image [7, 18]. In this method, image features are formed by a small number of representative colors. These colors are normally obtained using clustering and color quantization. The DCD in MPEG-7 is defined as:

$$F = \{(c_i, p_i, v_i), s\}, (i=1, 2, \dots, N), \quad (9)$$

where  $N$  is the total number of dominant colors in an image while a region in an image can be presented by a maximum of eight dominant colors.  $c_i$  represents a 3-D dominant colors vector, and  $p_i$  is the percentage of each dominant color. The color variance  $v_i$  and  $s$ , which are optional, show variation of the pixel color values in a corresponding representative color and the spatial coherency, respectively.

To extract DCD features of an image different algorithms such as GLA (Generalized Lloyd Algorithm) and FRCFE (Fixed Representative Colors Feature Extraction algorithm) have been used [24-26]. The GLA is the most extensively used algorithm while it has an expensive computation cost and long running time [23]. For the employed system in this paper, we used FRCFE that provides a better performance with less computational complexity in comparison with the GLA. In FRCFE, 38 perceptual colors are used to represent an image based on DCDs in RGB space [26]. Table II shows these colors selected from the RGB color space.

Table II. Perceptual colors to represent images based on dominant colors

S/N	Re	Gree	Blu	S/N	Re	Gree	Blu
o	d	n	e	o	d	n	e
1	255	0	0	20	0	204	255
2	255	102	0	21	200	200	200
3	255	255	153	22	210	210	175
4	255	255	0	23	255	0	255
5	128	0	0	24	255	153	204
6	255	153	0	25	0	255	255
7	255	153	51	26	153	51	0
8	0	51	0	27	153	102	51
9	51	153	51	28	102	128	51
10	51	51	0	29	204	102	0
11	128	128	51	30	0	0	0
12	51	153	102	31	255	255	255
13	153	204	0	32	51	102	153
14	0	255	0	33	0	51	102
15	0	128	0	34	0	128	128
16	51	102	255	35	51	51	153
17	153	204	255	36	153	51	102
18	0	0	255	37	204	153	255
19	0	0	128	38	85	75	60

Steps of FRCFE algorithm are as follow:

- Read the input image and separate  $R, G, B$  components of the image for each pixel.
- For each pixel
  - search for 38 colors and for nearest colors by computing a distance between the pixel color  $I$  represented by  $(P_r, P_g, P_b)$  and the colors in the colors table  $c_i$  as  $(C_{iR}, C_{iG}, C_{iB})$  using the follow equation:



$$c_d = \min \left( \sqrt{(P_r - C_{iR})^2 + (P_g - C_{iG})^2} + (P_b - C_{iB})^2 \right) \quad (10)$$

- Assign to the pixel an entry from RGB color table with minimum  $c_d$ .

- Create frequency table for assigned colors.
- Sort the frequency table in descending order.

The highest eight frequent colors and their percentages are then selected to create the description of the image. Result of the DCD color quantization using FRCFE for an image is illustrated in Fig. 7.



Fig. 7: DCD color quantization using FRCFE

The dissimilarity  $D(F1, F2)$  between the two descriptors of DCD can be computed as:

$$D^2(F1, F2) = \sum_{i=1}^{N_1} P_{1i}^2 + \sum_{j=1}^{N_2} P_{2j}^2 - \sum_{i=1}^{N_1} \sum_{j=1}^{N_2} 2a_{1i,2j} P_{1i} P_{2j} \quad (11)$$

where  $a_{1i,2j}$  is the similarity coefficient between two colors  $c_{1i}$  and  $c_{2j}$  that is defined in (11). It should be noted that  $c_{1i} \neq c_{2j}$  since they are from different sets of colors in two descriptors,

$$a_{1i,2j} = \begin{cases} 1 - \|c_{1i} - c_{2j}\| / \alpha T_d, & \|c_{1i} - c_{2j}\| \leq T_d \\ 0, & \|c_{1i} - c_{2j}\| > T_d \end{cases} \quad (12)$$

where  $\|c_{1i} - c_{2j}\|$  is the Euclidean distance between two colors  $c_{1i}$  and  $c_{2j}$  in CIELuv color space,  $T_d$  is the maximum distance for two colors to be considered as similar. A typical value for  $T_d$  is between 10 and 20 in the CIELuv color space and for  $\alpha$  is between 1.0 and 1.5. The above dissimilarity measure is very similar to the quadratic distance measure that is commonly used in comparing two color histogram descriptors [17]. In this paper, implementation of DCD algorithm carried out by  $T_d = 15$  and  $\alpha = 1.16$ .

#### F. Fuzzy Color Histogram

Color image histogram is an efficient tool, which is widely used in CBIR. However, using 3D color histogram is complicated and time consuming. Therefore, applying approaches to reduce three dimensions procedures to one is very efficient. Here, the Fuzzy linking method projects a 3D color histogram to 1D on the  $L^*a^*b^*$  color space and provides a histogram which contains only 10 bins. The Mamdani type of fuzzy interface is used for

fuzzification the inputs and output. The fuzzification of the inputs is established using the triangular shaped membership (MF) functions for three input components ( $L^*$ ,  $a^*$ ,  $b^*$ ) which are shown in Fig. 8. The output of system has only 10 equally divided MFs, as shown in Fig. 9. The final fuzzy histogram consists of only 10 bins that are respectively: black, dark grey, red, brown, yellow, green, blue, cyan, magenta, and white. The fuzzy value of three components is made according to 27 fuzzy rules which lead to the system output after defuzzification [26]. In Fig. 10, a sample image with its fuzzy histogram has shown.

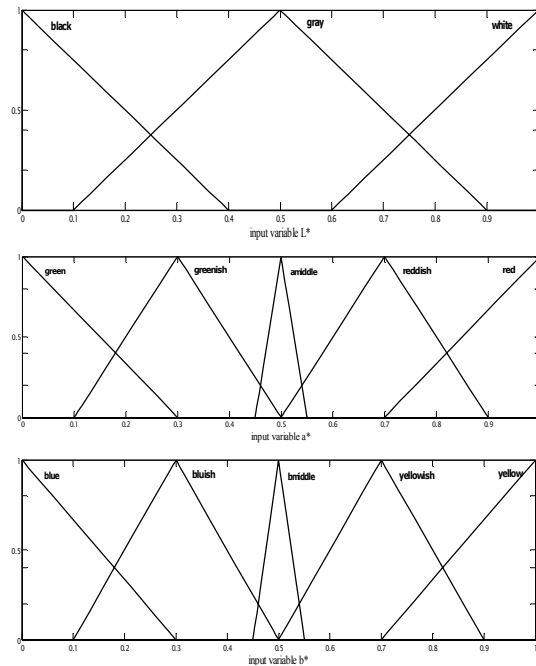


Fig. 8 Member functions of inputs ( $L^*$ ,  $a^*$  and  $b^*$ )

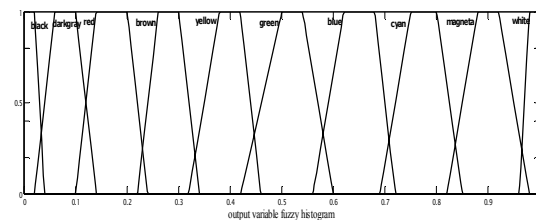


Fig. 9 Member function of the output of the fuzzy system

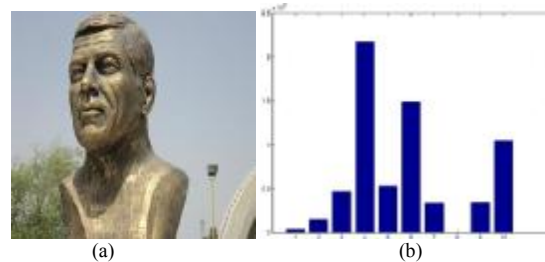


Fig. 10: (a) query image and (b) its fuzzy histogram



## III. EXPERIMENTAL RESULTS

## A. Evaluation

MPEG-7 group has defined an evaluation metric called *Average Normalized Modified Retrieval Rate* (ANMRR) to measure overall performance calculated by averaging the result from each query [28] as:

$$ANMRR = \frac{1}{NQ} \sum_{q=1}^{NQ} NMRR(q), \quad (13)$$

where,  $NQ$  is a number of query images and  $NMRR$  stands for *Normalized Modified Retrieval Rate* that is used to measure the performance of each query according to:

$$NMRR(q) = \frac{\left( \sum_{k=1}^{NG(q)} \frac{Rank(k)}{NG(q)} \right) - 0.5 - \frac{NG(q)}{2}}{K(q) + 0.5 - 0.5 \times NG(q)} \quad (14)$$

where  $NG(q)$  is the size of ground truth image set for a query image  $q$ ,  $Rank(k)$  is ranking of ground truth images retrieved by the retrieval algorithm and  $K(q)$  specifies a "relevance rank" for each query. A suitable  $K(q)$ , because of variation of the ground truth images, is determined by

$$K(q) = \min(4 \times NG(q), 2 \times GTM) \quad (15)$$

where  $GTM$  is the maximum of  $NG(q)$  over all queries.  $NMRR$  and  $ANMRR$  are in the range of (0, 1) and smaller values represent better retrieval performance.

## B. Database

In order to compare the performance of studied CBIR methods based on a common database specialized for tourism application, photographs from the attractions of Zahedan city and University of Sistan and Baluchestan were taken in different times of a day and from different angles and distances. Then, the database was prepared as a two sets, image set consisting of 1000 images, and query set consisting of 21 images including a ground truth set for each query. Sample images from query set are shown in Fig. 11.

## C. Performance Comparison

The explained CBIR algorithms were applied to our database to evaluate their performance. The performance evaluation was carried out in terms of ANMRR criteria and query running time. The reason for selecting the ANMRR evaluation metric is that ANMRR is defined in MPEG-7 standard to measure retrieval performance base on both, the number of correctly retrieved images and how highly these images are ranked which distinct ANMRR from other evaluation metrics such as precision and recall [18,29,30]. Moreover, the ANMRR measure approximately coincides with the results of subjective retrieval accuracy evaluation of search engines [28]. Table III shows the comparison results.

Furthermore, some advantages and disadvantages of the studied algorithms applied to the tourism database are listed in Table IV.

Reported running time in Table III, is only a time for extracting a query. In order to measure the running time of an algorithm in multitasking operating system environment precisely, first we measured the running time of algorithm applied to three subsets with 100, 200, and 400 images of tourism database. Then the measured running time of each subset was divided the number of its images in order to compute the running time for an image. Finally, an average over the three running times obtained by three subsets was computed as the running time of the algorithm. The obtained running times of studied algorithms are presented in Table III.

According to the results, the SCD algorithm provided the minimum running time (0.05 sec.) and the second best ANMRR (0.44). In the next stage, the CLD algorithm performed as the third best in terms of ANMRR and running time. These algorithms provide the most relevant retrieved images for a query as well. This means edge and color are the most important features in this application. Fig. 12 demonstrates retrieval results for a sample query image in our Tourism database applying some of the studied methods. In this figure, the left-upper image marked by a solid black line is a query (1<sup>st</sup> ranked image). Other images are displayed in a raster scan order according to their retrieval ranks.



Fig. 11: A number of sample query images in our tourism database

TABLE III: ANMRR and Running Time results of the studied algorithms applying our tourism database

Method	ANMRR	Time(sec.)
CLD	0.4954	0.0510
EHD	0.3444	0.6206
Co-occurrences	0.4814	0.1402
SCD	0.4464	0.0500
DCD	0.5577	4.0221
Fuzzy-Histogram	0.6636	5.7174

TABLE IV: Advantages and disadvantages of the studied algorithms applying our tourism database

Method	Advantages	Disadvantages
CLD	- Considering local color information - Less sensitivity to fast color variation - Low running time - Good accuracy	- Not considering edge information
EHD	- Considering local edge information - Best accuracy	- Not considering color information - High running time
Co-occurrences	- Considering texture information	- Not considering color & edge information
SCD	- Low running time - Considering color information - Good accuracy	- Not considering edge information - No local information
DCD	- Considering color information	- Not considering edge information - No local information - High running time - Low retrieval performance
Fuzzy-Histogram	- Considering color information - Small feature vector (only 10 bins histogram)	- Low retrieval performance - High running time - No local information

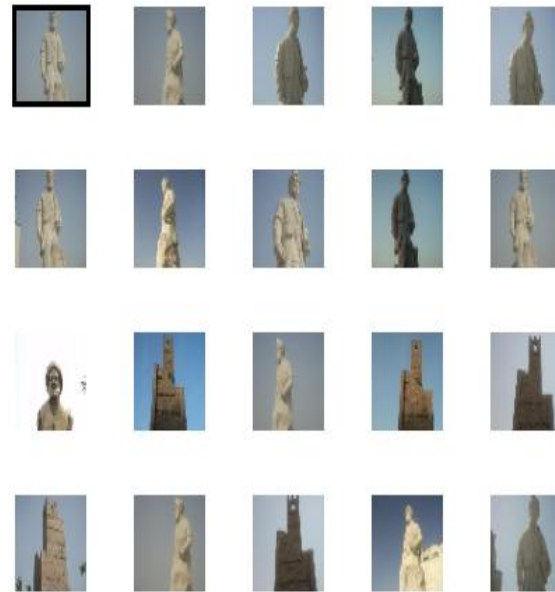


Fig. 12b: Retrieval results of MPEG-7 EHD (NMRR=0.1523). 14 of 28 ground truth images found in the first 20 retrievals.

IV. CONCLUSION

In this paper a scenario for using of CBIR in tourism application was proposed. Several CBIR algorithms were studied and applied for this application. An image database specialized for the proposed scenario was made and the performances of studied methods were evaluated and compared based on the ANMRR criteria and query time. Among the studied method the EHD and SCD methods presented best performance.

REFERENCES

- [1] J. Z. Wang, J. Li, and G. Wiederhold, "SIMPLcity: Semantics-sensitive integrated matching for picture libraries," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 9, pp. 947-963, Sep. 2001.
- [2] W. Bian and D. Tao, "Biased Discriminant Euclidean Embedding for Content-Based Image Retrieval", *IEEE Transaction on Image processing*, vol. 19, NO. 2, February 2010.
- [3] Y. Rui and T.-S. Huang, "Optimizing learning in image retrieval," presented at the IEEE Int. Conf. Computer Vision and Pattern Recognition, 2000.
- [4] F. Li, Q. Dai Dai, W. Xu, and G. Er, "Multilabel Neighborhood Propagation for Region-Based Image Retrieval", *IEEE Transaction on multimedia*, vol. 10, no. 8, December 2008.
- [5] M. Swain and D. Ballard, "Color indexing," *Int. J. Computer. Vis.*, vol. 7, no. 1, pp. 11-32, 1991.
- [6] H. Tamura, S.Mori, and T.Yamawaki, "Texture features corresponding to visual perception," *IEEE Trans. Syst., Man, Cybern.* vol. 8, no.SMC-6, pp. 460-473, 1978.
- [7] X. Wang, Y. Yu and H. Yang, "An effective image retrieval scheme using color, texture and shape features", *ELSEVIER Computer Standards & Interfaces*, vol. 33, pp.: 59-68, 2011.
- [8] Y. Deok Chun, N. Kim, and I. Hoon Jang, "Content-Based Image Retrieval Using Multiresolution Color and Texture Features", *IEEE Transaction on Multimedia*, vol. 10, no. 6, OCTOBER 2008.
- [9] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture Image Retrieval Using New Rotated Complex Wavelet Filters

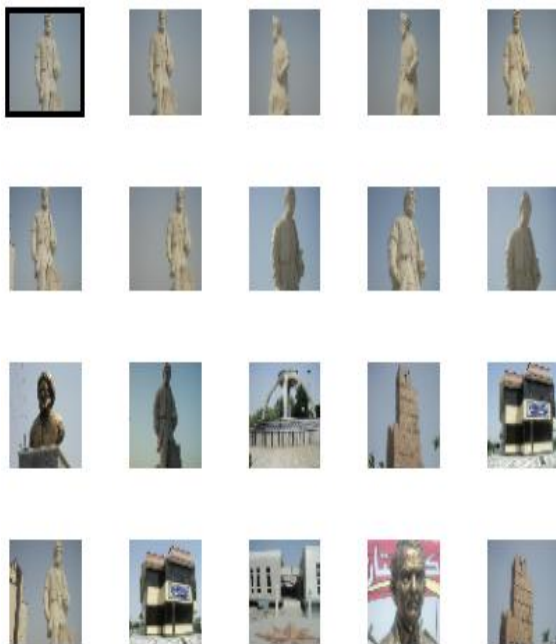


Fig.12a: Retrieval results of MPEG-7 DCD (NMRR=0.3070). 12 of 28 ground truth images found in the first 20 retrievals.



- [9] M. Kokare, P. K. Biswas, and B. N. Chatterji, "Texture Image Retrieval Using New Rotated Complex Wavelet Filters", *IEEE Transaction on systems, man and cybernetics, part B: Cybernetics*, vol. 35, NO. 6, December 2005.
- [10] M. Partio, B. Cramariuc and M. Gabbouj, "Texture Retrieval Using Ordinal Co-occurrence Features", *NORSIG-2004,-6<sup>th</sup> Nordic Signal Processing Symposium*, On Board Hurtigruten M/S Trollfjord, June 9 - 11, 2004, Finland.
- [11] ISO/IEC/JTC1/SC29/WG11: "Core Experiment Results for Edge Histogram Descriptor (CT4)," *MPEG document M6174*, Beijing, July 2000.
- [12] Y. Liu, D. Zhang, G. Lu, and W.-Y. Ma, A survey of content-based image retrieval with high-level semantics. *Pattern Recognition*. Vol. 40, No. 1, pp. 262-282, 2007.
- [13] A. Iftikhar, F. AlayaCheikh, B. Cramariuc, and M. Gabbouj, "Query by Image Content Using Mobile Information Device Profile (MIDP)", *Proc. of the 2003 Finnish Signal Processing Symposium*, Finsig'03, Tampere, Finland, May 2003.
- [14] A. Iftikhar, M. Gabbouj, "Content-based Interactive Image Retrieval from Java Enabled Mobile Devices", *proceeding of the 4<sup>th</sup> International Symposium on Communications, Control and signal Processing*, ISCCSP 2010, Limassol, Cyprus, 3-5 March 2010.
- [15] I. Ahmad, M. Gabbouj, "A generic content-based image retrieval framework for mobile devices", *Multimedia Tools and Applications*, Springer, 2011.
- [16] A. Yamada, and E. Kasutani, "The MPEG-7 Color Layout Descriptor: A Compact Image Feature Description for High speed Image/Video Segment Retrieval," *ICIP*, vol. 1, pp. 7-10, Oct. 2001.
- [17] Sung Min Kim et al., "Image Retrieval via Query-by-Layout Using MPEG-7 Visual Descriptors", *ETRI Journal*, Volume 29, Number 2, April 2007.
- [18] B. S. Manjunath, Jens-Rainer Ohm, Vinod V. Vasudevan, and Akio Yamada, "Color and Texture Descriptors", *IEEE transaction on circuits and system for video technology*, vol. 11, no. 6, June 2001.
- [19] B.S. Manjunath, Philippe Salembier, Thomas Sikora, "Introduction to MPEG-7", JOHN WILLEY & SONS, LTD, pp. 183-184, 2002.
- [20] ISO/IEC/JTC1/SC29/WG11: "Core Experiment Results for Edge Histogram Descriptor (CT4)," *MPEG document M6174*, Beijing, July 2000.
- [21] S. Nandagopalan, Dr. B. S. Adiga, and N. Deepak. "A Universal Model for Content-Based Image Retrieval", *World Academy of science, Engineering and Technology* 46 2008.
- [22] M. Eom, and Y. Choe, "Fast Extraction of Edge Histogram in DCT Domain based on MPEG7", *World Academy of science, Engineering and Technology* 9. 2005.
- [23] D.A. Cluasi, H. Deng, "Fusion of Gabor Filter and Co-occurrence Probability Features for Texture Recognition", *IEEE Transactions on Image Processing*, Vol. 14, No. 7, pp. 925-936, 2005.
- [24] A.Gersho, R.M. Grey, "Vector Quantization and Signal Compression," Kluwer, Norwell, Massachusetts, 1991.
- [25] N. Yang, C. Kuo, W. Chang and T. Lee, "A Fast Method for Dominant Color Descriptor with New Similarity Measure" *Journal of Visual Communication and Image Representation*, vol. 19, pp. 92-105, 2008.
- [26] A.Ibrahim and et al, "Fixed Representative Color Feature Extraction Algorithm for Moving Picture Experts Group-7 Dominant Color Descriptor", *Journal of Computer Science*, vol. 5 (11), pp. 773-777, 2009.
- [27] K. Konstantinidis, A. Gasteratos b, I. Andreadis, "Image Retrieval based on Fuzzy Color Histogram Processing", *ELSEVIER Optics Communication*, vol. 248, pp. 375-386, 2005.
- [28] P. Ndjiki-Nya, J. Restat, T. Meiers, J.-R. Ohm, A. Seyferth, and R. Sniehotta, "Subjective Evaluation of the

MPEG-7 Retrieval Accuracy Measure (ANMRR)," *ISO/WG11 MPEG Meeting, Geneva, Switzerland*, Doc. M6029, May 2000.

- [29] D.V.K Balasubramani.R,"Efficient use of MPEG-7 Color and Edge Histogram Descriptors in CBIR Systems"," *Global Journal of Computer Science and Technology*, pp. 157-163,2007.
- [30] W.K.Man,"Content Based Image Retrieval Using MPEG-7 Dominant Descriptor" Master of philosophy, Department of Electronic Engineering, CITY UNIVERSITY OF HONGKONG, Hongkong, 2004.



**Zobeir Raisi** received his B.Sc. degree in electronic engineering from University of Sistan and Baluchestan in 2009 and his M.Sc. degree in communication engineering (with honor) from University of Sistan and Baluchestan Zahedan, Iran in 2011. He is a member of the Computer Vision and Image Processing, IAEE and Iranian Fuzzy System Committees. His current research interests include pattern recognition, image processing, image retrieval, computer vision and digital signal processing (DSP).



**Farahnaz Mohanna** received her B.Sc. degree in electronics engineering from University of Sistan and Baluchestan, Iran in 1987, her M.Sc. degree in electronics engineering from Tehran University in 1992, and her Ph.D. degree in Image Processing from Surrey University, Guilford, UK, in 2002. Dr. Mohanna has been working as a research fellow at the Centre for Vision, Speech and Signal Processing (CVSSP) at the Surrey University, UK in 2003. Dr. Mohanna is an Assistant Professor and Faculty Dean of Electrical and Computer Engineering, University of Sistan and Baluchestan, Iran. Her research interests include corner detection, active contours, object tracking, and video database retrieval and communications. She has published several Journal and conference papers in these fields. She is a member of the computer vision and image processing societies.



**Mehdi Rezaei** (IEEE M'04) received his B.Sc. degree in electronics engineering from Amir Kabir University of Technology (Polytechnic of Tehran) in 1992, his M.Sc. degree in electronics engineering from Tarbiat Modares University of Tehran in 1996, and his PhD degree in Signal Processing from Tampere University of Technology, Finland, in 2008. Dr. Rezaei had a closed cooperation with the Nokia Research Center as researcher and senior researcher from 2003 to 2009. Dr. Rezaei is an Assistant Professor and also Deputy Dean for Education and Postgraduate Studies, in Faculty of Electrical and Computer Engineering, University of Sistan and Baluchestan, Iran. His research interests include multimedia signal processing and communications. He has published several journal and conference papers in these fields. He also holds several patents in related applications. Dr. Rezaei was the recipient of the Nokia Foundation Award in 2005 and also in 2006. He is a member of IEEE SP, COMSOC, CAS, and BT Societies.

