Image Retrieval System based on Color Global and Local Features Combined with GLCM for Texture Features

Jehad Q. Alnihoud

Department of Computer Science,
Al al-Bayt University

Abstract—In CBIR (content-based image retrieval) features are extracted based on color, texture, and shape. There are many factors affecting the accuracy (precision) of retrieval such as number of features, type of features (local or global), color model, and distance measure. In this paper, a two phases approach to retrieve similar images from data set based on color and texture is proposed. In the first phase, global color histogram is utilized with HSV (hue, saturation, and value) color model and an automatic cropping technique is proposed to accelerate the process of features extraction and enhances the accuracy of retrieval. Joint histogram and GLCM (gray-level co-occurrence matric) are deployed in phase two. In this phase, color features and texture features are combined to enhance the accuracy of retrieval. Finally, a new way of using K-means as clustering algorithm is proposed to classify and retrieve images. Two experiments are conducted using WANG database. WANG database consists of 10 different classes each with 100 images. Results of comparing the proposed approach with the most relevant approaches are promising.

Keywords—CBIR; color histogram; GLCM; K-means; WANG database

I. INTRODUCTION

Visual cues rather than textual description of images is used in CBIR to extract features automatically from query image and image database and the most similar images are ranked and retrieved. Color and texture features are considered as the most frequently used in CBIR that because it is easy to extract features based on these visual contents. Moreover, shape [1], and regions [2] are broadly considered to index image based on contents. These low level features of an image are directly related to the contents of the image. In [3], they propose CBIR system based on color distribution. The proposed system relies on the probability of identical pixel colors as compared with its neighbors and the difference among pixels. Moreover, they deployed K-means Clustering algorithm group pixels. The approach overcome problems related to image displacement and rotation.

There are three major methods for color analysis and retrieval. The most common method is global-based or global color histogram. Approaches relies on this method such as [4, 5] are efficient but suffers from lack of information related to spatial distribution of colors. The second method relies on dividing images to equal blocks or regions such as [1, 6] and utilizes local color histogram for each region to extract color

features. These approaches provides better information related to spatial distribution of colors. The third method is segmentation based, in which the division of the image is done using clustering algorithm such as [7, 3]. Complexity of features extraction is considered as the major drawback of such approaches.

Texture is used as an image discriminator utilizing visual properties of homogeneity rather than the existence of distinctive color [8]. There are many texture features such as contrast, correlation, energy, uniformity, density, coarseness, and homogeneity. As stated in [9], the most common texture descriptors are statistical model and transform-based. Statistical model extracts statistical features based on spatial dependence of texture's grey-level. The co-occurrence matrix used in [3, 10] is considered as an example of statistical model based. Discrete wavelet transform (DWT) is the most dominant example of transform based model. In [17], they propose CBIR system by integrating many features extracted using DWT, HSV color feature, cumulative color histogram, dominant color descriptor (DCD) and GLCM. Moreover, they propose using adjacency matrix of a bipartite graph to the sub-blocks of query image and image database. The proposed approach achieves better results as compared with the most relevant retrieval schemes. In [18], they propose CBIR system based on evolutionary algorithm. The proposed system discriminate images based on shape, color and texture features. The system is evaluated by precision-recall value of the retrieved results. In [19], they studied the effect of color spaces on CBIR using wavelet decomposition of each color channel. They recommended to use Lab color space and color features extracted using color moments. In [20], a two-leveled CBIR system is proposed. The proposed approach uses color features, edge detection, morphological operator (Dilate), and 2D-DWT (wavelet transform). WANG database is used to test the proposed CBIR system. Experimental results show superiority of the proposed approach as compared with other CBIR systems.

Most of the previously mentioned approaches used color and texture features. In this paper attempt is made to overcome some limitations related to spatial consideration of color distribution, disregard unrelated parts of image which may affect the accuracy of retrieval, and better utilizing of k means as clustering algorithm to enhance the retrieval accuracy.

In this paper a new and efficient CBIR system of two phases is proposed. In the first phase the emphases is given to global color features utilizing color histogram. Moreover, a new cropping technique is proposed to eliminate the least significant pixels within the image. This phase allows huge reduction of dissimilar images to be considered in the second phase. However, in the second phase images are sliced to three different regions and for each region the most significant features of color and texture based on joint histogram and gray level co-occurrence are calculated. Finally, a modified way of using K-means clustering is proposed to retrieve the final set of similar images as compared to the query image. The rest of the paper is organized as follows. Section 2 reviews the most important global and local color features to be extracted. Section 3 explains the texture features extraction using GLCM. Proposed approach is discussed in section 4. Section 5 presented similarity measures and experimental results. Finally, section 6 draws the conclusion.

II. COLOR FEATURE EXTRACTION

In CBIR number and type of features extracted from image play a significant role in increasing or decreasing the precision of image matching and retrieval. In this section the major extracted color features are discussed to enhance the understanding of the proposed system. As shown in [11], color is easy to implement and it considered as the most reliable visual content as compared with texture and shape. It is robust to background compilation and not affected by image size and orientation. Color histogram is the most prominent approach for color features extraction. Color histogram demonstrates the color distribution in image and it requires less computational cost as compared with other color approaches. Lack of spatial information is considered as the major drawback of color histogram approach.

As stated in [12], color histogram of image is represented using vector H= {H(0), H(1), …,H(j),…, H(N)}, where j denotes the color bin, N represents the number of bins, and H(j) denotes the total number of pixels at j^{th} color bin. In this research, images are converted to HSV color space first, then color histogram is used to extract global features of cropped images. Furthermore, statistical-based features (color moments) are extracted for each image regions of the cropped images. These statistical features as follows:

- *Mean***:** the mean value for each color channel (H, S, and V) is calculated as follows:

$$M_i = \frac{1}{N}\sum_{j=1}^{N} P_{ij} \qquad (1)$$

Where, N is the number of pixels in the image and P_{ij} is the jth pixel value of the image at the ith color channel

- *Standard deviation:* the standard deviation value for each color channel (H, S, and V) is calculated as follows:

$$\sigma_i = \sqrt{\left(\frac{1}{N}\sum_{j=1}^{N}(P_{ij} - M_i)^2\right)} \qquad (2)$$

Where, N is the number of pixels in the image and P_{ij} is the jth pixel value of the image at the ith color channel

- *Entropy:* the entropy value for each color channel (H, S, and V) is calculated as follows:

$$Entropy = -\sum_i P_i log_2 P_i \qquad (3)$$

Where, P_i is the probability of the difference between 2 adjacent pixels.

III. TEXTURE FEATURES EXTRACTION

As stated in [9], GLCM elements measures the relative of gray level occurrences between pairs of pixels within the image in different directions. The co-occurrence matrix is considered as useful way of extracting second order statistics. These statistics related to human view of texture.

In this research GLCM is utilized to extract texture statistical features (contrast, energy, correlation, and homogeneity). Moreover, GLCM is applied for images sub-regions in order to enhance accuracy of matching utilizing localization of features rather than relying on global texture features only. As an example, the GLCM is calculated to the following image in figure 1 using the following MATLAB code.

The MATLAB code is shown in figure 2.

```
%read image
I= imread('beach.jpg');
% convert rgb image to gray
R1 = rgb2gray(I);
% resize image
R2 = imresize(R1, [200 200]);
% create gray-level comatrix with default offset [0 1]
GL = graycomatrix(R2,'Offset',[0 1]);
%Extract properties of gray-level co-occurrence matrix
Ca = graycoprops
(GL,{'contrast','Correlation','Energy','homogeneity'})
```

Fig. 1.   GLCM calculation (MALLAB code)

Gray level co-occurrence matrix of the sample image is [8 x 8] as follows:

1034	149	3	0	0	0	0	0
137	3606	700	17	0	1	0	0
4	659	13833	738	78	22	5	0
1	29	654	7024	684	75	35	0
0	10	52	687	4796	404	63	3
0	1	26	81	409	1892	185	1
0	0	1	12	59	211	1397	9
0	0	0	0	1	8	4	0

Fig. 2. Sample image

Table 1 shows statistical features of the sample image:

TABLE I. STATISTICAL FEATURES

Contrast	Correlation	Energy	Homogeneity
0.2168	0.9404	0.1809	0.9192

IV. PROPOSED APPROACH

In this section a complete description of the proposed approach is presented.

A. Phase I: Global Color Histogram

a) ∀ Image (I) ∈ Db (database) do:

1) Convert the RGB images to HSV. Figure 3 shows the original image converted to HSV color model and the different histograms related to hue, saturation, and value. As stated in [13], RGB image is converted to HSV color space using the following formulas:

$$H = cos^{-1} \frac{\frac{1}{2}[(R-G)+(R-B)]}{\sqrt{(R-G)^2+(R-B)(G-B)}} \qquad (4)$$

$$S = 1 - \frac{3[min(R,G,B)]}{R+G+B} \qquad (5)$$

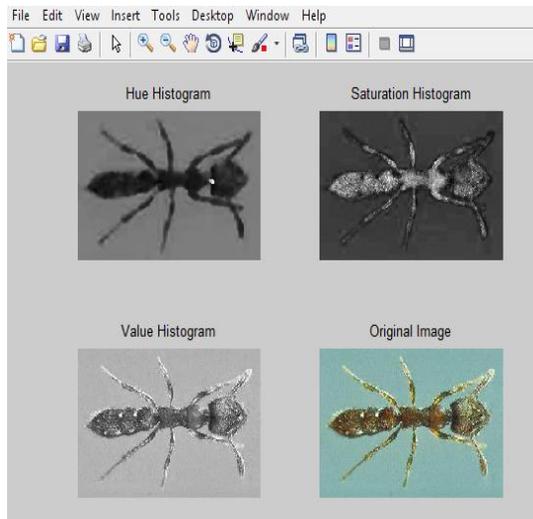$$V = \left[\frac{R+G+B}{3}\right] \qquad (6)$$

Fig. 3. RGB to HSV color conversion

2) Eliminate the least significant pixels of the original image (*I*) by applying the following steps to create rectangular mask with proper dimension. Figure 4 illustrates the cropping process.

- Identify cropping starting points (x , y) by applying the following formulas:

$$x = \frac{1}{8} * original_{width} \qquad (7)$$

$$y = \frac{1}{8} * original_{height} \qquad (8)$$

Note*:* different scaling factors were tested and $(\frac{1}{8})$ gives the best results.

- Identify the new width and new height using the following formulas:

$$newWidth = original_{width} - x \qquad (9) \qquad (9)$$

$$newHeight = Oroginal_{height} - y \ (10) \qquad (10)$$

- Crop the original image using (x, y, newWidth, and newHeight).
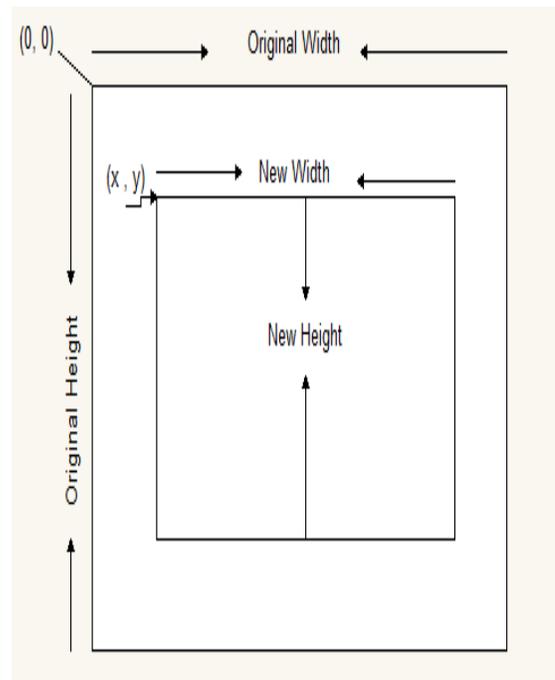

Fig. 4. Cropping process

3) Apply global color histogram to the cropped image in (2) using (H-8 bins, S-2 bins, and V-2 bins)

b) Apply steps (1-3) to query image.

c) Do apply bin to bin comparison between image query and image database then:

- Calculate the similarity based on Euclidean distance measure.

- Consider the most similar images- using proper threshold value as candidate images to be used in Phase-II.

B. Phase II: Joint Histogram and Gray Level Co-occurrence Matric (GLCM)

1) Resize candidate images (CI) retrieved via Phase-I to be 200 x 200 pixels.

2) Starting from the Top-Left Corner, divide each image in (1) to 3-regions as follows:

R1= 50 x 50 Pixels.
R2= 100 x 100 Pixels.
R3= 200 x 200 Pixels.

3) $\forall R_i \in CI,\ i = 1, 2, 3$. Do extract:

- $[\mu_h, \mu_s, \mu_v\ ]$, mean values for H, S, and V.

- $[\sigma_h, \sigma_s, \sigma_v]$, standard deviation values for H, S, and V.

- $[E_h, E_s, E_v\ ]$, entropy values for H, S, and V.

This step generates 9-D feature vector for each region. Total number of features is 27-D for each image. R3 represents the whole image.

4) Texture based retrieval :

$\forall R_i \in CI$ Do:

- Apply GLCM (gray-level co-occurrence matrix) to the image.

- Normalize GLCM.

- Extract the following features:

a) Contrast using the following formula:

$$Cont = \sum_{i,j=0}^{N-1} P_{ij}(i-j)^2 \qquad (11)$$

b) Correlation using the following formula:

$$Corr = \sum_{i,j=0}^{N-1} P_{ij}\frac{(i-\mu)(j-\mu)}{\sigma^2} \qquad (12)$$

c) Energy using the following formula:

$$En = \sum_{i,j=0}^{N-1}(P_{ij})^2 \qquad (13)$$

d) Homogeneity using the following formula:

$$Homog = \sum_{i,j=0}^{N-1} \frac{P_{ij}}{1+(i-j)^2} \qquad (14)$$

Where, P_{ij} is the probability of ij element of the normalized symmetrical GLCM, N is the number of gray-levels in the image, μ the GLCM mean, and σ is the variance of intensity.

This step generates 12-D feature vector. At the end of this step, each image is represented by 39-D feature vector generated by step 3 and 4.

5) Do apply steps (1-4) to query image.

6) Retrieve similar images by deploying a new way of using K-means clustering as shown in figure 5.
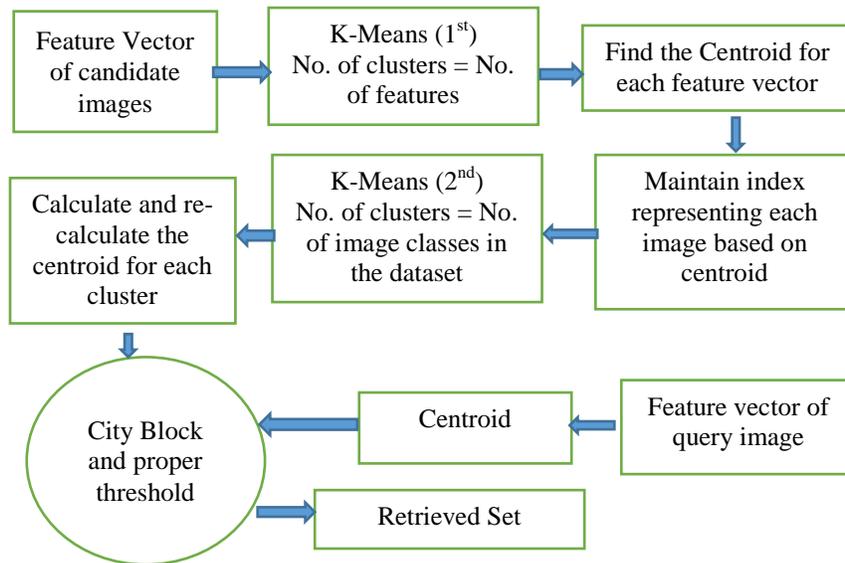

Fig. 5. Retrieving final set using K-means

As stated in [9], K-means as clustering algorithm is one of the unsupervised learning algorithms. It classifies data set into number of clusters in simple way. Users of K-means need to define k centroids based on number of chosen clusters. It is preferable to place these centroids away from each other. Each point in the data set is associated to the proper cluster by calculating the distance between this point and nearest centroid. The process continued by re-calculating k new centroid till no changes are done. To overcome some of problems related to user interference in original K-means

clustering, a two phases K-means clustering is proposed as previously shown in figure 5. In this research number of clusters is chosen based on number of features and image classes, then cluster centroid is calculated and re-calculated based on the feature vector of image data set. As for query, centroid is calculated based on its feature vector and the distance between the centroid of query image and different clusters is measured. Consequently, the image class is identified first and similar images are ranked based on distance from small to large.

V. SIMILARITY MEASURE AND EXPERIMENTAL RESULTS

A. Similarity Measure

The similarity between two feature vectors can be measured using one of the many available distance functions such as Euclidian distance, City block distance, Minkowski, and Mahalanobis [14]. Using different distance functions may slightly increases or decreases precision of retrieval. In this research, Euclidean distance and City block are chosen because of ease of use and they achieve better results as compared with other distance measure. As shown in [15], Euclidean distance function effectiveness and efficiency is noticeable because of that it is considered as one of the most common similarity measures used in image retrieval. It calculates the distance between two feature vectors as follows:

$$D_{euclidean} = \sqrt{\sum_{i=1}^{n}(|I_i - D_i|)^2} \qquad (15)$$

City block distance is calculated as:

$$D_{city} = \sum_{i=1}^{n}|I_i - D_i| \qquad (16)$$

Where, *I* and *D* are two feature vectors of size n.

To measure the effectiveness of any CBIR system, the precision and recall are the most common parameters to be considered [16]. Precision is defined as follows:

$$Precision = \frac{No. of\ Releveant\ Retrieved\ Images}{No. of\ All\ Retrieved\ Images}$$

While, recall is defined as:

$$Recall = \frac{No. of\ Relevant\ Retrieved\ Images}{No. of\ All\ Relevant\ Images\ in\ the\ Catogery}$$

Precision measures accuracy of the retrieved set as compared to the query image, while the recall measures the ability of the system to retrieve as much as possible of similar images related to query image from the database. Precision and recall are affected by the threshold values because number of retrieved images may increases or decreases based on these values. If the accuracy or the precise matching is our target, then the threshold value (tolerance) is chosen to be small. As a result, the number of retrieved images will be less. Whereas,

for the general purpose CBIR systems, the similarity rather than the precise matching is the target of such systems. Consequently, threshold values are chosen to be large. In the following section, precision and recall are measured based on different threshold values.

B. Experimental Results

The proposed system is implemented using MATLAB R2010a and WANG database [3]. WANG database containing 1000 JPEG Corel images of 10 categories, each with 100 images. These categories as buses, elephants, flowers, and horses represent wide range of image classes. Categories such as dinosaurs, and flowers contain objects in contrast with the background, while food and Africa contain objects that mixed with background. As stated in [13], this database is widely used because of reasonable size and classes categorizations.

1) Experiment One

In order to test the proposed approach many images randomly selected as query images. Figure 6 shows sample of test images belong to different categories available in WANG data set.

Fig. 6. Test images

Precision and recall are affected by the number of retrieved images which is increases or decreases based on the chosen threshold values. Figure 7 shows sample of retrieved set, image query which belongs to africa category is shown at the upper left corner. Precision of the retrieved set is 88% at recall level 21%.

Fig. 7. Sample of retrieved set (Africa)

Figure 8 shows sample of retrieved set, image query which belongs to dinosaur's category is shown at the upper left corner. Precision of the retrieved set is 92% at recall level 22%.

Fig. 8.   Sample of retrieved set (Dinosaurs)

Figure 9 shows the experimental results of the proposed approach at different recall levels. It is reasonable that the precision is affected by recall level since as recall level goes higher, the retrieval of unrelated images are increased.
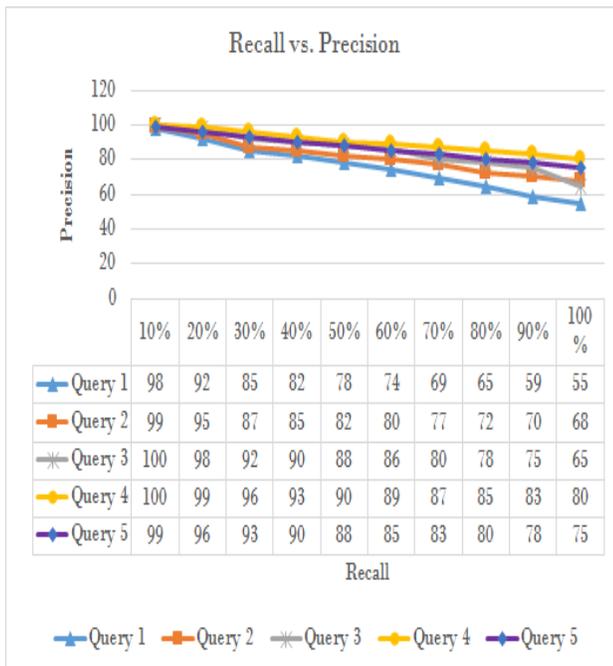

Fig. 9.   Recall vs. Precision

The average precision is calculated for each test image and the overall average precision is measured to show the retrieval accuracy of the proposed system. The result is compared with two of the most relevant approaches as it can be seen in table 2.

TABLE I.   1ST COMPARISON WITH SOME OF THE RELATED CBIR METHODS

	1st Method [17] Yossef *et. al* [2012]	2nd Method [18] Jadhav and Pail [2012]	Proposed Method
Average precision	0.703	0.525	0.84

Table 2 shows that the proposed approach achieves higher average precision rate as compared with [17] and [18]. Phase-I as part of the proposed approach enhances the precision dramatically since filtration process based on color global features normally precise and eliminates unwanted candidate images to be considered in Phase-II.

2) Experiment Two

Once again WANG database is used in the second experiment. In Giveki *et. al.* [19] and Jehad [20], they tested all categories of image database available in WANG. In order to be consistent with the proposed approaches in [19] and [20], all categories of image database are tested with the same test images used in [20]. These categories are (Buses, Elephants, Flowers, Horses, Dinosaurs, Buildings, Food, Mountains, Beaches, and Africa) each with 100 images. Test images are shown in Figure 10.
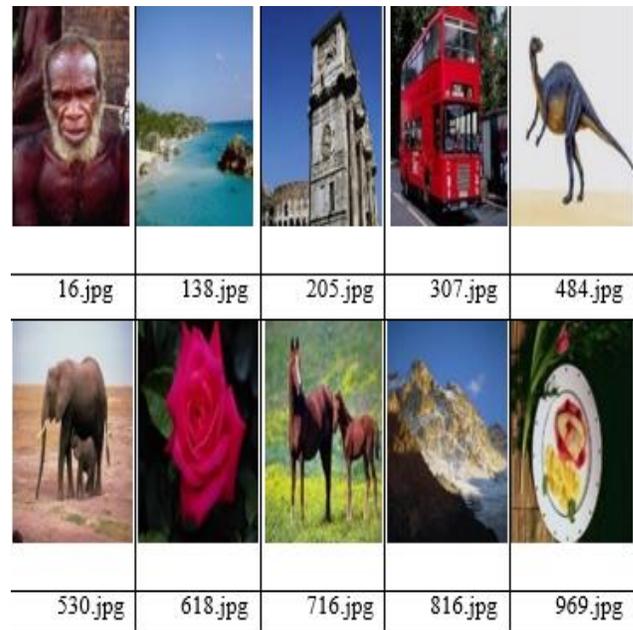

Fig. 10. Test images (2nd experiment)

Figure 11 shows sample of retrieved set, image query which belongs to horse's category is shown at the upper left corner. Precision of the retrieved set is 92% at recall level 22%.

Fig. 11. Sample retrieved set (2nd Experiment)

Figure 12 shows the precision of retrieval to the image query (138.jpg) which belongs to the Beach category. The precision was measured at different recall levels and the average precision is 80%.

Fig. 12. Image class (Beaches)

Figure 13 shows the precision of retrieval to the image query (816.jpg) which belongs to the Mountains category. The precision was measured at different recall levels and the average precision is 75%.
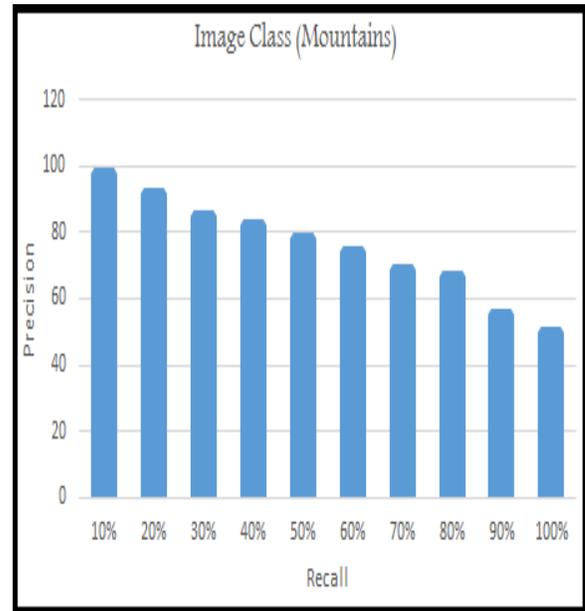

Fig. 13. Image class (Mountains)

Figure 14 shows the precision of retrieval to the image query (484.jpg) which belongs to the Dinosaurs category. The precision was measured at different recall levels and the average precision is 99%.

Fig. 14. Image class (Dinosaurs)

It is observed that, the proposed approach is succeed to achieve high precision rate with image categories in which objects are mixed with the image background such as Mountains, Beaches, and Africa and it achieves better precision rate with image categories in which objects are in contrast with the image background such as Buses, and Dinosaurs. Because of space limitations the results of testing are summarized in the following figure. Figure 15 presents the average precision of test images. Moreover, the overall average precision is shown.
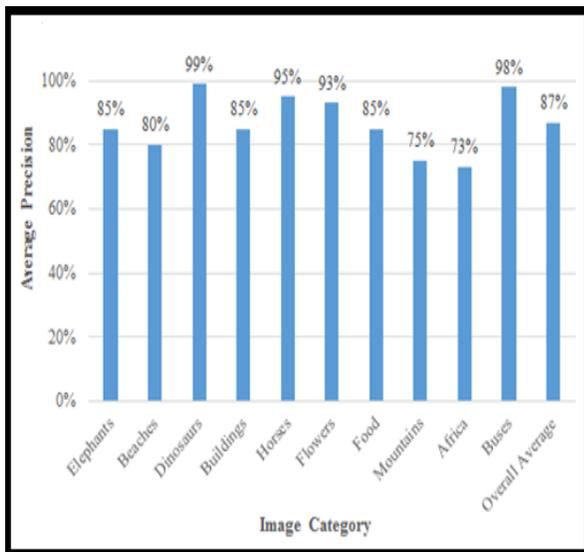
Fig. 15. Average precision of the proposed approach (2nd Experiment)

In [19], they propose CBIR system using wavelet transform. They tested different color models, the best results are obtained using HSV-wavelet with average precision 64%. In [20], a two levels CBIR approach is proposed utilizing color features and 2D-DWT (wavelet transform). The best results are obtained in the 2nd level with average precision 83%. Accordingly, the proposed approach in this paper achieves 87% overall average precision which is considered higher than the previous approaches.

VI.   CONCLUSION

In this paper a new CBIR system utilizing color and texture features is proposed. HSV color space is deployed and color global and local features are extracted in two phases. Initially, color histogram is used in the first phase with the cropped images. As a result the number of candidate images to be considered in phase two are reduced dramatically. Dividing images to regions or sub-images and extract features using joint histogram and gray-level co-occurrence matric (GLCM) is proposed in phase two. Lastly, K-means as clustering algorithm is utilized in an innovative manner to classify and retrieve images. WANG database is used to test the proposed approach and the results obtained is better as compared with the most relevant approaches. Exploring techniques related to shape features extraction and integrating them with the proposed system is left as a future work.

REFERENCES

[1]   K. Iqbal, M. O. Odetayo, and A. James, "Content-based image retrieval approach for biometric security using colour, texture and shape features controlled by fuzzy heuristics," Journal of Computer and System Sciences, vol. 78, no. 4, pp. 1258-1277, 2012.

[2]   M. Nascimento, V. Sridhar, and X. Li, "Effective and efficient region-based image retrieval," Journal of Visual Languages and Computing, vol. 14, no. 2, pp. 151-179, 2003.

[3]   C. Horng Lin, R. Tai Chen, and Y. Kuan Chan, "A smart content-based image retrieval system based on color and texture feature," Image and Vision Computing, vol. 27, no. 6, pp. 658-665, 2009.

[4]   C. Man Pun, and C. Fong Wong, "Fast and robust, color feature extraction for content- based image retrieval," International Journal of Advancements in Computing Technology, vol. 3, no. 6, 2011.

[5]   M. J. Swain, and D. H. Ballard, "Color indexing," International Journal of Computer Vision, vol. 7, no. 1, pp. 11-32, 1991.

[6]   M. Stricker, and M. Orengo, "Similarity of color images," In SPIE Storage and Retrieval for Image and Video Databases III, vol. 2420, pp. 381- 392. 1995.

[7]   R. O. Stehling, M. A. Nascimento, and A. X. Falcao, "Cell histograms versus color histograms for image representation and retrieval," Knowledge and information Systems, vol. 5, no. 3, pp. 315-336, 2003.

[8]   C. C. Gotlieb, and H. E. Kreyszig, "Texture description based on co–occurrence matrices," CVGIP, vol. 51, no. 1, pp. 70–86, 1990.

[9]   S. Sridhar, and M. Gowri, "Color and Texture Based Image Retrieval", ARPN Journal of System and Software, vol. 2, no. 1, pp. 1-6, 2012.

[10]   N. Jhanwar, G. Chaudhuri, G. Seetharaman, and B. Zavidovique, "Content based image retrieval using motif co-occurrence matrix," Image and Vision Computing, vol. 22, no. 14, p.1211-1220, 2004.

[11]   S. M. Youssef, "Ictedct-cbir: Integrating curvelet transform with enhanced dominant colors extraction and texture analysis for efficient content-based image retrieval," Computers and Electrical Engineering, vol. 38, no. 5, pp. 1358-1376, 2012.

[12]   R. Nag, M. Aatif, A. Kazi, and Z. Khan, "Color and Texture Based Image Retrieval," International Journal for Research & Development in Technology, vol. 7, no. 4, pp. 724-729, 2017.

[13]   Y. Mistry, D. T. Ingole, M. D. Ingole, "Content based image retrieval using hybrid features and various distance metric", Journal of Electrical Systems and Information Technology. 2017, http://dx.doi.org/10.1016/j.jesit.2016.12.009.

[14]   E. U. Scott, Computer vision and image processing. International Edition, Prentice-Hall, Inc, 1998.

[15]   F. Malik, and B. Baharudin, "Analysis of distance metrics in content-based image retrieval using quantized histogram texture features in the dct domain", Journal of King Saud University Computer and Information Sciences, vol. 25, no. 2, pp. 207-218, 2013.

[16]   H. Muller, W. Muller, S. Marchand-Maillet, T. Pun, and D. McG. Squire, "A framework for benchmarking in visual information retrieval," International Journal on Multimedia Tools and Applications, vol. 21, no. 1, pp. 55-73, 2003.

[17]   S. M. Youssef, Mesbah S., Mohmoud Y. M., "An efficient content-based image retrieval system integrating wavelet-based image sub-block with dominant colors and texture analysis", In 8th International Conference on Information Science and Digital content Technology (ICIT), vol. 3, IEEE, pp. 518-523, 2012.

[18]   S. M. Jadhav, V. Patil, "An efficient content-based image retrieval (CBIR) system based on evolutionary programming (EP)", IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT), pp. 310-315, 2012.

[19]   D. Giveki, A. Soltanshahi, F. Shiri, and H. Tarrah, "A new content based image retrieval model based on wavelet transform", Journal of Computer and Communications, vol. 3, pp. 66-73, 2015. http://dx.doi.org/10.4236/jcc.2015.33012.

[20]   J. Q. Alnihoud, "Image Retrieval Model Based on Color and Shape Features Utilizing 2D Wavelet Transform", World of Computer Science and Information Technology Journal (WCSIT), vol. 7, no. 4, pp.20-25, 2017.