Image Retrieval using Multi Cascaded Features

Hashem B. Jehlol
Al-Mustansiriyah University
Computer Center

Mayyadah Jabbar Gailan
Al-Mustansiriyah University
College of Touristic

Anwer Subhi Abdulhussein Oleiwi
Al-Mustansiriyah University
Avicenna Center for E-learning

ABSTRACT
This paper depends on Content Based Image Retrieval CBIR to retrieve desired images from a large images database. It is based on extracting cascade features from image such as color, shape and texture features. The Euclidean Distance is used to measure similarity of each feature and retrieve similar images from the image database. The proposed method contains three stages: in the first stage, the HSV color histogram from query image is found and the nearest 100 images are retrieved. In the second stage the shape features are extracted from edge detection and the nearest 50 image are retrieved. In the last stage the texture feature are found based on first order features and the nearest 10 images are retrieved. In this paper a mechanism for image retrieval based on cascading approach is developed to improve image retrieval performance and reduce the computational time required to retrieve images. It is found that cascading features with Euclidean distance give about 76% precession of image retrieval.

Keywords
CBIR, Euclidean Distance, HSV color histogram, shape features, texture feature.

1. INTRODUCTION
In recent years, the image database has a great role because of the tremendous development in Internet and computer technologies. With the fast evolution of networks and computers, the transmission and storage of a huge number of images became possible. In these databases, the need of seeking matching images for the query images are boring and time exhausted tasks [1]. In the last decade, the images retrieved are widely desirable instead of the retrieved text. Image retrieval of based content is considered as one of the more effectual ways to reach visible data [2]. Image retrieval of based content became a prominent subject matter because of the increasing image data and video in digital format. In traditional images database, image as text-annotate and images retrieval depend on keywords seeking [3]. Content Based Image Retrieval is used widely to describe the method of retrieving favorable images from a big collection of databases, on the base of syntactical images features texture, shape and color [4].

Content Based Image Retrieval CBIR denotes to image contents that are used to retrieve a desired image directly, by which the image with specific features or containing specific content will be searched in an image database. The image features are the basis for CBIR, which are certain visible characteristics of images. They are either local for a little group of pixels or global for the full images. The image features can be categorized into two levels: high level features and low level features. The high level features are used to remove the semantic gap among the information extracted from the visible data and the interpretations that the same information have for users in a given position. The low level features are used to remove the sensory gap among the objects in the world and the information in a description extract from a registration of that scene [5].The analysis of Image content by using low level features of image is a major idea of CBIR. Examples of low-level features are used in this paper color, shape, texture features.

In CBIR method, it is usual to classify the images features in three major classes: shape, color and texture features. Colors are by far the most public visible features used in Content Based Image Retrieval, mainly because of the ease of extracting colors feature from image. The color histogram explains the level-gray or distribution colors for a given image, colors computed efficient, but in general insensitive to tiny changes in camera location. Colors histogram also has several restrictions. Color histograms provide no spatial info; it simply explains which colors exist in the images, and in what amount. Shape is a significant visible feature and it is one of the fundamental features used to explain image contents. But shape description and representation is a hard task [6]. Textures are also considered as an attribute of the features extraction by several researchers. In spite of there is no official definition for texture, taking a look into this descriptor provides measuring of the characteristics such as regularity, smoothness and coarseness. Mostly the texture features extraction of images is analyzed during structural, spectral and statistical methods [4].

The organization of this paper is as follows. A brief overview about previous research in the field of image retrieval is presented in Section 2. Section 3 presented the proposed methods that used in the research. Section 4 presented an overview of the experimental setup. Section 5 contains experimental results and discussions of the results. Finally conclusions are presented in section 6.

2. RELATED WORK
Suhasini P. S. et al [7]. Proposed method aims to retrieval images using different method. They used three color histogram, fuzzy color histogram, invariant color histogram and Conventional Color Histogram (CCH) based on color extracted and compared between them. The computation of (CCH) is easy and simply. Fuzzy linking Color Histogram (FCH) is used by determine the problem of spatial relationship while invariant color histograms depended on color gradient is used by determine the problems of the translation and rotation.

Pujari J. et al [8]. Proposed a novel method based on color and shape features that used various extracted components from the images such as Lab and HSV. In terms of performance, this method is compared with RGB and gray method and gives best results. The training images in this work belong to WANG database.

Sridhar G. [9]. Proposed a method deals with color histogram and Gray-Level Co-occurrence Matrices texture features to extract features vectors from the images and used a K-mean to

clustered image and stored in the database to retrieve a similar image and compare with color and texture retrieved images efficiency.

P. V. N. Reddy et al [10]. Proposed a method depends on integrated color and texture for extracting features of images to improve retrieval method. The Multi Wavelet transform indicates the information of image texture and the color space (RGB) indicates the color features. The integration of texture and color features in the trial results is more efficient than the method used in the single visual feature.

R. Malini et al [11]. They proposed retrieval system gives high performance and efficiency. They based on combine color content by using average mean and feature extraction by using central tendency. The features vector is decreased in this method. WANG images database used in the training images and testing images. The results of the system were compared with others and gave better results.

Ramesh K. Lingadalli et al [12]. Proposed an approach based on images content and used features like color, texture and shape which enhance the performance and accuracy of images retrieval. The color features are extracted using color histograms and texture features were extracted using grey level co-occurrence matrix (GLCM). The fast retrieval of shape can be improved by considering approximate shape instead of accuracy shape.

3. PROPOSED METHOD

In this paper, the presented method consists of three cascade stages .In the first stage the color features are extracted according to HSV histogram values. Next stage, the shape features are extracted from images that resulted from first stage. The final stage in the proposed method is extracted first order features from images that resulted from second stage. The measure similarity is performed by using the Euclidean Distance. Figure 1 illustrates the block diagram of proposed method.

3.1 Color Features

Color histograms are frequently used in the comparison between images [7]. In this paper the histogram of image is computed depend on HSV color space. Therefore, the color images are converted to HSV image in the first step. Next step the histogram for HSV color space is calculated. Images are

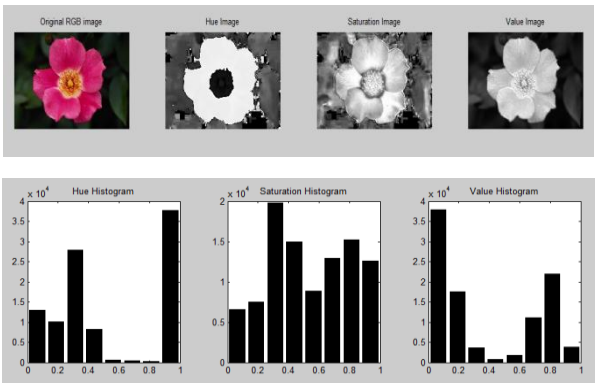

Figure 2: Convet RGB Image to Hue Image ,Suturation Image and Value Image and Found Histogram for Each.

retrieved from the database according to histogram values. Figure 2 show converting RGB image to HSV image and find Hue (H) Histogram, Saturation (S) Histogram and Values (V) Histogram.

The image retrieval depends on HSV histogram is illustrated in the following procedures:

Step1: Image is loaded from image database.

Step2: Image with RGB color space converted to image with HSV color space.

Step3: Finding HSV Color histogram by quantized into 8 levels each to Hue (H), Saturation (S) and Value (V).

Step4: The resulted HSV histogram is stored in a new vector.

Step5: Repeating steps from 1 to 4 for all images.

Step6: Query image is loaded.

Step7: Repeating steps from 2 to 4 to find HSV histogram of the query image.

Step8: Calculating the similarity of query image and image database.

Step9: Retrieving top 100 images.

Figure 3 represents the block diagram of convert RGB color image to HSV color image and finding HSV histogramfeatures.

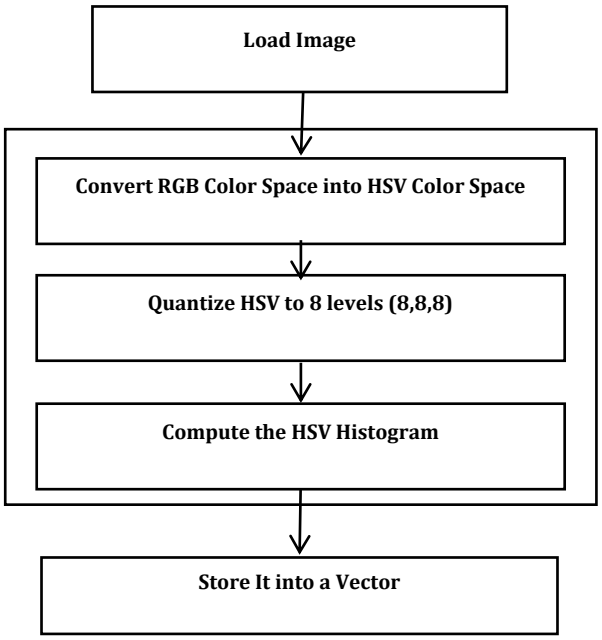

Figure 3: Extract Color Feature of Image

3.2 Shape Features

Shape is a very powerful and an important feature; therefore it is used to retrieve images from database's image [13]. Edge detection histogram is used to extract the shape feature. Canny edge detection is used in the proposed method to get image edge information. Figure 4 shows an example of image with Canny edge detection method.

a **b** **c**

Figure 4: a RGB image, (b) gray scale image , (c) canny image

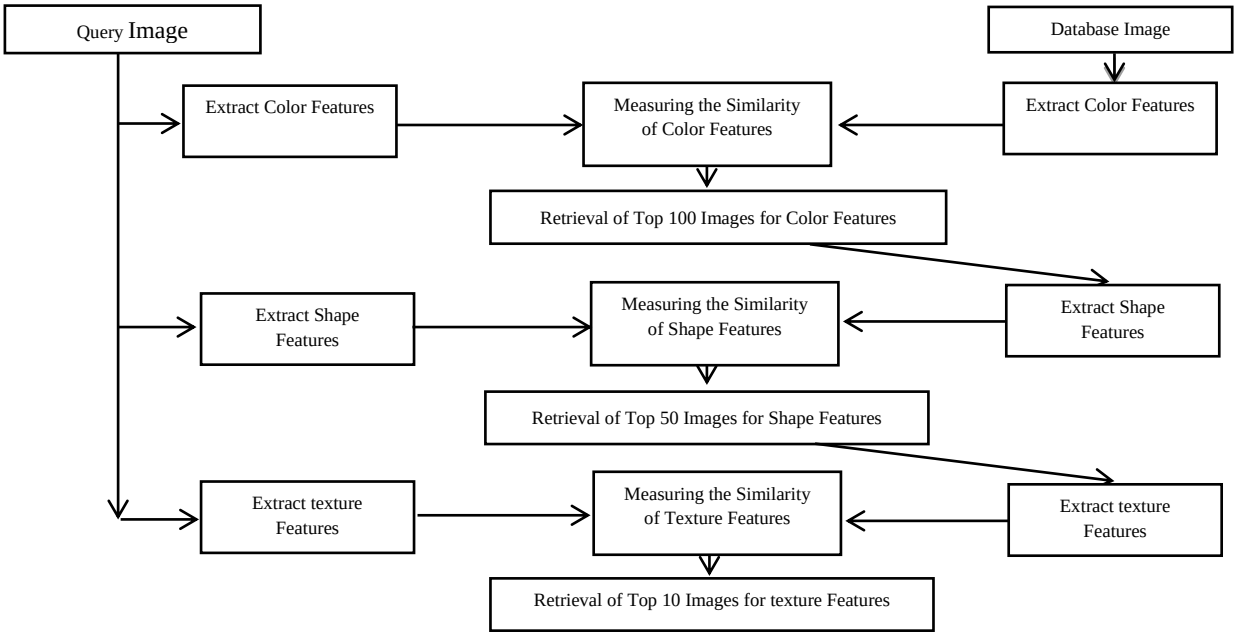

Figure 1: Block Diagram of Proposed Method

The image retrieval depends on the edge detection method is illustrated in the following procedures:

Step1: loading an image from 100 images that resulted from previous stage.

Step2: the image with RGB color space is converted into a grey scale image.

Step3: Canny edge detection technique is implemented to detect the edges and store it into a 2D matrix.

Step4: the sum values of each column and row is taken from the generated matrix.

Step5: storing the values in a new vector.

Step6: Repeating steps from 1 to 5.

Step7: Query image is loaded.

Step8: Repeating procedure steps from 2 to 5 to find shape features of the query image.

Step9: Calculating the similarity of query image and image database.

Step10: Retrieving top 50 images.

3.3 Texture Feature

The texture feature is considered one of the important properties of many types of images. In the proposed method, the texture feature that extraction from image based on first order Statistics. First order Statistics is computed from the original values of image. In the first-order features, the relationships with neighboring pixels are not considered. It gives many features include mean, standard deviation, kurtosis, skewness, and entropy as in the following equations.

$$mean(u) = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N} Ii(x,y)}{M*N} \quad (1)$$

$$standard\ deviation\ (\sigma) = \sqrt{\frac{\sum_{x=1}^{M}\sum_{y=1}^{N} Ii(x,y) - u}{M*N}} \quad (2)$$

$$kurtosis = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N} Ii(x,y) - u^4}{M*N*\sigma^4} \quad (3)$$

$$skewness = \frac{\sum_{x=1}^{M}\sum_{y=1}^{N} Ii(x,y) - u^3}{M*N*\sigma^2} \quad (4)$$

$$entropy = \frac{1}{M*N}\sum_{x=1}^{M}\sum_{y=1}^{N} Ii(x,y)(-\ln Ii(x,y)) \quad (5)$$

In the proposed algorithm, the texture features of a color image are found as follows:

Step1: reading the image from 50 images that resulted from second stage.

Step2: converting the image into a grey scale image.

Step3: Finding texture feature based on first order features.

Step4: Storing it in a new vector.

Step5: Repeating steps from 1 to 4 on all images in the database.

Step6: the query image is loaded.

Step7: performing the procedure 2-4 to find texture feature of the query image.

Step8: Calculating the similarity of query image and image database.

Step9: Retrieving top 10 images.

3.4 Matching Similar Features

Various feature extractions methods are used in this paper. The Distance Metric Measure is used for comparing the various features. To retrieve the similar images from the image database, many types of Distance Metric Measures like Chi-Square Distance, Euclidean Distance, and Weighted Euclidean are applied. In the proposed method Euclidean distance is used. It measures the distance between two vectors of images by calculating the square root of the sum of the squared absolute differences. Euclidean distance is defined as:

$$d = \sqrt{\sum_{i=1}^{k} (|Q_i| - |D_i|)^2} \quad (6)$$

Where Q represents vector features of query image and D represents vector features of database images.

4. EXPERIMENT AND RESULT

The database image used in proposed method is WANG database, and the Matlab program with version 7.3 has been implemented. The WANG database contains 1,000 images in JPEG format. The size of each image is 256x386 and 384x256. It consists of 10 classes such as (Africa, beach, monuments, buses, dinosaurs, elephants, flowers, horses, mountains and food) each class contains 100 images. Figure 5 shows sample of WANG database.

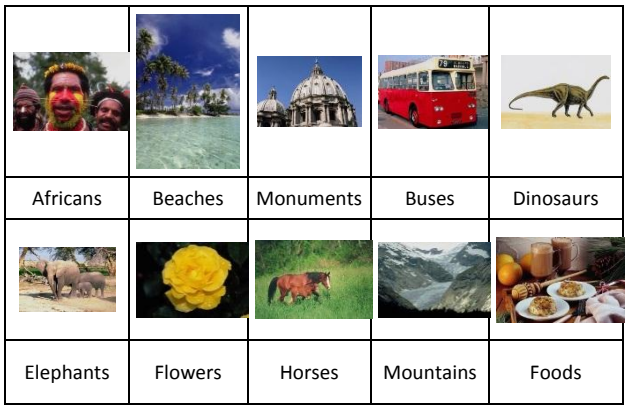

Figure 5: Example Image from each of the 10 Classes of WANG Database

The retrieval effectiveness of the proposed method is measured by using 5 different query images from each class. It is tested for 50 query images. In the beginning, the image size is changed to [128 128] to get a similar size for each image. It is very important to remove noise from the image to get good results in the retrieval images. The proposed method consist of three stages: in the first stage, after converting RGB image to HSV image, the HSV color quantization is implemented using color histogram with quantization schemes (8, 8, 8). The color features are extracted according to HSV histogram values. Euclidean distance is used to measure similarity between query image and database images. Top 100 images are retrieved depending on the minimum distance. Next stage, the shape features are extracted from 100 images that resulted from the first stage and similarity measurements is performed between shape features of query image and the100 images. Top 50 images are retrieved depending on the minimum distance. The final stage in the proposed method is extracted the first order features from 50 images and compared with query image. The nearest 10 images of the query image are retrieved from the image database.

The precision and recall are used to measure performance of retrieval. Recall is used to measure the system's ability to retrieve all the images that are relevant, while precision is used to measure the system's ability to retrieve only the images that are relevant. The equation of the recall and precision are illustrated in the following:

$$precision = \frac{Numer\ \ of\ relevant\ \ images\ \ retrieved}{Total\ \ number\ \ of\ images\ \ retrieved} = \frac{N}{N+B} \quad (7)$$

$$Recall = \frac{Number\ \ of\ relevant\ \ images\ \ retrieved}{Total\ \ number\ \ of\ relevant\ \ images} = \frac{N}{N+C} \quad (8)$$

Where the number of retrieved images is represented with N and the irrelevant images are represented with B, while the C represents the number of relevant images not retrieved. The similarity measurement used in this work is Euclidean distance. Table 1 shows the results of precession when using color, shape and texture features. The best results are obtained in using cascade color, shape and texture features.

Table 1:Average precision of all image classes when using color features only ,cascade color and shape features and cascade color, shape, texture features

Class Name	Precision using Color Features	Precision using Cascade Color and Shape features	Precision using Cascade Color , Shape and texture features
Africa	0.56	0.69	0.7
Beaches	0.62	0.66	0.63
Building	0.6	0.63	0.65
Buses	0.75	0.7	0.85
Dinosaurs	0.83	0.87	0.96
Elephants	0.66	0.62	0.69
Flowers	0.85	0.83	0.79
Horses	0.65	0.79	0.92
Mountains	0.55	0.59	0.66
Foods	0.74	0.65	0.77
Average	0.681	0.703	0.762

Figure 6 shows the precession obtained by the different methods like color only, cascade color and shape features and cascade color, shape and texture features. It shows that the cascaded three features gave the best result in image retrieval.

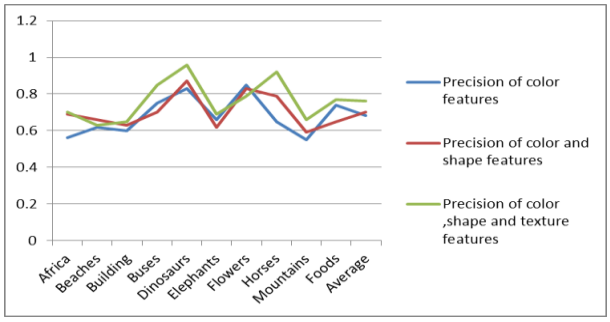

Figure 6: Precision of the Image Retrieval Using different methods

Figure 7 illustrates the Graphical user interface (GUI) of the proposed method to retrieve the images according to the cascaded features.

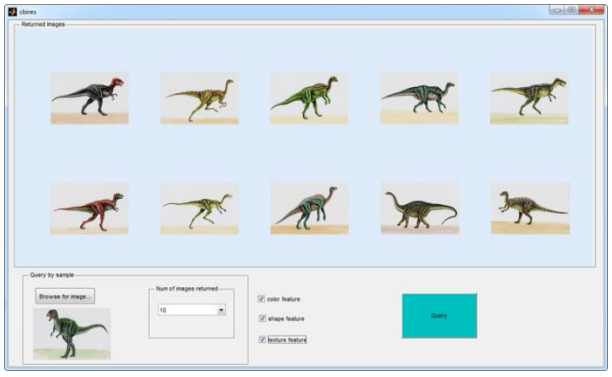

Figure 7: Graphical User Interface of proposed method

5. CONCLUSION

This work depended on the cascade method where three types of features are used. These features include HSV color histogram features, shape features and the texture features represented by six first-order statistical features. The Euclidean distance is used to retrieve similar images from the database. A mechanism for image retrieval based on cascading approach is developed to reduce the computational time and to increase the accuracy of retrieval. Three stages are used to select top 10 images. It is found that the cascading features with Euclidean distance give about 76% precession of image retrieval.

6. REFERENCES

[1] N.Puviarasan and R.Bhavani, "Retrieval of Images Using Weighted Features", International Journal of Advanced Computer Research (ISSN (print): 2249-7277 ISSN (online): 2277-7970) Volume-4 Number-1 Issue-14 March-2014.

[2] JunYue, Zhenbo Li , Lu LiuandZetian Fu,"Content Based Image Retrieval Using Color and Texture Fused Features", School of Information Science and Engineering, LUDONG University, 2010.

[3] S. Mangijao Singh, K. Hemachandran, "Content Based Image Retrieval Using Color Moment and Gabor Texture Feature", IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 5, No 1, September 2012.

[4] Manimala Singha and K.Hemachandran, "Content Based Image Retrieval Using Color and Texture", February 2012.

[5] R. Datta, D. Joshi, J. Li, J.Z. Wang, "Image Retrieval: Ideas, Influences, and Trends of The New Age", ACM Computing Surveys 40 (2) (2008) 1–60.

[6] Reshma Chaudhari and A. M. Patil, "Content Based Image Retrieval Using Color and Shape Features", November 2012.

[7] Suhasini P. S. , K. R Krishna and I. V. M. Krishna, "CBIR Using Color Histogram Processing", Journal of Theoretical and Applied Information Technology, Vol. 6, No.1, pp-116-122, 2009.

[8] Pujari. J, Pushpalatha S.N and Desai. P.D, "Content Based Image Retrieval using Color and Shape Descriptors", International Conference on Chennai, 2010.

[9] Sridhar and Gowr, "Color and Texture Based Image Retrieval", Journal of Systems and Software, VOL. 2, NO. 1, January 2012.

[10] P. V. N. Reddy and K. Satya Prasad, "Color and Texture Features for Content Based Image Retrieval", IJCTA, 2011.

[11] R. Malini and C.V asanthanayaki, "An Enhanced Content Based Image Retrieval System Using Color Features", International Journal of Engineering and Computer Science ISSN:2319-7242, Volume 2 Issue 12 Dec,2013.

[12] Ramesh. K. Lingadalli and N.Ramesh, "Content Based Image Retrieval Using Color, Shape and Texture", International Advanced Research Journal in Science, Engineering and Technology Vol. 2, Issue 6, June 2015.

[13] Raman Maini, Dr. Himanshu Aggarwal," Study and Comparision of Various Image Edge Detection Techniques", International Journal of Image Processing (IJIP), Volume (3): Issue (1), 2009.

[14] Wasim Khan, Shiv Kumar. Neetesh Gupta, Nilofar Khan, "A Proposed Method for Image Retrieval using Histogram values and Texture Descriptor Analysis", International Journal of Soft Computing and Engineering (IJSCE), Volume-I Issue-II, May 2011.