ISSN: 2278 – 7798
International Journal of Science, Engineering and Technology Research (IJSETR)
Volume 2, Issue 10, October 2013

A SPATIAL FEATURE BASED EFFICIENT CBIR SYSTEM FOR COLOR IMAGES

Miss. Laxmi Goswami
Asst. Professor
Dr.C.V. Raman University
Bilaspur (C.G.)

Miss.Amrita Singh
Asst.Professor
Dr.C.V.Raman University
Bilaspur (C.G.)

Aijaz Ur Rahman Khan
Asst. Professor
M.M. College of Technology
Raipur(C.G.)

Abstract-Row sum, column sum and histogram are some of the important spatial feature of any image and represent the pixel distribution of any image. Row sum , column sum and histogram feature of two similar images are nearly same and hence if these feature are combined then a very efficient CBIR system can be designed. This paper present a row sum, column sum and histogram based hybrid approach for designing a CBIR system. Different statistical parameter were computed to show the accuracy of this system.

Keywords**—Row sum, Column sum, Histogram, feature vector, Euclidean Distance.**

I. INTRODUCTION

Content based image retrieval system can be defined as the software tool which is used to search and access digital images in large database [1]. Feature vector of the image are basically low dimensional data which represent the content of the image[2].Some of the most widely used features of the image are color, texture,shape, mean, entropy, standard deviation etc. Since these feature has low dimension and contain all the necessary information about the image therefore these feature can be used for image searching and indexing in a large database of image.Due to low dimensionality,these feature require less storage and less computation power. In the past various CBIR system were proposed which are based on color feature [3][4].Some author also suggested an CBIR system based on computing color histogram[5][6]. J. Berens, G. D. Finlayson and G. Qiu presented a CBIR system based on color histogram in compressed transform domain[7].Some authors [8][9] used color mean or color average for designing CBIR system.In some images shape of the object[10-12] is also one of the important feature and has been used for CBIR system. Texure is also important feature and hence have been used for designing CBIR system[13].Some frequency domain approaches have also been proposed like discrete cosine transform based [14] and wavelet transform based approach[15].This paper describe a very simple and accurate image retrieval system which is based on computing the row sum and column sum of R,G B channel of color image separately along with the histogram of Red channel, Blue Channel and Green Channel. Since Row sum , column sum and histogram of each channel in color image contain very crucial information about the image therefore an efficient CBIR system can be designed using these features.

II.PROPOSED CBIR ALGORITHM

A. Image database Creation

In any CBIR system design, the first part is extracting the features of the image. In this algorithm, row sum, column sum and histogram are used as the spatial features. Since in this approach, coral database has been taken whose image dimension is 384x256 so in order to reduce the computational complexity, all the images are first resized to 256x256.

The steps for spatial feature extraction are as follows-

1. Input the image.
2. Resize the image to dimension 256x256.
3. Extract the Red channel, Green channel and blue channel.

4.Extract the row and column for each channel and store it in a matrix.

4.Calculate the Row sum and column sum for each channel.

5.Compute the histogram of each channel.

6.Prepare the feature database for each image.

Block diagram of database creation is shown in Figure 1.

Since in this algorithm, an image of size 256x256 has been taken therefore there are 256 row sum and 256 column sum is obtained for each channel and represented RSR(row sum for red channel),CSR(column sum for red channel),RSG(row sum for green channel),CSG(column sum for green channel),RSB(row sum for blue channel) and CSB(column sum for blue channel) by and the dimension of histogram for each channel is also 256 and is represented by HR,HG and HB for red, green and blue channel respectively. Feature vector for each image has a dimension 9x256 while the dimension of the image is 256x256.Image feature database for this algorithm is shown in Figure 2

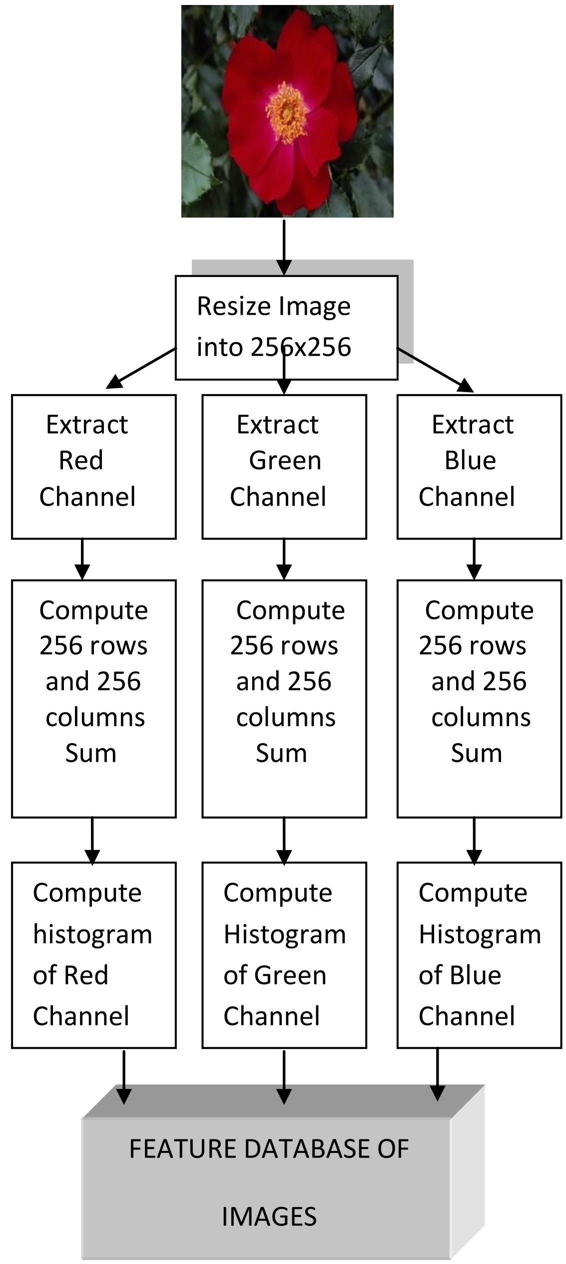


Figure.1 Block Diagram of Image Feature database Creation

Image1

RSR	CSR	HR	RSG	CSG	HG	RSB	CSB	HB
1	1	1	1	1	1	1	1	1
2	2	2	2	2	2	2	2	2
3	3	3	3	3	3	3	3	3
. . . .	. . . .	. . . .	. . . .	. . . .	. . . .	. . . .	. . . .	. . . .
256	256	256	256	256	256	256	256	256

Image2

RSR	CSR	HR	RSG	CSG	HG	RSB	CSB	HB
1	1	1	1	1	1	1	1	

..

Image n

RSR	CSR	HR	RSG	CSG	HG	RSB	CSB	HB
1	1	1	1	1	1	1	1	1
. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .	. . . . .
256	256	256	256	256	256	256	256	256

Image Feature Database

Figure.2 Illustration of Image Feature database

A. *Image Retrieval System*

Once the creation of feature database for all the images are carried out then the next step is to design a CBIR system. The block diagram of proposed method is shown in Figure 3.The proposed system starts by first taking the query image as input and resize it in to the dimension of 256x256.Then this system extract and separate the red channel, blue channel and green channel and the compute the row sum, column sum and histogram for each channel for this query image and prepare the feature vector for the query image. After that, this system compare the feature vector of the query image with the feature vectors stored in database and calculate the similarity measurement using Euclidean distance.All the images for which the Euclidean distance is less than a preset threshold T is considered as the right match for the query image.

III.METRIC USED FOR SIMILARITY MEASUREMENT

Similarity measurement has a very important role in any image retrieval system. Euclidian Distance is most widely used parameter for measuring the similarity between query image and database images. Euclidean Distance between features of query image and database image is computed using equation given below

$$Ed_{q,db} = \sqrt{\sum_{i=1}^{k}(FQ_i - Fdb_i)^2}$$

FQ_i = Feature vector of Query Image

Fdb_i = Feature vector of Database Image

k = No. of features

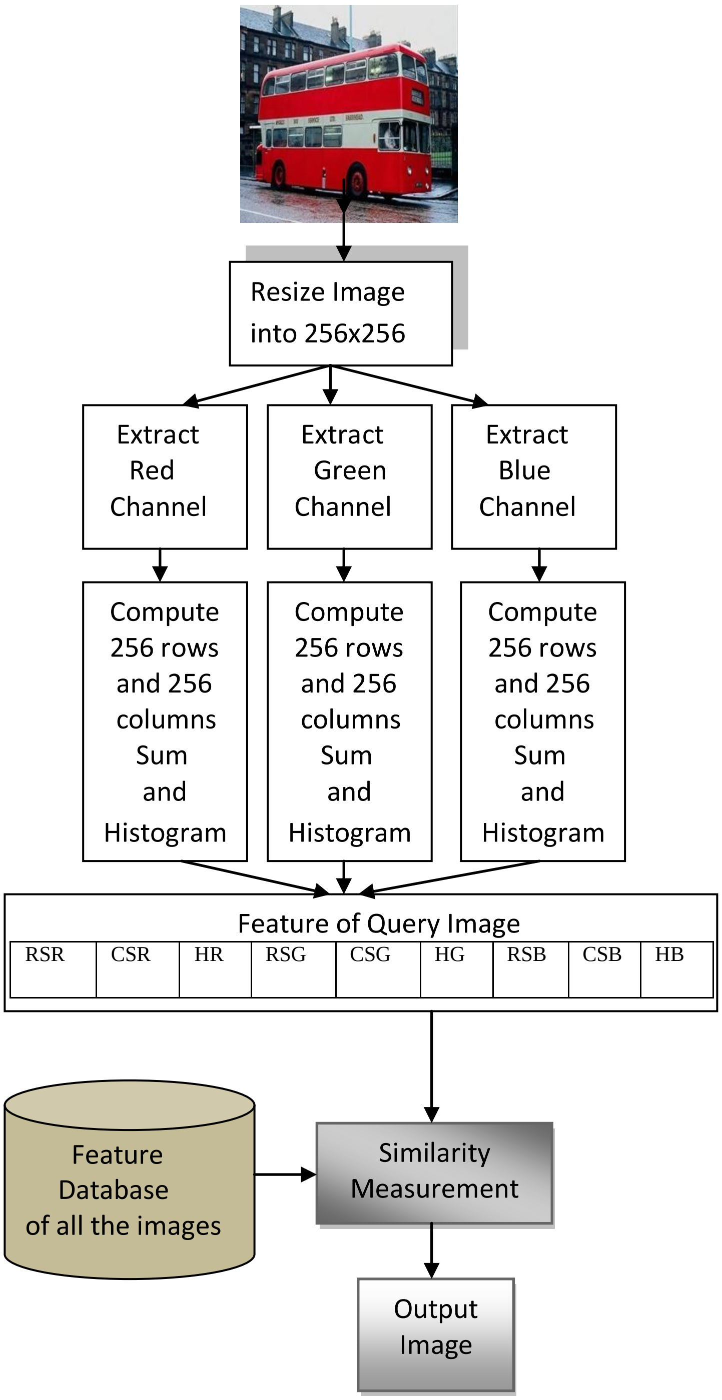


Figure 3 Image Retrieval Process

IV. PROPER THRESHOLD SELECTION

It is very important for any image retrieval system that a proper threshold value T of Euclidean distance must be set. In order to set the proper threshold value T, maximum ED distance between database and query image has been taken as a reference point to calculate the value of threshold T. It was found through extensive experimentation that taking 40% of maximum ED distance produce better result than other value and hence this value is chosen for threshold T.

V. EXPERIMENTAL RESULT

The proposed algorithm was implemented in a system having configuration core 2 duo processor, 2GB RAM.MATLAB (Ver 7.0) platform is used for programming.A database having as many as 800 images of different categories has been prepared from coral database.Images are divided into 8-different categories as shown in table I.

TABLE I

Group ID	Image category
1	Bus
2	Dinosaur
3	elephant
4	Tribal
5	Beaches
6	Flower
7	Horse
8	Food

All the images are first converted in to dimension of 256x256 then feature vectors of all the images has been computed and stored in feature vector matrix as described in section II.In the database,100 images for each category has been taken. A GUI has been designed for this method which take the query image as input and gives five closest match images as output.

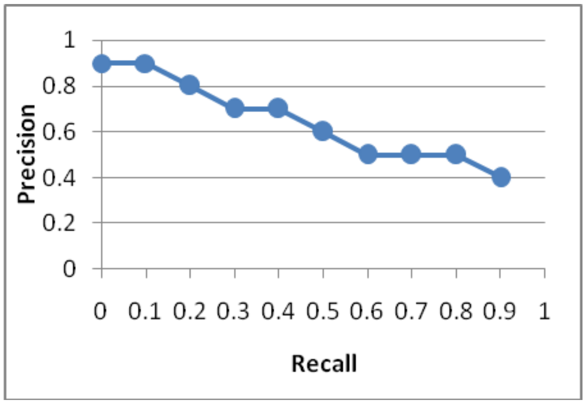


Figure.4 Precision Vs Recall Curve for proposed CBIR

Precision and recall are the two parameter which has been used for testing the performance of proposed system. Average precision has been computed for different image categories and a graph is plotted between average precision and image categories and shown in Figure 5.Moreover precision and recall have also be computed and a plot between precision and recall is drawn and shown in Figure 4.

$$\text{Precision}=\frac{\textit{Number of relevent images retrieved}}{\textit{Total number of images retrieved}}$$

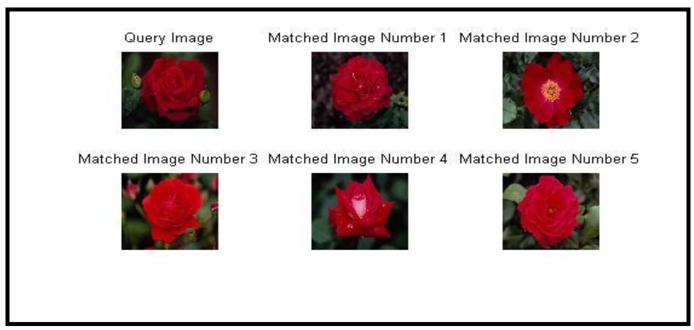
$$\text{Recall}=\frac{\textit{Number of relevent images retrieved}}{\textit{Total number of relevent images in database}}$$

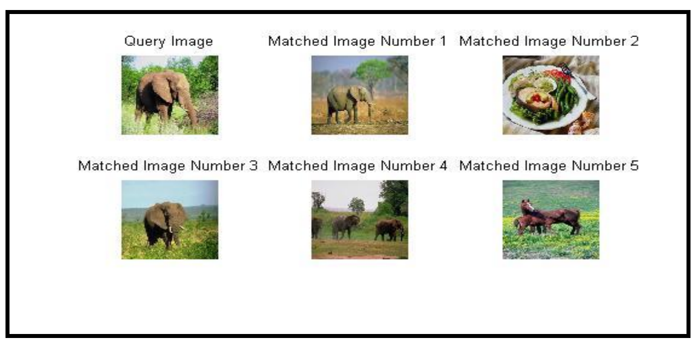
Figure.5 Average Precision Curve for proposed CBIR

The proposed CBIR system starts by inputting the query image. Then Euclidean distance between feature vector of query image and the feature vectors of database image is conputed. An image is considered as the right match for the query image or belong to the same category of the query image if the ED distance between query image and considered image is less than the preset threshold T

VI. CONCLUSION

Basically, the main criterion for judging the performance of any CBIR system is its efficiency and speed. From the result obtained by extensive experimentation and the computed statistical parameter, it is clear that the CBIR system proposed in this paper is able to retrieve the right image efficiently and accurately from the database of large image. It has also been observed that this system retrieve the right image in less than a 3 second from the database of 800 images which clearly shows that it is very fast in speed.





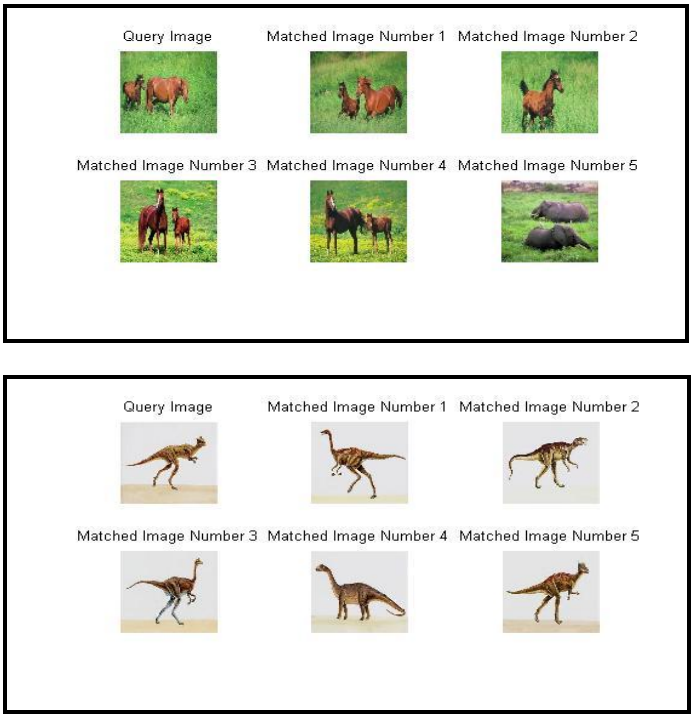


Figure 6 Some output of the proposed CBIR system

References

[1] MarjoMarkkula, Marius Tico, BemmuSepponen, KatjaNirkkonen and EeroSormunen, "A Test Collection for the Evaluation of Content-Based Image Retrieval Algorithms - A User and Task-Based Approach", Published in Information Retrieval 4(3/4), 275-294 (2001).

[2]. P. S. Hiremath, Jagadeesh Pujari, "Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement".

[3] Hossein Nezamabadi-pour and Saeid Saryazdi, "Object-Based Image Indexing and Retrieval in DCT Domain using Clustering Techniques", Vol. 3 JANUARY 2005 ISSN 1307-6884.

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

[5]. Hirata K. and Kato T. "Query by visual example – content-based image retrieval", In Proc. Of Third International Conference on Extending Database Technology, EDBT"92, 1992, pp 56-71.

[6]. John Eakins, Margaret Graham, "Content Based Image Retrieval", Chatpter 5.6, pg 36-40, University of Northrumbia at New Castle, October 1999.

[7] J. Berens, G. D. Finlayson and G. Qiu, "Image indexing using compressed color Histograms", IEEE Proc.-Vision Image Signal Process. Vol. 147, No. 4, August 2000.

[8]. Dr.H.B.Kekre, Sudeep D. Thepade, Archana Athawale, Anant Shah, Prathmesh Verlekar, Suraj Shirke,"Energy Compaction and Image Splitting for Image Retrieval using Kekre Transform over Row and Column Feature Vectors", International Journal of Computer Science and Network Security (IJCSNS),Volume:10, Number 1, January 2010, (ISSN: 1738-7906) Available at www.IJCSNS.org.

[9]. Dr.H.B.Kekre, Sudeep D. Thepade, Varun K. Banura, "Amelioration of Colour Averaging Based Image Retrieval Techniques using Even and Odd parts of Images", International Journal of Engineering Science and Technology (IJEST), Vol. 2, Issue 9, Sept. 2010. pp. (ISSN: 0975).

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

[11] A. K. Jain and A. Vailaya, "Image retrieval using color and shape", Pattern Recognition, 996, vol.29, no.8, pp.1233-1244.

[12] F. Mokhtarian and S. Abbasi, "Shape similarity retrieval under affine transforms", Pattern Recognition, 2002, vol. 35, pp. 31-41.

[13] 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.

[14] Jose A. Lay and Ling Guan, "Image Retrieval Based on Energy Histograms of the Low Frequency DCT Coefficients," IEEE 0-7803-5041-3/99, 1999.

[15] Stepan Obdrzalek and Jiri Matas, "Image Retrieval Using Local Compact DCT-based Representation." DAGM'03, 25th Pattern Recognition Symposium, September 10-12, 2003.