IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-5, Issue-10 E-ISSN: 2347-2693

Efficient Content Based Image Retrieval Using Fuzzy Approach

Dahale Sunil V[1*], Thorat S.B.[2], P.K. Butey[3], M.P. Dhore[4]

[1*] Mahatma Gandhi Mission's College of Computer Science and IT, Nanded, India
[2] Director, SSBES Institute of Technology and Management, Nanded, India
[3] Head of Department CS, Kamla Nehru Mahavidyalaya, Nagpur, India
[4] Dept. of Electronics and Computer Science, Rashtrasant Tukadoji Maharaj (Nagpur University), Nagpur, India

[*]*Corresponding Author: sunildahale@rediffmail.com, Tel.: +91-9860489300*

Available online at: www.ijcseonline.org

Received: 16/Sep/2017, Revised: 28/Sep/2017, Accepted: 18/Oct/2017, Published: 30/Oct/2017

***Abstract*—** content based image retrieval (CBIR) has been one of the significant research areas in computer science in the recent time. Various feature like Texture, color & shape are used in CBIR for image retrieval. Basically, color is the most striking feature in the content based image retrieval. Different extraction methods are used in content based image retrieval (CBIR). A retrieval method which is based on color histogram by using fuzzy inference system approach is proposed in this paper. We propose a fuzzy rule based approach with27 rules, for effective image retrieval. The experimental results show the robustness and the efficiency of the proposed system for CBIR.

***Keywords*—** CBIR, Fuzzy, Fuzzy Inference System ,Color Histogram

I. INTRODUCTION

Internet and networking technologies are very popular concepts among the users worldwide and growing at rapid speed. Internet is used various areas education, biomedicine, entertainment. In the world of internet use of images is very important for communicating information effectively. This is one reason that digital images are used in internet for various purposes like advertising, crime prevention, medical diagnosis, architecture, fashion designing publishing etc. use the images for them is very essential. We can retrieve the images from the database effectively if they are stored in systematic order. In early day's retrieval of image from the database is done by comparing keywords manually. It is not working well if image database is large. The image database sizes are growing rapidly on the internet. So, the problem of storing database is becoming a big headache. To overcome such problem image are stored and classified in various formats like jpeg, mpeg, tiff etc.

In 1990's the concept of Content-Based Image Retrieval (CBIR) has been developed by T. Kato. This technique (CBIR) of retrieving images from large image database is called as content based image retrieval (CBIR). This technique is able to retrieve relevant images effectively. Various techniques are used to store and retrieving images from large image database. CBIR uses query as an image for searching purpose. The searching based on query (i.e. image) improves the performance and accuracy in retrieving relevant images from large database. In this paper we proposed a color histogram with fuzzy inference system using Mamdani method.

II. RELATED WORK

Qin-Jun Qiu, Yong Liu, Da-Wei Cai, Jia-Zheng [1], proposed HSV based texton histogram (HSV-TH) method for CBIR. The proposed HSV-TH method is based on Julesz's textons theory, and it works directly on nature images as shape descriptor and a color texture descriptor. This method combines the advantages of co-occurrence matrix and histogram by representing the attribute of co-occurrence matrix using histogram. The result of this method is tested on the different image databases i.e. Corel and MIT VisTex. The result shows that a significant improvement in terms of their evaluation measures as compared to RGB-TH.

Szabolcs Sergyan [2] proposed a pre-processing algorithm for color histogram features based image classification in content-based image retrieval systems. Image classification for various images is analyzed. This new approach proposed, which based on low level image histogram features. The main advantage of this method is the very quick generation and comparison of the applied feature vectors.
The author proposed a color histogram based classification approach, which is efficient, quick and enough robust. The advantage of this approach is the comparison of histogram features is much faster and more efficient than of other commonly used methods.

R. Venkata Ramana Chary Padmasri Dr.B.V Raju [3] proposed a technique based on color to retrieve image from Large Set of Database. They use combinations of color

feature. The system mainly focuses on visual content of image like color, texture, shape& spatial layouts. The system extracted R, G, and B values for all images separately from the selected 1000 images from image database. Proposed system implemented features like color histogram, color projections Mathematical approaches are proposed for efficient retrieval.

Nikita Raina, Neeshu, Roshi, Rashmi Chauhan [4], Dr. R H Goudar, presented a fused features approach on Content-Based Image Retrieval based on Fuzzy rule-set. They proposed a fast and efficient CBIR approach which is based on color and texture features of an image. The algorithms like Block color histogram for color and Grey level Co-occurrence Matrix (GLCM) and Gabor filter for texture are used in this approach. The similarity between query image and the relevant image is calculated by Euclidean distance which is widely used due to its efficiency. Then, a set of fuzzy rules are applied to get relevant image which is given as a query.

Applications of CBIR

- Crime Prevention
- Medical Diagnosis
- Home Entertainment
- Education
- Advertising and Journalism
- Fashion Designing & Interior design
- Architectural and engineering design
- Web searching
- Military

Problem Description

Color is the main component feature vectors of an image are always considered to be an important attribute in content-based image retrieval system. The common problem in content based image retrieval (CBIR) is selection of features. Retrieving images based on content such as color and texture is still a challenging issue. Color is the abstract information embedded in an image. With the rapid development of technology, the traditional information retrieval techniques based on keywords are not sufficient, content - based image retrieval (CBIR) has been an active research topic. Retrieval efficiency becomes a bottleneck when the matching process is progressed over a large volume of image database. There are three main types of features that are extracted from images: color, texture and shape.

Some times when image is retrieved based on color, lot of images are retrieved. But, with the retrieved images so many irrelevant images are also retrieved this is the problem in content based image retrieval.

As color have been the main image attributes in content based image retrieval systems. We are combining color & soft computing features to obtain higher efficiency in image retrieval from the image database.

Proposed System

- Database created with 1000 images
- Input Query image
- Resize the image into 256*256
- Give the output to fuzzy inference system for color enhancement
- Create a color histogram of every image
- Similarity check of every image in database with query image.
- Retrieve similar images

Flowchart for CBIR

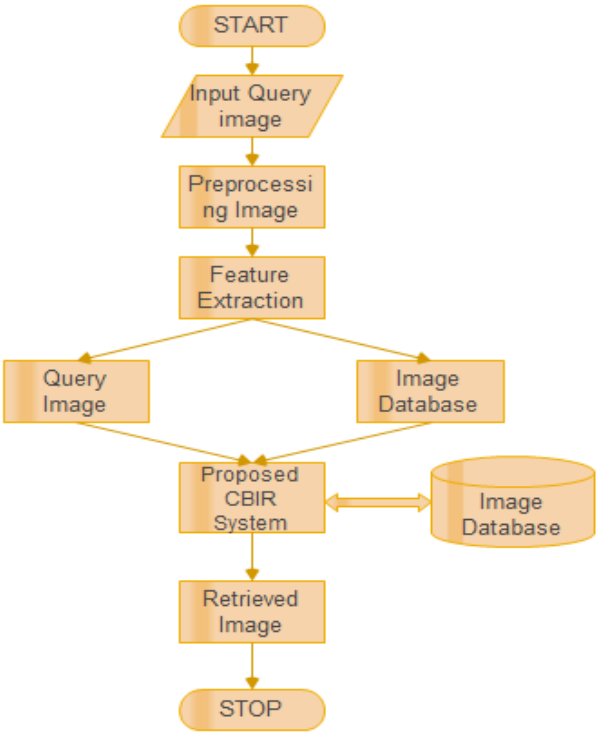

Fig 1: Flowchart of CBIR

Histogram

It was first introduced by Karl Pearson. A **histogram** is an accurate graphical representation of the distribution of numerical data. It is used to an estimate the probability distribution of a continuous variable . It is a type of bar graph.

To construct a histogram, the primary step is to decide the "bin" the range of values. These values divide the entire range of values into a series of intervals and then count how many values fall within which interval. The bins are consecutive and non-overlapping intervals of a variable. The bins must be adjacent, and are often of equal size. A histogram may also be normalized to display "relative" frequencies. [5][6]

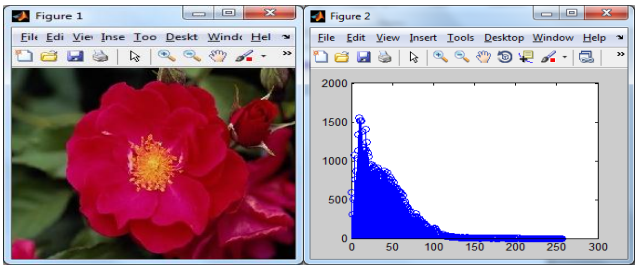

Fig 2: Image Histogram

I. COLOR HISTOGRAM

In image processing and photography, a color histogram is a representation of the distribution of colors in an image. For digital images, a color histogram represents the number of pixels that have colors in each of a fixed list of color ranges that span the image color space, the set of all possible colors. The color histogram can be built for any kind of color space like RGB or HSV. The color histogram is *N*-dimensional, with N being the number of measurements taken.

Statistically, a color histogram is a way to approximate the joint probability of the values of the three color channels. The most mutual form of the histogram is obtained by splitting the range of the data into equally sized bins. Then for each bin, the number of points from the data set here the color of the pixels in an image that fall into each bin is counted. There are different ways of describing the fundamental color distributions.

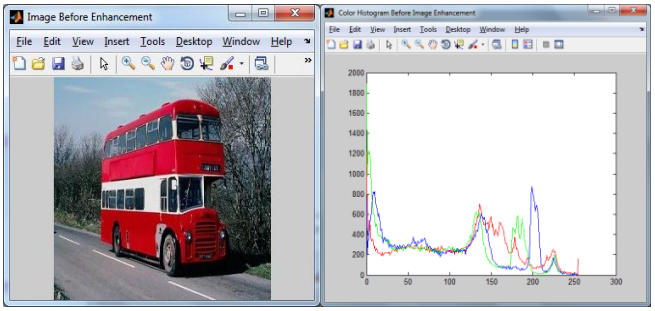

Fig 3: Color Image Histogram

WHAT IS FUZZY LOGIC?

In recent years, the number and variety of applications of fuzzy logic have increased significantly. The applications range from consumer products such as cameras, camcorders, washing machines, and microwave ovens to industrial process control, medical instrumentation, decision-support systems, and portfolio selection.

To understand why use of fuzzy logic has grown, you must first understand what is meant by fuzzy logic.

Fuzzy logic has two different meanings. In a narrow sense, fuzzy logic is a logical system, which is an extension of multivalued logic. However, in a wider sense fuzzy logic (FL) is almost synonymous with the theory of fuzzy sets, a theory which relates to classes of objects with un-sharp boundaries in which membership is a matter of degree. In this perspective, fuzzy logic in its narrow sense is a branch of FL. Even in its more narrow definition, fuzzy logic differs both in concept and substance from traditional multivalued logical systems. [7]

FUZZY INFERENCE SYSTEM

The simulating human reasoning based fuzzy concepts and of inferring with fuzzy actions by employing the rules of inference is capability of inference engine. The process of formulating the mapping form given input to an output using fuzzy logic is called fuzzy inference. Decision can be made on the basis of mapping of given input to an output provided to fuzzy system.

There are various types of fuzzy inference systems that can be implemented. Based on the fuzzy inference systems, fuzzy decision table has been formed through which the output of the fuzzy logic controllers determined.

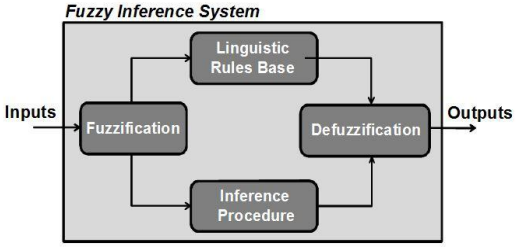

Fig 4: Fuzzy Inference System

The process of fuzzy inference involves:

- Membership Functions
- Logical Operations
- If-Then Rules

Two types of fuzzy inference systems available

- Mamdani-type
- Sugeno-type

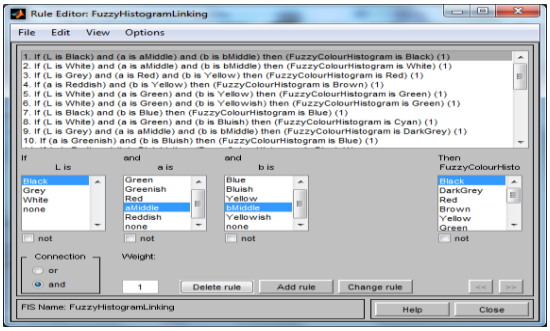

Fig5: Fuzzy Rules

Similarity Measurement:

A. Quadratic Distance Method

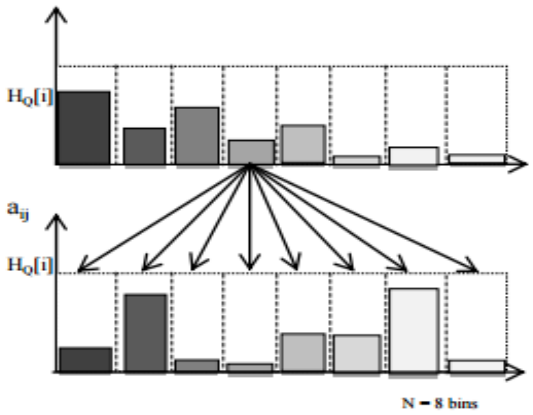

Fig 6: Quadratic Distance Method

Quadratic distance method has been commonly used in CBIR systems which are based on color histogram. We use Quadratic Distance Metric for image classification. By using this method the distance between two color histogram are calculated by following equation

$$d^2(Q, I) = (H_Q - H_I)^t A (H_Q - H_I)$$

.

The equation defines three terms.

The first term (HQ - HI) defines the difference between two color histograms or we can say that it indicates number of pixels differences in each bin. The number of bins in a histogram is defined by number of vector column. (HQ - HI) t defines vector transpose is denoted by the third term. Similarity matrix is defined by the middle term (A). Color distance between the two images is represented by d. The color distance (D) defines the distance between the two images.

When comparing the query image with stored database image, if the both images distance zero, we can get the similar images to the query image, otherwise less color similarity images will have greater distance from zero.

Instead of trying to find exact match, we calculate visual feature similarities between a query image and images in a large database. Therefore the result of retrieval image is not a single image but a list of images ranked by their similarities with the query image. [8][9][10][11]

I. PERFORMANCE MEASURES:

The measure of performance used in image retrieval borrows from the field of information retrieval and are based on two primary figures of metric which is precision and recall. Precision is the number of relevant documents retrieved. Recall is the number of relevant documents in the database which should have been retrieved. [12][13]

$$\text{Precision} = \frac{\text{Number of relevant image retrieved}}{\text{Total number of images retrieved}}$$

$$\text{Recall} = \frac{\text{Number of relevant image retrieved}}{\text{Total number of relevant images}}$$

Precision can be interpreted as a measure of exactness, whereas recall provides a measure of completeness. A perfect precision score of 1.0 means that every retrieved image is relevant but it does not provide any insight as to whether all relevant documents are retrieved. A perfect recall score of 1.0 means that all relevant images is retrieved but says nothing about how many irrelevant images might have also been retrieved.

The performance of the system can be evaluated by using a set of parameters such as

- **Time**
- **Accuracy**
- **Error rate**

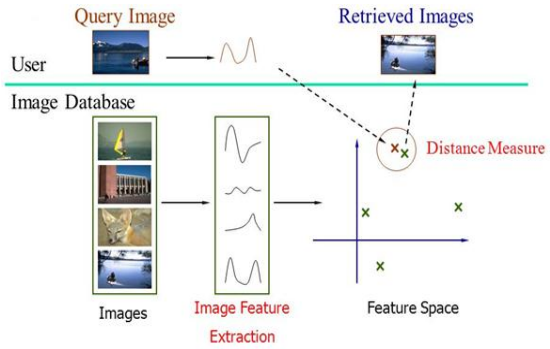

Fig 7: Similarity Measurement (Image courtesy- http://www.slideplayer.com)

Experimental Results:

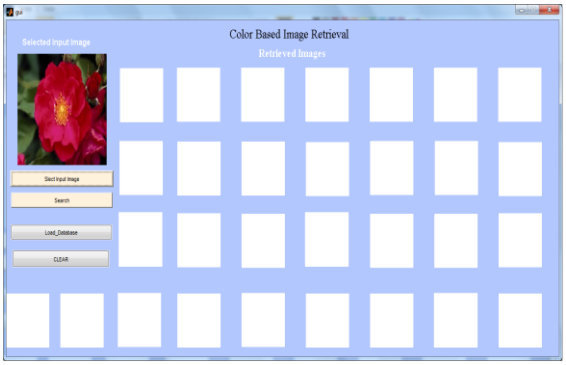

Fig 8: Selection of query Image

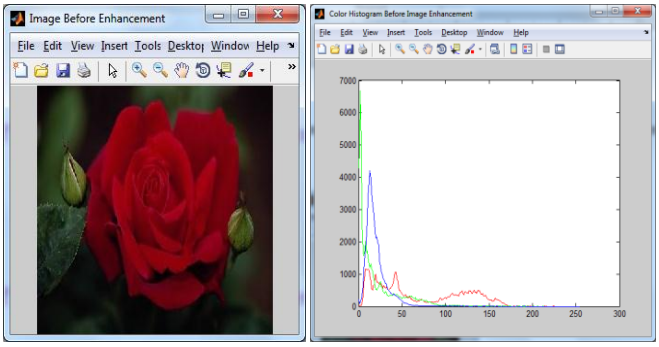

Fig 9: Color histogram before Image enhancement

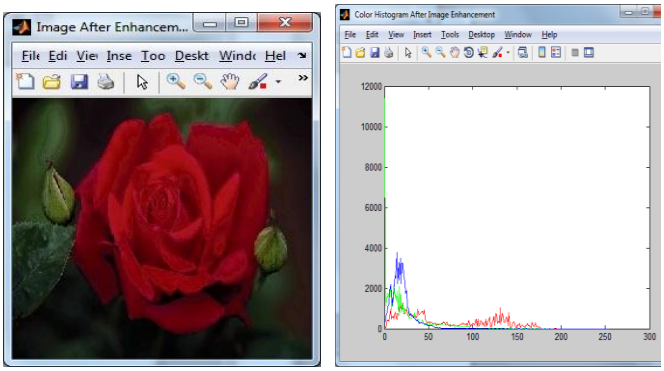

Fig 10: Color histogram after Image enhancement

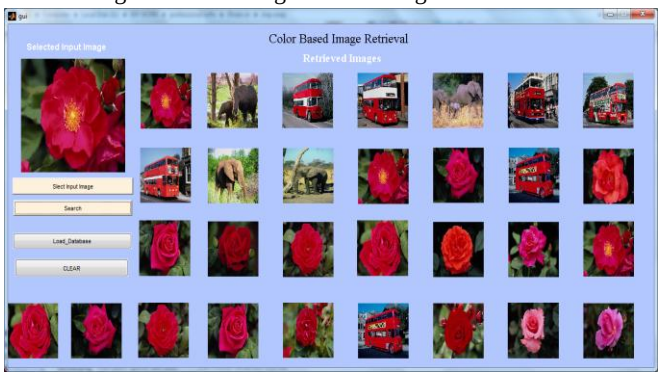

Fig 11: Relevant Image retrieval

Table 1: Comparisons with various techniques

Algorithm	Color Histogram		HSV Color Histogram		Color Histogram Equalization	
Image Type	*Acc (%)*	*Time (sec)*	*Acc (%)*	*Time (sec)*	*Acc (%)*	*Time (sec)*
African	80	10.929	96	62.277	25	31.03
Beach	25	10.071	60	63.598	21	29.755
Buildings	31	10.957	57	60.416	21	28.457
Buses	46	10.689	83	60.69	37	27.952
Dinosaurs	98	10.721	100	58.934	50	28.367
Elephants	39	10.364	47	71.741	35	30.741
Flowers	91	10.286	100	61.726	43	29.027
Horses	51	10.736	94	61.586	18	28.823
Mountains	25	13.572	25	61.258	20	28.595
Foods	42	13.962	85	62.206	25	28.151

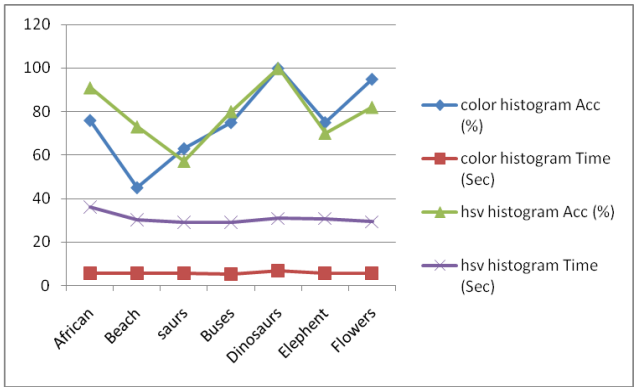

Fig 12 : Comparision chart

CONCLUSION

To improve the accuracy (precision) of a CBIR application by allowing the system to retrieve more images similar to the source image.

In our proposed work, we used color images that is widely used for CBIR. This database contains many images divided into 10 classes where some images for each class. We used it to evaluate the performance of our system by calculating the difference of color histogram between two images with precision and recall. We used the RGB color space and color histogram to represent the images.

This preprocessing step is important to prepare the image for the next step. In the preprocessing step, we make the image enhancements. To improve the performance of CBIR system we can use some additional attribute like shape with color.

FUTURE SCOPE

The recommendation and the future work appear from the limitations and the difficulties when we develop our system. The recommendations and the future works are as the following: ·

- To represent an image, more than one feature is needed. Our system uses the color feature. We recommend the combination of the texture, shape, and spatial features with the color feature to represent the image. This will give good results.
- To further improve the retrieval results, segmentation is a method to extract regions and objects from the image. The segmented regions are used for similarity matching
- To enhance the results, the system has to take the feedback from the user. The user checks the results and comments on them by some way. Then, the system recalculates the results with the advantage of the feedback.

REFERENCES

[1] Nadira T., Rehna K., Fepslin Athish Mon, "*A Study and Analysis on Feature Extraction in Content-Based Image Retrieval*", International Journal of Computer Sciences and Engineering, Vol.5, Issue.6, pp.305-307, 2017.

[2] Sergyan, Szabolcs. "*Color histogram features based image classification in content-based image retrieval systems"*, Applied Machine Intelligence and Informatics, 2008. SAMI 2008. 6th International Symposium on. IEEE, 2008.

[3] P. Rachana, S. Ranjitha, H.N. Suresh, "*Stride Towards Developing an CBIR System Based on Image Annotations and Extensive Multimodal Feature Set*", International Journal of Computer Sciences and Engineering, Vol.2, Issue.1, pp.18-22, 2014.

[4] Raina, Nikita, et al. "*A fused features approach on Content-Based Image Retrieval based on Fuzzy rule-set*", Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on. IEEE, 2014.

[5] www.wikipedia.com

[6] Shraddha S.Katariya and Ulhas B.Shinde, "*Review of Content Based Image Retrieval Using Low Level Features*", International Journal of Computer Sciences and Engineering, Vol.4, Issue.3, pp.91-97, 2016.

[7] http://in.mathworks.com

[8] Alladi Ujwala, Nagaratna P. Hegd, "*Image Retrieval Using Quadratic Distance Based On Color Feature And Pyramid Structure Wavelet Transform And Glcm Based On Texture Feature*", International Journal of Technical Research and Applications e-ISSN: 2320-8163, www.ijtra.com Volume 2, Issue 4 (July-Aug 2014), PP. 169-172

[9] *Zhang D, Lu G.,"Evaluation of similarity measurement for image retrieval",* International Conference on Neural Networks and Signal Processing, 2003. (Vol. 2, pp. 928-931)

[10] Mehtre BM, Kankanhalli MS, Narasimhalu AD, Man GC., "*Color matching for image retrieval*", Pattern Recognition Letters, Vol.16, Issue.3, pp.325-31, 1995.

[11] Aasish Sipani, Phani Krishna and Sarath Chandra, "*Content Based Image Retrieval Using Extended Local Tetra Patterns*", International Journal of Computer Sciences and Engineering, Vol.2, Issue.11, pp.11-17, 2014.

[12] Shambharkar MS, Tirpude MS. "*Content Based Image Retrieval Using Texture and Color Extraction and Binary Tree Structure*", International Journal of Computer Technology and Electronics Engineering (IJCTEE), Vol.5, Issue.6, pp.51-56, 2011.

[13] You FC, Zhang C., "*The technique of color and shape-based multi-feature combination of trademark image retrieval*", IEEE 2010 International Conference on Multimedia Technology, pp. 1-5, 2010.

