Comparison of CBIR Using Texture Feature And Histogram Technique

Mehak Garg*[1], Jyoti Rani[2#], Lalit Jindal[3#]
Computer Science Engineering Department
GZS-PTU Campus Bathinda
[1]mehakgarg2025@gmail.com
[2]csejyotigill@gmail.com
[3]lalitbti@gmail.com

Abstract— *In the present scenario image retrieval plays a vital role. The field of image retrieval has been an active research area for several decades and has been paid more and more attention in recent years as a result of the dramatic and fast increase in the volume of digital images. CBIR aims at finding image databases for specific images that are similar to a given query image based on its features. Users can query example images based on these features such as texture, color, region, shape and others. Target or close Images can be retrieved in a little fast if it is clustered in a right manner. For clustering, we use fuzzy- c mean clustering. In this way relevant images will be retrieved from database.*

Keywords— *content based image retrieval, Auto-correlation, RGB components, Query, Texture.*

I INTRODUCTION

Image retrieval is the process of searching and retrieving images from a large database. As the images grow complex, retrieve the right images become a difficult problem. Content Based Image Retrieval is a task of searching images from a database and retrieval of an image, which are looking to be visually similar to a given example or query image. Content-based image retrieval uses the visual contents of an image such as texture, color, and spatial layout to represent the image.

A. Image Retrieval

The advent of the World Wide Web (WWW) and the development of highly economical devices to store, capture and transmit images have led to the creation of huge image libraries. Thus, we are faced with the inevitable problem of having to retrieve useful information from these collections, both efficiently and effectively. This has led to a renewed interest in image retrieval and its practical applications.

1) Text based retrieval:

Traditional image retrieval employed text as the primary means by which to represent and retrieve images from databases. Images were stored along with string attributes – keywords prepared by an annotator that the content of the image. Although text-based image retrieval took advantage of already well-established information retrieval algorithms, its disadvantages as an effective tool to retrieve images became readily apparent.

2) Color based retrieval:

Since color is a low-level image feature that does not appear to classify images distinctly, few CBIR systems exist that utilize only color as the image retrieval feature. Yet color does have its advantages for image retrieval. It provides multiple measurements at a single pixel of the image, enabling categorization to be done without the need for complex spatial decision-making.

3) Content based retrieval:

Initial research in the retrieval of images based on their inherent features has been reported. Content-based image retrieval utilizes representations of features that are automatically extracted from the images themselves. Allmost all of the current CBIR systems allow for querying-by-example, a technique wherein an image (or part of an image) is selected by the user as the query. The system extracts the feature of the query image, searches the database for images with similar features, and exhibits relevant images to the user in order of similarity to the query.

B. Components of CBIR system

The CBIR system consists of the following components:

1) **Query image**:
It is the image to be found from the image database.

2) **Image database**:
It consists of n number of images depends on the user choice.

3) **Feature extraction**:
It separates visual information from the image and saves them as features vectors in a features database.

- Feature vectors: stores details about each image.
- Feature vectors are used to: compare the query image with the other images and retrieval.

4) **Image matching**:
Feature vectors of various images are compared with the feature vectors of query image which helps in measuring the similarity.

5) **Resultant retrieved images**:
Similar images having closest or same features as that of the query image are retrieved.

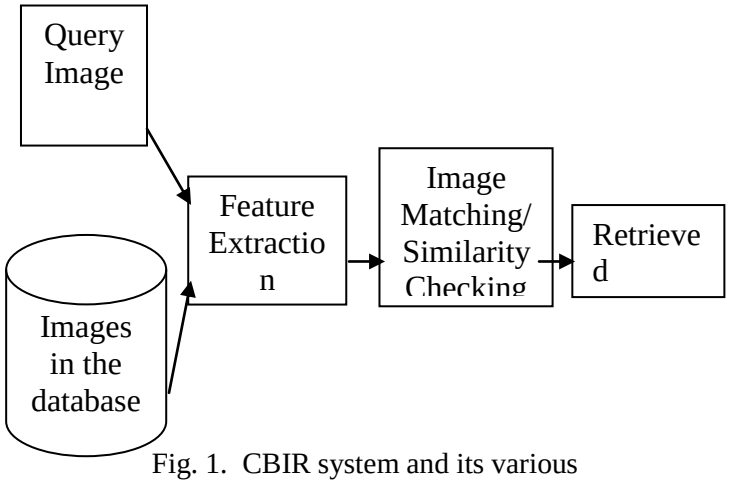

Fig. 1. CBIR system and its various components.

II APPLICATIONS OF CBIR SYSTEM

- Architectural and engineering design
- Art collections
- Crime prevention
- Medical diagnosis
- Military
- Photograph archives
- Retail catalogs
- Industrial area
- Fashion and graphic design

III PROBLEM DEFINITION

A. Image Database

First we have to create an image database. For this we need to search RGB components of images

B. RGB Components

In the color based image retrieval the RGB Color model is used. Color images normally are in three dimensional.RGB color components are taken from each and every image and their mean values are calculated and stored in the database and clustering is done based on these values. These three mean values for each image are deposited and considered as features.

C. Feature Extraction

The top ranked images are re-grouped according to their texture features. In which, the parameters are gathered on the basis of statistical approach. Statistical features of gray levels are one of the systematic methods to classify texture. E.g. entropy, contrast, dissimilarity, standard deviation, mean, Auto-correlation, and variance of both query image and target images are calculated. From the calculated values required image from the database is extracted.

The Gray Level Co-occurrence Matrix (GLCM) is used to extract second order statistics from an image. GLCMs have been used successfully for texture calculations.

D. Fuzzy-c Mean

In fuzzy-c means clustering method, every point has a degree of belonging to the clusters, as in fuzzy logic, rather than belonging completely to just one cluster. Thus, points on the edge of a cluster may be in the cluster to a smaller degree than points in the center of the cluster.

E. Similarity Comparison

Similarity comparison is done by selecting query image and then comparing this image with all the images in the database. Most relevant images are indexed at top. Then the top ranked images are retrieved from the database.

IV IMAGE RETRIEVAL SYSTEM

Image Retrieval from the image collections involves the following steps-

- Pre-processing
- Image Classification based on a true factor
- RGB Components processing
- Preclustering
- Texture feature extraction
- Similarity Comparison
- Target image selection

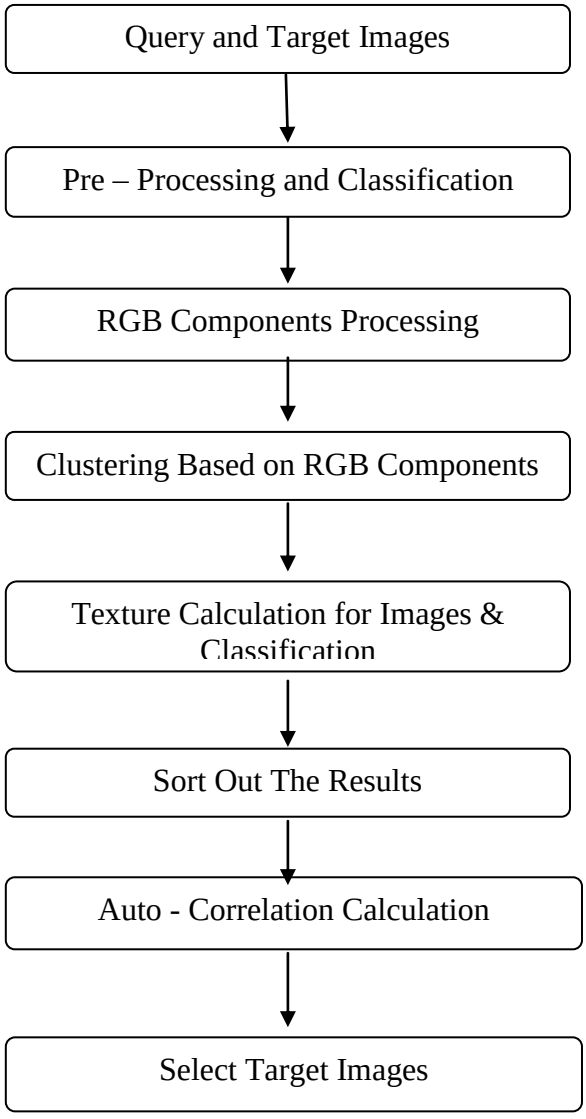

Fig. 2. Image retrieval system

A. Pre-Processing

Pre-processing is the name used for actions on images at the lowest level of abstraction. The main goal of the pre-processing is an improvement of the image that suppresses unwilling distortions or enhances some image features. This step focus on image feature processing.

Filtering is a technique used in pre-processing for modifying or enhancing an image. The noise in the image is filtered using linear and non-linear filtering techniques. Lucy - Richardson filtering is used here to reduce the noise.

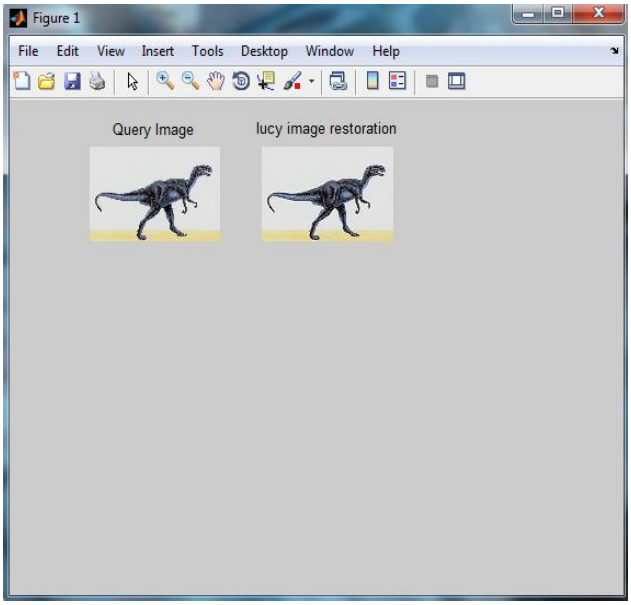

Fig.3. Pre-processing.

B. RGB Components Processing

An RGB colour image is an M*N*3 array of colored pixels, where each color pixel is a triplet equivalent to the red, green, and blue components of an image. An RGB image can be seen as the stack of three grayscale images that, when inserted into the red, green, blue inputs of a color monitor, generate the color image on the screen. By custom the three images form an RGB images are known as red, green and blue components.

The mean values for the RGB components are calculated for all images-

Red Mean (rm) = sum of all the red pixels in the image R(y) / No. of pixels in the image P(y)

Green Mean(gm) = sum of all green pixels in the image G(y) / No. of pixels in the image P(y)

Blue Mean(bm) = sum of all the blue pixels in the image B(y)

No. of pixels in the image P(y)

Where R (y) = RED component pixels,
G (y) = GREEN component pixels,
B (y) = BLUE component pixels,
P =No. of pixels in the image.

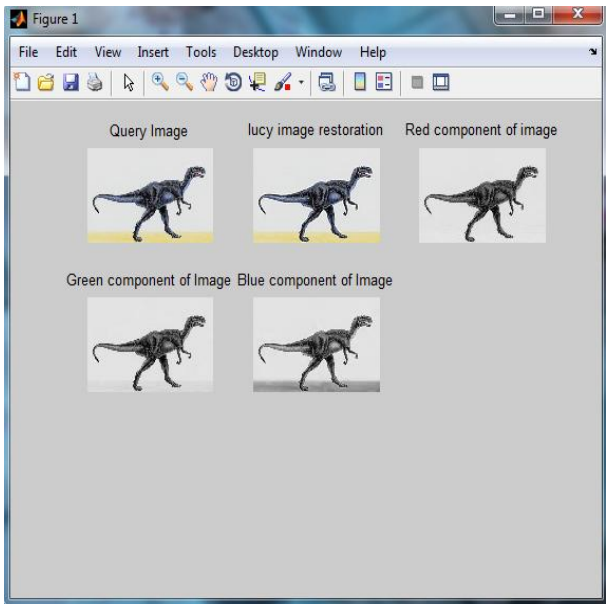

Fig. 4. RGB components processing.

After calculating the mean value of Red, Blue and Green components, the values are to be compared with each other to find the apogee value of the components. For example, if the value of Red component is Higher than the two, then we can conclude that the image is Red Intensity oriented image and which can be clustered into Red Group of Images.

Whenever the query image is already given, calculate the Red, Green, Blue components average values. Then compare this with the already stored values.

C. Image Clustering

Image Clustering will be a widely advantage for reducing the searching time of images in the database. Fuzzy c-means (FCM) is a technique of clustering which allows one piece of data to belong to two or more clusters. In fuzzy clustering data elements can belong to more than one cluster, and with each element a set of membership levels is associated. These indicate the strength of the corporation between that data element and a particular cluster. Fuzzy clustering is a technique of assigning these membership levels, and then using levels to assign data elements to one or more clusters. FCM sort data in specific number of clusters.

D. Auto-Correlation

The texture represents the energy content of the image. If an image contains high textures, then the energy will be high compared to the average and low texture images. There are various texture parameters to be considered. However, here, the texture parameter Auto-Correlation is highly focused and which is to be calculated for the query and target images.

Auto-correlation introduces to the correlation of a time series with its own past and future values.

E. Image Retrieval by Similarity Comparison

After clustering, similarity comparison is performed. To compare images with query image we find difference between the values of their red, green and blue components as well as their auto-correlation. Then we perform indexing to sort the retrieved images. While performing indexing the image from database with the smallest difference is ranked on the top and so on. Then the top ranked images from the database are retrieved.

F. Performance Measurement

Evaluation of retrieval performance is a crucial problem in Content-Based Image Retrieval (CBIR). Many different methods for measuring the performance of a system have been created and used by researchers. We will use the most common evaluation methods namely, Recall and Precision usually presented as a Precision vs. Recall graph. Recall and Precision alone contain insufficient information. With this, the following formulae are used for finding Precision and Recall values.

$$\text{Precision} = \frac{\text{No. of relevant images retrieved}}{\text{Total number of images retrieved}}$$

$$\text{Recall} = \frac{\text{Total no of relevant images in the database}}{\text{No of relevant images retrieved}}$$

V RELATED WORK

In this section we will look into the review of Content based image retrieval for image databases. It describes the previous work which had been done on a CBIR system using texture feature extraction and other techniques.

By Kannan in 2010[1] In this paper content based image retrieval method was proposed. It uses the feature of the image for its retrieval. The entropy texture feature is used here.

Kun-Che in 2009[2] In this paper Pixel-wised image characteristics were extracted and changed into a database like table which permits a variety of data mining algorithms to make explorations on it.

Silakari in 2009[3] In this paper a framework of unsupervised clustering of images based on the color feature of image. Clustering of images based on color moment and Block Truncation Coding to extract features from an image database is proposed. K-means clustering algorithm is conducted to group the dataset in various clusters.

Amanbir Sandhu, Aarti Kochhar in 2012[4] Presents a technique for content based image retrieval using texture, color and shape for image analysis. In this paper they worked with the three features i.e. texture, color and shape and its different combinations. The GLCM is used for texture feature extraction, histogram for Color feature extraction and for shape different factors are found like area, Euler No., eccentricity and Filled Area.

Saroj Shambharkar and Shubhangi Tirpude in 2011[5] Proposed a technique for image retrieval using fuzzy-c mean clustering. In this they said an optimization model or objective function must be devised to search for the optimal partition according to the chosen objective function. The way that most researchers have solved the optimization problem has been through an iterative locally optimal technique, called the FCM algorithm and hence they suggested a fuzzy-c mean algorithm.

Manimala Singha and K.Hemachandran in 2012 [6] Presents a technique for content based image retrieval using color and texture. In this they proposed two algorithms for image retrieval based on the color histogram and Wavelet-based Color Histogram. They presented a novel approach for Content Based Image Retrieval by combining the color and texture features called Wavelet-Based Color Histogram Image Retrieval (WBCHIR). Similarity between the images is ascertained by means of a distance function. The computational steps are effectively reduced with the use of Wavelet transformation.

Ray-I Chang, Shu-Yu Lin, Jan-Ming Ho, Chi-Wen Fann, and Yu-Chun Wang in 2012[7] Proposed a novel content based image retrieval system using K-means/KNN with feature extraction. This paper first combines segmentation and feature extraction module, grid module, K-means clustering and neighborhood module to build the CBIR system. The problem with this technique is that the system architecture and modules proposed in this paper are not optimized properly.

Peter Stanchev in 2003[8] Used image mining for image retrieval. In this paper after low level image properties extraction image mining was made for obtaining associate rules, describing the high level image semantic features. It deals with high level image semantic features which combine color, shape and texture features of an image.

Dr. Fuhui Long, Dr. Hongjiang Zhang and Prof. David Dagan Feng in 2003[9] Represents fundamentals of content-based image retrieval. They introduced some fundamental techniques for content based image retrieval, including visual content description, indexing scheme, similarity/distance measures, user interaction and system performance evaluation. Their emphasis is on visual feature description techniques.

VI CONCLUSION

The main objective of the image retrieval is to retrieve the images from database very fast and in an efficient manner. The images are pre-processed with various techniques and the texture calculation is highly focused. Here, the images are clustered based on RGB

Components, Texture values and Fuzzy C means Clustering algorithm. Clustering is very efficient and powerful technology to handle large data sets. It assists faster image retrieval and also allows the search for more relevant images in large image databases. Auto-correlation is used to compare the images and to improve the system performance.

REFERENCES

[1] A.Kannan, Dr.V.Mohan, Dr.N.Anbazhagan, "Image Clustering and Retrieval using Image Mining Techniques", *IEEE International Conference on Computational Intelligence and Computing Research, 2010*.

[2] Kun-Che Lu, Don-Lin Yang, "Image Processing and Image Mining using Decision Trees", *Journal of Information Science and Engineering 25, 989-1003, 2009.*

[3] Dr. Sanjay Silakari, Dr. Mahesh Motwani and Manish Maheshwari, "Color Image Clustering using Block Truncation Algorithm", *IJCSI International Journal of Computer Science Issues, Vol. 4, No. 2, 2009.*

[4] Amanbir Sandhu, Aarti Kochhar, "Content Based Image Retrieval using Texture, Color and Shape for Image Analysis", *International Journal of Computers & Technology ,Volume 3, No. 1, AUG, 2012* .

[5] Saroj Shambharkar, Shubhangi Tirpude, "Fuzzy C-Means Clustering For Content Based Image Retrieval System", *International Conference on Advancements in Information Technology With workshop of ICBMG 2011, IPCSIT vol.20 (2011) © (2011) IACSIT Press, Singapore, 2011.*

[6] Manimala Singha and K.Hemachandran, "Content Based Image Retrieval using Color and Texture", *Signal & Image Processing: An International Journal (SIPIJ) Vol.3, No.1, February 2012.*

[7] Ray-I Chang, Shu-Yu Lin, Jan-Ming Ho, Chi-Wen Fann, and Yu-Chun Wang, "A Novel Content Based Image Retrieval System using K-means/KNN with Feature Extraction", *ComSIS Vol. 9, No. 4, Special Issue, December 2012.*

[8] Peter Stanchev, "Using image mining for image retrieval", IASTED conf. "Computer science and Technology", *Cancum, Mexico, 214-218, May 19-21, 2003.*

[9] F. Long, H. Zhang, H. Dagan, and D. Feng, "Fundamentals of content based image retrieval," in D. Feng, W. Siu, H. Zhang (Eds.): "Multimedia Information Retrieval and Management. Technological Fundamentals and Applications," *Multimedia Signal Processing Book, Chapter 1, Springer-Verlag, Berlin Heidelberg New York, pp.1-26, 2003.*

[10] Nishchol Mishra, Dr. Sanjay Silakari, "Image Mining in the Context of Content Based Image Retrieval: A Perspective", *IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012.*

[11] 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*.

[12] Yong Rui and Thomas S. Huang, Shih-Fu Chang, "Image retrieval: current techniques, promising directions, and open issues", *Journal of Visual Communication and Image Representation* ***10****, 39–62, 1999.*

[13] Gurpreet Kaur, "A Review: Content Base Image Mining Technique for Image Retrieval Using Hybrid Clustering", *International Journal of Advanced Research in Computer Engineering & Technology (IJARCET), Volume No. 2, Issue No. 6, June 2013.*