**Research Paper** E-ISSN NO : 2455-295X | VOLUME : 3 | ISSUE : 4 | APRIL 2017

# AN IMPROVED CONTENT BASED IMAGE RETRIEVAL APPROACH USING LOCAL BINARY PATTERN (LBP)

Manju Rani [1] | Pawan Kumar Mishra [2]

[1] M.Tech, Scholar, UTU, Dehradun.

[2] Assistant Professor, UTU, Dehradun.

**ABSTRACT**

*The increased necessity of content based image retrieval (CBIR) technique can be found in a variety of different domains such as Data Mining (DM), Education, Medical Imaging (MI), Weather forecasting Crime Prevention, , Remote Sensing (RS) etc. An image retrieval system permits us to browse, search & retrieve the images. In past because of very huge image collections the manual annotation approach was more tedious. In order to conquer these difficulties content based image retrieval (CBIR) was introduced. This paper presents the content based image retrieval (CBIR), using local binary pattern (LBP). The local binary pattern encodes the relationship between the referenced pixel & its surrounding neighbors by computing the GLD (gray-level difference). The objective of the proposed methodology is to retrieve the best images from the stored database that resemble the query image with an optimized way.*

**Keywords:** *Content Based Image retrieval, CBIR, Image Processing, Color Histogram Techniques, Image feature Extraction.*

## *INTRODUCTION*

Digital images are currently widely used in medicine, fashion, architecture, face recognition, finger print recognition and bio-metrics etc. Hence, efficient image searching and retrieval are important. The earlier image retrieval systems were text based. Images were represented by using keywords. The keyword for the image was created by human operators. Manually entering keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Therefore, Content Based Image Retrieval (CBIR), based on the image content came into existence. Ying Liu et al, 2007 surveyed the CBIR system focusing on high level semantic concepts [1]. Discussed feature extraction, distance measures, classifier techniques such as neural network classifiers, K-nearest neighbor algorithm and the performance measures. The CBIR system relies on color, texture and shape which are low level image features. There are various possible low level features available in the literature. The low level features are extracted from the database images and stored in a feature database. Similarly, the low level features are extracted from the query image and the query image features are compared with the database image features using the distance measure. Images having the least distance with the query image are displayed as the result. The various popular distance measures reported in the literature survey. The main drawback of the CBIR system is that the images with similar low level features may vary from the query [2].

Digital images are currently widely used in medicine, fashion, architecture, face recognition, finger print recognition and bio-metrics etc. Hence, efficient image searching and retrieval are important. The earlier image retrieval systems were text based. Images were represented by using keywords. The keyword for the image was created by human operators. Manually entering keywords for images in a large database can be inefficient, expensive and may not capture every keyword that describes the image. Therefore, Content Based Image Retrieval (CBIR), based on the image content came into existence. Ying Liu et al, 2007 surveyed the CBIR system focusing on high level semantic concepts. Discussed feature extraction, distance measures, classifier techniques such as neural network classifiers, K-nearest neighbor algorithm and the performance measures. The CBIR system relies on color, texture and shape which are low level image features. Section 2 discusses the various possible low level features available in the literature. The low level features are extracted from the database images and stored in a feature database. Similarly, the low level features are extracted from the query image and the query image features are compared with the database image features using the distance measure. Images having the least distance with the Query image are displayed as the result. The main drawback of the CBIR system is that the images with similar low level features may vary from the query Content-Based Image Retrieval (CBIR) aims at Developing techniques that support effective searching And browsing of large image digital libraries on the Basis of automatically derived image features. The similarities (distances) between the feature vectors of the query images and the images stored in the Database are calculated and retrieval is performed by the indexing scheme. Indexing normally improves the Retrieval performance and hence it can be used to Effective retrieval process. Indexing relevant images as per the query image should retrieve images in the order Of similarity of the query image. Unsupervised learning can be applied to those types of problems where in the Main issue is to organize the data. So far, a major Breakthrough in image retrieval has not been achieved. Furthermore, there is no single universal approach to The content-based retrieval problems, even in a Restricted domain application [3] There are systems That discuss about the objects based on similarity Measures. Based on the similarity of each object set in each group (also called a cluster) or they are not Similar are put into different clusters. CLUE, cluster based Retrieval of images by unsupervised learning is an example of CBIR technique based on unsupervised Learning. However, content based image retrieval confront Many image types, some of them even have not a clear Object, so some strategies have to be dealt to reduce Such a problem. The main objective is to reduce the need for efficient content-based image retrieval has increased tremendously in many application areas such as biomedicine, military, commerce, education, and web image classification and searching. Currently, rapid and effective searching for desired images from large-scale image database becomes an important and challenging research topic [1] [2]. Content-based Image Retrieval (CBIR) technology overcomes the defects of traditional text-based image retrieval technology, such as heavy workload and strong subjectivity. It makes full use of image content features (color, texture, shape, etc.), which are analyzed and extracted automatically by computer to achieve the effective retrieval [4]. The stages of CBIR are shown in figure 1.

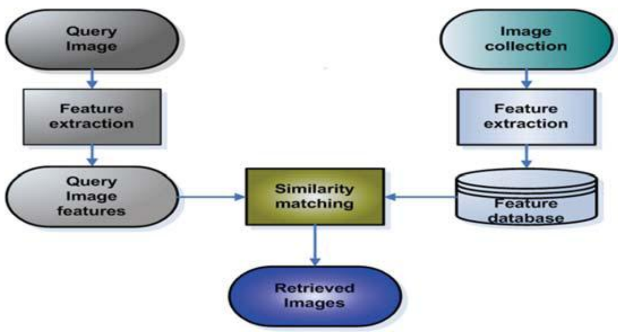


***Fig 1. Various Stages of CBIR***

## *LITERATURE REVIEW*

Content-Based Image Retrieval Based on Integrating Region Segmentation and Color “Histogram Duraisamy Yuvaraj” and “Shanmugasundaram Hariharan” proposed by Developments in multimedia technology, increasing number of image retrieval functions and capabilities has led to the rapid growth of CBIR techniques. Color histogram could be compared in terms of speed and efficiency.

An efficient CBIR system with better performance is presented by using the wavelets decomposition of image; they have generated the composite sub-band gradient and the energy distribution pattern string from the sub images of are generated by means of wavelet decomposition to the input image Kim D. and Chung C., Cluster Relevance Feedback For filtering out the undesired images a technique based on energy distribution pattern strings fuzzy matching is used. The resultant images are compared with query image after filtering. The system is tested on the database of 2400 images. Texture has no formal definition but intuitively it provides measure of properties such as coarseness, regularity and smoothness. It plays a role in human visual perception and interpretation. Texture describes in three approaches, namely, structural, statistical and spectral. In structural approach the texture is formed of small texture elements called ’Texel’s’ by following placement rules. The statistical method assumes texture by means of statistical grey level features of image pixels. The spectral approach is based on filtering theory in frequency domain and power density function. The structural approach is not so prevalent because majority of the natural structures contain asymmetrical shapes. [8].

It computes the probability of the incidence of identical pixel colors for each pixel and its neighbor pixels in the image. The difference among pixels is calculated and applied to the entire image. Li J., Wang J., and Wiederhold G., “IRM all pixels in the image are grouped by using K-means Clustering algorithm. The technique is applied to 3 databases and better results are reported. The technique is not suffering from image displacement and rotation. Analyzing images in real-time dynamic environment is not feasible. A system is proposed in that analyze images in real-time environment. It extracts shape and color to represents image contents. The system uses C-Means clustering algorithm to segment the images and extract its boundaries. Fast Fourier Transform is applied to provide an array of vector corresponding to a region. For color features HSI color model is used. Similarity Matching Algorithm using distance measures among the feature vector is applied to match the queried images. The capability of a CBIR approaches is fully dependent on the features retrieved from the image. Frequently it is observed that there is a semantic gap between the visual features and semantic content of an image. The semantic gap could be decreased by extraction of more effective features. This is a challenging area in CBIR research. [9]

To overcome the semantic gap different machine learning techniques are used. In SVM is used to extract the image features accurately and retrieve the desired image efficiently. In hierarchical methodology is used to retrieve an iris image is presented. This technique is based on an innovative indexing method for an iris database. Two distinct features are exploited to extract the contents of iris image i.e. iris color is used to form the basis for the indexing of image while texture of image is used to retrieve image from the indexed iris database. The undesired images are filtered out by using the color feature; the images having no similarity with the query image color are filtered out. The proposed technique is evaluated over noisy iris images. It produced better results. A novel image retrieval scheme (ICTEDCT-CBIR) based on curve let transform is presented in Lin Y ., Liu T., and Chen H.[10], this model integrates Curve let multi-scale ringlets with region-based vector codebook sub band clustering to extract dominant color feature and texture analysis.

In a new approach based on three popular algorithms that are: color histogram, texture and moments invariants is proposed. The three algorithms are used to ensure capturing of the regularity of image edges Mathew S., Balsa V., Zachariah P., and Samue P., [11].

A new approach to extract color and texture features of image for CBIR is proposed. They identified the low level features of color and texture for CBIR by using two color histogram function and their comparison. it combined texture, color and shape descriptors for image retrieval. The paper focuses on feature extraction and representation. In an image retrieval approach based on fuzzy KNN classifier is proposed. It allocates an initial semantic label to database images. In this method the assigned labels to the images are modified steadily by relevance feedback. Several measures based on similarity were observed for three kinds of visual features. Genetic algorithm was used to assign optimum weights for each type of feature and to find out their components. For the residual 800 images of the image database, the proposed approached assigned outputs classifier as initial weights to the associated links of the network. During the image contents retrieval session, the network weights are turned in to the relevance feedback from the users.

In a new descriptor for the retrieval of image is called micro-structure descriptor (MSD) is proposed. The MSD is defined on the basis of edge orientation similarity. This technique based on color, it extracts features and effectively combines them with shape, color, texture and color layout features entirely for image retrieval. The proposed method is verified on the database of 15000 images. The results are compared with Gabor features and multi-texting histograms. It produced better results. Feature selection is a common problem in CBIR.

## *EXISTING METHODOLOGY*

Techniques of color histogram CBIR could be compared in terms of speed and efficiency, and a modified approach based on a composite color image histogram processing is introduced. The proposed approach is fast and provides results comparable to those of much slower algorithms [14]. In the above paper a new histogram based image retrieval technique have been proposed which outperforms most previous methods in terms of speed and accuracy. Research of CBIR, visual signature based on region was attracted more attention. In order to get the signature based on region, the crucial step is image segmentation, and reliable image segmentation is also critical to get the image shape description. Unfortunately, it has been demonstrated that accurate image segmentation is still an open problem [1]. The authors have suggested some strategies for dealing with this problem to reduce dependence on accurate image segmentation for a practical image retrieval system. Due to the semantic gap, there are still many shortcomings for image retrieval system only with the low level visual features [9]. Based on the high dimension bio mimetic information geometry theory, we segment image into main region and margin region for cognition the whole image characteristic. A prototype image retrieval system was made using the color and texture features of regions. CBIR combining some percentage value of two features namely; color-texture features and color-shape features would be interesting and taking the union of these two features has gained sufficient interest [6]. The combination of features provides a robust feature set for image retrieval. Evaluation is then measured based on different precision value of the image retrieval on each category of image database. Earlier CBIR system based on unsupervised learning, combines all the features values namely shape, color and texture of an image for assigning a weight on different images (as a target images) in the image database with 60% features stores of each visual features [4]. It is investigated that CBIR based on unlabelled images and recommendations for improving the CBIR system using unlabelled images improves the retrieval results. More specific studies have been presented a survey in the field of CBIR, providing an overview of the most important aspects characterizing that kind of images. The image databases are classified into the labeled image as relevant, irrelevant and also unlabelled image [12].

## *EXISTING ALGORITHM*

The proposed CBIR framework is given as shown in the Figure 1. The sequence of steps is given below.

Step 1: Input A query image “I”.

Step 2: Compute image segmentation and then calculate features (colour, texture and edge) values of the image.

Step 3: Convert I to gray scale.

Step 4: Construct histogram of image.

Step 5: Similarity comparisons between input images.

Step 6: Finally, relevant images are retrieved with respect to corresponding query image I.

Step 7: Repeat step 1 to 6 for rest of the query images.

Step 8: End.

In the existing algorithm, the authors proposed a CBIR system based on unsupervised learning, where in, they combine all the features values namely shape, color texture, norms of residual values for an image for assigning a weight on different images (as a target images) in the image database with 60% features stores of each visual features. The authors experimented with a standard image database consisting of approximately 100 images to compare the performance of the existing systems by combining both shape-color features and color texture features. They have taken the union of these two approaches and experimentally, they found that the union of both gives the better performance at different residual values. In their experiments, they used histogram gray scale image as the similarity measure for computing the similarity of images in the database with a query image. Experimentally, they found that the CBIR systems after taking union outperforms at smaller residual at k (almost 100%) in almost all categories of an image database. Experimentally, they also found that the proposed CBIR system gives better results than the CLUE and UFM based CBIR systems.

## *PROPOSED METHODOLOGY*

An image histogram is a type of histogram that acts as a graphical representation of the tonal distribution in a digital image. It plots the number of pixels for each tonal value. 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's color space, the set of all possible colors. [14]Histograms plots how many times (frequency) each intensity value in image occurs z Example: z Image (left) has 256 distinct gray levels (8 bits) z Histogram (right) shows frequency (how many times) each gray level occurs Many cameras display real time histograms of scene z Helps avoid taking over-exposed pictures z Also easier to detect types of processing previously applied to image z E.g. K = 16, 10 pixels have intensity value = 2 z Histograms: only statistical information z No indication of location of pixels Different images can have same histogram z 3 images below have same histogram Histograms: only statistical information z No indication of location of pixels Intensity values Histograms z Different images can have same histogram z 3 images below have same histogram z Half of pixels are gray, half are white z Same histogram = same statistics z Distribution of intensities could be different z Can we reconstruct image from histogram? No! Histograms z So, a histogram for a grayscale image with intensity values in range would contain exactly K entries z E.g. 8-bit grayscale image, K = 28 = 256 z Each histogram entry is defined as: h(i) = number of pixels with intensity I for all 0 < i < K. z E.g. h(255) = number of pixels with intensity = 255 z Formal definition Histograms help detect image acquisition issues z Problems with image can be identified on histogram z Over and under exposure z Brightness z Contrast z Dynamic Range z Point operations can be used to alter histogram. E.g. z Addition z Multiplication z and Log z Intensity Windowing (Contrast Modification).

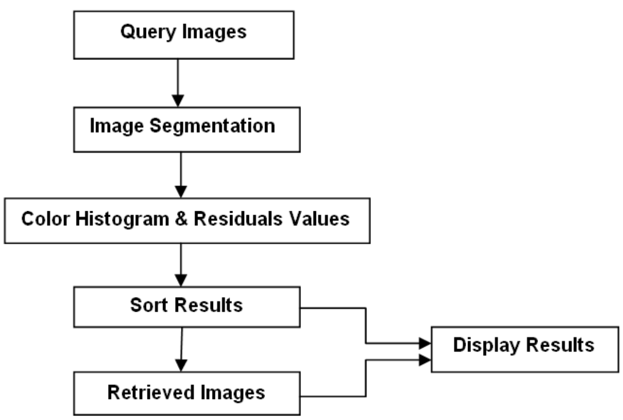


*Fig 2. Proposed Architecture for CBIR System*

Based on LBP operator, we propose a new descriptor denoted Upper-Lower of Local binary Pattern (UL-LBP). It reduces the dimensionality of the LBP operator and is robust against noise for describing texture analysis. The basic idea is to divide the eight neighboring pixels into two parts and then compute the code for each part separately. Part 1 consists the four pixels Lower Neighbors (LN) and part 2 contains the four pixels Upper Neighbors (UN) as illustrate in fig 1 above methodology has been implemented on the image database and one query image is chosen for getting images having almost same histogram. HSV histogram is used for comparison. Here steps of implementation have been shown [13].

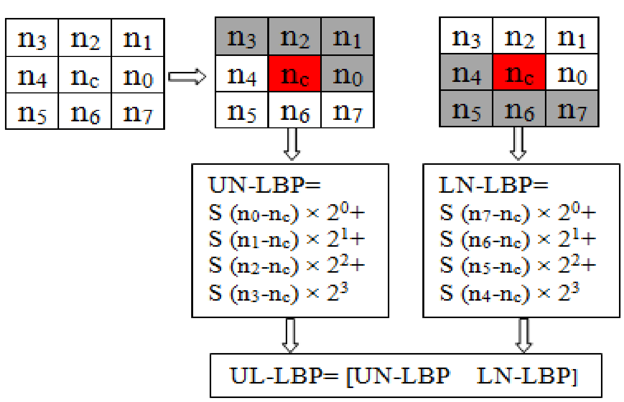


*Fig. 3 Computation of the UL-LBP Descriptor*

Instead of comparing each neighbor pixel with the central pixel as the original LBP operator, we propose to test each pixel of the UN (n0, n1, n2, n3) with the central pixel. The first descriptor proposed is called UN-LBP whereas the second is called LN-LBP which obtained by a comparison of LN pixels (n7, n6, n5, n4) with the central one. UN-LBP and LN-LBP descriptors are defined. By combining histograms of UN-LBP and LN-LBP, we obtain new descriptor named Upper-Lower of Local Binary Pattern (UL-LBP) with a size of 32 bins.

## *COMBINATION OF LOCAL AND GLOBAL FEATURES*

We present a method that combines local and global features in order to enhance results of image retrieval. Global feature represents texture-color feature for the whole image using UL-LBP descriptor from individual color channels (R, G and B). Local feature describes texture in an image region based on SIFT algorithm and extract the IPs features from each color image independently. SIFT algorithm provides a large set of feature per images. To solve this problem, we used a BOW model which is successfully used in CBIR. BOW quantizes the features descriptor into "visual words" and builds a compact histogram called BOW-SIFT (B-SIFT). Block diagram of our method is illustrated. The training image is done with local and global features.

## *RESULT DISCUSSION*

We, design a graphical user interface to show the result & analysis of existing Vs proposed algorithm. We show the result with the help of following steps- The Graphical User Interface for content based image retrieval is shown in figure 4. Image selection from database is shown in figure 5. Figure 6, 7, 8 and 9 show color histogram feature, edge histogram feature, edge histogram features from query image, edge direction respectively. The texture of query image is shown by figure 10. Figure 11 shows matching of input image with database, the result of number of images is shown in figure 12. The residual of norms of existing Vs proposed algorithm is shown in figure 13. The resulting values of existing Vs proposed algorithm is shown in figure 14. Finally, the values for 10 images using existing Vs proposed algorithm is shown in table 1.

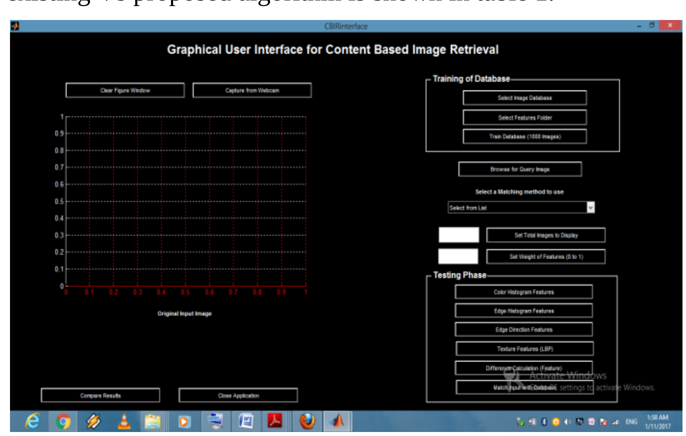


*Fig 4. Graphical User Interface for Content Based Image Retrieval*

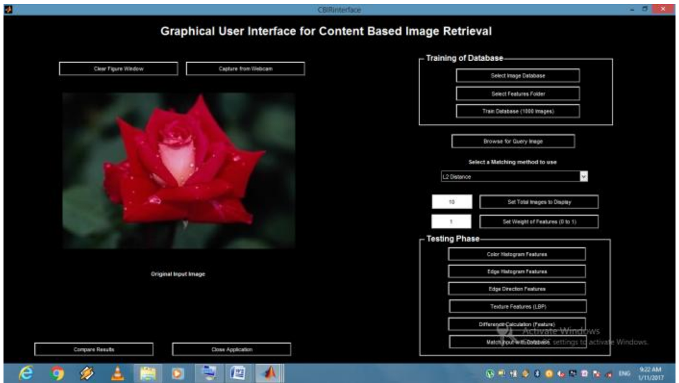


*Fig 5. Select Image from Database*

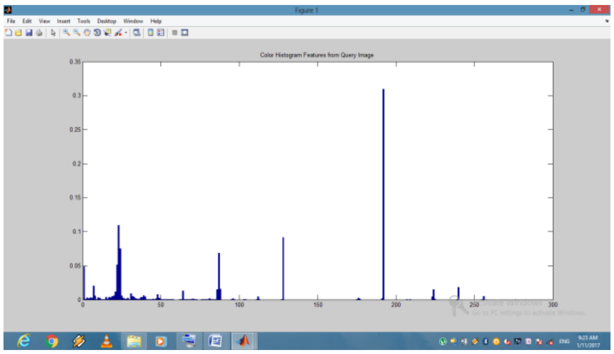

*Fig 6. Color Histogram Feature*

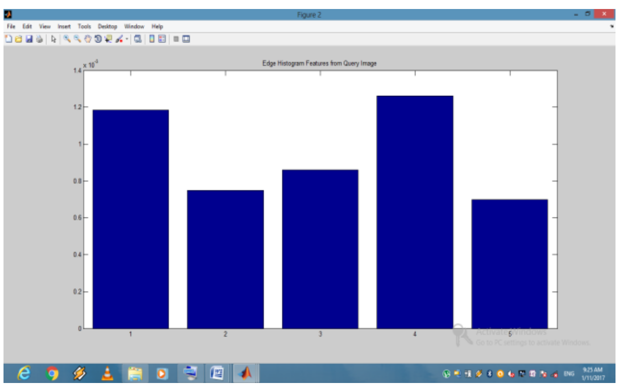

*Fig 7. Edge Histogram Feature*

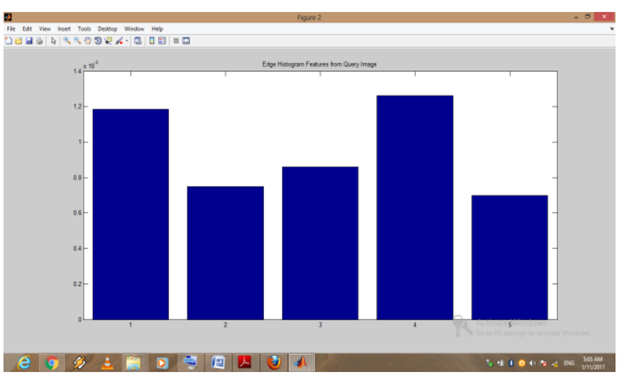

*Fig 8. Edge Histogram Features of Query Image*

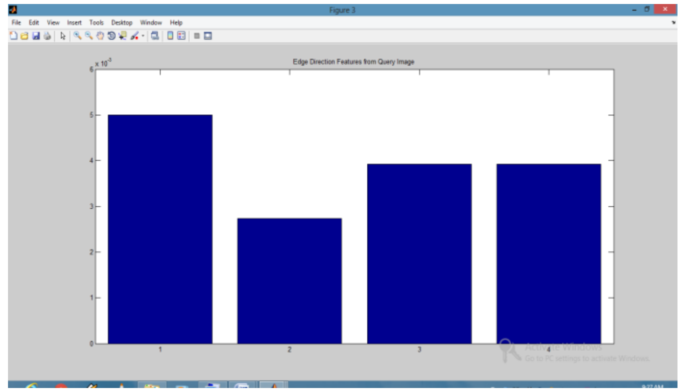

*Fig 9. Edge Direction*

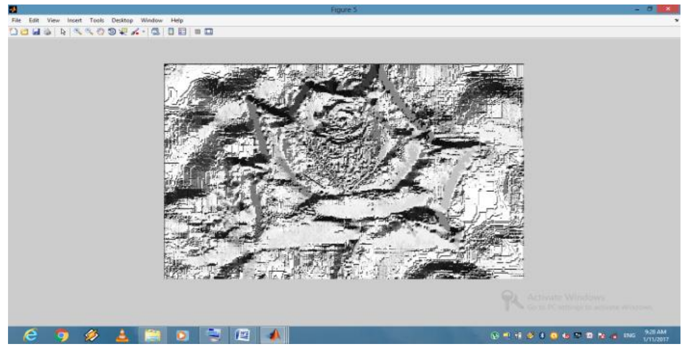

*Fig 10. Texture of Query Image*

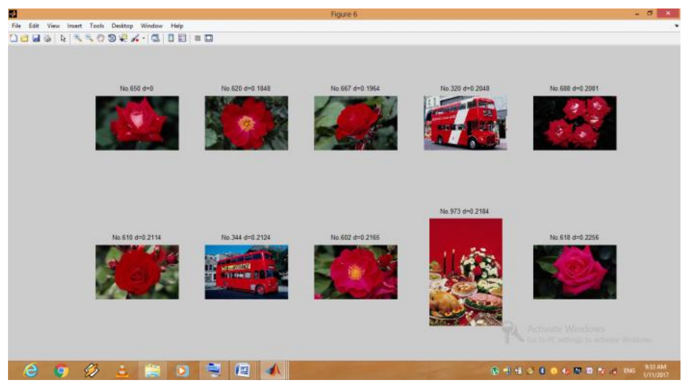

*Fig 11. Match Input Image with Database*

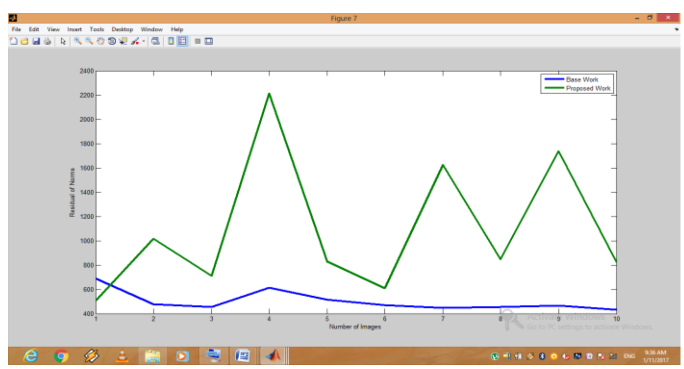

*Fig 12. Result of Number of Images*

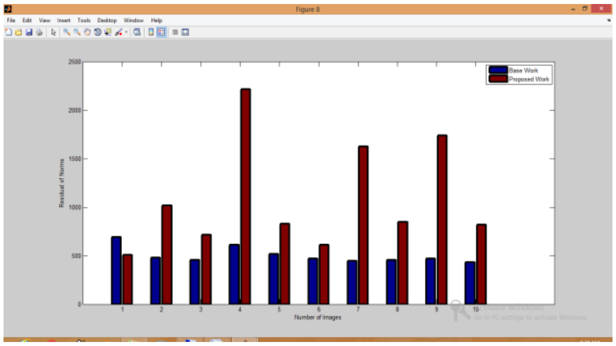


*Fig 13. Residual of Norms Comparison of Existing Vs Proposed Algorithm*

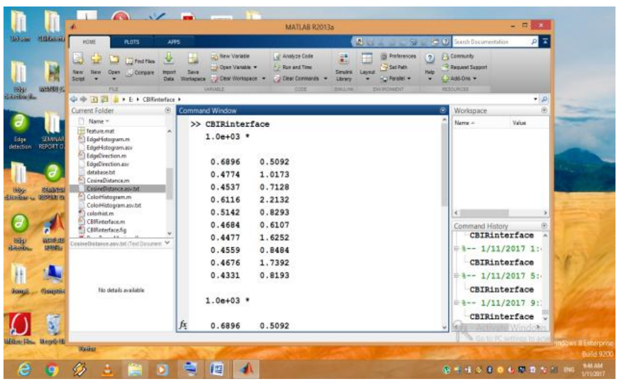

*Fig 14. Compare Retrieved Image Value of Existing Vs Proposed Algorithm*

***Table 1***

| *Image* | *Result Using Existing Methodology* | *Result Using Proposed Methodology* |
|---|---|---|
| I-1 | 0.6896 | 0.5092 |
| I-2 | 0.4774 | 1.0173 |
| I-3 | 0.4537 | 0.7128 |
| I-4 | 0.6116 | 2.2132 |
| I-5 | 0.5142 | 0.8293 |
| I-6 | 0.4684 | 0.6107 |
| I-7 | 0.4477 | 1.6252 |
| I-8 | 0.4559 | 0.8484 |
| I-9 | 0.4676 | 1.7392 |
| I-10 | 0.4332 | 0.8193 |
| ***AVERAGE*** | ***0.6896*** | ***0.5092*** |

### *CONCLUSIONS*

In this paper, we proposed CBIR system based on unsupervised learning, where in, we combine all the features values namely shape, color texture, norms of residual values for an image for assigning a weight on different images (as a target images) in the image database with 60% features stores of each visual features. We experimented with a standard image database consisting of approximately 100 images to compare the performance of the proposed systems by combining both shape-color features and color texture features. We have taken the union of these two approaches and experimentally, we found that the union of both gives the better performance at different residual values. In our experiments, we used histogram gray scale image as the similarity measure for computing the similarity of images in the database with a query image. Experimentally, we found that the CBIR systems after taking union outperforms at smaller residual at k (almost 100%) in almost all categories of an image database. Experimentally, we also found that the proposed CBIR system gives better results than the CLUE and UFM based CBIR systems [20].

### *FUTURE SCOPE*

This thesis concludes by just extending the texture feature extraction techniques into CBIR system. This area can be further explored and the techniques can be finely tuned with or without involving some pre or post processing works for increasing the retrieval efficiency. The fine-tuning may be done adding some shape and color information in well-determined form with the already existing texture information to suit the application this work can be further extended to some domain-based applications such as finger print recognition, retina identification, and object detection etc for large image database. Since texture analysis consumes a considerable amount of time for feature extraction, there is a scope for optimization also.

### ***REFERENCES***

*1. Alnihoud J., "Content-Based Image Retrieval System Based on Self Organizing Map, Fuzzy Color Histogram and Subtractive Fuzzy Clustering," the International Arab Journal of Information Technology, vol. 9, no. 5, pp. 452- 458, 1990.*

*2. Chen Y., Wang J., and Krovetz R., "Content Based Image Retrieval by Clustering," in Proceedings of 5th International ACM Workshop on Multimedia Information Retrieval, pp. 193-200, 1991.*

*3. Chen Y., Wang J., and Krovetz R., "An Unsupervised Learning Approach to Content-Based Image Retrieval," in proceeding of 7th International Symposium on Signal Processing and its Applications, pp. 197-200, 1995.*

*4. He X., Ma W., and Zhang H., "Learning an Image Manifold for Retrieval," in Proceedings of ACM International Conference on Multimedia, New York, USA, pp. 17-23, 1998.*

*5. Jain A. and Farrokhnia F., "Unsupervised Texture Segmentation using Gabor Filters," In Proceedings of IEEE International Conference on Systems Man and Cybernetic, Los Angeles, USA, pp. 14-19, 2000.*

*6. Jobay R. and Sleit A., "Quantum Inspired Shape*

*Representation for Content Based Image Retrieval," Journal of Signal and Information Processing, vol. 5, no. 2, pp. 54-62, 2002.*

*7. Duraisamy Yuvaraj and Shanmugasundaram Hariharan "Content-Based Image Retrieval Based on Integrating Region Segmentation and Color Histogram" The International Arab Journal of Information Technology, Vol. 13, No. 1A, 2003.*

*8. Kim D. and Chung C., "Cluster: Relevance Feedback using Adaptive Clustering for Content Based Image Retrieval," in Proceedings of the Content-Based Image Retrieval Based on Integrating Region Segmentation and Color Histogram 207 ACM SIGMOD'03, San Diego, USA, pp. 599- 610, 2005.*

*9. Li J., Wang J., and Wiederhold G., "IRM: Integrated Region Matching for Image Retrieval," in Proceedings of the 8th ACM International Conference on Multimedia, New York, USA, pp. 147-156, 2006.*

*10. Lin Y., Liu T., and Chen H., "Semantic Manifold Learning for Image Retrieval," in Proceedings of the 13th ACM International Conference on Multimedia, New York, USA, pp. 249-258, 2007.*

*11. Monika J. and Singh S., "An Experimental Study on Content Based Image Retrieval Based on Number of Clusters using Hierarchical Clustering Algorithm," International Journal of Signal Processing, Image Processing and Pattern Recognition, vol. 7, no. 4, pp. 105-114, 2009.*

*12. Pavani P. and Prabha S., "Content Based Image Retrieval using Machine Learning Approach," in Proceedings of the In ternational Conference on Frontiers of Intelligent Computing: Theory and Applications, vol. 247, pp. 173-179, 2011.*

*13. Rue Y., Huang T., and Mehrotra S., "Content- Based Image Retrieval with Relevance Feedback MARS," in Proceedings of the IEEE International Conference on Image Processing, Santa Barbara, USA, pp. 815-818, 2009.*

*14. Yang C., Dong M., and Fotouhi F., "Semantic Feedback for Interactive Image Retrieval," in Proceedings of the 13th annual ACM International Conference on Multimedia, New York, USA, pp. 415-418, 2012.*

*15. Zhou X., and Huang T., "Relevance Feedback in Image Retrieval: A Comprehensive Review," Multimedia Systems, vol. 8, no. 6, pp.536–544, 2008.*

*16. A. Csillaghy, H. Hinterberger, and A.O. Benz, "Content Based Image Retrieval in Astronomy," Information Retrieval, 3(3):229–241, 2013.*

*17. T. H. Painter, J. Dozier, D. A. Roberts, R. E. Davis, and R. O. Green, "Retrieval of sub-pixel snow-covered area and grain size from imaging spectrometer data," Remote Sensing of Environment, 85(1):64–77, 2014.*

*18. M. Schroder, H. Rehrauer, K. Seidel, M. Dutch, "Interactive learning and probabilistic Retrieval in remote sensing image archives," IEEE Trans. Geosciences and Remote Sensing, 38(5):2288–2298, 2015.*

*19. J.Z. Wang, J. Li, and G. Wiederhold, "Simplicity: Semantics-Sensitive Integrated Matching for Picture Libraries," IEEE Trans. Pattern Analysis and Machine Intelligence, 23(9), 947– 963, 2014.*

*20. H. Muller, W. Muller, D. M. Squire, S. Marchand-Maillet, and T. Pun, "Performance evaluation in content-based image retrieval: Overview and proposals," Pattern Recognition Letters, 22(5):593-601, 20016.*