

An Efficient Model for Content based Image Retrieval

M. Karthikeyan
Programmer (Senior Scale)
Department of Computer
Science and Engineering
Annamalai University
Chidambaram
India

P. Aruna
Associate Professor
Department of Computer
Science and Engineering
Annamalai University
Chidambaram
India

ABSTRACT

In the recent world with the advances in multimedia technologies such as compression, display, and visualization technologies and the increased emphasis on multimedia application, the production of image information has resulted in large volume of images that need to be properly indexed for retrieval in future. Hence, there is a need for Content Based Image retrieval application which makes the retrieval process very efficient. Current systems generally make use of low level features like colour, texture, and shape. In this paper, a novel approach for generalized image retrieval based on semantic contents is presented. A combination of two feature extraction methods namely colour and edge histogram descriptor is proposed. The retrieval efficiency is computed and compared by using four methods such as k-means, colour histogram, edge histogram and sobel method. For colour, the histogram of images is computed and for edge, edge histogram descriptors (EHD) are found. For retrieval of images, a novel idea is developed based on greedy strategy to reduce computational complexity. The proposed system stores the content of database images automatically and query image's content is extracted during runtime and it is used to match against those in database. The result of the query is a set of images that are similar to the query image.

Keywords

Content-based Image Retrieval (CBIR), Hue Saturation Value (HSV), Local Colour Histogram (LCH), Global Colour Histogram (GCH), Edge Histogram Descriptor (EHD).

1. INTRODUCTION

A fairly new branch in information retrieval is Content Based Image Retrieval (CBIR), the science of how images are indexed and retrieved based on its contents. An image consists of different materials than textual documents. A textual document consists of terms, phrases and with some extension hyperlinks. There are mainly four different ways of searching images. The first one is searching images based on attributes such as file name, date of creation and manually added categories such as subject and creator. The problem with this method is, auto-generated information about the images may not be enough to describe the content and adding categories manually is a very time consuming task. The second approach is to use automated object recognition, which is often used in specific domains such as medical images. Automated object recognition is time-consuming and very difficult to achieve in a general system. The third approach is to use free text to annotate the images, and use textual information retrieval techniques to retrieve images. This approach can be divided

into manual annotation and automatic annotation. With manual annotation, specialists annotate appropriate text to each image. The problem here is, in a massive image collection, this will be an enormous task. Automatic image annotation consists of two methods, partitioned and hierarchical. The fourth approach is perhaps what most people think of when they talk about CBIR. This approach uses the most fundamental elements of an image. Images can typically be divided in three different low-level features by using image content descriptors such as colour, shape and texture. The proposed system concentrates on Content Based Image retrieval method using image content descriptors colour, texture and shape for efficient retrieval of images.

The structure of this paper is as follows: Section 2 discusses some related research work regarding content based image retrieval. Section 3 describes system overview of content based image retrieval. Section 4 gives the implementation procedure for the proposed method. The experimental results are given in section 5. Finally some conclusion and discussion is given in section 6.

2. RELATED WORKS

In the early 1990s, as a result of advances in the Internet and new digital image sensor technologies, the volume of digital images produced by scientific, educational, medical, industrial and other applications available to users, increased dramatically. The difficulties faced by text-based retrieval system become more and more severe. The efficient management of the rapidly expanding visual information become an urgent problem. Tristan Glatard and John Montagnat [1] proposed a texture based medical image indexing and retrieval, which is based on the application to cardiac images. The goal of this paper is to perform clinically relevant queries on large image databases that do not require user supervision. The work shown by B.S.Manjunath, Jens-Rainer Ohm, Vinod V.Vasudevan and Akio Yamanda [2] was mixture of colour, texture and edge density for MPEG-7 standards. Zhe-Ming Lul, Su-Zhi Li, and Hans Burkhardt [3] proposed a CBIR Scheme in JPEG Compressed Domain. Minyoung Eom, and Yoonsik Choe [4] used edge histogram for the fast extraction of edge histogram in DCT Domain based on MPEG7.

Son Lam Phung and Abdesselam Bouzerdoum [5] explained a similar kind of approach based on edge density for detecting peoples and they expressed a new image feature for fast detection of people in Images. Paul Stefan and J.Kaufman [6] used the CBIR approach for the Segmentation of natural

Images and also the various segmentation technologies. P.S.Hiremath and Jagadeesh Purjari [7] proposed a work based on CBIR using colour, texture and shape features. Remco C. Velcamp, and Mirela Tanse [8] has given a detailed survey on CBIR Systems. The purpose of this survey is to provide an overview of the functionality of temporary image retrieval systems in terms of technical aspects. Mustafa Ozden and Ediz Polat [9] described a new colour image segmentation method based on low-level features including colour, texture and spatial information. This paper also introduces a method called wavelet frames that provide translation of invariant texture analysis. John Montagnat F. Bellet, H. Benoit-Cattin [10] proposed a method that explains Texture-based medical Image Indexing and Retrieval on Grids.

C.R. shyu [11] proved that local features based methods yields good results. For a successful CBIR, the indexing scheme should be efficient for searching the database. Recent retrieval systems have incorporated user's relevance feedback to modify the retrieval process in order to generate perceptually and semantically more meaningful retrieval results. Roger Weber and Michael Milivoncic [12] proposed an effective and efficient region based image retrieval. Current research works attempt to obtain and use the semantics of image to perform better retrieval. Towards this goal, segmentation of an image into regions has been used in recent years. S.L Phung and A.Bouzerdoun [13] proposed an edge density approach for detecting people in Images.

Bohyung Han and Changjiang, Yang Ramani Duraiswami and Larry Davis [14] Yang use a Bayesian Filtering and Integral image for Visual Tracking. They used a new method of tracking techniques for visual based objects. Vincent arvis and Christophe Debain [15] given a different approach for generalization of the co occurrence matrix for colour image application to colour texture classification. The method used in this paper is tested based on extensions of the co occurrence matrix method. Alberto Amato and Vincenzo Di Lecce [16] proposed edge detection techniques in Image retrieval. Thomas M.Lehman [17] proposed an automatic segmentation of medical images for CBIR and data-mining. Dong Yin, and Jia Pan [18] proposed medical image categorization based on Gaussian mixture model. Fuhui Long, Hongjiang Zhang and David Dagan Feng [19] explained the fundamentals of CBIR. To make image retrieval faster several indexing structures were designed.

3. SYSTEM OVERVIEW OF CONTENT BASED IMAGE RETRIEVAL

Figure 1 shows the block diagram and the steps involved in the proposed method.

3.1 Feature Extraction of Colour

Colour histogram is the most commonly used feature in image retrieval systems and it is, one of the straight-forward features utilized by human beings for visual recognition and discrimination. Statistically, it denotes the joint probability of the intensities of the three colour channels. Once the image is segmented, from each region the colour histogram is extracted. The major statistical data extracted are histogram mean, standard deviation, and median for each colour channel i.e. Red, Green, and Blue. So, totally $3*3=9$ features per segments are obtained. All the segments need not be considered, but only segments that are dominant may be

considered, because this would speed up the calculation and may not significantly affect the end result.

3.1.2 The RGB Colour Model

The RGB colour model uses three primary colours, red, green and blue, and it can able to reproduce other colours. In a true-colour image each pixel will have red, green and blue value ranging from 0 to 255 giving a total of 16777216 different colours. One disadvantage with the RGB model is its behaviour when the illumination when an image changes. The distribution of RGB values will change proportionally with the illumination, thus giving a very different histogram.

3.1.2 RGB Histogram

Colour histogram is one of the most common method used in image retrieval systems. The RGB histogram uses, as the name suggests, the RGB colour model. In Proposed system $8x8x8$ -bin histogram is used to quantize the images, giving a total of 512 different bins.

3.1.3 HSV Colour Model

HSV defines a colour by its hue saturation value rather than by primary colours. The hue ranges from 0 to 255 and indicates colour-type. Saturation defines the vibrancy of the colour and ranges from 0 to 100%. It describes colours similar to how humans perceive them. This perceptual relevance might make it a better choice than RGB for image retrieval. Calculation of HSV is done by a non-linear transformation of RGB.

3.1.4 Feature Extraction of colour by

K-Means Algorithm

K-Means clustering is an algorithm to group objects based on attributes/features into k number of groups where k is a positive integer. The grouping (clustering) is done by minimizing the Euclidean distance between data and the corresponding cluster centroid. Thus the purpose of k-means clustering is to cluster the data. K-means clustering algorithm, in its standard formulation consists mainly of four steps.

1. Initialize k centroids.
2. Compute the distance between each feature vector to the centroids.
3. Assign the feature vector to the centroid whose distance is minimum.
4. Re-calculate the centroids.

The aim of K-Means algorithm is the minimization of an objective function that samples the closeness between the data points and the cluster centres, and it is calculated as follows:

$$J = \sum ||X_i - C_i||^2 \quad (1)$$

Where $||X_i - C_i||^2$ is the distance between the data point and the cluster centre. As it can be easily observed by equation 1, the assignment of the data points may not be unique. In this case K-Means algorithm doesn't find the optimal solution corresponding to the global objective function J. In addition, it is sensitive to the initialization process that selects the initial cluster centres. Since the pixel assignment is performed only by evaluating the colour information in a certain colour space, the connection between the data point under evaluation and its neighbours is not taken into account; it will lead to a partition of the input data into regions that are not related to the scene object components but also the local colour-texture complexity, which allows to obtain homogenous clusters. Since the random selection of the initial cluster centres from image data is not an appropriate solution, a new scheme to perform the initialization for the k-means algorithm with the dominant colour components that are extracted from the colour histograms of the input image is developed. In this

regard, histogram for each colour channel are constructed and these are partitioned linearly into R sections (where R is a

fixed value, $R > K$ and n_k is the number of pixels contained in the bin k).

$$P_j = \arg \max(n_k) \quad (2)$$

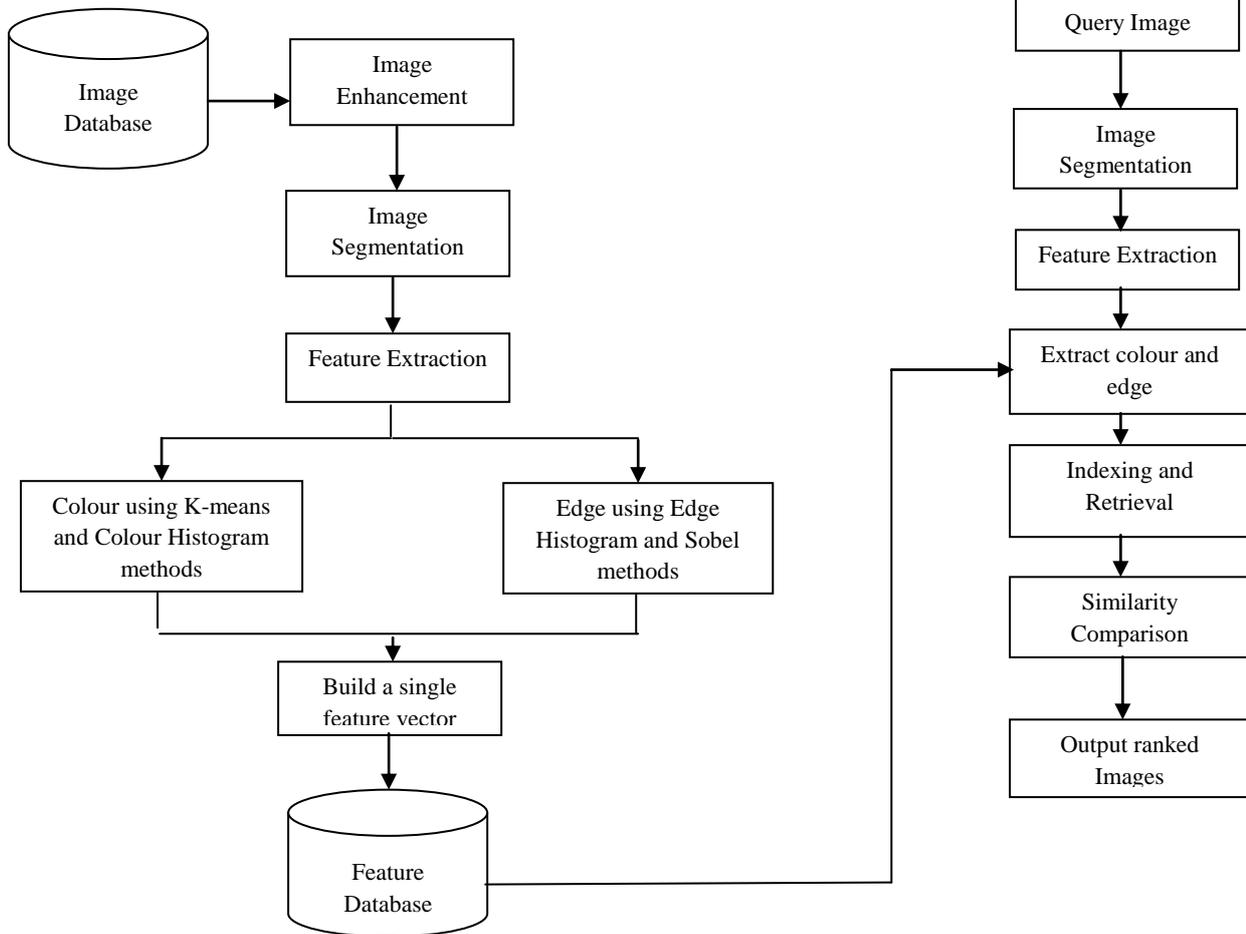


Fig 1: Block diagram of the proposed efficient model for CBIR

The process of feature extraction of colour using k-means clustering algorithm is summarized by the following pseudo-code sequence.

1. Construct the histograms for each colour channel.
2. Partition each histogram into R sections.
3. Compute the peaks in each section and rank the peaks p_1, p_2, \dots, p_r where p_r has the highest number of elements.
4. Start form the colour seeds for highest peak p_i . If ($p_i \rightarrow$ red) mark the pixels in the red channel and calculate gmean and bmean for marked pixels from green and blue channels. If ($p_i \rightarrow$ green) mark the pixels in the green channel and calculate the rmean and bmean for marked pixels from red and blue channels. ($p_i \rightarrow$ blue) mark the pixels in the blue channel and calculate the rmean and gmean for marked pixels from red and green channels.
5. From the colour seed and eliminate p_i from the list.
6. Repeat the steps 4 and 5 until the desired number of colour seeds has been reached.

3.2. FEATURE EXTRACTION OF EDGE

An edge is the boundary between two regions with relatively distinct gray-level properties. Edge detection is the most common approach for detecting meaningful discontinuities in gray-level. The reason is that isolated points and thin lines are not frequent occurrences in most practical applications.

Basically the idea underlying most edge-detection techniques is the computation of a local derivative operator. The coarseness of random texture can also be represented by the density of edge pixels. If an edge map is given, the edge density is measured by the average number of edge pixels per unit area.

3.2.1 Edge Histograms

Histogram is the most commonly used characteristic to represent the global feature composition of an image. It is invariant to translation and rotation of the images and normalizing the histogram leads to scale invariance. Exploiting the above properties, the histogram is considered to be very useful for indexing and retrieving images. Edge in the image is considered an important feature to represent the content of the image. Human eyes are known to be sensitive to edge features for image perception. In MPEG-7, there is a descriptor for edge distribution in the image. This edge histogram descriptor proposed for MPEG-7 consist only local edge distribution in the image. That is, since it is important to keep the size of the histogram as small as possible for the efficient storage of the metadata, the normative edge histogram for MPEG-7 is designed to contain only local edge distribution with 80 bins. These 80 histogram bins are the only standardized semantics for the MPEG-7 edge histogram descriptor. The normative part of the edge histogram

descriptor consists of 80 local edge histogram bins. To localize edge distribution to a certain area of the image, divide the image space into 4x4 sub-images. Then, for each sub-image, generate an edge histogram to represent edge distribution in the sub-image. To define different edge types, the sub-image is further divided into small square blocks called image-blocks. After the edge extraction from image-blocks, count the total number of edges for each edge type in each sub-image. Since there are five different edges, five histogram bins for each sub-image are defined. Then, since there are 4x4=16 sub-images, 6x5=80 bins available for edge histogram.

3.2.2 Normalization and Quantization of Bins

After generating local edge histograms for all 16 sub-images, normalization of each bin in the histogram is done by dividing it with the total number of image-blocks with an edge in the corresponding sub-image. Then, each histogram bin has a value ranging from 0 to 1. To represent the normalized bin values in binary form, quantization is done. Since the normalized bin values are distributed in a small range, bin values are non-linearly quantized. The quantization tables are obtained by adopting the Lloyd-Max algorithm. Then, assigning 3 bits per bin totally 3x80=240 bits are used to represent the local histogram.

3.2.3 Edge Extraction Method

For edge extraction, Edge Histogram Descriptor (EHD) algorithm is used. To extract both directional non-directional edge features, proposed system defines a small square image-block. That is, an image space is divided into non-overlapping square blocks and then edge information from each block is extracted. Note that, regardless of the image size, the sub-image is divided into a fixed number of image-blocks. That is, the size of the image-block is proportional to the size of original image to deal with the images with different resolutions. The size of image-block is assumed to be a multiple of 2. If it is not a multiple of 2, simply proposed system ignores some outmost pixels so that the image-block becomes a multiple of 2.

$$\text{block-size} = \frac{x}{3 \times 2} \quad (3)$$

Here, image width and image height represent the horizontal and vertical sizes of the image respectively. The desired number-of-block is given and fixed to cope with the various image resolutions. Edge feature is extracted from the image-block as shown in Figure 2. Here, the image-block is further divided into four sub-blocks. Then, the luminance mean values for the four sub-blocks are used for the edge detection. More specifically, mean values of the four sub-blocks are obtained, and they are convolved with filter coefficients to obtain edge magnitudes. More specifically, label the sub-blocks from 0 to 3 for the k^{th} ($k=0,1,2,3$) sub-block of the $(i, j)^{\text{th}}$ image block, proposed system calculates the average gray level $A_k(i, j)$. Adopting the same labeling order proposed system calculates coefficients of the vertical edge filter, horizontal edge filter, dia45 edge filter, dia135 edge filter and nond edge filter as follows.

$$\begin{aligned} \text{ver_edge_filter}(0) &= 1 \\ \text{ver_edge_filter}(1) &= -1 \\ \text{ver_edge_filter}(2) &= 1 \\ \text{ver_edge_filter}(3) &= -1 \\ \text{hor_edge_filter}(0) &= 1 \\ \text{hor_edge_filter}(1) &= 1 \\ \text{hor_edge_filter}(2) &= -1 \end{aligned}$$

$$\begin{aligned} \text{hor_edge_filter}(3) &= -1 \\ \text{dia45_edge_filter}(0) &= \text{sqrt}(2) \\ \text{dia45_edge_filter}(1) &= 0 \\ \text{dia45_edge_filter}(2) &= 0 \\ \text{dia45_edge_filter}(3) &= \text{sqrt}(2) \\ \text{dia135_edge_filter}(0) &= 0 \\ \text{dia135_edge_filter}(1) &= \text{sqrt}(2) \\ \text{dia135_edge_filter}(2) &= \text{sqrt}(2) \\ \text{dia135_edge_filter}(3) &= 0 \\ \text{nond_edge_filter}(0) &= 2 \\ \text{nond_edge_filter}(1) &= -2 \\ \text{nond_edge_filter}(2) &= -2 \\ \text{nond_edge_filter}(3) &= 2 \end{aligned}$$

Using the above five edge filters, proposed system obtain five edge strengths for the image block (i, j) as follows.

$$\begin{aligned} \text{ver_edge_stg}(i, j) &= 0^3 |A_k(i, j) \times \text{ver_edge_filter}(k)| \quad (4) \\ \text{hor_edge_stg}(i, j) &= 0^3 |A_k(i, j) \times \text{hor_edge_filter}(k)| \end{aligned}$$

$$\begin{aligned} \text{dia45_edge_stg}(i, j) &= 0^3 |A_k(i, j) \times \text{dia}[[45]]_(\text{edge_filter}(k) \text{ +} | | \quad (6) \end{aligned}$$

$$\begin{aligned} \text{dia}[[135]]_(\text{edge_stg}(i, j) &= 0^3 |A_k(i, j) \times \text{dia}[[135]]_(\text{edge_filter}(k) \text{)} | \quad (7) \end{aligned}$$

$$\begin{aligned} \text{nond_edge_stg}(i, j) &= 0^3 |A_k(i, j) \times \text{nond_edge_filter}(k) \text{)} | \quad (8) \end{aligned}$$

If the maximum value among five edge strengths obtained from equations (4) to (8) is greater than a threshold (Th_{edge}), then the image-block is considered to have the corresponding edge in it.

$$\begin{aligned} \text{max}\{\text{ver_edge_stg}(i, j), \text{hor_edge_stg}(i, j), \text{dia45_edge_stg}(i, j), \\ \text{dia135_edge_stg}(i, j), \text{nond_edge_stg}(i, j)\} > Th_{edge} \quad (9) \end{aligned}$$

4. IMPLEMENTATION PROCEDURE

4.1. Retrieval of Images Using K-Means

Algorithm

K-means clustering is used for the classification of feature set obtained from the histogram refinement method. Histogram refinement provides a set of features for proposed CBIR system. Global histograms are used for image retrieval, because of their effectiveness and insensibility. Figure 2 shows the retrieval of similar images using K-means algorithm.

4.2. Retrieval of Images Using Colour

Histogram Method

This approach is more frequently adopted for CBIR systems based on the conventional colour histogram (CCH). Each pixel is associated to a specific histogram bin only on the basis of its own colour. Colour similarities across different bins or colour dissimilarity in the same bin are not taken into account. Figure 3 shows the retrieval images using colour histogram method.

4.3. Retrieval of Images Using Sobel Method

The sobel operator is based on convolving the image with a small, separable, and integer valued filter in horizontal and vertical direction. In simple terms, sobel operator calculates the gradient of the image intensity at each point, giving the direction of the largest possible increase from light to dark and the rate of change in that direction. At each image point, the gradient vector points in the direction of largest possible intensity increase, and the length of the gradient vector corresponds to the rate of change in that direction. In the proposed system the shape feature vector is formed by applying slope magnitude method on Gradient of an image in vertical and horizontal directions. This feature vector is considered for finding the similarity of query image with database images. Figure 4 shows retrieval of similar images using sobel method.

4.4. Retrieval of Images Using Edge Histogram Method

To achieve a high retrieval performance, the local histogram alone may not be enough. Proposed system calculates edge distribution information for the whole image space and some horizontal and vertical semi-global edge distributions. Since there are five edge types, the global edge histogram also contains five bins. For the semi-global edge histograms, proposed system cluster four connected sub-images. There are 13 different clusters and for each cluster proposed system generates edge distributions for five different edge types. So, totally 80 bins (local) + 5 bins (global) + 65 bins (13x5, semi-global) = 150 bins. The bin values for all global and semi-global histograms can be obtained directly from the local histogram. Figure 5 shows retrieval of similar images using edge histogram method.

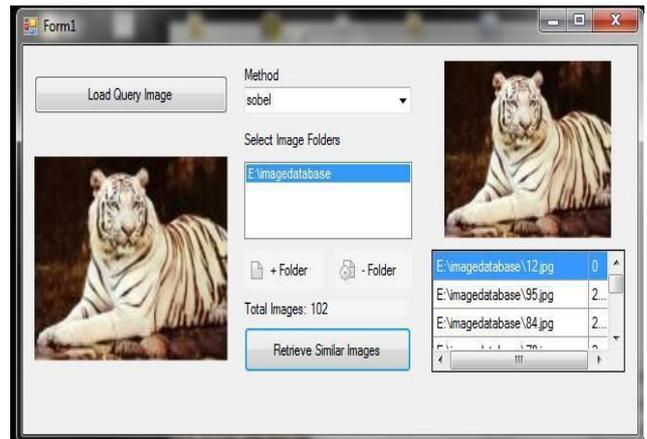


Figure 4. Image retrieval using sobel method

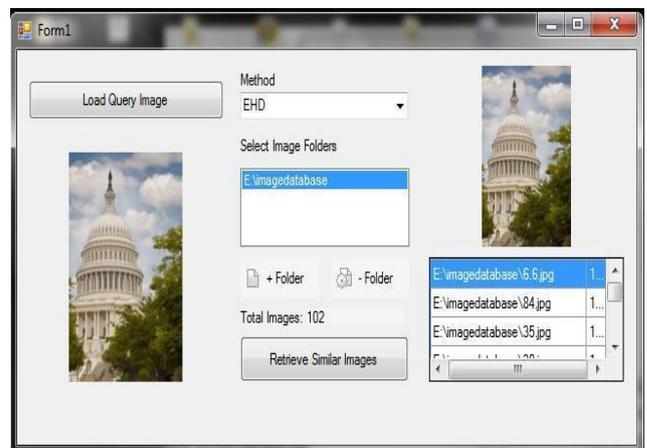


Figure 5. Image retrieval using EHD method

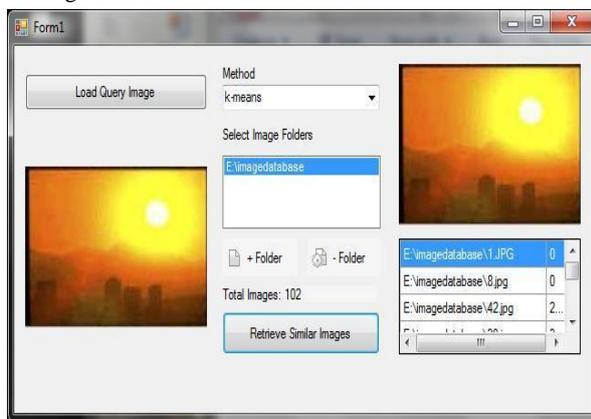


Figure 2. Image retrieval using K-means method

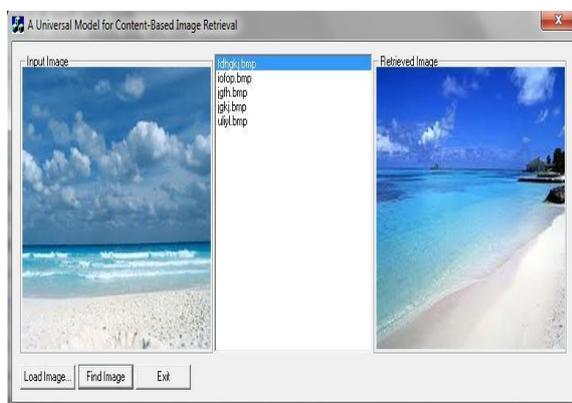
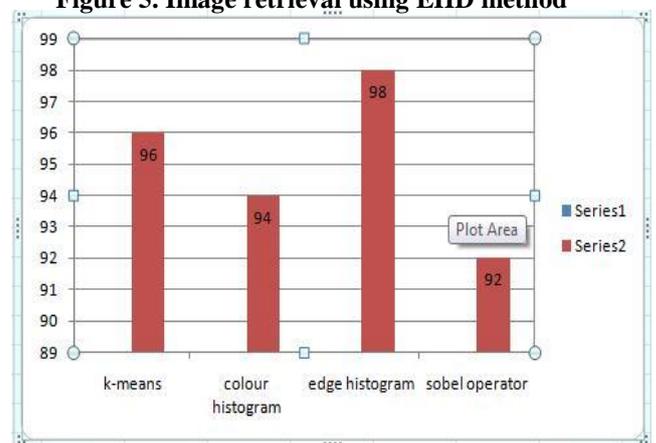


Figure 3. Image retrieval using Colour Histogram method



Graph 1. Performance Comparison of different methods used for Image retrieval

Graph 1 shows the performance in Y-axis and the different methods in X-axis. To assess the retrieval effectiveness, the precision and recall values are used. K-means, colour histogram, edge histogram descriptor and sobel methods are compared. Edge histogram descriptor method gives better performance and retrieval efficiency when compared with K-means, colour histogram and Sobel methods.

5. CONCLUSION

An efficient model for Content Based Image Retrieval System by combining the colour, and edge density features are proposed. Proposed system involves three modules feature extraction, image retrieval and performance comparison based on similarity. For colour extraction K-means clustering algorithm is used since it attempts to minimize the errors in the assignment of the data points into clusters by evaluating the local texture complexity. For the extraction of an edge an edge histogram descriptor algorithm is used. In image retrieval four methods are used such as k-means, colour histogram, edge histogram and sobel method. The advantages of global and local features together have been utilized for better retrieval efficiency. To assess the retrieval effectiveness, the precision and recall are used as statistical comparison parameters. From the experiments edge histogram descriptor method gives better performance. The results are quite good for most of the query images and it is possible to further improve by fine tuning the threshold and adding relevance feedback.

6. REFERENCES

- [1] Tristan Glatard and John Montagnat, "Texture based medical image indexing and retrieval: application to cardiac imaging", Proceedings of the 6th ACM SIGMM international workshop on Multimedia information retrieval, 2004, ISBN:1-58113-940-3.
- [2] B. S. Manjunath , Jens-rainer Ohm , Vinod V. Vasudevan , Akio Yamada, "Colour and Texture Descriptors", IEEE Transactions on Circuits and Systems for Video Technology, 1998.
- [3] Zhe-Ming Lul, Su-Zhi Li, and Hans Burkhardt, "A Content-Based Image Retrieval Scheme in JPEG Domain", International Journal of Innovative Computing, Information and Control ICIC International, Volume 2, Number 4, August 2006.
- [4] Minyoung Eom , and Yoonsik Choe , "Fast Extraction of Edge Histogram in DCT Domain based on MPEG7", World Academy of Science, Engineering and Technology 9 2005.
- [5] Son Lam Phung, Abdesselam Bouzerdoum, "Detecting People in Images: An Edge Density Approach", ICASSP (1) 2007, 1229-1232.
- [6] Paul Stefan and J.Kaufman, "Segmentation of Natural Images for CBIR", Department of Electrical and Electronics Engineering, University of Western Australia.
- [7] P. S. Hiremath and Jagadeesh Pujari, "Content Based Image Retrieval based on colour, texture and shape features using Image and its complement", International Journal of Computer Science and Security, Volume (1) : Issue (4).
- [8] Remco C. Veltcamp, and Mirela Tanse, "Content Based Image Retrieval Systems: A survey", 2000.
- [9] Mustafa Ozden and Ediz Polat, "Acolour image segmentation approach for content-based image retrieval", Pattern Recognition 40 (2007) 1318 – 1325.
- [10] J. Montagnat, F. Bellet, H. Benoit-Cattin, "Medical images simulation, storage, and processing on the European DataGrid testbed, Kluwer Academic Publishers, 2004.
- [11] J. G. Dy, C. E. Brodley, A. Kak, C. Shyu and L. S. Broderick, "The Customized-Queries Approach to CBIR Using EM".
- [12] Roger Weber and Michael Milivoncic, "Efficient region-based image retrieval", Proceedings of the twelfth international conference on Information and knowledge management, 2003.
- [13] S.L Phung and A.Bouzerdoum and Douglas Chai "Skin Segmentation Using Colour Pixel Classification: Analysis and Comparison", IEEE transactions on pattern analysis and Machine Intelligence, vol. 27, No. 1, January 2005.
- [14] Bohyung Han Changjiang Yang Ramani Duraiswami and Larry Davis, "Bayesian Filtering and Integral Image for Visual Tracking".
- [15] Vincent Arvis, Christophe Debain, Michel Berducot and Albert Benassi, "Generalization of the Cooccurrence Matrix for Colour images: Application to colour texture Classification", Image Anal Stereol 2004;23:63-72.
- [16] Alberto Amato and Vincenzo Di Lecce, "An Image Retrieval Interface Based On Dynamic Knowledge", Proceedings of IEEE Conference on Computational Intelligence for Modelling, Control and Automation, 2005.
- [17] Thomas M.Lehman, "Segmentation of medical images combining local, regional, global, and hierarchical distances into a bottom-up region merging scheme, Proceedings of SPIE symposium on Medical Imaging (MI'05):Image Processing, Vol 5747(56).
- [18] Dong Yin, and Jia Pan, "Medical Image Categorization Based on Gaussian Mixture Model", Proceedings of IEEE conference on Biomedical Engineering and Informatics, 2008, pp 128-131.
- [19] Fuhui Long, Hongjiang Zhang and David Dagan Feng, "Fundamentals of Content-Based Image Retrieval".