IJISET - International Journal of Innovative Science, Engineering & Technology, Vol. 2 Issue 1, January 2015.
www.ijiset.com
ISSN 2348 – 7968

Content Based Medical Image Retrieval Using Fuzzy C- Means Clustering With RF

Jasmine Samraj[#1], NazreenBee . M [*2]

[#]Associate Professor, Department of Computer Science,
Quaid-E-Millath Government college for Women, Chennai, Tamil Nadu, India

[*]Research Scholar
Quaid-E-Millath Government college for Women, Chennai, Tamil Nadu, India

Abstract—Content-based image retrieval is one of the techniques of image mining. Content-based image retrieval system (CBIR) has been proposed by the medical community to manage the storage and distribution of images to radiologists, physicians, specialists, clinics, and imaging centers. There are three fundamental steps for Content Based Image Retrieval. They are Visual Feature Extraction, Similarity Measurement and Retrieval System Design. Content-based image retrieval with relevance feedback (RF) schemes based on Fuzzy C-Means Clustering is used to retrieve the medical image effectively and efficiently. Data clustering is a process of separating similar data into groups. Fuzzy C-Means Clustering (FCM) is useful to mine complex and multi-dimensional data sets. This technique allows users to retrieve a similar query image from a database, thus ahigher retrieval performance can be achieved and also comparing c-means and k-means clustering techniques for MRI scan images

Keywords:— Content-based image retrieval, shape and texture extraction, fuzzy c-means clustering, similarity measurement, relevance feedback.

1. INTRODUCTION

CBIR uses the visual contents of an image such as color, shape, texture, and spatial layout to represent an index the image. In addition to their development, efforts are also being made to evaluate the performance of image retrieval systems. There are different ways to retrieve the images in CBIR

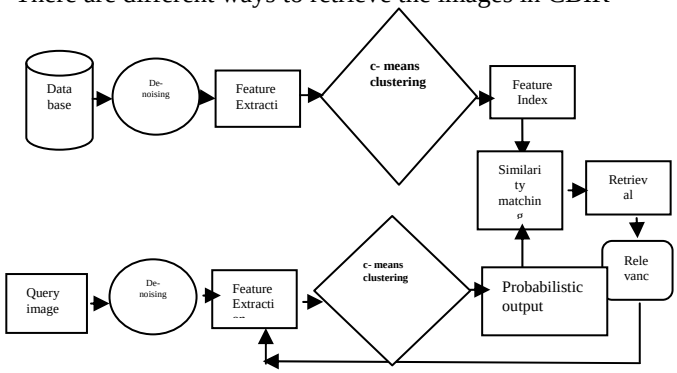

Fig 1. SYSTEM ARCHITECTURE

The oldest method is text annotation to images in the database [1]. Image annotation is a tedious task. To address such significant limitations, researchers have turned their attention to content-based image retrieval. In CBIR systems, low level image features are extracted based on visual content such as color, shape and texture [2& 3].

In order to reduce the gap between the low level features and high level concepts, relevance feedback was introduced into CBIR. The proposed system can be divided into two parts, off-line and on-line approach. In the on-line approach, a query image is De-noised using a non-linear filtering technique which is useful to find the edge of an image efficiently. An image can be corrupted during the transmission of image from one place to another place due to the moisture present in the environment. A De-noised image can effectively extract the image feature such as shape and texture. Then the image is classified using a Fuzzy C-Means Clustering (FCM)to separate a gray and white color present in the image. The clustered image is also used to measure the affected region of an image. In the off-line approach, the same steps can be processed and pre-classify a dataset using a FCM technique, which is useful to mine complex and multi-dimensional dataset present in an image database. Retrieval of image based upon a similarity between the query image and database image can be measured using the Euclidean distance measures.Relevance Feedback (RF) has been widely used to narrow the semantic gap between low-level visual features and high-level human perception.

2. SYSTEM ARCHITECTURE

Fig. 1 shows thescheme of image retrieval from a database using relevance feedback. In CBIR, each image that is stored in the database has its features extracted and compared to the features of the query image. It involves two steps:

ISSN 2348 – 7968

FEATURE EXTRACTION: The first step in the process is extracting image features to a distinguishable extent.

SIMILARITY MATCHING: The second step involves matching these features to yield a result that is visually similar.

3. PROPOSED WORK

The purpose of this proposed work was to improve the accuracy (precision) of a CBIR application by allowing the system to retrieve image similar to the source image. The proposed framework is given in Fig 2.

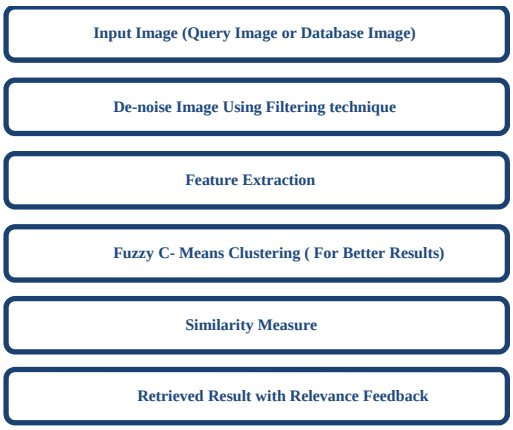

Fig 2. Proposed Framework Design

The input query image is taken for the process of retrieval of similar image from the large amount of medical images in the database.

3.1 DE-NOISING:

De-noising process can be done by filtering the corrupted image to remove noise of an image which is mostly introduced due to external noise.Focus on pre-processing the medical image is important since the noise issues changes image pixel value either on or off [14].

The median filter is a non-linear digital filtering technique, often used to remove noise. Such noise reduction is a typical pre-processing step to improve the results forlater processing. Median filter is better for medical image based upon MSE (Mean Square Error) and PSNR(Peak Signal Noise Ratio)value. Image quality is a characteristic of an image that measures the perceived image degradation (typical, compared to another filtering technique).The following plot Fig 3. illustrates that Median filter is better than compared to averaging and adaptive filter based on MSE and PSNR value.

Fig 3. Image Quality Measurement

3.2 FEATURE EXTRACTION:

The features extracted are carefully chosen and it is expected that the feature set will extract the relevant information from the input data in order to perform the desired task using this reduced representation instead of the full size input.

3.2.1 SHAPE EXTRACTION:

Shape is an important and most powerful feature used for feature extraction. The use of object shape is one of the most challenging problems in creating efficient CBIR. The object's shape plays a critical role in searching for similar image objects. In particular, image regions occupied by an object have to be found in order to describe its shape [5 & 6].For detecting edges Canny Edge Detection Algorithm is used and connecting the edges to form the shape of the image. Boundaries are linked edges that characterize the shape of an object

3.2.2 TEXTURE EXTRACTION:

Texture is a combination of repeated patterns with regular frequency [9,10& 11]. In visual interpretation texture has several types, for examples, smooth, coarse, fine etc. The Texture feature extraction is very computationally intensive for individual pixels.

An image texture is a set of metrics calculated in image processing designed to quantify the perceived texture of an image. Image Texture gives us information about the spatial arrangement of color or intensities in an image or selected region of an image. Image textures can be artificially created or found in natural scenes captured in an image.

3.3 CLUSTERING USING FUZZY C-MEANS CLUSTERING:

Fuzzy C- means Clustering (FCM) is a method of clustering which allows one piece of data to belong to one or more clusters as in Fig 4.

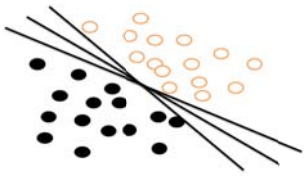

Fig.4 Clustering using FCM

This method is frequently used in pattern recognition [12, 13, & 15]. They are several methods for gray level image segmenting, based on this they are two approaches.

- **BASED ON DISCONTINUITY**
- **BASED ON SIMILARITY**

BASED ON DISCONTINUITY

First approach uses the discontinuities between gray level region to detect isolated points, edges and contours.

BASED ON SIMILARITY

Second approach includes clustering, thresholding, region growing, region splitting and merging.

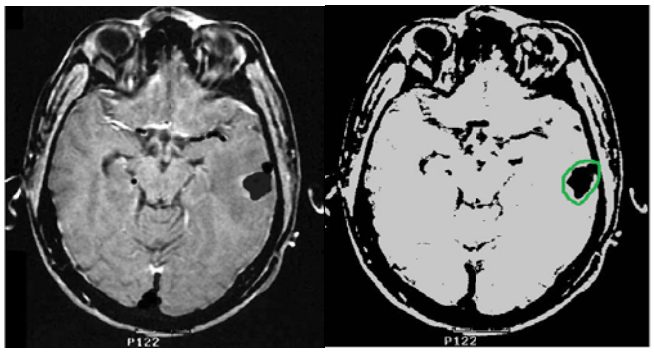

Fig.5 Segmenting using FCM

Additional process using Fuzzy C-means clustering is done which accurately clusters the image.

3.4CLASSIFICATION USING FUZZY C-MEANS CLUSTERING

Image classification is one of the important steps in the image retrieval process because it saves more time while searching the images from the huge volume of the database [17 &18]. Classification is identification of different regions of the image by which the retrieval efficiency of the system will be improved. Data Clustering is considered an interesting approach for finding similarities in data and putting similar data into groups. Clustering partitions a data set into several groups such that the similarity within a group is larger than that among groups. Fig.5 shows the segmenting using FCM.

` The Fuzzy C-Means clustering algorithm was first introduced by Dunn and later was extended by Bezdek. Fuzzy C-means clustering relies on the basic idea of Hard C-means clustering (HCM), with the difference that in FCM each data point belongs to a cluster to a degree of membership grade, while in HCM every data point either belongs to a certain cluster or not. The FCM objective function and its generalizations are the most heavily studied fuzzy models in Pattern Recognition. Clustering is a process for classifying objects or patterns in such a way that samples of the same group are more similar to one another than samples belonging to different groups. Many clustering strategies have been used, such as the hard clustering scheme and the fuzzy clustering scheme, each of which has its own special characteristics. The conventional hard clustering method restricts each point of the data set to exclusively just one cluster. As a consequence, with this approach the segmentation results are often very crisp, i.e., each pixel of the image belongs to exactly just one class.

\

3.5 SIMILARITY MATCHING:

The most common method for comparing two images in content-based image retrieval (typically an example image and an image from the database) is using an image distance measure [1]. An image distance measure compares the similarity of two images in various dimensions such as color, texture, shape, and others. For example a distance of 0 signifies an exact match with the query, with respect to the dimensions that were considered. As one may intuitively gather, a value greater than 0 indicates various degrees of similarities between the images. Euclidean Measure distance between point's p and q is the length of the line segment connecting them ($\overline{pq}$).

ISSN 2348 – 7968

3.6 RELEVANCE FEEDBACK:

Relevance feedback (RF) can be used to make a bridge between machine learning and human perception. The concept of relevance feedback was introduced into CBIR from the concept of text-based information retrieval in the 1998's and then has become a popular technique in CBIR. In RF, the user assesses the previously retrieved images as similar or dissimilar to the initial query and provides this assessment as feedback to the system [12 &19].

Relevance feedback is a supervised active learning technique used to improve the effectiveness of information systems. The main idea is to use relevant and irrelevant answer from the user to improve system performance. The typical algorithm for relevance feedback in CBIR is as follows.

Algorithm:
Begin
Obtain the retrieval results of CBIR system
From user interaction, obtain the feedback from the users on prior results. Feedback is in the form of relevant or irrelevant to request.
If
Results found to be not satisfied learn the system through a feedback algorithm and hence results are refined
Else Result found to be satisfied
End
End

RELATED WORK IN RELEVANCE FEEDBACK

3.7 PERFORMANCE METRICS:

CBIR is essentially an information retrieval problem. Two of the most popular evaluation measures are the precision and recall [14]. The precision measures the proportion of the total images retrieved which are relevant to the query

$$Precision = \frac{No.of relevant images retrieved}{Total no.of images Retrieved}$$

The recall measure is defined as the fraction of the all relevant images.

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

High precision means that less irrelevant images are returned or more relevant images are retrieved, while high recall indicates that few relevant images are missed.

4. RESULTS:

ON-LINE PROCESS IN IMAGE RETRIEVAL

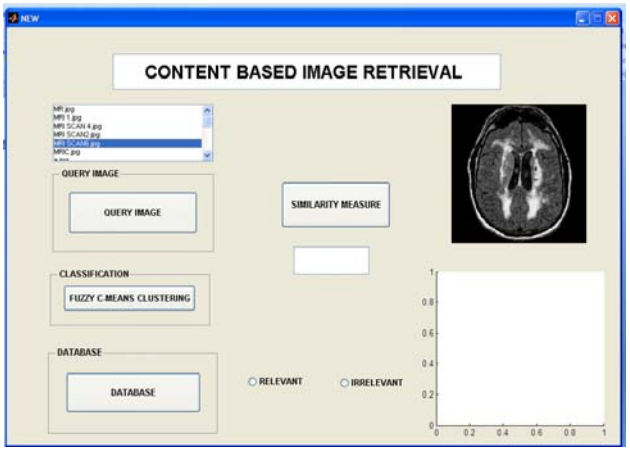

Fig.5 CBIR SYSTEM

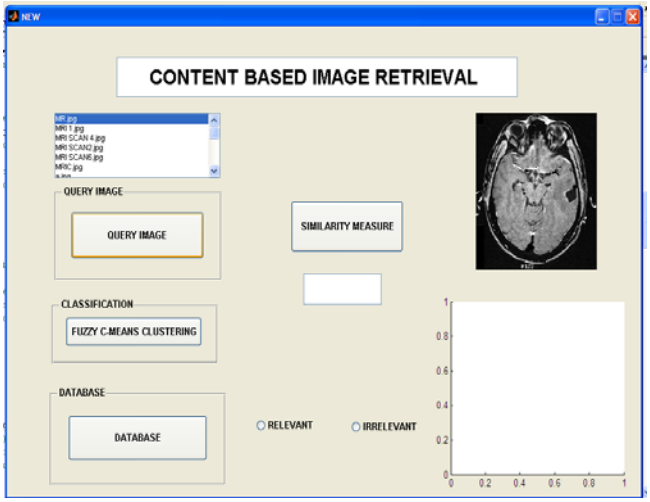

Fig. 6Inserting Query Image

OFF-LINE PROCESS IN IMAGE RETRIEVAL

ISSN 2348 – 7968

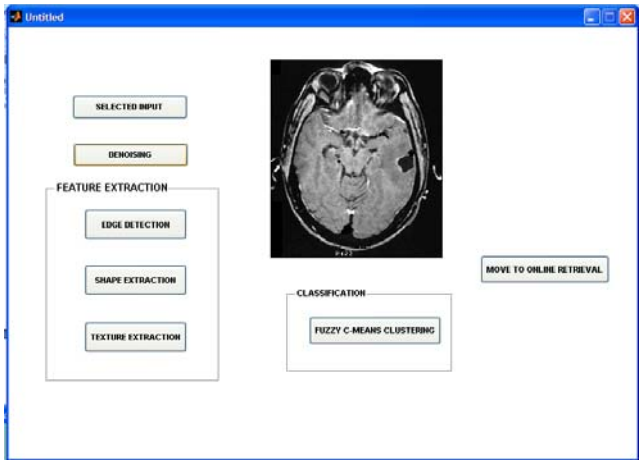

Fig.7De-noising

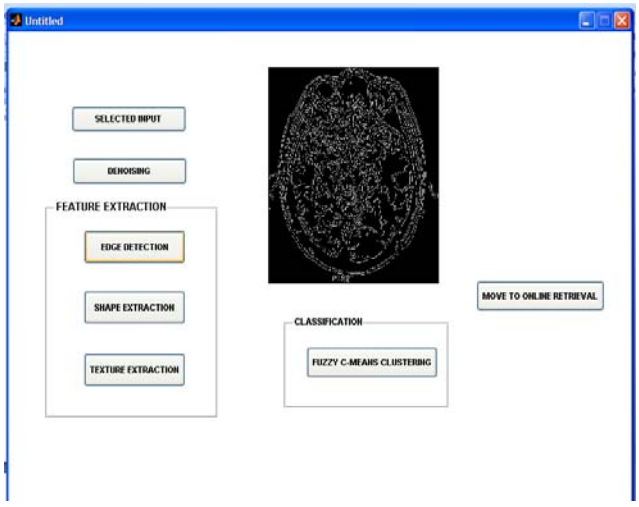

Fig.8 Edge Detection

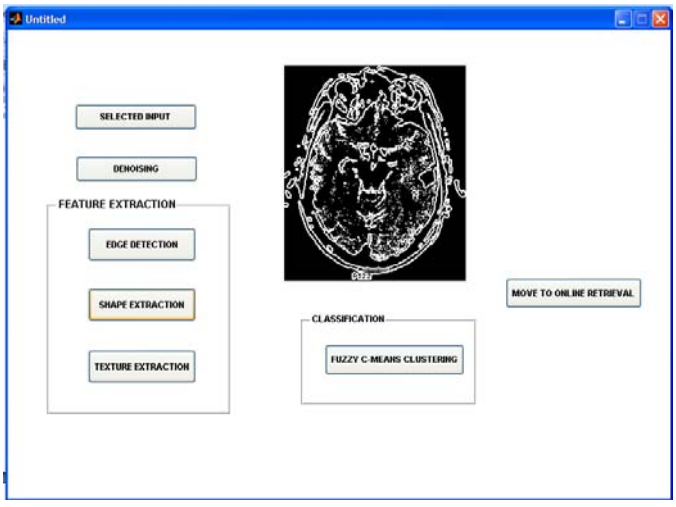

Fig. 9 Shape Feature Extraction

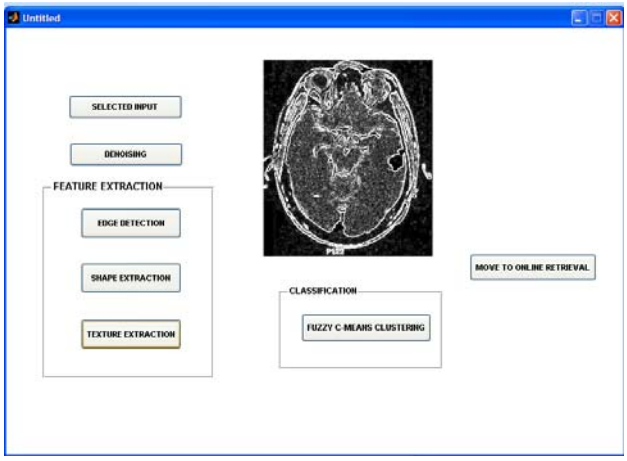

Fig.10Texture Feature Extraction

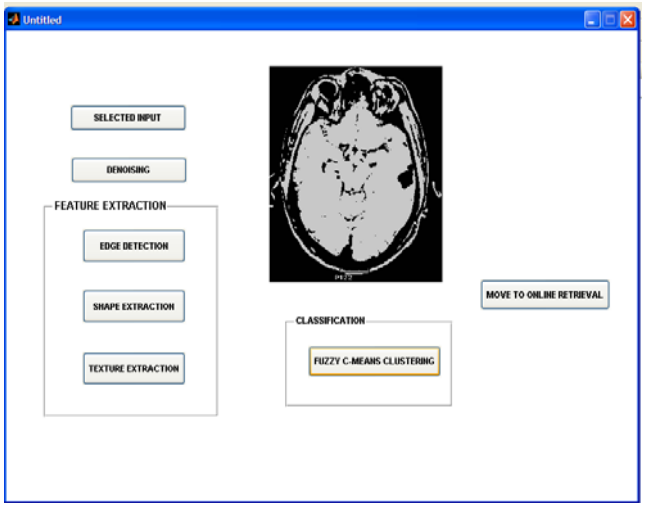

Fig.11 Clustering Using FCM

Move to On-line Process and Retrieval of similar Result from the pre-classified database images, the same off-line step is processed in the database images. The processed query image is compared with the pre-classified database image based upon the similarity measurement.

ISSN 2348 – 7968

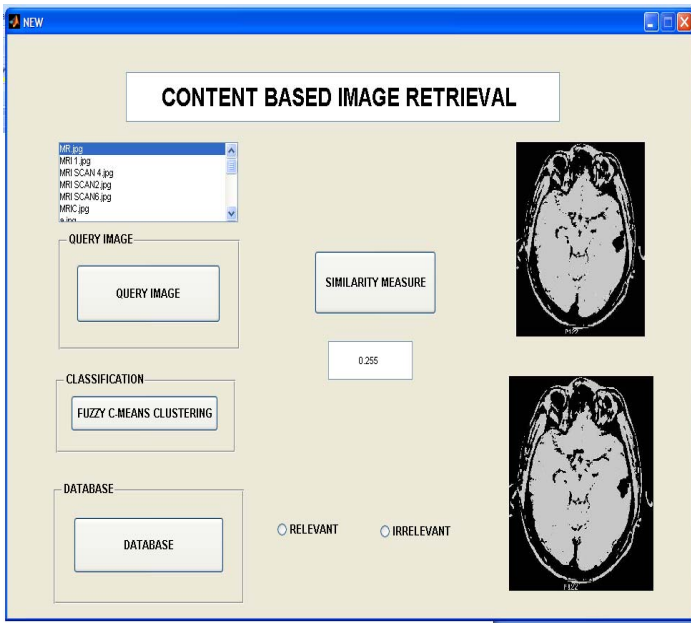

Fig.12Pre- Classified database Images Using FCM and the Retrieval of Similar Result based upon the Euclidean Distance measurement

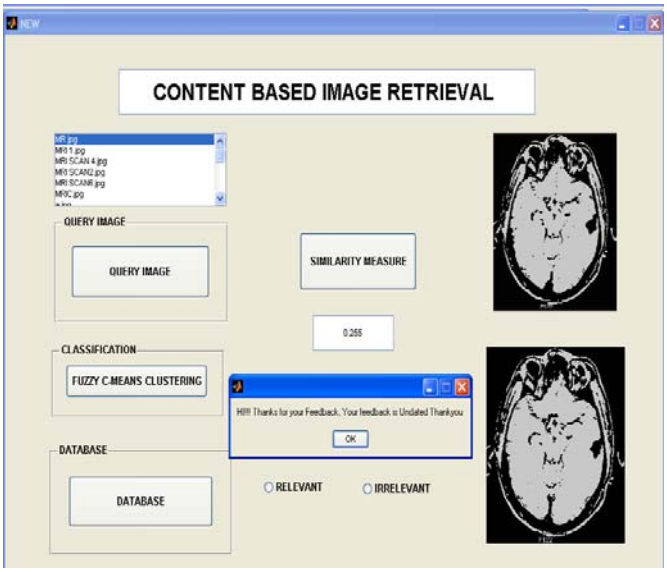

Fig. 13 Retrieval of Similar Image from a Database, measured similarity between the two images and Relevance Feedback can be updated

5. COMPARING C-MEANS AND K-MEANS CLUSTERING

Clustering the image accurately is valuable to radiologists in assessing medical images by identifying similar images in large archives that could assist with decision support[20].The accuracy of clustering is very important for classifying an image.

In K-Means partitioning based clustering algorithm required to define the number of final clusters (k) beforehand. Such algorithms are also having problems like susceptibility to local optima, sensitivity to outliers, memory space and an unknown number of iteration steps that are required to cluster [13& 16].

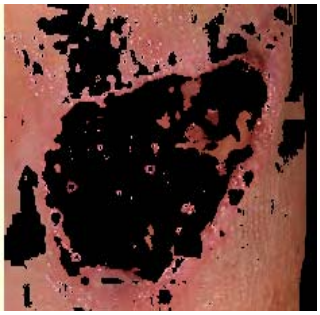

Fig.14clustering using K-means

In FCM is an extension of k-means Hierarchical Clustering. In FCM each data point can only be assigned in one cluster Fuzzy c-means allows data points to be assigned into more than one cluster, each data point has a degree of membership (or probability) of belonging to each cluster

Fig.15 Clustering with C-means

6. CONCLUSION:

In this proposed content based image retrieval system (CBIR), for Pre-processing uses De-noising, for Feature Extraction uses Shape and Texture Extraction,for Classification and Clustering uses FCM and for similarity measure uses Euclidean distance. When Implemented the proposed system can easily retrieve a similar image from a medical database using FCM techniques. The purpose of thiswork is to aid the doctor or a radiologist to easily judge from an image if it is a malignant ora benign tumor and would

ISSN 2348 – 7968

like to compare with previous cases to decide if this patient requires a dangerous operation. Content based image retrieval could be valuable to radiologists is assessing medical images by identifying similar images in a large database that could assist with decision support. In Fuzzy C-means Clustering (FCM), the data point is assigned membership to each cluster center as a result of which a data point can belong to more than one cluster center. FCM is useful to mine complex data set. Gray levels of an image are separated by using a Fuzzy C-means clustering. A Comparison between K-Means and FCM is also done.

Relevance feedback (RF) can be used to make a bridge between machine learning and human perception. In RF, the user assesses the previously retrieved images as similar or dissimilar to the initial query and provides this assessment as feedback to the system.

7. FUTURE ENHANCEMENT:

In this work, Image can be retrieved from a huge image database using a Fuzzy c-means clustering in medical application. In otherapplications, like geographical area, remote sensing etc. the modification of the penalty terms of fuzzy c-means clustering can help in the retrieval of the image very efficiently. As image collections grow in size the system may take a lot of time, and eventually reduce the query-retrieval process. To increase the speed and the user's interaction with image retrieval systems, the images to be accessed from the internet sources and the CBIR system can be implemented over the World Wide Web and applying proposed fuzzy C-means algorithm in a more efficient manner.

7. REFERENCES:

[1]. Sameer antani et al "Content- based image retrieval for large biomedical archives", 2004.

[2].Ceyhunburakakgül et al "Content-based image retrieval in radiology: current status and future directions", 2013.

[3]. S.meenachisundaresan et al "Design of image retrieval efficacy system based on cbir", 2013.

[4]. Pravin d. Pardhi et al " Introduction to query techniques for large cbir systems", 2012.

[5]. Dr.h.b.kekre et al "Image retrieval with shape features extracted using gradient operators and slope magnitude technique with btc", 2010.

[6]. Reshmachaudhari et al "Content based image retrieval using color and shape features", 2012.

[7]. Neetusharma et al"efficientcbir using color histogram processing", 2011.

[8]. Ms. K. Arthi et al "Content based image retrieval algorithm using colour models"- 2012.

[9]. S. Mangijaosingh "Content-based image retrieval using color moment and gabor texture feature", 2012.

[10]. Mandeepsingh et al "Texture analysis of ct scan images", 2008.

[11]. A. Sheppard and liwen et al "Integrated platform for image textural analysis and clustering of transrecal prostate ultrasound images", 2007.

[12]. Sujata t bhairnallykar et al [12] in their work have proposed a " Content based medical image retrieval with svm classification and relevance feedback"-2013.

[13]. Soumighosh et al " Comparative analysis of k-means and fuzzy c-means algorithms", 2013.

[14]. Abhishekarvindgulhane et al "Noise reduction of an image by using function approximation techniques", 2012.

[15]. R. Krishna priya et al "Fuzzy c-means method for colour image segmentation with l*u*v* colour transformation", 2011.

[16]. B. Rama murthy et al "Shape-based image retrieval using canny edge detection and k-means clustering algorithms for medical images", 2013.

[17]. Tarasaikumar et al "Colour based image segmentation using fuzzy c-means clustering", 2013.

[18]. Yong yang et al "Image segmentation by fuzzy c-means clustering algorithm with a novel penalty term", 2012.

[19].beijing et al "Improving relevance feedback in image retrieval by incorporating unlabelled images", 2012.

[20]. Anil z chitade et al "Colour based image segmentation using k-means clustering", 2010.