ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

Adaptive Query Image Searching Method with Low Level Feature Extraction and K-means Clustering

Devadasu Jahnavi
M.Tech Student,
Department of ECE,
Bhimavaram Institute of Engineering & Technology,
Pennada, Andrapradesh, India.

Sujatha Chavakula
HOD,
Department of ECE,
Bhimavaram Institute of Engineering & Technology,
Pennada, Andrapradesh, India.

Abstract:

In this thesis we had implemented a novel low level feature extractor and k-means clustering with bit wise Euclidean distance for similarity measurement between the feature vectors of query and database, which will be more secured and real time based application. By using the biorthogonal spline wavelet filtering method the computational time is high and the accuracy will be much lower. Hence, it can be extended to improve the efficiency and accuracy with inclusion of segmentation with k-means clustering (SKMC), which will represent the cluster pattern of a query and data base images. Therefore, to incorporate both cluster patterns and low level features of image, we proposed a novel scheme in which both of them will be represented in a single composite feature known as bi-orthogonal spline discrete wavelet with k-means clustering (BSDW-KMC). This addition work had improved the system accuracy as well as the precision time image retrieval system.

Key words:

CBIR, Bi-orthogonal spline wavelets, DWT, segmentation, K-means clustering and precision.

I.INTRODUCTION:

Recent years there is a rapid growth in searching engines such as Bing image search: Microsoft's CBIR engine (Public Company), Google's CBIR system, note: does not work on all images (Public Company), CBIR search engine, by Gazopa (Private Company), Imense Image Search Portal (Private Company) and Like.com (Private Company), image retrieval has become a challenging task.The interest in CBIR has grown because of the retrieval issues, limitations and time consumption in metadata based systems.

We can search the textual information very easily by the existing technology, but this searching methods requires humans to describe each images manually in the database, which is not possible practically for very huge databases or for the images which will be generated automatically, e.g. images generated from surveillance cameras. It has more drawbacks that there is a chance to miss images that use different equivalent word in the description of images. The systems based on categorizing images in semantic classes like "tiger" as a subclass of "animal" can debar the miscatergorization problem, but it will requires more effort by a use to identify the images that might be "tigers" , but all of them are categorized only as an "animal". Content-based image retrieval (CBIR) is a application of methods of acquisition, pre-processing, analyzing, representation and also understanding images to the image retrieval problem, that is the problem of exploring for digital images from large databases. The CBIR system is opposed to traditional approaches, which is known on concept based approaches i.e., concept based image indexing (CBII) [1].

II.RELATED WORK:

In the past decades several CBIR systems have been proposed, and still the researchers are focusing on developing extended CBIR systems with more effective results. The letter proposed in [4] gives a comparison of different approaches of CBIR based on similarity measures and image features to identify the similarity between the images, which provides accurate information for retrieving the relevant images from large database. Wan Siti et.al proposed in [5] compares the several medical image retrieval systems based on the feature extraction and to improve the effectiveness of the CBIR system for medical images such as magnetic resonance (MR) images and computed tomography (CT) images [10]. The major concept proposed in [5] is to help in the diagnosis such as to find the similar disease and monitoring of patient

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

progress continuously. B. S. Manjunath et.al presented in [6] is the combination of color, texture with inclusion of edge compactness for Motion Picture Expert Group (MPEG)-7 standards. Another approach proposed in [7] used different color spaces such as HSV and YCbCr explains a similar approach based on color and texture analysis. The work proposed in [8] introduces a new retrieval system which has done by using wavelet transformation with both color and texture features together and will perform better than existed state of art algorithms. Recently, retinal image retrieval system called CBIR for retinal and blood vessels extraction [9] has been analyzed by the histogram features of RGB color components. The multi resolution analysis has applied to the image to acquire the texture information. In addition to improve the performance, morphological operations are applied to study the shape of object. Swati Agarwal has proposed a new CBIR system in [11], which is by using discrete wavelet transform and edge histogram descriptor (EHD). Here the retrieval is based on color and texture features not by using color information in the image, input image first decomposes the input query image into several sub bands i.e., approximation coefficients and detail coefficients, where detail coefficients consists of horizontal (LH), vertical (HL) and also the diagonal information (HH) of the image. Afterwards, EHD is used to gather the information of dominant edge orientations. This mixture of 3D-DWT and EHD will improve the efficiency of the CBIR system.

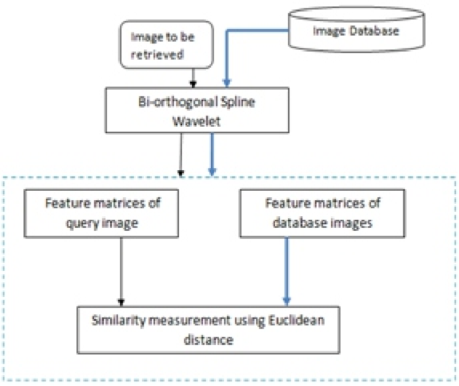


Fig1. Existing Bi-orthogonal Spline wavelet (BSW) filtering scheme for CBIR system.

Bi-orthogonal Spline Wavelet filtering :

When we use similar FIR filters for decomposition and reconstruction then symmetry and exact reconstruction will be incompatible. Hence, two wavelets will be used in biorthogonal filters instead of just one: $\tilde{\psi}$ is used as a one wavelet to analyze the signal coefficients using

$$\tilde{c}_{j,k} = \int I(x)\tilde{\psi}_{j,k}(x)dx \qquad (1)$$

Other wavelet used in the synthesis is Ψ.

$$I = \sum_{j,k} \tilde{c}_{j,k}\psi_{j,k} \qquad (2)$$

In this paper, a novel method for CBIR system "An adaptive query image searching using low level feature extraction and k-means clustering" has been proposed to improve the system efficiency and precision time. Here we have four main steps: First, segmentation will be done with the k-means clustering for both query and database images as well. Then, BSDW will be applied to get the low level features and patterns of images. Now, make a feature vectors using the low level features, patterns and clusters. Finally, calculate the distance between feature vector of the query image and data base images, if the distance is small enough then the corresponding image will be matched with the query image.

III.PROPOSED CBIR SYSTEM:

For image retrieval, classification and indexing both color and texture have been used widely in various applications. Histogram of a image is a graphical analysis of a image, which represents the color information of image. It is a first order statistical measure. The major drawback of this histogram based approaches is that the spatial distribution and local variations will be ignored. Local spatial variation of pixel intensity is commonly used to capture texture information in an image. Wavelet transform is a one of the best approach to find the low level features of image which is an extension for both fourier (FT) and short time fourier transforms (STFT). It uses an adaptive window to make sure that the every part of the signal or image information will be extracted by using the scaling property. Here, we used a discrete wavelet (DW) transform with bi-orthogonal spline filters to get the pattern classification of medical x-ray images.

3.1.Discrete Wavelet Transform (DWT):

Discrete Wavelet Transform (DWT) is a modified version of Continuous Wavelet Transform (CWT). DWT principles are very similar to the CWT however the wavelet scales and positions are based upon powers of two.

$$W(\tau, s) = \frac{1}{\sqrt{s}} \int_{-\infty}^{\infty} x(t)\psi\left(\frac{t-\tau}{s}\right)dt \qquad (3)$$

$$\int_{-\infty}^{\infty} \psi(t)dt = 0$$

$$\int_{-\infty}^{\infty} \left|(\psi(t))^2\right| dt < \infty$$

The basic principle of DWT is to pass the input signal through a group of filters i.e., low pass and high pass filters to get the low frequency (LF) and high frequency (HF) of source signal. Low frequency contents include LL and these coefficients are known as the approximation coefficients [18]. This means the approximations are obtained by using the high scale wavelets which corresponds to the low frequency. The high frequency components which are known as LH, HL and HH of the signal are called the details which will be obtained by using the low scale wavelets which corresponds to the high frequency. The process of DWT filtering includes, first the signal is fed into the wavelet filters. These wavelet filters comprises of both the high-pass and low-pass filter. Then, these filters will separate the high frequency content and low frequency content of the signal. However, with DWT the numbers of samples are reduced according to dyadic scale. This process is called the sub-sampling. Sub-sampling means reducing the samples by a given factor. Due to the disadvantages imposed by CWT which requires high processing power [11] the DWT is chosen due its simplicity and ease of operation in handling complex signals.

3.2.K-means Clustering :

•First we will select the number of centroids randomly i.e., depends on number of clusters
•Now, partition the objects within each cluster
•It finds partitions such that pixels within each cluster are as close to each other as possible, and as far from objects in other clusters as possible.
•The objects are in the cluster or not will be calculated by measuring the distance between the cluster pixels. When the calculated Euclidean distance has minima value then the pixels will be grouped with the respective cluster.
•Do the above process for remaining clusters also. Then, we will get three clusters with their similar pixels.
•Now, calculate the mean of each cluster and replace the mean values with the centroids
•Repeat the same process with these new centroids by giving the number of iterations until unless the convergence occurrence i.e., the mean value of clusters = cluster centroid value.

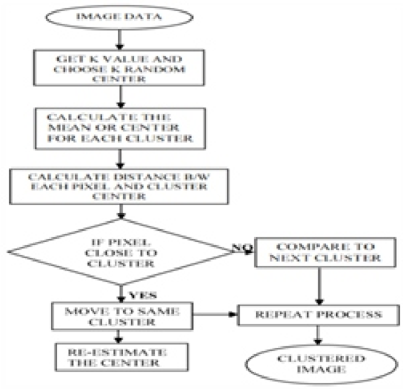


Fig2. Flow chart of K-means clustering

The steps involved in proposed scheme are as follows:
Step1: Select and read the query image from the database.
Step2: Apply segmentation with k-means clustering to the query image.
Step3: Now, select the bi-orthogonal spline wavelet filters and apply DWT to extract the low level features
Step4: Make feature vector by using the above three steps
Step5: Read all the database images and apply all the three steps which have been applied to the query image to find the feature vectors
Step7: Combine all the feature vectors and make it into a single feature vector for query image features and database images as well
Step8: Calculate the Euclidean distance for the similarity check, when the similarity between the query and database images will be more than 85%

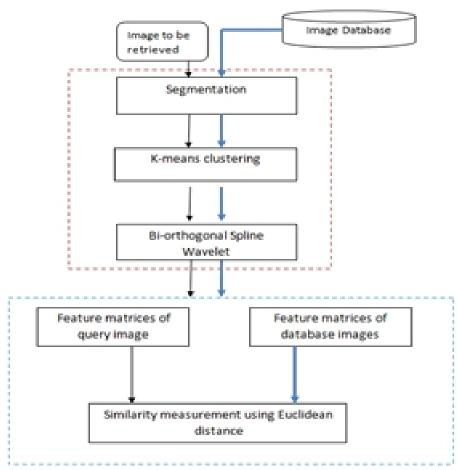


Fig.3. Proposed scheme of CBIR system

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

IV.SIMULATION RESULTS:

In this section we discussed the simulation results of CBIR system for bio-medical images based on the proposed and conventional schemes. The proposed algorithm has been tested with few databases and displayed the outputs in the below figures. Fig4 shows that the initial GUI model for retrieving images using existing and proposed CBIR schemes, selecting an image by browsing computer is shown in fig5. The output of existing BSW scheme has shown in fig6, which has only 2 relevant images from the four windows i.e., less accuracy in efficiency and precision. Proposed retrieval system has been shown in fig7, in which we got more precision time and accuracy. This CBIR system has been tested with few more medical images, the remaining figures shows that the retrieval of the given input relevant images. As a measure of performance we have used two widely used metrics of Precision and Recall [13]. Precision is a measure of ability of CBIR algorithm to retrieve only relevant images, while Recall decides the ability of CBIR algorithm to retrieve all relevant images as defined by (4) and (5) respectively.

$$P = \frac{Total\ numebr\ of\ relevant\ images\ retrieved}{Total\ numebr\ of\ retrieved\ images} \quad (4)$$

$$R = \frac{numebr\ of\ relevant\ images\ retrieved}{numebr\ of\ relevant\ images\ in\ database} \quad (5)$$

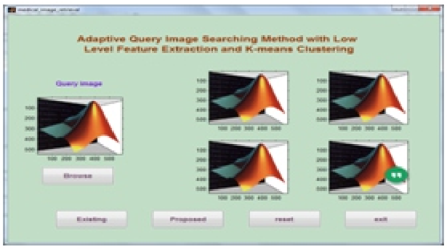


Fig.4 GUI model of CBIR system using existing and proposed algorithm

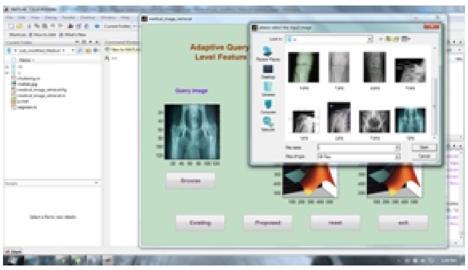


Fig.5 selecting an image from the query folder

Fig.6 Retrieved images using Existing CBIR scheme

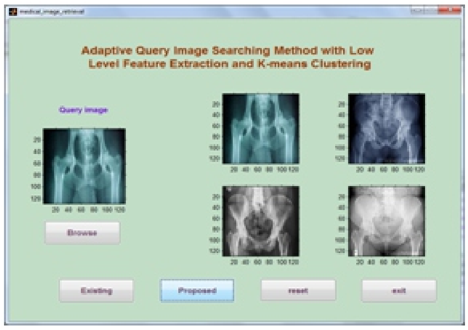
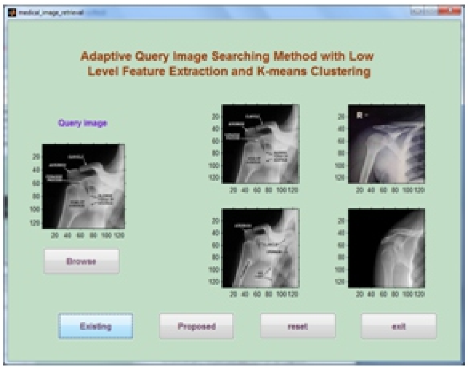
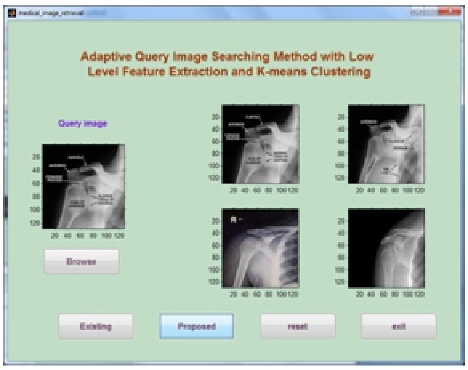


Fig7. Retrieved images using proposed CBIR system



(a)



(b)

Fig.8 retrieved images (a) Existing scheme and (b) Proposed scheme with another image

ISSN No: 2348-4845
International Journal & Magazine of Engineering, Technology, Management and Research
A Peer Reviewed Open Access International Journal

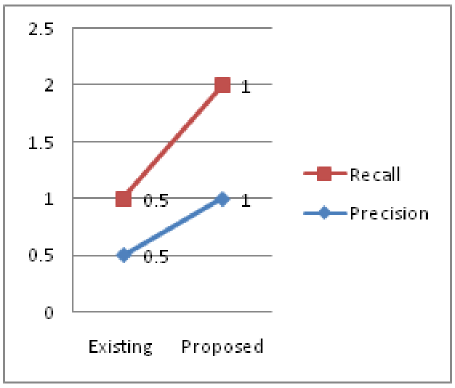


Fig.9 performance evaluation with precision and recall

Fig. 10 Retrieved images using existing and proposed schemes.

V.CONCLUSION:

In this letter we had proposed an adaptive CBIR scheme for large database systems. Feature vectors have been generated by considering multiple features of images with wavelet filtering and k-means clustering. Euclidean distance is used to measure the similarity between the query and database feature vectors to retrieve the relevant images. By using low level feature extraction and SKMC, the performance of the proposed CBIR system had improved in terms of the accuracy and computational complexity while improving the system efficiency. The proposed system has proven that this approach has got superior performance than the existing CBIR schemes.

REFERENCES:

1.Chua, T.-S., Hung-ICeng Pung; Guo-Jun Lu; Hee-Sen Jong, “A concept based image retrieval system” Proceedings of the Twenty-Seventh Hawaii International Conference on System Sciences, Jan. 1994.

2.Eakins, John; Graham, Margaret, “Content-based Image Retrieval”, University of Northumbria, New castle

3.X Wang, T S Chua and Al-Hawamdeh, “Probabilistic and semantic based retrieval in hypertext”, Proc. of the South-East Asia Regional Computer Conf., pp. 25.01 -25.17 1992

4.Khan, S.M.H., Hussain, A. ; Alshaikhi, I.F.T., “Comparative Study on Content Based Image (CBIR),” International Conference in Advanced Computer Science Applications and Technologies (ACSAT), 2012.

5.Wan Siti, H. Munirah, W. Ahmad, M. Faizal and A. Fauzi, “Comparison of Different Feature Extraction Techniques in Content Based Image Retrieval For CT Brain Images,” 10Th IEEE workshop on Multimedia Signal Processing, pp. 503-508, 2008.

6.B. S. Manjunath, J. R. Ohm, V. V. Vasudevan, A. Yamada, “Colour and Texture Descriptors,” IEEE Transactions on Circuits and Systems for Video Technology, 1998

7.P. S. Hiremath, S. Shivashankar, J. Pujari, “Wavelet Based Features for Color Texture classification with Application to CBIR,” International Journal of Computer Science and Network Security, Vol. 6, No.9A, September 2006

8.Tian Yumin, Mei Lixia, “Image Retrieval Based on Multiple Features Using Wavelet,” 5th IEEE International Conference on Computational Intelligence and Multimedia Applications (ICCIMA’03), 2003.

9.M. R. Zare, R. N. Ainon, W. C. Seng, “Content-based Image Retrieval for Blood Cells,” Third Asia International Conference on Modeling & Simulation, 2009

10.Prasad, B.G., Krishna, A.N., “Statistical Texture Feature Based Retrieval and Performance Evaluation of CT Brain Images” 3rd International Conference on Electronics Computer Technology (ICECT), Vol. 2, April 2011

11.Swati Agarwal, A.K. Verma, Preethvanti Singh, “Content Based Image Retrieval using Discrete Wavelet Transform and Edge Histogram Descriptor” International conference on Information Systems and Computer Networks, 2013

12.D. Ashok Kumar, J. Esther, “Comparative Study on CBIR based by Color Histogram, Gabor and Wavelet Transform,” International Journal of Computer Applications (0975 – 8887) Volume 17– No.3, March 2011.