ISSN (Online) : 2319 - 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

Volume 3, Special Issue 3, March 2014

2014 International Conference on Innovations in Engineering and Technology (IC444IET'14)

On 21st & 22nd March Organized by

K.L.N. College of Engineering, Madurai, Tamil Nadu, India

Medical Image Retrieval Using Integer Wavelet Transform

#1A.Kumari Sankari Mala, #2Dr.S.Allwin,M.E.,Ph.D.,

#1PG Scholar, Infant Jesus College of Engineering, Thoothukudi, Tamilnadu, India

#2Associate Professor, Infant Jesus College of Engineering, Thoothukudi, Tamilnadu, India

ABSTRACT— Image Retrieval is one of the most applicable image processing technique , which has been used extensively. Feature extraction is one of the most important procedures used for interpretation and indexing images in content - based image retrieval systems. Effective storage , indexing and managing a large number of image collections is a critical challenge in computer systems. There are many proposed methods to overcome these problems. However , the rate of accurate image retrieval and speed of retrieval is still an interesting field of research. In this study , we propose a new method based on texture based image retrieval system. It is particularly valuable in Histology image analysis for interpreting varying tissue composition and architecture into histological concepts. The proposed method consists of three steps. Texture features are retrieved using integer Wavelet , The retrieval is performed using a combination of simple and weighted (class membership based) distance metric in complete search space unlike the conventional classifier based retrieval techniques. The proposed technique also provides flexibility in reducing the search space in steps increasing the speed of retrieval at the cost of gradual reduction in accuracy. Experiments were performed on two different collections of histology images . The performance of the method is evaluated using three texture data sets varying in orientations, complexity and number of classes. Average normalized rank and combination of precision and recall are considered as metrics to evaluate and compare the proposed method against different methods.

KEYWORDS— Image retrieval , fuzzy logic , integer wavelet , medical image retrieval, histology.

I.INTRODUCTION

Image Retrieval techniques are useful in many image – processing applications. Content – based image retrieval systems work with whole images and searching is based on the comparison of the query. General techniques for image retrieval are color, texture and shape. These techniques are applied to get an image from the image database. They are not concerned with the various resolutions of the images , size and spatial color distribution. Hence all these methods are not appropriate to the art image retrieval. Moreover shape based image retrievals are useful only in the limited domain. The content and metadata based systems gives images using an effective image retrieval technique. Many other image retrieval systems use global features like color, shape and texture. But the prior results say there are too many false positives while using those global features to for similar images. The use of images in human communication is hardly new and the use of maps and buildings plan to convey information almost certainly dates back to pre-Roman times. But the twentieth century has witnessed unparallel growth in the number, availability and importance of images in all walks of life. Images now play a crucial role in fields as diverse as medicine , journalism , advertising ,design , education and entertainment. Technology , in the form of inventions such as photography and television , has played a major role in facilitating the capture and communication of image data. But the real engine of the imaging revolution has been the computer, bringing with in a range of techniques for digital image capture , processing, storage and transmission which would surely have started even pioneers like John Logie Baird. Once computerized imaging became affordable , it soon penetrated into areas traditionally depending heavily on images for communication, such as engineering, architechture and medicine . Photograph libraries , art galleries and museums', too , began to see the advantages of making their collections available in electronic form. The creation of the World – Wide Web enabling the users to access data in a variety of media from anywhere on the planet , has provided a further massive stimulus to the exploitation of digital images . The number of digital images available on the Web was recently estimated to be between 10 and 30 million in which some observers consider to be significant underestimate. The need for efficient storage and retrieval of images recognized by managers of large image collections such as picture libraries and design achieves for many years was reinforced by a workshop sponsored by the USA's National Science Foundation in 1992. After examining the issues involved in managing visual information in some depth, the participants concluded that images were indeed likely to play an increasingly important role in electronically – mediated communication. One of the main problems that was highlighted was the difficulty of locating a desired image in a large and varied collection.

II.RELATED WORK

The content – based histology image retrieval systems have shown great potential in supporting decision making in clinical activities , teaching, and biological research. In this, the feature combination plays a key role. It aims at enhancing the descriptive power of visual features corresponding to semantically meaningful queries. It is particularly valuable in histology image analysis where intelligent mechanisms are needed for interpreting varying tissue composition and architecture into histological concepts. The core of this approach is a multiobjective learning method which aims to understand an optimal visual semantic matching function by jointly considering the different preferences of the group of query image [1], [3] . The task is posed as an optimization problem and a multi objective optimization strategy is employed in order to handle potential contradictions in the query images associated with the same keyword. It improves the system of content based image retrieval by using appropriately defined multi feature fusion model, which takes careful consideration of the structure and distribution of visual features.

The reported works on Histology image retrieval has been often limited due to the unreliable outputs from the feature metrics [2]. This is because the systems often rely on single low level features , which are not always capable of interpreting visual oblects into varying and complicated semantic meanings. To tackle this problem , the combination of low level features for semantic image retrieval has been widely considered.

Storage , search and organization of digital images are highly in demand. Designing a search image mechanisms based on user requirements to find images related to user demand has become an important research topic in this field. Image retrieval system is able to find an appropriate image according to query image and human perception . There are many proposed methods and approaches for classification, indexing , search and retrieval of visual information based on the analysis of low level image features such as color, texture, shape and others. The combination of these features showed more efficient performance in image retrieval systems [7].

Although the methods of color feature extraction require low complexity for retrieval , they do not consider distribution of color location. Therefore the image retrieval based on texture is generally considered. Different methods have been proposed for description of image texture. They have divided the texture analysis into four groups such as statistical, geometrical, model based and signal processing. And signal processing has been given the better performance in comparison with other methods.

Color histogram has been extensively used as global color descriptors [6]. It is used to solve translation and rotation invariant problems. Besides , it is characterized by being easily implemented and accurate; particularly with small database size. Accordingly, many enhancements in histogram based approaches have been achieved . However such approaches have several drawbacks ; the basic one is its dependence on a static quantization method. That is , why it is used to reduce color space to make storage and time more reasonable. Static quantization methods suffer from low discrimination power. This is because many similar colors may be set to different bins; a matter many similarity measure between the two histogram inefficient.

To solve the static quantization problem in the color histogram ,a quadratic similarity was proposed. The proposed method is set to compute the similarity between the two images.

The quadratic distance (D_Q) between the two images can be computed as :

$$D_Q\ (X,Y) = (X-Y)^T A\ (X-Y). \qquad (1)$$

III.PROPOSED SYSTEM

In this section we develop the Texture Feature Extraction and Retrieval method .

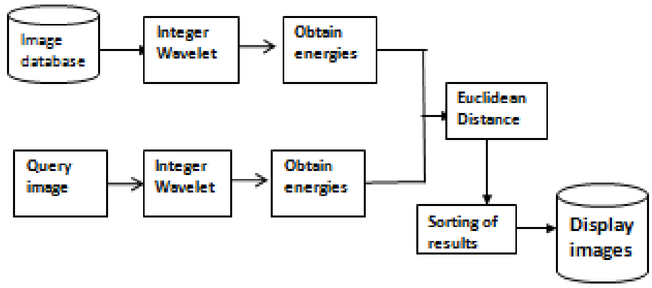

Fig 3.1 Proposed Framework

3.1 Texture Based Image Retrieval

The ability to retrieve images on the basis of texture similarity may not seem very useful. But the ability to match on texture similarity can often be useful in distinguishing between areas of images with similar colors (such as sky and sea, or leaves and grass). A variety of techniques have been used for measuring texture similarity; the best established rely on comparing values of what are known as second order statistics calculated from query and stored images [9].

Essentially, these calculate the relative brightness of selected pairs of pixels from each image. From these it is possible to calculate measures of image texture such as the degree of contrast , coarseness, directionality and regularity, or periodicity, directionality and randomness. Alternative methods of texture analysis for retrieval include the use of Gabor filters and fractals. Texture queries can be formulated in a similar manner to color queries, by selecting examp\les of desired textures from a palette, or by supplying an example query image. The system then retrieves images with texture measures most similar in value to the query. A recent extension of the technique retrieves textured regions in images.

Texture is another important property of images. Various texture representations have been investigated in pattern recognition and computer vision. Basically, texture representation methods can be classified into two categories :

1. Structural
2. Statistical

Structural methods, including morphological operator and adjacency graph, describe texture by identifying structural primitives and their placement rules. They tend to be most effective when applied to textures that are very regular. Statistical methods, including Fourier power spectra , co-occurrence matrices, shift in-variant principal component analysis (SPCA) , Tamura feature, Wold Decomposition, Markov random field, fractal model and multi resolution filtering techniques such as Gabor and wavelet transform, characterize texture by the statistical distribution of the image intensity [2], [5].

3.*2 Integer Wavelet Transform*

The basic idea of the wavelet transform is to represent any arbitrary function as a superposition of a set of such wavelets or basis functions. These basis functions or baby wavelets are obtained from a single prototype wavelet called the mother wavelet, by dilations or contractions (scaling) and translations (shifts). They have advantages over traditional fourier methods in analyzing physical situations where the signal contains such as image compression, turbulence, human vision, radar and earthquake predictions are developed in recent years. In wavelet transform the basis functions are wavelets. Wavelets tend to be irregular and symmetric.

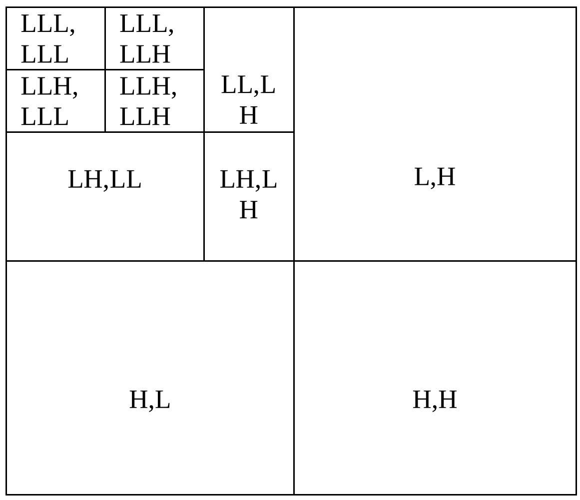

Fig 3.1 : Pyramid Wavelet Transform (Level 3)

A more efficient approach to wavelet transform is the use of integer transforms, such as Integer Wavelet Transform. The transform coefficients exhibit the feature of being exactly represented by finite precision numbers and this allows for truly loseless coding. Integer wavelet Transform is much faster than the floating point arithmetic in almost general purpose computers because the floating point wavelet transform demands for longer length than the integer wavelet transform does. Another benefit of using integer wavelet is the reversibility. That is, the image can be reconstructed
losslessly because all the coefficients are integers and can be stored without rounding off errors.

3.3 Algorithm for Texture Feature Extraction and Retrieval

The texture feature extraction involves the following steps when the query image is given as the input.

1) After reading the intensity or RGB plane, we apply DWT on each RGB with size N*N to generate approximation (low – low) , horizontal (low – high) , vertical (high – low) and diagonal (high – high) components. We use approximation components for the next step because wavelet transform analyses the signal at various frequency bands and gives higher frequency resolution and lower time resolution at lower frequencies .For a given image with size N*N , the two dimensional wavelet transform includes $\log_2^N$ stages. The first stage provides four sets of coefficients known as ving columnsy convolapproximation coefficients (CA1) , horizontal coefficients (CH1) , vertical coefficients (CV1), diagonal coefficients (CD1). These sets computed by convolving columns and rows of images with the low pass filter for approximation , and with the high pass filter , which are followed by down sampling. The length of these filters is 2n samples . Therefore if the length of an image is N, then the length of the output signal of low pass and high pass filters will be N+2n-1.
2) Then the modified approximation components has been constructed.
3) Construction of modified plane from step 2 by applying inverse wavelet transform with the modified approximation components , Zeroing horizontal, vertical, and diagonal components. A new image is constructed by using inverse wavelet whereas some information disappears because of removing horizontal, vertical and diagonal components.

4) To take alternative rows and columns by down sampling the output of step 3 with size of N/2 * N/2.

We generate feature vectors of all dataset image and store approximation components as feature vectors for each image. Euclidean distance is a simple and fast metric used to compute the match or similarity value to obtain the relevant images. If FVD and FVQ are the two dimensional feature vector of dataset image and query image, then the Euclidean distance if these feature vectors is obtained by

$$ED = \sqrt{\sum_{i=1}^{n}\left(\left(FVD_i - FVQ_{i)}\right)^2\right)} \qquad (2)$$

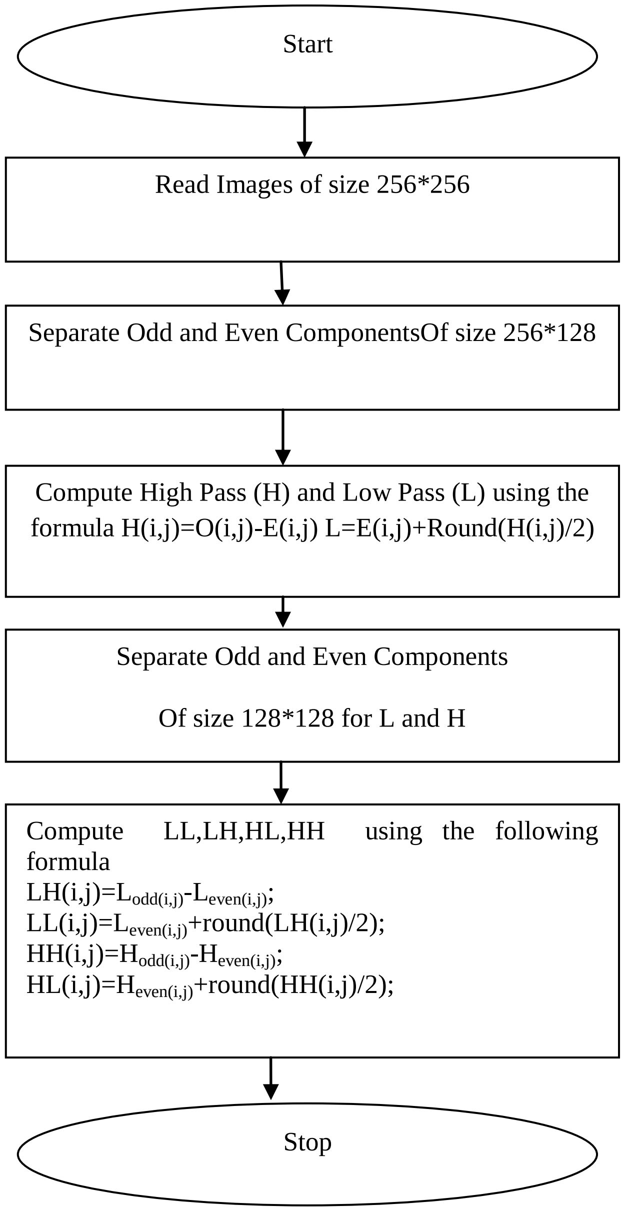

Fig 3.2 : Computation Of Integer Wavelet Transform

IV.EXPERIMENTAL RESULT

Evaluation of the proposed framework is done on a medical image dataset such as MRI, Mammograms, CT, ECG, X-RAY and Histology images. The similarity measure by a given query image involves searching the database for similar coefficients. Euclidean and quadratic distance is suitable and effective method which is widely used in image retrieval area. The retrieval results are a list of medical images ranked by their similarities measure with the query image.

The images in the database are ranked according to their distance *d* to the query image in ascending orders, and then the ranked images are retrieved. The computed distance is ranked according to closest similar; in addition, if the distance is less than a certain threshold set, the corresponding original images is close or match the query image. Precision *P* is defined as the ratio of the number of retrieved relevant images *r* to the total number of retrieved images *n*, i.e., $P = r/n$ Precision measures the accuracy of the retrieval.

$$\text{Precision} = \frac{\text{No.of relevant images retrieved}}{\text{Total no.of images retrieved}} = \frac{r}{n} \quad (3)$$

Recall is defined by *R* and is defined as the ratio of the number of retrieved relevant images *r* to the total number *m* of relevant images in the whole database, i.e.,$R=r/m$ [1]. Recall measures the robustness of the retrieval.

$$\text{Recalll} = \frac{\text{No.of relevant images retrieved}}{\text{Total no.of relevant images in db}} = \frac{r}{m} \quad (4)$$

The Average Recall Rate (AVRR) is given by equation

$$\text{AVRR} = \frac{1}{Q}\left\{\sum_{j=1}^{Q} \frac{\sum_{i=1}^{10} Ranki}{Nr}\right. \quad (5)$$

Where the rank of any of the retrieved images is defined to be its position in the list of retrieved image is one of the relevant images in the database. The rank is defined to be zero otherwise. Nr is the number of relevant images in the database, and Q is the number of queries performed. Therefore AVRR is defined in equation (12).In our case, the number of images retrieved was 10, and Nr was less than 10.

$$\text{AVRR} = (N_r + 1)/2 \quad (6)$$

Table 4.1 : Retrieval time and AVRR for Integer wavelet based Image retrieval systems

Approach	Total Retrieval Time for 100 images (in Seconds)	Average Recall Rate (%)
DCT	54	72
DWT	45	79
Integer Wavelet	32.91	84.17

Table 4.2-Recall and Precision for Integer wavelet based Image retrieval systems

Method	Recall in %	Precision in %
DCT	72	74
DWT	79	76
Integer Wavelet	82	82

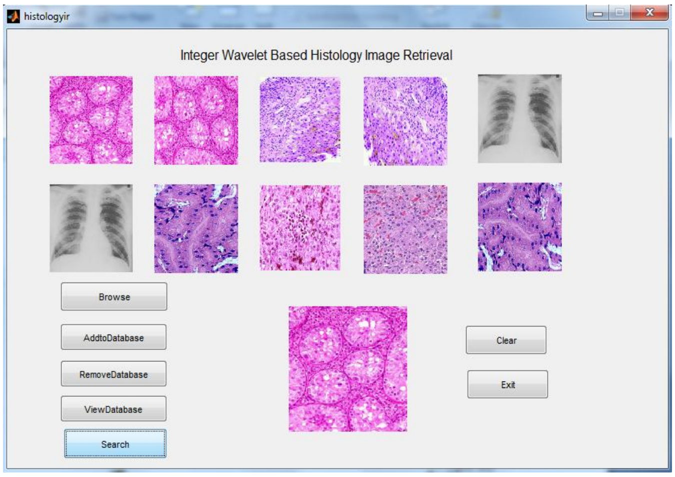

Fig 4.1 Integer wavelet based Histology Image retrieval

V.CONCLUSION

In this paper ,we analyze the texture features of image content descriptor for retrieval of Histology image database. This proposed method is compared with other retrieval algorithms in terms of recall rate and precision .The tabular column show that our method has better retrieval accuracy in terms of recall rate and precision The timing results for this approach is less and accurate ,the same algorithm could be used for retrieving other imaging modalities.

REFERENCES

[1] "*Histology Image Retrieval in Optimized Multifeature Spaces* " by Qianni Zhang and Ebroul Izquierdo , Senior Member of IEEE , published in IEEE journal of Biomedical and Health Informatics, Vol 17, No. 1, January 2013.

[2] "*Color and Texture Feature based Image Retrieval by using Hadamard matrix in discrete wavelet transform* " by Hasan Farsi , Sajad Mohamadzadeh , Published in IET Image Processing , January 2013.

[3] "*Content based texture Image Retrieval using fuzzy class membership* " by Sudipta Mukhopadhyay, Pattern Recognition Letters of Elsevier journal, January 2013.

[4] "*A performance Evaluation of Gradient field HOG descriptor for sketch based Image Retrieval*" by Rui Hu, John Collomosse , Computer Vision and Understanding of Elsevier journal , January 2013.

[5] "*Fuzzy shape Clustering for Image Retrieval* " by G. Castellano , A.M .Fanelli, F. Paparella, M.A. Torsello, Published in Pattern Recognition Letters of Elsevier journal , January 2012.

[6] "*Content based Image Retrieval using Dual Tree complex Wavelet Transform*" by Christina George Bab, D. Abraham Chandl. 2012 International Conference on Signal Processing , Image processing and Pattern Recognition.

[7] "*Content based Image Retrieval using Discrete wavelet Transform and Edge Histogram Descriptor* " by Swati Agarwal, A.K.Verma, Preetvanti Singh, 2011 International Conference on Information systems and Computer Networks.

[8] "A New Feature Set for Content based Image Retrieval" by M. Babu Rao, Ch. Kavitha , B. Prabhakara Rao , A.Govardhan, 2012 International conference on Signal processing, Image processing and Pattern Recognition.

[9] "*Combined Texture and Shape features for Content based Image Retrieval* " by M.Mary Helta Daisy, Dr. S. Tamilselvi, Js. GinuMol, 2012 International Conference on Circuits, Power and Computing Technologies.

[10] "*Fusion of Color, Shape and Texture Features for Content Based Image Retrieva*l" by Pratheep Anantharatnasamy, Vahissan Nandakumar and Sampath Deegalla, The 8th International Conference on Computer Science & Education (ICCSE 2011) April 26 – 28 , 2012, Colombo, SriLanka.