INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, 10(3):382-389, June, 2015.

CBMIR: Content Based Medical Image Retrieval Using Multilevel Hybrid Approach

B. Ramamurthy, K.R. Chandran

B. Ramamurthy*
Sri Ramakrishna Engineering College, Coimbatore, India
*Corresponding author: hod-ct@srec.ac.in

K.R. Chandran
PSG College of Technology, Coimbatore, India.

Abstract: The problem of retrieval and managing of medical images has become more important due to its scalability and rich information contained in it. In day to day activities, as more as medical images were converted into digital form. Due to its nature, the need for efficient and simple access to this data is essential one. This paper proposed a novel approach namely "content based medical image retrieval using multilevel hybrid approach" to manage and retrieval of this data.This work has been implemented as four levels. In each level, the retrieval performance of the work has been improved. The result of this work has been compared with some of the existing works that are made as literature review on this work.
Keywords: CBMIR, CBIR, euclidean method, precision, recall.

1 Introduction

The rapid expansion of digital image data has led to the need of rich descriptions and efficient retrieval tool. Due to this, many problems were often arises while retrieval of such data in specific applications such as accessing speed, scalability, efficiency etc. To improve this, still more and more research initiative is needed in the field of digital image retrieval system. In medical domain, the amount of image data generation is more and more in day to day activities such as CT, X-ray, and MRI etc. So the scalability of the medical image database has been increased and thereby need of this image retrieval system is essential one. Content based medical image retrieval (CBMIR) is an emerging technique that plays a pivot role in this domain. In this work, shape, texture and intensity contents were used to implementing the system.

2 Related work

CBMIR: shape-based image Retrieval using canny edge Detection and k-means Clustering algorithms for Medical images: In this work, the authors have proposed a shape based image retrieval system for medical images. To implement this, preprocessing, image segmentation, feature extraction, and classification steps were carried out. To retrieval, Euclidian distance calculated between query image and database images. Since the system used shape feature alone, it gives about 50% retrieval performance only [1].

Content Based Medical Image Retrieval with Texture Content Using Gray Level Co-occurrence Matrix and K-Means Clustering Algorithms: This system has used a texture based image retrieval system for medical image retrieval. It has capable of retrieving images based on the texture feature of the image. To implement this, the Preprocessing, feature extraction, Classification and retrieval steps are carried out. The main feature of the system is

to utilization of Gray Level Co-occurrence Matrix (GLCM) and k-means clustering algorithm to improve the retrieval efficiency of the system. Since the system used texture feature alone, it gives about 50% retrieval performance only [2].

CBMIR: Content Based Medical Image Retrieval Using Shape and Texture Content: The system proposed a shape and texture based image retrieval system for medical images. It has capable of retrieving images based on the shape and texture feature extraction combined together. The main objective of the system is to improve the retrieval efficiency when compared to single feature method. Since this system has used both shape and texture features combined together,it gives about 50%-90% efficiency [3].

MIRAGE: An E-repository of Medical Images for Learning Biomedical Informatics: This paper proposed an E-repository for medical images that offers great facilities to learn about the biomedical informatics. The facilities of domain-based, atlas-based, and content-based retrieval (CBIR) techniques are implemented to search the images in this developed repository. The uniqueness of the system is CBIR system for 3D is developed and coupled with 3D visualization that has used for educational material and as well as tele-education in future [4].

Computer-aided diagnostics of screening mammography using content-based image retrieval: This paper proposed a computer-aided diagnostics tool for screening the mammography using content based image retrieval technique. In this work, Support Vector Machine (SVM) classification technique was used for doubtful tissue pattern extraction in an image. Based on that, the retrieval of the system implemented using cbir for detecting the mammography of the image [5].

Content-based binary image retrieval using the adaptive hierarchical density histogram: The paper proposed a scheme for binary image retrieval. It has utilized black and white binary represented values as image feature and it named as the adaptive hierarchical density histogram that develops the allocation of the image points on a two-dimensional area. This technique uses the assessment of point density histograms of image regions that are computed by a pyramidal grid that is repeatedly simplified through the calculation of image geometric centroids. This extracted descriptor includes both global and local possessions that can be used in different types of binary image databases for the retrieval of images [6].

Improving the ranking quality of medical image retrieval using a genetic feature selection method: In this paper, the authors proposed a method for improving the ranking quality for medical image retrieval using genetic feature selection method. Here the authors have used single-valued genetic functions for evaluating the rankings to extend a group of feature selection methods based on the genetic algorithm approach to improve the precision of content-based image retrieval systems [7].

3 Proposed Model

The general process method of Content based medical image retrieval system is described as follows. Initially medical images were taken as input to the System and preprocessing of the images carried out in order to improve the flexibility of the images for further processing of the system. Then the output images of this preprocessing step are taken for feature extraction, feature vector construction and feature database construction processes respectively. This Similar process has been applied for query image also (up to feature extraction and feature vector construction process only). While searching the image, the query image feature vector compared with the database image feature vector by using Euclidean distance calculation method. The minimal Euclidean distance value is considered as closest distance image from the database image that ranked as first. The proposed CBMIR framework is shown in Figure 1.

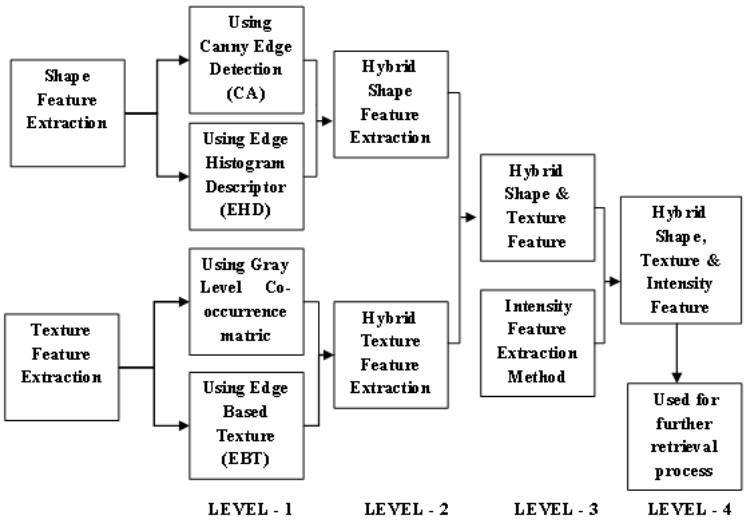

Figure 1: The proposed CBMIR framework

In this proposed approach, there are four levels of results have been derived based on the multilevel hybrid approach method. They are

Level 1. In this level, the content based medical image retrieval system has been implemented based on shape feature extraction method alone and texture feature extraction method alone (ie.single feature extraction method). For shape feature extraction, there are two methods have been implemented. First one is canny edge detection and another one is edge histogram descriptor. Similarly, for texture feature extraction, there are two methods have been implemented. First one is gray level co-occurrence matrix and another one is edge based texture feature extraction. For each method, separate results have been derived for further comparison with the next level in order to measure the performance of the system in each level.

Level 2. In this level, a hybrid feature extraction method implemented that includes a hybrid shape feature extraction method by combining Canny's Edge Detection method (CED) and Edge Histogram Descriptor (EHD) and hybrid texture feature extraction method by combining Gray Level Co-Occurrence Matrix (GLCM) and Edge based texture feature extraction(EBT) in order to extract the best pattern of the images. After obtaining this pattern, the image feature vector value of that pattern has been generated and stored in a database named 'feature database' for further searching and retrieval of the process. While retrieving the images, for each method (hybrid shape, hybrid texture), separate results have been derived for further comparison with the next level in order to measure the performance of the system.

Level 3. In this level, a hybrid shape and texture feature extraction method has been implemented by combining the output result of the hybrid shape feature extraction method and hybrid texture feature extraction method. After combining this, the fine tuned of the both shape and texture pattern of the image has been obtained and then the feature vector value of that pattern also has been generated and stored as a database named 'feature database' for further searching and retrieval of the process. While retrieving the images, the result also has been derived for further comparison with the next level in order to measure the performance of the system.

Level 4. In this level, hybrid intensity feature extraction method has been implemented by combining intensity features with hybrid shape and texture feature extraction method presented in level 3. After combining this, another level of fine tuned image pattern has been obtained

along with combined form of shape, texture and intensity features. The feature vector value for this obtained pattern has been generated and stored as a database named 'feature database' for further searching and retrieval of the process. While retrieving the images, the result also has been derived for further comparison with the next level in order to measure the performance of the system.

Similarly n number of image features can be used to develop n number of levels to measure retrieval accuracy of the system. In this system we have incorporated up to four levels only.

Similarity Comparison

In this approach (in all the levels), for retrieval of the images, similarity comparison technique has been used. For similarity comparison, we have used Euclidean distance, d using the following equation [8, 9].

$$d = \sqrt{\sum_{i=1}^{N}(F_Q[i] - F_{DB}[i])^2} \tag{1}$$

Where $F_Q[i]$ is the i^{th} query image feature and $F_{DB}[i]$ is the corresponding feature in the feature vector database. Here N refers to the number of images in the database.

4 Experimental setup and result analysis

A Intel ® Core 2 Duo CPU Workstation with 2GB RAM computer has been used for conducting the experiments. The MATLAB 7.2.0-Image Processing tool Box was used for developing User Interface components as front end, MATLAB 7.2.0-Image Processing tool Box-Workspace was used as feature database for storage as back end and for image processing work, other MATLAB 7.2.0 utilities were used.

4.1 Retrieval Efficiency

For retrieval efficiency, traditional measures namely precision and recall parameters were computed for several real time medical images [10–12]. Standard formulas that have been computed for determining the precision and recall measures. Precision and Recall parameters can be defined as follows.

Precision (P) is the ratio of the relevant images to the total number of images retrieved

$$P = \frac{r}{n1} \tag{2}$$

where,
r - number of relevant images retrieved
$n1$ - total number of images retrieved.

Recall(R) is the percentage of relevant images among all possible relevant images

$$R = \frac{r}{n2} \tag{3}$$

where,
r - number of relevant images retrieved
$n2$ - total number of relevant images in the database.

By randomly selecting some sample query images from the MATLAB-Image Processing tool Box-Workspace database, the system was tested and the results are shown in the following Tables for all the four levels presented in section 3.

Level 1: In level 1, there are four methods have been implemented as single feature extraction method. The result of all the methods in level 1 has been presented in the following tables. The shape feature extraction method using canny edge detection (CED) method result has been presented in table 1.

Table 1. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	50.00	58.33	68.26	56.95	53.33
Recall	37.50	36.84	51.42	66.66	34.00

The shape feature extraction method using edge histogram descriptor (EHD) method result has been presented in table 2.

Table 2. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	50.00	55.33	65.23	59.84	60.33
Recall	40.50	33.50	45.95	59.00	55.00

The texture feature extraction method using gray level co-occurrence matrix (GLCM) method result has been presented in table 3.

Table 3. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	55.71	65.71	53.33	57.00	56.00
Recall	45.00	48.94	51.42	56.66	60.00

The texture feature extraction method using edge based texture (EBT) method result has been presented in table 4.

Table 4. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	55.71	50.00	65.00	56.33	68.94
Recall	55.66	56.00	65.33	65.00	51.00

In this first level of multi-level hybrid approach, all the approaches such as CED, EHD, GLCM and EBT gives about 50% to 68.94% of the retrieval performance in terms of precision and 33.50% to 66.66% in term s of recall since all the works were in initial stage and also each method gives its own retrieval performance.

Level 2: In level 2, there are two methods have been implemented as hybrid feature extraction method. The result of these methods in level 2 has been presented in the following tables.

The hybrid shape feature extraction using CED and EHD method result has been presented in table 5.

Table 5. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	73.00	70.33	72.26	71.95	73.33
Recall	60.50	68.84	62.42	60.66	61.00

The hybrid texture feature extraction using GLCM and EBT method result has been presented in table 6.

Table 6. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	76.00	70.33	75.26	70.95	78.33
Recall	66.50	60.84	68.42	63.66	66.00

In this, second level of multi-level hybrid approach, the first approach such as the hybrid shape feature extraction using CED and EHD method gives about 70.33% to 73.33% of the retrieval performance in terms of precision and 60.50% to 68.84% in terms of recall. The hybrid texture feature extraction using GLCM and EBT method gives about 70.33% to 78.33% of the retrieval performance in terms of precision and 60.84% to 68.42% in terms of recall. Hence this second level gives better performance than the level 1 performance.

Level 3: In level 3, the hybrid shape & texture feature extraction method has been implemented using the combined form of hybrid shape feature extraction method (HBS) and hybrid texture feature extraction method (HBT).The result of this method in level 3 has been presented in the following table 7.

Table 7. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	94.73	94.44	93.93	92.59	87.50
Recall	90.00	89.47	88.57	83.33	84.00

In this, 3rd level, the performance of the retrieval drastically increased since both the shape and texture features combined together used for retrieval process. When compared to the level 2, this level gives its performance much better and also in terms of precision it gives about 87.50% to 94.73% and in terms recalls about 83.33% to 90.00% of the retrieval performance.

Level 4: In level 4, the hybrid shape, texture and intensity feature extraction method has been implemented by combining the hybrid shape and texture feature extraction method with intensity feature extraction method in order to extend the next level of the work. The result of this method has been presented in the following table 8.

Table 8. Precision and Recall values in %

Query Image	Image 1	Image 2	Image 3	Image 4	Image 5
Precision	95.75	95.00	94.93	93.58	94.65
Recall	91.66	90.10	90.57	92.00	87.00

In this, 4^{th} level, the performance of the retrieval highly increased since all the shape, texture, and intensity features combined together used for retrieval process. When compared to the level 3, this level outperforms well and also in terms of precision it gives about 93.58% to 95.75% and in terms of recall about 87.00% to 92.00% of the retrieval performance. Similarly n-number of image features can be integrated in order to improve the retrieval performance of the system since the image pattern has played a vital role in an image retrieval process.

4.2 Result Comparison and Discussion

The result comparison also made for the above described levels of the work. The following table 9 shows comparison results of the work.

Table 9. The Comparison of the result (Precision and Recall values in %)

Level 1								Level 2				Level 3		Level 4	
Shape (CED)		Shape (EHD)		Texture (GLCM)		Texture EBT		Hybrid Shape (CED+EHD)		Hybrid Texture (GLCM+EBT)		Hybrid Shape & Texture (HBS+HBT)		Hybrid Shape, Texture & Intensity (HBS+HBT+HBI)	
P	R	P	R	P	R	P	R	P	R	P	R	P	R	P	R
50.00	37.50	50.00	40.50	55.71	45.00	55.71	55.66	73.00	60.50	76.00	66.50	94.73	90.00	95.75	91.66
58.33	36.84	55.33	33.50	65.71	48.94	50.00	56.00	70.33	68.84	70.33	60.84	94.44	89.47	95.00	90.10
68.26	51.42	65.23	45.95	53.33	51.42	65.00	65.33	72.26	62.42	75.26	68.42	93.93	88.57	94.93	90.57
56.95	66.66	59.84	59.00	57.00	56.66	56.33	65.00	71.95	60.66	70.95	63.66	92.59	83.33	93.58	92.00
53.33	34.00	60.33	55.00	56.00	60.00	68.94	51.00	73.33	61.00	78.33	66.00	87.50	84.00	94.65	87.00

P - Precision, R - Recall

From the above table, the proposed framework has improved results in significantly for content based medical image retrieval system using a new approach called "multilevel hybrid approach". It has been shown that results are in level wise improvement. In level 1, experiments provide about 50% result performance. In level 2, it has been increased to in some extent that is about 25% higher than the levelb1. In level 3, the retrieval performance of the system drastically increased since both the shape and texture features combined together used in retrieval. In level 4, a new feature called 'intensity' has been used along with shape and texture method. It gives a better result than previous level 3. This work also provides an entire framework for inserting and deleting the features as user preferred.

5 Conclusion

On the whole, this proposed system gives its own performance and as well as flexibility for content based medical image retrieval system since the system supports n number of features for pattern recognition of the image. In each level, the retrieval performance of the system also significantly improved. The experimental results show that the medical images have been retrieved in an efficient manner using multilevel hybrid approach method. The system is designed in a flexible and consistent flow for easy understanding. Some of the goals that have been achieved by the developed system are accuracy of retrieved images by designing an efficient framework and incorporating it with the feature extraction techniques.

Bibliography

[1] B. Ramamurthy, K.R.Chandran (2011); CBMIR: Shape-based image Retrieval using canny edge Detection and k-means Clustering algorithms for Medical images, *International Journal of Engineering Science and Technology (IJEST)*, ISSN: 0975-5462, 3(2): 1864-1872.

[2] B. Ramamurthy, K.R.Chandran (2012); Content Based Medical Image Retrieval with Texture Content Using Gray Level Co-occurrence Matrix and K-Means Clustering Algorithms, *Journal of Computer Science*, ISSN 1549-3636, 8(7): 1070-1076.

[3] B. Ramamurthy, K.R.Chandran (2013); CBMIR: Content Based Medical Image Retrieval Using Shape and Texture Content, *Advances in Modeling B:* AMSE Press, France, 56(2): 84-95.

[4] Xiaohong Gao, Yu Qian (2012); MIRAGE: An E-repository of Medical Images for Learning Biomedical Informatics, eTELEMED 2012: *The Fourth International Conference on eHealth, Telemedicine, and Social Medicine, January 30, 2012 to February 4*, 2012, ISBN: 978-1-61208-179-3, Valencia, Spain, 209 - 214

[5] Thomas M. Deserno, Michael Soiron, Júlia E. E. de Oliveira, and Arnaldo de A. Araújo(2012); Computer-aided diagnostics of screening mammography using content-based image retrieval, *Proceedings of Society of Photo-Optical Instrumentation Engineers (SPIE)*, Vol 8315, Medical Imaging 2012: Computer-Aided Diagnosis,(February23,2012), San Diego, California, pp.831527-831527-9.

[6] Panagiotis Sidiropoulos, Stefanos Vrochidis, Ioannis Kompatsiaris (2011); Content-based binary image retrieval using the adaptive hierarchical density histogram, *Pattern Recognition*, 44(4): 739 - 750.

[7] Sérgio Francisco da Silva, Marcela Xavier Ribeiro,João do E.S. Batista Neto, Caetano Traina-Jr.,Agma J.M. Traina(2011); Improving the ranking quality of medical image retrieval using a genetic feature selection method, *Decision Support Systems*, 51(4): 810 - 820.

[8] Mohammed Eisa,Ibrahim Elhenawy, A. E. Elalfi and Hans Burkhardt(2006); Image Retrieval based on Invariant Features and Histogram Refinement, *International Journal on Graphics, Vision and Image Processing (ICGST)*, pp. 6: 7-11.

[9] M. Flickner, et al (1995); Query by Image and Video Content: The QBIC System, *IEEE Computer*, 28(9): 23-32.

[10] Xiaohui Yang, Lijun Cai (2014); Adaptive Region Matching for Region- based Image Retrieval by Constructing Region Importance Index, *IET Comput. Vis*, 8(2): 141-151.

[11] Bor-Chun Chen,Yan-Ying Chen et al (2013); Scalable Face Image Retrieval Using Attribute-Enhanced Sparse Codewords, *IEEE Transactions On Multimedia*, 15(5): 1163-1173.

[12] Subrahmanyam Murala et al (2014); Local Mesh Patterns Versus Local Binary Patterns: Biomedical Image Indexing and Retrieval, *IEEE Journal of Biomedical and Health Informatics*, 18(3):929-938.