ISSN: 2348-4098
Volume 3 Issue 2: 2015

CONTENT-BASED RETRIEVAL OF MEDICAL IMAGES

TIKESHWAR SAHU[1], SANDEEP PATIL[2]

[1]M.E. Scholar, Department of Electronics & Telecommunication Engineering, Shri Shankaracharya Technical Campus (SSTC), SSGI (FET), Junwani Bhilai C.G., Email: omtikeshwar@gmail.com

[2]Sr. Associate Professor, Department of Electronics & Telecommunication Engineering, Shri Shankaracharya Technical Campus (SSTC), SSGI (FET), Junwani Bhilai C.G.,, Email: patilsandeepb1212@gmail.com

ABSTRACT

Information retrieval has become emerging research areas in several applications such as medical imaging, underwater imaging and others. Content based image retrieval (CBIR) plays an important role in solving the difficulties in applications dealing with large data collections of real images. The medical CBIR system provides a tool to browse the representative images of each type. These systems rely on color, texture and property of image (bitdepth, width, height, size) features. Retrieval is made on the basis of image features such as histogram, color histogram, GLCM (gray level co-occurrence matrix) etc. Color histogram is an effective representation of the color content of an image. The minimum set of features is selected so as to extract particular information. Principal component analysis (PCA) is one of the most important examples that involves transforming the original data into a low dimension new coordinate system and creates a new dataset. The new coordinate system removes the superfluous data, and the new dataset represents the vital information in better way. PCA is used to extract the unique characteristic of the query image, which distinguishes it from the other images. Relevance feedback can originally be developed for improving the effectiveness of information retrieval systems. GUI design used is also important for medical CBIR, provide flexible interface to user. Hence comparison of the query image with the database of images will result in an exact match.

Keyword: *Medical image retrieval, Content-based approach (CBIR), visual based features extraction and retrieval, PCA (Principle component analysis).*

1. INTRODUCTION

Visual interpretation of picture content for indexing and retrieval is made using CBIR. An image retrieval system is a computer system for organizing, searching and retrieving images from a large database of medical images. A database is there having organized collection of data which is characteristically organized in a way to reality aspects model that chains processes requiring information. Most common methods of image retrieval use some method of adding metadata such as captioning, descriptions or keywords to the images so that retrieval can be performed over the annotation words. With the advancement in multimedia and internet technologies, a enormous amount of multimedia data in the form of images, audio and video has been used in many fields like medical treatment, digital forensics, surveillance system, satellite data, video and still images repositories. TBIR (Text Based Image Retrieval) systems are the most common retrieval systems, where manual or automatic annotation of images search is available. A conventional TBIR searches the database for the similar text surrounding the image as given in the query string. But the expressions of the whole visual content of images in words are complex and hence the producing irrelevant results of TBIR may end up. To find the alternative way of searching and overcoming the limitations imposed by TBIR systems, content based image retrieval systems (CBIR) were developed.

Recently, a large number of digital medical images have been produced in medical diagnostic centers. Comprehensive image databases are prepared having various images, including MRI (magnetic resonance imaging), X-ray, CT (computed tomography), US (ultrasound) etc. By using a description of image content, the powerful retrieval of the desired images is the most important aspect of image database management. CBIR refers to the retrieval of images from a database using information directly derived from the content of images themselves [1]. Content here includes textures, color or any other information that can be derived from the image itself. Content-based image retrieval (CBIR) makes use of image features, such as low-level and semantic features, for indexing the images with minimal human involvement. The feature content of images is a powerful and direct query which can be used to search for other images containing similar content.

2. LITERATURE REVIEW

Several research papers were studied on CBIR and summary of few of them is reported here. Wei et al. (2006) implemented content-based Approach to Medical Image Database Retrieval method. The paper addressed the problems and challenges of medical image retrieval and reviewed the existing medical CBIR systems. It also provided greater details on the key components of content-based image retrieval systems for medical imaging applications. The methodology of CBIR systems for digital mammograms was described and potential research issues in the future research agenda were discussed. However, CBIR systems mainly rely on low-level features and most medical retrieval systems are designed for one particular type of medical images and

ISSN: 2348-4098
Volume 3 Issue 2: 2015

current research on medical CBIR rarely evolves the relationship between CBIR and user interface design. The image retrieval community still awaits the construction and implementation of a scientifically-valid evaluation framework and standard test bed [1]. Müller et al. (2004) suggested review of CBIR systems in medical applications such as clinical benefits and future directions. The propositions for the use of image retrieval in medical practice and the various approaches are explained. The techniques used in the implemented system are described. Possible clinical benefits of image retrieval systems in clinical practice as well as in research and education are identified. However no comparison of interfaces is reported and no real usability studies have been published. Performance comparisons for different feature sets have never been performed [2].

Akgul et al. (2010) proposed in which an overview of the CBIR technology along with a general description of its key components is provided. State-of-the-Art Medical CBIR Systems present an analysis of existing medical CBIR systems. The challenges faced by the generic CBIR technology in the radiology domain have been pointed out. Some of the opportunities to advance CBIR in context have been described. There is lack of tools for data collection. Early detection of hepatocellular carcinoma is of low priority. Paper focuses lightly to the variety of data types and complex database structures rather it specific focus is required for efficiently and effectively searches for similarities [3]. Atnafu et al. (2002) implemented a Content-Based and Metadata Retrieval in medical image. In which global image data model support Metadata and low level descriptions have been implemented. Image data repository model capture all the data described has been presented. Novel image data model for managing image data has also been implemented. Detailed study and implementation of the similarity based operation was described. However, there was possible integration of temporal concept. Evolutionary content of medical images are important issue. The development of a query optimization model and possible integration of temporal concept were needed [4]. Avi kak et al. (2002) suggested CBIR in which the feature extraction strategies were discussed. The nature of the low level features was described. New interactive scheme for contour extraction was described. This is too computationally expensive for physician Interaction on the workstation of the day. It is not possible to use the same set of parameters and decision threshold for all the images in the database [5].

Luz et al. (2006) analyzed DICOM features in Content Based medical image Retrieval using a multi layer approach in which a hybrid approach was used to perform content based retrieval on medical image database was presented. New procedures to execute medical image retrieval have been presented. The material and image processing method and the various similarity modules were described. Actual base of cases was still lacking of some image modalities and types [6]. Akakin et al. (2012) implemented CBIR for Microscopic Images and the feature extraction techniques employed to the image in Database was explained. Two-Tier retrieval approach for multi Image Queries was suggested. The Dataset and Experimental Results was presented. However, investigation of more effective texture and color feature extraction methods is required. Improvement to the robustness of the system is needed. Evaluation of the performance of the system on automatically selected HPF regions for the Query [7]. Khapli et al. suggested CBIR system for biomedical Images highlighting the challenges and open issues in which the general CBIR system for large medical Database was explained. The challenges specific to this area along with the open issues was described. Difference in general CBIR and medical CBIR was discussed. However, Retrieval reliability necessary for biomedical images was not discussed in CBIR technique used so far for stock images. Fully automated feature extraction from medical images was very challenging problem. Dimensionality of feature represented highly affects the quality and retrieval efficiency [8].

Willy et al. (2004) implemented an intelligent Retrieval system for handling multiple organs of interest in which an experimental design of an intelligent retrieval system for handling multiple organs of interest was presented. This was proposed to overcome a problem caused by the subjectivity of a user's perception. It was needed to apply the CBMIR to real medical images. Future work aimed at adding Genetic Algorithms to the CBMIR system [9]. Shyu et al. suggested Local Versus Global Features for CBIR in which an empirical evaluation of the implementation illustrating that local features significantly improve performance over using only global features was presented. A sensitivity study that shows that physician Subjectivity has little impact on retrieval performance was discussed. The reasoning architecture of ASSERT, the implementation of the approach was reviewed. However, the database needs more images from IUMC and other sources. Incorporating user feedback into the design of the indexing scheme. Increasing the speed of retrieval was necessary [10]. Marchiori et al. (2001) suggested a CBIR for medical images. Flaws in the trial setup and how the trial led to a new model for retrieval was discussed. However, evaluation trial did not precisely model the conventional diagnosis process, the result justify further investigation [11]. Noorzaie et al. (2004) implemented a system for distributed image acquisition in which the system architecture was described. Data storage and transfer functionalities were explained. The accuracy of the system can be further improved by using more sophisticated shape features. The retrieval efficiency can be enhanced by exploiting a multidimensional indexing structure. It can be extended to multiple significant objects by using flexible region matching method [12].

Celebi et al. (2004) implemented CBIR using Models of Human Perception and an overview of retrieval system for medical images was presented. The dimensionality reduction and feature selection were described. The human perception of similarity experiments was described. Optimization of the similarity function using a genetic Algorithm was presented. However, the improvement was needed for Image retrieval based on general similarity. Human judge similarity information was not used to improve the performance of CBIR system [13]. Kumar et al. (2013) implemented a Content-

ISSN: 2348-4098
Volume 3 Issue 2: 2015

Based Medical Image Retrieval and a survey of Applications to Multidimensional and Multimodality Data was presented. A review of state-of-the-art medical CBIR approaches in five main categories has been presented. Many challenges were there related to feature selection, retrieval visualization and interpretation from various modalities and efficient image processing [14].

3. PROBLEM IDENTIFICATION

The related research reports following problems in existing research contributions:

- Current medical CBIR systems mainly rely on single features.
- No comparison of interfaces is reported and no real usability studies have been published.
- The development of a query optimization model and possible integration of temporal concept were needed.
- Improvement to the robustness of the system is required.
- Retrieval reliability necessary for biomedical images was not met by CBIR technique.
- Dimensionality of feature represented highly affects the quality and retrieval efficiency.
- The accuracy of the system can be further improved by using more sophisticated features.
- The retrieval efficiency can be enhanced by exploiting a multidimensional indexing structure.

4. METHODOLOGY

In Content-based retrieval, the contents of images are used to represent and access the images. A typical content-based retrieval system is divided into two parts: off-line feature extraction and online image retrieval. A conceptual framework for content-based image retrieval is shown in Fig. 1.

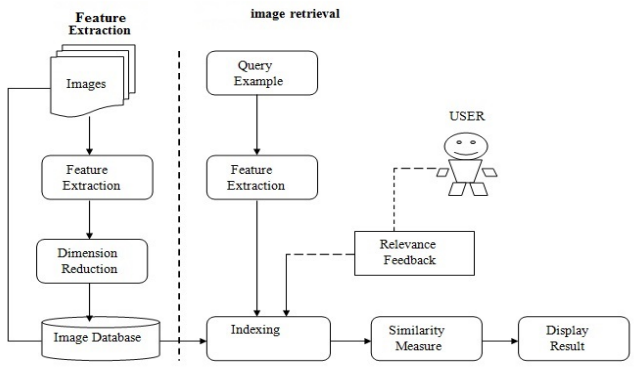

Figure 1: Content-based image retrieval.

In feature extraction, the contents of the images in the database are extracted using histogram and described with a multi-dimensional feature vector. Feature dataset stored in the database is composed the feature vectors of the image. The visual attributes (color, texture, and spatial information) of each image are extracted and stored in a feature database. For extraction of texture feature the GLCM (gray level co-occurrence matrix) has been used. This method is a way of extracting statistical texture features, has been used in a several of applications. In image retrieval, in search of desired images the user submits a query example to the retrieval system. PCA is used to extract the unique characteristic of the query image, which distinguishes it from the other images. Hence comparison of the query image with the database of images will result in an exact match. The main use of PCA is to reduce the dimensionality of a data set while retaining as much information as possible. The concept of PCA to recognize images by extracting their principal components is used. The distances (i.e., similarities) between the feature vectors of the query example and those of the media in the feature dataset are then computed and ranked by using PCA classifier. The system ranks the search results and then returns the results that are most similar to the query image. The user can provide relevance feedback to the retrieval system if the user is not satisfied with the search results which have a mechanism to learn the user's information needs.

5. EXPERIMENTAL SETUP AND RESULT

An experiment has been setup using MATLAB tool with operating system windows7.There has been a database of 60 images of three different types of (X-ray, CT-scan, MRI) medical images. There are 20 trial hits have been taken for the retrieval of medical images with the query images in which 15 hits gave the retrieved images and 18 trial hits gave the relevant results. Relevance judgment plays the vital role or part of performance evolution. Criteria for measurement of performance evolution are as follows:

Table1: Criteria for measurement of Performance evolution of CBIR.

Score	Criteria
1.0	The retrieved images belong to the class of query image.
0.5	The retrieved image belongs to one of the abnormal classes, But not the class of query image.
0	The retrieved image does not belong to any abnormal class.

For evaluating the effectiveness of information retrieval system basic measures used are Precision and Recall. Precision (P) is the ratio of the score retrieved hit (S_i) to the total no. of hits retrieved (N). Recall (R) is the ratio of relevant retrieved (R_n) to the total no. of relevant records in the database.

$$P= \frac{\Sigma_{i=1}^{n} S_i}{N} \quad , \quad R= \frac{R_n}{T_n}$$

The main contribution of this work is to present a sound CBIR methodology for different medical images. The methodology follows the two steps in which first is feature extraction and the second is image retrieval stage. The purpose of the feature extraction is to collect feature samples from the database, then apply it for the feature extraction in the image retrieval stage. A medical CBIR system is based on effective color histogram and GLCM (gray level co-occurrence matrices) technique. These approaches were also proposed for improvement and the effectiveness of normalization. The best

ISSN: 2348-4098
Volume 3 Issue 2: 2015

Precision rate (P) and Recall rate (R) are 87.5% and 66.66% respectively were achieved with the experiment. Relevance feedback can originally be developed for improving the effectiveness of information retrieval systems. The accuracy of the system can be improved by using more sophisticated features. The best effective approach Color histogram used, represent the efficient color content of an image. GUI design used, is also important for medical CBIR, provide flexible interface to user.

Some GUI results are:

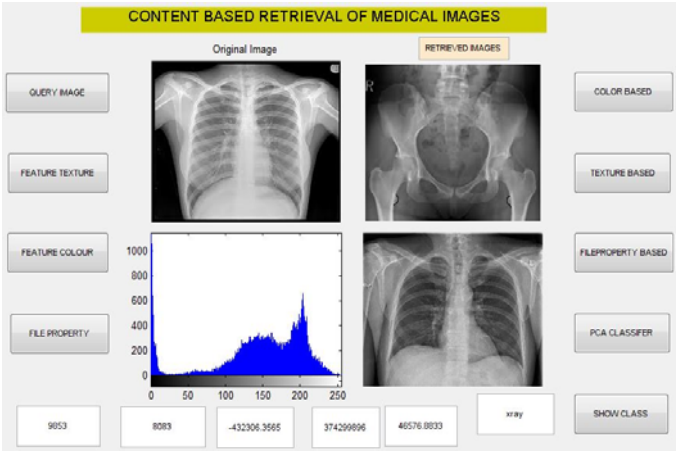

Figure 2: Retrieved image of class of X-ray Image with Histogram plot

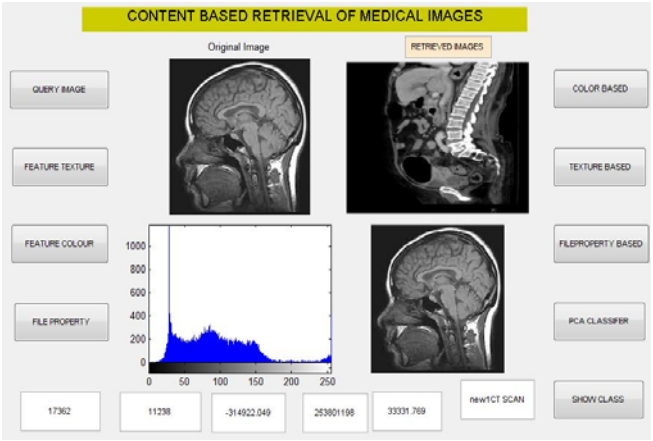

Figure3: Retrieved image of class of CT- Scan Image with Histogram plot.

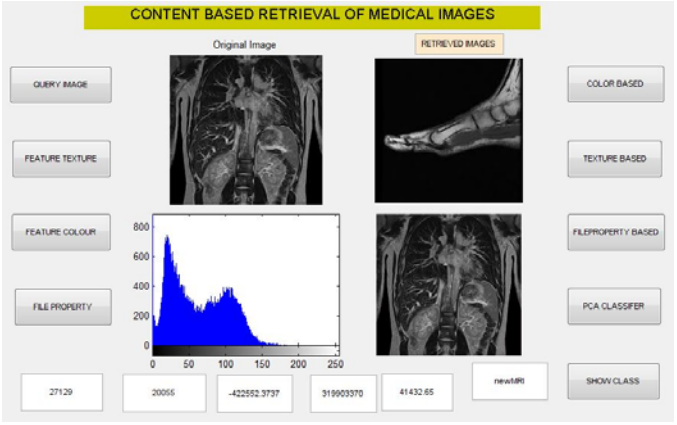

Figure4: Retrieved image of class of MRI Image with Histogram plot.

6. CONCLUSIONS

CBIR is used for an organizing, browsing, and indexing large collections of medical images. CBIR uses an intelligent system having potential to identify and appreciate the composite content of medical images. The flexible method uses for image retrieval to achieve high efficiency. Retrieval can be used for large database of medical images.

REFERENCES

[1] Wei, Chia-hung, Chang-Tsun Li and Roland Wilson. " A Content- Based Approach To Medical Image Database Retrieval." page 258-291, 2006.

[2] Muller*, Henning, et al.: "A review of content based image retrieval systems in medical Applications-clinical benefits and future directions." International journal of medical Informatics (2004), doi: 10.1016, 2004.

[3] Akgul[1] Ceyhun Burak, et al: "Content Based Image Retrieval in Radiology: Current Status and Future Directions." Journal of Digital Imaging. Published April 08, 2010, Doi: 10.1007/s10278-010-9290-9, 2010.

[4] Atnafu, Solomon, Richard Chbeir, Lionel Brunie. "Content Based and Metadata Retrieval in Medical Image Database." Proceeding of the 15th IEEE symposium on computer-based medical systems (CBMS 2002) 1063-7125/02, IEEE, 2002.

[5] Kak, Avi and Christina Pavlopoulou. "Content Based Image Retrieval from Large Medical Databases." Proceeding of the 1st international symposium on 3D data processing visualization and transmission (3DPVT'02) 0-7695-1521-5/02, IEEE, 2002.

[6] Jr. da, Luz Antonio, Daniel D. Abdala, Aldo v. Wangenheim, Eros Comunello: "Analyzing DICOM and non-DICOM Features In ContentBased Medical Image Retrieval: A Multi-Layer Approach." IEEE, 2006.

[7] Akakin, Hatice Cinar and Metin N. "Content-Based Microscopic Image Retrieval System for Multi-Image Queries." IEEE Transactions on Information Technology in Biomedicine, Vol.16, No.4, July 2012.

[8] Khapli[1] Vidya R. and Anjali S. Bhalachandra[2]. "CBIR System for Biomedical Images: Challenges and Open Issues."INDIA: University of Pune, ref. BCUD/578, January 22, 2007.

[9] Willy, Paul Miki and Karl-Heinz Kufer. "Content – Based Medical Image Retrieval (CBMIR): An Intelligent Retrieval System For Handling Multiple Organs Of Interest."Proceeding of the 17th IEEE symposium on computer-based medical systems (CBMS'04)1063- 7125/04, IEEE, 2004.

[10] Shyu, C.R., C.E. Brodly, A.C. kak, A kosaka: "Local versus Global Features for Content based Image Retrieval." National science foundation under Grant No. IRI9711535.

[11] Marchiori A., C. Brodley, L.Broderick, A.M. Aisen.: "CBIR for Medical Images-An Evaluation Trial." Page 89-93, IEEE, 2001.

[12] Noorzaie, Ibrahim, Y.Alp Aslandogan, M.Emre Celebi. "A system for Distributed image Acquisition, Content – Analysis and similarity Retrieval." Page 168-173, IEEE, 2004.

ISSN: 2348-4098
Volume 3 Issue 2: 2015

[13] Celebi, M. Emre and Y.Alp Aslandogan. "Content-Based Image Retrieval Incorporating Models of Human Perception." Proceeding of the international conference on information Technology: Coding and Computing (ITCC'04)0-7695-2108- 8/04, IEEE, 2004.

[14] Kumar, Ashnil, Jinman Kim, Weidong Cai, Michael Fulham, dagan feng.: "Content- Based Medical Image Retrieval: A Survey of Applications to Multidimensional and Multimodality Data." DOI 10.1007/s10278-013-9619-2, springer, 2013.