IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012
ISSN (Online): 1694-0814
www.IJCSI.org

300

Content Based Image Retrieval System for Medical Databases (CBIR-MD) - Lucratively tested on Endoscopy, Dental and Skull Images

Ashish Oberoi and Manpreet Singh

Department of Computer Science & Engineering, M.M. Engineering College, M.M. University
Mullana, Ambala, Haryana, PIN-133 207, India

Abstract

In medical field, digital images are produced in ever increasing quantities and used for diagnostics and therapy. The swift expansion of digital medical images has enforced the requirement of efficient Content-based image retrieval system for retrieving medical images that are visually similar to query image. Such systems provide great assistance to doctors in clinical care and research. In this paper, we have designed a Content Based Image Retrieval System for Medical Databases (CBIR-MD) based on various techniques like Fourier descriptor, Euclidean distance, Haar Wavelet transformation, Canberra distance and analyzed its performance on Endoscopy, Dental and Skull images.

Keywords: CBIR-MD, Fourier Descriptor (FD), Haar Wavelet (HW), Euclidean Distance (ED), Canberra Distance (CD).

1. Introduction

Content-based image retrieval (CBIR) is the application of computer vision techniques to the problem of digital image search in large databases. CBIR enables to retrieve the images from the databases [1, 2]. Medical images are usually fused, subject to high inconsistency and composed of different minor structures. So there is a necessity for feature extraction and classification of images for easy and efficient retrieval [3]. CBIR is an automatic retrieval of images generally based on some particular properties such as color composition, shape and texture [4, 5]. Every day large volumes of different types of medical images such as dental, endoscopy, skull, MRI, ultrasound, radiology are produced in various hospitals as well as in various medical centres [6]. Medical image retrieval has many significant applications especially in medical diagnosis, education and research fields. Medical image retrieval for diagnostic purposes is important because the historical images of different patients in medical centres have valuable information for the upcoming diagnosis with a system which retrieves similar cases, make more accurate diagnosis and decide on appropriate treatment. The main objective of this research work is to retrieve the similar images matching the query image from medical databases by using feature extraction and similarity measurement techniques.

This paper is organized as follows: the existing methods and its related literature survey is presented in section 2. The proposed method is presented in section 3. Implementation and Result analysis is described in section 4. Finally, this paper is concluded in section 5.

2. Related Work

In picture archiving and communication system (PACS), image information is retrieved by using limited text keyword in special fields in the image header (e.g. patient identifier). Content-based image retrieval (CBIR) has received significant attention in the literature as a promising technique to facilitate improved image management in PACS system [7, 8]. The Image Retrieval for Medical Applications (IRMA) project [8,9] aims to provide visually rich image management through CBIR techniques applied to medical images using intensity distribution and texture measures taken globally over the entire image. This approach permits queries on a heterogeneous image collection and helps in identifying images that are similar with respect to global features e.g. all chest x-rays in the AP (Anterior-Posterior) view. The IRMA system lacks the ability for finding particular pathology that may be localized in particular regions within the image. In contrast, the Spine Pathology and Image Retrieval System (SPIRS) [10, 11, 12] provides localized vertebral shape-based CBIR methods for pathologically sensitive retrieval of digitized spine x-rays and associated person metadata. Image Map [14] is so far, the only existing medical image retrieval that considers how to handle multiple organs of interest and it is based on spatial similarity. Consequently, a problem caused by user subjectivity is likely to occur, and therefore, the retrieved image will represent an unexpected organ. ASSERT [15] (Automatic Search and Selection Engine with Retrieval Tools) is a content–based retrieval system focusing on the analysis of textures in high resolution Computed Tomography (CT) scan of the lung. In WebMIRS [16] system, the user manipulates GUI tools to create a query such as, "Search for all records for people over the age of 65 who reported chronic back pain. Return the age, race, sex and age at pain onset for these people." In response, the system return values for these four fields of all matching records along with a display of the associated x-ray images. CervigramFinder

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012
ISSN (Online): 1694-0814
www.IJCSI.org

301

system [17] operates on a subset of the cervigram database. To use this system, the user defines a query by marking a region of interest on an image through GUI. SPIRS-IRMA [18] is a CBIR system is based on the merits of two already existing systems (SPIRS & IRMA). So there is a need of absolute error free, efficient and automatic CBMIR system which can really helpful in medical stream.

3. Proposed System

In proposed system like other CBIR system, images are represented by appropriate feature vector in feature space. Such feature vector gives meaningful information of image properties. Fig. 1 reflects the working as well as milestone achieved after completion of each discrete process in proposed CBIR-MD system.

3.1. Metrics for Feature Extraction

Fourier Descriptors

Fourier transform is used to generate the feature vectors based on the mean values of real and imaginary parts of complex numbers of polar coordinates in the frequency domain. Fourier Descriptors (shape based) can be used as a dominant feature for boundaries and object representation [21]. Consider a M point digital boundary, starting from an arbitrary point (x_0, y_0) then $(x_1, y_1),...,(x_{M-1}, y_{M-1})$ can be generated. These coordinates can be represented in a complex form as:

$$q(m) = x(m) + j\,y(m)\,,$$
$$m = 0, 1, 2...M-1$$

The Discrete Fourier Transform (DFT) of q(m) gives

$$b(k) = \frac{1}{M} \sum_{m=0}^{M-1} q(m)\, e^{-j2\pi km/M}$$
$$k = 0,1,2,...M-1$$

The complex coefficients b (k) are called Fourier descriptors of the boundary.

Haar Wavelet

Haar Wavelets [23] are fastest to compute and simplest to implement. In addition user queries tend to have large constant-colored regions, which are well represented by this basis. The technical disadvantage of Haar Wavelet is that it is not continuous and therefore not differentiable.

In CBIR-MD, both Fourier Descriptor and Haar Wavelet are used for feature extraction.

3.2 Metrics for Similarity Comparison

Distance metric is the main tool for retrieving similar images from large medical databases. In CBIR-MD, Euclidean distance [2] and Canberra distance [24] are used for the purpose of similarity comparison.

$$ED = \sqrt{\sum_{i=1}^{N}\left(fq(i)-fdb(i)\right)^2}$$

Where $f_q(i)$ stands for i^{th} query image feature and $f_{db}(i)$ for corresponding feature vector database. Here N refers to number of images in database.

$$\text{Canberra Distance(CD)} = \frac{\sum_i |u_i - v_i|}{\sum_i |u_i| + |v_i|}$$

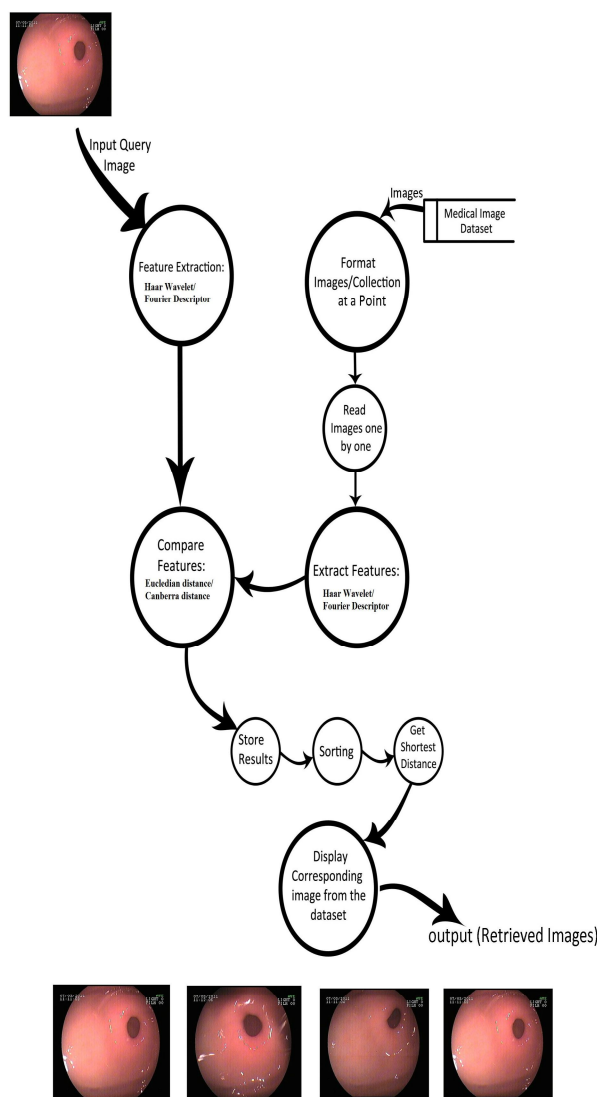Where u and v are both n-dimensional vectors.

Fig. 1 Data Flow Diagram of Proposed Retrieval System

3.3 Retrieval Process

The following steps are performed in the retrieval process:

Step 1: Input query medical image.

Step2: Extract features by using Fourier Descriptor (FD) or Haar Wavelet (HW) .

Step 3: Format/Collect the medical images from

the medical databases at a point.

Step 4: Read medical images one by one.

Step 5: Extract features by either of the above mentioned techniques.

Step 6: Compare features of the query medical image with medical images from the database by Euclidean Distance (ED)/ Canberra Distance (CD) technique.

Step 7: Store the result.

Step 8: Perform sorting of the result.

Step 9: Display the corresponding medical images.

4. Experimental Study

4.1 Dataset for the Experiment

The functional code of proposed system is implemented using MATLAB 7.8 on an Intel Core 2 duo, 2 GHz window based laptop. The system is tested on three different dataset having 200 dental images, 1100 endoscopy images and 50 skull images respectively.

4.2 Performance Parameters

Precision and Recall (P-R): The images are retrieved and measured against P-R [22] as:

$$P = \frac{\text{Number of relevant images retrieved}}{\text{Total number of images retrieved}}$$

$$R = \frac{\text{Number of relevant images retrieved}}{\text{Total number of relevant images in the database}}$$

where P is the ratio to measure accuracy and R is used to measure robustness.

4.3 Experiments and Results

Subject Test on Endoscopy Images Dataset: The retrieval of images is observed for the database of 1100 Endoscopy images. The retrieval accuracy with FD/CD, FD/ED, HW/CD and HW/ED is observed and shown in Table 1.

Table1: Precision-Recall Against Various Descriptors for Endoscopy Images

The results obtained from these descriptions can be seen through output screens of the developed system for any query image (represented by Fig. 2) in Fig. 3, Fig. 4, Fig. 5 and Fig. 6.

Feature Extraction Technique	Distance Calculation Technique	Delay (Output) in seconds	Precision %	Recall %
FD	CD	4-6	80	82
FD	ED	22-26	74.2	74.0
HW	CD	11-16	79.3	72
HW	ED	4-8	79.8	80

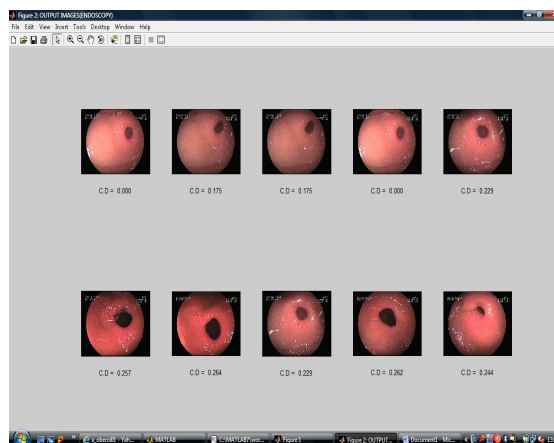Fig. 2 Query Image for Endoscopy Dataset

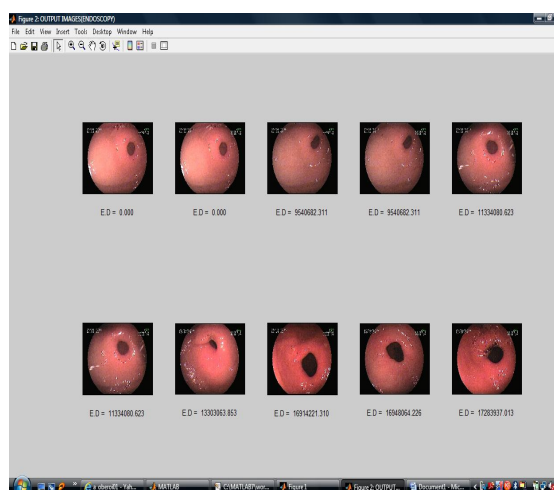Fig. 3 Retrieved Images from Endoscopy Dataset for the Query using FD/CD

Fig. 4 Retrieved Images from Endoscopy Dataset for the Query using FD/ED
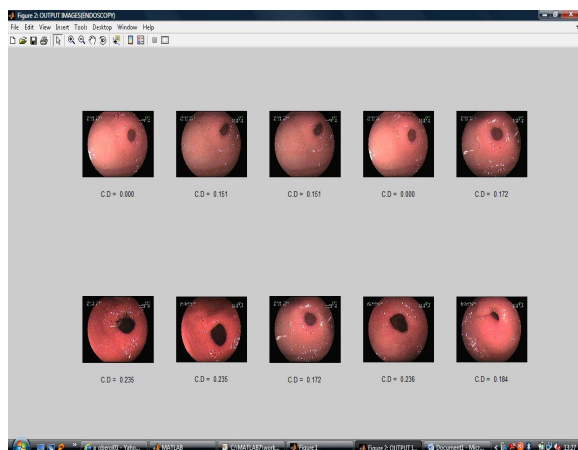
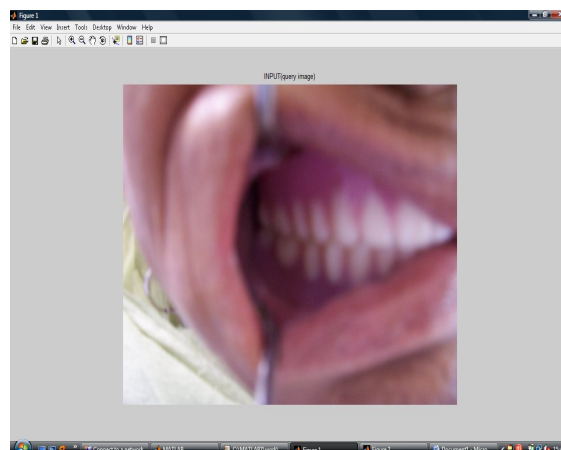Fig. 5 Retrieved Images from Endoscopy Dataset for the Query using HW/CD
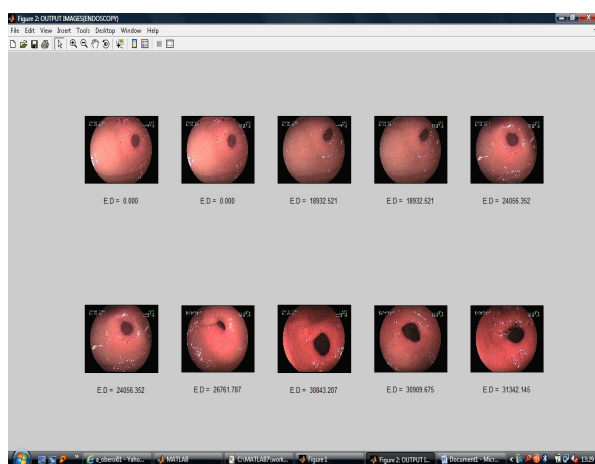

Fig. 7 Query Image for Dental Dataset

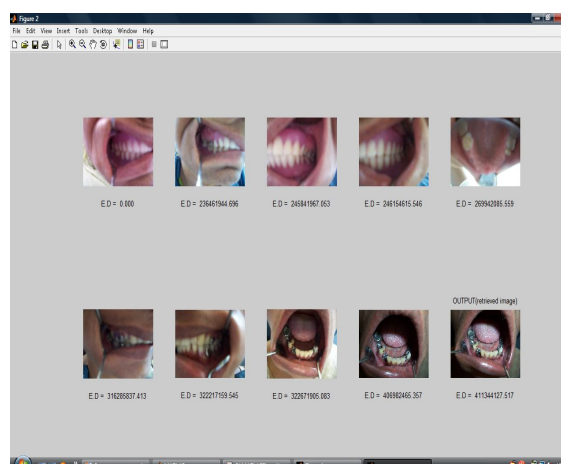Fig. 6 Retrieved Images from Endoscopy Dataset for the Query using HW/ED

Fig. 8 Retrieved Images from Dental Dataset for the Query using FD/ED

Subject Test on Dental Images Dataset: The retrieval of images is observed for the database of 200 dental images. The retrieval accuracy with FD/CD, FD/ED, HW/CD and HW/ED is observed and shown in Table 2.

Table 2: Precision-Recall Against Various Descriptors for Dental Images.

Feature Extraction Technique	Distance Calculation Technique	Delay (Output) in seconds	Precision	Recall
FD	CD	4-6	74.6	72.2
FD	ED	22-26	70.1	69.8
HW	CD	11-16	73.1	70.2
HW	ED	4-8	71.2	69.2

The results obtained from these descriptors can be seen through output screens of the developed system for any query image (represented by Fig. 7) in Fig. 8, Fig. 9, Fig. 10 and Fig. 11.
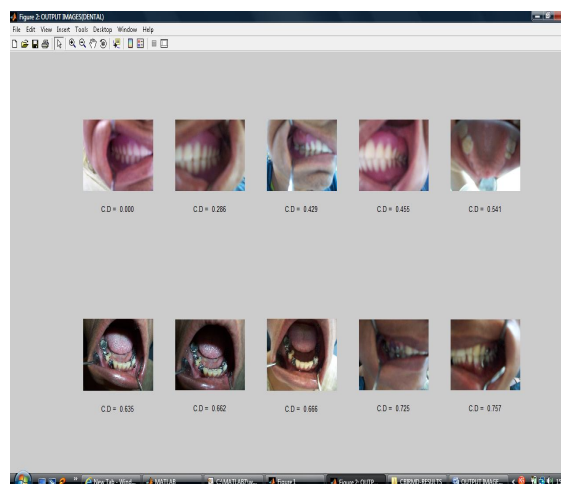

Fig. 9 Retrieved Images from Dental Dataset for the Query using FD/CD

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012
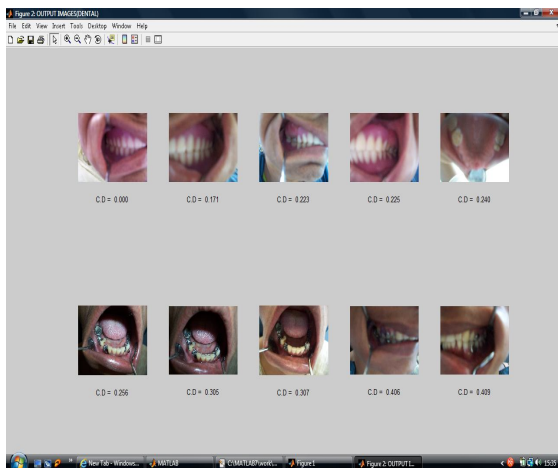ISSN (Online): 1694-0814
www.IJCSI.org

304

Fig. 10 Retrieved Images from Dental Dataset for the Query using HW/CD
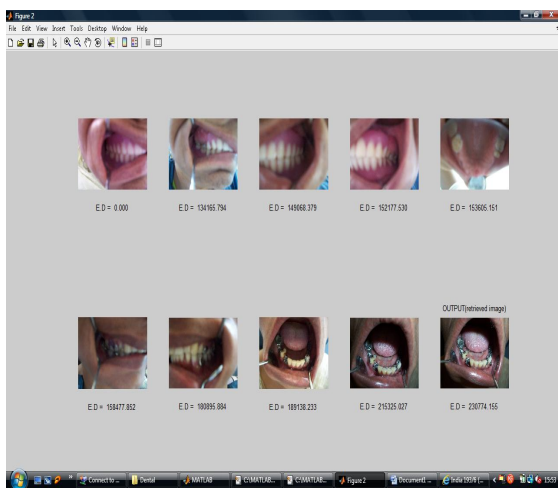

Fig. 12 Query Image for Skull Dataset

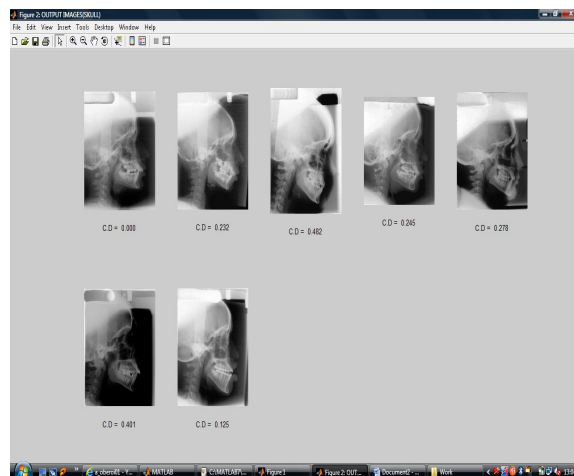Fig. 11 Retrieved Images from Dental Dataset for the Query using HW/ED

Fig. 13 Retrieved Images from Skull Dataset for the Query using FD/CD

Subject Test on Skull Images Dataset: The retrieval of images is observed for the database of 50 Skull images. The retrieval accuracy with FD/CD, FD/ED, HW/CD and HW/ED is observed and shown in Table 3.

Table 3: Precision-Recall Against Various Descriptors for Skull Images.

Feature Extraction Technique	Distance Calculation Technique	Delay (Output) in seconds	Precision	Recall
FD	CD	4-6	75.6	72.3
FD	ED	22-26	70	68.2
HW	CD	11-16	73.2	70.7
HW	ED	4-8	74.1	72

The results obtained from these descriptors can be seen through output screens of the developed system for any query image (represented by Fig. 12) in Fig. 13, Fig. 14, Fig. 15 and Fig. 16.
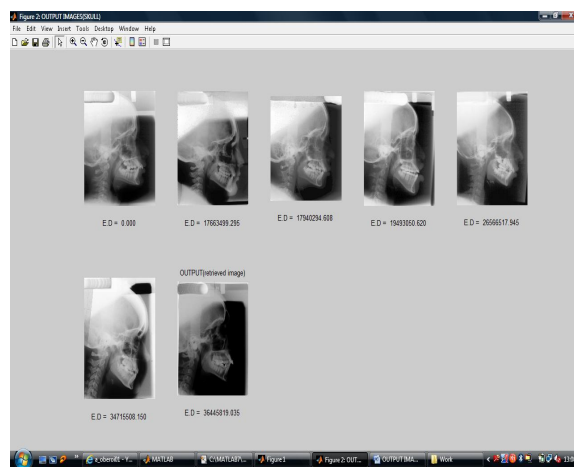

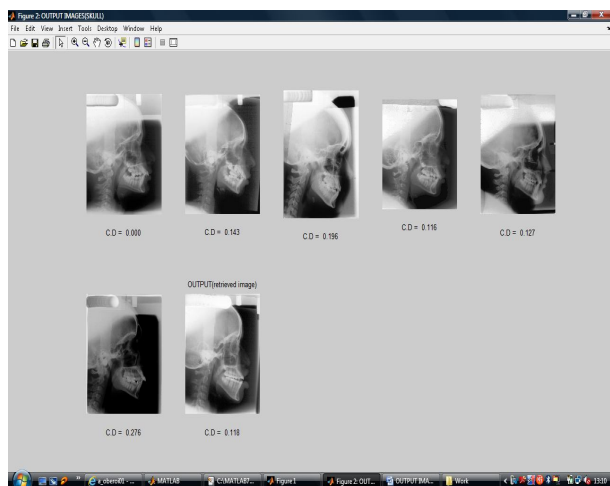Fig. 14 Retrieved Images from Skull Dataset for the Query using FD/ED

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012
ISSN (Online): 1694-0814
www.IJCSI.org

305

Fig. 15 Retrieved Images from Skull Dataset for the Query using HW/CD
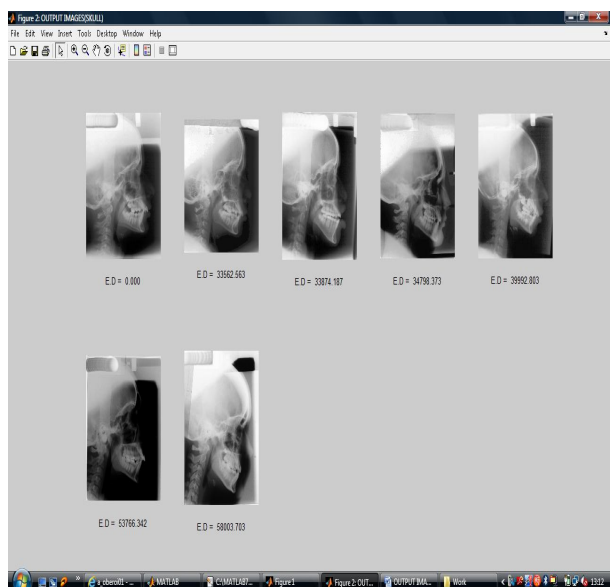

Fig. 16 Retrieved Images from Skull Dataset for the Query using HW/ED

5. Conclusion

A technique that effectively use most of the information from image is backbone of an efficient content-based image retrieval system for medical databases. In this paper, we have developed an image retrieval system based on various techniques for feature extraction and similarity measurement. The experiment is performed on three different datasets in order to measure the accuracy and robustness of the system. It has been observed from the experimental study that FD and CD combination gives better result in terms of delay, precision and recall. Future work may be carried out in the the field of image enhancement and there is a need for GUI based CBIR-MD for creating better user interface to interact and work efficiently with the system.

References

[1] M.Smeulders, Worring, and M. Santini, "Content-based image Retrieval at The End of Early Years", IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, No.12, 2000, pp. 1349-1380.

[2] V.S. Murthy, E.Vamsidhar, J.N.V.R. Swarup Kumar, and P. Sankara Rao,"Content based Image Retrieval using Hierarchical and K-means Clustering Techniques", International Journal of Engineering Science and Technology, Vol. 2, No. 3, 2010, pp. 209-212.

[3] B. Ramamurthy, and K.R. Chandran, "CBMIR:Shape-based Image Retrieval using Canny Edge Detection and K-means Clustering Algorithms for Medical Images", International Journal of Engineering Science and Technology, Vol. 3, No. 3, 2011, pp. 209-212.

[4] Roberto Parades, Daniel Keysers, Thomas M. Lehman, Berthold Wein, Herman Ney, and Enrique Vidal,"Classification of Medical Images Using Local Representation", Workshop Bildverarbeitung fur die Medizin, 2002, pp.171-174.

[5] Wei Zhang, Sven Dickinson, Stanley Sclaroff, Jacob Feldman, and Stanley Dunn,"Shape – Based Indexing in a Medical Image Database", Biomedical Image Analysis, 1998, pp. 221-230.

[6] Monireh Esnaashari, S. Amirhassan Monadjami, and Gholamali Naderian,"A Content-based Retinal Image Retrieval Method for Diabetes- Related Eye Diseases Diagnosis", International Journal of Research and Reviews in Computer Science(IJRRCS), Vol. 2, No. 6, 2011, pp. 1222-1227.

[7] H. Müller, N. Michoux, D. Bandon, and A. Geissbuhler," A review of content-based image retrieval systems in medical applications-Clinical benefits and future directions", International Journal of Medical Informatics, Vol. 73, No. 1, 2004, pp. 1-23.

[8] T.M. Lehmann, M.O. Guld, C Thies,B Fischer , K. Spitzer, and D. Keysers," Content-based image retrieval in medical applications", Methods of Info in Med, IOS Press , Vol. 43, No. 4, 2004, pp. 354–361.

[9] C. Thies, M.O. Guld, B Fischer, and T.M. Lehmann,"Content-based queries on the CasImage database within the IRMA framework", Lecture Notes in Computer Science,Springer 3491, 2005, pp. 781–792.

[10] S. Antani, L.R. Long, and G.R. Thoma, "Content-based image retrieval for large biomedical image Archives", Proceedings of 11th World Congress Medical Informatics, 2004, pp. 829–833.

[11] L.R. Long, S.K. Antani, and G.R. Thoma, "Image informatics at a national research center", Computer Medical Imaging & Graphics (ELSEVIER), Vol. 29, 2005, pp. 171–193.

[12] G.R. Thoma, L.R. Long, and S.K. Antani, "Biomedical imaging research and development: knowledge from images in the medical enterprise",Technical Report Lister

IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 3, No 1, May 2012
ISSN (Online): 1694-0814
www.IJCSI.org

306

Hill National Centre for Biomedical Communications, 2006.

[13] R. Chbeir, Y. Amghar, and A. Flory,"MIMS: A Prototype for Medical Image Retrieval", Proceedings of 6th Conference on Content-Based Multimedia Information Access, 2000, pp. 54-59.

[14] E.G.M. Petrakis, and C. Faloutsos, "ImageMap: An Image Indexing Method Based on Spatial Similarity", IEEE Transaction on Knowledge and Data Engineering, 2002, pp. 979–987.

[15] Chi-Ren Shyu, Carla E. Brodley, Avinash C. Kak, and Akio Kosaka," ASSERT:A Physician-in-the-Loop Content-Based Retrieval System for HRCT Image Databases", Computer Vision and Image Understanding, Vol. 75, No. 1, 1999, pp. 111–132.

[16] L.R. Long, S.R. Pillemer, R.C. Lawrence, G-H Goh, L. Neve, and G.R. Thoma,"WebMIRS: Web-based Medical Information Retrieval System" , Proceedings of SPIE Storage and Retrieval for Image and Video Databases VI, SPIE , Vol. 3312, 1998, pp. 392-403.

[17] Z. Xue, L.R. Long, S. Antani, J. Jeronimo, and G.R. Thoma,"A Web-accessible content-based cervicographic image retrieval system", Proceedings of SPIE medical imaging, Vol. 6919, 2008, pp. 1-9.

[18] S.K. Antani, T.M. Deserno, L.R. Long, M.O. Guld, L. Neve, and G.R. Thoma,"Interfacing global and local CBIR systems for medical image retrieval", Proceedings of the workshop on Medical Imaging Research, 2007, pp.166-171.

[19] M. Henning, G. Antoine, M. Johan, L. Christian , and R. Patrick ,"The Use of MedGIFT and EasyIR for ImageCLEF", Presented in Annual workshop on CLEF, 2005.

[20] Papandreou George, and Petros Maragos, "Adaptive and Constrained Algorithms for Inverse Compositional Active Appearance Model Fitting", Proceeding of IEEE International conference on Computer Vision and Pattern Recognition (CVPR-2008), 2008, pp. 111-117.

[21] Quing chen, Emil Petriu, and Xiaoli Yang,"A comparative Study of Fourier Descriptors and Hu's Seven Moment Invariants for Image Recognition", Proceeding of International conference CCECE, 2004, pp. 103-106.

[22] M. Henning, R. Antoine , and Jean-Paul, "Comparing Feature Sets for content-based Image Retrieval in a Medical Case Database", Proceeding of SPIE Conference on Medical Imaging, 2004, pp. 99-109.

[23] James S. Walker, "A primer on Wavelets and Scientific Applications", 2nd Edition, ISBN: 1584887451, CRC, 2011.

[24] "SciPy Reference guide Release 0.7.dev", written by SciPy community, 2008, pp. 257-260.

Ashish Oberoi is presently serving as Asst. Professor in Department of Computer Science and Engineering , M. M. Engineering College, M.M. University, Mullana (Ambala), India. He obtained his M.Tech (CSE) from Kurukshetra University Kurukshetra and pursing PhD in the field of digital Image processing from M.M. University. He has about 6 years of teaching experience. He has published 16 research papers in International and Indian Journals and Conferences.

Dr. Manpreet Singh is presently serving as Professor and Head, Department of Computer Science and Engineering , M. M. Engineering College, M.M. University, Mullana (Ambala), India. He has about 13 years of experience in teaching and research. He has guided two PhD research scholars and four others are under his supervision. He has published 30 research papers in International and Indian Journals and Conferences.

IJCSI
www.IJCSI.org