Asian Journal of Information Technology 5 (4) : 448-453, 2006

Intelligent Fractured Image Retrieval From Medical Image Databases

[1]H. Khanna Nehemiah, [1]A. Kannan and [2]D. Senthil Kumar
[1]Department of Computer Science And Engineering,
Anna University, Chennai B 600 025, India
[2]Department of Information Technology, M.N.M. Jain Engineering College,
Thorapakkam, Chennai-600096, India

Abstract: The retrieval of stored medical images matching an input medical image is an imperative form of content-based retrieval. For efficient similarity image retrieval and integration, the medical images should be processed systematically to extract a representing feature space vector for each member image. This study explains a system, which takes a fractured image as a query image and retrieves the similar images from the image database using distance metrics and also provides the radiologists with details about the type of fracture and the treatment recommended. The key objective of present research is to retrieve similar X-ray images of fractured reports using K-Nearest Neighbor. Images are matched using color in gray level and texture attributes. Similarity between images is established based on the respective numeric values (Signature). Features are extracted from X-ray images. Indexing is also performed on extracted features using a k-d tree data structure for images and is stored in a backend database for effective retrieval.

Key words: Content based retrieval, IFIR, indexing, feature extraction, image database

INTRODUCTION

In this study, we discuss about a system that is capable of retrieving similar fractured images from the image database using distance metrics and also provide the radiologists with details about the type of fracture and the treatment recommended using content-based retrieval techniques. One of the capabilities of content based retrieval of images is the ability to retrieve images that are similar to a given query image. Our approach called Similar Fractured Image Retrieval technique (IFIR) uses intrinsic visual features of images, such as texture and color in gray level. The content-based retrieval of visual data requires a paradigm that differs significantly from both traditional databases and text-based systems. Present study focuses on study of fractured images, their representation and comparison, based on image texture and color in gray level. Similarity between images is established based on the respective distance metric. The first step in the process is extracting image features to a distinguishable extent. The second step involves matching these features to yield the results that are visually similar. X-ray images of various fractures are used in present research for implementation. Indexing is performed on extracted features using a k-d tree for the images and uses Oracle8i, as a back end, which enables storage, retrieval and management of images. Using K-Nearest Neighbor similar images are extracted from the image database for a given input query image. Our system extracts the primitive features of a query image first and then compares the extracted features with the features of the fractured images in the image database. The image features under consideration are texture and color in gray level. We use matching and comparison algorithms to match and compare the color in gray and texture of one image with the corresponding features of another image.

MATERIALS AND METHODS

Agma Traina, Natalia A. Rosa, Caetano Traina Jr have discussed a prototype system that can perform the retrieval of images through similarity queries based on their contents[1]. Joaquim Cezar Felipe, Agma Traina, Caetano Traina Jr have discussed a statistical approach called co-occurrence matrix to extract the feature texture[2]. Patricia G.Foschi, Deepak Kolippakkam, Huan Liu and Amit Mandvikar have discussed the extraction of the feature color in gray scale[3]. Rinie Egas, Nies Huijsmans, Michael Lew, Nice Sebe have discussed k-d tree indexing technique for images[4]. Melliyal Annamalai, Rajiv Chopra, Susan Mavris, Samuel DeFazio have built a product called Visual Information Retrieval on top of Oracle 8i intermedia which enables storage, retrieval and management of images[5]. Myrian R.B.Araujo, Caetano Traina Jr, Agma

Corresponding Author: H. Khanna Nehemiah, Department of Computer Science And Engineering, Anna University, Chennai B 600 025, India

Traina, Josiane M.Bueno, Humberto L.Razente discussed how to support images in a relational database, so it can fulfill the requirements to be used as the storage mechanism of a Picture Archiving and Communication System (PACS)[6]. Manesh Kokare, B.N.Chatterji and P.K.Biswas have discussed techniques for matching and retrieval of images based on different distance metric such as Manhattan metric, Euclidean metric[7]. Maleq Khan, Qin Ding and William Perrizo have discussed a common similarity function based on the Euclidean distance between two data tuples for finding K-neighbors nearest to the query image[8]. Comparing with all these works, present research is different in the following ways. First, we use a Knowledge Based approach for effective decision making using backward chaining rules. Second, we constructed the knowledge base in consultation with a domain expert. Third, we used rule matching and rule firing techniques for matching rules with facts available in the knowledge base. Finally, we used the k-d tree storage structure for effective indexing of fractured images.

SYSTEM ARCHITECTURE

The perspective of present research is to enable a radiologist to retrieve similar X-ray images from an image database for a given input query image along with description, fracture and treatment given for the retrieved similar images. This artifact will provide a basis to perform analysis of similar medical cases and also aims to reduce the access time to retrieve similar X-ray images by indexing the images stored in the image database. The architecture of the IFIR system is illustrated in Fig. 1. The major components as illustrated in Fig. 1 are stated and discussed as follows.

- Feature Extraction Engine
- Indexing Engine
- Matching and Retrieval Engine

Feature extraction deals with extracting average color in gray level, texture and it returns a feature vector. Indexing is performed on the feature vector for effective retrieval Matching and retrieval is performed using the Euclidian similarity function for k-nearest neighbor and similar images are extracted from the image database. Our system handles the features color in gray level and texture for every image. Image and its corresponding feature vector are stored in the image database. While storing the images in the database indexing is performed based on the image feature vector. The image size must be ensured to be constant (115×115) pixels and the image is a monochrome image stored in JPEG format.

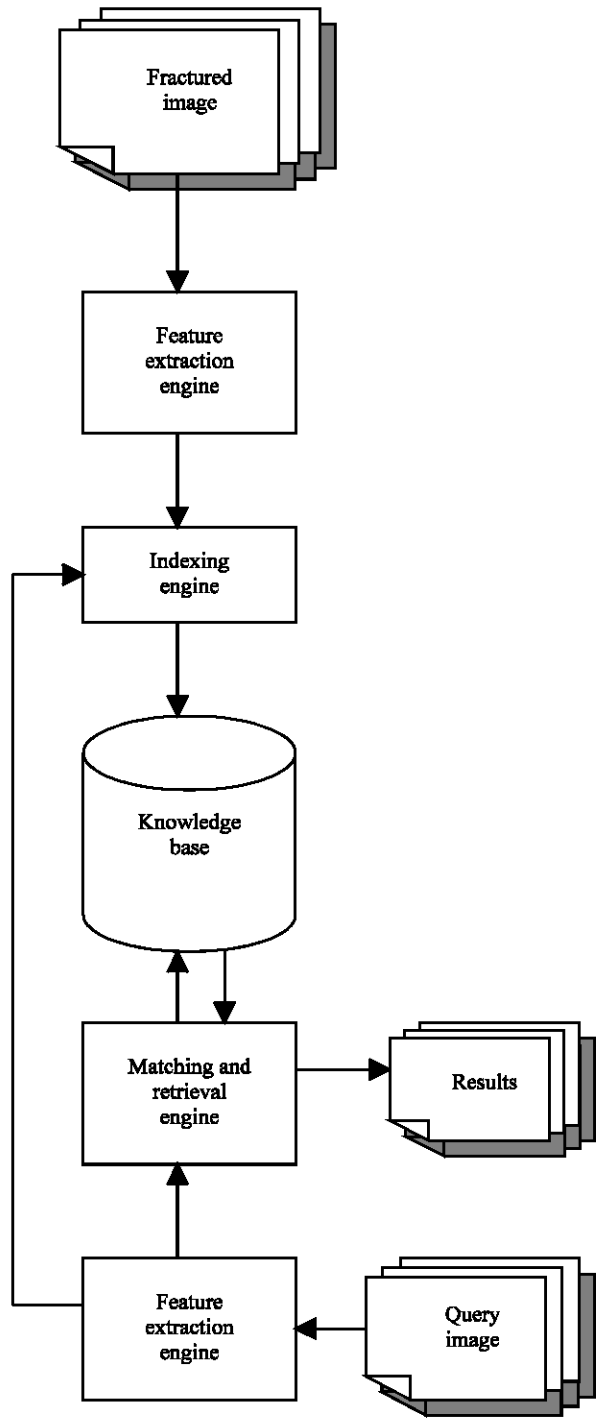


Fig. 1: System architecture

Feature extraction: Feature extraction is the technique of capturing visual content of image for indexing and retrieval. The input image is characterized by signature acquired through computations on the gray (brightness) levels of the image. The input to the feature extraction

0	0	0	1	2
1	1	0	1	1
2	2	1	0	0
1	1	0	2	0
1	0	1	0	0

0	1	2	
6	8	2	0
8	4	5	1
2	5	0	2

Fig. 2: Image gray levels

engine is fractured image in JPEG format of size 115×115 and output is feature vector.

Texture feature extraction: Inter-relationship between the gray (bri-ghtness) levels define the image texture. A statistical approach called Co-occurrence matrix, which is a two-dimensional array of gray levels for a pair of pixels, which are separated by a fixed spatial relationship is used to extract texture feature[2]. The pixel values are used to construct numerical structures, which are associated to the texture pattern of the image. This pattern is based on the inter-relationship between one pixel and its neighbors. Fig. 2 shows an image quantified in three gray levels and sampled at 5×5 pixels and its corresponding Co-occurrence matrix for direction = 90° and distance = 1.

In present research, we have constructed the Co-occurrence matrix as indexes of rows and columns which represent the given range of the image gray levels quantified in 15 and sampled at 16×16 pixels and the value P(i,j) stored at the Position(i,j) is the frequency that gray levels i and j occur with, at a given distance and at a given direction. For instance, consider Fig. 2, the value of Co-occurrence P(0,1) = 8 (value stored at the position (i,j)) is calculated by scanning the gray level matrix and, for each pixel with value 0, its top and bottom neighbors are checked and P(0,1) is incremented whenever a value 1 is found. In present research the gray levels of the image are reduced to 16. Similarly the Co-occurrence matrices for 0°, 45°, 90° and 135° and distances 1, 2, 3, 4 and 5 are generated. This produces 20 matrices of 16X16 integer element per image. For each co-occurrence matrix, the values for the texture descriptors variance, entropy and energy are calculated. The resulting descriptor values are stored in matrix where rows represent the directions (0°, 45°, 90° and 135°) and column represents distances 1, 2, 3, 4 and 5. The final descriptor values are calculated from the descriptor matrix by averaging the values obtained with the different distances for each direction. The descriptor variance is the level of contrast of the image, entropy is suavity of the image and energy is the uniformity of the image. P(i,j) is the value stored at position(i,j) in the Co-occurrence matrix

Color feature extraction: Gray is a color commonly seen in nature. Mixing of white and black in different proportions creates gray color. Depending on the amount of light, the human eye can interpret the same object as either gray or some other color. Two colors are called complementary colors if they produce gray when combined. The psychological primary colors are:

- Black and White
- Blue and Yellow
- Red and Green

Gray scale images are typically composed of shades of gray, varying from black at the weakest intensity to white at the strongest, though in principle the samples could be displayed as shades of any color, or even coded with various colors for different intensities. Grayscale images are distinct from black-and-white images, which in the context of computer imaging are images with only two colors, black and white; grayscale images have many shades of gray in between. In most contexts other than digital imaging, however, the term black and white is used in place of "grayscale"; for example, photography in shades of gray is typically called black-and-white photography. The term monochromatic in some digital imaging contexts is synonymous with grayscale and in some contexts synonymous with black-and-white. Grayscale images are often the result of measuring the intensity of light at each pixel. Grayscale images intended for visual display are typically stored with and 8 bits per sample, which allows 256 intensities (i.e., shades of gray) to be recorded, typically on a non-linear scale.

Extracting Average color in gray scale of an image is performed as discussed in[3].

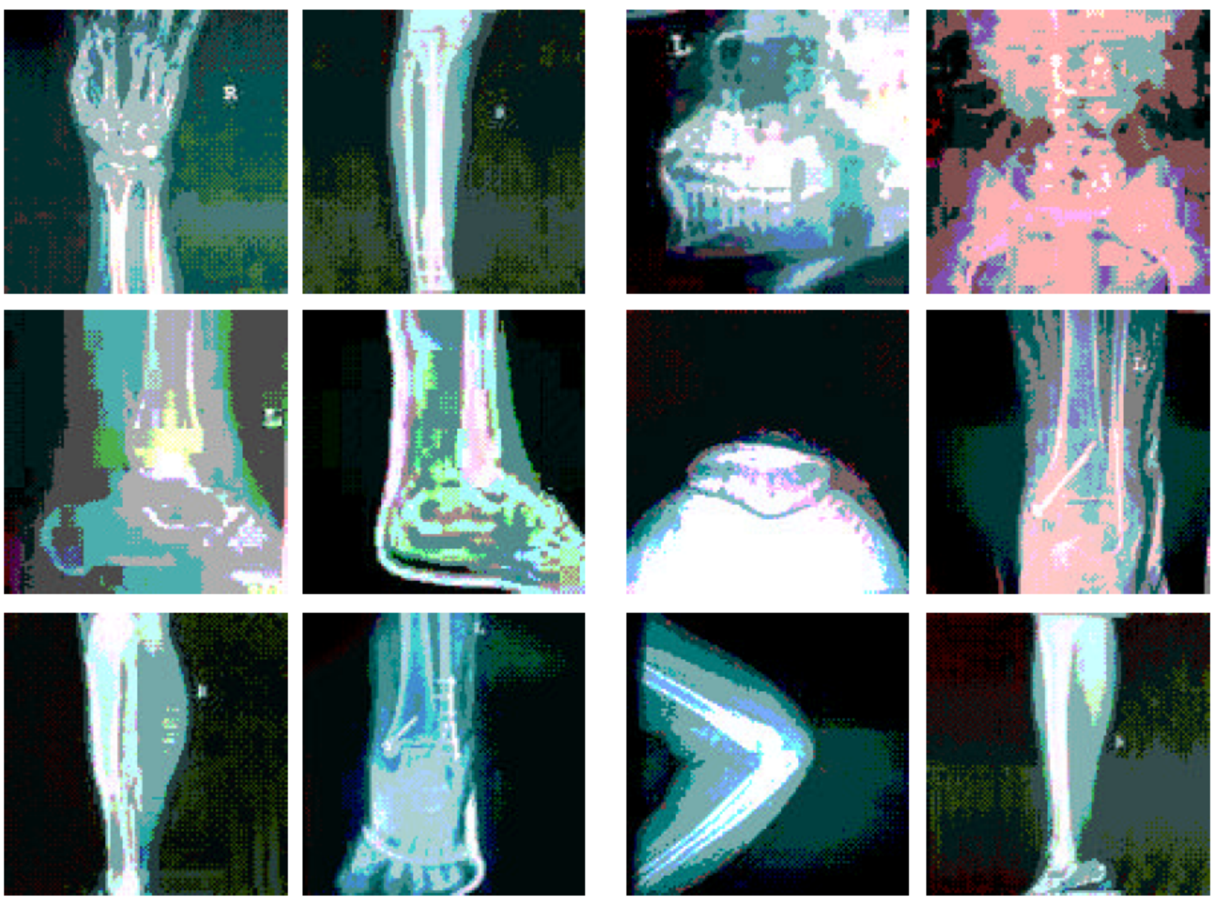
Indexing: The extracted feature values are sent to indexing engine which links the features with the image stored in the database. Indexing is performed based on k-d tree indexing using feature vector[4]. k-d tree is efficient for finding nearest neighbor in high dimensional spaces. Variance and Entropy are to be considered to construct k-d tree with the value of k is 2.

Matching and retrieval: Reading large sets of full-featured image files is a costly operation, So K-Nearest Neighbor extracts values from indexed feature of images as they are being stored in the system. The query image will be more similar to the database images if the distance is smaller. This module takes input as query image and produces output as set of similar images with doctor inference and treatment given. Euclidian similarity function compares two images using the extracted features and returns a nonnegative value that is smaller as more similar the two images are[7,8].

Knowledge base: The knowledge base consists of two major components namely the database and the rule base. The attributes imageid, image which the original image stored in blob format, description which gives description of an image, energy, variance, entropy are texture features that are numerical values derived from the original image, average color is color is sum of intensity of all pixels in the current block of an image divided by total pixels in the

Description: Leg lateral view showing mid shaft of tibia
Fracture: Mid shaft of tibia
Treatment: Interlocking with nail fixation

Fig. 3: A Sample Fractured Image with Description, Fracture and Treatment



Query image: **Query image:**

Fig. 4: Subset of our data set used for experimenting present research

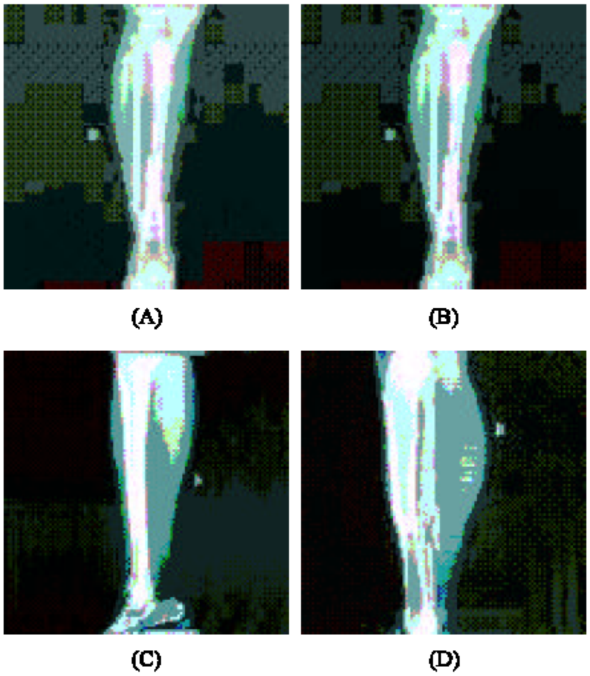


Fig. 5: (A). Input query image; (B, C, D) Retrieved similar images

block, fracture is the inference drawn by a medical expert and treatment which is the action taken by the medical expert for treating the fracture are stored in the database. The rule base has been constructed using the rules generated from the past data and experts advice.

RESULTS

Present research ISIS was developed using Java and Oracle 8i. A set of eighty fractured images was used for present study. Radiographic images collected were converted in to 115×115 pixels in JPEG format. Feature extraction was performed. The attributes that play a major role in computing the distance metric and retrieving similar images were computed and stored in the image database along with inference made by the radiological experts. We initially faced bottle necks in retrieving similar fracture images for a given input query image. It was noted that selection of the threshold value based on distance metric for matching and retrieving similar images played a major role in effective and error free retrieval of similar images. Through trial and error we were able to fix an appropriate threshold value for the distance metric and finally we were successful in achieving our goal, which consumed 15% of our development time. The initial proto-type was implemented with out indexing. Once we were successful in our goal indexing was performed based on feature values using k-d tree. It was noted that indexing improved the response time of our system for a given input query. For each new image similar images are retrieved, a unique image identifier is gene-rated and the attributes description, fracture and treatment are also updated in the image database. A subset of the data set that was used in carrying out present research is being shown in Fig. 3, 4 and 5.

CONCLUSIONS

Present research as discussed above can retrieve and integrate similar fractured images through content-based retrieval techniques and provides information about the type of fracture and the treatment given for similar cases. This aids the radiologist in decision-making. Moreover, this facility enables a better usage of digitalized information to aid in health procedures, as searching for similar medical cases to compare and make decisions. Indexing using k-d tree improves the response time of the system. We have carried out a comparative study between retrieval of similar images with indexing and without indexing has been done and it has been inferred that there is an improved performance when the experiment was performed using indexing. It is proposed to develop Content-based Image Retrieval (CBIR) systems for various medical applications using fuzzy set theory that would need to handle vagueness in medical images and uncertainty in decision-making.

REFERENCES

1. Agma, J.M., Traina, Natalia A. Rosa, Caetano Traina Jr, 2003. AIntegrating Images to Patient Electronic Medical Records thro-ugh Content-based Retrieval Techniques Proceedings of the 16th IEEE symposium on Computer Based Medical Systems (CBMS'2003), New York, pp: 163-168.
2. Joaquim, Cezar Felipe, Agma Traina, Caetano Traina Jr, 2003. ARet-rieval by Content of Medical Images Using Texture for Tissue Identification, Proceedings of the 16th IEEE Symposium on Computer Based Medical Systems (CBMS' 2003), New York, pp: 26-27.
3. Patric, G. Foschi, Deepak Kolippakkam, Huan Liu and Amit Mandvikar, 2002. A Feature Extraction for Image Mining, Proceedings of the International Workshop on Multimedia Information Systems (MIS 2002), Temple, Arizona, USA, pp: 103 B 109.
4. Rinie Egas, Nies Huijsmans, Michael Lew, Nice Sebe, 1999. A Adapting k-d Trees to Visual Retrieval, Proceedings of the 3rd International Conference, VISUAL'99, Amsterdam, Netherlands.

5. Melliyal Annamalai, Rajiv Chopra, Susan Mavris, Samuel DeFazio, 539. A Indexing Images in Oracle8i, Proceedings of the 2000 ACM SIGMOD International Conference on Mana-gement of Data, Texas, USA, pp: 539 B 547.
6. Myrian, R.B. Araujo, Caetano Traina Jr, Agma Traina, Josiane M.Bueno, Humberto L.Razente, 2002. A Extending Relational Databases to support Content Based Retrieval of Medical Images, Proceedings of the 15th IEEE Symposium on Computer Based MedicalSystems (CBMS' 2002), Maribor, Solvenia, pp: 303-308.
7. Manesh Kokare, B.N. Chatterji and P.K. Biswas, 2003. A Comparing of similarity Metrics for Texture Image Retrieval., TENCON 2003, Bangalore.
8. Maleq Khan, Qin Ding and William Perrizo, 2002. AK-Nearest Neighbor Classification on Spatial Data Streams using P-trees, PAKDD 2002, Taiwan.