ISSN:0975-9646

Amjan Shaik et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 2 (3) , 2011, 1336-1341

An Image Retrieval System Using Spatial Database: An Image Processing Context With Software Metrics

1. Amjan.Shaik, *CSE ,Ellenki College Of Engineering and Technology(ECET), Hyderabad, India.*
2. Shaik.Meeravali, *ECE ,DVR College Of Engineering and Technology(DVRCET),Hyderabad, India.*
3. G. Krishna Reddy, *ETM , G.Narayanamma Institute of Technology and Science For Women, Hyderabad. India.*
4. S.V.Achuta Rao , *CSE and IT , DJR Institute of Engineering and Technology (DJRIET), Vijayawada, India.*
5. Mohd Mukarram Uddin , *CSE, Muffakham Jah College of Engineering and Technology(MJCET), Bajara Hills, Hyderabad, India.*

Abstract

Image processing is any form of signal processing for which the input is an image, such as a photograph or video frame; the output of image processing may be either an image or, a set of characteristics or parameters related to the image.In this paper, symbolic images are represented in the form of symbolic image database (SID) which is an invariant to image transformations and this SID is used for advanced match retrieval of queried images. The conventional pair-wise spatial relationships (9DLT matrix) approach is not robust as that of image invariant transformations. For invariant image transformations, a new concept called advanced direction of reference is introduced which preserved by a set of triples and they are represented by using principal component analysis (PCA). The problem of using principal component transformation is that it produces identical principal component vectors (PCV)s even for different images which consumes additional memory. In our methodology a distinct and unique key is computed for each distinct triple. The mean and standard deviation of the set of keys are computed for a symbolic image and stored along with the total number of keys as the representatives of the same object image. The proposed advanced match retrieval design is based on a binary search technique and, thus, requires O(log n) search time in the worst case, where n is the total number of symbolic images in the symbolic image database. Images acquired through any modern sensors consist of variety of noises, resulting from stochastic variations and deterministic distortions or shading. In order to extract image data, smoothing algorithms should be applied initially to reduce noises before further analysis and processing. We empirically analyzed with our software tool and also got the favorable results.

Keywords: Symbolic image database, Image Retrieval, PCA, Software Metrics.

1.INTRODUCTION

How to perceive spatial relationships among objects in a picture is an important criterion to retrieve the pictures in a symbolic picture database. In 1991, Chang proposed an indexing structure, called the 9DLT matrix, to logically represent symbolic pictures by pair wise spatial relationships. The corresponding 9DLT matrix preserves the spatial relationships among objects in a symbolic picture. we demonstrated an advanced match retrieval of pictures in a symbolic picture database by the 9DLT matrix is also discussed. The proposed scheme is based upon a famous advanced direction of reference technique and principal component analysis. By the proposed scheme, it requires O(log *n*) search time in the worst case for the advanced match retrieval of symbolic pictures, where *n* is the total number of pictures in the data base. Retrieval of images from a symbolic image database (SID) is a challenging and motivating research issue today. A SID is a system in which a large amount of symbolic data and their related information are represented by both logical images and physical images. A logical image, also referred as symbolic image, can be regarded as an abstract of physical image, while the physical image (at pixel) is the real image itself. Various image processing and understanding techniques are used to identify the domain objects (components) and their locations in the image. Though this task is computationally expensive and difficult, it is performed only once at the time of image insertion. Moreover, this task may be carried out in a semi automated fashion or in an automated fashion depending upon the domain and complexity of images. An intermediate image called symbolic image is then obtained by associating each of the identified domain objects. For example, in an interior design application, various furniture's and decorative items in an image, constitute the domain objects [1]. Similarly an image of a human face may include: eyes, nose, mouth, ears, and jaw as the domain objects. Retrieval of images with a desired content from a symbolic image database (SID) is a challenging and motivating research issue today to effectively represent/retrieve an image in/from a SID, the attributes such as symbols/ icons and their relationships which are rich enough to describe the corresponding symbolic image are necessary. Thus, many researchers have highlighted the importance of perceiving spatial relationships existing among the components of an image for efficient representation/retrieval of symbolic images in/from a SID. In fact, the perception of spatial relationships preserves the reality being embedded in physical images besides making the system intelligent, fast, and flexible.

1.1 Special Image Retrieval Systems

Basically, there are two types of retrieval: similarity retrieval and Advanced match retrieval. In similarity retrieval, the task is to retrieve from the SID all those images that are similar to a given query image, while the advanced match retrieval retrieves only the image exactly identical (i.e., 100% similar) to a query image from the SID. In fact, advanced match retrieval is a special case of similarity retrieval and more precisely, is an image recognition problem finding applications in industrial automation, biomedicine, social security, architectural design, and many more robotics/computer vision applications, particularly when image registration takes place. There have been several attempts made to scatter the demands in the design of efficient, invariant, flexible, and intelligent image archival and retrieval systems based on the perception of spatial relationships. To make image retrieval, visualization, and traditional image database operations more flexible and faster, the data structure should be object oriented. The design of such object oriented data structures began with the discovery of 2D strings[2][3]. In this connection software metrics can play vital role.

1.2 Need for Software Metrics

Now days lot of software's are developed by the developers. Many of the software's are very big in code size. So generally to maintain the quality of the code, developers need to is tribute the code in small pieces or parts. But how to divide the software is also an important task as it can lead to various problem of inter module communication therefore this modularized code should also be checked for the quality. There are problems in removing the errors of non modularized code. Particularly in object oriented software development developer needs to use a lots of object oriented concepts which may introduced the inter dependency of the various units of the software e.g. Inheritance. A software metric is a measure of some property of a piece of software or its specifications. Therefore software metrics suite is needed [11] [12]. We are concentrating on the same issue and providing the software metrics for this modularized object oriented code.

1.3 Image Retrieval Techniques

Using the concept of 2D strings, in order to retrieve similar symbolic images from a SID, based on the longest common subsequence matching were proposed. Though these iconic image representation schemes offer many advantages, the linear string representation given to the spatial relations existing among the components takes nondeterministic polynomial time complexity during the process of string matching, in addition to being not invariant to image transformations, especially to rotation. In order to reduce the search time and to avoid string matching, hash-oriented methodologies for similarity retrieval based upon the variations of 2D string. However, hash function-based algorithms require O(m2) retrieval time in the worst case, where m is the number of iconic objects. Chang proposed[14] a symbolic indexing approach called nine directional lower triangular (9DLT) matrixes to encode symbolic images. A 9DLT matrix was used to preserve the spatial relationships existing among objects in an image. The pair-wise spatial relationships existing between iconic objects were preserved with the help of nine directional codes and were then represented in a 9DLT matrix[15]. The first principal component direction of the set of triples representing the 9DLT matrix of a symbolic image was computed and stored as the representative of the symbolic image in the SID. The first principal component direction of all symbolic images were stored in a sorted sequence, there by enabling the retrieval process to consume only O(log n) search time even in the worst case, with the help of the binary search technique, where n is the number of symbolic images stored in the SID. The method is not robust to take care of image transformations especially, in rotation.

In view of this, we have addressed the problem of advanced match retrieval of symbolic images by introducing the concept of advanced direction of reference. The spatial relationship was perceived with respect to the advanced direction of reference and preserved with the help of triples. The triples were then given a compact representation using principal component analysis (PCA). Due to the limitation of principal component transformation, there is a possibility of getting identical first principal component vectors (PCV)s for different symbolic images. Hence, two entirely different symbolic images may have the same PCV. Under such conflicting situations, one has to employ the second PCV for resolution and go up to the third PCV. If all the PCVs associated with two or more symbolic images are the same, then the conflict in discriminating such symbolic images could be resolved by storing the associated triplets themselves, which inevitably entails additional spatial memory. With this backdrop, an efficient spatial data structure, specifically for advanced match retrieval, invariant to image transformations, is suggested in this paper. In order to take care of image transformations, the concept of advanced direction of reference is used as a basis. The advanced direction of reference is conceptually aligned with the X-axis of the coordinate system and the spatial relationships existing among the objects present in the image are perceived relative to the advanced direction of reference and then are preserved by the use of triples. A distinct and unique key is computed for each distinct triple. The mean and standard deviation of the set of keys computed for a symbolic image are stored along with the total number of keys as the representatives of the corresponding image. A binary search technique-based scheme for exact match retrieval is used. The presented retrieval scheme requires O(log n) search time in the worst case, where n is the total number of symbolic images in the SID.

1.4 Scope

- creating symbolic image database (SID), which is invariant to image transformation
- Image relies on primitive features such as color, shape and texture.

Primitive features can be automatically extracted from the images themselves image features. In general, image features that were most promising were color, texture and shape. After the features have been identified, a step-wise procedure extracts the features and combines them using rules in the image. In the individual image each pixel corresponds to a particular intensity value. These sizes were considered after experimenting with various other sizes. If the size is too small, then texture features cannot be described. If it is too large, then the image is divided into blocks.

2.RESEARCH BACKGROUND

2.1 Image features

Generally the images can have one or more objects. For each object is segmented from image based on its features vectors like colour, texture and shape. Thus, for an image with n objects, there will be n colour and n texture/shape feature vectors being generated, for each object. Colour and texture/shape feature vectors are stored separately for each object. The saved feature values are then used to build the indexing tree for image objects. By indexing, the attributes of regions, such as sizes, locations and visual features, a wide variety of complex joint spatial and feature queries are efficiently computed.

Algorithm. Creation_of_SID

Input: Set of symbolic images
Output: SID, Symbolic Image Database
Method:
Step 1: For each symbolic image S to be archived in the SID do
For each rotated instance Rs of S do
(i) Apply TSR and obtain a set of quadruples Pq preserving TSR among the components present in Rs.
(ii) For each quadruple in Pq, compute a unique TSR key
(iii) Compute the vector D = (N,l,r) where N is the total number of TSR keys, l is the mean and r is the standard deviation of the TSR keys generated.
For end
Compute the representative vector Cs (for the symbolic image S) which is the centroid of all Ds computed for S.
For end
Step 2: Store the centroids obtained for all images in a sorted sequence.
Creation_of_SID ends[1].

Algorithm. Advanced match retrieval

Input: Q, a symbolic query image
SID, Symbolic Image Database
Output: Desired image
Method:
Step 1: Preserve TSR existing among the components of Q by the use of quadruples.
Step 2: For each quadruple, compute the corresponding TSR key
Step 3: Compute the vector Dq = (N,l,r)
Step 4: Employ the modified binary search technique to find out the two adjacent vectors Di and Di+1, such that Di 6 Dq 6 Di+1.
Step 5: Find out the distances, d1 and d2 of Dq to Di and Di+1 respectively.
Step 6: Retrieve the symbolic image corresponding to the index i if d1 < d2,
i + 1 otherwise:
Advanced match retrieval ends[1].

2.2 Color feature

Pixel

The picture elements that make up an image, similar to grains in a photograph or dots in a halftone. Each pixel can represent a number of different shades or colors, depending upon how much storage space are allocated for it.

8-bit image

A digital image that can include as many as 256 possible colors. In this kind of image, 8 bits are allocated for the storage of each pixel, allowing 2 to the power of 8 (or 256) colors to be represented.

24-bit image A digital image that can include approximately 16 million possible colors. In this kind of image, 24 bits are allocated for the storage of each pixel, allowing 2 to the power of 24 (or more than 16 million) colors to be represented.

Palette

The set of colors that appear in a particular digital image. Becomes part of a color look-up table.

Adaptive palette

Image-specific set of colors chosen to most closely represent those in the original source. Part of a custom color look-up table. Several methods for retrieving images on the basis of color similarity, Each image compute a color histogram, which shows the proportion of pixels of each color within the image. The color histogram for each image is then stores in the database. At each time, the user can either specify the desired proportion of each color (&75% olive green and 25% red, for example), or submit an example image from which a color histogram is calculated. Either way, the matching process then retrieves those, which a color histogram is calculated.

A color space is defined as a model for representing color in terms of intensity values. The color space models can be different as hardware-oriented and user-oriented. The hardware-oriented color spaces, including RGB, HLS, HCV, HSV and HSB, are based on the human percepts of colors, i.e., hue, saturation, and brightness. The color space can also be distinguished as uniform and non-uniform depending in differences in color space as perceived by humans.

2.3 RGB color model

The RGB color space is defined as a unit cube with Red, Green and Blue axes; hence a color in an RGB color space is represented by a vector with three Co-ordinates. When all three values are set to zero, the corresponding color is Black. When all three values are set to one, the corresponding color is white. The RGB model uses the Cartesian coordinate system as shown in Figure.1(a). Notice the diagonal from (0,0,0) black to (1,1,1) white which represents the grey-scale. Figure.1(b) is a view of the RGB color model looking down from "White" to origin. Commonly used image formats such as JPEG's, GIF's and BMP's always store and show colors in RGB color space. For current computer systems, the RGB color space is usually represented by a 24 bit true color system. In a 24 bit color system a color is specified by 3 integers: (red, green and blue), and these 3 integers range from 0 to 28-1 respectively, so giving 16,777,216 (224) colors in total.

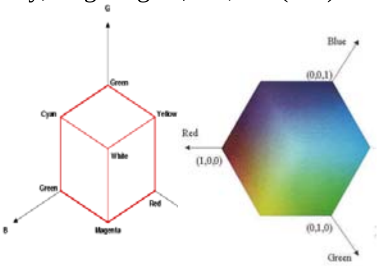

Figure 1 (a) RGB coordinates system

Figure 1 (b) RGB color model

2.4 Modules in Advanced Match Retrieval System

Base on indexing and CBIR , In advanced match image retrieval system has , mainly two steps to perform operations. They are image database creation (figure:2) and image retrieval from the storing database (figure:3). In this mainly they are working based on image features (color, texture and shape) and segmentation.

Database creation module

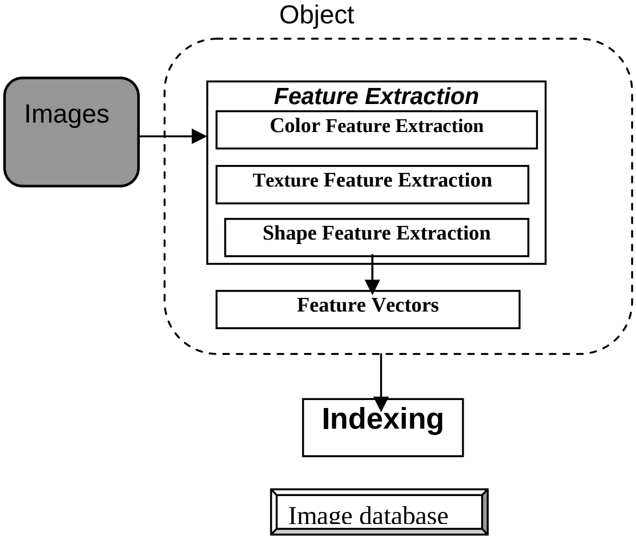

Figure 2: Image database creation by using indexing concept

Image Retrieval Module

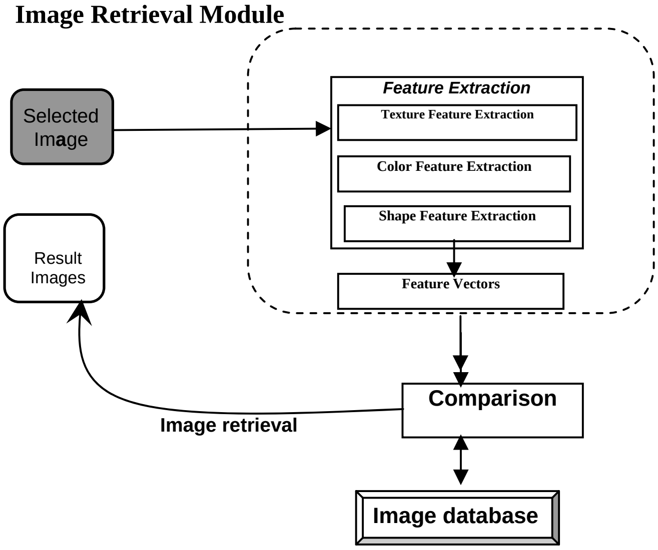

Figure 3: Image Retrieval using indexing concept

2.5 Spatial Image

Researchers have developed different techniques for spatial image, based on the locations of objects. These approaches compare symbolic images in which the regions or objects have been defined. Spatial image does not incorporate features such as color, texture, shape, and so... It is difficult to extend spatial image to include feature measures of the objects in two types of spatial images, i.e. relative and absolute. In a relative spatial image, the images are matched based upon the relative locations of symbols. For example, a relative spatial query may ask for images in which symbol A is to the left of symbol B. In absolute spatial image query, the images are matched based upon fixed positions in the images.

2.5.1 Spatial and feature image query

In the spatial and feature query system, regions feature and spatial attributes are first extracted from the images. The comparison of the images considers the similarities of the feature and spatial attributes of the regions as depicted in Figure:4. The overall match score between images is computed by summing the weighted distances between the best matching regions in terms of spatial locations, sizes and features. The relative spatial locations of regions in the target images are also compared to those in the query image to determine the image matches.

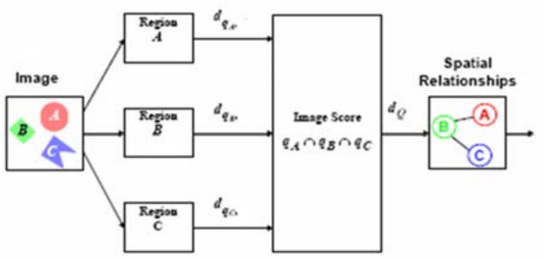

Figure 4: Object extraction and spatial composition of objects

2.5.2 Spatial Query

In computing matches between regions(figure:5), we consider the absolute spatial locations and sizes of the query (**Q**) and target (**T**) regions. The spatial attributes of each region are determined by the spatial centroids of the region (x; y) and the region width and height (w, h). The spatial distance between region centroids is measured using the Euclidean distance as follows:

$$d^{s}_{q,t} = [(x_q - x_t)^2 + (y_q - y_t)^2]^{\frac{1}{2}}.$$

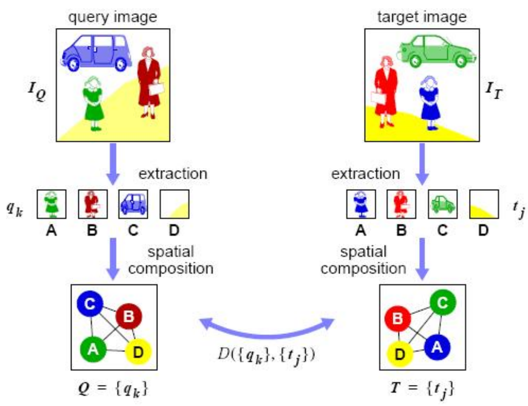

Figure 5: Example for object representation

2.5.3 Special spatial relations

A relative spatial query that specifies "nearness" of regions is specified by the user placing the regions within a single grid box and by choosing relative spatial matching. The overlap of regions requires the determination of nearness followed by the test for overlap of the MBRs by examining the region spatial attributes: (x; y) and (w; h), directly approximate rotation invariance. Approximate rotation invariance around the image center point by providing two projections of each image, one normal projection, and another based upon a rotation by 45 degrees with projection onto the rotated axis.

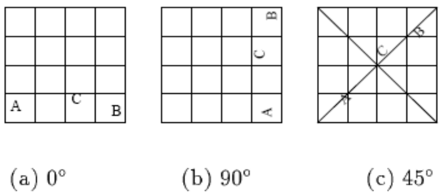

Figure 6: Rotation invariance

(a) 0^0 (A < C < B;ABC),

(b) (b) 90^0 (ABC;A < C < B),

(c) (c) 45^0 (A < C < B;ABC)

3. SYSTEM ARCHETECTURE

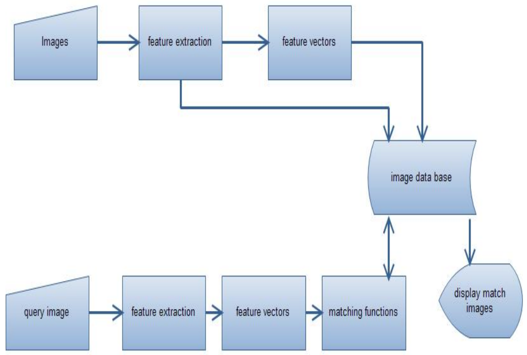

Figure 7: System Architecture

3.1 Workflow of Image Retrieval System

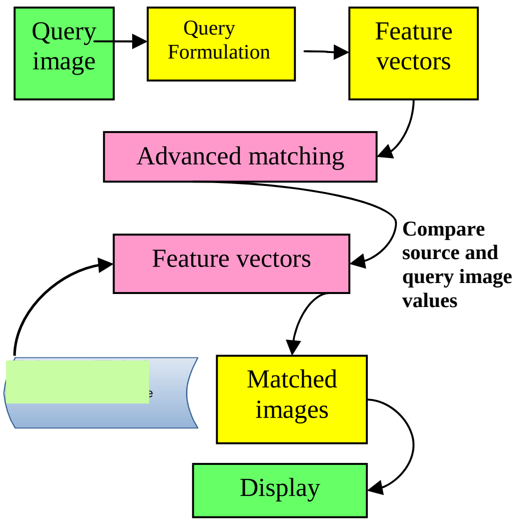

Figure 8: Workflow of Image Retrieval System

ACKNOWLEDGEMENTS

The authors thankful to the Auroras Scientific Technological and Research Academy and Ellenki College of Engineering and Technology-Research and Development Cell for collecting the data and preparing the paper.

CONCLUSION

In this paper, we made a flourishing attempt in exploring a model which overcomes the above mentioned shortcomings and best suits advanced match retrieval. The paper demonstrates a way of representing a symbolic image in SID invariant to image transformations through the perception of spatial relationships. The spatial relationship is perceived with respect to the advanced direction of reference and preserved in terms of triples, which are then mapped onto unique keys. The mean and standard deviation of the set of keys are computed and stored in the SID along with the total number of keys as the representative vector of the image in a sorted sequence. This representation not only makes the task of retrieval easier but also achieves a reduction in memory requirement at two levels: at the level of key computation and also at the level of computation of representative triples. The proposed model integrates the representation of an image with the retrieval of an image. Unlike other models, our model automatically takes care of additional information such as angles and is invariant to image transformations. In addition, it takes care of multiple instances of objects, which is considered to be a major problem in most of the existing methodologies. In retrieval stage, query image is also represented as a feature vector and the similarity between the query vector and stored feature vectors is computed. The similarly measure is used to determine the distance between the query image and stored images. After that images are ranked according to the distance and retrieved. It is concluded that the variation of complexity depends on the measure of DIT (Depth of Inheritance) and NOC (Number of Children) in the context of image retrieval, while build SID the measure of Depth of Inheritance (DIT) with respect to Number of children (NOC) place a central role, which is evidence from the fact that the complexity depends on the depth of inheritance (DIT) with respect to Number of children (NOC). Both DIT and NOC directly relate to the design of the class hierarchy. In this approach , Classes with high DIT values are associated with a higher number of defects.

REFERENCES

[1].P. Punitha , D.S. Guru," An invariant scheme for exact match retrieval of symbolic images: Triangular spatial relationship based approach" Pattern Recognition Letters 26 (2005) Page:893–907
[2]Bhatia, S.K., Sabharwal, C.L., 1994. A fast implementation of a perfect hash function for picture objects. Pattern Recognition 27 (3), 365–375.
[3]Chang, C.C., 1991. Spatial match retrieval of symbolic pictures. Inf. Sci. Eng. 7, 405–422.
[4] Lee, S.Y., Shan, M.K., Yang, W.P., 1989. Similarity retrieval of iconic image databases. Pattern Recognition 22 (6), 675–682.

[5].P. Punitha , D.S. Guru," An invariant scheme for exact match retrieval of symbolic images: Triangular spatial relationship based approach" Pattern Recognition Letters 26 (2005) Page:893–907

[6]V.N. Gudivada, "R-string: A Geometry-Based Representation for Efficient and Effective Retrieval of Images by Spatial Similarity," IEEE Trans. Knowledge and Data Eng., vol. 10, no. 3, pp. 504-512, May/June 1998.

[7] S.K. Chang, Q.Y. Shi, and C.W. Yan, "Iconic Indexing by 2D Strings," IEEE Trans. Pattern Analysis and Machine Intelligence, vol. 9, no. 5, pp. 413-428, Sept. 1987.

[8] S.K. Chang, E. Jungert, and Y. Li, "Representation and Retrieval of Symbolic Pictures Using Generalized 2D Strings," SPIE Proc. Conf. Visual Comm. and Image Processing, pp. 1360-1372, 1989.

[9] C.C. Chang and T.C. Wu, "Retrieving the Most Similar Symbolic Pictures from Pictorial Databases," Information Processing and Management, vol. 28, no. 5, pp. 581-588, 1992.

[10] P.W. Huang and Y.R. Jean, "Using 2D C+ Strings as Spatial Knowledge Representation for Image Database Systems," Pattern Recognition, vol. 27, no. 9, pp. 1249-1257, 1994.

[11] Amjan Shaik,Dr.C.R.K Reddy"Statiscal Analysis for Object Oritented Design Software Metrics", International Journal of Engineering Science and Technology(IJEST),Vol. 2(5), 2010, 1136-1142.

[12] Amjan Shaik, C. R. K. Reddy, Bala Manda, Prakashini. C, Deepthi. K" An Empirical Validation of Object Oriented Design Metrics in Object Oriented Systems" International Journal of Emerging Trends in Engineering and Applied Sciences (IJETEAS) 1 (2): 216-224 (ISSN: 2141-7016).

[13] X.M. Zhou, C.H. Ang, and T.W. Ling, "Image Retrieval Based on Object's Orientation Spatial Relationship," Pattern Recognition Letters, vol. 22, pp. 469-477, 2001.

[14] E.G.M. Petrakis and S.C. Orphanoudakis, "A Methodology for the Representation, Indexing and Retrieval of Images by Content," Image and Vision Computing, vol. 8, no. 11, pp. 504-521, 1993.

[15] E.G.M. Petrakis, "Design and Evaluation of Spatial Similarity Approaches for Image Retrieval," Image and Vision Computing, vol. 20, pp. 59-76, 2002.

[16] D.S. Guru and P. Punitha, "An Invariant Scheme for Exact Match Retrieval of Symbolic Images Based Upon Principal Component Analysis," J. Pattern Recognition Letters, vol. 25, no. 1, pp. 73-86, 2004.

[17] S.Y. Lee and F.J. Hsu, "Picture Algebra for Spatial Reasoning of Iconic Images Represented in 2D C String," Pattern Recognition Letters, vol. 12, no. 7, pp. 425-435, 1991.

[18] Y.I. Chang and H.Y. Ann, "A Note On Adaptive 2D-H Strings," Pattern Recognition Letters, vol. 20, pp. 15-20, 1999.

[19] S.Y. Lee and M.K. Shan, "Access Methods of Image Databases," Pattern Recognition and Artificial Intelligence, vol. 4, no. 1, pp. 27-44, 1990.

[20] C.L. Sabharwal and S.K. Bhatia, "Image Databases and Near Perfect Hash Table," Pattern Recognition, vol. 30, no. 11, pp. 1867- 1876, 1997.

[21] C.C. Chang, "Spatial Match Retrieval of Symbolic Pictures," Information Science and Eng., vol. 7, pp. 405-422, 1991.

[22] C.C. Chang and T.C. Wu, "An Exact Match Retrieval Scheme Based upon Principal Component Analysis," Pattern Recognition Letters, vol. 16, pp. 465-470, 1995.

ABOUT THE AUTHORS

Amjan Shaik is working as a Professor and Head, Department of Computer Science and Engineering at Ellenki College of Engineering and Technology (ECET), Hyderabad, India. He has received M.Tech. (Computer Science and Technology) from Andhra University. Presently, he is a Research Scholar of JNTUH, Hyderabad, India. He has been published and presented 34 Research and Technical papers in International Journals , International Conferences and National Conferences. His main research interests are Software Engineering, Software Metrics, Image Processing, Software Quality and Object Oriented Design.

Shaik Meeravali is working as an Associate Professor, Department of Electronics & Communication Engineering at D.V.R. College of Engineering and Technology, Hyderabad, India.. He has received B.Tech. (ECE) from Gulbarga University, Gulbarga and M.Tech.(Digital Systems and Computer Electronics) from JNTU, Hyderabad, India. He has presented number of technical papers in International, National Conferences and International Journals.His research interests are Image Processing, Information Security and Wireless Networks.

G. Krishna Reddy is working as an Associate Professor, ETM Department, G.Narayanamma Institute of Technology and Science For women, Shaikpet, Hyderabad.India.. He has received B.Tech. (ECE) from Gulbarga University,Gulbarga and M.Tech.(Digital Systems and Computer Electronics) from JNTU, Hyderabad, India. He has presented number of technical papers in International, National Conferences and International Journals. His research interests are Image Processing, Speech Recognition, Information Security and Digital Signal Processing .

S.V. Achuta Rao is working as a Professor and Head, Department of CSE and IT at DJR Institute of Engineering and Technology (DJRIET), Vijayawada, India. He has received M.Tech. (Computer Science and Engineering) from JNTU, Kakinada, India. Presently, he is a Research Scholar of Rayalaseema University (RU), Kurnool, India. He has been published and presented good number of Research and technical papers in International and National Conferences. His main research interests are Data Mining, Networking, Image Processing, Software Engineering and Software Metrics.

Mohd Mukarram Uddin is working as an Asst Professor, Department of Computer Science and Engineering at Muffakham Jah College of Engineering and Technology(MJCET), Bajara Hills, Hyderabad, India. He has received M.Tech (Information Technology) from ASTRA, Affiliated to JNTUH Hyderabad. He has presented number of technical papers in International and National Conferences. His research interests are Mobile Computing, Image Processing, Information Security and Software Engineering.