Improving Web Search for Images Using Social Bookmarking

M. Praveen Kumar

M-Tech Research Scholar

Avr and Svr Engineering College

Kurnool, Andhrapradesh – India

Abstract: Image Retrieval system is an effective and efficient tool for managing large image databases. A content based image retrieval system allows the user to present a query image in order to retrieve images stored in the database according to their similarity to the query image. In this paper content based image retrieval method is used on digital image data set. The main objective of this paper is to evaluate the retrieval system based on Hybrid features. The texture features are extracted by using pyramidal wavelet transform and the shape features are extracted by using Fourier descriptor. And the hybrid technique is the combination of both texture and shape. The major advantage of such an approach is that little human intervention is required. It is ascertained that the performance is superior when the image retrieval based on the Hybrid features, and better results than primitive set.

Keywords: component; formatting; style; styling; insert (Minimum 5 to 8 key words must be phrases)

I. INTRODUCTION

Image Retrieval aims to provide an effective and efficient tool for managing are image databases. There is a significant amount of increase in the use of medical images in clinical medicine and disease research. Image retrieval (IR) is one of the most exciting and fastest growing research areas in the field of medical imaging [2]. The goal of CBIR is to retrieve images from a database that are similar to an image placed as a query. In CBIR, for each image in the database, features are extracted and compared to the features of the query image. A CBIR method typically converts an image into a feature vector representation and matches with the images in the database to find out the most similar images. A comparative study has given on multiple databases [3,6,10] it is concluded that performance of DRD image is less compared to other database images
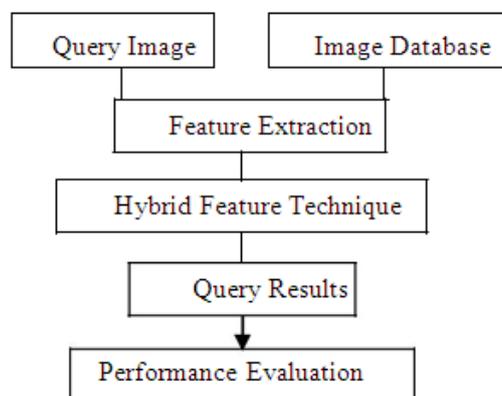

Figure 1. Block diagram of Image Retrieval System

The goal of CBIR is to retrieve images from a database that are similar to an image placed as a query. In CBIR, for each image in the database, features are extracted and compared to the features of the query image. A CBIR method typically converts an image into a feature vector representation and matches with the images in the database to find out the most similar images. In various studies different databases have been used to compare the study. The similarity between features was to be

calculated using algorithms used by well known CBIR systems such as IBM's QBIC[28]. For each specific feature there is a specific algorithm for extraction and another for matching.

In this paper evaluation of retrieval system based on hybrid features is carried out. The major advantage of this approach is that little human intervention is required. The databases used here is digital data of 200 images with multiple contexts.

Figure 1 show the basic block diagram used in this work. Shape and texture features are extracted for both query image and images in the database. The distance (ie. similarities) between the features vectors of the query image and database are then computed and ranked. The database images that have highest similarity to the query image are retrieved. Then the performance analysis is carried out using precision and recall.

II. FEATURE EXTRACTION

Feature Extraction is the process of creating a representation, or a transformation from the original data. The images have the primitive features like color, texture, shape, edge, shadows, temporal details etc. The features that were most promising were color, texture and shape/edge. The reasons are color can occur in limited range of set. Hence the picture elements can be compared to these spectra. Texture is defined as a neighborhood feature as a region or a block. The variation of each pixel with respect to its neighboring pixels defines texture. Hence the textural details of similar regions can be compared with a texture template. shape/edge is simply a large change in frequency. The three feature descriptors mainly used most frequently during feature extraction are color, texture and shape

2.1 Texture Feature Extraction

Texture is that innate property of all surfaces that describes visual patterns, each having properties of homogeneity. It contains important information about the structural arrangement of the surface, such as; clouds, leaves, bricks, fabric, etc. It also describes the relationship of the surface to the surrounding environment. In short, it is a feature that describes the distinctive physical composition of a surface.The most popular statistical representations of texture are: Co-occurrence Matrix, Tamura Texture and Wavelet Transform.

2.2 Feature Extraction Methods for Texture Analysis

In Texture Analysis the main feature extraction methods are spectral approach, structural (or syntactic) approach and statistical approach. The spectral approach to texture analysis deals with images in the frequency domain. Therefore, this approach requires Fourier transform to be carried out on the original images to acquire their corresponding representations in the frequency space. Fourier transform based methods usually perform well on textures showing strong periodicity, however their performance deteriorates as the periodicity of textures weakness. The classification of all methods among the two other approaches: structural and statistical.

The statistical point of view, an image is a complicated pattern on which statistics can be obtained to characterize these patterns. The techniques used within the family of statistical approaches make use of the intensity values of each pixel in an image, and apply various statistical formulae to the pixels in order to calculate feature descriptors. Texture feature descriptors, extracted through the use of statistical methods, can be classified into two categories according to the order of the statistical function that is utilized: First-Order Texture Features and Second Order Texture Features. First Order Texture Features are extracted exclusively from the information provided by the intensity histograms, thus yield no information about the locations of the pixels. Another term used for First-Order Texture Features is Grey Level Distribution Moments.

In contrast, Second-Order Texture Features take the specific position of a pixel relative to another into account. The most popularly used of second-order methods is the Spatial Grey Level Dependency Matrix (SGLDM) method. The method roughly consists of constructing matrices by counting the number of occurrences of pixel pairs of given intensities at a given displacement.

2.3 Shape Feature Extraction

Shape may be defined as the characteristic surface configuration of an object; an outline or contour. It permits an object to be distinguished from its surroundings by its outline. Shape representations are two categories boundary-based and region-based. Boundary-based mathematical representations are Polygonal Models, boundary partitioning, Fourier Descriptors, Splines, higher order constructs, Curvature Models. Region-based mathematical representations are Super quadrics, Fourier Descriptors, Implicit Polynomials, Blum's skeletons.The most successful representations for shape categories are Fourier Descriptor and Moment Invariants. The Fourier Descriptor is used to boundary as the shape feature and the Moment invariants is used to region-based shape feature.

Edge detection is used to detect sharp changes in image brightness is to capture important events and changes in properties of the world. It can be shown that under rather general assumptions for an image formation model, discontinuities in image brightness are likely to correspond to discontinuities in depth, discontinuities in surface orientation, changes in material properties and variations in scene illumination. Canny considered the mathematical problem of deriving an optimal smoothing filter given the criteria of detection, localization and minimizing multiple responses to a single edge. This method showed that the optimal filter given these assumptions is a sum of four exponential terms. It also showed that this filter can be well approximated by first-order derivatives of Gaussians. This paper emphasizes on the image retrieval based on the extraction of features of texture and shape of a specific image database.

2.4 Texture Feature Extraction

Texture is that innate property of all surfaces that describes visual patterns, each having properties of homogeneity. It contains important information about the structural arrangement of the surface, such as; clouds, leaves, bricks, fabric, etc. It also describes the relationship of the surface to the surrounding environment. In short, it is a feature that describes the distinctive physical composition of a surface. The most popular statistical representations of texture are: Co-occurrence Matrix, Tamura Texture and Wavelet Transform.

2.5 Feature Extraction Methods for Texture Analysis

In Texture Analysis the main feature extraction methods are spectral approach, structural (or syntactic) approach and statistical approach. The spectral approach to texture analysis deals with images in the frequency domain. Therefore, this approach requires Fourier transform to be carried out on the original images to acquire their corresponding representations in the frequency space. Fourier transform based methods usually perform well on textures showing strong periodicity, however their performance deteriorates as the periodicity of textures weakness. The classification of all methods among the two other approaches: structural and statistical.

The statistical point of view, an image is a complicated pattern on which statistics can be obtained to characterize these patterns. The techniques used within the family of statistical approaches make use of the intensity values of each pixel in an image, and apply various statistical formulae to the pixels in order to calculate feature descriptors. Texture feature descriptors, extracted through the use of statistical methods, can be classified into two categories according to the order of the statistical function that is utilized: First-Order Texture Features and Second Order Texture Features. First Order Texture Features are extracted exclusively from the information provided by the intensity histograms, thus yield no information about the locations of the pixels. Another term used for First-Order Texture Features is Grey Level Distribution Moments.

In contrast, Second-Order Texture Features take the specific position of a pixel relative to another into account. The most popularly used of second-order methods is the Spatial Grey Level Dependency Matrix (SGLDM) method. The method roughly consists of constructing matrices by counting the number of occurrences of pixel pairs of given intensities at a given displacement.

Shape may be defined as the characteristic surface configuration of an object; an outline or contour. It permits an object to be distinguished from its surroundings by its outline. Shape representations are two categories Boundary-based and Region-based.

Boundary-based shape representation only uses the outer boundary of the shape. This is done by describing the considered region using its external characteristics; i.e., the pixels along the object boundary. Region-based shape representation uses the entire shape region by describing the considered region using its internal characteristics; i.e., the pixels contained in that region. Boundary-based mathematical representations are Polygonal Models, boundary partitioning, Fourier Descriptors, Splines, higher order constructs, Curvature Models. Region-based mathematical representations are Super quadrics, Fourier Descriptors, Implicit Polynomials, Blum's skeletons. The most successful representations for shape categories are Fourier Descriptor and Moment Invariants. The Fourier Descriptor is used to boundary as the shape feature and the Moment invariants is used to region-based shape feature.

Edge detection is a terminology in image processing and computer vision, particularly in the areas of feature detection and feature extraction . The purpose of detecting sharp changes in image brightness is to capture important events and changes in properties of the world. It can be shown that under rather general assumptions for an image formation model, discontinuities in image brightness are likely to correspond to discontinuities in depth, discontinuities in surface orientation, changes in material properties and variations in scene illumination. The most successful representations for shape categories are Fourier Descriptor and Moment Invariants: The main idea of Fourier Descriptor is to use the Fourier transformed boundary as the shape feature. The main idea of Moment invariants is to use region-based moments, which are invariant to transformations as the shape feature.

Canny Edge detection is an optimal smoothing filter given the criteria of detection, localization and minimizing multiple responses to a single edge. This method showed that the optimal filter given these assumptions is a sum of four exponential terms. It also showed that this filter can be well approximated by first-order derivatives of Gaussians. Canny also introduced the notion of non-maximum suppression, which means that given the pre-smoothing filters, edge points are defined as points where the gradient magnitude assumes a local maximum in the gradient direction.

Sobel edge detection operations are performed on the data and the processed data is sent back to the computer. The transfer of data is done using parallel port interface operating in bidirectional mode. All the digital logic implemented and verified on the field programmable gate array kit was described using the Verilog® Hardware Description Language and the target was Xilinx Spartan 3 Family device XC3S400.

For estimating image gradients from the input image or a smoothed version of it, different gradient operators can be applied. The simplest approach is to use central differences:

$$L_x(x,y) = -1/2 \cdot L(x-1,y) + 0 \cdot L(x,y) + 1/2 \cdot L(x+1,y).$$

$$L_y(x,y) = -1/2 \cdot L(x,y-1) + 0 \cdot L(x,y) + 1/2 \cdot L(x,y+1).$$

corresponding to the application of the following filter masks to the image data:

$$L_x = \begin{bmatrix} -1/2 & 0 & 1/2 \end{bmatrix} * L \quad \text{and} \quad L_y = \begin{bmatrix} +1/2 \\ 0 \\ -1/2 \end{bmatrix} * L$$

The well-known and earlier Sobel operator is based on the following filters:

Given such estimates of first- order derivatives, the gradient magnitude is then computed as:

$$|\nabla L| = \sqrt{L_x^2 + L_y^2}$$

while the gradient orientation can be estimated as

$$\theta = \operatorname{atan2}(L_y, L_x)$$

2.6 Thresholding and linking

Once this project computed a measure of edge strength (typically the gradient magnitude), the next stage is to apply a threshold, to decide whether edges are present or not at an image point. The lower the threshold, the more edges will be detected, and the result will be increasingly susceptible to noise, and also to picking out irrelevant features from the image. Conversely a high threshold may miss subtle edges, or result in fragmented edges.

If the edge thresholding is applied to just the gradient magnitude image, the resulting edges will in general be thick and some type of edge thinning post-processing is necessary. For edges detected with non-maximum suppression however, the edge curves are thin by definition and the edge pixels can be linked into edge polygon by an edge linking (edge tracking) procedure. On a discrete grid, the non-maximum suppression stage can be implemented by estimating the gradient direction using first-order derivatives, then rounding off the gradient direction to multiples of 45 degrees, and finally comparing the values of the gradient magnitude in the estimated gradient direction.

A commonly used approach to handle the problem of appropriate thresholds for thresholding is by using thresholding with hysteresis. This method uses multiple thresholds to find edges and begin by using the upper threshold to find the start of an edge. Once it has a start point, it then trace the path of the edge through the image pixel by pixel, marking an edge whenever it has above the lower threshold. It stops marking our edge only when the value falls below our lower threshold. This approach makes the assumption that edges are likely to be in continuous curves, and allows us to follow a faint section of an edge it has previously seen, without meaning that every noisy pixel in the image is marked down as an edge. Still, however, this project has the problem of choosing appropriate thresholding parameters, and suitable thresholding values may vary over the image.

2.7 Second-Order Approaches to Edge Detection

Some edge-detection operators are instead based upon second-order derivatives of the intensity. This essentially captures the rate of change in the intensity gradient. Thus, in the ideal continuous case, detection of zero-crossings in the second derivative captures local maxima in the gradient. The early Marr-Hildrenth operator is based on the detection of zero-crossings of the Laplacian operator applied to a Gaussian-smoothed image. It can be shown, however, that this operator will also return false edges corresponding to local minima of the gradient magnitude. Moreover, this operator will give poor localization at curved edges. Hence, this operator is today mainly of historical interest.

2.8 Differential Edge Detection

A more refined second-order edge detection approach, which also automatically gives edges with sub-pixel accuracy, is by using the following differential approach of detecting zero-crossings of the second-order directional derivative in the gradient direction: Following the differential geometric way of expressing the requirement of non-maximum suppression proposed by Lindeberg, let us introduce at every image point a local coordinate system (u,v), with the v-direction parallel to the gradient direction. Assuming that the image has been presmoothed by Gaussian smoothing and a scale-space representation L(x,y;t) at scale t has been computed, this project can require that the gradient magnitude of the scale-space representation, which is equal to the first-order directional derivative in the v-direction Lv, should have its first order directional derivative in the v-direction equal to zero

$$\partial_v(L_v) = 0$$

while the second-order directional derivative in the v-direction of Lv should be negative, i.e.,

$$\partial_{vv}(L_v) \leq 0$$

written out as an explicit expression in terms of local partial derivatives Lx, Ly ... Lyyy, this edge definition can be expressed as the zero-crossing curves of the differential invariant

$$L_v^2 L_{vv} = L_x^2 L_{xx} + 2 L_x L_y L_{xy} + L_y^2 L_{yy} = 0,$$

that satisfy a sign-condition on the following differential invariant

$$L_v^3 L_{vvv} = L_x^3 L_{xxx} + 3 L_x^2 L_y L_{xxy} + 3 L_x L_y^2 L_{xyy} + L_y^3 L_{yyy} \leq 0$$

where Lx, Ly ... Lyyy denote partial derivatives computed from a scale-space representation L obtained by smoothing the original image with a Gaussian kernel. In this way, the edges will be automatically obtained as continuous curves with subpixel accuracy. Hysteresis thresholding can also be applied to these differential and subpixel edge segments. In practice, first-order derivative approximations can be computed by central differences as described above, while second-order derivatives can be computed from the scale-space representation L according to: Higher-order derivatives for the third-order sign condition can be obtained in an analogous fashion.

A recent development in edge detection techniques takes a frequency domain approach to finding edge locations. Phase congruency (also known as phase coherence) methods attempt to find locations in an image where all sinusoids in the frequency domain are in phase. These locations will generally correspond to the location of a perceived edge, regardless of whether the edge is represented by a large change in intensity in the spatial domain. A key benefit of this technique is that it responds strongly to Mach bands, and avoids false positives typically found around roof edges. A roof edge is a discontinuity in the first order derivative of a grey-level profile.

III. IMAGE RETRIEVAL

Image retrieval algorithms based on Image Primitive features like color, texture and shape. Combinational features are specified to give good performance. On each feature more efficient algorithms are used to retrieve the information from data set.

3.1 Image Retrieval Based on Texture

3.1.1. Pyramid-Structured Wavelet Transform

It is suitable for signals consisting of components with information concentrated in lower frequency channels. Due to the innate image properties that allows for most information to exist in lower sub-bands, the pyramid-structured wavelet transform is highly sufficient.

Using the pyramid-structured wavelet transform, the texture image is decomposed into four sub images, in low-low, low-high, high-low and high-high sub-bands. At this point, the energy level of each sub-band is calculated .This is first level decomposition. Using the low-low sub-band for further decomposition, we reached fifth level decomposition, for our project. The reason for this is the basic assumption that the energy of an image is concentrated in the low-low band. For this reason the wavelet function used is the Daubechies wavelet.

For this reason, it is mostly suitable for signals consisting of components with information concentrated in lower frequency channels. Due to the innate image properties that allows for most information to exist in lower sub-bands, the pyramid-structured wavelet transform is highly sufficient.

3.1.2. Energy Level

Energy Level Algorithm:

» Decompose the image into four sub-images

Calculate the energy of all decomposed images at the same scale

where M and N are the dimensions of the image, and X is the intensity of the pixel located at row i and column j in the image map

» Repeat from step 1 for the low-low sub-band image, until index ind is equal to 5. Increment ind.

Using the above algorithm, the energy levels of the sub-bands were calculated, and further decomposition of the low-low sub-band image. This is repeated five times, to reach fifth level decomposition. These energy level values are stored to be used in the Euclidean distance algorithm.

3.1.3. Euclidean Distance

Euclidean Distance Algorithm:

- Decompose query image.

- Get the energies of the first dominant k channels.

- For image i in the database obtain the k energies.

Calculate the Euclidean distance between the two sets of energies Using the above algorithm, the query image is searched for in the image database. The Euclidean distance is calculated between the query image and every image in the database. This process is repeated until all the images in the database have been compared with the query image. Upon completion of the Euclidean distance algorithm, we have an array of Euclidean distances, which is then sorted. The five topmost images are then displayed as a result of the texture search.

IV. RESULTS AND ANALYSIS

The image database has used to retrieve the relevant images based on query image. The test image database contains 200 images on various categories. Image retrieval has performed on hybrid feature texture and shape. The results of Image Retrieval are shown below:
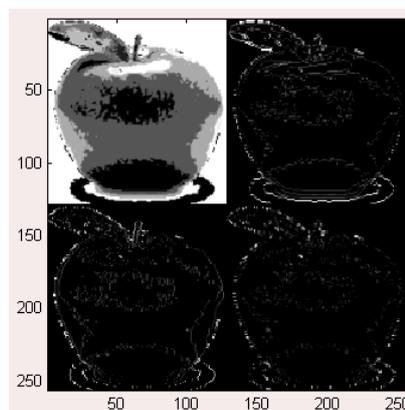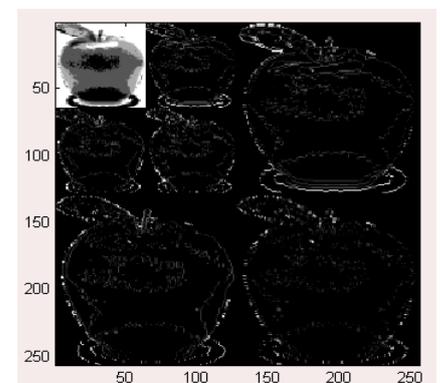

Figure 2: Query Image of Apple   *Figure 3: First Decomposition Figure*
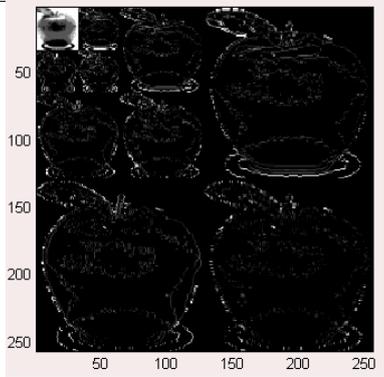
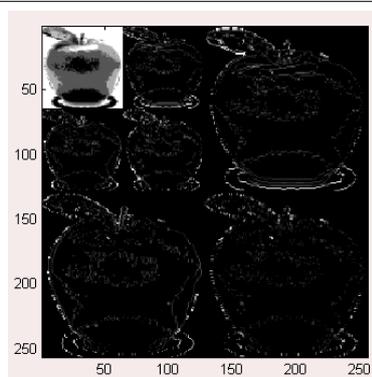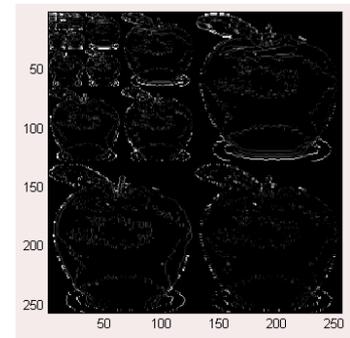*Figure 4: Second Decomposition*   *Figure 5: Third Decomposition*   *Figure 6: Fourth Decomposition*

Figure 7: Fifth Decomposition of Apple Image

From the figure, the application called decomposition uses the wavelet decomposition technique to decompose the query image that passes through the five levels of decomposition, this wavelet decomposition technique shows that the texture features of a selected image. After decomposition of an image the user will get the desired images, which has same texture and shape features, for this purpose the user using wavelet decomposition technique.

V. PERFORMANCE ANALYSIS

Based on commonly used performance measures in information retrieval, two statistical measures were computed to assess system performance namely Recall and Precision. For good retrieval system ideal values for recall is 1 and precision is of low value.

5.1 Precision

Precision gives the accuracy of the retrieval system. Precision is the basic measures used in evaluating the effectiveness of an information retrieval system. Precision consists of the proportion of relevant images that are retrieved.

Precision = (Number of Relevant Images Retrieved) / (Total Retrieved Images)

5.2 Recall

Recall consists of the proportion of target images that have been retrieved among all the relevant images in the database. Recall gives the measurement in which how fast the retrieval system works. It also measures how well the CBIR system finds all the relevant images in a search for a query image Recall = (Number of Relevant Images Retrieved) / (Total Number of Relevant Images).

It gives the average recall for each class of image. Recall 1 means all the relevant image in the database are retrieved. From the table it can be observed that almost all the relevant images are retrieved from the database when the first seven images are retrieved.

VI. CONCLUSION AND FUTURE ENHANCEMENTS

This paper elucidates the potentials of extraction of features of the image using texture and shape for retrieving the images from the specific image databases. The images are retrieved from the given database of images by giving the query image using texture and shape features. These results are based on various digital images of dataset. The performance of the image retrieval was assessed using the parameters recall rate and precision. It was ascertained that the recall rate and precision are high when the image retrieval was based on the Hybrid Features texture and shape than primitive features alone. This work can be extended further on huge data bases for retrieving relevant image. This can be extended for pixel clustering to obtain objects using different combination of weight for color and texture and shape features. It can maximize the performance by choosing the best combination between these weights

References

1. "Adaptation in the Quincunx wavelet filter bank with in image denoising,"in Proc. of the 2004 Int. TICSP Workshop on Spectral Methods and Multirate Signal Processing, SMMSP 2004, pp. 245- 252., 2004.

2. Aigrain, P et al (1996) "Content-based representation and retrieval of visual media – a state-of-the-art review" Multimedia Tools and Applications

3. Aliaa.A.A Yousif, A.A Darwish and R.A. Mohammad, "Content based medical image retrieval based on pyramid structure wavelet" IJCSNS International Journal of Computer Science and Network Security, Vol.3. March 2010.

4. Alsuth, P et al (1998) "On video retrieval: content analysis by Image Miner" in Storage and Retrieval for Image and Video Databases

5. Androutsas, D et al (1998) "Image retrieval using directional detail histograms" in Storage and Retrieval for Image and Video Databases

6. Bjarnestam, A (1998) "Description of an image retrieval system", presented at The Challenge of Image Retrieval

7. Carson, C., Thomas, M., Belongie, S., Hellerstein, J. M., Malik, J., (1999), "Blobworld: A system for region-based image indexing and retrieval," Third International Conference on Visual Information Systems, Springer.

8. Castelli, V. and Bergman, L. D., (2002), "Image Databases: Search and Retrieval of Digital Imagery", John Wiley & Sons, Inc.

9. Craig Nevill-Manning (Google) and Tim Mayer (Fast). Anatomy of a Search Engine, Search Engine Strategies Conference, 2002

10. Daubechies, I., (1992), "Ten Lectures on Wavelets," Capital City Press, Montpelier, Vermon.t

11. Drimbarean A. and Whelan P.F. (2001) Experiments in color texture analysis, Pattern Recognition Letters, 22: 1161-1167.

AUTHOR(S) PROFILE

M PRAVEEN KUMAR Completed M-Tech in computer science stream from AVR&SVR Engineering College, Nannur, Kurnool Affiliated to JNTU Anantapur, Andhra Pradesh, India & completed B.Tech in Computer Science and engineering from Alfa college of Engineering & technology, Allagadda, Kurnool.