IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Volume-4, Special Issue-7, Dec 2016 ISSN: 2347-2693

Indexing of Voluminous Data Using K-D Tree with Reference to CBIR

Jayashree Das and *Minakshi Gogoi*
Department of Computer Science and Engineering
Girijananda Choudhury Institute of management and Technology, Guwahati, India
minakshi_cse@gimt-guwahati.ac.in

Available online at: www.ijcseonline.org

***Abstract*— This paper proposes a fast and efficient indexing technique that can be used in an image indexing and retrieval system for voluminous image data. The proposed technique is based on K-d tree which uses multi-dimensional features. At first the colour feature of a set of images are extracted. Then an index tree is generated with K-d tree index based on these colour features. After indexing is done the efficiency of the method is tested against search time for the collected dataset. The validation of the method is also tested with and without indices for the said dataset.**

***Keywords*— Indexing; k-d tree; multi-dimensional feature; colour moment; haar wavelet.**

I. INTRODUCTION

With the increase in database volume, fast indexing and retrieval techniques has been considered as demanding for the enhanced network and multimedia technologies. Earlier, managing large spatial databases was used in geosciences and computer aided design (CAD). Later, the newfound applications in computer vision and robotics, computer visualization, geographical information processing, automated mapping and facilities management etc has increased the need for fast and efficient indexing and retrieval of spatial data. Image indexing is nothing but a way of arranging the particular image data in some ordered fashion so that we can retrieve them easily. To fulfil the requirement of indexing of images many techniques have evolved in the last few decades. One way is the traditional image database indexing and retrieval approach which is text based. Here the image data is fully converted into an electronic presentation [1]. But with the increase in popularity of the internet and enhancement in multimedia technologies, this approach is not favourable due to some factors such as lower quality text and higher cost. Difficulties of traditional indexing approach has led to raise the interest level in enquiring and developing the techniques for retrieving images automatically by using content features such as colour, shape and texture etc.

The different existing techniques that can be used for indexing of spatial image data are K-d tree, MapReduce, R+-tree, score-based etc. A K-d tree (short for k-dimensional tree) is a space-partitioning data structure for organizing points in a k-dimensional space. K-d trees are a useful data structure for several applications, such as searches involving a multidimensional search key (e.g. range searches and nearest neighbour searches). MapReduce is a programming model and an associated implementation for processing and generating large data sets with a parallel, distributed algorithm on a cluster.

There are some issues or challenges that are faced while developing an efficient algorithm for image indexing. Some of them are- Scalability, Robustness, platform independence and miss-categorization.

The concept of scalability is desirable in technology settings. Robustness is the ability of an indexing system to cope with errors during execution or processing. Platform independence in this context means that the indexing system should be independent of the specific technological platform used to implement it. Textual information about images can be easily searched using existing technology, but this requires humans to manually describe each image in the database. This can be impractical for voluminous databases or for images that are generated automatically. Also, sometimes it is not possible to select images that use different synonyms in their descriptions.

In this work, we have concentrated on indexing of voluminous data. The method exploits a K-d tree based indexing technique with respect to two feature extraction techniques based on colour feature and texture feature. For colour features colour moment is used and for texture features haar wavelet is used. 3D tree construction method is provided based on the extracted feature. The performance of the method was established using benchmark dataset and result has been found satisfactory. In the next section, we are describing the prior related work.

II. PRIOR RELATED WORK

From the related studies performed [2], it is found that many researches has provided indexing mechanism of different types of data or work. In the Table I, a brief description of the same is provided.

TABLE I. RELATED STUDIES ON DIFFERENT INDEXING APPROACHES

Approach/ Method used	Authors/Description
Hilbert	Lawder, J. K., and King, P. J. H., had used Hilbert

Space-Filling Curve [3]	curve as it has superior clustering properties while comparing it with Z-order curve.
K-d tree with Feature Level Fusion [4]	Jayaraman, U., Prakash, S., and Gupta, P., had first normalized a multi-dimensional feature vector of each trait and then, they projected it to a lower dimensional feature space. Then the reduced dimensional feature vectors are fused at feature level and the fused feature vectors are used to index the database by forming K-d tree.
Graph based Approach [5]	Gogoi, M., and Bhattacharya, D. K., had concentrated on a graph-theoretic approach to analyze the process of fingerprint comparison to give a feature space representation of minutiae and to produce a lower bound on the number of detectably distinct fingerprints. Their method provided a graph based index generation mechanism of fingerprint biometric data.
Pyramid Technique [6]	Li, Jia and Wang, Cheng had presented a new approach for fast indexing the hyper-spectral data based on the Pyramid Technique, and further propose new algorithms to process k Nearest Neighbour queries and spectral partial feature queries.
Feature extraction using Colour moments [7]	Tian, D. P. had extracted image features by using techniques such as colour moments, colour histogram, colour coherence vector etc prior to indexing.

Basically there are two ways to index an image for retrieving –

1. The traditional approach which is text based and
2. The modern approach that is based on image content.

Now-a-days importance is given on different indexing techniques which use image features such as colour, texture, shape etc for indexing. But for indexing spectral data, the main thing that we need to deal with is dimesionality. Spectral data such as a hyperspectral image is multidimensional. Li, Jia and Wang, Cheng in their research used the pyramid technique which is an effective high-dimensional data mapping method to map high-dimensional space point into one-dimensional space to exploit one-dimensional indexing structure such as the B+-tree [6].

Again, multi-model data means data with multiple features can be indexed first by extracting features and then applying technique such as K d-tree [4]. The technique they used is based on K-d tree with feature level fusion which uses the multi-dimensional feature vector. Here a multi-dimensional feature vector of each trait is first normalized and then, it is projected to a lower dimensional feature space. Then the reduced dimensional feature vectors are fused at feature level and the fused feature vectors are used to index the database by forming K-d tree. This method reduces the data retrieval time along with possible error rates. Comparing the performance with indexing based on score level fusion they found that this technique performs better.

Another approach for indexing is graph based. In this approach indexing is done based on the characteristics of a graph. A minutiae graph is created from each fingerprint image and then an index is generated for each fingerprint which is unique [5]. Here, the index is generated based on four parameters number of vertex, degree of each vertex, highest degree and number of vertices with same degree.

How to extract ideal features that can reflect the intrinsic content of the images as complete as possible is still a challenging problem in computer vision. However, very little research has paid attention to this problem in the last decades. As we have known, the most common visual features include colour, texture and shape, *etc.* [8-16], and most image annotation and retrieval systems have been constructed based on these features. However, their performance is heavily dependent on the use of image features. In general, there are three feature representation methods, which are global, block-based, and region-based features.

A number of colour spaces have been used in literature, such as RGB, LUV, HSV and HMMD [9]. A number of important colour features have been proposed in the literatures, including colour histogram [17], colour moments (CM) [18], colour coherence vector (CCV) [19] and colour correlogram [20], *etc.*

Table II provides a summary of different colour methods excerpted from the literature, including their strengths and weaknesses. Here DCD, CSD and SCD denote the dominant colour descriptor, colour structure descriptor and scalable colour descriptor respectively.

TABLE II. CONTRAST OF DIFFERENT COLOUR DESCRIPTORS

Colour Method	Pros.	Cons.
Histogram	Simple to compute, intuitive	High dimension, no spatial info, sensitive to noise
CM	Compact, robust	Not enough to describe all colours, no spatial info
CCV	Spatial info	High dimension, high computation cost
Correlogram	Spatial info	Very high computation cost, sensitive to noise, rotation and scale
DCD	Compact, robust, perceptual meaning	Need post-processing for spatial info
CSD	Spatial info	Sensitive to noise, rotation and scale

SCD	Compact on need, scalability	No spatial info, less accurate if compact

III. BACKGROUND OF THE WORK

A. K-d tree

A k-d tree (short for k-dimensional tree) is a space partitioning data structure for organizing points in a k-dimensional space. K-d trees are a useful data structure for several applications, such as searches involving a multidimensional search key (e.g. range searches and nearest neighbour searches). K-d trees are a special case of binary space partitioning trees.

The k-d tree is a binary tree in which every node is a k-dimensional point. Every non-leaf node can be thought of as implicitly generating a splitting hyperplane that divides the space into two parts, known as half-spaces. Points to the left of this hyperplane are represented by the left sub tree of that node and points right of the hyperplane are represented by the right sub tree. The hyperplane direction is chosen in the following way: every node in the tree is associated with one of the k-dimensions, with the hyperplane perpendicular to that dimension's axis. So, for example, if for a particular split the "x" axis is chosen, all points in the sub tree with a smaller "x" value than the node will appear in the left sub tree and all points with larger "x" value will be in the right sub tree. In such a case, the hyperplane would be set by the x-value of the point, and its normal would be the unit x-axis.

B. Feature Extraction

Since the content based indexing of images is mostly based on the content features of images, so the feature extraction is a vital phase in image indexing. Before we proceed with the indexing of images, the images are gone through the feature extraction phase.

1) Colour features

Colour is one of the most important features of images. Colour features are defined subject to a particular colour space or model. Once the colour space is specified, colour feature can be extracted from images or regions. Among the feature extraction techniques, colour moment (CM) is one of the simplest yet very effective features [7,18]. The common moments are mean, standard deviation and skewness, the corresponding calculation can be defined as follows:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} f_{ij} \quad (1)$$

$$\sigma_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^2\right)^{\frac{1}{2}} \quad (2)$$

$$\gamma_i = \left(\frac{1}{N}\sum_{j=1}^{N}\left(f_{ij} - \mu_i\right)^3\right)^{\frac{1}{3}} \quad (3)$$

Where, f_{ij} is the colour value of the i^{th} colour component of the j^{th} image pixel and N is the total number of pixels in the image. μ_i, σ_i, γ_i (i=1,2,3) denote the mean, standard deviation and skewness of each channel of an image respectively.

2) Texture features

Texture is a very useful characterization for a wide range of image. It is generally believed that human visual systems use texture for recognition and interpretation. In general, colour is usually a pixel property while texture can only be measured from a group of pixels. Based on the domain from which the texture feature is extracted, the techniques can be broadly classified into spatial texture feature extraction methods and spectral texture feature extraction methods. For the former approach, texture features are extracted by computing the pixel statistics or finding the local pixel structures in original image domain, whereas the latter transforms an image into frequency domain and then calculates feature from the transformed image. As the most common method for texture feature extraction, Gabor filter has been widely used in image texture feature extraction.

3) Shape features

Shape is known as an important cue for human beings to identify and recognize the real-world objects, whose purpose is to encode simple geometrical forms such as straight lines in different directions. Shape feature extraction techniques can be broadly classified into two groups, viz., contour based and region based methods. The former calculates shape features only from the boundary of the shape, while the latter method extracts features from the entire region.

4) Wavelet transform for features

The Wavelet transform plays an important role in image processing by its good quality of multi resolution representation of an image. Haar Wavelet is the simplest category of wavelets which are related to the mathematical operations of Haar transform and used for image compression as well as feature extraction [21]. Haar wavelet transform has number of advantages such as it could able to represent the core pattern if the image, thus by removing the redundant information and it is computationally efficient in terms of both time and space complexity.

C. Proximity measure

The proximity measure is calculated by using the two distance formulas i.e. Euclidian distance and City Block distance. They are described below.

1) Euclidian distance

The Euclidian distance between two points (*x1*, *y2*) and (*x1*, *y2*) is given by the following formula as,

$$D = \sqrt[2]{(x1 - x2)^2 + (y1 - y2)^2} \quad (4)$$

2) City Block distance

The City Block distance between two points (*x1*, *y2*) and (*x1*, *y2*) is given by the following formula as,

$$D = |x1 - x2| + |y1 - y2| \quad (5)$$

IV. METHOD USED

As per my study of different research works related to the indexing of image data, the different methods that are used for image indexing till the time are pyramid technique, K-d tree, Map Reduce, R-tree, R+ tree, score based method etc. In my approach, I have implemented the K-d tree as the main stage of image indexing.

The proposed system of my work "indexing using K-d tree" is shown in Fig. 1.

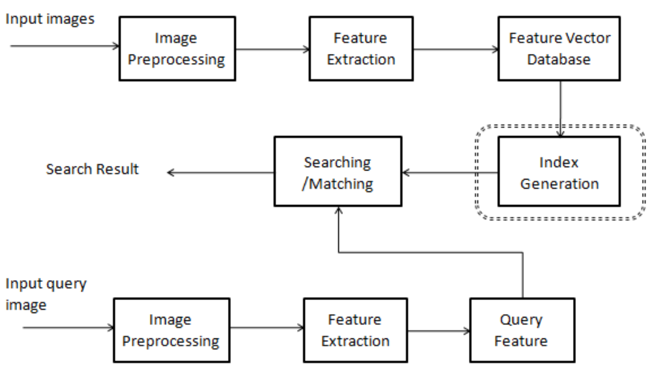

Fig. 1. Proposed Image Indexing System based on K-d tree

A. Pre-processing

The first stage of my proposed indexing system is image pre-processing. The input to this stage is the set of RGB (red, green, blue) colour images. In this stage the pre-processing of images is done as by resizing the images into a fixed size. A dataset of 127 RGB images of different sizes (based on number of pixels) is taken as input for the pre-processing stage. The images are then resized into a fixed size (400 by 600 pixels). The output of the pre-processing stage is a set of RGB images having identical dimensions that are the input to the next stage which is the feature extraction stage.

B. Feature Extraction

The feature extraction of the set of images is done once based on the colour content of the images using colour moment and once using the haar wavelet transform.

1) Colour Moment

In this stage the RGB colour features of each colour levels of RGB (Red, Green, and Blue) for each image are extracted using the technique called as colour moments (CM). In my approach, I have used the mean colour moment. The corresponding calculation for the mean colour moment is given by the equation (1) in the earlier section.

The output of this stage will be a set of 3 values (colour moment of red, green and blue respectively) for each image in the dataset.

2) Haar Wavelet Transform

The discrete wavelet transform for single level haar wavelet is used for feature extraction of the pre-processed images (after converting the RGB images to gray) for testing it against the other method.

In the wavelet toolbox of MATLAB there is a function called as DWT2( ), which is a single level 2-D discrete wavelet transform. DWT2( ) performs a single-level 2-D wavelet decomposition with respect to either a particular wavelet ('wname') or particular wavelet filters (Lo_D and Hi_D) that we specify.

The function is given as,

[CA,CH,CV,CD] = DWT2(X,'wname')

Where, 'wname' is a string containing the wavelet name. So it can be Haar or any other wavelet transform.

The output of this wavelet transform is 4 matrices viz. the approximation coefficients matrix CA and details coefficients matrices CH, CV, CD, (horizontal, vertical and diagonal respectively) obtained by a wavelet decomposition of the input matrix X.

C. Proximity Measure

The proximity measure is calculated by using the distance formula called as Euclidian distance which is given by the equation (4).

D. Indexing and Searching

The colour moment values which we have got as the output from the feature extraction stage is provided as input for the next stage which is index generation. Here the K-d tree data structure is used for indexing the images using the set of feature values.

Now the next step here is to construct the k-d tree to generate the indices. The K-d tree is constructed by array implementation of the nodes. The set of colour moment values for each image is considered as a node and provided as input to the k-d tree. The three colour moments for red, green and blue colour component represents the three dimensions to the k-d tree. In each step one dimension is chosen and the median value among the colour moment values (for a particular colour component) is calculated. The image possessing this set of feature values is then placed as a node into the k-d tree.

The CM value of Red, Green and Blue component is considered as the three dimensions X, Y and Z respectively. The first node or the root node is selected based on the median value of X dimension and levelled as first. Then in the second level the children of the root node is inserted to the tree based on the median value of Y dimension and so on. A node in a particular level is inserted in the left sub tree if the CM value for that particular dimension is less than that of its parent, otherwise if it is more it is inserted in the right sub tree.

After indexing is done, we need to check whether this helps in searching of any image data in an efficient manner. So, now we again need to go through the two stages namely, pre-processing and feature extraction for the query image.

Then, the query feature which is obtained from the above stages is compared with the features of the dataset by computing the proximity measure between the query image and other corresponding images in the dataset. We name this phase as the searching or matching stage. The result will provide the information whether the image is present or not in the dataset and also the time elapsed during searching.

I have used a 3-dimensional k-d tree. I have proposed an algorithm for indexing of an image dataset. The steps of my proposed method are as given below:

Algorithm : K-d_image ( )

Input: images

Output: index of the images

Step1. Input an RGB image.

Step2. Resize the image into a fixed size.

Step3. Compute the colour moments (CM) for each colour of RGB for each image in the dataset.

Step4. Compute the median of the colour moment values of each colour.

Step5. The set of RGB median colour moment values in a level is inserted as a node to the k-d tree.

Step6. The root node of the k-d tree is indexed as 1.

Step7. Every internal and leaf node is indexed as 2 times that of its parent node for left node and 2 times plus 1 for right node.

Step8. Find the Euclidian distance between the query image and the images in the sub tree.

Step9. The distance values are compared to give the search result.

V. EXPERIMENTAL RESULTS

A. Dataset Used

The dataset used are colour (RGB) images of different sizes (in terms of number of pixels) collected from the internet and also from other sources. These images include colour images of various flowers, birds, animals (dogs, tigers, bears etc.), fruits etc. The name of the website is McGill Calibrated Colour Image Database and the URL is http://tabby. vision.mcgill.ca/html/ browsedownload.html/.

B. Result

The stepwise output of the implementation of the above algorithm is shown in the following Figures and Tables.

127 images are used out of the dataset for indexing. The input image is an RGB image, which is first resized into a fixed size (in terms of number of pixels) image as shown in the figures below. Fig. 2. shows the original image and Fig. 3. shows the resized image.

Fig. 2. Original image

Fig. 3. Resized image

After resizing, the mean colour moment of an image is calculated and stored in another matrix. Each row of this matrix will contain the moment values for only a single image. If there are 'n' no of images then 'n' rows will be formed.

Considering the set of moment values as nodes the K-d is constructed and then the indexing is performed. Table III shows the K-d indices and respective original indices of a collection of 10 images along with their feature values.

The K-d tree constructed out of the feature values of Table III containing 7 nodes can be depicted as in Fig. 4.

TABLE III. INDICES ALONG WITH FEATURE VALUES OF A COLLECTION OF 10 IMAGES

K-d Index	Colour moment of Red component	Colour moment of Green component	Colour moment of Blue component	Original Index
1	36	41	24	116
2	34	26	18	89
3	40	38	44	117
4	25	26	17	32
5	35	31	23	41
6	44	33	23	25
7	41	44	34	65
8	23	19	14	97
9	26	22	18	66
10	30	29	11	119

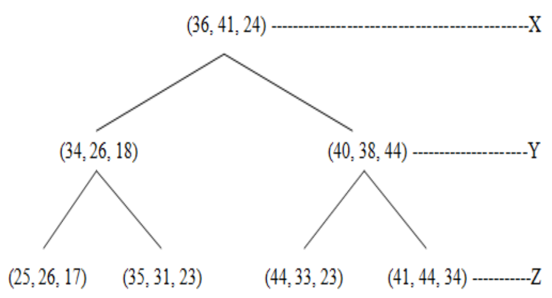

Fig. 4. Constructed 3-d tree

TABLE IV. COMPARISON OF SEARCH TIME OF METHODS WITH AND WITHOUT K-D TREE

Indexing method	Colour moment	Haar wavelet transform
Without K-d tree	0.00406024	0.204625
With K-d tree	0.00046927	0.000124

After providing the input query image, the indexing system will perform searching and gives the result in terms of search time and also displaying a message as the image is found or not found in the dataset. The searching is done with and without indexing using K-d tree including different feature extraction techniques and the results thus obtained are compared in the table IV.

The following is a screenshot of MATLAB GUI of the indexing system after selecting the input query images for both with and without K-d tree indexing.

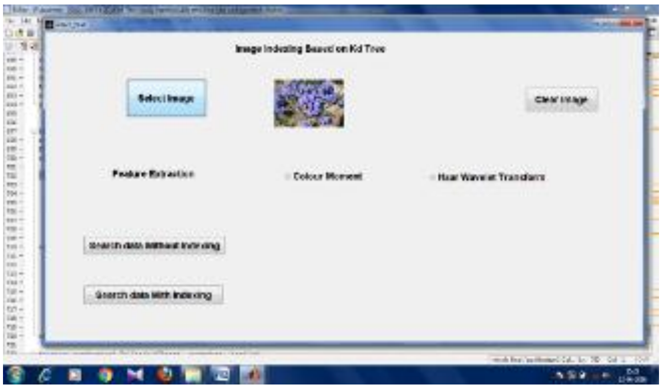

Fig. 5. Starting GUI

The time elapsed for the different methods after searching is performed (when the image is present in the database) is measured and showed in the following GUI.

Fig. 6. GUI displaying image found

Again the time elapsed for the different methods after searching is performed (when the image is not present in the database) is measured and showed in the GUI as in Fig. 7.

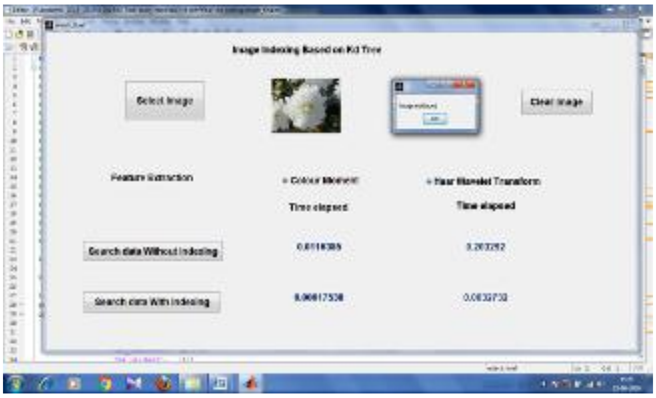

Fig. 7. GUI displaying image not found

VI. CONCLUSION

The K-d tree is implemented for indexing multidimensional image data. For a given query image searching is also performed to find a match of the input image with the images that are indexed (image dataset). The proposed algorithm worked nicely. I compared this with other methods without using the K-d tree. The previous one gives a better result in terms of searching time which is a very small fraction of second as compared with the other methods. Thus searching in a K-d tree takes only O (log n) time.

The K-d tree can perform better than a quad tree. It requires less time in searching despite large dimensionality. It can also find the nearest neighbour efficiently.

VII. FUTURE SCOPE

The K-d tree implementation for image indexing is done using a colour image dataset. So in near future, one can focus in the design of algorithms to implement the K-d tree using different type of data other than spatial image data such as hyperspectral or multispectral data. Again, since Insertion, deletion in a K-d tree are costly so special measure has to be taken in this aspect and lesser searching time should be achieved.

REFERENCES

[1] Md. K. I. Rahmani, and R. Sharma, "Image Indexing and Retrieval," International Journal of Software and Web Sciences (IJSWS), ISSN (Print): 2279-0063, ISSN (Online): 2279-0071.

[2] M. Gogoi, J. Das, "Indexing of Voluminous Data: Its Needs and Challenges," International conference on Electronic Devices, Circuits, Applied Electronics and Communication Technology, 2015.

[3] Li, Jia and Wang, Cheng, "Indexing Method for Hyperspectral Data Fast Retrieval by Pyramid Technique," International Conference on Computer Science and Software Engineering, 2008.

[4] U. Jayaraman, S. Prakash, and P. Gupta, "Indexing Multimodal Biometric Databases Using Kd-Tree with Feature Level Fusion."

[5] M. Gogoi, and D. K. Bhattacharya, "An Effective Fingerprint Verification Technique," Journal of Computer Science and Engineering, Volume 1, Issue 1, May 2010.

[6] J. K Lawder, and P. J. H. King, "Querying Multi-dimensional Data Indexed Using the Hilbert Space-Filling Curve."

[7] D. P. Tian, " A Review on Image Feature Extraction and Representation Techniques ," International Journal of Multimedia and Ubiquitous Engineering , Volume 8, No 4, July 2013.

[8] T. K. Shih, J. Y. Huang, C. S. Wang, et al., "An intelligent content-based image retrieval system based on colour, shape and spatial relations", In Proc. National Science Council, R. O.C., Part A: Physical Science and Engineering, vol. 25, no. 4, pp. 232-243, 2001.

[9] P. L. Stanchev, Jr. D. Green, and B. Dimitrov, "High level colour similarity retrieval", International Journal of Information Theories and Applications, vol. 10, no. 3, pp. 363-369, 2003.

[10] N. C. Yang, W. H. Chang, C. M. Kuo, et al., "A fast MPEG-7 dominant colour extraction with new similarity measure for image retrieval", Journal of Visual Comm. and Image Retrieval, vol. 19, pp. 92-105, 2008.

[11] M. M. Islam, D. Zhang, and G. Lu, "A geometric method to compute directionality features for texture images", In Proc. ICME, pp. 1521-1524, 2008.

[12] S. Arivazhagan, and L. Ganesan, "Texture classification using wavelet transform", Pattern Recognition Letters, vol. 24, pp. 1513-1521, 2003.

[13] S. Li, and S. Shawe-Taylor, "Comparison and fusion of multi-resolution features for texture classification", Pattern Recognition Letters, vol. 26, no. 5, pp. 633-638, 2005.

[14] W. H. Leung, and T. Chen, "Trademark retrieval using contour-skeleton stroke classification", In Proc. ICME, pp. 517-520, 2002.

[15] Y. Liu, J. Zhang, and D. Tjondronegoro, et al., "A shape ontology framework for bird classification", In Proc. DICTA, pp. 478-484, 2007.

[16] C. F. Tsai, "Image mining by spectral features: A case study of scenery image classification", Expert Systems with Applications, vol. 32, no. 1, pp. 135-142, 2007.

[17] A. K. Jain, and A. Vailaya, "Image retrieval using colour and shape", Pattern Recognition, vol. 29, no. 8, pp. 1233-1244, 1996.

[18] M. Flickner, H. Sawhney, and W. Niblack, et al., "Query by image and video content: the QBIC system", IEEE Computer, vol. 28, no. 9, pp. 23-32, 1995.

[19] G. Pass, and R. Zabith, "Histogram refinement for content-based image retrieval", In Proc. Workshop on Applications of Computer Vision, pp. 96-102, 1996.

[20] J. Huang, S. Kuamr, and M. Mitra, et al., "Image indexing using colour correlogram", In Proc. CVPR, pp. 762-765, 1997.

[21] K. Usha, M. Ezhilarasan, "Haar-Wavelet Transform based Finger Knukle Print Recognition," International Conference on Recent Trends in Information Technology, 2014.

[22] J. Barros, J. French, W. Martin, P. Kelly, and J. M. White, "Indexing multispectral images for content-based retrieval."

[23] T. G. Kulcsar, Sarossy, G. Bereznai, R., Auer, and J. Abonyi, "Visualization and Indexing of Spectral Databases," World Academy of Science, Engineering and Technology, vol. 6, July 2012.

[24] T. Sellis, N. Roussopoulos, and C. Faloutsos, "The R+-Tree: A Dynamic Index for Multi-Dimensional Objects," Proceedings of 13th International Conference on Very Large Data Bases, September 1987.

[25] A. Paliwal, U. Jayaraman, and P. Gupta, "A score based indexing scheme for palmprint databases," Proceedings of 2010 IEEE 17th International Conference on Image Processing, September 26-29, 2010, Hong Kong.

[26] S. Berchtold, C. Böhm, and H. P. Kriegel, "The Pyramid-Technique: Towards Breaking the Curse of Dimensionality," Proceedings of the ACM SIGMOD International Conference on Management of Data, pp. 142-153, 1998.

[27] P. Lu, G. Chen, B. C. Ooi, H. T. Vo, and S. Wu, "ScalaGiST: Scalable Generalized Search Trees for MapReduce Systems [Innovative Systems Paper]."

[28] A. Mhatre, S. Chikkerur, and V. Govindaraju, "Indexing Biometric Databases using Pyramid Technique."

[29] T. W. S. Chow, and M. K. M. Rahman, "A new image classification technique using tree-structured regional features", Neurocomputing, vol. 70, no. 4-6, pp. 1040-1050, 2007.

[30] C. F. Tsai, and W. C. Lin, "A comparative study of global and local feature representations in image database categorization", In Proc. 5th International Joint Conference on INC, IMS & IDC, pp. 1563-1566, 2009.

[31] H. Lu, Y. B. Zheng, and X. Xue, et al., "Content and context-based multi-label image annotation", In Proc. Workshop of CVPR, pp. 61-68, 2009.

Author Profile

Jayashree Das received her M. Tech degree in Computer Science and Engineering (CSE) from Girijananda Chowdhury Institute of Management and Technology (GIMT), Guwahati under Assam Science and Technology University, Guwahati in 2016 and received her B. E. degree in CSE from GIMT, Guwahati under Gauhati University in 2012. Presently she is working as a guest faculty, Department of CSE, GIMT, Azara, Guwahati. Her research areas include software testing and Image processing. She has published and presented one survey paper in an International conference in 2015.

Dr Minakshi Gogoi did her M. Tech in 2005 and PhD in 2015 from dept. of Computer Science and Engineering, Tezpur University, Tezpur. Presently she is serving as an Asst. Professor, in the Computer Science & Engineering Department at GIMT, Guwahati. Her research areas include Biometrics authentication, Machine learning and Image Processing. She has published 16 research papers in International/ National Journals and conferences. She has also served as a reviewer of many internationally reputed journals. She is acting as a member of professional bodies of India like Institute of Engineers, India and Indian Society for Technical Education (ISTE).

