An Experiential Survey on Image Mining Tools, Techniques and Applications

C. Lakshmi Devasena,
Dept. of Software Systems,
Karpagam University,
Coimabatore, India

T.Sumathi
Dept. of Software Systems,
Karpagam University,
Coimabatore, India

Dr. M. Hemalatha
Dept. of Software Systems,
Karpagam University,
Coimabatore, India

Abstract – Digitization in every sector leads to the growth of digital data in a tremendous amount. Digital data are not only available in the form of text but it is also available in the form of images, audio and video. Decision making people in every field like business, public sector, hospital, etc. are trying to get useful and implicit information from the already existing digital data bases. Image mining is the concept used to extract implicit and useful data from images stored in the large data bases. Image mining is used in variety of fields like medical diagnosis, space research, remote sensing, agriculture, industries and even in the educational field. This paper elaborates the research works already done in image mining and also summarizes different tool developed, algorithms emerged and the applications of image mining used to extract the useful images in various fields.

Keywords-Image Mining for WWW; Mining Color Images; Medical Diagnosis.

I. INTRODUCTION

Progresses in image acquirement and storage technology have led to implausible growth in very large and detailed image databases. These images, if analyzed, can expose useful information to the human users. Image mining deals with the extracting inherent and embedded knowledge, image data relationship, or other patterns which is not explicitly found in the images [24]. Image mining is more than just an expansion of data mining to image domain. It is an interdisciplinary challenge that draws upon proficiency in computer vision, digital image processing, image extraction, data mining, machine learning, databases, and artificial intelligence. Wynne Hsu, Mong Li Lee and Ji Zhang [60] examines the research issues in image mining, developments in image mining, predominantly, image mining frameworks and suggests some future research guidelines for image mining.

The fundamental challenge in image mining is to reveal out how low-level pixel representation enclosed in a raw image or image sequence can be processed to recognize high-level image objects and relationships. Ji Zhang, Wynne Hsu and Mong Li Lee [23] proposed an efficient information-driven framework for image mining. In that they made out four levels of information: Pixel Level, Object Level, Semantic Concept Level, and Pattern and Knowledge Level. To achieve that High-dimensional indexing schemes and retrieval techniques are incorporated in the framework to maintain the flow of information among the levels. Ji Zhang, Wynne Hsu and Mong Li Lee [24] highlighted the need for image mining in the era of rapidly growing amounts of image data and pointed out about the unique characteristics of image databases. In addition, it is also examined function-driven and information-driven frameworks for image mining.

Image mining normally deals with the study and development of new technologies that allow accomplishing this subject. Image mining is not only the simple fact of recovering relevant images; the aim is the innovation of image patterns that are noteworthy in a given collection of images. J. Fernandez, N. Miranda, R. Guerrero, F. Piccoli [17] shows how a natural source of parallelism provided by an image can be used to reduce the cost and overhead of the whole image mining process(Figure 1). The images from an image database are first preprocessed to improve their quality. These images then undergo various transformations and feature extraction to generate the important features from the images. With the generated features, mining can be carried out using data mining techniques to discover significant patterns.

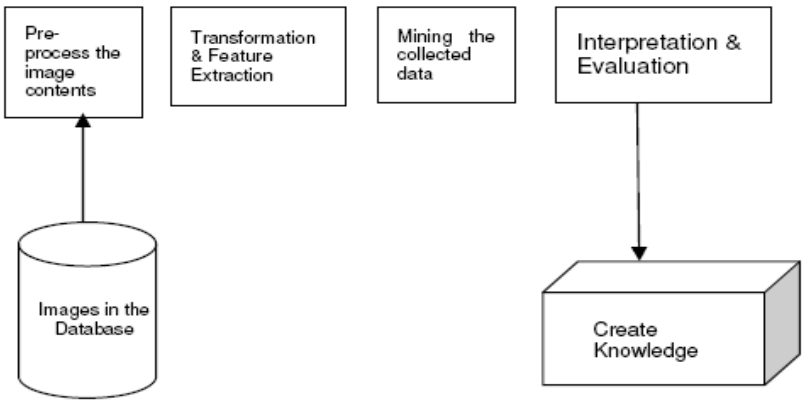

Figure1: Processes in Image Mining

The resulting patterns are evaluated and interpreted to obtain the final knowledge, which can be applied to applications. [23]

II. IMAGE MINING LANGUAGES, SYSTEMS AND TOOLS

Hassan H. Malik [19] provides a simple, easy to learn language that facilitates mining association rules from images called iARM. iARM is a scripting language that makes it easier to dig out association rules from images. It lets us to make a list of source image files and customize association rule parameters similar to number of terms, filters, support and confidence. Using iARM, we can extract association rules by writing simple, easy understanding code that can be written and maintained by end users who need no programming knowledge. iARM can be customized according to the end-user needs and can extract flexible rules utilizing both textual and signal features. Extracted rules symbolize implicit knowledge contained in images, and implicit relationship that exists in a set of images. This extracted data could further assist in classification and clustering of images.

Pham, N.K. Morin, A. Gros, P. [43] proposed an interactive graphical tool, CAViz, which allows us to exhibit and to extort knowledge from the results of a Correspondence Analysis (CA) on images. It is a descriptive technique designed to analyze simple two-way and multi-way tables having some measure of correspondence between the rows and columns.

Keiji Yanai [10], proposed a new system to mine visual knowledge on the Web. This system introduced a latest image recognition technique, the bag-of-key points representation to Web image-gathering task and the SVM classifier. By the experiments it is proved that the proposed system outperforms their previous systems and Google Image search greatly.

Sangkyum Kim, Xin Jin, Jiawei Han[51] proposed a novel image representation method *B2S* (Bag to Set) that keeps all frequency information and is more discriminative than traditional histogram based bag representation. Based on *B2S*, they have constructed two different image classification approaches. First they applied *B2S* to a *state-of-the-art* image classification algorithm *SPM in* computer vision. Second, they designed a framework *DisIClass* (Discriminative Frequent Pattern-Based Image Classification) to utilize data mining algorithms to categorize images, which was hardly done before due to the built-in differences between the data of data mining and computer vision fields. *DisIClass* acclimatizes the locality property of image data, and apply sequential covering method to persuade the most discriminative feature sets from a closed frequent item set mining method.

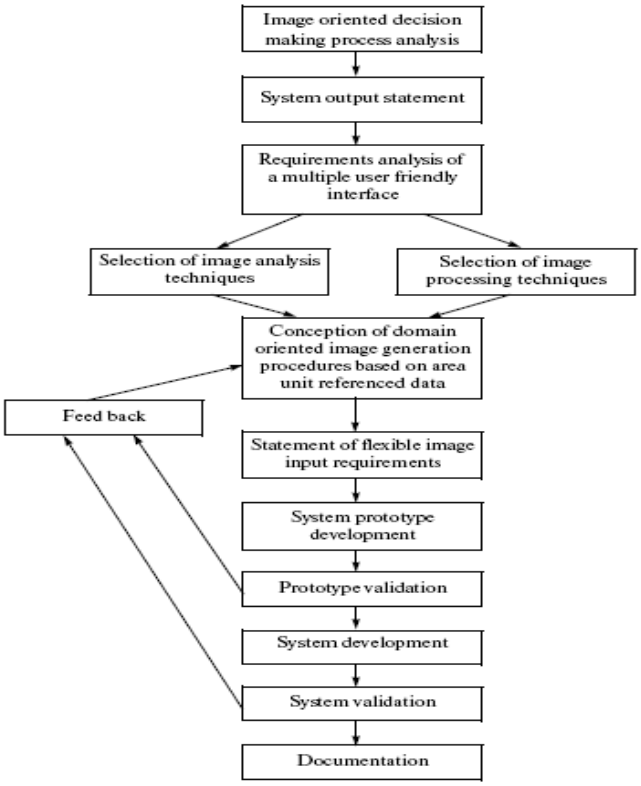

Figure 2: Image Mining oriented Geo-processing System (IMOGS).

R. Guadagnina, L. Santanab, E. Fernedaa, and H. Pradoa [18] proposed an integration of Geoprocessing and Image Mining to support image based decisions in several domains such as healthcare. The IMOGS model is shown in Figure 2.

Ana B. Benitez, Mandis Beigi, and Shih-Fu Chang [1] describes MetaSEEk, a meta-search engine used for extracting images based on their visual content on the Web as shown in Figure 3. MetaSEEk is designed to astutely select and interface with multiple on-line image search engines. This is done by ranking their performance for diverse classes of user queries. User response is also integrated in the ranking refinement. MetaSEEk has been built up to discover the issues involved in querying large, distributed, on-line visual information system sources.

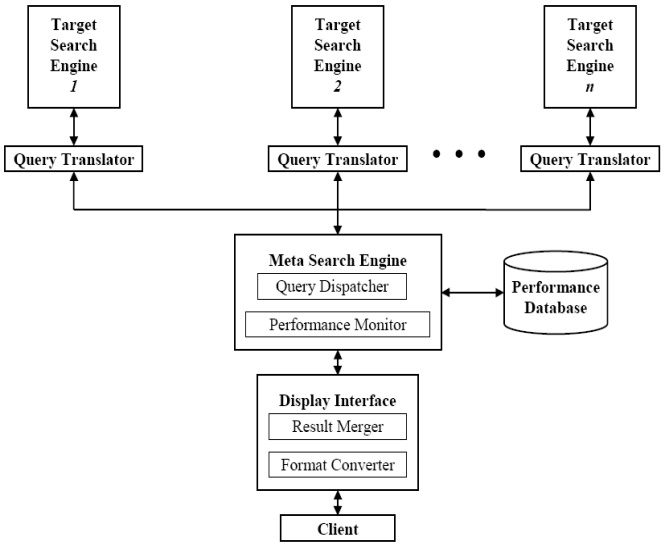

Figure 3: Basic components of a meta-search engine.

III. IMAGE MINING FOR WWW

World Wide Web, which is the repository of very huge amounts of data, provides access to every body in the world who are starving for information. From web we can easily obtain digital images of various kinds of real world scenes. Presently, however, classification /recognition of generic real world images are far from practical due to a multiplicity of real world scenes. Keiji Yanai [28] has proposed a project to deal with such diversity. It is an ongoing project aims at generic image classification using images automatically-gathered from the Web as training images instead of hand-made image collections.

Keiji Yanai [29] describes a generic image classification system with an automatic knowledge acquisition mechanism from the Web. They have used three steps for the processing. First step, the gathering stage, gathers images related to given class keywords from the Web automatically. Second step, the learning stage, extracts image features from gathered images and associates them with each class. Third step, the classification stage, classifies an unknown image into different classes corresponding to class keywords by using the relationship between the image features and the classes.

Nick Morsillo, Chris Pal, Randal Nelson [40] presented a technique that allows a user to reduce noisy search results and characterize a more precise visual object class. This approach is based on semi-supervised machine learning in a novel probabilistic graphical model made of both generative and discriminative elements.

Keiji Yanai [30] propose a new method to select relevant images to the given keywords from images acquired from the Web based on the Probabilistic Latent Semantic Analysis (PLSA) model.It is proved that this method can select more various images compared to the existing SVM-based methods.

Bingbing Ni, Zheng Song, Shuicheng Yan [10] presented an *automatic* image mining system in web which builds a *universal* human age estimator based on facial information, which can be used to all racial groups and various image qualities. In this system an outlier removal step with PCA refines the images and then a robust multi-instance regressor learning algorithm is projected to learn the kernel-regression based human age estimator under the scenarios with possibly noisy bags.

Keiji Yanai[31] proposed a new method for automated large scale gathering of Web images relevant to specified concepts. In that, good quality candidate sets of images for each keyword are collected as a function of analysis of the surrounding HTML text. The gathered images are then segmented into regions of images by using a Gaussian mixture.

Keiji Yanai, Kobus Barnard [15] proposed a measuring “visualness” of concepts with images on the Web, which measures the level of extent of visual content in a document / image. And this concept might be used for image annotation.

The relations between different modalities of Web images could be very useful for Web image retrieval. Ruhan He Wei Zhan [49] investigates the multi-modal associations between two basic modalities of Web images, i.e. keyword and visual feature clusters, by multi-model association rule.

Automatic image classification is a demanding research topic in Web image mining. Rong Zhu, Min Yao and Yiming Liu [47] formulated image classification problem as the calculation of the distance measure between training manifold and test manifold. They proposed an improved nonlinear dimensionality reduction algorithm based on neighborhood optimization to decrease feature dimensionality and complexity.

Since the web images are not annotated, it very difficult to get user intended image from web. Zheng Chen, Liu Wenyin, Feng Zhang, Mingjing Li, Hongjiang Zhang [61] presented an effective approach to and a prototype system for image retrieval from the Internet using web mining. This system can serve as a web image search engine. Key idea in this approach is to extract the text information on the web pages to semantically describe the images. The text description is then pooled with other low-level image features in the image similarity assessment.

Even though many mining algorithms and approaches are there to retrieve images from web, mining semantically relevant images are the challenging one. Chunjie Zhang, Jing Liu, Hanqing Lu, Songde Ma [15] proposed a concept in that first Google Image Searcher is used to find the relevant images and then the output image set is explored to learn sensitive Markov stationary feature (C-MSF) to represent images by the algorithm of random walk with restart (RWR), in which the spatial co-occurrence of the bag-of-words representation and the concept information are integrated and classified with a SVM based classifier for web image mining.

Of late, e-business is a peak business and the companies need to keep their product details online. So the web sites for e-business and shopping mall deal with a lot of image information. To locate a specific image from these image sources, we habitually use web search engines or image database engines which rely on keyword based retrievals or color based retrievals with limited search capabilities. Seong-Yong Hong [53]

presented an intelligent web e-catalog image retrieval system IIMS using metadata and user log. In that, the texture and color based image classification and indexing techniques like bit vector indexing are used to represent schemes of user usage patterns. The system keeps track of user's preferences by generating user query logs and robotically adds more search information to successive user queries.

Chu-Hong Hoi, Michael R. Lyu [13] proposed an algorithm using support vector machines for attacking the problems to learn Web images for searching the semantic concepts in large image databases.

Mohamed Y. Eldib and Hoda M. Onsi [38] introduced a fully automated age estimation engine that is competent of collecting images using human age related text queries from Flickr photo sharing website. They used the Active Shape Model for robust face detection; it acts also as a removal step for non-face images. After that, the bio-inspired features (BIF) to extract the facial aging information are used. Then to label Flickr images automatically, a universal labeler algorithm is introduced. Finally, they have used the web image collection as a training dataset, and the standard databases as testing datasets presents the superiority of the proposed image web mining algorithm. The entire process is shown in Figure 4.

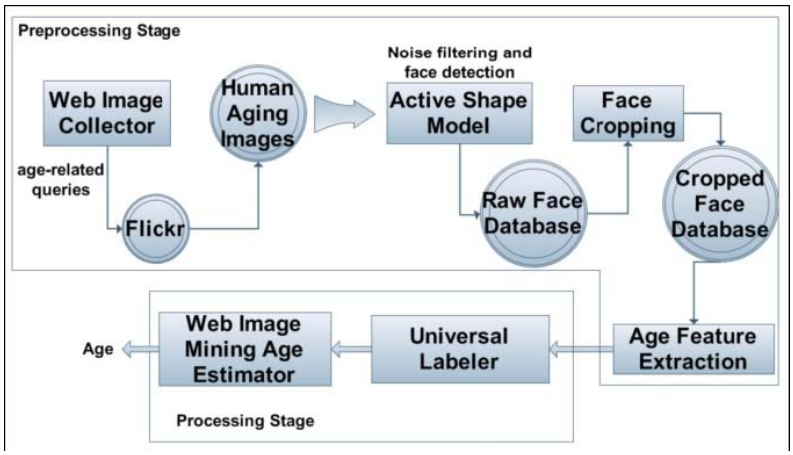

Figure 4: Web image mining using universal labeler algorithm for efficient and effective image labeling.

Wichian Premchaiswadi, Anucha Tungkatsathan [58], presented an On-line Content-Based Image Retrieval (CBIR) System using joint query and relevance feedback concept based on high-level and low-level features. This approach also used fast and efficient color feature extraction namely auto color correlogram and correlation (ACCC) based on color correlogram (CC) and autocorrelogram (AC) algorithms, for extracting and indexing low-level features of images. To integrate an image analysis algorithm into the text-based image search engines without debasing their response time, the framework of multi-threaded processing is proposed.

IV. MINING COLOR IMAGES

Aura Conci., Everest Mathias M. M. Castro [7] proposed a framework for mining images by colour content. Their framework provides the possibility of use 5 distance function for evaluation of similarity among images and 2 type of quantization. D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos [2] presented a scheme which implements a recursive *HSV*-space segmentation technique to identify perceptually prominent color areas. The average color vector of these extracted areas is then used to build the image indices, requiring very little storage.

Dr. Sanjay Silakari, Dr. Mahesh Motwani and Manish Maheshwari [50] proposed a framework focuses on color as feature using Color Moment and Block Truncation Coding (BTC) to extract features for image dataset. Then K-Means clustering algorithm is conducted to group the image dataset into various clusters. Lukasz Kobylinski and Krzysztof Walczak [35] proposed an application of Binary Thresholded Histogram (BTH), a color feature description method, to the creation of a metadatabase index of multiple image databases.

Rajshree S. Dubey [46] described about an Image mining techniques which is based on the Color Histogram, texture of that Image. The query image is taken then the Color Histogram and Texture is produced and based on this the resultant Image is found. They have investigated a histogram-based search methods and color texture methods in two different color spaces, RGB and HSV. Histogram search differentiate an image by its color distribution. It is shown that images retrieved by using the global color histogram may not be semantically associated even though they share similar color distribution in some results.

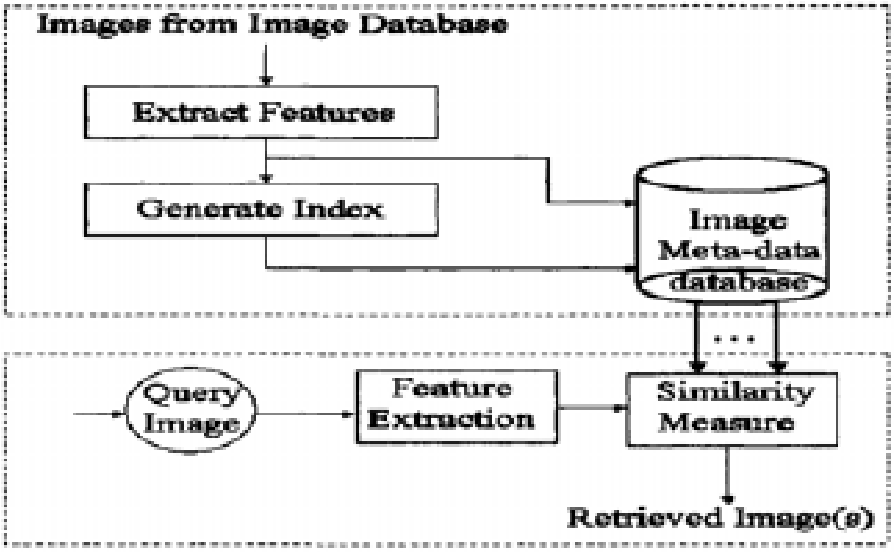

Figure 5: Content Based Image Retrieval (CBIR) System Architecture.

Carlos Ordonez and Edward Omiecinski [12] presented a data mining algorithm to find association rules in 2-dimensional color images. The algorithm has four main steps: feature extraction, object identification, auxiliary image creation and object mining and it does not rely on any type of domain knowledge. This resembles the CBIR system architecture shown in Figure 5.

Vitorino Ramos, Fernando Muge [57] formulated the segmentation problem upon images as an optimization problem and espouses evolutionary strategy of Genetic Algorithms for the grouping of small regions in colour feature space. The approach uses k-Means unsupervised clustering methods into Genetic Algorithms, for guiding Evolutionary Algorithm in the search for finding the finest or best data partition, task that as we know, requires a non-trivial search because of its intrinsic NP-complete nature. To solve this, the appropriate genetic coding is also discussed.

Dr.V.Mohan and A.Kannan [39] proposed a new technique Color Image Classification and Retrieval using a Image (Figure 6) for improving user interaction with image retrieval systems by fully exploiting the similarity information. In that technique retrieving the images from the image collection involves the steps like Preprocessing, Color image classification, Preclustering, Texture feature extraction, Similarity comparison and Neighboring target image selection.

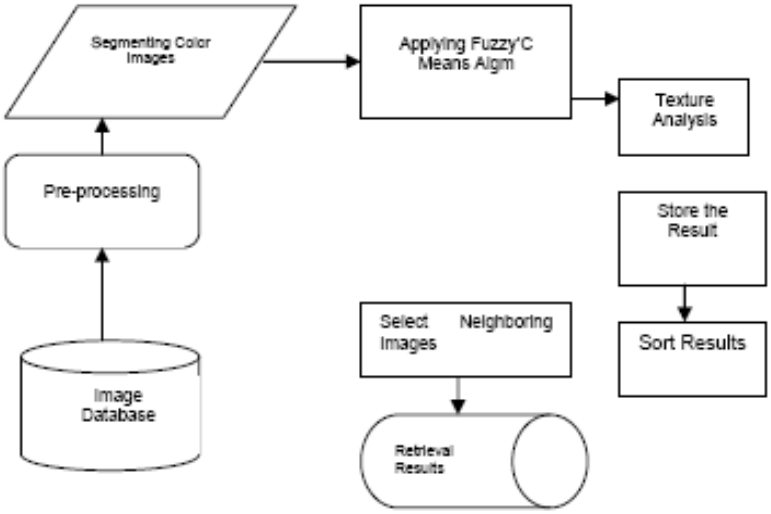

Figure 6: Block Diagram of Color Image Classification and Retrieval System

V. IMAGE MINING FOR MEDICAL DIAGNOSIS

Magnetic resonance imaging (MRI) of the brain, followed by automated segmentation of the Corpus Callosum in midsagittal sections has momentous applications in neurology and neurocognitive study since the size and shape of the Corpus Callosum are shown to be interrelated to sex, age, neurode generative diseases and various lateralized behaviour in people. The segmentation of the Corpus Callosum is considered as a crucial step in image mining frameworks to classify the brain MR images. Ashraf Elsayed, Frans Coenen1, Marta García-Fiñana and Vanessa Sluming [5] proposed a segmentation algorithm operates by first extracting regions satisfying the statistical characteristics (gray level distributions) of the Corpus Callosum that have relatively high intensity values. This is then processed using graph analysis and classification procedures.

Wynne Hsu, Mong Li Lee, Kheng Guan Goh [59] proposed an system called IRIS, an Integrated Retinal Information system, has been built up to afford medical professionals trouble-free and unified access to the screening, trend and development of diabetic-related eye diseases in a diabetic patient database.

P. Rajendran, M. Madheswaran [45] proposed a method deals with the detection of brain tumor in the CT scan brain images. First preprocessing technique applied on the images eliminates the inconsistent data from the CT scan brain images. Then feature extraction process is applied to extort the features from the brain images. A Novel Fuzzy Association Rule Mining (NFARM) applied on the image transaction database contains the features extracted from the CT scan brain images. The NFARM gives the diagnosis keywords to physicians for making a better diagnosis system.

Automated detestion of tumers in different medical images is motivated by the inevitability of high accuracy when we dealing with human life. L. Jeba Sheela, Dr. V.Shanthi [21] proposed a system using image mining techniques to categorize the images either as normal or abnormal and then classify the tissues of the abnormal brain MRI to identify brain related diseases.

P. Rajendran, M.Madheswaran [44] proposed a method concerned with the classification of brain tumor in the CT scan brain images. The main steps involved in the system are: pre-processing, feature extraction, association rule mining and hybrid classifier. The pre-processing is done using the median filtering and edge features are extracted using canny edge detection technique. The combination of two image mining approach is been proposed. The frequent patterns from the CT scan images are produced by frequent pattern tree (FP-Tree) algorithm that mines the association rules. The decision tree method is used to categorize the medical images for diagnosis. This system improves the classification process to be more accurate. The hybrid method enhances the efficiency of the proposed method than the traditional image mining methods.

Aswini Kumar Mohanty, Saroj Kumar Lenka [6] applied image mining in the domain such as breast mammograms to classify and detect the cancerous tissue. A hybrid approach of feature selection using fast branch and bound algorithm shown in Figure 7 and a hybrid genetic algorithms are used which approximately reduces 75% of the features and new decision tree is used for classification and provide promising results.

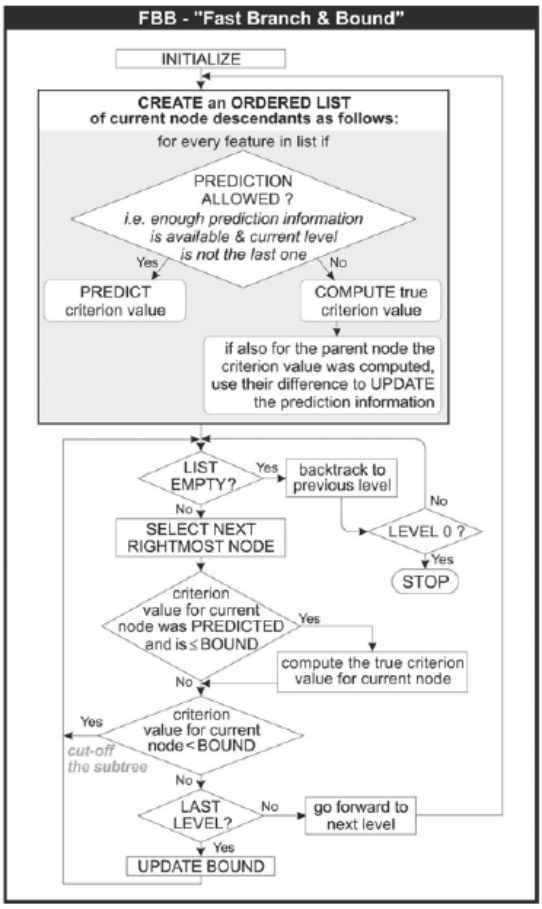

Figure 7: Fast Branch and Bound Algorithm

VI. IMAGE MINING IN VARIETY OF APPLICATIONS

Image Mining in Natural Scene Recognition: Aura Conci., Everest Mathias M. M. Castro [9] presented a method that considers patch appearances and its relationships in the form of adjectives and prepositions for natural scene recognition. For example, a 'beach' scene can be characterized by a 'sky' region above 'sand', and a 'water' region between 'sky' and 'sand'. This approach represents each image as a spatial pyramid, from which we attain a collection of patch appearances with spatial layout information. Many of the existing scene categorization approaches only employ patch appearances or co-occurrence of patch appearances to conclude the scene categories, but the relationships among patches remain ignored. In this approach, each image is represented as a spatial pyramid, from which a collection of patch appearances with spatial layout information is obtained. Then they apply a feature mining approach to get discriminative patch combinations. The extracted patch combinations can be interpreted as adjectives or prepositions, which will be used for scene understanding and recognition.

Image Mining in Remote Sensing: Information extraction using mining techniques from remote sensing image (RSI) is rapidly gaining attention among researchers and decision makers because of its potential in application oriented studies. Dr. C. Jothi Venkateswaran, S. Murugan, Dr. N. Radhakrishnan [26] has discussed that combination of various techniques such as statistical, decision, parametric and association rules help in extracting information effectively from RSI rather than applying any single method.

Shah, V.P. Younan, N.H. Durbha, S.S. King, R.L. [52] is presented a new feature set, obtained by integrating independent component analysis and wavelet transformation for image information mining in geospatial data.

Manoranjan Dash, Deepak K[37] discussed about an image mining application of Egeria detection. Egeria is a type of wild plant found in various lands and water regions over San Joaquin and Sacramento deltas. The challenge is to locate a view to accurately detect the weeds in new images. Their solution contributes two new aspects to image mining. (1) Application of view selection to image mining: View selection is appropriate

when a specific learning task is to be learned. For example, to look for an object in a set of images, it is useful to select the appropriate views (a view is a set of features and their assigned values). (2) Automatic view selection for accurate detection: They have used association rule mining to automatically select the best view. And their results show that the selected view outperforms other views including the full view.

T. H. Manjula Devi, H.S. Manjunatha Reddy, K. B. Raja, K. R Venugopal and L. M. Patnaik [36] proposed a Universal Steganalysis using Histogram, Discrete Fourier Transform and SVM (SHDFT). The stego image has irregular statistical characteristics as compare to cover image. Using Histogram and DFT, the statistical features are generated to train One-Class SVM to discriminate the cover and stego image. The block diagram and the algorithm for SHDFT are shown in Figure 8 and 9.

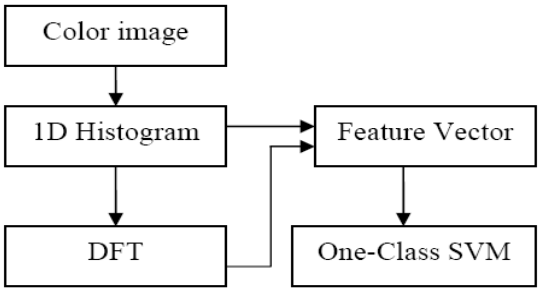

Figure 8: Block Diagram of SHDFT

Algorithm for SHDFT

- **Input**: The test image
- **Output**: SVM classifies the test image as cover or stego image.

1. Separation of the color planes, $c \in \{r, g, b\}$ of the color image.
2. Build Histogram (*1D*) for each color planes $c \in \{r, g, b\}$. Compute the 1^{st} and 2^{nd} moments i.e., mean and variance of the histogram coefficients. This yields *6D* statistics.
3. Build Discrete Fourier Transform for histogram of each color planes, $c \in \{r, g, b\}$. Compute the 1^{st}, 2^{nd}, 3^{rd} *and* 4^{th} moments i.e., mean variance, skewness and kurtosis of the DFT coefficients. This yields *12D* statistics. Also compute the total energy for the DFT coefficients. This yields *3D* more statistics.
4. The 1^{st} order moment Mean of the difference of histogram and DFT is computed for each color channel. This yields *3D* statistics.
5. The feature vectors obtained from steps *2, 3* and *4* yields *24D* features.
6. The feature vectors obtained in steps 5 are used to train SVM to classify test image.

Figure 9: SHDFT Algorithm

Currently government and private agencies use remote sensing imagery for ample applications from military application to farm development. The images may be sensitive to light and some might not. Remote sensing image classification is one of the most significant application worlds for remote sensing. Some of the image classification algorithms have proved good precision in classifying remote sensing data. K Perumal and R Bhaskaran [41] compared the different classification methods and their performances and concluded that Mahalanobis classifier; a supervised learning classification technique performed the best.

Nuclear Magnetic Resonance or Magnetic Resonance Imaging is the traditional method to perform fruit grading. However image processing technique can avoid the depredation of fruit comparing to traditional methods, applying image processing on fruit grading is current trend. Chu-Hui Lee and Meng-Ting Liu [14] modified primary k-means algorithm which perform efficiently at image segmentation. and investigated how to select better number of clusters automatically and fine-tuning in k-means algorithm especially for apple grading.

VII. TECHNIQUES IN IMAGE MINING

R. Brown, B. Pham [11] described in detail a general hierarchical image classifier approach, and illustrated the ease with which it can be trained to find objects in a scene using support vector machine concept. Ji Zhang, Wynne Hsu and Mong Li Lee [23] proposing new representation schemes for visual patterns and to facilitate fast and effective access we need to frame efficient content-based image indexing and retrieval techniques.

Dr.S.P.Victor, V.Narayani and S.Rajkumar [56] proposed an efficient retrieval technique for images using enhanced univariate transformation approach. In this method they have treated images as a compilation of the representative prototypes selected from the training image corpus, and then used the resulted distribution in the descriptor space as a characterization of the image. The words in the images, called visual words, are calculated to form a vocabulary of N words. Then each image was represented by a word histogram. The construction of visual words is evaluated in two steps: 1) Computation of local descriptors for a set of images 2) Cluster the previous descriptors by K-means. SIFT algorithm is then employed to extract the features due to its impressive performance in image recognition.

Peter Stanchev [42] proposed a new method for image retrieval using high level semantic features is proposed. It is based on removal of low level color, shape and texture characteristics and their translation into high level semantic features using fuzzy production rules derived using an image mining technique.

One of the leading frameworks for image object mining is the bag-of-words (BOW) approach. The idea is to encode an image as a collection of visual words of the quantized local patches. Objects in the image can then be retrieved through inferring the semantic topics associated with the set of visual words. However, the visual BOW mining framework is apt to suffer from the so-called term-mismatch problem (vocabulary problem). Jen-Hao Hsiao, Chu-Song Chen and Ming-Syan Chen [22] proposed a novel language model- based approach with pseudo-relevance feedback for tackling the vocabulary problem in visual BOW mining.

Due to improvements in image acquisition and storage technology, terabyte-sized of image databases leads to two basic problems: how to exploit images (image mining)? how to make it accessible to human beings (image retrieval)? The specificity of image mining /retrieval techniques operates on the whole collection of images, not a single one. Under these circumstances, it is natural that the time complexity of related algorithms plays an important role. T.Iwaszko, Mahmoud Melkemi and L.Idoumghar [20] suggested a novel general approach applicable to image mining and retrieval, using compact geometric structures which can be pre-computed from a database.

Rosalina Abdul Salam and Abdullah Zawawi Hj. Talib [48] proposed a technique to test the capability of producing an automatic shape recognition system by mining relevant image features. The method has the capability to be extended to three-dimensional objects, which is currently under investigation. Color, depth and texture can be grouped together to form a set of new features.

A.Kannan, Dr.V.Mohan, Dr.N.Anbazhagan[27] formulated a new technique called Image retrieval based on optimal clusters is proposed for improving user interaction with image retrieval systems by fully exploiting the similarity information. The technique involves steps as shown in Figure 10.

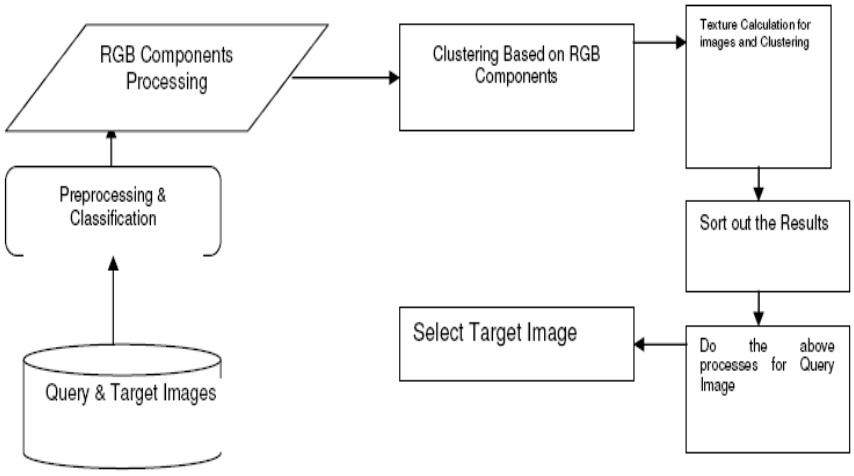

Figure 10: Block Diagram: Image Retrieval System

Dipesh Dugar M , Dinesh Kumar Jain N , Raaja Sarabhoje Gopal , Haritha Ravi , Baskaran Alagappan [16] proposed a novel discriminative learning framework based on canonical correlations for object recognition with image sets.

Anthony J.T. Lee, Ruey-Wen Hong, Wei-Min Ko, Wen-Kwang Tsao, Hsiu-Hui Lin [3] proposed a novel spatial mining algorithm, 9DLT-Miner. This is used to mine the spatial association rules from an image database. In the image database e every image is stored in 9DLT representation. Two phases are there in this method. First phase finds all lengthier frequent patterns. Second phase generates all candidate (k + 1)-patterns using frequent k-patterns (k P1). For each candidate pattern generated, the frequent supporting patterns in the database were counted. The steps are repeated until no more frequent patterns can be found.

Kun-Che Lu And Don-Lin Yang [34] adopted the decision tree induction to recognize relationships between attributes and the target label from image pixels, and constructed a model for pixel-wised image processing according to a given training image dataset. And it is proved that the proposed model can be very competent and effectual for image processing and mining.

S.P. Victor, S. John Peter [55] propose minimum spanning tree based clustering algorithm using weighted Euclidean distance for edges to segment the image.

S. John Peter [25] propose a novel algorithm, Minimum Spanning Tree based Structural Similarity Clustering for Image Mining with Local Region Outliers (MSTSSCIMLRO) to segment the given image and to detect anomalous pattern (outliers). In that he used weighted Euclidean distance for edges, which is key element in building the graph from image. The algorithm uses a cluster validation criterion based on the geometric properties of data partition of the data set in order to find the proper number of segments. Two phased involved in this algorithm. First phase creates optimal number of clusters/segments, Second phase segments the optimal number of clusters/segments and detect local region outliers.

Dr. Ashok N. Srivastava [4] presented a methodology for automatic knowledge driven image mining using the theory of Mercer Kernels. Mercer Kernels are extremely nonlinear symmetric positive definite mappings from the original image space to a very high, probably infinite dimensional feature space.

Rupali Sawant [54] presented a framework of image mining based on concept lattice and cloud model theory shown in Figure 11. The methods of image mining from image texture and shape features are introduced here, which include the basic steps: pre-processing the images, using cloud model to extract concepts, and then using concept lattice to extract a series of image knowledge.

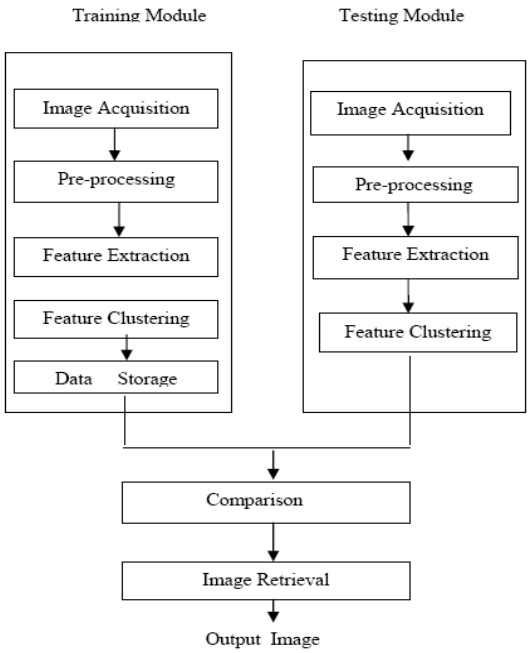

Figure 11: Framework of Image Mining System based on concept lattice and cloud model theory.

V. Balamurugan and P. Anandhakumar [8] introduced a neuro-fuzzy based clustering approach for content based image retrieval using 2D-wavelet transform (2D-DWT) as shown in Figure 12.

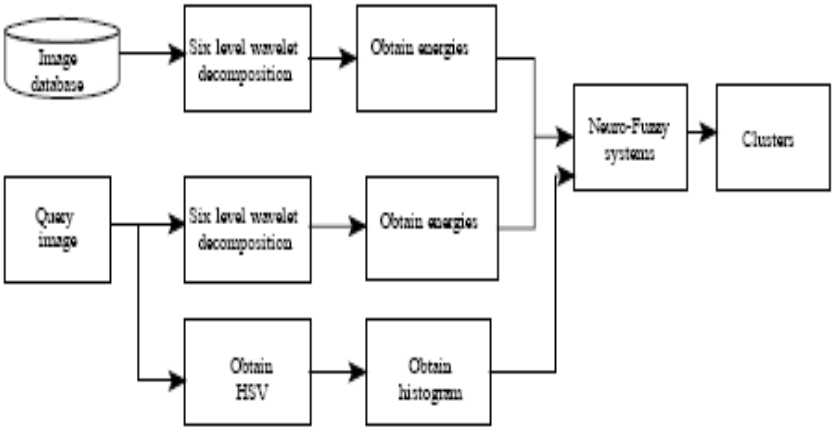

Figure 12: Neuro-Fuzzy based clustering approach for image retrieval using 2D- WT.

VIII. REFERENCES

[1] Ana B. Benitez, Mandis Beigi, and Shih-Fu Chang "A Content-Based Image Meta-Search Engine using Relevance Feedback"

[2] D. Androutsos, K. N. Plataniotis and A. N. Venetsanopoulos "A Novel Vector-Based Approach to Color Image Retrieval Using a Vector Angular-Based Distance Measure" Computer Vision and Image Understanding Vol. 75, Nos. 1/2, July/August, pp. 46–58, 1999

[3] Anthony J.T. Lee, Ruey-Wen Hong, Wei-Min Ko, Wen-Kwang Tsao, Hsiu-Hui Lin "Mining spatial association rules in image databases" Information Sciences 177 (2007) 1593–1608

[4] Dr. Ashok N. Srivastava, Dr. Nikunj oza "Knowledge Driven Image Mining With Mixture Density Mercer Kernels" Manuscript for the ESA-EUSC 2004

[5] Ashraf Elsayed, Frans Coenen1, Marta García-Fiñana and Vanessa Sluming "Segmentation for Medical Image Mining: A Technical Report"

[6] Aswini Kumar Mohanty, Saroj Kumar Lenka, "Efficient Image Mining Technique for Classification of Mammograms to Detect Breast Cancer" Special Issue of IJCCT Vol. 2 Issue 2, 3, 4; 2010 for International Conference [ICCT-2010], 3rd-5th December 2010

[7] Aura Conci., Everest Mathias M. M. Castro "Image Mining By Color Content "

[8] V. Balamurugan and P. Anandhakumar "Neuro-Fuzzy Based Clustering Approach For Content Based Image Retrieval Using 2DWavelet Transform" International Journal of Recent Trends in Engineering, Vol. 1,No.1, May 2009

[9] Bangpeng Yao, Juan Carlos Niebles, Li Fei-Fei "Mining Discriminative Adjectives and Prepositions for Natural Scene Recognition "

[10] Bingbing Ni, Zheng Song, Shuicheng Yan "Web Image Mining Towards Universal Age Estimator"

[11] R. Brown, B. Pham "Image Mining and Retrieval Using Hierarchical Support Vector Machines "Proceedings of the 11th International Multimedia Modelling Conference (MMM'05) 1550-5502/05 © 2005 IEEE

[12] Carlos Ordonez and Edward Omiecinski "Image Mining: A New Approach for Data Mining"1998

[13] Chu-Hong Hoi, Michael R. Lyu "Web Image Learning for Searching Semantic Concepts in Image Databases"

[14] Chu-Hui Lee and Meng-Ting Liu "Color Image Segmentation For Apple Grading" supported by National Science Council under Grant NSC 97-2218-E-324-004

[15] Chunjie Zhang, Jing Liu, Hanqing Lu, Songde Ma "Web Image Mining Using Concept Sensitive Markov Stationary Features" 978-1-4244-4291-1/09 ©2009 IEEE ICME 2009

[16] Dipesh Dugar M , Dinesh Kumar Jain N , Raaja Sarabhoje Gopal , Haritha Ravi , Baskaran Alagappan "Object Recognition of Image Sets Using Discriminative Canonical Correlations" Preceedings of the International MultiConference of Engineers and Computer Scientists 2010 Vol I, IMECS 2010.Hong Kong.

[17] J. Fernandez, N. Miranda, R. Guerrero, F. Piccoli "Appling Parallelism in Image Mining" March 26, 2007

[18] R. Guadagnin*a*, L. Santana*b*, E. Fernedaa, and H. Prado*a* "Improving Image Mining through Geoprocessing" *ISSN 1054_6618, Pattern Recognition and Image Analysis, 2010, Vol. 20, No. 1, pp. 81–85. © Pleiades Publishing, Ltd., 2010.*

[19] Hassan H. Malik "iARM: Image Association Rule Mining Language"2005.

[20] T. Iwaszko, Mahmoud Melkemi and L. Idoumghar "A Geometric Data Structure Applicable to Image Mining and Retrieval "

[21] L. Jeba Sheela, Dr. V.Shanthi, "Image Mining Techniques for Classification and Segmentation of Brain MRI Data" 2005 – 2007 JATIT.

[22] Jen-Hao Hsiao, Chu-Song Chen and Ming-Syan Chen "A Novel Language-Model-Based Approach for Image Object Mining and Re-Ranking" 2008 Eighth IEEE International Conference on Data Mining

[23] Ji Zhang, Wynne Hsu and Mong Li Lee "An Information-Driven Framework for Image Mining" Database and Expert Systems Applications in Computer Science, 2001, Volume 2113/2001, 232-242, DOI: 10.1007/3-540-44759-8_24

[24] Ji Zhang Wynne Hsu Mong Li Lee " Image Mining: Issues, Frameworks and Techniques"

[25] S. John Peter "Minimum Spanning Tree-based Structural Similarity Clustering for Image Mining with Local Region Outliers" International Journal of Computer Applications (0975 – 8887) Volume 8– No.6, October 2010

[26] Dr. C. Jothi Venkateswaran, S. Murugan, Dr. N. Radhakrishnan "An Useful Information Extraction using Image Mining Techniques from Remotely Sensed Image (RSI)" Dr.C.Jothi Venkateswaran et al / (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 08, 2010, 2577-2580

[27] A.Kannan, Dr.V.Mohan, Dr.N.Anbazhagan "An Effective Method of Image Retrieval using Image Mining Techniques" The International journal of Multimedia & Its Applications (IJMA) Vol.2, No.4, November 2010

[28] Keiji Yanai "Web Image Mining: Can We Gather Visual Knowledge for Image Recognition from the Web?"

[29] Keiji Yanai "Web Image Mining toward Generic Image Recognition"

[30] Keiji Yanai "Automatic Web Image Selection with a Probabilistic Latent Topic Model"

[31] Keiji Yanai "Probabilistic Web Image Gathering"

[32] Keiji Yanai, "Image Collector III: A Web Image-Gathering System with Bag-of-Keypoints" , May 2007, ACM 978-1-59593-654-7/07/0005

[33] Keiji Yanai, Kobus Barnard "Finding Visual Concepts by Web Image Mining"
[34] Kun-Che Lu And Don-Lin Yang "Image Processing and Image Mining using Decision Trees" Journal Of Information Science And Engineering 25, 989-1003 (2009)
[35] Lukasz Kobylinski and Krzysztof Walczak "Color Mining of Images Based on Clustering" Proceedings of the International Multiconference on Computer Science and Information Technology pp. 203–212 ISSN 1896-7094 2007 PIPS
[36] T. H. Manjula Devi, H.S. Manjunatha Reddy, K. B. Raja, K. R Venugopal and L. M. Patnaik "Detecting Original Image Using Histogram, DFT and SVM" International Journal of Recent Trends in Engineering Vol. 1, No. 1, May 2009
[37] Manoranjan Dash, Deepak Kolippakkam "Automatic View Selection: An Application to Image Mining"
[38] Mohamed Y. Eldib and Hoda M. Onsi "Web Image Mining Age Estimation Framework" ICGST-GVIP Journal, Volume 11, Issue 1, March 2011
[39] Dr.V.Mohan and A.Kannan "Color Image Classification and Retrieval using Image mining Techniques" International Journal of Engineering Science and Technology Vol. 2(5), 2010, 1014-1020.
[40] Nick Morsillo, Chris Pal, Randal Nelson "Mining the Web for Visual Concepts Univ. of Rochester Comp. Sci. Dept. Technical Report TR-2008-931 March 18, 2008
[41] K Perumal and R Bhaskaran "Supervised Classification Performance Of Multispectral Images" Journal Of Computing, Volume 2, Issue 2, February 2010, ISSN 2151-9617
[42] Peter Stanchev "Using Image Mining For Image Retrieval"
[43] Pham, N.K. Morin, A. Gros, P. "CAViz, an interactive graphical tool for image mining "June 2008 IEEE, pages 371 - 376 ISSN: 1330-1012 ISBN: 978-953-7138-12-7
[44] P. Rajendran, M.Madheswaran "Hybrid Medical Image Classification Using Association Rule Mining with Decision Tree Algorithm" Journal Of Computing, Volume 2, Issue 1, January 2010, ISSN 2151-9617
[45] P. Rajendran, M. Madheswaran " Novel Fuzzy Association Rule Image Mining Algorithm for Medical Decision Support System" 2010 International Journal of Computer Applications (0975 - 8887) Volume 1 – No. 20
[46] Rajshree S. Dubey "Image Mining using Content Based Image Retrieval System" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 07, 2010, 2353-2356
[47] Rong Zhu, Min Yao and Yiming Liu, "Image Classification Approach Based on Manifold Learning in Web Image Mining "Advanced Data Mining and Applications 2009, Volume 5678/2009, 780-787, DOI: 10.1007/978-3-642-03348-3_83
[48] Rosalina Abdul Salam, Abdullah Zawawi Hj. Talib, "Feature Extraction In Automatic Shape Recognition System"
[49] Ruhan He Wei Zhan "Multi-Modal Mining in Web image retrieval"Computational Intelligence and Industrial Applications, 2009. PACIIA 2009. IEEE Issue Date: 28-29 Nov. 2009 Volume: 2 On page(s): 425 - 428 ISBN: 978-1-4244-4606-3
[50] Dr. Sanjay Silakari, Dr. Mahesh Motwani and Manish Maheshwari "Color Image Clustering using Block Truncation Algorithm" IJCSI International Journal of Computer Science Issues, Vol. 4, No. 2, 2009 ISSN (Online): 1694-0784 ISSN (Print): 1694-0814
[51] Sangkyum Kim, Xin Jin, Jiawei Han "DisIClass: Discriminative Frequent Pattern-Based Image Classification"
[52] Shah, V.P. Younan, N.H. Durbha, S.S. King, R.L. "Feature Identification via a Combined ICA–Wavelet Method for Image Information Mining "Geoscience and Remote Sensing Letters, IEEE Issue Date: Jan. 2010 Volume: 7 Issue: 1 On page(s): 18 - 22 ISSN: 1545-598X
[53] Seong-Yong Hong "Web Image Information Mining System Using Uup For E-Business Intelligence" Issues in Information Systems Volume X, No. 2, 2009
[54] Rupali Sawant "Image Mining Using Texture and Shape Feature" (IJCSE) International Journal on Computer Science and Engineering Vol. 02, No. 09, 2010, 2808-2815
[55] S.P. Victor, S. John Peter "A Novel Minimum Spanning Tree Based Clustering Algorithm for Image Mining" European Journal of Scientific Research ISSN 1450-216X Vol.40 No.4 (2010), pp.540-546
[56] Dr.S.P.Victor, V.Narayani and S.Rajkumar "Efficient Retrieval Techniques For Images Using Enhanced Univariate Transformation Approach" International Journal of Engineering Science and Technology Vol. 2(8), 2010, 3762-3768
[57] Vitorino Ramos, Fernando Muge "Image Colour Segmentation by Genetic Algorithms"
[58] Wichian Premchaiswadi, Anucha Tungkatsathan "On-line Content-Based Image Retrieval System using Joint Querying and Relevance Feedback Scheme" Wseas Transactions On Computers Issue 5, Volume 9, May 2010 ISSN: 1109-2750
[59] Wynne Hsu, Mong Li Lee, Kheng Guan Goh "Image Mining in IRIS: Integrated Retinal Information System"
[60] Wynne Hsu, Mong Li Lee and Ji Zhang "Image Mining: Trends and Developments" Journal of Intelligent Information Systems Volume 19, Number 1, 7-23, DOI: 10.1023/A:1015508302797
[61] Zheng Chen, Liu Wenyin, Feng Zhang, Mingjing Li, Hongjiang Zhang "Web Mining for Web Image Retrieval"