

# A Study on Image Mining Methods and Techniques

Ankita Tripathi<sup>1</sup>, Hitesh Jangir<sup>2</sup>

M.Tech Student, Dept. of C.S., Mody University of Science and Technology, CET, Lakshmanagarh, Rajasthan, India<sup>1</sup>

Assistant Professor, Dept. of C.S., Mody University of Science and Technology, CET, Lakshmanagarh, Rajasthan, India<sup>2</sup>

**ABSTRACT:** Advancement in image accretion and tremendous growth in significantly gigantic image database led to the requirement of image mining. The process of image mining is all about to unfold and extirpate knowledge relevant to images. Numerous techniques are developed in previous many researches and thereafter acknowledge useful facts and information for the requirements of individual and organization, but it still requires further development. It is an incorporative venture that magnetizes dexterity in image retrieval and processing, mining data, expert system learning, database and artificial intelligence. Analogous to data mining task, image mining task can also be assorted into classification, clustering, association rules and characterization. In the paper, we will explore image mining method, issues, reassess image mining techniques based on various neural network and association rules. We will also pinpoint further research dimensions in the area image mining at the end of paper.

**KEYWORDS:** Image Mining, Data Mining, Image Retrieval, Image Indexing, Object Recognition, Image Classification and Clustering, Association Rule Mining.

## I. INTRODUCTION

With the boom of digitalization in every sector, tremendous amount of digital data is stored every day. Digital data is diversified in various forms of text, audio and video. Nowadays, numerous deviant databases like datasets of images are used for mining. The process of extrication of internal knowledge, data relationship among various images and various patterns that are not explicitly stored in images and many ideas from computer vision are used, image processing and retrieval, mining data, expert system learning, databases and Artificial intelligence is known as Image mining.

Image mining process comprised of following steps as shown in fig.1.: pre-processing, transformations, feature extraction, mining significant patterns and features, evaluation, interpretation and finally extracting knowledge.

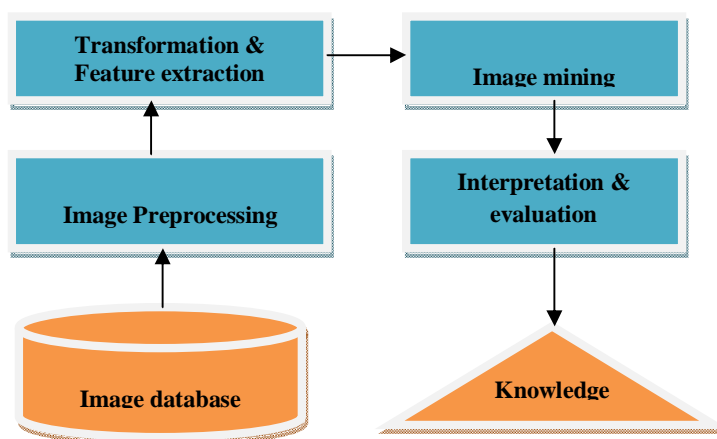


Fig. 1 The image mining process



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

*Image Preprocessing:* According to the users and application requirements, the preprocessing of image data can be done at region level and edge level. Proximate features retrieval and extrication should be performed either automatically or manually so that the image can be best captured. Preprocessing of images is performed to improve quality.

*Image Transformation and Feature Extraction:* Preprocessed images transformed in varying ways and features are extricated to produce the useful and desired features from image content. Shapes, colour, texture the traits that are generally used for mining to extract features.

*Image Mining:* The extricated features than used for mining using various mining techniques to discover significant patterns. Image mining is not just to apply currently existing data mining techniques or algorithms, it is much more than that. There difference between image and relation database and they are-

- 1) In image database, the data values are not relatively significant while in relational database, the data values are significant.
- 2) In image database, it is critical to interpret implicit spatial information as the position is dependent and interrelated while in relational database there is no difficulty in such interpretation because position is independent.

*Interpretation, Evaluation and Knowledge discovery from images:* After mining, patterns are obtained and these patterns finally evaluate and interpret the knowledge that is required. The knowledge retrieved can be used by individual or organization for various purposes to make predictions and profitable output further

## II. IMAGE MINING ISSUES

Image mining till today, is not such a culminated field and is still under examination and development and there is a great scope in area. It differs from low level processing techniques as image mining targets on enormously large database to extricate interesting pattern.

Conceivably image mining and pattern recognition are synonyms, is a big misconception. They are different in manner that pattern recognition is about extracting significant pattern whereas image mining is introducing significant pattern from the knowledge that we have. Other than these aspects, few more issues are to be resolved and they are-

- 1) To retain various information like contextual, spatial & important image characteristics in the representation scheme, how an image pattern can be represented.
- 2) For the meaningful visualization of the discovered patterns in the mining, what are the important images features to be used.
- 3) In a visually efficient environment, how the patterns extricated are represented in front of users.

## III. METHOD FOR IMAGE MINING

The algorithms required to perform image mining follows four major steps:

1. *Extracting feature:* Segment images into various identifiable regions.
2. *Object identification and creation:* Based on preprocessing algorithm objects of one image are compared to other image objects and mark each object with an id and is known as preprocessing algorithm.
3. *Create complementary images:* All those images that contains objects which can be identified by objects id are retrieved to evaluate the association rules.
4. *To produce object association rules data mining algorithm is applied:* To do this, we select number of images that make a collection of specified area. After collection, the objects are extracted from images and then images are indexed in the database. As images are indexed then desirable features are extracted and finally mined and prediction for knowledge discovery.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

## III. EXISTING TECHNIQUES OF IMAGE MINING

Various techniques that are used and how they are implemented in image mining can be explained by object recognition, image retrieval, classification, clustering, association rule, neural network. Fig.2, shows image mining techniques.

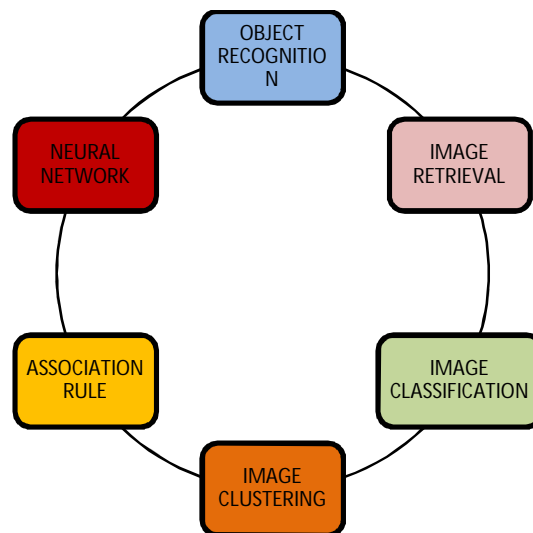


Fig. 2 Image mining techniques

*Object Recognition:* It has been a focused technique for image processing among various active research fields. Object models and object recognition system find objects from real applied world from images that belong to image dataset. This is the main task in image mining. Useful information extrication can only be implemented when few objects are identified and recognized by machine. The problem is referred as a supervised labeling problem based on object models of known objects.

*Image Retrieval:* Image mining demands retrieval of images based on some pre-described features or requirements. The features or requirement can be categorized as-

**LEVEL 1-** Considering primary features like shape, color, texture to retrieve images.

**LEVEL 2-** Considering logical features like objects of given type or separate objects.

**LEVEL 3-** Considering high level reasoning by retrieving attributes based on purpose or scenes depicted.

On the other hand, image retrieval can broadly follow criteria-

- 1) *Text based Image Retrieval:* The process is based on image retrieval through descriptions. Description can be like size of images (20 KB), type of images, date and time of image when it was captured, identity of the owner, keywords mapping of the image. The description for each image is given manually, because the automatic keyword description for the images is not possible without information regarding visuality and feature to be extracted. Based on these contents, required image features are mined and indexed.
- 2) *Query Based Image Retrieval:* Based on visual features and index of query image, images from the database are retrieved that matches the query image. While comparing two images, similarities between them are measured by computing euclidean distance between the feature vectors and k images are retrieved.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

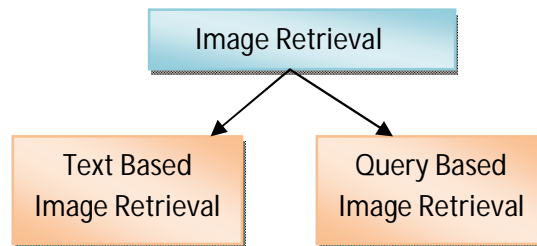


Fig. 3 Type of Image Retrieval

*Image Indexing:* The goal of image indexing is to extract matching images from an image dataset from a given input query image. Each image has its varying and unique feature. Therefore, image mining can be done by comparing the features, which are retrieved from the images. The similarity measure based on images may be features such as color, intensity, texture, position and shape, and various other image attributes. Image indexing techniques can be classified as:

- 1) *Textual:* Textual image indexing is a very efficient and easy technique for image retrieval. Basically, the approach emphasizes on keywords given for a specific image. Text data present in images contain useful information for, keyword additions, Standard titles of content, caption indexing, etc. There are differences in text style, text size, text orientation and alignment and all these can be used for image indexing[16]. Fig. 4, shows architecture of text classification.

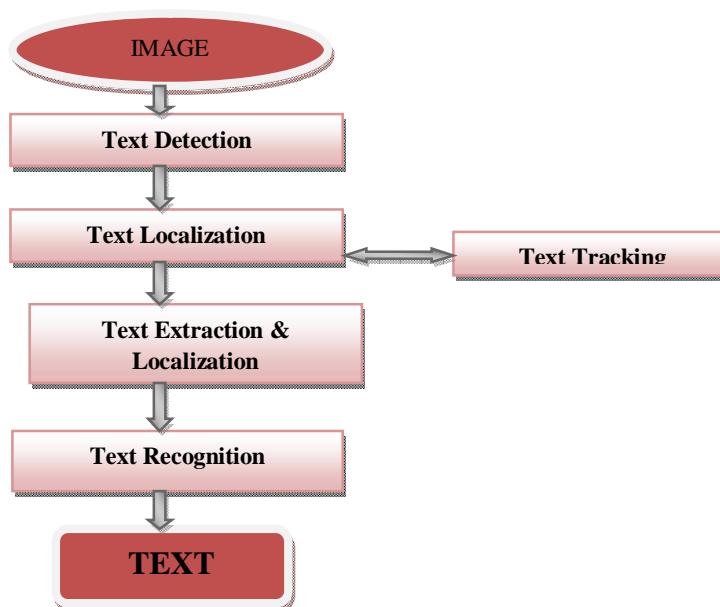


Fig. 4 Architecture of Text Extraction

- 2) *Content-based:* Content based image retrieval (CBIR) technique indexes images based on their content like color, shape, texture etc. ,image retrieval in CBIR covers image processing and information retrieval mechanism. The CBIR system performs 2 main tasks:
  - a) *Feature Extraction:* where feature vector (set of features) are generated to accurately represent the content of each image in the database.

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

- b) *Similarity Measurement*: where distance between the query image and other images in the stored database based on feature vectors is used to retrieve the “closest” images.[15]

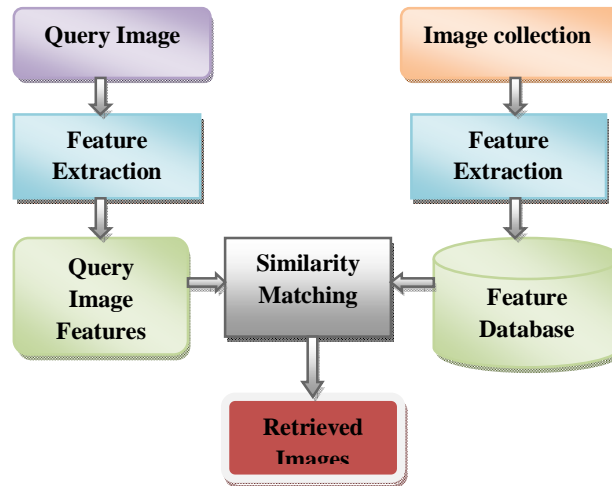


Fig. 5 CBIR Process

*Image Classification and Image Clustering*: Supervised and unsupervised classification of images into groups respectively is basically performed on the basis of classification and clustering. The main motive for performing classification or clustering on images is to retrieve knowledge that the users demand from the image group stored in the image database associated with the image. In supervised classification, the pool or set of labeled images which are used to label newly unlabeled images that are encountered. In image clustering, the main motive is to combine a number of unlabeled images into meaningful groups called clusters. Based on image content, these clusters are formed without any prior knowledge.

*Association Rule Mining*: The association rule mining is used as an efficient tool for pattern recognition in knowledge discovery and data mining. Its main motive is to extricate useful information from large data sets. By implementing association rules retrieved from images, the data from images can be easily analyzed, and the information required for classification can be obtained. Association rule can be expressed as  $X \Rightarrow Y$ , where  $X$  and  $Y$  can be any discrete entity. In image database,  $X$  and  $Y$  can be some feature elements mined from images. The meaning of  $X \Rightarrow Y$  is: Assuming an image database  $D$ , for each image  $I \in D$ ,  $X \Rightarrow Y$  expresses that whenever an image  $I$  contains  $X$  then  $I$  will probably holds  $Y$ . The support of association rule is defined as the probability  $p(X \subseteq I, Y \subseteq I)$  and the confidence of association rule is defined as the conditional probability  $p(X \subseteq I | Y \subseteq I)$ . [9]

*Neural Networks*: Neural Networks are used in image retrieval and image mining. They are basically used for computational through simple processing units called neurons. These neurons are organized into layers with full or partial connections. The main function of neuron is to receive input values from its neighbors, evaluate an output based on its weighted input and send the output to its neighbors. Neural networks are useful for pattern recognition, prediction and fault tolerance.

## V. IMAGE MINING REAL-WORLD APPLICATION

In this section, we will describe real-world application of image mining. Image mining can be used in various image retrieval based application. Some of them are described -



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 4, April 2016

*Monsoon and Rainfall Prediction:* The prediction of rainfall is one of the major studies in field of image mining. In India, where agriculture is major occupation which is dependent on rainfall, the time and amount of rainfall holds high importance and thus mining useful data relevant to this is on high demand.

*Satellite Image Mining:* The satellite images contain information for weather forecasting and early prediction of different natural calamities like typhoon, hurricanes etc. Satellite Images Contains various parameter which can be used for information retrieval. Other parameters like humidity, linear cloud, typhoon can be extracted from satellite images to get a useful and efficient knowledge.[11]

*Textile Image Retrieval Using Color as parameter:* All textile industries aim to produce large scale of textile depends mainly on designs and quality of the dresses produced. Every day, numerous textile images are being generated such as images of shirts, jeans, t-shirts and sarees.. Images play an important role as a picture is worth thousand words in the field of textile design and marketing. A retrieving of images needs special concepts such as image annotation, context, and image content and image values. [12]

## VI.CONCLUSION

Image mining is a progressive field that retrieves images that best matches from the images in the image dataset, on the basis of query image. The process model represents about storing the various type/level of information already present in image dataset and pointing the issues of analyzing and retrieving useful patterns/knowledge from each type/level. In this paper, we have pinpointed the need for image mining due to boom in growing image datasets. We have also pointed the unique features in the area of image mining that brings a whole new set of challenging research issues to be resolved. We have also discussed techniques that are used in image mining, namely, object recognition, image retrieval, image indexing, image classification and clustering, association rule mining and neural network. Finally, the applications of image mining that have a very dynamic and deep future scope. Many researchers have done lot of research in described applications but there is lack of efficiency that is required at current stage.

## REFERENCES

- [1] T.Y.Gajjar, N.C.Chauhan,"A review on image mining frameworks and techniques" International journal of computer science and information technologies, vol3 (3), 2012.
- [2] Shaikh Nikhat Fatma," Image Mining Method and Frameworks", International Journal of Computational Engineering Research (ijceronline.com), Vol. Issue. 8, 2012.
- [3] S. Gibson Intelligent Mining in Image Databases, with Applications to Satellite Imaging and Web Search, Data Mining and Computational Intelligence, Springer-Verlag, Berlin, 2001.
- [4] Prof. Dr.Hilal, M. Yousif, Dr.Abdul- Rahman Al-Hussaini, Mohammad A. Al-Hamami," Using Image Mining to Discover Association Rules between Image Objects".
- [5] Mahip M.Bartere , Dr.Prashant and R.Deshmukh, "An Efficient Technique using Text & Content Base Image Mining Technique for Image Retrieval", International Journal of Engineering Research and Applications (IJERA) Vol. 2, Issue 1, pp.734-739,Jan-Feb 2012.
- [6] Ji Zhang Wynne Hsu Mong Li Lee," Image Mining: Trends and Developments", School of Computing, Singapore, 2012.
- [7] A.Kannan, Dr.V.Mohan and 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.
- [8] Prabhjeet Kaur, Kamaljit Kaur ,"Review of Different Existing Image Mining Techniques" International Journal of Advanced Research in Computer Science and Software Engineering, Volume 4, 2014.
- [9] J.Han and M.Kamber. Data Mining: Concepts and Techniques, Morgan Kaufmann, USA, 2001.
- [10] Dinu John, K. K. Sindhu, B. B. Meshram, "Two Stage Data Mining Technique for Fast Monsoon Onset Prediction", International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 2,2012.
- [11] Jay Narayan Thakre, Divakar Singh, "A Survey on Knowledge Discovery from the Satellite Image Using Association Rule Mining", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, 2013.
- [12] S.Balan, T.Devi," Design and Development of an Algorithm for Image Clustering In Textile Image Retrieval Using Color Descriptors", of International Journal of Computer Science, Engineering and Applications (IJCEA) Vol.2, No.3, 2012.
- [13] Nishchol Mishra and Dr. Sanjay Silakari "A Perspective Image Mining in the Context of Content Based Image Retrieval" IJCSI International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012.
- [14] Md. Farooque,"Image Indexing and Retrieval", DRTC Workshop on Digital Libraries: Theory and Practice March DRTC, 2002.



ISSN(Online) : 2320-9801  
ISSN (Print) : 2320-9798

# International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

**Vol. 4, Issue 4, April 2016**

## **BIOGRAPHY**



Ms. Ankita Tripathi received Bachelors of Technology from Shri Ram Murti Smarak Women's College of Engineering & Technology, Bareilly, U.P in 2014 and pursuing Masters of Technology in Computer Science from Mody University of science & Technology, Lakshmangarh, Sikar. Her research interests are in the fields of Data Mining, Database, Information retrieval, Image mining.



Mr. Hitesh Jangir received Masters of Engineering in Computer Science from University Of Bedfordshire, UK and pursuing PhD from Mody University of science & Technology, Lakshmangarh, Sikar. He is presently working as an Assistant Professor in Department of Computer Science, Mody University of science & Technology, Lakshmangarh, Sikar. His current research includes Data mining, Image mining, online database application, Wireless ad hoc networks, Intelligent Agents, Project Management.