

# A Survey on Image Mining Techniques for Image Retrieval

**J. Priya**

Research Scholar, Department of Computer Science,  
NGM College, Pollachi, India.

**Dr. R. Manicka Chezian**

Associate Professor, Department of Computer Science,  
NGM College, Pollachi, India.

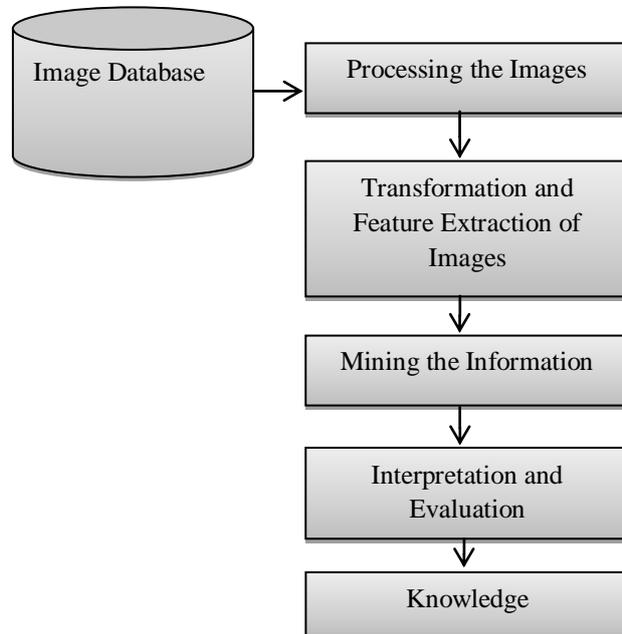
**Abstract** - Image mining is an expansion of data mining in the field of image processing. Image mining handles with the hidden knowledge extraction, image data association and additional patterns which are not clearly gathered in the images. It is an interdisciplinary field that incorporates techniques like Image Processing, Data Mining, Machine Learning, Database and Artificial Intelligence. The most important function of the mining is to generate all significant patterns without prior information of the patterns. This paper presents a survey of various image mining techniques.

**Keywords** - Data Mining, Image Mining, Feature Extraction, Image Retrieval.

## I. INTRODUCTION

Image mining is the process of searching and discovering valuable information and knowledge in large volumes of data. Fig. 1 shows the Typical Image Mining Process. Some of the methods used to gather knowledge are, Image Retrieval, Data Mining, Image Processing and Artificial Intelligence. These methods allow Image Mining to have two different approaches. One is to extract from databases or collections of images and the other is to mine a combination of associated alphanumeric data and collections of images. In pattern recognition and in image processing, feature extraction is a special form of

dimensionality reduction. When the input data is too large to be processed and it is suspected to be notoriously redundant, then the input data will be transformed into a reduced representation set of features. Feature extraction involves simplifying the amount of resources required to describe a large set of data accurately. Several features are used in the Image Retrieval system. The popular amongst them are Color features, Texture features and Shape features.



**Fig.1. Image Mining Process**

## II. FEATURE EXTRACTION

Feature selection is an important problem in object detection, and demonstrates that Genetic

Algorithm (GA) provides a simple, general and powerful framework for selecting good sets of features, leading to lower detection error rates. Zehang Sun et al., [13] discuss to perform Feature Extraction using popular method of Principle Component Analysis (PCA) and Classifications using Support Vector Machines (SVMs). GAs are capable of removing detection-irrelevant Features. The methods are on two difficult object detection problems, Vehicle detection and Face Detections. The methods boost the performance of both systems using SVMs for Classification.

Patricia G. Foschi [10] discuss that Feature selection and extraction is the pre-processing step of Image Mining. Obviously this is a critical step in the entire scenario of Image Mining. The approach to mine from Images is to extract patterns and derive knowledge from large collections of images which mainly deals with identification and extraction of unique features for a particular domain. Though there are various features available, the aim is to identify the best features and thereby extract relevant information from the images.

Increasing amount of illicit image data transmitted via the internet has triggered the need to develop effective image mining systems for digital forensics purposes. Brown, Ross A et al., [3] discuss the requirements of digital image forensics which underpin the design of our forensic image mining system. This system can be trained by a hierarchical SVM to detect objects and scenes which are made up of components under spatial or non-spatial constraints. Bayesian networks approach used to deal with information uncertainties which are inherent in forensic work.

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; but also the innovation of image patterns that are noteworthy in a given collection of images. Fernandez. J et al., [4] show 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. The images from an image database are first pre-processed 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.

#### A. Color Feature

Image mining presents special characteristics due to the richness of the data that an image can show. Effective evaluation of the results of image mining by content requires that the user point of view is used on the performance parameters. Aura Conci et.al, [2] proposed an evaluation framework for comparing the influence of the distance function on image mining by colour. Experiments with colour similarity mining by quantization on colour space and measures of likeness between a sample and the image results have been carried out to illustrate the proposed scheme.

Lukasz Kobyliński and Krzysztof Walczak [9] proposed a simple but fast and effective method of indexing image metadatabases. The index is created by describing the images according to their color characteristics, with compact feature vectors, that represent typical color distributions. Binary Thresholded Histogram (BTH), a color feature description method proposed, to the creation of a metadatabase index of multiple image databases. The BTH, despite being a very rough and compact representation of image colors, proved to be an adequate method of

describing the characteristics of image databases and creating a metadatabase index for querying large amounts of data.

Ji Zhang, Wynne Hsu and Mong Li Lee [8] 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.

#### B. Texture Feature

The image depends on the Human perception and is also based on the Machine Vision System. The Image Retrieval is based on the color Histogram, texture. The perception of the Human System of Image is based on the Human Neurons which hold the 10<sup>12</sup> of information; the Human brain continuously learns with the sensory organs like eye which transmits the Image to the brain which interprets the Image. Rajshree S. Dubey et.al, [12] examines the State-of-art technology Image mining techniques which are based on the Color Histogram, texture of Image. The query Image is taken then the Color Histogram and Texture is taken and based on this the resultant Image is output.

Janani. M and Dr. Manicka Chezian. R [7] discusses Image mining is a vital technique which is used to mine knowledge from image. The development of the Image Mining technique is based on the Content Based Image Retrieval system. Color, texture, pattern, shape of objects and their layouts and locations within the image, etc are the basis of the Visual Content of the Image and they are indexed.

#### C. Shape Feature

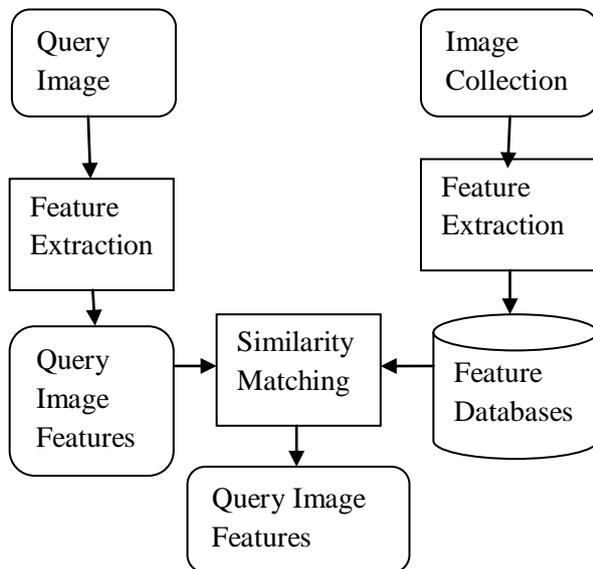
Peter Stanchev [11] proposed a new method for image retrieval using high level semantic features is proposed. It is based on extraction of low level color, shape and texture

characteristics and their conversion into high level semantic features using fuzzy production rules, derived with the help of an image mining technique. Dempster- Shafer theory of evidence is applied to obtain a list of structures containing information for the image high level semantic features. Johannes Itten theory is adopted for acquiring high level color features.

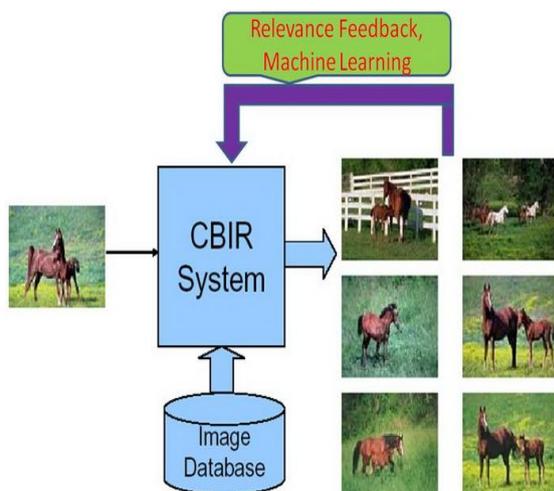
Harini. D. N. D and Dr. Lalitha Bhaskari. D [5] discuss Image Retrieval, which is an important phase in image mining, is one technique which helps the users in retrieving the data from the available database. The fundamental challenge in image mining is to reveal out how low-level pixel representation enclosed in a raw image or image sequence can be processed to recognize high-level image objects and relationships.

Hiremath. P. S and Jagadeesh Pujari [6] discuss Color, texture and shape information has been the primitive image descriptors in content based image retrieval systems. The image and its complement are partitioned into non-overlapping tiles of equal size. The features drawn from conditional co-occurrence histograms between the image tiles and corresponding complement tiles, in RGB color space, serve as local descriptors of color and texture. This local information is captured for two resolutions and two grid layouts that provide different details of the same image. An integrated matching scheme, based on most similar highest priority (MSHP) principle and the adjacency matrix of a bipartite graph formed using the tiles of query and target image, is provided for matching the images. Shape information is captured in terms of edge images computed using Gradient Vector Flow fields. Invariant moments are then used to record the shape features. The combination of the color and texture features between image and its complement in conjunction with the shape features provide a robust feature set for image retrieval. Fig.

2 and Fig. 3 show the Content Based Image Retrieval System(CBIR) Architecture and an example for it.



**Fig.2. Content Based Image Retrieval System Architecture**



**Fig.3. An Example for Content Based Image Retrieval (CBIR) System**

Anil K. Jain and Aditya Vailaya [1] propose that the Color and Shape based queries provide better performance than either of the individual feature based queries. Combination of simple features which can be easily extracted is more efficient. The speed of retrievals can be also

increased by using Branch and Bound to compute the nearest neighbour for the query image without affecting the robustness of the system.

### III. CONCLUSION

Image mining is an expansion of data mining in the field of image processing. This paper presents a survey on various image mining techniques that was proposed earlier by researcher. This overview of image mining focuses on image mining implementations, usability and challenges in various fields.

### REFERENCES

- [1]. Anil K. Jain and Aditya Vailaya, "Image Retrieval using color and shape", In Second Asian Conference on Computer Vision, pp 5-8. 1995.
- [2]. Aura Conci, Everest Mathias M. M. Castro, "Image mining by Color Content", In Proceedings of 2001 ACM International Conference on Software Engineering and Knowledge Engineering (SEKE), Buenos Aires, Argentina Jun 13-15, 2001.
- [3]. Brown, Ross A., Pham, Binh L., and De Vel, Olivier Y, "Design of a Digital Forensics Image Mining System", in Knowledge Based Intelligent Information and Engineering Systems, pp 395-404, Springer Berlin Heidelberg, 2005.
- [4]. Fernandez. J, Miranda. N, Guerrero. R and Piccoli. F, "Applying Parallelism in Image Mining", In IX Workshop de investigadores en Ciencias de la Computacion, 2007. ([www.ing.unp.edu.ar/wicc2007/trabajos/PDP/120.pdf](http://www.ing.unp.edu.ar/wicc2007/trabajos/PDP/120.pdf)).
- [5]. Harini. D. N. D and Dr. Lalitha Bhaskari. D, "Image Mining Issues and Methods Related to Image Retrieval System", International Journal of Advanced Research in Computer Science, Volume 2, No. 4, 2011.

- [6]. Hiremath. P. S and Jagadeesh Pujari, “Content Based Image Retrieval based on Color, Texture and Shape features using Image and its complement”, International Journal of Computer Science and Security, Volume (1) : Issue (4).
- [7]. Janani. M and Dr. Manicka Chezian. R, “A Survey On Content Based Image Retrieval System”, International Journal of Advanced Research in Computer Engineering & Technology, Volume 1, Issue 5, pp 266, July 2012.
- [8]. Ji Zhang, Wynne Hsu and Mong Li Lee, “An Information-Driven Framework for Image Mining” Database and Expert Systems Applications in Computer Science, pp 232 – 242, Springer Berlin Heidelberg, 2001.
- [9]. Lukasz Kobylinski and Krzysztof Walczak, “Color Mining of Images Based on Clustering”, Proceedings of the International Multi conference on Computer Science and Information Technology pp. 203–212, 2007.
- [10]. Patricia G. Foschi, Deepak Kolippakkam, Huan Liu and Amit Mandvikar, “Feature Extraction for Image Mining”, Proceedings in International workshop on Multimedia Information System, pp 103-109, 2002.
- [11]. Peter Stanchev, “Image Mining for Image Retrieval”, In Proceedings of the IASTED Conference on Computer Science and Technology, pp 214-218, 2003.
- [12]. Rajshree S. Dubey, Niket Bhargava and Rajnish Choubey, “Image Mining using Content Based Image Retrieval System”, International Journal on Computer Science and Engineering, Vol. 02, No. 07, 2353-2356, 2010.
- [13]. Zehang Sun, George Bebis, Ronald Miller, “Boosting Object Detection Using Feature Selection”, Proceedings of IEEE Conference on Advanced video and Signal Based Surveillance, pp 290-296, 2003.

**AUTHORS:**

**J.Priya** received MCA degree from Bharathiar University, Coimbatore. Currently she is doing M.Phil Degree in Computer Science at Bharathiar University, Coimbatore. Currently she is working as an Assistant Professor in Computer Science at NGM College, Pollachi, India. She has got 2 years of teaching experience. Her research interest lies in the area of image mining.



**Dr. R. Manicka chezian** received his M.Sc., degree in Applied Science from P.S.G College of Technology, Coimbatore, India in 1987. He

completed his M.S. degree in Software Systems from Birla Institute of Technology and Science, Pilani, Rajasthan, India and Ph D degree in Computer Science from School of Computer Science and Engineering, Bharathiar University,

Coimbatore, India. He served



as a Faculty of Maths and Computer Applications at P.S.G College of Technology, Coimbatore from 1987 to 1989.

Presently, he has been working as an Associate Professor of

Computer Science in N G M College (Autonomous), Pollachi under Bharathiar University, Coimbatore, India since 1989. He has published thirty papers in international/national journal and conferences: He is a recipient of many awards like Desha Mithra Award and Best Paper Award. His research focuses on Network Databases, Data Mining, Distributed Computing, Data Compression, Mobile Computing, Real Time Systems and Bio-Informatics.