# An Information-driven Framework for Image Mining

Ji Zhang Wynne Hsu Mong Li Lee

School of Computing, National University of Singapore
{zhangji, whsu, leeml}@comp.nus.edu.sg

**Abstract.** Image mining systems that can automatically extract semantically meaningful information (knowledge) from image data are increasingly in demand. The fundamental challenge in image mining is to determine how low-level, pixel representation contained in a raw image or image sequence can be processed to identify high-level spatial objects and relationships. To meet this challenge, we propose an efficient information-driven framework for image mining. We distinguish four levels of information: the Pixel Level, the Object Level, the Semantic Concept Level, and the Pattern and Knowledge Level. High-dimensional indexing schemes and retrieval techniques are also included in the framework to support the flow of information among the levels. We believe this framework represents the first step towards capturing the different levels of information present in image data and addressing the issues and challenges of discovering useful patterns/knowledge from each level.

## 1. Introduction

An extremely large number of image data such as satellite images, medical images, and digital photographs are generated every day. These images, if analyzed, can reveal useful information to the human user. Unfortunately, there is a lack of effective tools for searching and finding useful patterns from these images. Image mining systems that can automatically extract semantically meaningful information (knowledge) from image data are increasingly in demand. Image mining deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the images and between image and other alphanumeric data. It is more than just an extension of data mining to image domain. It is an interdisciplinary endeavor that draws upon expertise in computer vision, image processing, image retrieval, data mining, machine learning, database, and artificial intelligence [6]. Despite the development of many applications and algorithms in the individual research fields, research in image mining is still in its infancy. The fundamental challenge in image mining is to determine how low-level, pixel representation contained in a raw image or image sequence can be processed to identify high-level spatial objects and relationships.

In this paper, we propose an information-driven framework for image mining. We distinguish four levels of information: (1) the Pixel Level comprises the raw image information such as image pixels and the primitive image features such as color, texture, and shape; (2) the Object Level deals with object or region information based on the primitive features in the Pixel Level; (3) the Semantic Concept Level takes into consideration domain knowledge to generate high-level semantic concepts from the identified objects and regions; (4) the Pattern and Knowledge Level incorporates domain related alphanumeric data and the semantic concepts obtained from the image data to discover underlying domain patterns and knowledge. High-dimensional indexing schemes and retrieval techniques are also included in the framework to support the flow of information among the levels. This framework represents the first step towards capturing the different levels of information present in image data and addressing the question of what are the issues and work that has been done in discovering useful patterns/knowledge from each level.

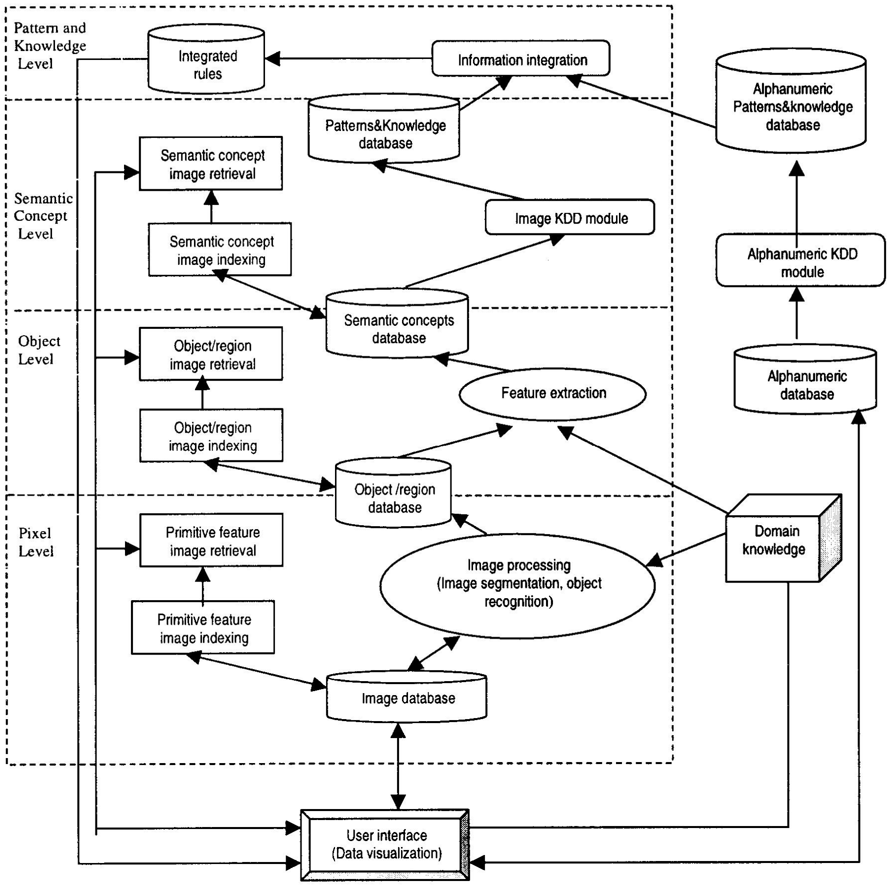
The rest of this paper is organized as follows: Section 2 presents an overview of the proposed information-driven image mining architecture. Section 3 describes each of the information level. Section 4 discusses how each of the information level can be organized and indexed. Section 5 gives the related work and we conclude in Section 6.

## 2. Information-Driven Image Mining Framework

The image database containing raw image data cannot be directly used for mining purposes. Raw image data need to be processed to generate the information that is usable for high-level mining modules. An image mining system is often complicated because it employs various approaches and techniques ranging from image retrieval and indexing schemes to data mining and pattern recognition.

Such a system typically encompasses the following functions: image storage, image processing, feature extraction, image indexing and retrieval, patterns and knowledge discovery. A number of researchers have described their image mining framework from the functional perspective [6,25,37]. While such functional-based framework is easy to understand, it fails to emphasize the different levels of information representation necessary for image data before meaningful mining can take place.

Figure 1 shows our proposed information-driven framework for image mining. There are four levels of information, starting from the lowest Pixel Level, the Object Level, the Semantic Concept Level, and finally to the highest Pattern and Knowledge Level. Inputs from domain scientists are needed to help identify domain specific objects and semantic concepts. At the Pixel Level, we are dealing with information relating to the primitive features such as color, texture, and shape. At the Object Level, simple clustering algorithms and domain experts help to segment the images into some meaningful regions/objects. At the Semantic Concept Lever, the objects/regions identified earlier are placed in the context of the scenes depicted. High-level reasoning and knowledge discovery techniques are used to discover interesting patterns. Finally, at the Pattern and Knowledge Level, the domain-specific alphanumeric data are integrated with the semantic relationships discovered from the images and further mining are performed to discovered useful correlations between the alphanumeric data and those found in the images. Such correlations discovered are particularly useful in the medical domain.




## 3. The Four Information Levels

In this section, we will describe the four information levels in our proposed framework. We will also discuss the issues and challenges faced in extracting the required image features and useful patterns and knowledge from each information level.

### 3.1 Pixel Level

The Pixel Level is the lowest layer in an image mining system. It consists of raw image information such as image pixels and primitive image features such as color, texture, and edge information.

**Color** is the most widely used visual feature. Color is typically represented by its RGB values (three 0 to 255 numbers indicating red, green, and blue). The distribution of color is a global property that does not require knowledge of how an image is composed of component objects. Color histogram is a structure used to store the proportion of pixels of each color within an image. It is invariant to under translation and rotation about the view axis and change only slowly under change of view angle, change in scale, and occlusion [32]. Subsequent improvements include the use of cumulative color histogram [31], and spatial histogram intersection [30].

**Texture** is the visual pattern formed by a sizable layout of color or intensity homogeneity. It contains important information about the structural arrangement of surfaces and their relationship to the surrounding environment [27]. Common representations of texture information include: the co-occurrence matrix representation [12], the coarseness, contrast, directionality, regularity, roughness measures [33], the use of Gabor filter [22] and fractals [17]. [22] develop a texture thesaurus to automatically derive codewords that represent important classes of texture within the collection.

**Edge information** is an important visual cue to the detection and recognition of objects in an image. This information is obtained by looking for sharp contrasts in nearby pixels. Edges can be grouped to form regions.

Content-based image retrieval focus on the information found at the Pixel Level. Researchers try to identify a small subset of primitive features that can uniquely distinguish images of one class from another class. These primitive image features have their limitation. In particular, they do not have the concept of objects/regions as perceived by a human user. This implies that the Pixel Level is unable to answer simple queries such as “retrieve the images with a girl and her dog” and “retrieve the images containing blue stars arranged in a ring”.

### 3.2 Object Level

The focus of the Object level is to identify domain-specific features such as objects and homogeneous regions in the images. While a human being can perform object recognition effortlessly and instantaneously, it has proven to be very difficult to implement the same task on machine. The object recognition problem can be referred to as a supervised labeling problem based on models of known objects. Given a target image containing one or more interesting objects and a set of labels corresponding to a set of models known to the system, what object recognition does is to assign correct labels to regions, or a set of regions, in the image. Models of known objects are usually provided by human input a priori. An object recognition module consists of four components: model database, feature detector, hypothesizer and hypothesis verifier [15]. The model database contains all the models known to the system. The models contain important features that describe the objects. The detected image primitive features in the Pixel Level are used to help the hypothesizer to assign likelihood to the objects in the image. The verifier uses the models to verify the hypothesis and refine the object likelihood. The system finally selects the object with the highest likelihood as the correct object.

To improve the accuracy of object recognition, image segmentation is performed on partially recognized image objects rather than randomly segmenting the image. The techniques include: “characteristic maps” to locate a particular known object in images [16], machine learning techniques to generate recognizers automatically [6], and use a set of examples already labeled by the domain expert to find common objects in images [10]. Once the objects within an image can be accurately identified, the Object Level is able to deal with queries such as “Retrieve images of round table” and “Retrieve images of birds flying in the blue sky”. However, it is unable to answer queries such as “Retrieve all images concerning Graduation ceremony” or “Retrieve all images that depicts a sorrowful mood.”

### 3.3 Semantic Concept Level

While objects are the fundamental building blocks in an image, there is "semantic gap between the Object level and Semantic Concept level. Abstract concepts such as happy, sad, and the scene information are not captured at the Object level. Such information requires domain knowledge as well as state-of-the-art pattern discovery techniques to uncover useful patterns that are able to describe the scenes or the abstract concepts. Common pattern discovery techniques include: image classification, image clustering, and association rule mining.

## (a) Image classification

Image classification aims to find a description that best describe the images in one class and distinguish these images from all the other classes. It is a supervised technique where a set of labeled or pre-classified images is given and the problem is to label a new set of images. This is usually called the classifier. There are two types of classifiers, the parametric classifier and non-parametric classifier. [7] employ classifiers to label the pixels in a Landset multispectral scanner image. [37] develop a MM-Classifier to classify multimedia data based on given class labels. [36] proposed IBCOW (Image-based Classification of Objectionable Websites) to classify websites into objectionable and benign websites based on image content .

**(b) Image clustering**

Image clustering groups a given set of unlabeled images into meaningful clusters according to the image content without a priori knowledge [14]. Typical clustering techniques include hierarchical clustering algorithms, partitioning algorithms, nearest neighbor clustering, and fuzzy clustering. Once the images have been clustered, a domain expert is needed to examine the images of each cluster to label the abstract concepts denoted by the cluster.

**(c) Association rule mining**

Association rule mining aims to find items/objects that occur together frequently. In the context of images, association rule mining is able to discover that when several specific objects occur together, there is a high likelihood of certain event/scene is being described in the images. An association rule mining algorithm works in two steps. The first step finds all large itemsets that meet the minimum support constraint. The second step generates rules from all the large itemsets that satisfy the minimum confidence constraint. [25] present an algorithm that uses association rule mining to discover meaningful correlations among the blobs/regions that exists in a set of images. [37] develop an MM-Associator that uses 3-dimensional visualization to explicitly display the associations in the Multimedia Miner prototype.

With the Semantic Concept Level, queries involving high-level reasoning about the meaning and purpose of the objects and scene depicted can be answered. Thus, we will able to answer queries such as: "Retrieve the images of a football match" and "Retrieve the images depicting happiness". It would be tempting to stop at this level. However, careful analysis reveals that there is still one vital piece of missing information – that of the domain knowledge external to images. Queries like: "Retrieve all medical images with high chances of blindness within one month", requires linking the medical images with the medical knowledge of chance of blindness within one month. Neither the Pixel level, the Object level, nor the Semantic Concept level is able to support such queries.

### 3.4 Pattern and Knowledge Level

To support all the information needs within the image mining framework, we need the fourth and final level: the Pattern and Knowledge Level. At this level, we are concerned with not just the information derivable from images, but also all the domain-related alphanumeric data. The key issue here is the integration of knowledge discovered from the image databases and the alphanumeric databases. A comprehensive image mining system would not only mine useful patterns from large collections of images but also integrate the results with alphanumeric data to mine for further patterns. For example, it is useful to combine heart perfusion images and the associated clinical data to discover rules in high dimensional medical records that may suggest early diagnosis of heart disease.

IRIS, an Integrated Retinal Information System, is designed to integrate both patient data and their corresponding retinal images to discover interesting patterns and trends on diabetic retinopathy [13]. BRAin-Image Database is another image mining system developed to discover associations between structures and functions of human brain [23]. The brain modalities were studied by the image mining process and the brain functions (deficits/disorders) are obtainable from the patients' relational records. Two kinds of information are used together to perform the functional brain mapping.

By ensuring a proper flow of information from low level pixel representations to high level semantic concepts representation, we can be assured that the information needed at the fourth level is derivable and that the integration of image data with alphanumeric data will be smooth. Our proposed image mining framework emphasizes the need to focus on the flow of information to ensure that all levels of information needs have been addressed and none is neglected.

## 4. Indexing of Image Information

While focusing on the information needs at various levels, it is also important to provide support for the retrieval of image data with a fast and efficient indexing scheme. Indexing techniques used range from standard methods such as signature file access method and inverted file access method, to multi-dimensional methods such as K-D-B tree [26], R-tree [11], R*-tree [3] and R+-tree [29], to high-dimensional indexes such as SR-tree [18], TV-tree [20], X-tree [4] and iMinMax [24].

Searching the nearest neighbor is an important problem in high-dimensional indexing. Given a set of n points and a query point Q in a d-dimensional space, we need to find a point in the set such that its distance from Q is less than, or equal to, the distance of Q from any other points in the set [19]. Existing search algorithms can be divided into the following categories: exhaustive search, hashing and indexing, static space partitioning, dynamic space partitioning, and randomized algorithms. When the image database to be searched is large and the feature vectors of images are of high dimension (typically in the order of $10^2$), search complexity is high. Reducing the dimensions may be necessary to prevent performance degradation. This can be accomplished using two well-known methods: the Singular Value Decomposition (SVD) update algorithm and clustering [28]. The latter realizes dimension reduction by grouping similar feature dimensions together.

Current image systems retrieve images based on similarity. Euclidean measures may not effectively simulate human perception of a certain visual content. Other similarity measures such as histogram intersection, cosine, correlation, etc., need to be utilized. One promising approach is to first perform dimension reduction and then use appropriate multi-dimensional indexing techniques that support Non-Euclidean similarity measures [27]. [11] develop an image retrieval system on Oracle platform using multi-level filters indexing. The filters operate on an approximation of the high-dimension data that represents the images, and reduces the search space so that the computationally expensive comparison is necessary for only a small subset of the data. [12] develop a new compressed image indexing technique by using compressed image features as multiple keys to retrieve images.

Other proposed indexing schemes focus on specific image features. [21] present an efficient color indexing scheme for similarity-based retrieval which has a search time that increases logarithmically with the database size. [34] propose a multi-level R-tree index, called the nested R-trees for retrieving shapes efficiently and effectively. With the proliferation of image retrieval mechanisms, a performance evaluation of color-spatial retrieval techniques was given in [35] which serves as guidelines to select a suitable technique and design a new technique.

## 5. Related work

Several image mining systems have been developed for different applications. The MultiMediaMiner [37] mines high-level multimedia information and knowledge from large multimedia database. [8] describes an intelligent satellite mining system that comprises of two modules: a data acquisition, preprocessing and archiving system which is responsible for the extraction of image information, storage of raw images, and retrieval of image, and an image mining system, which enables the users to explore image meaning and detect relevant events. The Diamond Eye [6] is an image mining system that enables scientists to locate and catalog objects of interest in large image collections. These systems incorporate novel image mining algorithms, as well as computational and database resources that allow users to browse, annotate, and search through images and analyze the resulting object catalogs. The architectures in these existing image mining systems are mainly based on module functionality. In contrast, we provide a different perspective to image mining with our four level information image mining framework. [6, 25] primarily concentrate on the Pixel and Object level while [37] focus on the Semantic Concepts level with some support from the Pixel and Object levels.

It is clear that by proposing a framework based on the information flow, we are able to focus on the critical areas to ensure all the levels can work together seamlessly. In addition, with this framework, it highlights to us that we are still very far from being able to fully discover useful domain information from images. More research is needed at the Semantic Concept level and the Knowledge and Pattern level.

## 6. Conclusion

The rapid growth of image data in a variety of medium has necessitated a way of making good use of the rich content in the images. Image mining is currently a bourgeoning yet active research focus in computer science. We have proposed a four-level information-driven framework for image mining systems. High-dimensional indexing schemes and retrieval techniques are also included in the framework to support the flow of information among the levels. We tested the applicability of our framework by applying it to some practical image mining applications. The proposal of this framework is our effort to provide developers and designer of image mining systems a standard framework for image mining with an explicit information hierarchy. We believe this framework represents the first step towards capturing the different levels of information present in image data and addressing the question of what are the issues and challenges of discovering useful patterns/knowledge from each level.

## References


1. Annamalai, M and Chopra, R.: Indexing images in Oracles8i. ACM SIGMOD, (2000)
2. Babu, G P and Mehtre, B M.: Color indexing for efficient image retrieval. Multimedia Tools and applications, (1995)
3. Beckmann, N, Kriegel, H P, Schneider, R and Malik, J.: The R*-tree: An efficient and robust access method for points and rectangles. ACM SIGMOD, (1990)
4. Berchtold, S, Keim, D A and Kriegel, H P.: The X-tree: An index structure for high dimensional data. 22nd Int. Conference on Very Large Databases, (1996)
5. Bertino, E, Ooi, B C, Sacks-Davis, R, Tan, K L, Zobel, J, Shilovsky, B and Catania, B.: Indexing Techniques for Advanced Database Systems. Kluwer Academic Publisher (1997)
6. Burl, M C et al.: Mining for image content. In Systems, Cybernetics, and Informatics / Information Systems: Analysis and Synthesis, (1999)
7. Cromp, R F and Campbell, W J.: Data mining of multi-dimensional remotely sensed images. International Conference on Information and Knowledge Management (CIKM), (1993)
8. Datcu, M and Seidel, K.: Image information mining: exploration of image content in large archives. IEEE Conference on Aerospace, Vol.3 (2000)
9. Eakins, J P and Graham, M E.: Content-based image retrieval: a report to the JISC technology applications program. (http://www.unn.ac.uk/iidr/research/cbir/report.html), (1999)
10. Gibson, S et al.: Intelligent mining in image databases, with applications to satellite imaging and to web search, Data Mining and Computational Intelligence, Springer-Verlag, Berlin, (2001)
11. Guttman, A.: R-trees: A dynamic index structure for spatial searching. ACM SIGMOD. (1984)
12. Haralick, R M and Shanmugam, K.: Texture features for image classification. IEEE Transactions on Systems, Man, and Cybernetics, Vol 3 (6) (1973)
13. Hsu, W, Lee, M L and Goh, K G.: Image Mining in IRIS: Integrated Retinal Information System, ACM SIGMOD. (2000)
14. Jain, A K, Murty, M N and Flynn, P J.: Data clustering: a review. ACM computing survey, Vol.31, No.3. (1999)
15. Jain, R, Kasturi, R and Schunck, B G.: Machine Version. MIT Press. (1995)
16. Jeremy S. and Bonet, D.: Image preprocessing for rapid selection in “Pay attention mode”. MIT Press. (2000)
17. Kaplan, L M et al.: Fast texture database retrieval using extended fractal features. Proc SPIE in Storage and Retrieval for Image and Video Databases VI (Sethi, I K and Jain, R C, eds). (1998)
18. Katayama, N and Satoh, S.: The SR-tree: An index structure for high-dimensional nearest neighbour queries. ACM SIGMOD. (1997)

19. Knuth, D E.: Sorting and searching, the Art of Computer Programming, Vol.3. Reading, Mass. Addison-Wesley (1973)
20. Lin, K, Jagadish, H V and Faloutsos, C.: The TV-tree: An index structure for high-dimensional data. The VLDB Journal, 3 (4). (1994)
21. Ma, W Y and Manjunath, B S.: A texture thesaurus for browsing large aerial photographs, Journal of the American Society for Information Science 49(7) (1998)
22. Manjunath, B S and Ma, W Y.: Texture features for browsing and retrieval of large image data, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18, (1996)
23. Megalooikonomou, V, Davataikos, C and Herskovits, E H.: Mining lesion-deficit associations in a brain image database. ACM SIGKDD. (1999)
24. Ooi, B C, Tan, K L. Yu,S and Bressan. S.: Indexing the Edges - A Simple and Yet Efficient Approach to High-Dimensional Indexing, 19th ACM SIGMOD-SIGACT-SIGART Symposium on Principles of Database Systems (2000).
25. Ordonez, C and Omiecinski, E.: Image mining: a new approach for data mining. IEEE. (1999)
26. Robinson, J T.: The K-D-B tree: A search structure for large multidimensional dynamic indexes. ACM SIGMOD. (1981)
27. Rui, Y, Huang, S T et al.: Image retrieval: Past, present and future. Int. Symposium on Multimedia Information Processing. (1997)
28. Salton, J and McGill, M J.: Introduction to Modern Information Retrieval. McGraw-Hill Book Company. (1983)
29. Sellis, T, Roussopoulous, N and Faloutsos.: C. $R^+$-tree: A dynamic index for multi-dimensional objects. 16th Int. Conference on Very Large Databases. (1987)
30. Stricker, M and Dimai, A.: Color indexing with weak spatial constraints. Proc SPIE in Storage and Retrieval for Image and Video Databases IV. (1996)
31. Stricker, M and Orengo, M.: Similarity of color images. Proc SPIE in Storage and Retrieval for Image and Video Databases III. (1995)
32. Swain, M J and Ballard, D H.: Color indexing. International Journal of Computer Vision **7**(1). (1991)
33. Tamura, H et al.: Textural features corresponding to visual perception. IEEE Transactions on Systems, Man and Cybernetics 8(6). (1978)
34. Tan, K L, Ooi, B C and Thiang, L F.: Retrieving Similar Shapes Effectively and Efficiently. Multimedia Tools and Applications, Kluwer Academic Publishers, accepted for publication, 2001
35. Tan, K L, Ooi, B C and Yee, C Y.: An Evaluation of Color-Spatial Retrieval Techniques for Large Image Databases, Multimedia Tools and Applications, Vol. 14(1), Kluwer Academic Publishers. (2001)
36. Wang, J Z, Li, J et al.: System for Classifying Objectionable Websites, Proceedings of the 5th International Workshop on Interactive Distributed Multimedia Systems and Telecommunication Services (IDMS'98), Springer-Verlag LNCS 1483, (1998)
37. Zaiane, O R and Han, J W.: Mining MultiMedia Data. CASCON: the IBM Centre for Advanced Studies Conference (http://www.cas.ibm.ca/cascon/), (1998)