



Image Mining to Assist Image Retrieval

C. Kanimozhi, D. Bayya, S. Gokulavani

PG Scholar, Manakula Vinayagar Institute of Engineering and Technology,
Puducherry, India

Abstract--- Image mining is the process of mining image from a large database (i.e.), it is similar to data mining. While gathering images, it handles knowledge hidden data and additional patterns are extracted. All significant patterns are generated without any prior information about the mining. A high level semantic feature is also proposed for image retrieval. By extracting different characteristics and it is converted in high level feature using production of fuzzy rules.

Keywords: Object Recognition, Image Retrieval, Image Indexing, Feature Extraction.

I. INTRODUCTION

The process of searching and discovering valuable information and knowledge from large volumes of data is image mining. Fig 1 depicts the image mining process.

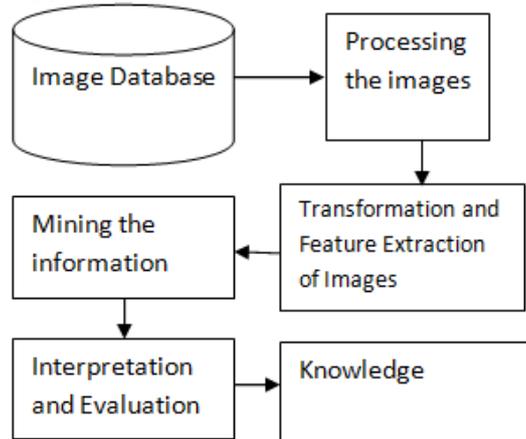


Fig 1. General image mining process

Two different approaches are used in image mining. One is to extract image from a huge database and other one is to extract a group of alphanumeric data and collection of images. In image processing feature extraction plays a vital role and it is a form of dimensionality reduction.

When the input entered is too large, it is transformed into lower representation of feature sets. Amount of resources are simplified, which helps in describing a large set of data in feature extraction. Image retrieval is done with help of several features. It includes shape, colour, and texture. Colour features are represented by pixel colours and histogram values based on the intensity. Shapes are represented by circularity and area measurements. The most used texture features are Gabor filters and texture measurements are based on the difference in histogram and run length features.

II. IMAGE MINING TECHNIQUES

To mine data from image information, it follows certain frameworks:

A.) Object Recognition

Object recognition plays a major role in image mining. Assigning a label to a region in recognition is represented to object detection and that label corresponds to a model containing lot of interesting objects. It is also called as supervised labelling problem, which is based on the model of known objects. Each object recognition system consist of a model database, which contains system known models

B.) Image Retrieval

Images are retrieved based on the requirement specification. It is categorized into three levels.

- a.) Image level 1 comprises retrieval with help of primitive features.
- b.) Image level 2 comprises retrieval with help of logical features.
- c.) Image level 3 comprises retrieval with help of attributes.

C.) Image Indexing

Image retrieval is supported with help of indexing, while focussing at various levels of information needs. For example if you are searching for an image in a large database, its feature vectors are based on high dimension and the complexity value for searching the image will be high.

In indexing two approaches are used: dimensionality reduction and high dimensional data indexing. Reduction is performed with help of algorithm and clustering. Indexing is performed with help of trees.

D.) Image classification and clustering

Image classification is based on supervised classification, where the images are available collectively and hence the labelled images are used for machine learning.

Clustering is based on unsupervised classification. Clustering is done with help of feature attributes.

E.) Association Rule Mining

An implication of the form $X \rightarrow Y$ is to depict association rule mining, where X refers an antecedent value and Y refers a consequent rule. Rule mining is based on two consecutive steps:

- 1.) In first find the largest item sets to meet the minimum constraints
- 2.) In second step, it generates rules from all the largest data sets that satisfy the minimum constraints.

III. FEATURE EXTRACTION

Feature extraction plays a vital role in object detection and it uses genetic algorithm. Good set of features are selected with help of this algorithm and to lead low error rates. The mechanism of feature extraction is done by converting low level features into high level semantics with help of fuzzy rules.

Irrelevant features are detected and removed with help of genetic algorithms. While mining an image from a database, first the image is pre-processed to convert the quality from low to high. Then the image is forwarded to various transformations. And then important features are extracted from the image with help of feature extraction.

A.) Colour Characteristics

In image mining, effective evaluation and performance parameters are based on the content requirement of the user. The colour image is first segmented to extract its features. First the RGB image is converted into an image format with an extended chromaticity (i.e. $l*u*v$) based on its luminance, redness-greenness, blueness-yellowness. Here twelve colours are used. They are blue, green, red, orange, purple, yellow and other six colours are made from the combination of the mentioned six colours. Different characteristics are obtained with help of luminance and saturation level. Based on the characteristic every colour is transformed into its reference colour. After this step the segmented image retains its original image with help of k-means algorithm. After retaining each segment is considered as N region to define its chromaticity space.

B.) Texture Characteristics

Image mining is generally based on the texture and colour histogram. To define the texture characteristics filters are used in the image. Energy value is calculated for each block in the image and the value is defined by different combination of frequencies and different orientations. Then for each block average magnitude value is computed to define its texture feature.

C.) Shape Characteristics

The shape of a particular image is determined by its pattern and shape of object. It is based on the primitive image descriptors. The primitive descriptors are colour, texture and shape. Characteristics are defined with help of content based image retrieval system.

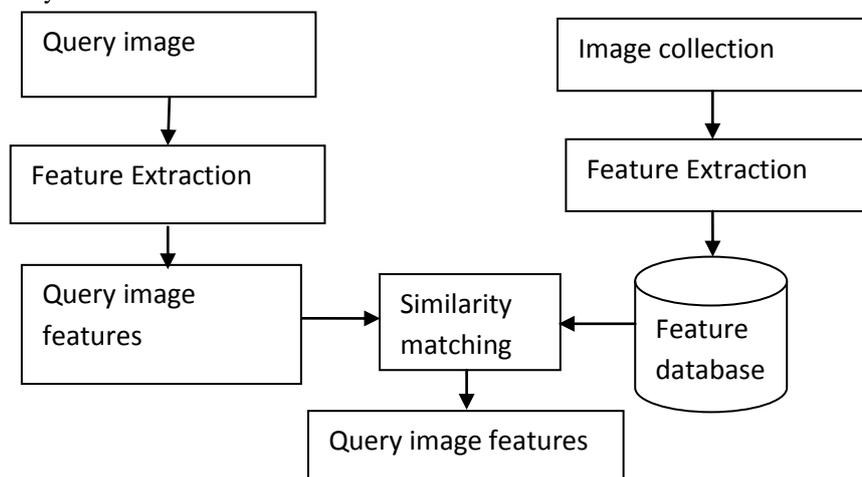


Fig 2. Content Based Image Retrieval System

First the image is partitioned into block of equal size. The features that are drawn from the histograms between RGB colour space, serve as the local information for the colour and texture feature. Different details of the same image are captured from the resolution and layouts. Information regarding shape of the image is captured from the edge images.

IV. CONCLUSION

In this proposed system, the process of retrieving the high level image semantics is used to make results efficient. Statistical characteristics can also be obtained from retrieval of certain properties of an image.

REFERENCES

- [1] Sethi, I., Coman, I., Day, B., Jiang, F., Li, D., Segovia-Juarez, J., Wei, G., and You, B., *Color-WISE: A System for Image Similarity Retrieval Using Color*, Proceedings of SPIE Storage and Retrieval for Image and Video Databases, Volume 3312, February 1998, 140-149.
- [2] Grosky W., Stanchev P., *An Image Data Model*, in *Advances in Visual Information Systems*, Laurini, R. (ed.), Lecture Notes in Computer Science 1929, 2000, 14-25.
- [3] Manjunath, B., Ma, W., *Texture features for browsing and retrieval of image data*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(8), 1996, 837-842.
- [4] Rabbitti, F and Stanchev, P., *GRIM_DBMS: a graphical image database management system*, in *Visual Database Systems*, Kunii, T. (ed.), Elsevier, Amsterdam, 1989, 415-430.
- [5] Oliva, A et al., *Real-world scene categorization by a self-organizing neural network*, Perception, supp 26, 19, 1997.
- [6] Chang, S., et al., *Semantic visual templates: linking visual features to semantics*, in IEEE International Conference on Image Processing (ICIP'98), Chicago, Illinois, 1998, 531-535.
- [7] Corridoni, J, Bimbo A., Vicario E., *Image retrieval by color semantics with incomplete knowledge*, Journal of the American Society for Information Science 49(3), 1998, 267-282.
- [8] Zhang Ji, Hsu, Mong, Lee, *Image Mining: Issues, Frameworks And Techniques*, Proceedings of the Second International Workshop on Multimedia Data Mining (MDM/KDD'2001), in conjunction with ACM SIGKDD conference. San Francisco, USA, August 26, 2001.
- [9] Ordonez C. and Omiecinski E., *Discovering association rules based on image content*. Proceedings of the IEEE Advances in Digital Libraries Conference (ADL'99), 1999.
- [10] Megalooikonomou, V., Davataikos C. and Herskovits, E., *Mining lesion-deficit associations in a brain image database*. KDD, San Diego, CA USA, 1999.
- [11] T. Sellis, N. Roussopoulos and C. Faloutsos. *The R+ Tree: A Dynamic Index for Multidimensional Objects*. In Proc 12th VLDB, 1987.
- [12] J. R. Smith and S. F. Chang. *VisualSEEK: A Fully Automated Content-based Image Query System*, in Proceedings of the 4th ACM International Multimedia Conference and Exhibition, Boston, MA, USA, November 1996.
- [13] W. Strelilein and A. Waxman. *Fused Multi-sensor Image Mining for Feature Foundation Data*. Proceedings of the Third International Conference on Information Fusion (FUSION 2000), Volume: 1, 2000.

AUTHOR BIOGRAPHY



Kanimozhi. C, completed Bachelor of degree in Information Technology at 2012 and Pursuing Master degree in Computer Science and Engineering under Pondicherry University, MIT (Manakula Vinayagar Institute of Technology)



Bavya. D, Bachelor of degree in Information Technology at 2012 and Pursuing Master degree in Computer Science and Engineering under Pondicherry University, MIT (Manakula Vinayagar Institute of Technology)



Gokulavani. S, completed Bachelor of degree in Computer Science and Engineering at 2012 and Pursuing Master degree in Computer Science and Engineering under Pondicherry University, MIT (Manakula Vinayagar Institute of Technology)