International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 27, 2017 ISSN: 2348-6600
Reference ID: IJCS-289 PAGE NO: 1898-1903

AN INTRINSIC STUDY ON IMAGE MINING AND ITS CHALLENGES

Ms.M.Tharini
Ph.D Research Scholar,
Rathinam College of Arts and Science,
Coimbatore.

Dr. R. Muralidharan,
Head, Department of Computer Science,
Rathinam College of Arts and Science,
Coimbatore.

Abstract - Image mining is the process of searching and discovering valuable information and knowledge in large volumes of data. Image mining is simply 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 accumulated in the images. Advances in image acquisition and storage technology have led to tremendous growth in significantly large and detailed image databases. These images, if analyzed, can reveal useful information to the human users. Image mining deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the images. Image mining is more than just an extension of data mining to image domain. With significantly large and increasing multimedia database often the users have to mine the available data to retrieve the relevant information. 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 increase in number of images and image databases has given way for the need for image mining techniques. Image mining is an extended branch of data mining that is concerned with the process of knowledge discovery concerning digital images. The main aim of this paper is to present an overview of the various image mining applications like image retrieval, Matching, Pattern recognition etc.

Keywords: Data mining, Image mining, Feature extraction, Image matching, Image retrieval.

1. INTRODUCTION

Data mining concept is combined with large databases such as Data repository and Data warehouse and its aim is to extract useful unknown information from raw data. Although like other concepts of information technology, it evokes several meanings such a data mining, information technology for different people; if it is applied accurately it can be a complex analytical tool for discovering useful patterns automatically among the data of a data repository. In fact, data mining is the advanced form of decision support that contrary to passive query tools generates templates, trends,

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 27, 2017 ISSN: 2348-6600
Reference ID: IJCS-289 PAGE NO: 1898-1903

and planned rules without requiring the user to generate questions.

Image mining in large set of image is a new approach in the field of research on the one hand, and image database and data mining researches on the other hand. Although, recently this discussion has caused the precise concept of image mining remain a challenge, researchers, particularly in recent years, have proposed different definitions of image mining, as well as various methods under this topic. Image mining focuses on the extraction of patterns from large collections of images while the emphasis of image processing and machine vision is on the understanding of certain characteristics of a specific image. A high volume of images, such as satellite images, medical images and digital photos produced on a daily basis. In case of the analysis of these images, a lot of useful information can be gained. The pixels shown in a raw image or series of images in order to detect objects and the relationship among them is the most fundamental challenge in the mining picture.

Image mining is a technique that explores information, images' data dependence and unambiguous patterns stored in the images. There are two basic techniques in this field, the first technique do the exploration in an extensive range of independent pictures. The second technique explores a series of integrated and linked images. Image mining is not only the simple fact of recovering relevant images but is the innovation of image patterns that are noteworthy in a given collection of images. The establishment of image mining system is frequently a complicated process because it implies joining diverse techniques ranging from image retrieval and indexing schemes up to data mining and pattern recognition.

This paper presents views about retrieval, matching, pattern recognition which will be very useful while extracting features like shape, color, size, texture, imprint etc from large image databases. The number of features required to represent an image can be very huge. Using all available features to recognize objects can suffer from curse dimensionality. Feature selection and extraction is the pre-processing step of image mining. Main issues in analyzing images are the effective identification of features and another one is extracting them. Due to the richness of information present in an image, Image mining comprises of various techniques in extracting the vital information from an image. The important activities in image mining are searching and retrieval of images based on the features and similarity of a given input query image from the image database. Image mining uses some operations of image processing such as image acquisition, image pre-processing, feature extraction and stored feature matching.

2. FEATURES USED FOR IMAGE MINING

There are main three features of image that are used to extract and store into database for matching with query. These are color, shape and texture.

1. Color Feature: A computer image is a matrix of pixels. The value of each pixel is proportional to the brightness of corresponding point in the scene. Color images are represented by three intensity component. These components generally

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 27, 2017 ISSN: 2348-6600
Reference ID: IJCS-289 PAGE NO: 1898-1903

correspond to red, green and blue (RGB). Image can associate an integer value with each pixel that can be used as an index to a table that stores the intensity of each color component. The histogram plots the number of pixels with a particular brightness level against the brightness level.

2. Shape or edge feature. Edge is simply a large change in frequency. Many approaches to image interpretation are based on edges, since analysis based on edge detection is insensitive to change in the overall illumination level. Edge features are particularly important for some of the darker images. Fortunately, the training image was of normal quality and hence it did not use the edge feature. However, it does use it for some of the darker images in the set for testing. Edge detection highlights image contrast. Detecting contrast, which is difference in intensity, can emphasize the boundaries of features within an image.

3. Texture feature. Texture is defined as a neighborhood feature as a region or a block. The variation of each pixel with respect to its neighboring pixels defines texture. Texture is a very general notion that can be attributed to almost everything in nature. For a human, the texture relates mostly to a specific, spatially repetitive structure of surfaces formed by repeating a particular element or several elements in different relative spatial positions. Generally, the repetition involves local variations of scale, orientation, or other geometric and optical features of the elements. It is almost impossible to describe textures in words, although each human definition involves various informal qualitative structural features, such as fineness - coarseness, smoothness, granularity, lineation, directionality, roughness, regularity - randomness, and so on.

2.1. Feature Extraction

Measuring features of an image is a basis factor to distinguish and categorize an image. The machine vision research is providing modals of objects and scenes of an image to extract image properties for developing decision rules, and then analyze and describe observed image. We use the image processing methods, clustering and measuring image properties for this purpose.

Developing imaging techniques according to image revival system is based on content. Color, texture, style, object shape, arrangement and their situations inside image and etc. are all bases of visual contents of an image and an image is indexed based on these properties. If properties and characteristics are selected correctly, they can express much useful information about an image. Features extraction methods analyze properties, objects and images to extract significant features indicating different classes of objects. Properties are given to categorization as an input to distinguish a class to which the object is related. In the three features, always texture plays an important role.

Texture is one of the most important features that can be extracted from images. Texture is referred to informational patterns or structural arrangement observed in an image. Texture is kind of vision features that it does not depend on color, severity and reflections in natural phenomenon in images. Texture is a collection of all natural features in a

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 27, 2017 ISSN: 2348-6600
Reference ID: IJCS-289 PAGE NO: 1898-1903

surface and for this reason we use from this feature widely in image processing. Many objects are distinguished via only texture and without any additional data. First, texture analysis was based on first order statistics or second order statistics. There are different methods to measure images textural features such as co – occurrence matrix, fractals, Gabor filters, and microwave converter socializations. Also many techniques were developed to describe local patterns via textural spectrum. We can use co-occurrence matrix and edges data to describe a texture. After extracting the image using any one of the feature, image should be verified to match.

2.2 Image Matching

Image matching, or comparing images in order to obtain a measure of their similarity, is a fundamental aspect of many problems in computer vision, including object and scene recognition, content-based image retrieval, stereo correspondence, motion tracking, texture classification and video data mining. It is a complex problem that remains challenging due to partial occlusions, image deformations, and viewpoint or lighting changes that may occur across different images. Therefore, according to this definition, image matching problem is accomplished by transforming one of the images in such a way that the similarity with the other image is maximized in some sense. The 3D nature of real-world scenarios makes this solution complex to achieve, especially because images can be taken from arbitrary viewpoints and in different illumination conditions. Instead, the similarity may be applied to global features derived from the original images. However, this is not the more efficient solution. Besides, these global statistics cannot usually deal with real-world scenarios because they do not often give adequate descriptions of the local structures or discriminating features which are present on the image.

Image matching refers to the automatic identification and measurement of corresponding image points that are located on the overlapping area of multiple images. The various image matching methods can be divided into three categories including:

- Area based matching
- Feature based matching
- Relation based matching

Using image matching methods we verify the image. If found correct then we proceed to Image retrieval otherwise all the steps have to be checked again and again. Even then we don't find the exact match, pattern should be changed.

2.3 Image Retrieval

An image retrieval system is a computer system for browsing, searching and retrieving images from a large database of digital images. Most traditional and common methods of image retrieval utilize some method of adding metadata such as captioning', keywords, or descriptions to the images so that retrieval can be performed over the annotation words. Information retrieval is the activity of obtaining information resources relevant

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 27, 2017 ISSN: 2348-6600
Reference ID: IJCS-289 PAGE NO: 1898-1903

to an information need from a collection of information resources. Searches can be based on full-text or other content-based indexing.

There are several approaches for image retrieval. The first approach is text-based. The image is described as a set of key words or free text. The queries are based on exact or probabilistic match of query text. The alternative way allows the retrieval query to be based on the visual content of an image such as: image patterns, colors, textures, shapes of image object and location information. The content-based image retrieval can be characterized by the ability to retrieve relevant images to the user defined image query, based on the semantic content of the images.

Conclusion

Image Mining is the advanced field of Data Mining technique .The main objective of the Image Mining is to remove the data loss and extracting the meaningful information to the human expected needs. This paper highlights the need for image mining in view of the fast growing amounts of image and the problems faced by the users. Many challenges and research problems are available in image mining. These problems can be solved by developing new algorithms, concepts and techniques for extracting hidden knowledge from the image data bases. This paper discussed the image mining basic concepts and techniques. The major problem Image Retrieval is also discussed in this paper.

References

1. Archana B. Waghmare, "Low-Level Feature Extraction for Content-Based Image Retrieval", International Journal of Advances in Computing and Information Researches ISSN: 2277-4068, Volume 1–No.2, April 2012.
2. Bi, J, and J. Liang. "Multiple Instance Learning of Pulmonary Embolism Detection with Geodesic Distance along Vascular Structure." Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'07). 2007.
3. 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.
4. 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.
5. James Dowe. Content based retrieval in multimedia imaging. In Proc. SPIE storage and Retrieval for Image and Video Databases, 1993.
6. Dr.V.Mohan, A.Kannan, "Color Image Classification and Retrieval using Image mining Techniques", International Journal of Engineering Science and Technology,Vol. 2(5), 2010.

International Journal of Computer Science
Scholarly Peer Reviewed Research Journal - PRESS - OPEN ACCESS
ISSN: 2348-6600
http://www.ijcsjournal.com Volume 5, Issue 1, No 27, 2017 ISSN: 2348-6600
Reference ID: IJCS-289 PAGE NO: 1898-1903

7. W. Niblack, R. Barber, and et al. The QBIC project: Querying images by content using color, texture and shape. In Proc. SPIE Storage and Retrieval for Image and Video Databases, Feb 1994.
8. Patricia G. Foschi, Deepak Kolippakkam, "Feature Extraction for Image Mining".
9. U.Ravindran,T.Shakila ,"Content Based Image Retrieval For Histology Image Collection Using Visual Pattern Mining", International Journal of Scientific & Engineering Research, Volume 4, Issue 4, April-2013.
10. Y. Rui, T. Huang and S. Chang. Image retrieval: current techniques, promising directions and open issues. Journal of Visual Communication and Image Representation, 10(4): 39-62, April 1999.
11. Sharlee Climer, Sanjiv K. Bhatia. Image Database indexing using JPEG coefficients. The journal of the Pattern Recognition Society, Pattern Recognition 35 (2002) 2479-2488.
12. S. F. Chang and J.R. Smith. Extracting Multi-Dimensional Signal Features for Content- Based Visual Query. SPIE Symposium on Visual Communications and Signal Processing, May 1995.
13. Swati V. Sakhare , Vrushali G. Nasre, "Design of Feature Extraction in Content Based Image Retrieval (CBIR) using Color and Texture", International Journal of Computer Science & Informatics,Volume-I, Issue-II, 2011.
14. Wu, D, J Bi, and K Boyer. "A Min-Max Framework of Cascaded Classifier with Multiple Instance Learning for Computer Aided Diagnosis." Proceedings of IEEE International Conference on Computer Vision and Pattern Recognition (CVPR'09). 2009.
15. 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.