Techniques of Image Mining: A Review

Ms. Shraddha D. Jumade

Student, Computer Science and Engineering, Sipna C.O.E.T., Amravati, India
shraddhajumade@gmail.com

Prof. Sheetal S. Dhande

Associate Professor, Computer Science and Engineering, Sipna C.O.E.T., Amravati, India
dhande_123@rediffmail.com

Abstract— Nowadays, a large portion of information is in visual form; it is essential and certainly pleasing to search for images. Image mining is used to discover the knowledge from the image dataset. Image mining is an extended branch of data mining that is concerned with the process of knowledge discovery concerning digital images. Image mining is a class of analytical techniques that examine a large amount of image data. Analyzed images reveal useful information to the human users. Image mining manages with the extraction of knowledge, image data relationship, or other patterns which is not stored in the images. Image mining has a variety of applications in various sectors like medical diagnosis, biology, remote sensing, space research, etc. The main aim of this paper is to present a survey of the various techniques used for image mining applications like image retrieval, Matching, Pattern recognition given by different researchers.

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

I. INTRODUCTION

The image mining was introduced to extract implicit knowledge, image and data relationship. In text based image retrieval system only find out the images those are just concerned with the accurate text that is described by human or relevant query, instead without looking into the content of related images[14].Image mining is an interdisciplinary endeavour that draws upon expertise in various fields like computer vision, image retrieval, matching and pattern recognition. Some methods allow image mining to have two different approaches. First method extracts images from image databases or collection of images. Second method mines a combination of associated alphanumeric data and collection of images. Research in Image mining can be broadly classified in two main directions (1) Domain specific applications (2) General applications. Both are used to extract most relevant image feature and later to generate image patterns.

Image mining is still at the experimental stage and growing field of research. Lack of understanding in the research issues of image mining is the obstacle to rapid progress. Image data plays vital role in every aspect of the systems like business, hospitals, engineering and so on. Image mining normally deals with the study and development of new technologies that allow easy analysis and interpretation of the 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 and complicated process because it implies joining diverse techniques ranging from image retrieval and indexing schemes up to data mining and pattern recognition [12].

II. OVERALL PROCESS OF IMAGE MINING

Image mining is just at its infancy, however**,** observing from some of the existing image mining systems, overall process can be divided into the following parts[14]:-

1. *Data preprocess*:

A lot of dirty and noisy data exist in large image databases, for instance, images that are extremely unclear. Those data often cause chaos in mining process and give birth to worse mining results, so it is necessary to preprocess data, clean up noisy, dirty data to highlight the features if that image

2. *Extracting multi-dimensional feature vectors*:

Using image processing technologies such as image segmentation, picking up the edge to extract task related feature vectors, form multi-dimensional feature vectors.

3. *Mining on vectors and acquire high-level knowledge*:

Various methods such as object recognition, image indexing and retrieval, image classification and clustering, neural network are used on feature vectors for mining and acquiring hidden and valuable high level knowledge, then evaluate and explain that exact query related knowledge.

III. LITERATURE REVIEW

Numerous researches have been carried in Image mining, in this section of the paper presenting a survey on various images mining techniques those were proposed earlier.

Wynne Hsu, Mong Lee and Ji Zhang(2002)[1], examined there search issues in image mining, development in image mining. They proposed an information driven framework for image mining. In that they made out four levels of information: Pixel level, Object level, semantic concept level, pattern and knowledge level. This approach does not addressed noise redundancy.

Aura Conci ,Everest Mathias, M.Castro(2002)[2], proposed a framework for mining images by color content. Their framework provides the possibility of use five distance function for evaluation of similarity among images and two types of quantization. The procedure here presented considers only retrieval aspects. Considerations like complexity or time performance are not treated here. The ideas presented are only a small step in a very rich research direction. Others visual features such as texture, shape, and use of compressed images can be identified for further extension of this problem.

Peter Stanchev(2003)[3], using image mining in image retrieval, described a new method for image retrieval using high level semantic features. 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. The main advantage of this method is the possibility of retrieval using high level image semantic features. After the full system realization it will be able to obtain statistic characteristics about the usefulness of the suggested method.

Md. Farooque(2003)[4], With the advancement in the world wide web ,a large amount of data on many different fields has become available online. User retrieved images by an efficient and effective manner .Many techniques have been developed to solve the image retrieval problem on the basis of image features such as color ,texture and shape. This technology called as Content Based Image Retrieval. It plays an important role in Image Indexing and Retrieval.

Herbert Daschiel , Mihai Datcu(2005)[5], demonstrated the concepts of a prototype of a Knowledge-driven content-based information mining system produced to manage and explore large volumes of remote sensing image data. The system consists of a computationally intensive offline part and an online interface. The offline part aims at the extraction of primitive image features, their compression, and data reduction, the generation of a completely unsupervised image content-index, and the ingestion of the catalogue entry in the database management system.

Sanjay T. Gandhe, K. T.Talele, Avinash G. Keskar (2007)[6], put forth an image mining technique using wavelet transform. The author proposed an image mining approach using wavelet transform. It uses common pattern identical, pattern identification and data mining models with the intention that a real life scene/image can be associated to a particular category, assisting in different prediction and forecasting mechanisms. It is a three-step procedure i.e. image gathering, learning and classification. The

conception of image mining as a consequence can be competently used for weather forecasting so that one can know the natural disasters that may occur in advance.

L. Jaba Sheela , V. Shanthi (2007)[7], described the image mining approaches for categorization and segmentation of Brain MRI data. Image segmentation plays a vital role in several medical imaging applications by computerizing or assisting the description of anatomical arrangements and additional regions of interest. Automatic recognition of tumors in several medical images is encouraged by the requirement of better accuracy when handling with a human life. It has been confirmed that double reading of medical images possibly will show the way for enhanced tumor detection. But the cost implied in double reading is extremely huge, that's why better software to assist humans in medical institutions is of vast interest at the present time. In their approach they developed a system which uses image mining approaches to categorize the images either as normal or abnormal and then divide the tissues of the anomalous Brain MRI to recognize brain related diseases.

P.Rajendran, M. Madheswaran (2009)[8], discussed an improved image mining technique. An enhanced image mining technique for brain tumor classification using pruned association rule with MARI algorithm is presented in their paper. The method proposed makes use of association rule mining technique to classify the CT scan brain images into three categories namely normal, benign and malign. It combines the low-level features extracted from images and high level knowledge from specialists. The developed algorithm can lend a hand to the physicians for well-organized classification with multiple keywords per image to get better the accuracy. The method proposed in this paper classifies the brain CT scan images into three categories: normal, benign and malignant. The experimental result on pre-diagnosed database of brain images showed 96% and 93% sensitivity and accuracy respectively.

Rajshree S. Dubey (2010) [9], illustrated about an Image mining methods which is dependent on the Color Histogram, texture of that Image. The query image is considered, then the Color Histogram and Texture is created and in accordance with this the resultant Image is found. They have examined a histogram-based search techniques and color texture techniques in two different color spaces, RGB and HSV. Histogram search distinguish an image through its color distribution. It is revealed that images retrieved by using the global color histogram possibly will not be semantically related although they share comparable color distribution in some results.

Shaikh Nikhat Fatma, Madhu Nashipudimath (2011) [10], presented association rule for Image mining. Association rule is deals with the extraction of image pattern from a large database of images. This method help us for Prediction. We will discuss this with an example, if sky contains black clouds so there are 64% chances it will rain.
The method is as follows:-
It segmented the images into blobs (region descriptor) where blob is equal to an object. Compare blob with all other blobs with an id. This works as a pre-processing algorithm after that create Auxiliary images with identified objects. Apply data mining techniques to produce object association rule. Basically this technique used for selecting images for a particular field (eg. Weather, Medical Images).

Nishchol Mishra , Dr. Sanjay Silakari(2012)[11], emergence and proliferation of social network sites such as Facebook and Linkedin and other multimedia networks such as Flicker has been one of the major events of this century. The networks have acquired immense popularity and have become a part of the daily lives of millions of people. Many of these network sites are thus extremely rich in content, and contain a tremendous amount of multimedia content waiting to be mined and analyzed. Analyzing this huge amount of multimedia data to discover useful knowledge is a challenging task. It has opened up opportunities for research in Multimedia Data Mining (MDM). Multimedia Data Mining can be defined as the process of finding interesting patterns from media data such as audio, video and text that are not ordinarily accessible by basic queries and associated results. This paper mainly focused on Image Mining techniques and how Content-based Image Retrieval(CBIR) can be helpful for Image Mining.

Mahip M.Bartere , Dr.Prashant R.Deshmukh(2012)[12], image mining presents special characteristics due to the richness of the data that an image can show. Effective results of image mining by content require that the user point of view is used on the performance parameters. Comparison between different mining by similarity systems is particularly challenging owing to the great variety of methods implemented to represent likeness and the dependence that the results present of the used image set. In this paper we proposed an evaluation framework for comparing the influence of the distance function on image mining by color and also a way to mine an image from its name. Experiments with color similarity mining by quantization on color space and measures of likeness between a sample and the image results have been carried out.

IV. CONCLUSION

This paper presents a survey on various image mining techniques such as image retrieval, image matching, pattern recognition etc, that was proposed earlier by researcher. This overview of image mining focuses on image mining implementations, usability and challenges. It also delivers conceptual overview of methodology. Image mining has its applicability in almost all the fields like medicine, remote sensing, entertainment, cyber forensics, DNA, Classification of objects etc. Effective evaluation of the results of image mining by content requires that the user point of view is used on the performance parameters. The query image is compared to each of database images to determine whether they are equivalent or not by comparing with all features. Main goal of image mining is the discovery of image patterns that are significant in a given collection of images. This paper deals with a brief study of the various approaches dealt with different researchers in all the phases related to image retrieval and mining. It is up to the user to choose the appropriate method based on the application.

REFERENCES

[1] Wynne Hsu, Mong Li Lee and Ji Zhang , "Image Mining: Trends And Developments", Journal Of Intelligent Information Systems, Vol.19, no.1, pp.7-2, Jan 2002.

[2] Aura Conci , Everest Mathias M. M. Castro , "Image Mining By Content", Expert Systems With Applications, Vol. 23, no. 4, pp. 377-383,April 2002.

[3] Peter Stanchev,"Using Image Mining In Image Retrieval", IASTED conference,"Computer Science and Technology"2003.

[4] Md. Farooque,"Image Indexing And Retrieval", DRTC Workshop on Digital Libraries: Theory and Practice March DRTC,2003.

[5] Herbert Daschiel , Mihai Datcu, " Information Mining in Remote Sensing Image Archives: System Evaluation", IEEE Transactions On Geoscience and Remote Sensing, Vol 43, no. 1, January 2005.

[6]Sanjay T. Gandhe, K. T. Talele and Avinash G. Keskar , "Image Mining Using Wavelet Transform", Knowledge-Based Intelligent Information and Engineering Systems, Springer link book chapter, pp. 797-803,2007.

[7]L. Jaba Sheela , V. Shanthi , "Image Mining Techniques For Classification And Segmentation Of Brain MRI Data," Journal of Theoretical and Applied Information Technology," vol. 3, no. 4, pp.115-121, 2007.

[8]P. Rajendran , M. Madheswaran, "An Improved Image Mining Technique For Brain Tumour Classification Using Efficient classifier", International Journal of Computer Science and Information Security, vol. 6, no. 3, pp. 107-116, 2009.

[9]Rajshree S. Dubey , "Image Mining using Content Based Image Retrieval System", International Journal on Computer Science and Engineering, Vol. 02, No. 07, pp. 2353-2356, 2010.

[10] Shaikh Nikhat Fatma, Madhu Nashipudimath, "Image Mining Using Association Rule," World Congress on Information and Communication Technologies, IEEE. pp 587-593, 2011.

[11]Nishchol Mishra, Dr. Sanjay Silakari, "A Perspective Image Mining In The Context Of Content Based Image Retrieval", International Journal of Computer Science Issues, Vol. 9, Issue 4, No 3, July 2012.

[12]Mahip M.Bartere ,Dr.Prashant R.Deshmukh, "An Efficient Technique Using Text And Content Base Image Mining Technique For Image Retrieval",International Journal of Engineering Research and Applications,Vol. 2, Issue 1, pp.734-739,Jan-Feb 2012.

[13]A. Hema ,E. Annasaro, "A Survey In Need Of Image Mining Techniques", International journal of Advanced Research In Computer and Communication Engineering, Vol. 2,Issue 2,Feb. 2013.

[14]Gurpreet Kaur, "A Review : Content Based Image mining Technique For Image Retrieval Using Hybrid Clustering", International journal of Advanced Research In Computer Engineering and Technology", Vol. 2,Issue 6,June 2013.