

A Survey of Image Mining & its Applications

^{#1}Ashlesh N. Mandke, ^{#2}Dr Rajesh S. Prasad



¹ashleshmandke1@gmail.com
²rajesh.prasad@sinhgad.edu

^{#12}Computer Engineering Department,

Savitribai Phule Pune University Pune-411037, India

ABSTRACT

The sole purpose of this review paper is the importance of Image data used in enormous fields like medical science, forensics, robotics, industries and geography as well as engineering. Due to Big data as a new concept, a chunk of data forms of datasets in different forms has been available for people. This paper aims for the goal at surveying the present situation of image mining and as well as at detailing its new problems and giving new directions of the future scope in Image Mining.

Keywords—analysis; preprocessing; image organization; color histogram; clustering.

ARTICLE INFO

Article History

Received: 20th December 2016

Received in revised form :

20st December 2016

Accepted: 22nd December 2016

Published online :

24th December 2016

I. INTRODUCTION

System Day by day, technology is advancing in many enormous fields. Due to concept of big data, there is very less chance of redundancy of data. This is not the issue, but rather huge data is available for people. So maintaining databases has become most essential for such loads of data. [1]

Over the last 10 years, Data mining has expanded rapidly due to big data. Day by day data mining, data processing are getting to be more simple, easy and more accurate. Besides text mining, image mining some other concepts like web mining and their algorithms etc. have come up with image mining is becoming popular so as putting weight in Big Data. [1]

Image data plays an essential role in different systems like business, hospitals, engineering and so on. Image mining normally deals with the study and development of new technologies that allows easy study and explanation of images. The extraction of all useful knowledge, image data its attributes or other patterns are store in the image. [1] Image mining is not simply fact of recovering right images but is the discovery of image patterns that are worth in a given collection of images. [2]

II. COMPONENTS OF IMAGE MINING

A. IMAGE STUDY

Image study is the first important and essential step in every Image Mining procedure. Its objective is to analyze each relevant aspect or attribute or basic feature required representing the actual image.

B. IMAGE PRE-PROCESSING

Preprocessing of image is the most important and mandatory step in Image analysis. Using different tools and software the image is extracted raw, then converted into using format. The image is made better in every way like improving resolution of the image and especially noise reduction is performed for making images normal. Fig 1 gives the proper explanation of image preprocessing procedure where all steps are carried out.

C. DATA CLUSTERING

Data clustering can be defined as grouping of similar objects in one cluster and dissimilar objects into another cluster. Fig 1 shows the graphical explanation of unclustered and clustered data. [1]

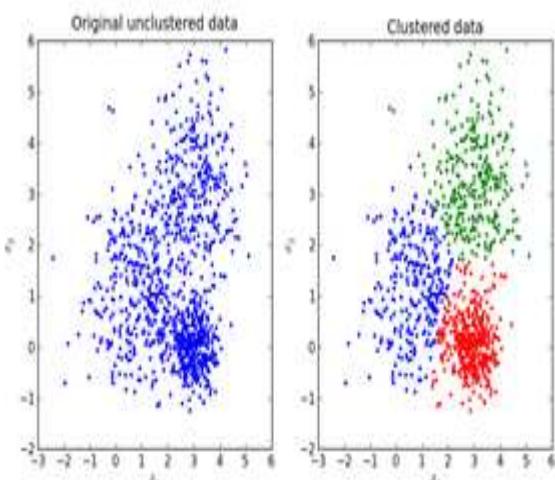


Fig 1: - un-clustered and clustered data

D. Image Organization

The main objective of image organization is to group objects discovered in an image.

Image classification is done in two ways: -

I. Image Clustering:

According to standard organization methods, image clustering describe grouping of objects. In other words, objects are grouped into clusters based on similarities and dissimilarities. The difference or the similarities are calculated based on evaluation of features such as color, shape, size etc. [1]

II. Image Observation:

Image observation is also called as the original approach of the respected image. The main objective of the supervised classification is to divide the objects into predefined group via detection tool. [1]

E. Color Histogram

Understanding the concept of color histogram is very essential for successful image processing. Color Histogram is defined as evenly distribution of color over the image is called as Color Histogram. It also represents the composition of different primary colors and no of pixels in each type of color appearing in that specific image. Fig 2 explains an image of sea wave and fig 3 its color histogram likewise color distribution. [1]



Fig 2:- image of a sea wave



Fig 3:- sea wave color histogram

F. Data management

As we all know image contains large amount of information. Also it needs a lot of space or size for storing the image. For retrieval of the respected image, it should be stored in image databases it can be extracted from the databases at any time instant. For storing the images, there are different compression images available such as jpeg, mpeg 2, di com etc. [2]

III.DISCUSIONS OF DATA MANAGEMENT

NASA is very interested in data mining and database management. The reason behind this keen interest is in spacecraft or manned rocket return back from the dedicated mission which were told to the respective astronauts or robots there they find large amount of data captured outside in space. It is said that placing or deploying good as well as intelligent algorithms in spacecraft has led to more economical missions and works with more efficiency. [3]

Mining of such data is useful upon expertise in computer vision, image understanding, data mining, machine learning, databases, distributed / parallel computing, software design and artificial intelligence (AI).

Prioritization and focusing can be another use of data mining. For e.g. if an area is a volcano prone region then this image processing system could give the entire geographical prone region and can alert all human community of evacuating the entire place of having high probability of volcanic eruptions. [2]

For deploying the intelligent algorithm for image mining onto the spacecraft it needs to be operated from the view that algorithm must be thoroughly tested and proved in ground based to gain acceptance. [2]

IV. APPLICATIONS

IM has its own applications in space explorations, database etc. It also has its vast applications in the field of medical sciences where MRI (Magnetic Resonance Imaging) for viewing internal structures of a human or animal body. [4]

The MRI is such a machine to view and understand internal structures of body in detail especially the soft tissues of brain thereby locating brain tumors. All types of cancers, smoking, drinking, tumors, infection of HIV of genetic risks can be detected by use of MRI i.e. use of small scale x-rays. [4]

But the issue of image processing efficiency is always degraded by noise i.e. unwanted interfering which misleads or misguides to false reports. The image is pre-processed by using a technique called DE noising to pull out noise of useful information. [4] Image Quality can be improved by many techniques such as prioritization and filter.

For making the image efficient in every better way, we use prioritization and filtration. As per the author, we come to know about the de noising technique, which is very essential for cutting the noise. [4]

The enhancement techniques for improving the image quality and to reduce noise. They are as follows:

a) Median Filter

Median filtering is an operation often used in image processing to reduce noise i.e. unwanted signal. A median filter is more effective than twisting when the goal is to simultaneously reduce noise and preserve edges. Median is calculated by rearranging the pixel value. [4]

b) Resolution Enhancement

Resolution of an image is always a problem in medical field. Resolution is the amount of detail information in that image. More the resolution high the image details.

The DWT (Discrete Wavelet Transform) is used preserving the edges and the outline information. The capacity of resolution enhancement technique is calculated using Peak Signal to Noise Ratio. [4]

Firstly, the image is preliminary processed using DE noising technique. After denoising, it ends in noise reduction and loss of quality at the image edges. For reduction of noise, resolution enhancement technique is very important. However, deriving the PNSR value is also very important.

Peak Noise to Sound ratio also pronounced PNSR for short. It can be defined as ratio of maximum noise to maximum signal. Noise also hampers the quality of the image. [4] The proposed methods by the author are applied in many of the applications like MRI where two enhancement techniques

were very effective namely, the median filter and resolution enhancement. In these techniques, resolution enhancement is much effective w.r.t median filter as the quality of an image is enhanced by the use of resolution enhancement method [1], [4].

V. CONCLUSION

The sole purpose of this paper is to introduce and discover the unknown patterns in image sets. The goal of this review paper or survey is to highlight the fact that users like doctors, teachers, students etc. need to face large volume of images from internet or any databases.

Main aim of Image Mining is to discover as many patterns in a given collection of images.

Deploying intelligent algorithm for efficient working of Image mining is very essential for extraction of useful data from large volumes of Big Data. [1][2][3][4]

VI. FUTURE SCOPE

- Knowledge gained or obtained on computer-driven processing images has enormous weight or we call it as potential.
- Giving appropriate indexing is also a concern.
- Encoding the contextual hidden information in image is very important for successful image mining.
- Organization of obtained patterns is mandatory step in Image Processing.

VII. ACKNOWLEDGEMENT

I express my true sense of gratitude to my guide Dr Rajesh S. Prasad for his perfect valuable guidance, all the time support and encouragement that he gave me. During the long journey of this study, he supported me in every aspect. He was the one who helped and motivated me to propose research in this field and inspired me with his enthusiasm on research, his experience and his lively character.

REFERENCES

- [1] Barbora Zahradnikova, Sona Duchovicova and Peter Schreiber, "Image Mining: Review and New Challenges" (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 6, No. 7, 2015.
- [2] M. C. Burl, C. Fowlkes, and J. Roden, "Mining for image content," Systemics, cybernetics, and informatics/information systems: analysis and synthesis, 1999.

- [3] A. Hema and E. Annasaro, "A survey in need of image mining techniques," International Journal of Advanced Research in Computer and Communication Engineering (IJARCCE), 2013
- [4] S. Rajeshwari and T. S. Sharmila, "Efficient quality analysis of MRI image using preprocessing techniques," in Information & Communication Technologies (ICT), 2013 IEEE Conference on. IEEE, 2013, pp. 391–396
- [5] Ankit Lodha, Agile: Open Innovation to Revolutionize Pharmaceutical Strategy, Vol-2, Issue-12, 2016
- [6] Ankit Lodha, Analytics: An Intelligent Approach in Clinical Trail Management, Volume 6 ,Issue 5 , 1000e124
- [7] A. Farruggia, R. Magro, and S. Vitabile, "A text based indexing system for mammographic image retrieval and classification," Future Generation Computer Systems, vol. 37, pp. 243–251, 2014.
- [8] "Understanding how the image colors are distributed," <http://wikipedia.org/>
- [9] https://en.wikipedia.org/wiki/Color_histogram