www.ijecs.in
International Journal Of Engineering And Computer Science ISSN:2319-7242
Volume 7 Issue 3 March 2018, Page No. 23755-23760
Index Copernicus Value (2015): 58.10, 76.25 (2016) DOI: 10.18535/ijecs/v7i3.17

A Big Data Based Edge Detection Method for Image Pattern Recognition - A Survey

S.Dhivya[1], Dr.R.Shanmugavadivu[2]

Research Scholar, Department of Computer Science, PSG College of Arts and Science, Coimbatore, India[1].

Assistant Professor, Department of Computer Science, PSG College of Arts and Science, Coimbatore, India[2].

Abstract:
In Today's era Big Data is one of the most well-known research area that try to solve many research problems. The focus is mainly on how to come out those problems of Big Data and it could be handling in recent systems. Image mining and genetic algorithm is used to automate the process of images, patterns, data sets and etc. Image mining is used to extract the hidden images from the set of images. Genetic algorithm is also quite effective in solving certain optimization and intelligence problems and it is used in many applications, including image pattern recognition. The survey paper reviews of Big Data with edge detection methods on various types of images. In edge detection image pattern recognition is to choose the best images from the group of images by using both image mining and genetic algorithm techniques.

Keywords: Edge Detection, Big Data, Image mining, Genetic Algorithm, Image Pattern Recognition.

1. Introduction:

Big Data is used both structured and unstructured data that is so large it is difficult to process using traditional database and software techniques. A data is classified as Big Data if it effectively satisfies any one of the requirements of Big Data, namely; Volume, Velocity and Variety. On the other hand, Big Data also arises with many challenges, such as difficulties in data capture, data storage, data analysis and data visualizations. Big data sets are used for minimizing risk, identifying unfamiliar objects and uncover hidden patterns. Data mining and Big Data are related to each other. Both containing more number of data's and the data's are extracted from the different data sets. Here, we have taken Big Data as an image. The clustering technique is used for grouping the images. After clustering the images, applies edge detection based genetic algorithm techniques like selection, crossover and mutation. In edge detection images are converted into gray scale images. For extracting the images, mining techniques will be used. Image mining is used to extract the images patterns from the large set of images, which draws the basic principles from concepts in databases, machine learning, statistics, pattern recognition and 'soft' computing. In image processing, edge detection is the one of the image processing technique for identifying the objects within the images.

There are many techniques used in edge detection but genetic algorithm is determined by a fitness function define to evaluate a solution's ability to deal with a given task, ending up in bringing out a best possible solution. These paper discuses various techniques used for Edge Detection based Genetic Algorithms.

2. Literature Review

Various algorithms and techniques are used for agricultural image processing method. Here, we studied various plant Phenotyping image segmentation methods. Pascal Fua and Yvan G. Leclerc, describes standard edge detection method and it fails to find most of the relavent edges so author proposed the new method for images in initial segmentation [9]. Massimo Minervini , Andreas Fischbach , Hanno Scharr , Sotirios A. Tsaftaris, describes a collection of benchmark datasets of raw and annotated top-view color images of rosette plants which describe plant material, imaging setup and procedures for different experiments: [10]. Eren Erdal Aksoy, Alexey Abramov, Florentin Wörgötter, Hanno Scharr, Andreas Fischbach, Babette Dellen, preprocessing, leaf segmentation, and leaf tracking. Leaf-shape models are applied to improve leaf segmentation, and further used for measuring leaf sizes and handling occlusions [11]. J.Jeya Priyankha and K.Suresh kumar, describes the preprocessed image is subjected to K means clustering to get infected part of the leaf which is subjected to morphological processing to expanding the infected area. The author uses the HOG and SVM classifier algorithms[12]. Arti Singh, Baskar Ganapathysubramanian, Asheesh Kumar Singh, Soumik sarkar, describes the various methods about plant phenotyping and how the leaf images are extracted by using various machine learning and data mining techniques[13].

led Bhanu et al. to adopt a GA to determine the parameter set that optimise the output of an existing segmentation algorithm under various conditions of image acquisition and is namely Phoenix segmentation algorithm[18]. Another situation wherein GAs may be useful tools is illustrated by the work of Yoshimura and Oe [21]. For example, Bhandarkar et al. defined a multiterm cost function, which is minimized using a GAevolved edge configuration. The idea was to solve medical image problems, namely edge-detection. In their approach to image segmentation, edge detection is cast as the problem of minimising an objective cost function over the space of all possible edge configurations and a population of edge images is evolved using specialised operators. Fuzzy GA fitness functions were also considered by Chun and Yang, mapping a regionbased segmentation onto the binary string representing an individual, and evolving a population of possible segmentations[18]. Other GA approaches for image segmentation include manually-traced contours by Cagnoni et al., methods by Andrey, artificial ant colonies by Ramos, Koza's genetic programming paradigm, Poli's GP work, etc.,

3. Data Mining

Data Mining plays the very important role in the analysis of data. Data Mining is used to extract the hidden data from the large set of database. It is define as a "process of discovering new significant correlation, patterns into massive amount of information stored in warehouse [2]. Data Mining is also known as "Knowledge Discovery from data" (KDD) which consists of the following steps data selection, data cleaning, data integration, data transformation, pattern evaluation and knowledge presentation [5]. Knowledge discovery has been applied in many fields like health care and insurance, medicine, agricultural, assessment of credit card applications and so on. The following figure shows the process of both knowledge discoveries from Fig1.

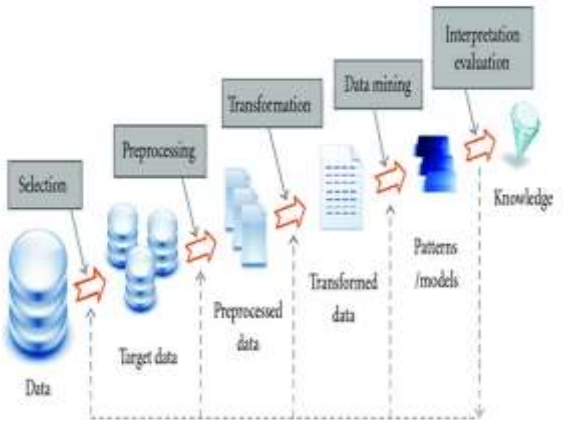

Fig1.1 Process of KDD or Data Mining

4. Clustering

Clustering is the unsupervised method and there is no target field which also uses the bottom-up approach. Clustering is a method for grouping the objects based on its attributes found in the data describing the objects or their relationships. A cluster is a collection of objects in a group which are "similar" to one other and different from the objects in other clusters. An image can be segmented based on its keyword (metadata) or its content (description)[8][7]. The clustering uses the various methods like partitioning, hierarchical, density-based, grid-based, model-based and constraint-based methods and Clustering algorithms can be classified into two main categories Linear clustering algorithms and Non-linear clustering algorithms.

5. Image Edge Detection

Edge detection is the method of identifying edges in the images. Edge detection is used for image segmentation and data extraction in areas such as image processing, computer vision, and machine vision. The use of edge detection method is to reduce unnecessary information in the images and also extract the significant characteristics of the images such as corners, lines and curves. Edge generally occurs between objects and backgrounds, objects and objects, primitives and primitives. Edge is a boundary between two homogeneous regions.

In image processing, most of the shape information of an image is enclosed in edges. So, first we detect these edges in images by using gradient and zero-crossing method and then enhancing those detected areas of image which contains edges, sharpness of the image will increase and image will become clearer. The basic idea of edge detection is as follows: First, use edge enhancement operator to highlight the local edge of the image. Then, define the pixel "edge strength" and set the threshold to extract the edge point set. However, because of the noise and the blurring image, the edge detected may not be continuous [3]. The following flow chart describes the steps for image processing in edge detection Fig2.

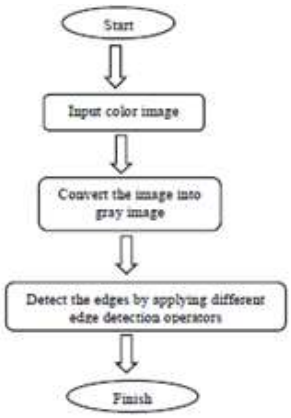

Fig 2: Flow chart for image edge detection

6. Edge Detection Methods

The edge detection uses Gradient and zero-crossing methods.

Gradient Method

The gradient method is one of the methods in image processing techniques. The gradient method detects the edges by looking for the maximum and minimum in the first derivative of the image [1]. Some of the gradient methods are [15][16].

- Prewitt Operator
- Sobel Operator
- Canny Edge Detection
- Roberst Edge Detection

Prewitt Operator

Prewitt operator is used for detecting edges horizontally and vertically. It is limited to 8 possible directions; however knowledge shows that most direct direction estimates are not much more perfect. In Fig3, This gradient based edge detector is estimated in the 3x3 neighborhood for eight directions.

-1	0	+1
-1	0	+1
-1	0	+1

Gx

+1	+1	+1
0	0	0
-1	-1	-1

Gy

Fig 3: Prewitt Operator

Sobel Operator

Sobel edge detection detects all edges of the images which also has some advantage of providing both a differencing and smoothing effect. It is implemented as the sum of two directional edge enhancement operations. The operator consists of a pair of 3×3 convolution kernels as shown in Fig4. The sobel operator is very similar to Prewitt operator. It is also a derivate mask and is used for edge detection. It also calculates edges in both horizontal and vertical direction.

-1	0	+1
-2	0	+2
-1	0	+1

Gx

+1	+2	+1
0	0	0
-1	-2	-1

Gy

Fig 4: Sobel Operator

Canny Edge Detection

Canny used the multi stage algorithm to detect the wide range of edges in the images. The algorithm runs in 5 separate steps:

1. Smoothing: Blurring of the image to remove noise.

2. Finding gradients: The edges should be marked where the gradients of the image has large magnitudes.

3. Non-maximum suppression: Only local maxima should be marked as edges.

4. Double thresholding: Potential edges are determined by thresholding.

5. Edge tracking by hysteresis: Final edges are determined by suppressing all edges that are not connected to a very certain (strong) edge.

Roberts Edge Detection

In fig5, It performs a simple, quick to compute, 2-D spatial gradient measurement on an image. This method emphasizes regions of high spatial frequency which often correspond to edges. The input to the operator is a greyscale image the same as to the output is the most common usage for this technique. This is very similar to sobel operator.

+1	0
0	-1

Gx

0	+1
-1	0

Gy

Fig 5: Roberts Detection

Zero-crossing Method

- Laplacian Operator.

Laplacian Operator

Laplacian Operator is also a derivative operator which is used to find edges in an image. Laplacian is a second order derivative mask. It can be further divided into positive laplacian and negative laplacian. The Laplacian *L(x,y)* of an image with pixel intensity values *I(x,y)* is given by:

$$L(x,y)= \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2}$$

All these masks find edges. Some find horizontally and vertically, some find in one direction only and some find in all the directions.

7. Edge Dtection Based Genetic Algorithm

Genetic algorithms are excellence for searching through large and complex data sets. They also considered capable of finding reasonable solution to complex issues as they are highly capable of solving unconstraint and constraint optimization issues. The GA uses the tree main types of rules at each step to create the next generation from the current population i.e Selection rules, Crossover and Mutation. The selection evaluates each individual and keeps only the fittest ones in the population [19]. In addition to those fittest individuals, some less fit ones could be selected according to a small probability [20]. The others are removed from the current population. The crossover recombines two individuals to have new ones which might be better. The mutation operator induces changes in a small number of chromosomes units. Its purpose

is to maintain the population diversified enough during the optimization process [18]. One reason for using GA is to deal large, complex search spaces where only minimum knowledge is available about the objective function

8. Conclusion

In this paper, studied the most commonly used edge detection techniques of Gradient-based and Laplacian based Edge Detection. Choosing a suitable method for edge detection is based on the some environmental conditions. Each technique has its own advantages and disadvantages. This review paper will be helpful for the researchers in understanding the concept of edge detection.

REFERNCES

[1] Raman Maini and Dr. Himanshu Aggarwal, "Study and Comparison of Various Image Edge Detection Techniques".

[2] Khalid Raza, "Application of data mining in Bioinformatics'', Centre for Theoretical Physics, Jamia Millia Islamia, New Delhi-110025, India, Vol 1 No 2, 114-118.

[3] Wenshuo Gao, Lei Yang, Xiaoguang Zhang, Huizhong Liu, "An Improved Sobel Edge Detection", 2010 IEEE, 978-1-4244- 5540-9/10.

[4] R.V.V.Krishna, Member, IAENG and S. Srinivas Kumar, "Hybridizing Teaching Learning Based Optimization with Genetic Algorithm for Colour Image Segmentation", the International Multi Conference of Engineers and Computer Scientists 2017 Vol I, IMECS 2017, March 15 - 17, 2017, Hong Kong.

[5] "An Overview of Knowledge Discovery Database and Data Mining Techniques", International Journal of Innovative Research in Computer and Communication Engineering, Vol.2,March 2014.

[6] Priyansh Sharma and Jenkin Suji, "A Review on Image Segmentation with its Clustering Techniques", International Journal of Signal Processing, Image Processing and Pattern Recognition Vol.9, No.5 (2016), pp.209-218.

[7] S. Bhowmik and V. Datta, "A Survey on Clustering Based Image Segmentation", International Journal of Advanced Research in Computer Engineering & Technology, vol. 1, no. 5, (2012) July, pp: 280-284.

[8] Santanu Bhowmik, Viki Datta, "A Survey on Clustering Based Image Segmentation", International Journal of Advanced Research in Computer Engineering & Technology Volume 1, Issue 5, July 2012.

[9] Pascal Fua and Yvan G. Leclerc, "Model Driven Edge Detection", Artificial Intelligence Center, SRI International, 333 Ravenswood Avenue, Menlo Park, California, USA

[10] Massimo Minervini , Andreas Fischbach , Hanno Scharr , Sotirios A. Tsaftaris, "Finely-grained annotated datasets for image-based plant phenotyping", Pattern Recognition and Image Analysis Research Unit, IMT Institute for Advanced Studies, 55100 Lucca, Italy, Institute of Bio- and Geosciences: Plant Sciences (IBG-2), Forschungszentrum Jülich GmbH, 52425 Jülich, Germany and Institute for Digital Communications, School of Engineering, The University of Edinburgh, Edinburgh EH9 3JL, UK.

[11] Abramov, Florentin Wörgötter, Hanno Scharr, Andreas Fischbach, Babette Dellen, " Modeling leaf growth of rosette plants using infrared stereo image sequences", Georg-August-Universität Göttingen, BCCN, Department for Computational Neuroscience, Inst. Physics-3, Friedrich-Hund Platz 1, D-37077 Göttingen, Germany, IBG-2: Plant Sciences, Forschungszentrum Jülich, 52425 Jülich, Germany and Institut de Robòtica i Informàtica Industrial (CSIC-UPC), Llorens i Artigas 4-6, 08028 Barcelona, Spain.

[12] J.Jeya Priyankha1 and K.Suresh kumar2, "Crop Disease Identification Using a Feature Extraction HOG Algorithm", Asian Journal of Applied Science and Technology (AJAST) Volume 1, Issue 3, Pages 35-39, April 2017.

[13] Arti Singh, Baskar Ganapathysubramanian and Asheesh Kumar Singh, "Machine Learning for High-Throughput Stress Phenotyping in Plants".

[14] Jyoti Deshmukh and Udhav Bhosle, "Image Mining using Association Rule for Medical Image dataset", International Conference on Computational Modeling and Security (CMS 2016).

[15] Muthukrishnan.R and M.Radha, "Edge Detection Techniques for Image Segmentation", International Journal of Computer Science & Information Technology (IJCSIT) Vol 3, No 6, Dec 2011.

[16] Pooja Sharma, Gurpreet Singh , Amandeep Kaur, " Different Techniques Of Edge Detection In Digital Image Processing", International Journal of Engineering Research and Applications (IJERA) ISSN: 2248-9622.

[17] M.K. Lee, S.W. Leung, T.L. Pun, H.L. Cheung, A.M.K. LEe, "Edge detection by genetic algorithm", IEEE **ISSN:** 1522-4880.

[18] M. Abdulghafour, "Image segmentation using Fuzzy logic and genetic algorithms", Journal of WSCG, vol. 11, no.1, 2003

[19] A. Borji, and M. Hamidi, "Evolving a Fuzzy Rule-Base for Image Segmentation", International Journal of Intelligent Systems and Technologies, 2007, pp.178-183.

[20] Xian Bin Wen, Hua Zhang and Ze Tao Jiang, "Multiscale Unsupervised Segmentation of SAR Imagery Using the Genetic Algorithm", Sensors, vol.8, 2008, pp.1704-1711.

[21] George Karkavitsas and Maria Rangoussi," Object Localization in Medical Image Using Genetic Algorithms", International Jnl. of Signal Processing, 2005, pp.204-20