Comparitive Analysis Of Scaling Factor In Aerial View Using DBscan And Data Grid Algorithm

M.Mayilvaganan, R.Manickam

Associate Professor, Dept.of Computer Science, P.S.G College of Arts and Science, Coimbatore dev7aki@gmail.com
Research Scholar(Ph.D), Dept.of Computer Science, P.S.G College of Arts and Science, Coimbatore.

ABSTRACT- The key problem in achieving efficient and develop user-friendly retrieval in the domain of image processing and also to estimate the density (height)of the snowfall in cross section view with scaling factors and density based algorithms. In this research work, an aerial digital images can be gathered from various background situations with 7.2MP resolution. After pre-processing the snowfall image, the proposed methodology can be implemented in order to identify the snowfall region using DBSCAN and data grid clustering algorithms. Using the clustering algorithm, the snowfall regions are clustered and segmented according to their size. After clustering the snowfall area, using binarization technique the total number of black pixels(1's) snowfall area are computed automatically. Final objective is to identify the suitable cluster density for landing the flight with respect to minimum density region and size for different sets of images in the snowfall occurrence area. The efficiency of the proposed methodology have been compared with DBSCAN and Data Grid algorithm with respect to memory usage and running time of the algorithms for different sets of aerial images.

Keywords- Binarization, clustering, DBSCAN , Data Grid algorithm, scaling factors.

I.INTRODUCTION

Contemporary methods for assessing the presentation of image mining algorithms are based on comparisons of one algorithm over the previous using the identical image datasets. This has led to diverse conclusions where now and again one of the algorithms is accessible as the best, while in other publications that equivalent algorithms performed in a different way. It is understood that some algorithms are finest appropriate to a particular type of image and that they will perform better when tested on these images. The projected study will based its evaluation on the use of diverse sets of images. Image Mining can do on dissimilar type of images like real time image, satellite image, and also in medical images. The Image Processing involves an assortment of steps namely; Image pre processing, Restoration, Analysis and Compression. Pre processing includes numerical correlation and radiometric correlation. The associated image is then fed for re-establishment task. In this research work, an aerial digital images can be gathered from various flight run way occurrence with snowfall as background situations.

The organization of the paper is structured as follows. Chapter 2 explains about the research problem and Data for Research .Chapter 3 demonstrates the proposed methodology. Chapter 4 reveals the Results and Discussion for image clustering and scaling factors techniques. Finally, Chapter 5 concludes the paper.

II.DATA FOR RESEARCH

In this research work, an aerial digital images can be gathered from various background situations with 7.2MP resolution. The original images where resized to a lower resolution of approximately 457x630 pixels so the algorithms chosen can process them more efficiently.

Figure 1 (a) and (b) shows the image datasets used for this study.

ISSN 2091-2730

Fig 1(a) set1 image with snow

Fig 1(b) set2 image with snow

III. PROPOSED METHODOLOGY - GRID DENSITY PARAMETER ESTIMATION

The Grid Density is a base algorithm of density of cells based clustering. It requires user specified two global input parameters i.e. MinPts and Eps. The density of an object is the number of objects in its Eps-neighborhood of that object. Grid Density can specify upper limit of a core object i.e. how much objects may present in its Eps-neighborhood. So due to this, the grid density of clusters are detected by it, are having wide variation in local density and forms clusters with any arbitrary shape. Such clusters may be represented by several smaller clusters so that each cluster may have reasonably uniform density.

Grid Clustering Flowchart

The following diagram(2) represents the flowchart for Grid density based clustering for grouping of snowfall region according to black and white pixels in the flight runway area.

<u>Distance between Black pixels(1's)</u>

To clusters a dataset, the Grid Density implementation starts by identifying the k nearest neighbours black pixels(1's) of each point and identify the farthest k nearest neighbour (in terms of Euclidean distance ki). Let's consider O1 and O2 be two black pixels intensity values from the aerial digital images and k be the minimum distance between black pixels from the universe of possible objects. The distance between O1 and O2 is denoted by distance (O1,O2) or d(O1,O2) .

$$\text{Euclidean distance} = (O_i,O_j) = \text{sqrt}(\Sigma^n(O_ik - O_jk)^2$$

$$= \text{sqrt}(5-3)2 + (6-9)2 + (4-3)2 + (9-2)2 \quad = 8.25 \text{ cm}$$

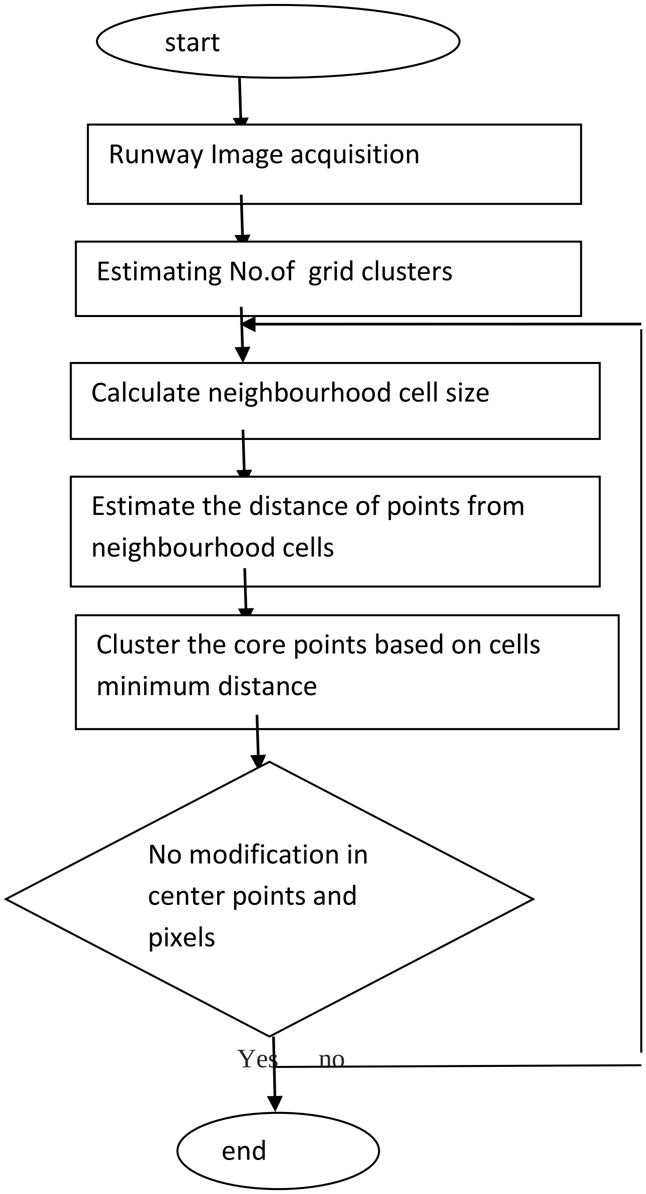

Fig 2: density clustering flowchart

After determining the average distance(Eps) values, there is a need to estimate the value of the minimum density (MinPts) is the immediate and urgent task. So firstly, the number of average data objects in Eps neighborhood of every point in dataset is calculated one by one.

In the fig (3) the joining or tree clustering method uses the black pixels when forming the clusters with respect to size, shape and density of pixels. It can be represented using Euclidean distance as follows.

The mathematic expectation of all these data objects is calculated, which is the value of MinPts.

Minpts = 1/n $\Sigma^{n}_{i=1} P_i$

Where pi is the number of points in Eps neighborhood of point i. So for each different value of Eps we will get corresponding Minpts value.

Distance (Oi,Oj) = 1 / n | $O_{ik} - O_{jk}$ |

= 1 / 4 (|5 – 3| +|6-9|+|4-3|+|9-2|

= 2.06 cm

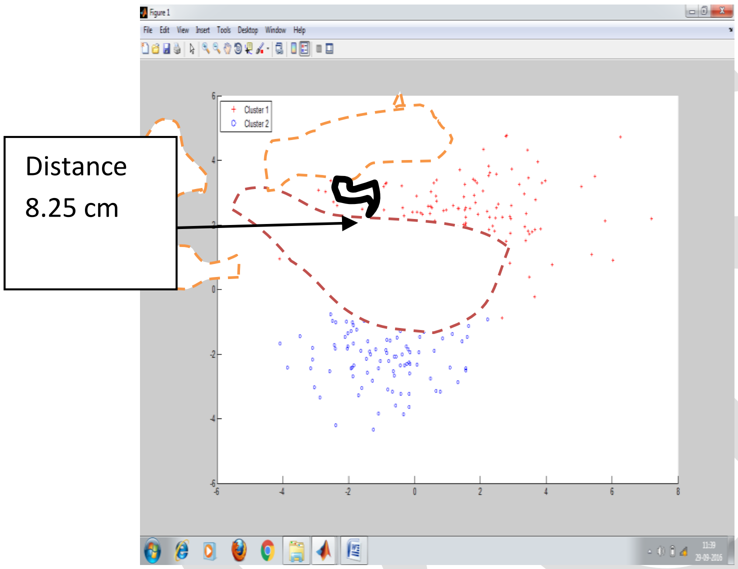

Fig 3 distance between two pixels

Average distance can be estimated the intensity difference across black and white pixels. It can be represented using the following figure(4) and formula. Consider the total no.of image data set n = 4 images and distance between two pixels $O_{ik} - O_{jk}$ respectively.

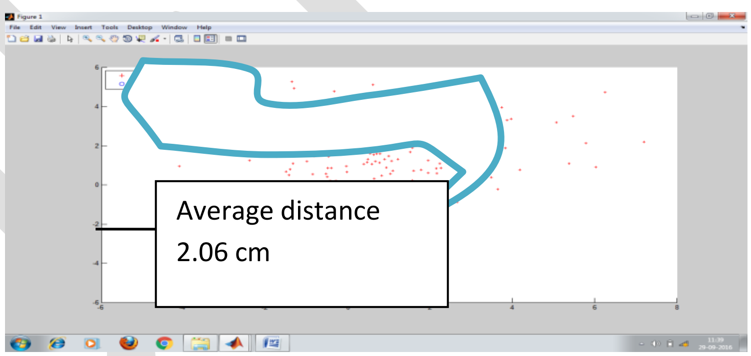

Fig 4 average distance between pixels

IV. IMAGE SCALING FACTORS

After estimating the total no. of black pixels, then image scaling factors can be estimated using the focal length of the camera and total no of black pixels in the image. It can be represented in the using the following image scaling factors method.

regionQuery(P, eps)

Total occurrence of snowfall Real area = primary area * m

M = 1 / image scaling factor

Image scaling(I) = focal length(fl) of the camera / height(h) of snowfall clustering.

return all points within P's eps-neighborhood (including P)
Image scaling(I) = focal length(fl) of the camera / height(h) of snowfall clustering.

= 100 mm / 14,326 = 0.0024250 mts

Co-efficient of the Image

After estimating the image scaling factors then co-efficient of the image can be estimated using image scaling factor. It can be estimated using the following method.

regionQuery(P, eps)

Total occurrence of snowfall Real area = primary area * m

M = 1 / image scaling factor.

Image scaling(I) = focal length(fl) of the camera / height(h) of snowfall clustering.Co-efficient of the image can be calculated using the following formula,

M = 1 / image scaling factor == 1 / 0.0024250 = 412.37 metres

The following diagram(5) illustrates the distance from the runway run way edge along with the image scaling factors. The coefficient of the image scaling factors can be estimated using focal length of the camera and the height of the snowfall region. In the following diagram the image scaling factors are estimated as 41.237 meters.

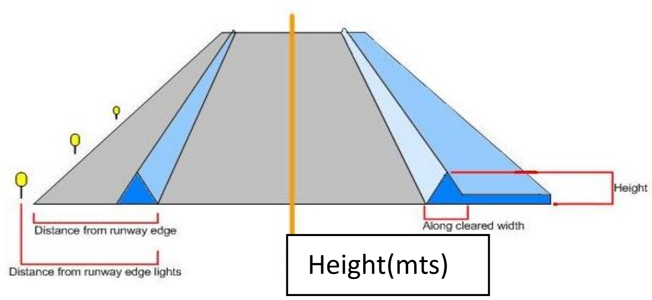

Fig 5: run way width estimation using image scaling factors.

Overall coverage of snowfall area

After estimating the co-efficient of the image then compute the total primary area coverage using the image scaling factor. The height of the snowfall region can be estimated using the Minpts(least density) and Eps(average distance between pixels)regionQuery(P, eps) can be represented in the figure(6).

Total occurrence of snowfall Real area = primary area * m

M = 1 / image scaling factor

Image scaling(I) = focal length(fl) of the camera / height(h) of snowfall clustering.

Hence the overall snowfall occurrence area can be calculated using the following formula,

Total occurrence of snowfall Real area = primary area * m = 412.37 * 0.0024250 = 0.989688 hectares

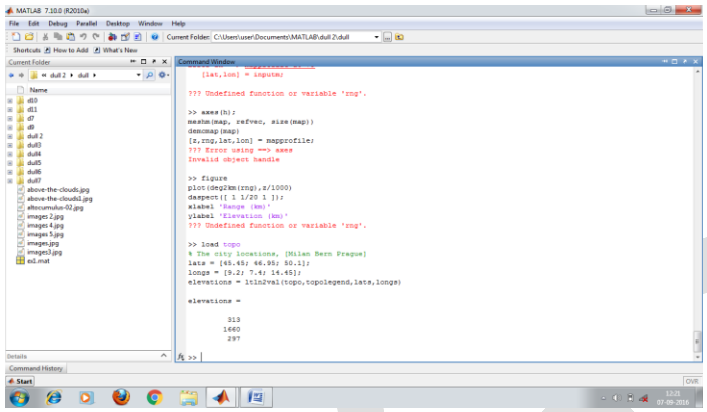

Fig 6 DBSCAN estimation of snowfall density and primary area coverage

Range of clusters shaped is compared with the size of image. DBSAN is performed when the volume of any of the clusters created is greater than 1/10th of image . if the probability density mixture of black pixel cluster size become less than 1/10th of image the iteration stops. The figure (7) represents the black and white pixel formation of cluster using dbscan clusteing. Blue and green circle represents the black pixels in the snowfall area and yellow and red colour represents the snowfall density height and total coverage area of snowfall region. The circumference of the total snowfall area are calculated using image scaling factors.By simply calculating the primary area and co-efficient 'm' ,the total snowfall occurrence area is obtained in 0.989688 hectares with respect to least density 0.002 mts.

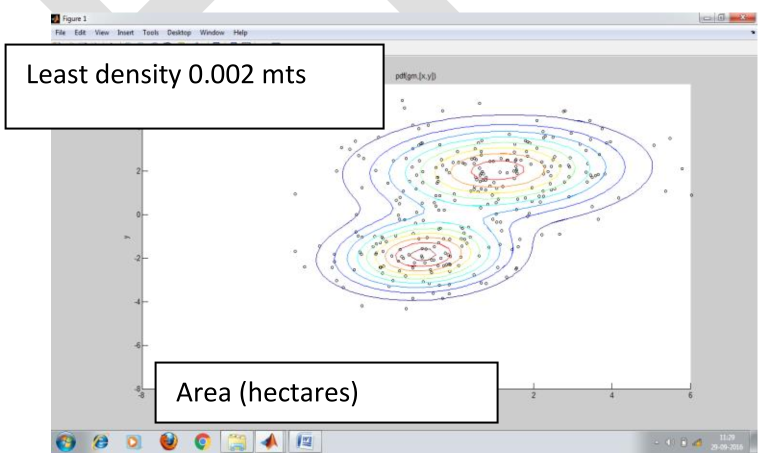

Fig 7representation of least density and total area coverage of snowfall

V CONCLUSION & FUTURE WORK

This proposed work which is useful for analysing and estimating the cluster segmentation of snowfall area using Datagrid algorithm. To evaluate DBSCAN algorithm, the snowfall occurrence aerial image was taken in flight runway area using the camera with the focal length 100mm.The proposed design can be used for analyzing the density of snowfall occurrence and estimated the spread over circumference area of snowfall region , and a framework of methodology has been developed for analyzing the aerial image sequence for a step by step process. The proposed technique is very useful, in order to know the snowfall density and also estimate the primary area coverage of snowfall with respect to segmentation and clustering. Regarding the possibilities of future research on the same lines, the current research can be extended to and evaluate the work in the following areas, Extend the proposed methodology focal length of the camera above 100mm.This promises a great scope for further research on these lines.

REFERENCES:

[1]. Density-Based Clustering Algorithms, research work, Lahore College for Women University Lahore, Pakistan.

[2]. Pasi Fränti. Number of clusters (validation of clustering), Speech and Image Processing Unit School of Computing University of Eastern Finland.

[3] R. Agrawal, J. Gehrke, D. Gunopulos, and P. Raghavan.
Automatic subspace clustering of high dimensional data for data mining applications. SIGMOD'98.

[4]. C. Rother, V. Kolmogorov, T. Minka, and A. Blake. Cosegmentation of image pairs by histogram matching – incorporating a global constraint into mrfs. In *CVPR*, 2006.

[5]. Russell, W. T. Freeman, A. A. Efros, J. Sivic, and A. Zisserman. Using multiple segmentations to discover objects and their extent in image collections. In *CVPR*, 2006.

[6]. J. Shawe-Taylor and N. Cristianini. *Kernel Methods for Pat-tern Analysis*. Camb. U. Press, 2004.

[7]. J. Shi and J. Malik. Normalized cuts and image segmentation. *IEEE Trans. PAMI*, 22(8):888–905, 1997.

[8]. J. Winn and N. Jojic. Locus: learning object classes with unsupervised segmentation. In *ICCV*, 2005.

[9]. L. Xu, J. Neufeld, B. Larson, and D. Schuurmans. Maximum margin clustering. In *NIPS*, 2005.

[10].K. Mumtaz1 and Dr. K. Duraiswamy. A Novel Density based improved k-means Clustering Algorithm – Dbkmeans. (IJCSE) International Journal on Computer Science and Engineering, Vol. 02, No. 02, pp. 213-218, 2010.

[11]. Duchenne, I. Laptev, J. Sivic, F. Bach, and J. Ponce. Automatic annotation of human actions in video. In *ICCV*, 2009.

[12]. P. F. Felzenszwalb and D. P. Huttenlocher. Efficient graphbased image segmentation. *International Journal of Computer Vision*, 59, 2004.

[13].M. X. Goemans and D. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM*, 42:1115–1145, 1995.

[14].T. Hastie, R. Tibshirani, and J. Friedman. *The Elements of Statistical Learning*. Springer-Verlag, 2001.