

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

## Survey on Image Segmentation using Improved DBSCAN Algorithm and Edge Detection

Hrishikesh Kamble<sup>1</sup>, Sanket Rane<sup>2</sup>, Prince Raj<sup>3</sup>, Yogesh Pawar<sup>4</sup>, Shruti Choudhari<sup>5</sup>

Student, Department of Information Technology, D. Y. Patil Institute of Engineering & Technology, Pune, India<sup>1</sup>

Student, Department of Information Technology, D. Y. Patil Institute of Engineering & Technology, Pune, India<sup>3</sup>

Student, Department of Information Technology, D. Y. Patil Institute of Engineering & Technology, Pune, India<sup>3</sup>

Professor, Department of Information Technology, D. Y. Patil Institute of Engineering & Technology, Pune, India<sup>4</sup>

Professor, Department of Information Technology, D. Y. Patil Institute of Engineering & Technology, Pune, India<sup>5</sup>

**ABSTRACT:** The process of partitioning a digital image into multiple segments is called as image segmentation. Segmentation can be correlated with clustering which does the task of grouping objects with similar properties. Edge detection is one application of segmentation. Edge detection aims at finding regions in the digital image where the intensity or brightness changes sharply. One of the algorithms extensively used for clustering is the Density-Based Spatial Clustering of Applications with Noise i.e DBSCAN algorithm, which is able to recognize clusters of any size or shape. An improvement suggested in this algorithm is to enhance the computational time by making use of grid cells. So in this paper, we propose to make use of Depth First Search(DFS) to improve the computational time as well as memory efficiency of the algorithm.

**KEYWORDS:** DBSCAN(Density-Based Spatial Clustering of Applications with Noise), DFS(Depth First Search), Clustering, Segmentation, Grid, Edge Detection.

### I. INTRODUCTION

Data clustering refers to grouping of data based on similarities. Such grouping is useful for analysis of data for future in-sights. One of the algorithms extensively used is the Density-Based Spatial Clustering of Applications with Noise i.e DBSCAN algorithm, which is able to recognize clusters of any size or shape. One of the main advantages of this algorithm is that we need not specify the number of clusters in advance which is a drawback of some traditional clustering algorithms.

The process of partitioning a digital image into multiple segments is called as image segmentation. Basically image segmentation is done to obtain a more meaningful representation of the image. The goal of segmentation is to simplify and/or change the representation of an image into something that is more meaningful and easier to analyse.[1][2] Image segmentation is typically used to locate objects and boundaries (lines, curves, etc.) in images. The set of partitions obtained after segmentation cover the entire image and each partition contains pixels that are similar to each other with respect to some property such as color or intensity.

Thus segmentation can be correlated with clustering which does the same task of grouping objects with similar properties. Segmentation is generally done to locate different objects or boundaries. Edge detection is one application of segmentation. Edge detection aims at finding regions in the digital image where the intensity or brightness changes sharply.

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

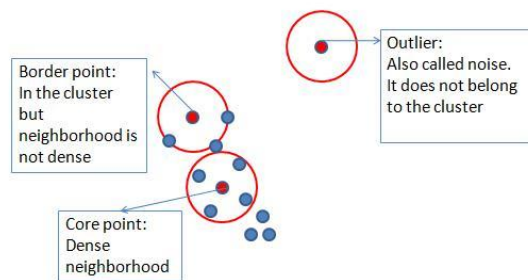
Thus the combined application of above concepts would be Segmentation of Image using modified DBSCAN clustering algorithm and edge detection.

## II. ORIGINAL DBSCAN

DBSCAN [3] is the first density based clustering algorithm. It was proposed by Ester et al. in 1996, and it was designed to cluster data of arbitrary shapes in the presence of noise in spatial and non-spatial high dimensional databases. The key idea of DBSCAN is that for each object of a cluster the neighbourhood of a given radius (Eps) has to contain at least a minimum number of objects (MinPts), which means that the cardinality of the neighbourhood has to exceed some threshold.[4]

Input parameters to the algorithm are:

- Eps: Maximum Radius of the neighbourhood.
- MinPts: Minimum number of points in the Eps-Neighbourhood.



(Fig.1)

### A. Some related definitions

1. Cluster: Cluster is a group of objects that belongs to the same class. In other words, similar objects are grouped in one cluster and dissimilar objects are grouped in another cluster.
2. A density-based notion of a cluster: A cluster is defined as a maximal set of density-connected points.
3. Eps-neighbourhood of a point p: Is the circular area with p as center and with Eps as radius, p belongs to dataset D. The set of points included in the Eps-neighbourhood is denoted by  $N_{Eps}(p) = \{q \in D \mid \text{dist}(p, q) \leq Eps\}$  [5]
4. Density of point p: Is the number of points within Eps-neighbourhood of point p, p belongs to dataset D. [5]
5. Core point: Is the point of which the density is greater than or equal to MinPts, as MinPts is a threshold which is user specified. [5]
6. Border Point: Border point q is the point which is not core point but is contained in the Eps-neighbourhood of the other core point. [5]
7. Noise point (outlier): Is the point which is neither core point nor Border point. [5]
8. Directly density-reachable: let q be a core point, A point p is directly density reachable from a point q if p belongs to the eps-neighbourhood of point q, which is denoted by  $(p) N_{eps}(q)$   $jN_{eps}(q) \geq M \text{ in } P \text{ ts } [5]$
9. Density Reachable: A point p is density Reachable from point q if there is a chain of points  $p_1; p_2; \dots; p_n$  such that  $p_1 = q; p_n = p$   $p_i \in D, 1 \leq i \leq n$ , and  $p_{i+1}$  is directly density-reachable from  $p_i$ . [5]
10. Density Connected: a point p is density connected to point q for the given Eps and MinPts if there is a point o such that both p and q are density-reachable from o for the given Eps and MinPts. [5]

## International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018

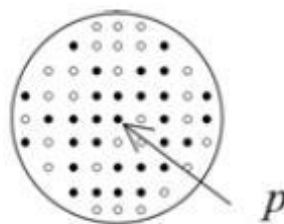
### B. Algorithm

1. Arbitrarily select a point  $p$ .
2. Retrieve all points density-reachable from  $p$  w.r.t Eps and MinPts
  - If  $p$  is a core point, a cluster is formed
  - If  $p$  is a border point, no points are density-reachable from  $p$ , and DBSCAN visits the next point of the database description
3. Continue the process until all of the points have been processed

### III. DBSCAN FOR IMAGE SEGMENTATION

DBSCAN for color image segmentation [6] is as follows:

- The spatial neighbourhood of a given pixel is called the spatial Eps of the pixel. It is named as SpatialEps. SpatialEps is set as a circle. Fig 2 illustrates the SpatialEps of a pixel  $p$ . The size of SpatialEps is the number of all pixels in the circle. Blackpoints in the circle represent pixels which are color similar with  $p$ .
- The neighborhood in color space of a given pixel is called color Eps of the pixel. It is named ColorEps. ColorEps is set as an ellipsoid in Munsell (HVC) color space. Fig 3 illustrates the ColorEps of a pixel  $p$ . ColorEps is used to judge if pixels are similar in color with  $p$  or not. Pixels within the ellipsoid are color similar with  $p$  and pixels outside are color dissimilar with  $p$ .
- If the SpatialEps of a pixel contains at least a minimum number, MinPts, of pixels similar in color, then the pixel is called a core pixel. Its SpatialEps forms a core region. When we are using density-based clustering to find clusters in an image, we need to know there how many pixels that are color similar with  $p$  in its SpatialEps. Then we project all the pixels in  $p$ 's SpatialEps to its ColorEps, if the number of pixels within the ColorEps is larger than MinPts,  $p$  is called a core pixel and its SpatialEps forms a core region.



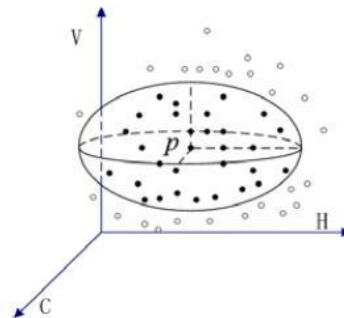
(Fig.2) SpatialEps

# International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

Vol. 7, Issue 12, December 2018



(Fig.3) ColorEps

- For the set of pixels in an image,  $I$ , we say that a pixel  $p$  is directly density-reachable from pixel  $q$  if  $p$  is within the SpatialEps of  $q$ , and  $q$  is a core pixel (Fig. 3).
- Density reachability and density-connectivity for pixels have the same definitions as in section 2. The color similarity between color pixels is transitive as density reach-ability is transitive. Transmissibility of color similarity makes it possible that pixels that are gradually changing in color and connected in spatial will be clustered into the same region. This is consistent with human perception.

## A. Algorithm[6]

Based on the above definitions and specifications, we prescribe our segmentation method as follow.

1. Search the unlabeled pixels in an image in order for current core pixel and current core region. The order is from the top left corner to the bottom down corner of the image.
2. If a core pixel  $p$  is found, a new cluster is created. Then, we iteratively collect unlabeled pixels that are density-connected with  $p$ , and Label these pixels with same cluster label.
3. If there are still existing core pixels in the image, goto 2.
4. For the pixels that are not included in any clusters, merge them with the cluster that is adjacent to them and has the highest similarity in average color value with them.
5. Label each cluster we find in the image as a segmentation region. After the above segmentation process, pixels are put into different clusters and form different segmented regions of an image.

## IV. MEDIAN FILTERING

The median filter is a nonlinear digital filtering technique which is used for removing noise from an image. It is a typical pre-processing step to improve the results of later processing which may be edge detection on an image. Median filtering is extensively used in digital image processing because, under certain conditions, it preserves edges while removing noise. [7]

Median filter runs through the image pixel by pixel, replacing each pixel with the median of neighbouring pixels. The pattern of neighbours is called the “window” that slides pixel by pixel, over the entire image.

It is a type of smoothing technique. Generally all smoothing techniques are effective at removing noise from the image, but largely affect edges. Usually alongside removing the noise, it is necessary to preserve the edges. Edges make a huge difference to the visual appearance of images. Median Filter considerably works well on speckle noise and salt-and-pepper noise (impulsive noise) and hence it is very widely used in digital image processing. [8]

# International Journal of Innovative Research in Science, Engineering and Technology

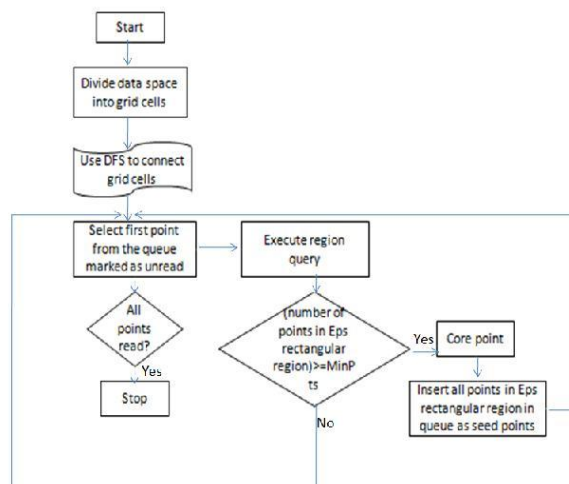
(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: [www.ijirset.com](http://www.ijirset.com)

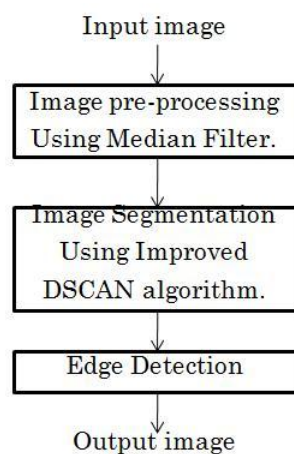
Vol. 7, Issue 12, December 2018

## V. PROPOSED SYSTEM

Image segmentation finds its applications in medical field for tumour detection, for facial recognition and many such fields. Over the years, there have been many modifications in the DBSCAN algorithm. We propose to use a modified grid based DBSCAN algorithm for segmenting a image. Median filtering would be used for preprocessing followed by segmentation of the image and finally edge detection.



(Fig. 4) proposed DBSCAN algorithm



(Fig. 5) proposed system

## VI. CONCLUSION

Image segmentation is done to obtain a more meaningful representation of the image. Furthermore, Edge Detection enhances this representation. DBSCAN can find clusters with arbitrary shape and size without needing to specify the

# International Journal of Innovative Research in Science, Engineering and Technology

*(A High Impact Factor, Monthly, Peer Reviewed Journal)*

Visit: [www.ijirset.com](http://www.ijirset.com)

**Vol. 7, Issue 12, December 2018**

number of clusters in advance. DBSCAN based on grid cells promises to give a better efficiency by using Depth First Search for connecting the grid cells.

## REFERENCES

- [1] Linda G. Shapiro and George C. Stockman "Computer Vision", pp 279-325, New Jersey, Prentice-Hall, ISBN 0-13-030796-3.
- [2] Barghout, Lauren, and Lawrence W. Lee. "Perceptual information processing system. ", Paravue Inc. U.S. Patent Application 10/618,543, filed July 11, 2003.
- [3] M.Ester, H.P.Krigel, J.Sander, and X.Xu, "A Density Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise", Proc. of the 2nd International Conference on Knowledge Discovery and Data Mining, Portland, WA, 1996, pp. 226-231.
- [4] Kamran Khan, Saif Ur Rehman, Simon Fong , S.Sarasvady, "DBSCAN: Past, Present and Future", 978-1-4799-2259-14 2014.
- [5] LI Mengao, MENG Dongxue, GU Songyuan, LIU Shufen, "Research And Improvement of DBSCAN Cluster Algorithm ", in 2015 7th International Conference on Information Technology in Medicine and Education, 978-1-4673-8302-8/15.
- [6] Qixiang Ye, Wen Gao, Wei Zeng "Color image segmentation using density based clustering ", ISBN 0-7803-7663-/03, IEEE 2003.
- [7] wikipedia "[https://en.wikipedia.org/wiki/Median filter](https://en.wikipedia.org/wiki/Median_filter) "
- [8] G. R. Arce "Nonlinear Signal Processing: A Statistical Approach ", Wiley:New Jersey, USA, 2005