An Improvement of DBSCAN Algorithm to Obstacles Detection for Stop&Go

Wanglong CHENG[1,a], Jian WU[2*]

[1]State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun, 130022, China

[2]State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun, 130022, China

[a]email: chengwangl@126.com, *Corresponding author, email: wujian@jlu.edu.cn

Keywords: DBSCAN; Stop&Go; Laser radar; Obstacles detection

Abstract. Obstacles detection is one of key technologies to realize Stop&Go cruise control. This paper presents an improved DBSCAN(Density-Based Spatial Clustering of Application with Noise) clustering algorithm for obstacles detection, which based on two-dimensional laser radar. The algorithm by improving the traditional DBSCAN data point search way and key parameter setting, improve the operation speed and anti-noise performance of the algorithm. Real vehicle test show that the algorithm in the actual application effect is good.

Introduction

Traditional Adaptive Cruise Control (ACC) is mainly in view of the high speed working condition, and in the face of urban low speed, high density conditions the system could not work effectively. Therefore developed Stop&Go Cruise Control for the working condition of city. The key technologies for Stop&Go including environmental information collection and processing, safe distance calculation and cruise control algorithm. Obstacle detection is an important part in the environmental information collection. With wide ranging , the advantages of high precision, suitable for real-time processing [1], the laser radar sensor can well meet the requirements of Stop&Go, so this paper uses the SICK LMS511 two-dimensional laser radar as information sensing sensor. Clustering is an important tool in the field of data mining, in image processing, pattern recognition, and other fields have a wide range of applications. The principles of clustering is to divide similar samples into several categories. DBSCAN clustering algorithm can discover clusters of arbitrary shape, but there are defects such as sensitive parameter setting, so this paper proposes an improved DBSCAN algorithm for clustering of the radar data, and realize the obstacles detection.

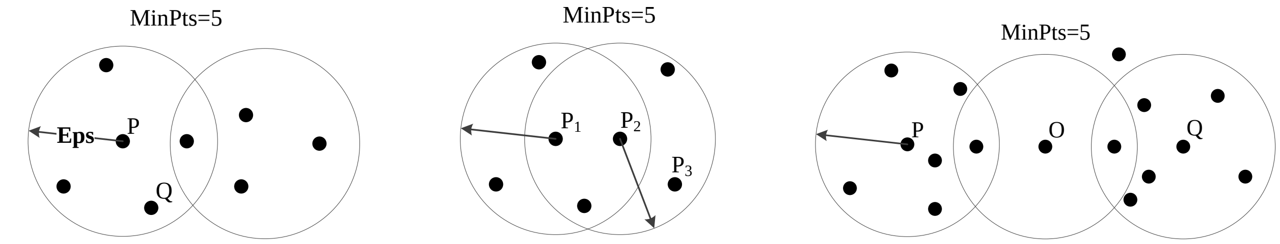
DBSCAN

DBSCAN is a density-based clustering algorithm. DBSCAN will divided area which have enough samples into different clusters, and discover arbitrary clusters in the spatial database which has noise[3-4]. DBSCAN algorithm used two important input parameters Epsilon(Eps) and minimum point(MinPts). The basic ideas of DBSCAN involve a number of definitions, which are presented below[5] and shown in Figure 1.

- ε-neighborhood: The neighborhood within a radius ε(Eps) of any point, P, is called the ε-neighborhood of P, is defined as following:
  $$N_{Eps}(p) = \{q \in D \mid dist(p,q) \le Eps\} \tag{1}$$
- Core Point: If the ε-neighborhood of an given point, Q, is greater than or equal to MinPts, then Q is called a core point.
- Directly Density-reachable: If P is within ε-neighborhood of Q, and Q is a core point, then we say that P is directly density-reachable from Q.
- Density-reachable: If there is a chain of points P_1, ... ,P_n, P_1=P and P_n=Q meanwhile P_i is

directly density-reachable from P_{i+1}($1 \le i \le n$), then we say P is density-reachable from Q.

- Density-connected: P is density-connected to Q express that there is a point, O, is density-reachable from both P and Q with respect to Eps and MinPts.
- Cluster: Cluster is a maximal set of density-connected points.



(a) Directly Density-reachable (b) Density-reachable (c) Density-connected

Figure 1. Definitions of DBSCAN

DBSCAN searches for core points by checking the ε-neighborhood of each point in the database. If the point P is a core point, then looking for all density-connected points from P, meanwhile find out all core points within the ε-neighborhood of P and all points which density-connected to these core points. The next step is clustering these points to a cluster. After that, DBSCAN checks the remaining point and repeats the above steps. The process terminates when no new points can be added to any cluster. The points not belong to any clusters is considered to be noises.

DBSCAN has several advantages, such as any shape of clusters can be represented, noise can be detected well, and the speed of clustering is fast. But there are some disadvantages. The algorithm is sensitive to parameters. For the low density part in database, the value of Eps is too small will classify the useful data as noise, on the contrary, it could divide the noise into a cluster. So it works badly for the uneven distribution of database. On the other, for a large database the neighborhood query becomes time consuming and affects the real-time performance.

Improvement of DBSCAN

The data characteristics of two-dimensional laser radar which used in this paper called LMS 511 as following:

- With the change of environment ,the number of clusters is random.
- The data in polar coordinate and each angle corresponds to a distance value, so there is no data overlap. The number of data in a single pass is fixed as the angular resolution is determined.
- The density of data varies with the distance between obstacles and the host car.

Based on the characteristics of DBSCAN and the data of LMS511, this paper improves DBSCAN algorithm by improving the query mode and parameter setting.

Improvement of query mode. DBSCAN need to query neighborhood constantly, so reducing the frequency of neighborhood query can improve the speed of clustering. This paper improve the query mode based on this principle. First, the data are sorted by scanning angle. The new algorithm query neighborhood start from the initial point, P. If P is a core point, then create a cluster marked by p and class all points within ε-neighborhood of P to the cluster. Next, query neighborhood of Q, which has not been marked, there are several situations as following:

- If Q is a core point and the ε-neighborhood of P and Q without overlapping, then the algorithm creates a new cluster marked by **q**.
- If Q is a core point and the ε-neighborhood of P and Q overlap, meanwhile there are some core points in the overlapping area, then class all points within ε-neighborhood of Q to the cluster **p**.
- If Q is a core point and the ε-neighborhood of P and Q overlap, meanwhile there is no point in the overlapping area, then the algorithm creates a new cluster marked by **q** and class points which nearer Q in the overlapping area to **q**.
- If Q is not a core point and the ε-neighborhood of P and Q overlap, meanwhile there are some core points in the overlapping area, then class the point Q to cluster **p**.
- If Q is not a core point and the ε-neighborhood of P and Q without overlapping, or there is no point in the overlapping area, then Q marked as noise temporary.

Repeat the process until all points have been processed.

Improvement of parameter. A fixed Eps can not suit to all database, hence, according to the characteristics of two-dimensional laser radar data, this paper presents the Eps value will be calculated as following:

$$Eps = \alpha \frac{\pi[\mathrm{int}(\rho_i / L) + 1]}{180} \theta_{res} \tag{2}$$

Where α is a weighting factor, int() means taking the integer portion,ρi denotes the distance between the point i and the origin of the laser radar, L signifies clustering step length and θres is angular resolution. The improved algorithm calculates Eps value according to the distance between the data points and the origin of the laser radar, so that it can suit to different database.

The algorithm flow chart as show in figure 2.

Real Vehicle Tests

In this paper, the two-dimensional laser radar which is shown in Figure 3 is installed on Volkswagen Touran. The mounting position on the vehicle is shown in Figure 4. The angular resolution is set to 0.5°, the scan range from -5° to 185° and the radar frequency is 50Hz. Under the environment of campus, the experiment is designed to test the practical effect of this algorithm. Experimental parameters setting: MinPts = 5, α= 200. A random frame of data intercepted from trials was processed by Matlab to validate the algorithm. A rectangular coordinate was built, in which radar's position was regarded as the origin, horizontally to the right as the positive direction of X-axis, vehicle's driving direction as the positive direction of Y-axis.

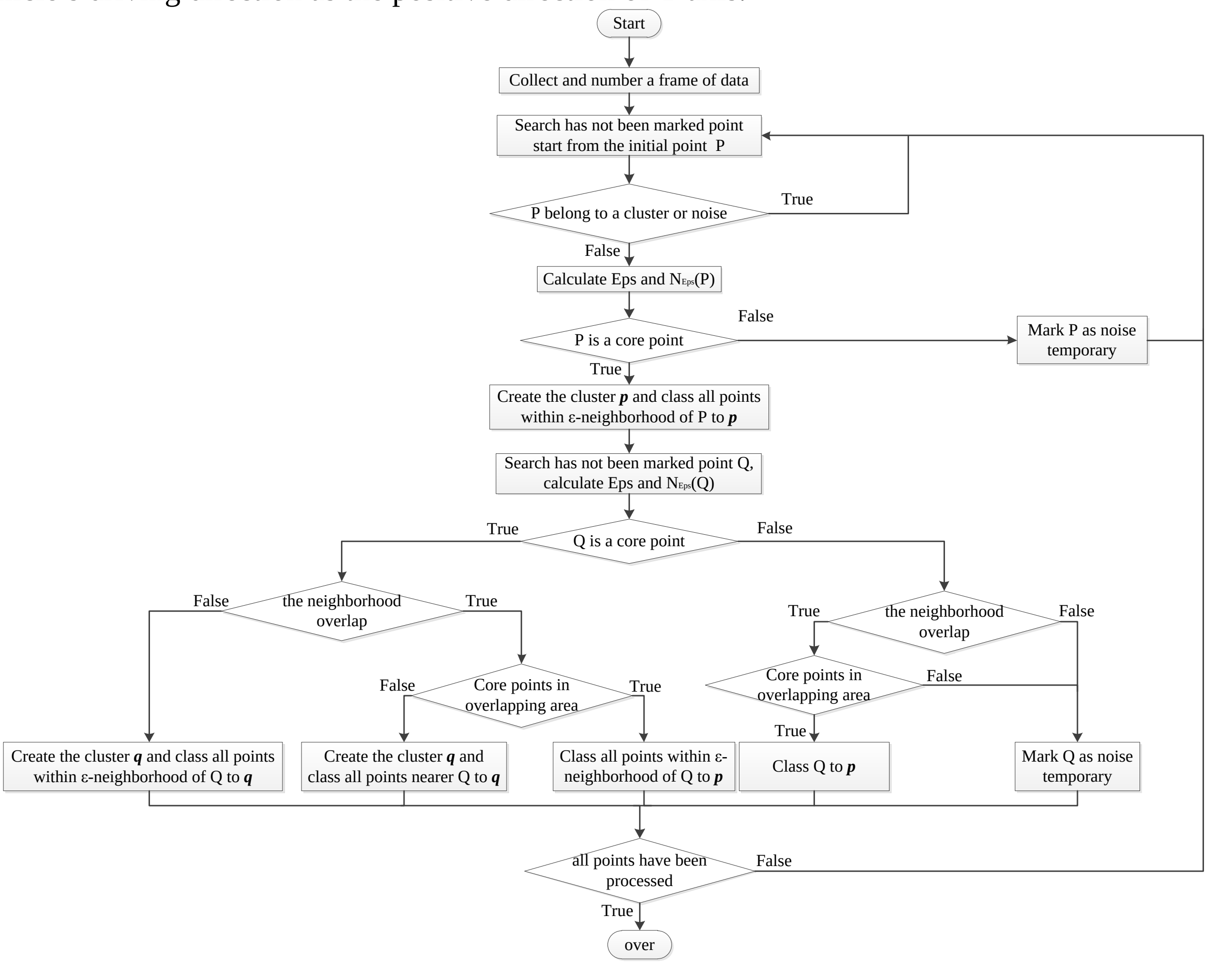
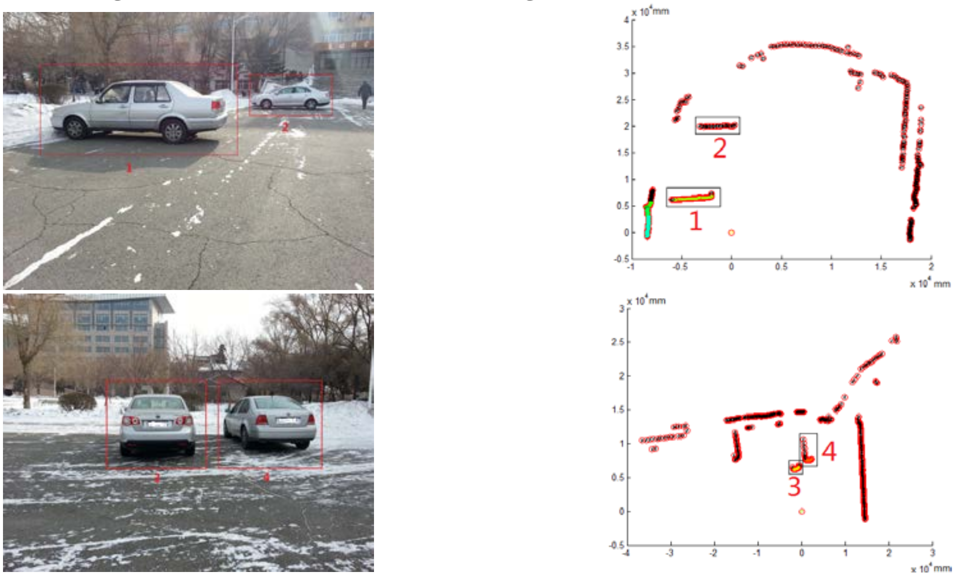


Figure 2. The improved DBSCAN flow chart

The experimental results under different environment are given in Figure 5. The origin of coordinates(0,0) in Figure 5(b) is the radar position. Vehicles which are shown in Figure 5(a) are marked by 1,2,3,4 and Figure 5(b) shows detected results. The other clusters in Figure 5(b) correspond to buildings, snow and other obstacles in the test scenario.

Figure 3. LMS511

Figure 4. Installation location of laser radar



(a) Test scenarios of obstacle detection

(b) Display of clustering results

Figure5. Real vehicle tests

Conclusion

This paper described the traditional DBSCAN and made an analysis of the deficiency of traditional algorithm in practical application, then put forward the improved DBSCAN clustering algorithm, used for obstacle detection based on two-dimensional laser radar, and completed the real vehicle experiment. Compared with traditional DBSCAN, this algorithm reduced the complexity of clustering and improved the accuracy. Results show that the improved algorithm can correctly identify obstacles.

Acknowledgement

In this paper, the research was sponsored by Chinese National High-tech R&D Program (863 Plan) under grant 2012AA110904, and National Science Foundation of China under grant 51175215.

References

[1] Damianos Gavalas, Grammati Pantziou, Charalampos Konstantopoulos, Basilis Mamalis. LIDAR:a protocol for stable and energy-efficient clustering of adhoc multihop networks[J].Telecommunication Systems.2007(1-3)

[2] Tung A K H, Hou J, Han J. Spatial Clustering in the Presence of Obstacles[C]//In Proc of 2001 Int Conf on Data Engineering. Heidelberg, Germany:[s.n.],2001:359-367.

[3] Ester M, Hans2peter K R Iegel,Sander J, et al. A density based algorithm for discovering

clusters in large spatial databases with noise[C]//Proceeding the 2nd international conference on Knowledge discovery and data mining(KDD),Portland,1996.266-231.

[4] Bing Liu, “A Fast Density Based Clustering Algorithm For Large Databases”, In: Proc. of IEEE Fifth International Conference on Machine Learning and Cybernetics, Dalian, August 2006

[5] Jie Zhao, Liu Yang.An Improved Density-based Clustering Algorithm[J]. Microelectronics & Computer,2009 26(11):189-192.

[6] Chaohua Guan, Danyong Chen, Huiyan Chen, Jianwei Gong.Improved DBSCAN Clustering Algorithm Based Vehicle Detection Using a Vehicle- Mounted Laser Scanner.Transactions of Beijing Institute of Technology,2010 6(30):732-736.