Research of Adaptive Grid Index Map Matching

Yong Qin [1], Qing Zhang [2,b] Yanfang Yang[2,a] ,Honghui Dong[3]and Xuewen Shi[2]

[1] State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University,

No.3 Shangyuancun, Haidian District, Beijing, 100044, China

[2]School of traffic and transportation, Beijing Jiaotong University,

No.3 Shangyuancun, Haidian District, Beijing, 100044, China

[3] Beijing Research Center of Urban Traffic Information Sensing and Service Technologies,

Beijing Jiaotong University   100044,China

[a] yanfang_4@163.com,[b]15125808@bjtu.edu.cn

Keywords: Map-matching; adaptive; grid index; calculus of differences method

Abstract. Map matching is a research hotspot and difficulty in transportation. The construction technology of topological space for GPS data of massive floating cars has been established for a long time in matching, however, its speed and effectiveness in database searching has always been a bottleneck. Based on the rasterization of Map-matching, this paper puts forward a method of improving the rasterizing Map-matching—completing the traversal of database resources through the adaptive size of occupancy grids indexing. First, to filter the collected GPS data. Second, summarize the characteristics of the method of grid index and then proposes a method of adaptive size of occupancy grid index, giving its process and rules of ID naming, neighborhood traversal, grade division and so on. Moreover, the correction of car is achieved through the calculus of differences method and the match is finished better. At last, judgement is made through the pros and cons of algorithm and the effect of matching.

Introduction

Nowadays, the tenure of vehicles has been increasing rapidly, causing the conflicts between the traffic flows and the road facilities , so it is more and more urgent to reduce the traffic jam and arrange the traffic induced demand reasonably. Facing interference of positioning error and map error, the map-matching algorithm aims to make the user's GPS location value correctly be matched to the actual way of the electronic map. The implementation of the Map-matching algorithm has a close relationship with the electronic map, and the electronic map must have the correct network topology and the precision high enough to complete the map matching, otherwise it will lead to false matches. In recent years, studies on algorithm of Map-matching have developed rapidly. Liu Weiming (2011) presented a fuzzy discriminant method of Map-matching based on the real-time traffic state of urban road [1]. Zhu Zhenxing (2011) proposed an large-scale incremental precise map-matching algorithm [2] (LSIPMM). Zhou Pu (2006) proposed an improved closest point estimation map matching algorithm and space vehicle average speed algorithm based on the analysis system of the GIS and floating car on the road traffic [3]. Yanying Li and Mike McDonald used the fuzzy model policy and put forward the model of estimating the road travel time though the matched floating cars [4]. David Anthony applied neural network as the estimation model to further improve the accuracy of estimating traffic parameters [5]. Huabei Liu of Chicago ,Wolfson and others proposed a weight-based map-matching algorithm of the trajectory of the vehicle abstracted into (x, y, t) and its matching sequence to the map sections. [6] .Overall, the current Map-matching in selecting the candidate roads generally takes the method of diffusing rectangular temporary road area [2]to store the candidate matching roads, or to determine the position error area by computer when the vehicle is in dynamic motion [7] The selection of area determines the number, matching speed and the accuracy of the candidate roads, and the relatively fixed regional will make the candidate roads miscellaneous in dense urban areas. This thesis proposes a

method of adaptive size of occupancy grid indexing to make the vehicles choose the best matching section in the grid and make the system memory allocation evenly through the dynamic grid to identify the matching area, increasing the matching speed and accuracy.

Method of Adaptive Grid Index

Data Screening

In this paper, the amount of GPS data covers 24 hours a day, sampled at 30-second intervals. More than 24 million pieces of data are received every day. That is, 1.67 thousand pieces of data are disposed by the system every minute. The original data exists anomalous data which is against the characteristics of traffic flow ,so it needs to be deleted through speed and positioning.

Road Interpolation

Before matching, the feature points need to be added to the roads. The road feature points is the point where the road used to describe the shape, location and other characteristics of road section. The feature points could be added through ArcMap, but the software added feature point to the small curvature of the section of road is significantly less than the road curvature of large sections. The algorithms related to matching used in this thesis require high distribution uniformity of the feature points. Thus, the feature points of roads need to be interpolated. Table 1 is the interpolation method.

Table 1. Interpolation Method and Parameters

F	After statistical calculation of the distance between adjacent two feature points, found in about 60 meters of data distribution of the most , This thesis sets the distance threshold between the feature points to F = 60 meters
P_1, P_2	P_1, P_2 is two adjacent feature points on the same road, interpolation is completed between P_1, P_2 .
$Distance（P_1, P_2）$	That is a function of the distance between $P_1 P_2$ computing,. Below are the mathematical formulas. $a = rad(P1_{Lat}) - rad(P2_{Lat})$ $b = rad(P1_{Lng}) - rad(P2_{Lng})$ $Distance(P1, P2) = 2\arcsin\sqrt{\sin^2\dfrac{a}{2} + \cos(P1_{Lat}) \times \cos(P2_{Lat}) \times \sin^2\dfrac{b}{2}} \times 6378$
N	the number of the inserted points between P_1 and P_2. Take the integer result only. N= Distance（P1,P2）/ F
$IncLat$ $IncLng$	$IncLat$ is the difference of latitude between P_1 and P_2. $IncLat = (P2_{Lat} - P1_{Lat}) / (N+1)$ $IncLng$ is the difference of longitude between P_1 and P_2. $IncLng = (P2_{Lng} - P1_{Lng}) / (N+1)$
Nx_{Lat} Nx_{Lng}	Nx_{Lat} indicates the insertion point in the sequence x-th point of latitude $P1_{Lat} + x * IncLat$ Nx_{Lng} indicates the insertion point in the sequence x-th point of longitude $P1_{Lng} + x * IncLng$

Method of Grid Index

The basic thought of grid matching[10] is to adjust the storage structure of the master map, reduce the times of calculating the distance between GPS coordinates and GIS coordinates as well as contrasting, and then improve the speed of the algorithm. Figure One shows some main methods of selecting roads from static areas in transportation nowadays. Peng Fei once proposed a cost function based map matching algorithm for GPS/DR integrated navigation systems. His main idea was to firstly fix the area of positioning error and then to extract the information of geographical position of the roads. Define the area of positioning error according to the probability. The area of positioning error must contain the actual position of the vehicles. This method is shown in Figure1 (1). In 2001, Jia Xianbo[8] proposed a grid index map algorithm of large-scale and complete coverage of urban road based on GPS equipped floating car technology. First, select the coordinates in the northeast and southwest and then create a rectangular region. Second, think how many grids are needed and then determine the proper size of the grid. Third, split the map. Mark every point on the GIS map and finds out its correspondent grid. The coordinates of the grid are its ID card. This method is illustrated in Figure 1(2). Zhu Zhenxing utilized the GPS to gain the information of the latitudes and longitudes and established a temporary area. In this area, the difference of longitude between the point on the map and its actual position is 0.01 degree. The difference of latitude is also 0.01 degree. Through contrasting the coordinates of the GPS point and the coordinates of the node, one can judge whether the temporary area contains the candidate road. As shown in figure 1.
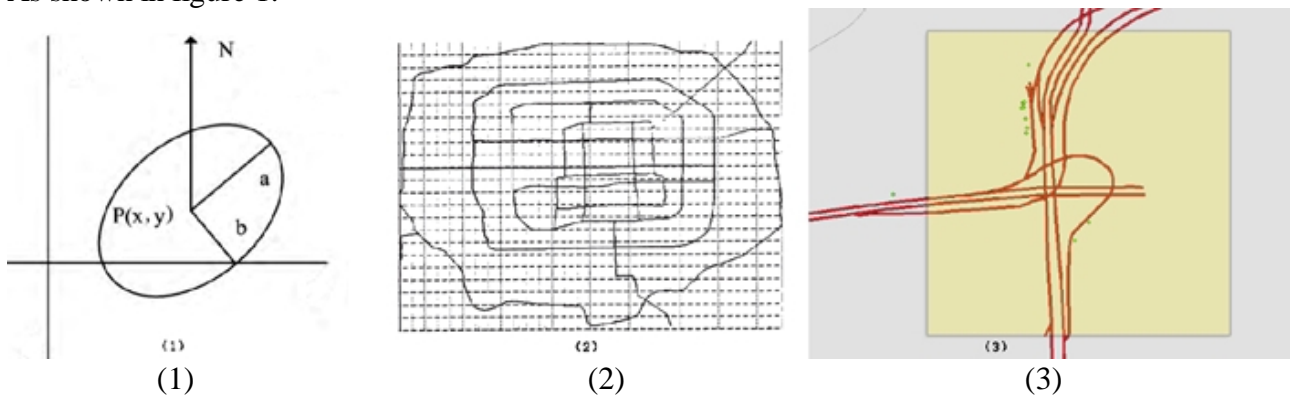

| (1) | (2) | (3) |

Figure1 Select Candidate Sections in Static Areas

Method of Adaptive Grid Index Based on Quadtree

The static grids should be divided artificially. The urban roads distribute unevenly. In the center of the city, the road net density is high. Thus there are many roads which are fit for selecting. In contrast, the road net density is not high in urban areas. So only some grids contain the candidate roads and the system memory does not distribute evenly.

Therefore, this thesis proposes a method of adaptive size of occupancy grid indexing based on quadtree[9]. This method, a modified raster map method, introduces a notion of dynamic grids based on raster map and makes the candidate roads in the grids distribute evenly.The grids are different in size, but the number of points to be matched is almost the same in each grid, which greatly accelerates matching speed. The steps of building adaptive quadtree grid are as follow, with an example of map of Beijing:

a)   Building the coordinate system by setting the northwest vertex at the map of Beijing as origin.

b)   Dividing the map into different layer grids. The number of matching points in the first layer grid is limited to 10, and the grip will be divided to second layer when the points outnumber 10. The first layer needs to be divided into about 300 square grids in the horizontal direction, with the grid's length calculated by the formula: D=(MaxLng-MinLng)/300. The second layer is divided by the method of quadtree, each grid equally divided into four parts, like Chinese character "田"。

c)   Building the coordinate system of grids in the form of M_N_Q. M represents abscissa and N represents ordinate, increasing along X-axis and Y-axis respectively. Q represents layer, and 0 is the

first layer. 1~4 are numbers in the second layer grids. Second layer numbered from the "田" upper-left corner counterclockwise starting at 1, as shown in figure 1.

    d)   When the vehicle searches the matching points the grids, it still needs to search the points in neighborhood areas, and then store the result in the database, as shown in figure 2.
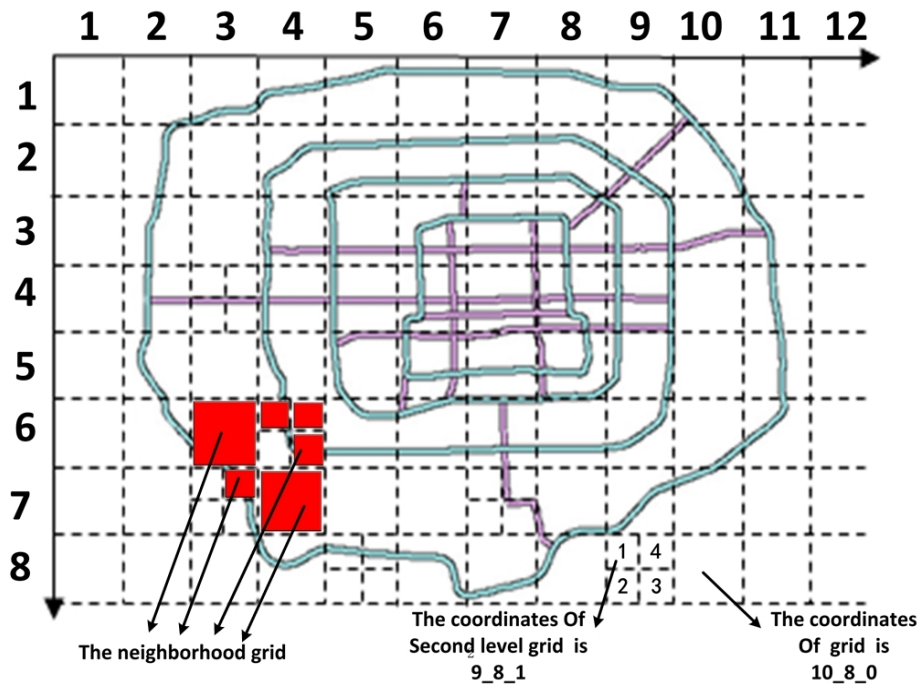

Figure2 Select Candidate Sections in Static Areas

Calculus of differences method

calculation principle

    As the vehicle location data error expectancy has certain gradient features and the real-time character of vehicle location data, so the vehicle's position data can be corrected by analyzing the deviation between real-time position data and actual position. Relevant variables and their correlations in the difference method are shown in figure 3:
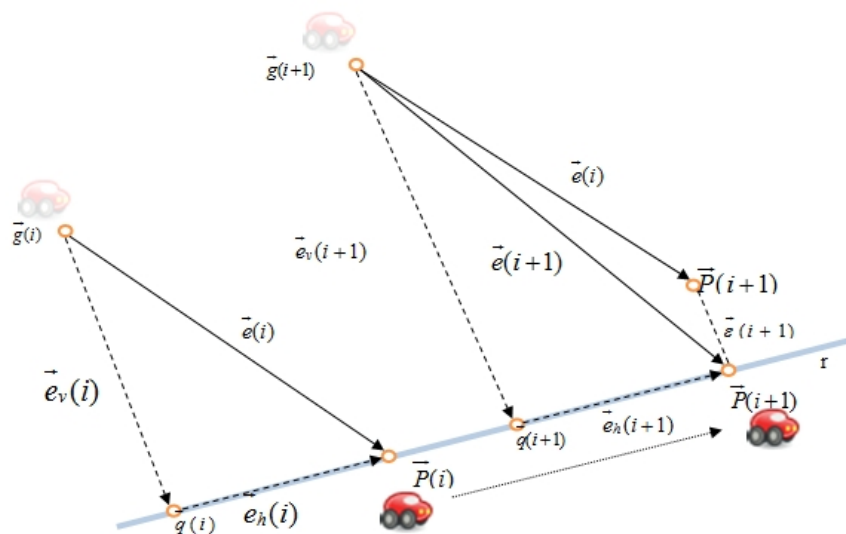

Figure 3 Difference Principle

Meanings and definitions of symbols in the above picture are as follow:

(1) $\overset{\centerdot}{g}$ is the actual GPS position coordinate of vehicle where the vehicle's GPS coordinate locates in GIS map.

(2) $\overset{\centerdot}{p}$ is the actual position of the vehicle, and also is the vehicle's position calculated by matching algorithm.

(3) $\overset{\centerdot}{e}$ is matching correction the between GPS position coordinate and vehicle's actual position. $\overset{\centerdot\centerdot}{e_h}$ is the horizontal component of $\overset{\centerdot}{e}$ that parallels the road, and $\overset{\centerdot\centerdot}{e_v}$ is the vertical component of $\overset{\centerdot}{e}$ that is perpendicular to the road.

(4)r is the road.

Steps to Apply Difference Method

At the time of 0, $\overset{\centerdot}{e}(0)=\overset{\centerdot\centerdot}{e_v}(0)$. To get more accurate matching correction, we can take the average of first three data to calculate $\overset{\centerdot}{e}(i)$. The formula is equation 1 as follow:

$$\overset{\mathbf{r}}{e}(i)=\frac{\sum_{j=0}^{i=2}[\overset{\centerdot\centerdot}{p}(i-j)-\overset{\centerdot\centerdot}{g}(i-j)]}{3}$$

(1)

Assuming current time is i+1, and the Specific steps are shown in figure 4:

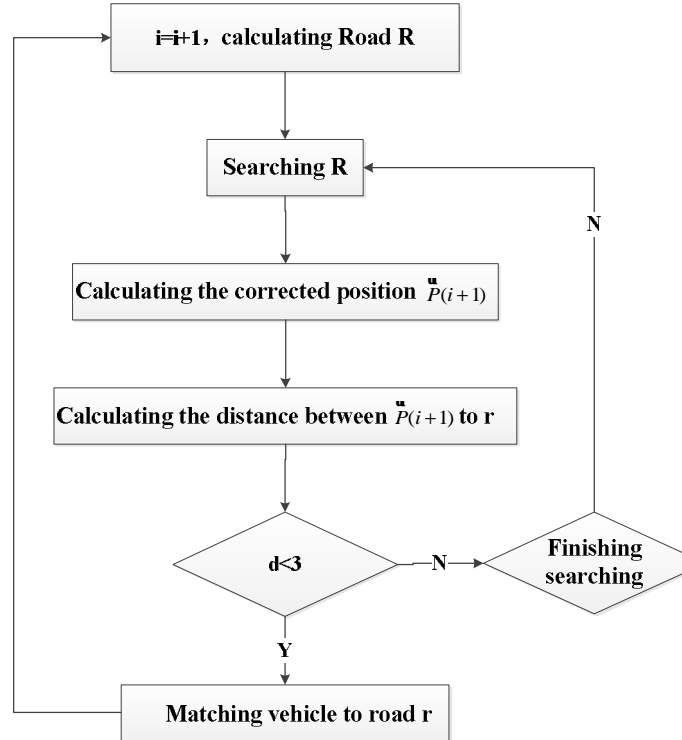

Figure 4 Flow Chart of Difference Method

The Algorithm Implementation

The Advantages and Disadvantages of the Algorithm

The verification and matching test of this Method is based on the road system of Beijing city. The GPS data sampling period is 30s, and the accuracy is within 20 to 50 meters. The result presentation and management of this method is realized mainly through ArcMap. Since this algorithm is initially devised to deal with mass data quickly, its accuracy is less emphasized. The adaptive grid index method can divide the map into grids, and thus has reduced the workload and quicken the speed of matching. The adaptive structure of the secondary level enables the grids to be more dynamic, which can

disassemble intelligently according to the density of the road network, thus enhancing the speed of data selection. The disadvantage of this method is that it is not accurate enough. Now that it uses the dot-to-dot model, it inevitably leaves out some complex cross roads, and can't differentiate the main roads from the side ones.

The matching effect

The Figure 5 is the matching effect of our study of Dongzhimen South Street. The red spots stand for the initial vehicle GPS data points, while the green ones, the matched points completed by using the adaptive grid index method and the calculus of differences method. We can see from the picture that the initial data points deviate from the road, while the ones matched using the algorithm are much more accurate.
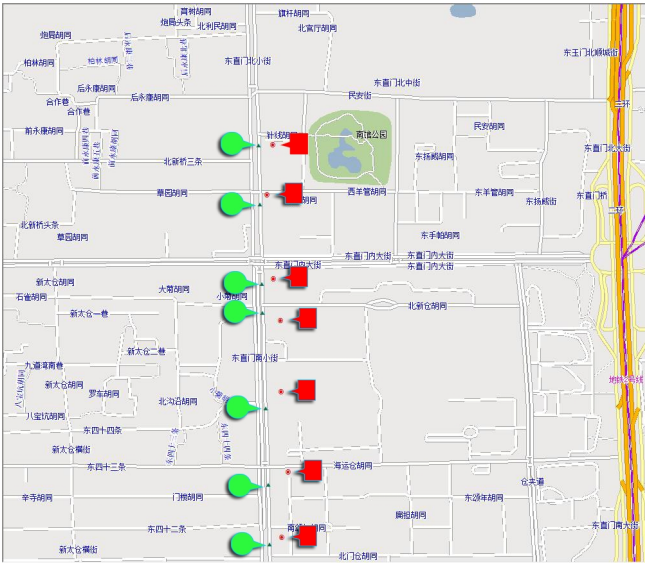

Figure 5 Matching Effect

Through observation can be found, the algorithm is not compute match point coordinates by direct projection. The point is the original GPS point on the road in the closest point on the road. Refer to table 2 for the differences of the two matching algorithms.

Table 2 .Differences of the Two Matching Algorithms

Algorithms	matching number		correct rate/%	average computing time/ms
	correct	wrong		
grid index	1433	101	93.4	60
adaptive grid index	1458	76	95.0	26

Therefore, through the adaptive grid index method and calculus of differences method the results meet the technical requirements.

Conclusion

This paper presents a method of adaptive grid index, how to build two layers that can automatically divide according to density of road. It assesses the above from two aspects, advantages and disadvantages of algorithm and matching effects . There are only two layers in the algorithm, so the more complex cases need to be optimized.

Acknowledgements

References

[1] Liu Weiming & Tian Shiyan. *Fuzzy Synthetic Judgement of Real-Time Traffic Conditions Urban Road Based on Map Matching,* Journal of Highway and TransportationResearch and Development *,2011(10)*

[2] Zhu Zhenxing. *Urban Traffic Data Integration and Traffic Processing Method for Transportation* [d] Masters 2011.

[3] Zhou Pu *Systematic Research of Traffic Analysis Based on GIS and Floating Vehicles*. [d] master 2006

[4] Yanying Li,Mike McDonald.*Link Travel Time Estimation Using Single GPS Equipped Probe Vehicle*[C]. Proc of the Intelligent Transportation Systems, IEEE the 5th International Conference, 2002. 932-937

[5] Faria David Anthony.*A Framework to Transform Real-time GPS Data Derived from Transit Vehicles to Determine Speed-flow Characteristics of Arterials*[D].Dissertation Abstracts International, 2002

[6] Huabei Yin,Wolfson.*A weight-based map matching method in moving objects. Scientific and Statistical Database Management*,2004.Proeeedings.16th International Conference on,2004:437-438

[7] Peng Fei& Liu Zhongji, *a Fuzzy Logic Based Map Matching Algorithm for GPS/DR Integrated Navigation Systems*[J], Telemetry and Telecontrol, 2001,22(1), P32-36.

[8] Jia Xianbo, *The Research of Large-Scale and Complete Coverage Urban Road Traffic Status Collection Based on GPS Equipped Floating Car Technology*[D],2010.

[9] Samet H, Rosenfeld A, Shaffer C A, et al. *A geographic information system using quadtrees*[J]. PatternRecognition, 1984, 17(6): 647-656.

[10] Lenser S,Veloso M.*Sensor resetting localization forpoorly modeled mobile robots*[C]. Proc of the IEEE IntConf on Robotics and Automation. San Francisco: IEEEPress, 2000: 1225-1232.