2022 IEEE 25th International Conference on Computer Supported Cooperative Work in Design (CSCWD) | 978-1-6654-0527-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/CSCWD54268.2022.9776281

A Driving Safety Assessment Method Based on Vehicle Trajectory Geometry

Puyao Guo[1], Ying Zeng[1], Yan Wang[1,2]*, Yuying Li[1]
[1]College of Computer Science, Inner Mongolia University,Hohhot, 010021, China
[2]Inner Mongolia Engineering Lab of Cloud Computing and Service Software, Hohhot, 010021, China

Abstract—**Drivers are autonomous and random in the process of vehicle operation, and their micro-driving behavior characteristics are related to traffic safety. Vehicle trajectory is the direct data that intuitively reflects the driver's micro-driving behavior, and drivers with different safety awareness and driving habits will generate different trajectory geometry. In this paper, through analyzing a large number of vehicle trajectory data to find drivers' irregular or dangerous driving behavior, a driving safety assessment method based on vehicle trajectory geometry is proposed. Firstly, the invalid point of vehicle trajectory is eliminated using the K-Means algorithm. And then the characteristics of vehicle are analyzed based on the curvature change of trajectory geometry, the Hoff transformation and other methods to establish the trajectory structure with linear elements. Finally， a similarity method of vehicle trajectory is proposed to quantify a trajectories by linear elements, and evaluate the safety index of the trajectory, which can provide safety reference for regulating drivers' driving habits.**

Keywords—trajectory geometry, micro-driving behavior, safety assessment, trajectory structure, similarity

I. Introduction

With the development of economy and the increase of the number of motor vehicles, the number of drivers has increased rapidly. The autonomy and randomness of drivers have become the most important factors affecting road safety among the four elements that make up the road traffic system (people, vehicles, roads and environment). According to the survey, a driver receives more than 300 messages per kilometer. In order to ensure safe driving, drivers has to make judgments and decisions on about 25% of these messages. Due to the differences in the gender and age of drivers, their driving behaviors are affected by their psychological factors, driving experience and driving proficiency. The driver's safety awareness and driving characteristics directly determine the degree of road traffic safety. How to regulate the driving behaviors of drivers to ensure the safety of the intricate road system to a greater extent becomes a hot issue of concern.

In recent years, a large number of studies have focused on the mining and analysis of driver behavior characteristics to identify dangerous driving, so as to discover and predict the dangerous states that may cause traffic accidents. The main research can be divided into two categories. One is to obtain real-time driving videos of drivers through machine vision, and then to discover safety hazards that may cause traffic accidents such as fatigue driving and irregular driving by analyzing the driver's behavior and state. The other is to identify abnormal trajectories by analyzing the spatio-temporal characteristics of vehicle trajectories. In the process of detecting abnormal trajectories, clustering and deep learning methods are mostly used[1]. The methods discover outliers and anomalies of the trajectories based on the supervised or unsupervised machine learning process of large amounts of trajectory data, so as to detect data with differences from most trajectories. The common outlier detection algorithms can be broadly classified into statistical-based algorithms, depth-based algorithms, distance-based algorithms, density-based algorithms and algorithms for high-dimensional data[2]. Fewer studies have focused on the geometry features of a trajectory. However, trajectories, as a kind of time-series data consisting of many location points, are inherently characterized by a large amount of points and significant correlation relationships between points. A trajectory formed by the driver reflects his driving habits to a certain extent, and the relationship between a trajector and its corresponding road can also reflects whether the vehicle has the implication of dangerous driving. Moreover, vehicle trajectory is represented in macroscopic, but still reflect sparsity in micro-local[3], and traditional machine learning methods directly affect the correctness of detection results when historical data are scarce. Therefore, this paper studies a driving safety assessment method based on vehicle trajectory geometry. Compared with machine learning methods based on massive trajectory data analysis, our method has a great efficiency advantage in judging the security attributes of trajectories and solves the problem of sparse data, which provides a valuable reference for portraying a driver's driving habits and regulating his/her driving behaviors.

II. Related Works

Traditional mathematical models and machine learning models are two important approaches to perform traffic trajectory anomaly detection.

The traditional mathematical models mostly use statistical probability methods. The basic idea of these methods is to make comprehensive judgments on different motion trajectories by a multivariate detection fusion strategy, including Hidden Markov Statistical Models (HMM), Bayesian Decision Probability Models, linear regression and function approximation fitting models [4-6]. Zhang et al. measured the similarity between trajectories and then used a Gaussian model to match the new trajectory data to detect whether or not the trajectory was anomalous [7]. However, the correlation of data within the trajectory is not taken into account, so the detection efficiency is not high. Li et al. processed historical trajectories by feature classification and used an improved Bayesian inference method to measure the correlation between trajectories as a way to improve prediction accuracy [8]. Due to the variability of inter-trajectory features, Venkatesan et al.

* Corresponding author

considered the global features of trajectories, including direction, velocity and distance, and then found anomalous trajectories in the dataset by measuring the distance difference, velocity variation and direct correlation between trajectory points[9]. For the limitation that the above methods cannot make comprehensive judgments on local trajectories and global trajectories under different situations and cannot obtain better processing results, a trajectory detection model combining time series approximate fitting and the linear regression method was established, which integrated the correlation of historical trajectory data, external interference and real trajectory data, making the model fitting results closer to the actual data[10-12]. Liu et al. proposed a two- part TRAOD algorithm for trajectory anomaly detection, which first partitioned a trajectory into multiple trajectory segments using the shortest description length principle, and then combined the Haffman distance (Hausdorff) and density between trajectory segments to detect anomalous trajectory segments, making up for the inability of previous studies to detect local anomalous sub-trajectories [13].

Machine learning models have the advantages of fast processing speed and strong learning ability in processing nonlinear mapping data such as traffic trajectories[14-15]. Kong et al. used a particle swarm algorithm to optimize the parameters of support vector machines by considering factors such as latitude and longitude of the trajectory points and vehicle travel routes. Based on this model, they performed the prediction of vehicle speed and congestion, and gave regional congestion level evaluation index[16]. Yang et al. extracted the feature information of trajectory points under different labels by training the trajectory data with existing labels, and then detected abnormal trajectories in real-time trajectory data[17]. However, the method is dependent on labels and has poor applicability, Mikluscak et al. used elastic back propagation method to train artificial neural networks, which improved the model autonomous learning efficiency and avoided the dependence on labels, but the convergence speed of the model training process is slow[18]. Lu et al. used a self-organizing feature mapping network (SOM) to learn historical trajectory sequences, identify the characteristic attributes of normal frequent trajectories, and use superior neuron representation to determine anomalies by comparing the response functions of normal trajectories with those of the trajectory points to be detected[19]. However, these methods simply consider the features of the trajectory points without considering the trajectory sequence continuity. De et al. decomposed the GPS data into multiple trajectory sequences with indeterminate length and use the B-RNN neural network to identify the overall features of historical trajectories, and their detection results are satisfactory [20].

From the above work, the current domestic and international research on trajectory prediction and anomaly detection has less consideration in the efficiency problem of mathematical modeling methods, while the accuracy of machine learning models with fast processing speed is greatly affected by the size of the data set. Therefore, this paper proposes a mathematical model with the reduced computational effort to overcome the algorithmic efficiency problem when identifying driving safety for trajectory data. Firstly, the trajectory line shape is detected by Hough transform. Secondly, the obtained straight line set as well as the circular curve set are fitted using the least-squares method. And then the trajectory feature vector is generated, the data structure of the whole trajectory is established. Finally, the traffic safety risk corresponding to the trajectory is assessed based on the similarity analysis of the trajectory to identify the abnormal features of the trajectory.

III. TRAJECTORY GEOMETRY DETECTION

A. Trajectory pre-processing

In order to realize the matching of the collected trajectories with the road network, we first need to perform some basic processing on the trajectory data, including trajectory coordinate conversion and elimination of invalid trajectory points. The trajectory data is collected based on an international geocentric coordinate system (WGS-84). In order to convert the latitude and longitude coordinates of the trajectory points into Baidu coordinates, we first convert the WGS-84 coordinate to the GCJ-02 Mars coordinate, and then convert the GCJ-02 Mars coordinate to the Baidu coordinate. Since the coordinate points after cartesian coordinate transformation are dense, it is not convenient for post-processing. For this, we process the local results after map matching to enhance the visibility of the processing results by extracting key coordinate points, and eliminate the invalid trajectory points from the extracted points. The process is as follows: 1) The K-Means clustering is performed on the collected trajectory points to eliminate the invalid points; 2) The Bresenham algorithm is used to generate a fixed-radius circular region with the clustered trajectory points as the midpoint; 3) The seed filling is applied to each region to obtain a more ideal trajectory map. The processed results are shown in Fig.1.

Fig.1 The trajectory preprocessing effect

B. Road geometry analysis

The main elements of road geometry are straight lines, gentle curves and circular curves. The linear combination of the gentle curve and the straight line is more gentle, which is in line with the urban planning design principles, as shown in Fig.2. The formal definition of each element is given below.

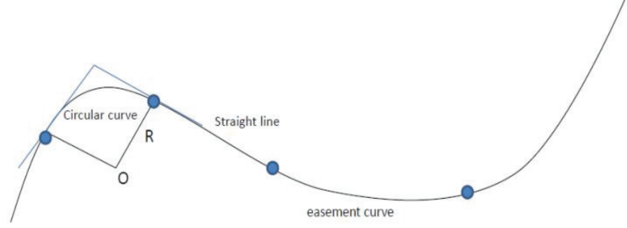


Fig.2 Basic elements of the road geometry

Definition 1 Curvature The degree of curvature of a line, which can be calculated by the ratio of the arc tangent angle to the arc length in the limited condition where the arc length tends to a point. Formally,

$$K = \lim_{\Delta s \to 0} \left| \frac{\Delta \alpha}{\Delta s} \right| \tag{1}$$

where Δa is the angle between the two tangents of the curve from midpoint of the curve MM' to M'and Δs is the difference in length between the start and endpoints of the curve, as shown in Fig.3.

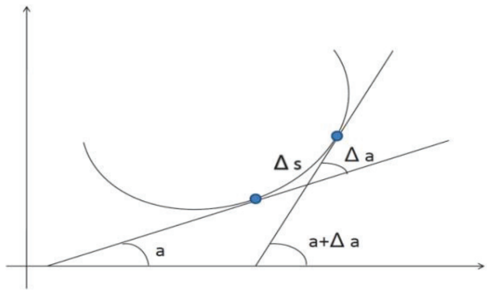


Fig.3 Definition of curvature

Definition 2 Straight line A line in a road with K=0.

Because urban roads are inevitably affected by buildings and other things that make straight lines don't fit well into urban road construction. And drivers who keep driving on a straight-line road may be distracted and even more likely to cause traffic accidents. Therefore, the combination of various road line shapes has become necessary for road construction. The definition of curved shape is given below.

Definition 3 Circular Curve A circular curve is a common line shape with K= 1/R, denoted as is a triple (L, T, O). L is the length of the curve, T is the tangent length of the curve and O is the outside distance of the curve.

Definition 4 Gently Curve A gentle curve is between a straight line and a circular curve, achieving a perfect transition between a straight line and a gentle curve, denoted as a triple (L, R, C). l is the length of the curve, R is the radius of the curve, and C is the parameter of the gyration curve.

This type of road shape avoids the sudden appearance or disappearance of centripetal forces at a certain road point and improves the stability and comfort of driver.

C. Trajectory geometry element detection

Since a trajectory is generally consistent with the road shape elements, we can divide the trajectory geometric shape into straight lines, gentle curves and circular curves, and evaluate the risk level of the trajectory by detecting the elements of the trajectory. However, since each trajectory contains at least ten thousand track points, it is difficult to find the curvature directly for the discrete points. In additon, the discrete points are scattered, so the measurement error is inevitable. The discrete point measurement errors can lead to inaccurately calculated curvature. We found that theintegral change of curvature along the line shape is consistent with the rate of azimuthal change of the line shape, so the calculation of trajectory curvature can be simplified by calculating the azimuthal change. At this point, the curvature of the straight line, circular curve and gentle curve is 0, constant and variable, respectively, that is, theri azimuthal change values are 0, $\Theta = L / R$ and $\Theta = \left(L^2\right) / \left(2C^2\right)$. The formula for calculating the azimuth angle between adjacent discrete points M (X_M,Y_M) and N (X_N,Y_N) is as follows:

$$\tan \Theta = \left| Y_M - Y_N \right| / \left| X_M - X_N \right| \tag{2}$$

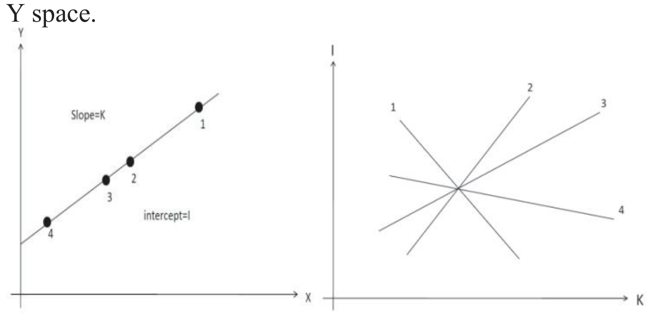
The formula for calculating the azimuth between adjacent discrete points M（X_M,Y_M）and N（X_N,Y_N） is as follows.

$$K = \left| \tan \Theta_M - \tan \Theta_N \right| / \tan \Theta_M \tag{3}$$

In this paper, the Hough transform is used to detect the trajectory line shape. The Hough transform maps a point in the right-angle coordinate system to a sine curve or cone in the polar coordinate system, and analyzes the points with the same properties in the image space. It can judge the trajectory line shape by the change in azimuth of the trajectory, which is close to 0 for a straight line, and close to L/R for a circular curve line, and defaults to a gentle curve line.

(1) Detecting straight line of trajectories

As shown in Fig.4, the trajectory points in the X-Y coordinate space in Fig.4 (a) are mapped to the K-I curvature space in Fig.4 (b), such that the point on a straight line in the X-Y space corresponds to the same point in the K-I space, and the point with the most occurrences in the K-I space is statistically derived as the most probable straight line in the X-Y space.



(a) Trajectories in X-Y space (b) Trajectories in X-Y space

Fig.4 Shape detection for trajectories

However, there are limitations in mapping from X-Y space to K-I space, where there are some line shapes with K=0 in X-Y space, such that a straight line under X-Y space cannot be represented in K-I space. Based on such considerations, we map the points in X-Y space to the points under polar coordinate space, thus avoiding the above cases. The Hough transformation uses this idea to transform the coordinates in X-Y space to those in polar coordinate space, i.e., a point in the cartesian coordinate system is mapped to a sine curve in the polar coordinate system. Different points on the line are mapped a cluster of sine curves in the parameter space, where the intersecting point is the most probable line in X-Y space.

(2) Detecting curve line of trajectories

Firstly, all coordinate points under the same circumference in the right-angle coordinate system are processed by using the Hough transform, and the coordinate points in the right-angle coordinate system are mapped to their cumulative counterparts in polar coordinate space. Such that, the right-angle coordinates (x,y) are converted to polar coordinates (a,b,r), where (a,b) is the center coordinate and r is the radius. If the ordinary Hough transform is used, it will result in a very large number of operations and a very undesirable spatio-temporal complexity. In order to minimize the space-time complexity, we should first determine the coordinates of the circle center point more accurately, after which the radius will be determined in a certain range, which can greatly improve the operational performance of the system. In this paper, we use a three-point method to determine the center of the circle, i.e., the intersection point of the mid-pipeline of two points is the center of the circle curve, and then the possible range of the radius T is obtained by taking the distance from the center to the point on the circle. Finally, the radius of the circle curve is further accurately determined by the Hough transform in the range of [r-T,r+T]. The detection effect is shown in Fig. 5.

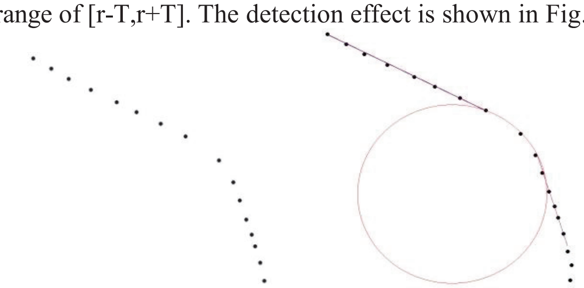


Fig.5 Hough transform detection of trajectory curve

D. Trajectory fitting

The line elements of each trajectory are determined after the initial processing of several discrete-continuous data points by the Hough transform, but in order to obtain a better fit, we use the least-squares method to fit the obtained set of straight lines as well as the set of circular curves. The method calculates the sum of the error distances from the obtained trajectory points to the line shapes generated by the Hough transform and obtains the best line shape matching function by minimizing the error. The basic process is described as follows.

Suppose the trajectory data set is S=(x_i,y_i),(i=1,2,.... ,n,n is the number of data points in the dataset), and the data points in S are distributed near the polylines generated by the Hough transform.

For direct trajectory fitting, the Hough transform equation is rewritten into the slope-intercept Equation (5) using Equation (4), and the sum of all the error is calculated by Equation (6). The slope-intercept with the smallest error is the best parametric straight line.

$$p_k = x_i \cos\Theta k + y_i \sin\Theta k \quad (i=1,2,\ldots n; k=1,2,\ldots R) \tag{4}$$

$$y_i = a_k x_i + b_k \tag{5}$$

$$d_{ki} = \left|a_k x_i + b_k - y_i\right| / \sqrt{(1+a_k * a_k)}$$
$$(x_i, y_i) \in S, (i=1,2,\ldots,n; k=1,2,\ldots R) \tag{6}$$

where k is the kth line shape generated by the Hough transform, $a_k = -\cos\Theta k / \sin\Theta k, b_k = -p_k / \sin\Theta k$.

For curve estimation fitting, the curve fitting error is calculated by Equation (7). Set Q(*a,b,c*) is the sum of squares of these errors by Equation (8). Through (9)-(11), we can obtain the best values of fitting parameters *a,b* and *c*.

$$D^2 - R^2 = \left(X_i - A\right)^2 + \left(Y_i - B\right)^2 - R^2 = X_i + Y_i + aX_i + bY_i + c \tag{7}$$

$$Q(a,b,c) = \Sigma\left(X_i^2 + Y_i^2 + aX_i + bY_i + c\right)^2 \tag{8}$$

$$\frac{\partial Q(a,b,c)}{\partial a} = \Sigma 2\left(X_i^2 + Y_i^2, +aX_i + bY_i + c\right)X_i = 0 \tag{9}$$

$$\frac{\partial Q(a,b,c)}{\partial b} = \Sigma 2\left(X_i^2 + Y_i^2, +aX_i + bY_i + c\right)X_i = 0 \tag{10}$$

$$\frac{\partial Q(a,b,c)}{\partial c} = \Sigma 2\left(X_i^2 + Y_i^2, +aX_i + bY_i + c\right)X_i = 0 \tag{11}$$

IV. TRAJECTORY RISK ASSESSMENT

In this paper, we adopt the similarity analysis method of trajectory geometry to evaluate its traffic safety risk. Based on the fitting results of the trajectory, a trajectory is represented as a finite number of feature points, based on which the trajectory feature vector is generated to establish the data structure of the whole trajectory. And then, the similarity analysis of the trajectory based on this structure is used to identify the abnormal features of the trajectory and evaluates the safety risk level of the driver.

A. Description of trajectory geometry

(1) Trajectory feature points

Each segment of the trajectory can be represented by a finite number of key points. A straight line can be uniquely determined by the two coordinate points of the endpoints, and a circular curve can be uniquely determined by the first and last endpoints of the arc segment as well as the center of the circle, and these key points play an important role in trajectory similarity processing. In order to represent a trajectory effectively, we process each segment of the trajectory and represent its line shape using representative key feature points. For linear trajectories, we take the midpoint of the linear trajectory as the feature point of the linear trajectory. For the circular curve trajectory, we make an inner tangent circle to the endpoint and the center of the circular curve trajectory, and take the center of the inner tangent circle as the feature point of the circular curve trajectory. After the above processing, we can get three trajectory feature points. After establishing the feature points, finding the most representative feature point as the datum of the trajectory becomes the key to the trajectory representation. The selection of the feature points is important because it has a great influence on the establishment of the trajectory feature vector and the final similarity between the trajectories. For this, we find the inner tangent circle center of the trajectory feature points, and the inner tangent circle center obtained is used as the reference feature point of the whole vehicle trajectory.

(2) Trajectory feature vectors

On the basis of the feature points of a trajectory, the feature vector is defined as a vector of feature points and datum points. In order to facilitate the calculation and trajectory processing, the feature vector is represented by the datum point and the nearest point to it. As shown in Fig.6, the feature point A is the datum point of the trajectory, and feature point B is the closest to feature point A. Therefore, vector AB is the feature vector of the trajectory.

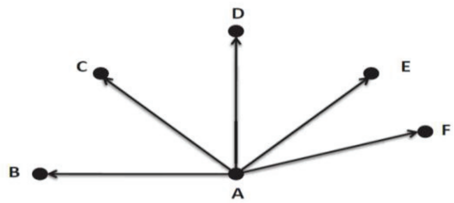


Fig.6 The trajectory feature vectors

(3) Trajectory data structure

The trajectory data is processed in a normalized manner, and the data structure of the trajectory feature points is established including the cosine of the angle between the trajectory feature vector and the base feature vecto and the length of the feature vector. The data structure makes the processing of the trajectory independent of the direction and position of the trajectory, so that the algorithm highly reliable.

B. Trajectory anomaly detection

In this paper, we use the similarity analysis method based on trajectory feature vectors to identify anomaly trajectories. Trajectory similarity mainly refers to the similarity of trajectories in terms of topology, geometry and expression function. Traditional similarity analysis uses different distance metrics according to the data characteristics, which are mainly divided into point-based similarity analysis and shape-based distance similarity analysis. Point-based similarity analysis mainly includes Euclidean distance similarity analysis, dynamic time regularized similarity analysis, longest common subsequence similarity analysis, edit distance similarity analysis, etc. Shape-based distance similarity analysis mainly includes Hausdorff distance similarity analysis, Fréchet distance similarity analysis, one-way distance similarity analysis,etc. Since this paper collects GPS trajectory data, they are the discrete points on the corresponding matching map. So the similarity of trajectory cannot be judged by its topology. The cosine similarity is mostly applied in the similarity comparison of abstract feature vectors. In Section 3, we have obtained multiple sets of dimensionless data sets by building the data structure for trajectories, which are not related to trajectory locations. In this section, the cosine similarity measure is used to measure the morphological similarity and dissimilarity of different trajectories. The formula for measuring the similarity between trajectory X and trajectory Y is shown as Equation (10).

$$sim(X,Y) = \cos\theta = \frac{x \cdot y}{\| x \| \cdot \| y \|} = \frac{\sum xy}{\sqrt{\left(\sum x^2\right)}\sqrt{\left(\sum y^2\right)}} \tag{10}$$

Sim(X, Y) is a unitless measure, and the smaller the value of the similarity measure indicates the larger the difference between vectors, and the larger the value of the similarity measure indicates the smaller the difference. It is proved that the use of cosine similarity can achieve a more ideal comparison result.

In the study of the classification of driving hazards, it is crucial to find a reasonable and appropriate number of classes. Too few classes cannot accurately reflect the danger of a vehicle driving state at a certain moment, and too many classes can produce redundant information, and may even lead to the phenomenon of small differences between adjacent classes. Therefore, this paper provides a detailed and accurate classification of vehicle hazard levels.

The process of this method for hazard level rating is as follows: all accident trajectories are detected well in advance and their data structures are saved, and then the vehicle trajectories to be evaluated are compared with all accident trajectories. For example, we have prepared 100 accident trajectories . A tested trajectory can be compared with all the accident trajectories. The results show that the similarity between this trajectory and 75 accident trajectories is distributed in the range of [0,0.249], accounting for 75% of the total number of accident trajectories; the similarity between this trajectory and 21 accident trajectories is distributed in the range of [0.25,0.499], accounting for 21% of the total number of accident trajectories; the similarity between this trajectory and 3 accident trajectories is distributed in the range of [0.50, 0.749], accounting for 3%, and the similarity this trajectory and 3 accident trajectories is distributed in the range of [0.75,1], accounting for 1% of the total number of accident trajectories.According to the similarity distribution ratio, the co rresponding hazard level is calculated based on the similarity level regulations, which are shown in the following table 1. We divide three hazard warning levels: 1-3 we define as slight hazard, 4-7 we define as general hazard, and 8-10 we define as extreme hazard.

TABLE 1 THE HAZARD CLASS REGULATIONS

Hazard class	**[0.75, 1]**	**[0.50, 0.749]**	**[0.25, 0.499]**	**[0, 0.249]**
10	≥1%	-	-	-
10	<1%	≥50%	-	-
9	<1%	≥40% & <50%	-	-
8	<1%	≥30% & <40%	-	-
7	<1%	≥20% & <30%	-	-
6	<1%	≥10% & <20%	≥50%	-
5	<1%	≥10% & <20%	<50%	-
4	<1%	≥5% & <10%	≥40%	-
3	<1%	<5%	<40%	-
2	<1%	<1%	≥10%	-
1	<1%	<1%	<10%	-

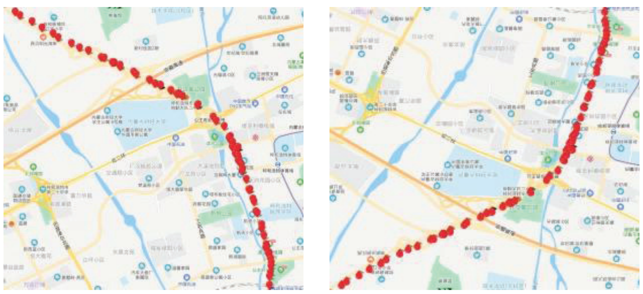
V. CASES AND ANALYSIS

In this section, we will validate the detection method for trajectory geometry and the assessment method for Trajectory geometry element detection safety proposed in this paper based on real GPS trajectory data. In the specific test, we selected some representative trajectories in the city as incident trajectories and compared them with the trajectory to be tested.

Through the experiment, we find the tested trajectory has a high similarity with one of the trajectory, ranging from [0.75, 1].The two trajectories are shown in Fig.7. Fig.7 (a) shows the tested trajectory, and Fig.7 (b)shows the most similar trajectory that is found by our method. The specific detection results of the two trajectories are shown in Table 2. As the similarity distribution between the tested trajectory and all accident trajectories exceeds 1% in the range [0.75,1], the final hazard level of the tested trajectory to be tested is 10 according to table 2. It is proved that the two trajectories are the driving trajectories of the same driver in different time periods.

TABLE 2 COMPARISON RESULTS OF TRAJECTORY SIMILARITY

Trajectory data	**Detected line segments**	**The Fitting equations**	**Trajectory feature vectors**
The tested trajectory	(37,18,484,246)-LINE (484-246,624,380)-Circle (624,380,715,668)-LINE	y=0x+0 y=3x+1612	$\begin{bmatrix} 260 & 132 \\ 507 & 360 \\ 669 & 524 \end{bmatrix}$
The found trajectory	(46,917,612,643)-LINE (612,643,789,482)-Circle (789,482,904,135)-LINE	y=0x-935 y=-3x-2922	$\begin{bmatrix} 329 & 780 \\ 645 & 501 \\ 846 & 308 \end{bmatrix}$



(a)The tested trajectory (b)The found trajectory

Fig.7 Map matching results of the tested trajectory and the found trajectory

VI. CONCLUSION

The micro driving behavior of drivers has great influence on traffic safety. For this problem, this paper proposes a traffic safety risk assessment method based on the analysis of drivers' driving trajectory.The main contributions include the following two aspects:(1) The K-means algorithm, Bresenham algorithm and seed filling algorithm are used to preprocess the trajectory, and a sequence of trajectory points reflecting the geometric characteristics of the trajectory is obtained; (2) By analyzing the trajectory geometry, the trajectory feature vector is established to form the data structure representing the trajectory geometric features, and the trajectory safety risk assessment method based on trajectory geometry is proposed. The effectiveness of our method in detecting drivers' driving habits is proved by an example. This method has practical significance to analyze the safety of drivers' micro driving behavior and regulate drivers' habits.

ACKNOWLEDGMENTS

This work was supported in part by Natural Science Foundation of China under Grants 62162047, Natural Science Foundation of Inner Mongolia under Grants 2019MS06029, Inner Mongolia Science and Technology Plan Project under Grant 2021GG0155, the Open-topic of Inner Mongolia Big Data Laboratory for Discipline Inspection and Supervision (IMDBD2020012, IMDBD2021014), Inner Mongolia Engineering Laboratory for Big Data Analysis Technology.

REFERENCES

[1] Huang Shi-chen, Shao Chun-fu, Li Juan, et al. Vehicle Trajectory Reconstruction and Anomaly Detection Using Deep Learning[J]. Journal of Transportation Systems Engineering and Information Technology,2021,21(03):47-54.

[2] Li Yan, Li Hao, Qian Xiao-Lu,et al. A review and analysis of outlier detection algorithms[J]. Computer Engineering,2002(06):5-32.

[3] Mao Jiang-Yun, Wu Hao, Sun Wei-Wei. Vehicle Trajectory Anomaly Detection in Road Networkvia Markov Decision Process[J]. Chines Journal of Computers,2018,41(08):1928-1942.

[4] Xin Jirong, Luo Source. A pure orientation trajectory detection method based on hidden Markov model[J]. Systems Engineering and Electronics Technology, 2016, 38(07): 1496-1501.

[5] Zhao Qianqian. Research on Vehicle Detection Algorithm Based on Target Feature Point Tracking and Clustering[D]. Xi'an: Chang'an University, 2015.

[6] Kim Soyeun, Lee Keumjin. En-route Trajectory Prediction via Weighted Linear Regression[J]. Journal of the Korean Society for Aviation and Aeronautics, 2016, 24(4): 44-52.

[7] Zhang Xian Yang, Zhu Xiaoyu, Lin Haoshen et al. A trajectory prediction algorithm based on Gaussian hybrid-variational self-encoder[J]. Computer Engineering, 2020, 46(07): 50-57.

[8] Li Wangao, Zhao Xuemei, Sun Dezhuang. Research on trajectory prediction algorithm based on improved Bayesian method[J]. Computer Applications, 2013, 33(7): 1960-1963.

[9] Venkatesan Kanagaraj. Trajectory Data and Flow Characteristics of Mixed Traffic[J].Transportation Research Record, 2015, 2491(1): 1-11.

[10] Vasileios Maroulas. Error analysis of stochastic flight trajectory prediction models[J]. Journal of Applied Statistics, 2012, 39(8): 1825-1841.

[11] Zhang Yu, Chen Jing,Shen Lincheng.Real-time trajectory planning for UCAV air-to-surface attack using inverse dynamics optimization method and receding horizon control[J]. Chinese Journal of Aeronautics, 2013, 26(04): 1038-1056.

[12] Xuedong Zhang, et al. Evaluation of multi-source forcing datasets for drift trajectory prediction using Lagrangian models in the South China Sea[J]. Applied Ocean Research, 2020, 104: 1-20.

[13] Zhipeng Liu, Dechang Pi, et al. Density-based trajectory outlier detection algorithm[J]. Journal of Systems Engineering and Electronics, 2013, 24(02): 335-340.

[14] Shuming Sun, Juan Chen, Jian Sun. traffic congestion prediction based on GPS trajectory data[J]. International Journal of Distributed Sensor Networks, 2019, 15(5): 1-18.

[15] Karlaftis M G, Vlahogianni EI. Statistical methods versus neural networks in transportation research:Differences, similarities and some insights[J]. Transportation Research Part C Emerging Technologies, 2011, 19(3): 387-399.

[16] Xiangjie Kong, Zhenzhen Xu, et al. Urban traffic congestion estimation and prediction based on floating car trajectory data[J]. Future Generation Computer Systems, 2016, 61: 97-107.

[17] Yang W Q, Gao Y, et al. TRASMIL: A local anomaly detection framework based in trajectory segmentation and multi-instance learning[J]. Computer Vision and Image Understanding, 2013, 117: 1273-1286.

[18] Mikluscak T, Gregor M, Janota A. Using Neural Networks for Route and Destination Prediction in Intelligent Transport Systems[J]. Communications in Computer & Information Science, 2012, 329(329): 380-387.

[19] Lu, Yunfei, Chen, Linqiang. SOM-based pedestrian anomaly trajectory detection[J]. Journal of Hangzhou University of Electronic Science and Technology (Natural Science Edition), 2015, 35(02): 45-49.

[20] De Brebisson A, Simon E, Auvolat A, et al. Artificial Neural Networks Applied to Taxi Destination Prediction[J].Computer Science, 2015: 1-18