引文格式：周琪姗,杨璨,李健,等. 基于轨迹方向特征的道路交叉口提取算法[J]. 地理空间信息, 2023, 21(7): 7-10. doi:10. 3969/j.issn. 1672-4623.2023.07.002

基于轨迹方向特征的道路交叉口提取算法

周琪姗[1]，杨　璨[1*]，李　健[2]，颜　凯[2]
（1. 武汉大学 遥感信息工程学院，湖北 武汉 430079；2. 重庆长安汽车股份有限公司，重庆 401120）

摘　要：作为城市路网骨架的定位节点，道路交叉口在城市交通地理信息服务中扮演着重要角色。浮动车GPS轨迹数据包含丰富的路网信息，具有成本低、时效性高、获取方便等优点，已被广泛应用于道路交叉口提取的相关研究中。以浮动车GPS轨迹数据为研究对象，提出了一种道路交叉口提取算法；同时考虑交叉口区域浮动车行驶方向变化和行驶方向的差异性，分别从单条轨迹和全部轨迹出发提取交叉口附近的转向点集，并对这两组转向点集进行聚类和整合，得到道路交叉口位置信息。该算法在公开的浮动车轨迹数据集上进行了测试，结果表明算法能有效识别道路交叉口。

关键词：道路交叉口；GPS轨迹数据；行驶方向
中图分类号：P228.4　　**文献标志码：**B　　**文章编号：**1672-4623（2023）07-0007-04

Road Intersection Extraction Algorithm Based on Trajectory Direction Features

ZHOU Qishan[1], YANG Can[1], LI Jian[2], YAN Kai[2]
(1. School of Remote Sensing and Information Engineering, Wuhan University, Wuhan 430079, China;
2. Chongqing Chang'an Automobile Co., Ltd., Chongqing 401120, China)

Abstract: As the positioning nodes of urban road network skeleton, road intersection plays an important role in urban traffic geographical information service. The GPS trajectory data of floating car contains rich road network information, has the advantages of low cost, high timeliness and convenient acquisition, and has been widely used in the road intersection extraction research. In this paper, we proposed a road intersection extraction algorithm based on GPS trajectory data of floating cars. Considering the change of driving direction and the difference of driving direction of floating cars in the intersection area at the same time, we extracted the set of turn-points near the intersections from a single trajectory and all trajectories respectively, and obtained the location information of road intersections by clustering and integrating these two sets of turn-points. Finally, we tested the algorithm on a publicly available floating car trajectory data set. The result shows that the algorithm can effectively identify road intersection.

Key words: road intersection, GPS trajectory data, driving direction

道路交叉口是整个城市路网的枢纽，是定位路网骨架的关键节点，并承担着连接不同道路段的重任。目前已有很多学者采用不同方法实现了道路交叉口识别，如通过高分辨率遥感影像[1]、车载激光点云[2]提取道路交叉口，但通常存在价格高、不易获取、更新周期长[3]等缺点，无法适应交叉口信息实时更新的需求。装载了GPS定位装置的浮动车为城市道路信息的获取提供了新思路。随着定位技术的发展，GPS轨迹数据变得越来越方便获取，且其价格低廉、覆盖范围广、即时性高，已被广泛应用于道路交叉口的提取工作中。利用GPS轨迹数据提取交叉口的方法主要分为两类。①先提取道路中心线信息，再通过中心线的交点得到道路交叉口，通常通过数学形态学思想来实现，如胡瀚[4]等将轨迹数据栅格化得到道路中心线，再通过统计每个像素8邻域内像素点的个数来确定交叉口位置。②根据交叉口附近轨迹点的特征设计算法对交叉口进行提取，如Fathi A[5]等通过创建形状描述符并训练分类器的方式区分交叉点和非交叉点；CHEN C[6]等利用基于图的聚类技术和道路先验知识来发现道路段，再通过监督学习框架来识别交叉口；王飞[7]等采用行驶方向角之差识别转弯片段并得到其中心点，同时设计了一种自适应聚类半径的DBSCAN算法识别交叉口；Karagiorgou S[8]等通过车辆在速度和方向上的变化来识别转弯点，再利用转弯点提取路网中的交叉口；CHEN B[9]等利用角度阈值和速度阈值从输入轨迹中选择转向点，并对转向点进行聚类，再利用

收稿日期：2022-05-19；修回日期：2022-06-08。
项目来源：国家重点研发计划资助项目（2020YFC1512003）；重庆市技术创新与应用发展专项资助项目（cstc2021jscx-dxwtBX0023）。
第一作者简介：周琪姗（1997—），硕士研究生，研究方向为时空数据分析与挖掘，E-mail：qszhou@whu.edu.cn。
通信作者：杨璨（1990—），博士研究生，研究方向为时空数据挖掘，E-mail：yangcan@whu.edu.cn。

细化算法对交叉口进行分类。上述算法都获取了道路交叉口的位置信息，但数学形态学方法在计算过程中往往会丢失移动对象的运动属性信息；监督方法则需先验知识来支撑交叉口分类过程；考虑了运动属性的方法更多地关注单个移动对象的转向信息，可能会遗失原本应被识别为交叉口的位置，导致识别结果不完整。因此，如何充分利用浮动车GPS轨迹数据，实现更全面的道路交叉口位置信息提取，仍是研究的热点。本文以浮动车行驶方向特征为切入点，分别从单条轨迹和整体轨迹的角度出发，设计了一种同时考虑了浮动车行驶方向变化和行驶方向差异性的道路交叉口提取算法，无需先验知识支持，可保留轨迹数据运动属性信息，并多角度利用浮动车行驶方向特征，获得了更加完整的道路交叉口提取结果。

1 基于轨迹方向特征的道路交叉口提取

本文的技术路线主要包括3个步骤：①轨迹数据预处理，对原始轨迹数据构建KD-tree索引和数据清洗；②转向点提取，分别根据浮动车的行驶方向变化和行驶方向差异性提取转向点；③交叉口识别，将两种方式得到的转向点分别聚类，再整合聚类结果，得到道路交叉口位置信息。

1.1 轨迹数据预处理

浮动车轨迹数据可理解为浮动车在道路上行驶过程中经过的位置点样本，对于任意轨迹点 P_i，通常由定位点的位置信息和定位时刻等属性组成。基于这些已知信息，本文对轨迹数据进行预处理，包括轨迹数据的索引构建和数据清洗两个部分。

1.1.1 轨迹数据索引构建

轨迹数据作为一种包含位置和时间属性的空间数据，可通过构建索引来提高算法的查询效率。本文对轨迹数据建立基于KD-tree的空间索引。KD-tree支持范围查询和K最近邻（KNN）查询。本文利用范围查询搜索以某一个轨迹点为圆心、一定距离为半径的圆形缓冲区内的轨迹点，利用KNN查询查找某个轨迹点周围最近的 k 个轨迹点，分别在轨迹数据清洗和转向点提取中使用。

1.1.2 轨迹数据清洗

本文从两个方面对轨迹数据中的冗余和噪声数据进行清洗：①限制浮动车行驶过程中的最大速度；②限制轨迹点周围包含的最少轨迹点数，将不符合要求的轨迹点视为噪声点并剔除。

1） P_i 处浮动车的行驶速度 V_i 记为 P_{i-1} 到 P_{i+1} 的平均速度。本文通过箱线图的方式计算得到速度正常值的上限，记为 $V_{\max}$，对于轨迹数据集中的任一轨迹点，若其速度 $V_i > V_{\max}$，则认为该点为漂移点，进行剔除。

2）以轨迹点 P_j 为圆心，以一定距离 R_{drift} 为半径作圆形缓冲区，并设置圆形缓冲区内的最小点数为 $p_\text{num}_{\min}$，对轨迹点 P_j 进行基于KD-tree的范围查询，获取该圆形缓冲区内的轨迹点信息，设缓冲区内轨迹点的个数为 p_num_j，若 $p_\text{num}_j < p_\text{num}_{\min}$，则认为 P_j 为噪声点，进行剔除。

1.2 转向点提取

道路交叉口位置上的轨迹点在行驶方向变化和行驶方向分布等方面与道路中间段位置区别较大，据此对浮动车行驶过程中的转向点进行提取。对于数据集中的任一轨迹定位点 P_i，其行驶方向角 θ_i 的定义为：由定位点 P_i 指向下一个定位点 P_{i+1} 的向量 $\overrightarrow{P_iP_{i+1}}$ 与 X 轴正方向的夹角。

本文从两个角度对浮动车行驶过程中可能发生转向行为的位置进行识别，并提取转向点信息。从单条轨迹的角度，当浮动车两相邻定位点的行驶方向角之差大于一定阈值时，考虑在该位置附近发生了转向行为；从整体轨迹的角度，当某个定位点与其周围几个定位点之间的行驶方向角有较大差异时，考虑该位置附近发生了转向行为。

1.2.1 基于行驶方向变化的转向点提取

浮动车在行驶过程中的行驶方向角变化可在一定程度上反映道路的转向情况，通过计算相邻两个GPS定位点处浮动车行驶方向的变化情况，可获取浮动车的转向点信息。

1）转向角计算。转向角为两个相邻轨迹点的行驶方向角之差。相邻轨迹点 P_i 和 P_{i+1} 的方向向量可构成相加为360°的两个角 $\Delta\theta_1$ 和 $\Delta\theta_2$，取二者中较小的一个作为 P_i 到 P_{i+1} 的转向角，并记为 $\Delta\theta_{ii+1}$。

2）转向点提取。设浮动车在非交叉口道路上行驶的最大转向角阈值为 $\Delta\theta_{\max}$，若行驶过程中相邻两点 P_i 和 P_{i+1} 的转向角 $\Delta\theta_{ii+1} \geqslant \Delta\theta_{\max}$，则认为浮动车在这两个定位点之间可能发生了转向，$P_i$ 和 P_{i+1} 为一对转向点对。

图1a中 P_t 更能代表 P_i 到 P_{i+1} 的转向过程，因此本文将 P_t 定义为转向点，并将 O、O' 的中点记为 P_t 的位置；图1b中 O 为 P_i 与 P_{i+1} 连线的中点，O' 为过 P_i 和 P_{i+1} 的垂线和水平线的交点，当点 P_{i+1} 的行驶方向角向量位于 $\overrightarrow{P_iP_{i+1}}$ 连线的右下方时，从轨迹点 P_i 到 P_{i+1} 的行驶路径如绿色线条所示，则转向点为 O'_1

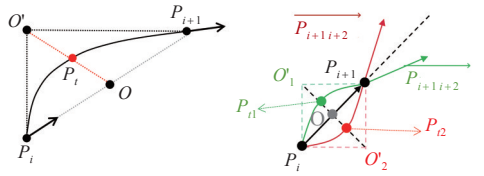


图1　转向点计算

与 O 的连线中点 P_{t1}；同理，当点 P_{i+1} 的行驶方向角向量位于 $\overrightarrow{P_iP_{i+1}}$ 连线的左上方时，转向点为 P_{t2}。计算所有通过行驶方向变化得到的转向点，并记为 $\mathrm{Set}P_t$，则 $\mathrm{Set}P_t=\{P_{t1}, P_{t2}, \cdots\}$。

1.2.2　基于行驶方向差异性的转向点提取

在道路交叉口位置，不同浮动车的行驶方向角之间往往存在一定差异性。对于行驶在道路中间段的车辆，行驶方向通常集中在一个方向或两个相反的方向，而交叉口位置的行驶方向则可能沿各个方向不均匀分布。基于这种区别，本文通过分析某个轨迹点与其周围轨迹点行驶方向之间的差别，提取可能位于转向区域的轨迹点，并将其记为转向点。

1）中心点与近邻点集。对于数据集中的任一轨迹定位点 P_i，设其为一次搜索过程的中心点，通过KNN查询得到与 P_i 距离最近的 k 个轨迹点组成的近邻点集，并记为 Points_k；分别计算点集 Points_k 中每一个点 P_j 与中心点 P_i 的行驶方向角之差，并记为 $\Delta\alpha_j$，方法与转向角计算相同，得到以 P_i 为中心点的行驶方向角差值序列，记为 $\mathrm{Set}\Delta\alpha_i$，则 $\mathrm{Set}\Delta\alpha_i=\{\Delta\alpha_1, \Delta\alpha_2, \cdots, \Delta\alpha_k\}$。

2）转向点提取。考虑到城市道路一般具有双向行驶规则，因此行驶方向角相同或互补的两个轨迹点均不认为其行驶方向相异。设两个轨迹点行驶方向角之差的最大阈值为 $\Delta\alpha_{\max}$，若满足 $\Delta\alpha_j\in(\Delta\alpha_{\max}, 180°-\Delta\alpha_{\max})$，则将点 P_j 记为与中心点 P_i 行驶方向相异的点。

对于点集 Points_k，设其中与 P_i 行驶方向相异的轨迹点个数为 k'，那么该点集中行驶方向相异点的占比为 $\gamma=k'/k$，设 γ 的最小值阈值为 γ_0，若 $\gamma\geqslant\gamma_0$，则认为中心点 P_i 与其周围轨迹点在行驶方向特征上有较大差异，并将 P_i 视为转向点，记为 P_k。计算得到所有转向点，并将通过行驶方向差异性得到的转向点集，记为 $\mathrm{Set}P_k$，则 $\mathrm{Set}P_k=\{P_{k1}, P_{k2}, \cdots\}$。

1.3　交叉口识别

本文分别对转向点集 $\mathrm{Set}P_t$ 和 $\mathrm{Set}P_k$ 进行聚类，得到两组聚类中心，由于这些转向点大多集中分布在浮动车发生明显转向行为的区域，因此可认为这两组聚类中心为道路交叉口的候选点。由于交叉口的数量和位置均是未知的，且越靠近交叉口中心位置转向点的分布越密集，因此本文采用DBSCAN聚类算法对轨迹点进行基于密度的聚类，通过设定聚类参数最小点数Min*Pts*和扫描半径Eps得到两组聚类中心点集，并将最终结果分别记为 $\mathrm{Set}C_t$ 和 $\mathrm{Set}C_k$，其中 $\mathrm{Set}C_t$ 中的每个聚类中心 C_t 包含该点的位置坐标 (x_t, y_t) 和对应聚类簇内转向点的数量 nC_t，$\mathrm{Set}C_k$ 同理。本文对这两组聚类中心进行整合，并将整合结果作为道路交叉口的中心点。

设两组候选点之间的距离最大阈值为 $\mathrm{Dist}_{\max}$，对于 $\forall C_{t_i}\in\mathrm{Set}C_t$，$\forall C_{k_j}\in\mathrm{Set}C_k$：若 C_{t_i} 和 C_{k_j} 之间的距离 $\mathrm{Dist}_{ij}\geqslant\mathrm{Dist}_{\max}$，则将两个候选点均保留为交叉口中心点；若 $\mathrm{Dist}_{ij}<\mathrm{Dist}_{\max}$，则需合并两个候选点，转向点坐标为两个候选点坐标的加权质心，权重由每个候选点所在聚类簇内的转向点个数决定，合并后得到中心点 C_{ij}。候选点合并后得到最终的交叉口中心点数据集，记为SetNode，则 $\mathrm{SetNode}=\{\mathrm{Node}_1, \mathrm{Node}_2, \cdots, \mathrm{Node}_n\}$。其中，每个交叉口中心点 $\mathrm{Node}_i(1\leqslant i\leqslant n)$ 由交叉口编码NodeId和交叉口中心位置坐标 $C_N(x_N, y_N)$ 构成。

2　实验与结果分析

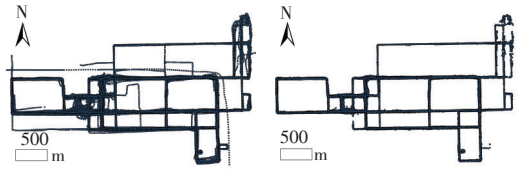
2.1　实验概况

本文采用的实验数据来自芝加哥市浮动车轨迹数据集[10]。该数据集记录了13辆浮动车共计约889条轨迹数据，全长约2 869 km，分布范围约为9.4 km^2，采样时间跨度约118 h，采样时间间隔在1~29 s不等。

1）数据预处理实验。首先对轨迹数据建立KD-tree索引，再分别基于速度和邻域限制清洗噪声点。在速度限制方面，利用箱线图的方式计算得到 $V_{\max}=18.29$ m/s，清理速度高于该数值的点；在缓冲区内点数限制方面，根据轨迹点分布情况，设置不同的 R_{drift} 和 $p_\mathrm{num}_{\min}$ 参数值，从而在清理噪声点的同时保留更多的正常轨迹点。轨迹数据清洗前后的分布形态如图2所示。

2）转向点提取实验。本文基于行驶方向特征，从两个角度提取转向点：①基于行驶方向变化的转向点提取，参数设置为 $\Delta\theta_{\max}=60°$；②基于行驶方向差异性的转向点提取，参数设置为 $\Delta\alpha_{\max}=45°$、$k=10$、$\gamma_0=0.6$，得到两组转向点集 $\mathrm{Set}P_t$ 和 $\mathrm{Set}P_k$。

3）交叉口识别实验。利用DBSCAN聚类算法分别计算两组转向点 $\mathrm{Set}P_t$、$\mathrm{Set}P_k$ 的聚类中心。聚类参数设置为Eps=20 m，Min*Pts*=5，最终基于行驶方向



a 清洗前　　b 清洗后

图2　轨迹数据清洗前后对比

变化的转向点聚类得到了44个交叉口中心候选点，基于行驶方向差异性的转向点聚类得到了34个交叉口中心候选点。对两种方法识别得到的候选点进行整合处理：保留仅被一种方法识别出来的候选点，合并被两种方法同时识别出来且 $Dist_{ij} < Dist_{max}$ 的候选点。设 $Dist_{max} = 25$ m，得到51个道路交叉口中心点（图3）。

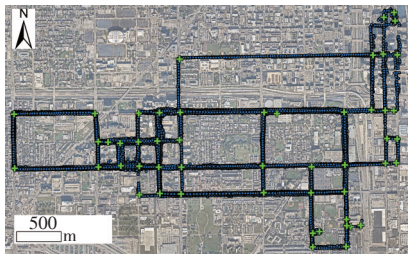


图3　交叉口提取结果

2.2　实验结果分析

本文将识别结果与人工标注的50个道路交叉口位置进行对比，采用精确率和召回率进行精度评价，并以40 m为阈值，评价结果见表1，可以看出，两种方法的整合结果在精确率和召回率上均有提升，召回率提升显著，得到了更加完整全面的交叉口识别结果。

表1　道路交叉口提取结果评价/%

识别方法	精确率	召回率
基于行驶方向变化	93.18	82.00
基于行驶方向差异性	91.18	62.00
两种方法整合	94.12	96.00

此外，本文将识别的转向点与参考文献[8]、[9]中识别的转向点进行比较，结果见表2，相较于参考文献[8]中通过车辆在速度和方向上的变化识别的转向点，本文方法在精确率和召回率上均有提高；相较于参考文献[9]中利用角度阈值和速度阈值识别的转向点，本文方法的精确率略低，但召回率显著提高。

表2　转向点识别结果评价

指标	参考文献[8]方法	参考文献[9]方法	本文方法
转向点个数/个	8 991	3 951	5 591
精确率/%	89.80	100.00	91.59
召回率/%	88.00	74.00	88.02

3　结　语

本文提出了一种基于浮动车行驶方向特征的道路交叉口提取方法，充分利用GPS轨迹数据包含的浮动车在城市道路上行驶时的运动状态信息，从浮动车行驶方向的变化和差异性两个方面分别提取转向点，然后通过转向点聚类和聚类结果整合，得到道路交叉口的位置信息，并在该过程中通过KD-tree索引提高了轨迹点查询的效率。最后通过芝加哥市浮动车轨迹数据集对本文方法的可行性和有效性进行了验证。相较于只考虑单一行驶方向特征，本文方法显著提高了交叉口识别的召回率，得到了更加完整的识别结果；但本文方法仅考虑了交叉口的位置信息而未考虑形态信息，今后的研究方向是对交叉口结构进行提取，从而得到更加完整的交叉口几何信息。

参考文献

[1] 李润生,司毅博,马德宝,等. 基于语义规则的高分辨率影像道路交叉口提取方法[J]. 测绘科学技术学报,2017,34(2):168-174

[2] 方莉娜,王康. 基于车载激光点云的道路交叉口检测与识别[J]. 南京信息工程大学学报(自然科学版),2021,13(6):635-644

[3] 汪剑云,刘岩,李兵,等. 基于车载GPS技术的道路更新系统设计[J]. 地理空间信息,2012,10(3):97-99

[4] 胡瀚,向隆刚,王德浩. 出租车轨迹数据的道路提取[J]. 测绘通报,2018(7):53-57

[5] Fathi A, Krumm J. Detecting Road Intersections from GPS Traces[C]. International Conference on Geographic Information Science. Springer, Berlin, Heidelberg, 2010:56-69

[6] CHEN C, LU C, HUANG Q, et al. City-scale Map Creation and Updating Using GPS Collections[C]//Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining,2016:1 465-1 474

[7] 王飞,郭庆胜,徐杏琳. 基于轨迹数据的道路交叉口及其结构提取方法[J]. 测绘科学,2022,47(1):212-218

[8] Karagiorgou S, Pfoser D. On Vehicle Tracking Data-based Road Network Generation[C]//Proceedings of the 20th International Conference on Advances in Geographic Information Systems,2012:89-98

[9] CHEN B, DING C, REN W, et al. Extended Classification Course Improves Road Intersection Detection from Low-frequency GPS Trajectory Data[J]. ISPRS International Journal of Geo-information,2020,9(3):181

[10] Ahmed M, Karagiorgou S, Pfoser D, et al. A Comparison and Evaluation of Map Construction Algorithms[J]. GeoInformatica,2014(9):601-632