



(12) 发明专利

(10) 授权公告号 CN 110570658 B

(45) 授权公告日 2022. 02. 01

(21) 申请号 201911012550.4

G06K 9/62 (2006.01)

(22) 申请日 2019.10.23

(56) 对比文件

(65) 同一申请的已公布的文献号  
申请公布号 CN 110570658 A

CN 107240264 A, 2017.10.10  
CN 109410586 A, 2019.03.01  
CN 109242024 A, 2019.01.18  
US 2019304298 A1, 2019.10.03  
CN 108571975 A, 2018.09.25  
CN 108665704 A, 2018.10.16  
CN 107103754 A, 2017.08.29  
CN 108986453 A, 2018.12.11

(43) 申请公布日 2019.12.13

(73) 专利权人 江苏智通交通科技有限公司  
地址 210006 江苏省南京市秦淮区应天大街388号晨光1865科技创意产业园E10幢三楼

韩旭.基于车辆轨迹多特征的聚类分析及异常检测方法的研究.《中国优秀硕士学位论文全文数据库》.2018,

(72) 发明人 吕伟韬 周东 张子龙 李璐

审查员 赵云峰

(74) 专利代理机构 南京正联知识产权代理有限公司 32243

代理人 张玉红

(51) Int. Cl.

G08G 1/01 (2006.01)

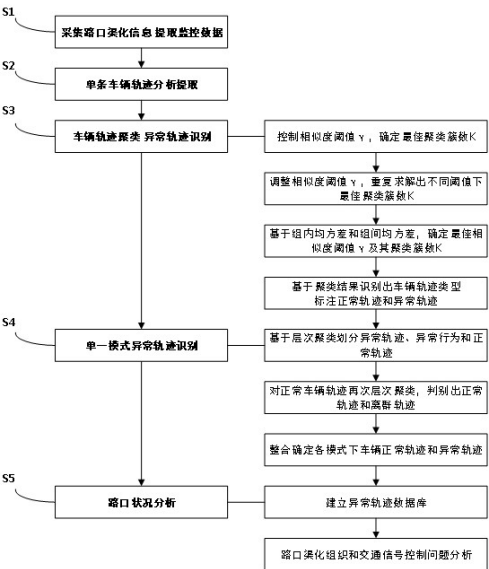
权利要求书4页 说明书13页 附图7页

(54) 发明名称

基于层次聚类的路口异常车辆轨迹识别分析方法

(57) 摘要

基于层次聚类的路口异常车辆轨迹识别分析方法,利用LCSS算法、DTW算法和层次聚类算法进行车辆轨迹聚类分析,基于路口渠化信息实现车辆轨迹模式划分,识别出路口异常车辆轨迹,进而分别对各模式下车辆轨迹再次聚类分析,精细化识别各流向异常车辆轨迹,从而建立异常轨迹数据库,为路口渠化设计和信控方案合理性评估提供支撑依据。本发明提高车辆轨迹层次聚类效果,能够有效识别出路口车辆轨迹类型和异常车辆轨迹,进一步对各模式类型下车辆轨迹分析,对各流向异常车辆轨迹识别提取,从而为路口交通冲突安全管理和拥堵管理提供有效支撑;通过对统计时间段内异常轨迹分析,有效识别出路口渠化组织的问题,对交通信号配时方案的优化调整提供支撑依据。



1. 基于层次聚类的路口异常车辆轨迹识别分析方法, 其特征在于: 包括如下步骤:

步骤1, 采集路口渠化信息, 提取路口监控视频数据完成数据清洗;

步骤2, 针对不同车辆ID实现单条轨迹分析, 提取数据特征点, 确定车辆轨迹;

步骤3, 基于LCSS算法和层次聚类算法实现路口车辆轨迹聚类, 划分出车辆轨迹模式类型, 识别出正常车辆轨迹和异常车辆轨迹;

所述步骤3中, 控制相似度阈值, 确定簇数的范围, 基于层次聚类算法确定相似度阈值下最佳聚类簇数以及其聚类结果; 重复步骤, 调整相似度阈值, 重复求解出不同相似度阈值下的最佳聚类簇数; 基于最佳聚类簇数求解出不同相似度阈值的组内均方差和组间均方差, 确定最优阈值及其最优组数; 基于最佳阈值和最优聚类簇数, 确定车辆轨迹的聚类结果, 识别出车辆轨迹模式类型, 确定正常车辆轨迹和异常车辆轨迹;

步骤4, 基于步骤3模式分析的车辆轨迹类型提取出单一模式正常车辆轨迹, 进一步对单一模式车辆轨迹分析, 识别出单一模式下异常车辆轨迹和正常车辆轨迹;

所述步骤4中, 提取出单一模式下正常车辆轨迹, 基于正常车辆轨迹的轨迹特征点数值, 以轨迹的轨迹相似度、加速度方差、弧长比作为特征数值进行层次聚类, 划分出异常行驶轨迹、异常行为轨迹和正常车辆轨迹; 整合正常车辆轨迹再次进行层次聚类, 判别出正常车辆轨迹和离群轨迹; 重复步骤, 对路口各模式下车辆轨迹分析, 划分各流向中异常行驶轨迹、异常行为轨迹、离群轨迹和正常车辆轨迹, 其中将异常行驶轨迹、异常行为轨迹和离群轨迹均默认为异常车辆轨迹;

步骤5, 基于步骤3和4中的异常车辆轨迹, 对路口状况进行分析。

2. 根据权利要求1所述的基于层次聚类的路口异常车辆轨迹识别分析方法, 其特征在于: 所述步骤1中, 具体包括如下分步骤:

步骤1-1, 对路口类型及其渠化信息进行采集;

步骤1-2, 基于单位时间段内路口监控视频根据不同车辆ID提取出原始轨迹点, 记为 $P(f, x, y)$ , 其中 $f$ 表示帧数,  $x$ 和 $y$ 表示轨迹点坐标数值;

步骤1-3, 基于原始轨迹点 $P$ 在二维坐标内绘制出车辆原始轨迹点, 从中确定缺失重要特征数据和不正确数据并进行剔除, 实现原始轨迹点清洗; 其中缺失重要特征数据为偏离数据, 其可组成短路径但偏离线路走向, 不正确数据即无法平滑连接的密集散点集, 其无法构成路径轨迹。

3. 根据权利要求1所述的基于层次聚类的路口异常车辆轨迹识别分析方法, 其特征在于: 所述步骤2中, 具体包括如下分步骤:

步骤2-1, 基于同一车辆ID下的原始轨迹数据 $P$ 进行分析, 在统计时间段 $T$ 内对单位时间提取定量数据实现轨迹特征点提取;

步骤2-2, 基于提取的特征点进行轨迹绘制, 根据 $f$ 帧数大小排列, 确定车辆轨迹, 记为 $TR = \{P|p_i, 1 \leq i \leq n, n \text{ 为轨迹点数}\}$ 。

4. 根据权利要求1所述的基于层次聚类的路口异常车辆轨迹识别分析方法, 其特征在于: 所述步骤3中, 具体包括如下分步骤:

步骤3-1, 控制相似度阈值 $\gamma$ , 确定簇数 $K$ 的范围, 基于层次聚类算法确定相似度阈值 $\gamma$ 下最佳聚类簇数 $K_{opi}$ 以及其聚类结果;

步骤3-2, 重复步骤3-1, 调整相似度阈值 $\gamma$ , 重复求解出不同相似度阈值 $\gamma$ 下的最佳聚

类簇数 $K_{opi}$ ,记为 $(\gamma_l, K_{opi}^l)$ ,其中 $l$ 表示相似度阈值序号,其中相似度阈值 $\gamma_l$ 数值取整数;同时各 $K_{opi}^l$ 下含有最优簇数的组内距均值 $\bar{\tau}_k^l$ 和组内距 $\varphi_k^l$ ;

步骤3-3,基于 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组内均方差 $S_\tau^l$ 和组间均方差 $S_\phi^l$ ,根据组内均方差 $S_\tau^l$ 和组间均方差 $S_\phi^l$ 确定最优阈值 $\gamma_{opi}$ 及其最优组数 $K_{opi}$ ;

步骤3-4,基于最佳阈值 $\gamma_{opi}$ 和最优聚类簇数 $K_{opi}$ ,确定车辆轨迹的聚类结果,识别出车辆轨迹模式类型,确定正常车辆轨迹和异常车辆轨迹;

基于步骤3-3确定的最佳阈值 $\gamma_{opi}$ 和最优聚类簇数 $K_{opi}$ 进行车辆轨迹层次聚类,根据步骤1的路口渠化信息,对聚类的车辆轨迹模式进行分类,将根据路口渠化流向下划分以外的车辆轨迹默认为异常车辆轨迹,其余则为正常车辆轨迹。

5.根据权利要求4所述的基于层次聚类的路口异常车辆轨迹识别分析方法,其特征在于:所述步骤3-1中,包括如下分步骤:

步骤3-1-1,控制相似度阈值 $\gamma$ ,通过LCSS算法确定两两轨迹之间的最长公共子序列和最长公共子序列相似度距离;

步骤3-1-2,根据上一步骤的最长公共子序列相似度距离列出相似度矩阵 $S[a][b]$ ,即为邻近矩阵;

步骤3-1-3,确定聚类簇数 $K$ 的范围;具体来说,根据步骤1的路口渠化信息确定聚类簇数 $K$ 的最小数值;

步骤3-1-4,给定簇数 $K$ ,不断重复步骤3-1-1和3-1-2得到层次聚类的聚类结果;

步骤3-1-5,根据步骤3-1-3给定的簇范围确定出控制相似度阈值 $\gamma$ 一定时 $K$ 值不同的聚类情况,根据不同 $K$ 值聚类下的组内距和组间距建立评价体系,从而确定最佳聚类簇数 $K_{opi}$ ,如下:

根据步骤3-1-4的聚类结果提取出组内轨迹的所有特征点 $P_i$ ,根据特征点 $P_i$ 确定划分 $K$ 组情况下 $k$ 个聚类簇中心 $\{C_1, C_2, C_3, \dots, C_k\}$ , $1 < k \leq K$ ,其中簇中心 $C_k$ 为该组内所有特征点的中心;

根据 $k$ 组簇中心分别求得 $k$ 组内的组内距 $\tau_k$ 以及组间距 $\varphi_k$ ;

基于不同组数 $K$ 下的组内距均值 $\bar{\tau}_k$ 和组间距 $\varphi_k$ ,确定最佳聚类簇数 $K_{opi}$ ,具体来说,将比值 $\alpha = \frac{\varphi_k}{\bar{\tau}_k}$ 最大时的 $k$ 值默认为 $K_{opi}$ 。

6.根据权利要求4所述的基于层次聚类的路口异常车辆轨迹识别分析方法,其特征在于:所述步骤3-3中,包括如下分步骤:

步骤3-3-1,根据 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组内均方差 $S_\tau^l$ ;

步骤3-3-2,根据 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组间均方差 $S_\phi^l$ ;

步骤3-3-3,基于不同的相似度阈值 $\gamma$ 及其 $K_{opi}$ 下的组内均方差 $S_\tau^l$ 和组间均方差 $S_\phi^l$ ,确定最优阈值 $\gamma_{opi}$ ,将比值 $\beta = \frac{S_\phi^l}{S_\tau^l}$ 最大时的 $\gamma_l$ 默认为最优阈值。

7.根据权利要求1所述的基于层次聚类的路口异常车辆轨迹识别分析方法,其特征在

于:所述步骤4中,具体包括如下分步骤:

步骤4-1,提取出单一模式下正常车辆轨迹,基于正常车辆轨迹的轨迹特征点数值,以轨迹的轨迹相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 作为特征数值进行层次聚类,划分出异常行驶轨迹、异常行为轨迹和正常车辆轨迹;

步骤4-2,整合上一步骤分析得到的正常车辆轨迹,在DTW算法下对上一步骤的正常车辆轨迹再次进行层次聚类,判别出正常车辆轨迹和离群轨迹;

步骤4-3,重复步骤4-1至4-2,对路口各模式下车辆轨迹分析,划分各流向中异常行驶轨迹、异常行为轨迹、离群轨迹和正常车辆轨迹,其中将异常行驶轨迹、异常行为轨迹和离群轨迹均默认为异常车辆轨迹。

8.根据权利要求7所述的基于层次聚类的路口异常车辆轨迹识别分析方法,其特征在于:所述步骤4-1中,具体包括如下分步骤:

步骤4-1-1,对车辆轨迹 $TR_j$ 的轨迹相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 进行求解;

根据车辆轨迹特征确定相似度计算公式,如LCSS算法或DTW算法,求得车辆轨迹相似度 $\lambda$ ;

基于车辆轨迹 $TR$ 的轨迹特征点确定各车辆轨迹的加速度方差 $\alpha^2$ ,即:

$$\alpha_{i+1} = \frac{\sqrt{\|p_{i+1} - p_i\|^2}}{(f_{i+1} - f_i)^2}$$

$$\bar{\alpha} = \frac{\alpha_{i+1} + \dots + \alpha_n}{n-1}$$

$$\alpha^2 = \frac{\sum (\alpha_{i+1} - \bar{\alpha})^2}{n-1}$$

式中, $\alpha_{i+1}$ 表示特征点 $i+1$ 的加速度,其中 $f_i$ 和 $f_{i+1}$ 表示帧数, $\sqrt{\|p_{i+1} - p_i\|^2}$ 表示特征点 $i$ 与特征点 $i+1$ 之间的欧式距离, $p_i$ 和 $p_{i+1}$ 表示特征点;

基于车辆轨迹 $TR$ 确定车辆轨迹弧长比 $\sigma$ ,即:

$$\sigma = \frac{\sum_{i=1}^{n-1} \sqrt{\|p_{i+1} - p_i\|^2}}{\sqrt{\|p_n - p_1\|^2}}$$

式中, $p_1$ 、 $p_i$ 、 $p_{i+1}$ 、 $p_n$ 均表示车辆轨迹内的特征点; $\sqrt{\|p_{i+1} - p_i\|^2}$ 表示车辆轨迹特征点 $i$ 和特征点 $i+1$ 之间的欧式距离; $\sqrt{\|p_n - p_1\|^2}$ 表示车辆轨迹特征点 $n$ 与特征点 $1$ 之间的欧式距离;

步骤4-1-2,基于上一步骤车辆轨迹 $TR$ 的相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 数值作为特征数据进行层次聚类,划分为异常行驶轨迹、异常行为轨迹和正常车辆轨迹;以相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 数值作为特征数据进行层次聚类,将车辆轨迹划分为三组数据,根据三组数据的数据量及各组数据离散程度确定数据类型;

计算三组车辆轨迹的轨迹数量及其与总轨迹数量的比值,若组内轨迹数量最少则那组车辆轨迹数据默认为异常行驶轨迹;

基于车辆轨迹的特征点分别求剩余两组车辆轨迹数据的离散程度 $\varepsilon$ ,即:

$$\varepsilon = \frac{\sum dist(p_i - \bar{p})^2}{N}$$

$$dist(p_i - \bar{p}) = \sqrt{\|p_i - \bar{p}\|^2}$$

$$\bar{p} = \frac{\sum p_i}{N}$$

式中, $N$ 为组内所有车辆轨迹特征点的总个数; $p_i$ 为特征点, $\bar{p}$ 为簇中心点; $dist(p_i - \bar{p})$ 表示 $p_i$ 点到簇中心点距离;

将离散程度 $\varepsilon$ 较大的那组数据默认为异常行为轨迹,较小的那组数据默认为正常车辆轨迹,同时实现正常车辆轨迹提取。

9. 根据权利要求1所述的基于层次聚类的路口异常车辆轨迹识别分析方法,其特征在于:所述步骤5中,具体包括如下分步骤:

步骤5-1,整合步骤3和4中的异常车辆轨迹,对异常车辆轨迹发生原因进行分析,同时将异常车辆轨迹与异常原因相对应,建立异常轨迹关联数据库;

步骤5-2,基于统计时间段内异常轨迹数目比例及其行为原因对路口渠化组织问题和交通信号控制问题进行分析,识别出不合理的路口渠化和不合理的路口信号方案。

10. 根据权利要求9所述的基于层次聚类的路口异常车辆轨迹识别分析方法,其特征在于:所述步骤5-2中,具体如下:

若统计时间段内异常轨迹数目大于路口异常状况阈值,则分析其渠化是否存在问题;

若同一流向下,统计时间段放行相位阶段内,异常轨迹数目大于异常状况阈值,则分析该相位阶段信号方案配置是否合理。

## 基于层次聚类的路口异常车辆轨迹识别分析方法

### 技术领域

[0001] 本发明涉及交通管控领域中的车辆轨迹识别领域和路口合理性分析领域,具体涉及一种基于层次聚类的路口异常车辆轨迹识别分析方法。

### 背景技术

[0002] 随着机动车数据的集聚增长,全国各城市都面临着日益严峻的交通拥堵问题,因此,对车辆轨迹的研究对于交通管理与疏导是十分必要的。迄今为止,已经有较多的学者对车辆轨迹相似性及异常轨迹进行了一定研究,如裴剑(裴剑,彭敦陆.一种基于LCSS的相似车辆轨迹查询方法[J].2016)在采集大量车辆行驶GPS数据基础上,基于数据清理形成轨迹,通过Ramer-Douglas-Peucker算法对原始轨迹轮廓抽取,基于LCSS算法求相似子轨迹。

[0003] 另一方面,现阶段城市道路上布设有大量的电警、卡口、球机等监控摄像头,产生了巨大的结构化数据,而将这些数据运用实现车辆轨迹分析更成了目前研究的主流之一。如发明CN201710492719.5提出一种车辆非有效行驶轨迹识别方法,通过对车辆轨迹的识别的层次聚类,实现最优路径推荐,针对停车问题进行轨迹路径规划;发明CN201510159009.1提出一种基于广域分布交通系统的异常轨迹检测方法,通过无监督聚类确定异常交通轨迹点和异常轨迹。

[0004] 现阶段的车辆轨迹分析研究主要集中在两个方面,一是基于GPS定位数据/手机移动通信数据等位置数据实现的单一车辆轨迹分析;二是对整个路网,长距离车辆行驶异常轨迹的判定,目前暂缺对路口车辆轨迹的有效分类和判别,以及车辆轨迹的路口交通管理运用。

### 发明内容

[0005] 本发明提出一种基于层次聚类的路口异常车辆轨迹识别分析方法,对路口球机监控设备历史视频数据进行提取,利用LCSS算法、DTW算法和层次聚类算法实现车辆轨迹聚类分析,基于路口渠化信息划分车辆轨迹模式,识别出路口异常车辆轨迹,进而分别对各模式下车辆轨迹再次聚类分析,精细化识别各流向异常车辆轨迹,分析异常原因,从而建立异常轨迹数据库,为路口渠化设计和信控方案合理性评估提供支撑依据。

[0006] 基于层次聚类的路口异常车辆轨迹识别分析方法,包括如下步骤:

[0007] 步骤1,采集路口渠化信息,提取路口监控视频数据完成数据清洗;

[0008] 步骤2,针对不同车辆ID实现单条轨迹分析,提取数据特征点,确定车辆轨迹;

[0009] 步骤3,基于LCSS算法和层次聚类算法实现路口车辆轨迹聚类,划分出车辆轨迹模式类型,识别出正常车辆轨迹和异常车辆轨迹;

[0010] 步骤4,基于步骤3模式分析的车辆轨迹类型提取出单一模式正常车辆轨迹,进一步对单一模式车辆轨迹分析,识别出单一模式下异常车辆轨迹和正常车辆轨迹;

[0011] 步骤5,基于步骤3和4中的异常车辆轨迹,对路口状况进行分析。

[0012] 进一步地,所述步骤1中,具体包括如下分步骤:

[0013] 步骤1-1,对路口类型及其渠化信息进行采集;

[0014] 步骤1-2,基于单位时间段内路口监控视频根据不同车辆ID提取出原始轨迹点,记为 $P(f, x, y)$ ,其中 $f$ 表示帧数, $x$ 和 $y$ 表示轨迹点坐标数值;

[0015] 步骤1-3,基于原始轨迹点 $P$ 在二维坐标内绘制出车辆原始轨迹点,从中确定缺失重要特征数据和不正确数据并进行剔除,实现原始轨迹点清洗;其中缺失重要特征数据为偏离数据,其可组成短路径但偏离线路走向,不正确数据即无法平滑连接的密集散点集,其无法构成路径轨迹。

[0016] 进一步地,所述步骤2中,具体包括如下分步骤:

[0017] 步骤2-1,基于同一车辆ID下的原始轨迹数据 $P$ 进行分析,在统计时间段 $T$ 内对单位时间提取定量数据实现轨迹特征点提取;

[0018] 步骤2-2,基于提取的特征点进行轨迹绘制,根据 $f$ 帧数大小排列,确定车辆轨迹,记为 $TR = \{P|p_i, 1 \leq i \leq n, n \text{ 为轨迹点数}\}$ 。

[0019] 进一步地,所述步骤3中,具体包括如下分步骤:

[0020] 步骤3-1,控制相似度阈值 $\gamma$ ,确定簇数 $K$ 的范围,基于层次聚类算法确定相似度阈值 $\gamma$ 下最佳聚类簇数 $K_{opi}$ 以及其聚类结果;

[0021] 步骤3-2,重复步骤3-1,调整相似度阈值 $\gamma$ ,重复求解出不同相似度阈值 $\gamma$ 下的最佳聚类簇数 $K_{opi}$ ,记为 $(\gamma_l, K_{opi}^l)$ ,其中 $l$ 表示相似度阈值序号,其中相似度阈值 $\gamma_l$ 数值取整数;同时各 $K_{opi}^l$ 下含有最优簇数的组内距均值 $\bar{\tau}_k^l$ 和组内距 $\phi_k^l$ ;

[0022] 步骤3-3,基于 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组内均方差 $S_r^l$ 和组间均方差 $S_\phi^l$ ,根据组内均方差 $S_r^l$ 和组间均方差 $S_\phi^l$ 确定最优阈值 $\gamma_{opi}$ 及其最优组数 $K_{opi}$ ;

[0023] 步骤3-4,基于最佳阈值 $\gamma_{opi}$ 和最优聚类簇数 $K_{opi}$ ,确定车辆轨迹的聚类结果,识别出车辆轨迹模式类型,确定正常车辆轨迹和异常车辆轨迹;

[0024] 具体来说,基于步骤3-3确定的最佳阈值 $\gamma_{opi}$ 和最优聚类簇数 $K_{opi}$ 进行车辆轨迹层次聚类,根据步骤1的路口渠化信息,对聚类的车辆轨迹模式进行分类,将根据路口渠化流向下划分以外的车辆轨迹默认为异常车辆轨迹,其余则为正常车辆轨迹。

[0025] 进一步地,所述步骤3-1中,包括如下分步骤:

[0026] 步骤3-1-1,控制相似度阈值 $\gamma$ ,通过LCSS算法确定两两轨迹之间的最长公共子序列和最长公共子序列相似度距离;

[0027] 步骤3-1-2,根据上一步骤的最长公共子序列相似度距离列出相似度矩阵 $S[a][b]$ ,即为邻近矩阵;

[0028] 步骤3-1-3,确定聚类簇数 $K$ 的范围;具体来说,根据步骤1的路口渠化信息确定聚类簇数 $K$ 的最小数值;

[0029] 步骤3-1-4,给定簇数 $K$ ,不断重复步骤3-1-1和3-1-2得到层次聚类的聚类结果;

[0030] 步骤3-1-5,根据步骤3-1-3给定的簇范围确定出控制相似度阈值 $\gamma$ 一定时 $K$ 值不同的聚类情况,根据不同 $K$ 值聚类下的组内距和组间距建立评价体系,从而确定最佳聚类簇数 $K_{opi}$ ,具体如下:

[0031] 根据步骤3-1-4的聚类结果提取出组内轨迹的所有特征点 $P_i$ ,根据特征点 $P_i$ 确定确

定划分K组情况下k个聚类簇中心  $\{C_1, C_2, C_3, \dots, C_k\}$ ,  $1 < k \leq K$ , 其中簇中心  $C_k$  为该组内所有特征点的中心;

[0032] 根据k组簇中心分别求得k组内的组内距  $\tau_k$  以及组间距  $\varphi_k$ ;

[0033] 基于不同组数K下的组内距均值  $\bar{\tau}_k$  和组间距  $\varphi_k$ , 确定最佳聚类簇数  $K_{opi}$ , 具体来说, 将比值  $\alpha = \frac{\varphi_k}{\bar{\tau}_k}$  最大时的k值默认为  $K_{opi}$ 。

[0034] 进一步地, 所述步骤3-3中, 包括如下分步骤:

[0035] 步骤3-3-1, 根据  $(\gamma_l, K_{opi}^l)$  求解出不同相似度阈值  $\gamma_l$  的组内均方差  $S_\tau^l$ ;

[0036] 步骤3-3-2, 根据  $(\gamma_l, K_{opi}^l)$  求解出不同相似度阈值  $\gamma_l$  的组间均方差  $S_\varphi^l$ ;

[0037] 步骤3-3-3, 基于不同的相似度阈值  $\gamma$  及其  $K_{opi}$  下的组内均方差  $S_\tau^l$  和组间均方差  $S_\varphi^l$ , 确定最优阈值  $\gamma_{opi}$ 。具体来说, 将比值  $\beta = \frac{S_\varphi^l}{S_\tau^l}$  最大时的  $\gamma_l$  默认为最优阈值。

[0038] 进一步地, 所述步骤4中, 具体包括如下分步骤:

[0039] 步骤4-1, 提取出单一模式下正常车辆轨迹, 基于正常车辆轨迹的轨迹特征点数值, 以轨迹的轨迹相似度  $\lambda$ 、加速度方差  $\alpha^2$ 、弧长比  $\sigma$  作为特征数值进行层次聚类, 划分出异常行驶轨迹、异常行为轨迹和正常车辆轨迹;

[0040] 步骤4-2, 整合上一步骤分析得到的正常车辆轨迹, 在DTW算法下对上一步骤的正常车辆轨迹再次进行层次聚类, 判别出正常车辆轨迹和离群轨迹;

[0041] 步骤4-3, 重复步骤4-1至4-2, 对路口各模式下车辆轨迹分析, 划分各流向中异常行驶轨迹、异常行为轨迹、离群轨迹和正常车辆轨迹, 其中将异常行驶轨迹、异常行为轨迹和离群轨迹均默认为异常车辆轨迹。

[0042] 进一步地, 所述步骤4-1中, 具体包括如下分步骤:

[0043] 步骤4-1-1, 对车辆轨迹  $TR_j$  的轨迹相似度  $\lambda$ 、加速度方差  $\alpha^2$ 、弧长比  $\sigma$  进行求解;

[0044] 根据车辆轨迹特征确定相似度计算公式, 如LCSS算法或DTW算法, 求得车辆轨迹相似度  $\lambda$ ;

[0045] 基于车辆轨迹  $TR$  的轨迹特征点确定各车辆轨迹的加速度方差  $\alpha^2$ , 即:

$$[0046] \quad \alpha_{i+1} = \frac{\sqrt{\|p_{i+1} - p_i\|^2}}{(f_{i+1} - f_i)^2}$$

$$[0047] \quad \bar{\alpha} = \frac{\alpha_{i+1} + \dots + \alpha_n}{n-1}$$

$$[0048] \quad \alpha^2 = \frac{\sum (\alpha_{i+1} - \bar{\alpha})^2}{n-1}$$

[0049] 式中,  $\alpha_{i+1}$  表示特征点  $i+1$  的加速度, 其中  $f_i$  和  $f_{i+1}$  表示帧数,  $\sqrt{\|p_{i+1} - p_i\|^2}$  表示特征点  $i$  与特征点  $i+1$  之间的欧式距离,  $p_i$  和  $p_{i+1}$  表示特征点;

[0050] 基于车辆轨迹  $TR$  确定车辆轨迹弧长比  $\sigma$ , 即:



$$[0051] \quad \sigma = \frac{\sum_{i=1}^{n-1} \sqrt{\|p_{i+1} - p_i\|^2}}{\sqrt{\|p_n - p_1\|^2}}$$

[0052] 式中,  $p_1, p_i, p_{i+1}, p_n$  均表示车辆轨迹内的特征点;  $\sqrt{\|p_{i+1} - p_i\|^2}$  表示车辆轨迹特征点  $i$  和特征点  $i+1$  之间的欧式距离;  $\sqrt{\|p_n - p_1\|^2}$  表示车辆轨迹特征点  $n$  与特征点  $1$  之间的欧式距离;

[0053] 步骤4-1-2, 基于上一步骤车辆轨迹  $TR$  的相似度  $\lambda$ 、加速度方差  $\alpha^2$ 、弧长比  $\sigma$  数值作为特征数据进行层次聚类, 划分为异常行驶轨迹、异常行为轨迹和正常车辆轨迹; 具体来说, 以相似度  $\lambda$ 、加速度方差  $\alpha^2$ 、弧长比  $\sigma$  数值作为特征数据进行层次聚类, 将车辆轨迹划分为三组数据, 根据三组数据的数据量及各组数据离散程度确定数据类型;

[0054] 计算三组车辆轨迹的轨迹数量及其与总轨迹数量的比值, 若组内轨迹数量最少则那组车辆轨迹数据默认为异常行驶轨迹;

[0055] 基于车辆轨迹的特征点分别求剩余两组车辆轨迹数据的离散程度  $\varepsilon$ , 即:

$$[0056] \quad \varepsilon = \frac{\sum dist(p_i - \bar{p})^2}{N}$$

$$[0057] \quad dist(p_i - \bar{p}) = \sqrt{\|p_i - \bar{p}\|^2}$$

$$[0058] \quad \bar{p} = \frac{\sum p_i}{N}$$

[0059] 式中,  $N$  为组内所有车辆轨迹特征点的总个数;  $p_i$  为特征点,  $\bar{p}$  为簇中心点;  $dist(p_i - \bar{p})$  表示  $p_i$  点到簇中心点距离;

[0060] 将离散程度  $\varepsilon$  较大的那组数据默认为异常行为轨迹, 较小的那组数据默认为正常车辆轨迹, 同时实现正常车辆轨迹提取。

[0061] 进一步地, 所述步骤5中, 具体包括如下分步骤:

[0062] 步骤5-1, 整合步骤3和4中的异常车辆轨迹, 对异常车辆轨迹发生原因进行分析, 同时将异常车辆轨迹与异常原因相对应, 建立异常轨迹关联数据库;

[0063] 步骤5-2, 基于统计时间段内异常轨迹数目比例及其行为原因对路口渠化组织问题和交通信号控制问题进行分析, 识别出不合理的路口渠化和不合理的路口信号方案。

[0064] 进一步地, 所述步骤5-2中, 具体如下:

[0065] 若统计时间段内异常轨迹数目大于路口异常状况阈值, 则分析其渠化是否存在问题;

[0066] 若同一流向下, 统计时间段放行相位阶段内, 异常轨迹数目大于异常状况阈值, 则分析该相位阶段信号方案配置是否合理。

[0067] 本发明达到的有益效果为:

[0068] 1. 目前的LCSS车辆轨迹研究依托GPS数据和路网信息, 将GPS点位信息与路网地图匹配实现数据初步处理和清洗, 通过GPS数据特征点实现聚类, 将LCSS相似度长度距离替换欧式距离实现层次聚类, 但无法进行聚类效果的评判, 本发明运用了视频号牌数据代替GPS

数据,利用LCSS算法对车辆轨迹相似程度分析,利用路口渠化特征为无监督学习打上标签,不断迭代确定最优簇数和最优阈值,从而提高了车辆轨迹层次聚类效果。

[0069] 2.对比传统的路网车辆轨迹研究(对所有车辆轨迹一体化聚类分析),本发明以单个路口和单个流向车辆轨迹进行分析,通过LCSS层次聚类算法对路口车辆轨迹分类,有效识别出路口车辆轨迹类型和异常轨迹,进一步对各模式类型下车辆轨迹分析,对各流向车辆异常轨迹识别提取,从而为路口交通冲突安全管理和拥堵管理提供有效支撑。

[0070] 3.本发明创新的建立异常轨迹数据库(一一对应的异常车辆轨迹类型和轨迹异常原因),根据路口渠化信息和交通信号方案对异常车辆轨迹进行分析,通过对统计时间段内异常轨迹分析,有效识别出路口渠化组织的问题和信号方案问题,对路口渠化优化和交通信号配时方案的优化调整提供支撑依据。

## 附图说明

[0071] 图1为本发明所述车辆轨迹识别分析方法的步骤流程图。

[0072] 图2为本发明实施例中绘制出车辆原始轨迹点的示意图。

[0073] 图3为本发明实施例中筛选后绘制出的车辆轨迹示意图。

[0074] 图4为本发明实施例中根据聚类信息用颜色标注后的车辆轨迹示意图。

[0075] 图5为本发明实施例中提取出的异常车辆轨迹示意图。

[0076] 图6为本发明实施例中对南进口车道左转车辆轨迹求解出各车辆轨迹的相似度、弧长比和加速度方差列表。

[0077] 图7为本发明实施例中通过层次聚类将车辆轨迹划分为正常轨迹、异常行为和异常轨迹的三种车辆轨迹示意图。

[0078] 图8为本发明实施例中划分后的正常轨迹和离群轨迹示意图。

[0079] 图9为本发明实施例中所有车辆轨迹的示意图。

## 具体实施方式

[0080] 下面结合说明书附图对本发明的技术方案做进一步的详细说明。

[0081] 基于层次聚类的路口异常车辆轨迹识别分析方法,可以对视频结构化数据分析,实现路口轨迹类型识别和分类,有效划分出异常车辆轨迹,为异常轨迹建立轨迹数据库,辅助实现交通渠化和信控方案评估分析。具体包括如下步骤:

[0082] 步骤1,采集路口渠化信息,提取路口监控视频数据完成数据清洗。

[0083] 步骤1-1,对路口类型及其渠化信息进行采集。如十字路口,则对东、南、西、北四个进口道的车道数目、宽度以及非机动车车道宽度进行渠化信息采集;如天桥路口,则对岔路口车道信息采集。

[0084] 步骤1-2,基于单位时间段内路口监控视频根据不同车辆ID提取出原始轨迹点,记为 $P(f, x, y)$ ,其中 $f$ 表示帧数, $x$ 和 $y$ 表示轨迹点坐标数值。一般情况下,监控视频源自电子警察、智能卡口、球机监控等,以角度高视野好的球机视频为主。

[0085] 步骤1-3,基于原始轨迹点 $P$ 在二维坐标内绘制出车辆原始轨迹点,从中确定缺失重要特征数据和不正确数据并进行剔除,实现原始轨迹点清洗。其中缺失重要特征数据为偏离数据,其可组成短路径但偏离线路走向,不正确数据即无法平滑连接的密集散点集,其

无法构成路径轨迹。

[0086] 步骤2,针对不同车辆ID实现单条轨迹分析,提取单条轨迹的数据特征点,确定车辆轨迹。

[0087] 步骤2-1,基于同一车辆ID下的原始轨迹数据P进行分析,在统计时间段T内分别从各单位时间t内选取n个数据,实现轨迹特征点提取。一般情况下,单位时间t选取5-6秒,n选取25-35个,同时为方便聚类分析,提高聚类效果,每条车辆轨迹的总特征点可均取一个数值。

[0088] 步骤2-2,基于提取的特征点P进行轨迹绘制,根据f帧数大小排列,确定车辆轨迹,记为 $TR = \{P | p_i, 1 \leq i \leq n, n \text{ 为轨迹点数} \}$ 。

[0089] 步骤3,基于LCSS算法和层次聚类算法实现路口车辆轨迹聚类,划分出车辆轨迹模式类型,识别出正常车辆轨迹和异常车辆轨迹。

[0090] 步骤3-1,控制相似度阈值 $\gamma$ ,确定簇数K的范围,基于层次聚类算法确定相似度阈值 $\gamma$ 下最佳聚类簇数 $K_{opi}$ 以及其聚类结果。

[0091] 步骤3-1-1,控制相似度阈值 $\gamma$ ,通过LCSS算法确定两两轨迹之间的最长公共子序列和最长公共子序列相似度距离,即:

$$[0092] \quad LCSS(TR_1, TR_2) = \begin{cases} 0, & \text{if } TR_1 = \emptyset \text{ or } TR_2 = \emptyset \\ 1 + LCSS(TR_1^{i-1}, TR_2^{j-1}), & \text{if } dist(p_i, q_j) < \gamma \\ \max(LCSS(TR_1^{i-1}, TR_2^j), LCSS(TR_1^i, TR_2^{j-1})), & \text{otherwise} \end{cases}$$

$$[0093] \quad dist(p_i, q_j) = \sqrt{\|p_i - q_j\|^2}$$

$$[0094] \quad D_{LCSS}(TR_1, TR_2) = 1 - \frac{LCSS(TR_1, TR_2)}{\min(len_{TR_1}, len_{TR_2})}$$

[0095] 式中 $TR_1$ 和 $TR_2$ 分别为两条长度为m、n的轨迹; $TR_1 = [p_1, p_2, \dots, p_{m-1}, p_m]$ ,  $TR_2 = [q_1, q_2, \dots, q_{n-1}, q_n]$ , 其中 $p_i$ 和 $q_j$

分别代表轨迹特征点坐标,  $TR_1^i = [p_1, p_2, \dots, p_{i-1}, p_i]$ ,  $TR_1^{i-1} = [p_1, p_2, \dots, p_{i-1}]$ , 同理

$TR_2^j = [q_1, q_2, \dots, q_{j-1}, q_j]$ ,  $TR_2^{j-1} = [q_1, q_2, \dots, q_{j-1}]$ , 其中 $1 \leq i \leq m, 1 \leq j \leq n$ ,  $\emptyset$ 表示轨迹为空;

$dist(p_i, q_j)$ 表示两个坐标点的欧氏距离, $\gamma$ 为相似度阈值; $LCSS(TR_1, TR_2)$ 表示 $TR_1$ 和 $TR_2$ 两条轨迹的最长公共子序列长度; $D_{LCSS}(TR_1, TR_2)$ 为轨迹 $TR_1$ 和 $TR_2$ 的相似度距离; $\min(len_{TR_1}, len_{TR_2})$ 表示轨迹 $TR_1$ 长度和 $TR_2$ 长度的较小值。

[0096] 步骤3-1-2,根据上一步骤的最长公共子序列相似度距离列出相似度矩阵S[a][b],即为邻近矩阵。

[0097] 具体来说,将传统层次聚类中的欧式距离替换成LCSS相似度距离,分别列出两两轨迹之间的相似度距离,其中以第i行为例,其行内数值为最长公共子序列的相似度距离,即为 $\{D_{LCSS}(TR_j, TR_1), D_{LCSS}(TR_j, TR_2), \dots, D_{LCSS}(TR_j, TR_N)\}$ ,式中 $1 \leq j \leq N$ ,表示第i条车辆轨迹。

[0098] 步骤3-1-3,确定聚类簇数K的范围(最小值)。

[0099] 具体来说,根据步骤1的路口渠化信息确定聚类簇数K的最小数值。如一个十字路

口,其存在东、南、西、北四个方向的左转、直行和右转,且路口允许掉头,因此簇数K从16开始取值。

[0100] 步骤3-1-4,给定簇数K,不断重复步骤3-1-1和3-1-2得到层次聚类的聚类结果。具体包括如下内容:

[0101] 1) 根据最长公共子序列相似度数不断合并最接近的两条轨迹,具体来说,取 $\min(S[a][b])$ 进行合并,其中将轨迹重合情况排除(其数值为零)。

[0102] 如 $TR_1$ 、 $TR_2$ 、 $TR_3$ 、 $TR_4$ 四条车辆轨迹,求解出其最长公共子序列相似度距离,邻近性矩阵 $S[a][b]$ 为:

$$[0103] \quad \begin{bmatrix} 0 & 1 & 1 & 0.6 \\ 1 & 0 & 0.8 & 1 \\ 1 & 0.8 & 0 & 1 \\ 0.6 & 1 & 1 & 0 \end{bmatrix}$$

[0104] 则可将 $TR_1$ 和 $TR_4$ 进行合并(其数值为0.6)作为新一个簇并进行 $S[a][b]$ 求解和聚类。

[0105] 2) 将合并的两条轨迹默认为新簇,重复步骤直到达到设定的K簇数为止。具体来说, $TR_j$ 和 $TR_{j-1}$ 根据上一轮聚类结果已合并,则进行第二轮层次聚类时,需要重新计算 $D_{LCSS}$ 数值,且列出新的 $S[a][b]$ 矩阵。

[0106] 步骤3-1-5,根据步骤3-1-3给定的簇范围确定出控制相似度阈值 $\gamma$ 一定时K值不同的聚类情况,根据不同K值聚类下的组内距和组间距建立评价体系,从而确定最佳聚类簇数 $K_{opi}$ 。具体包括如下内容:

[0107] 1) 根据聚类结果提取出组内轨迹的所有特征点 $P_i$ ,根据特征点 $P_i$ 确定划分K组情况下k个聚类簇中心 $\{C_1, C_2, C_3, \dots, C_k\}$ ,  $1 < k \leq K$ ,其中簇中心 $C_k$ 为该组内所有特征点的中心。

[0108] 2) 根据k组簇中心分别求得k组内的组内距 $\tau_k$ ,即:

$$[0109] \quad \tau_k = \frac{\sum_{i=1}^n dist(P_i, C_k)}{n}$$

$$[0110] \quad dist(P_i, C_k) = \sqrt{\sum_{i=1}^n (P_i - C_k)^2}$$

[0111] 式中: $n$ 为第k组内的样本数量,其中 $1 \leq k \leq K$ ;  $dist(P_i, C_k)$ 表示k组内样本点到簇中心点距离; $P_i$ 为轨迹i的样本信息,其中 $1 \leq i \leq n$ ;  $C_k$ 为k组的簇中心点。

[0112] 进一步,求得K组情况下的组内距均值,记为 $\bar{\tau}_k$ ,即:

$$[0113] \quad \bar{\tau}_k = \frac{\tau_k}{K}$$

[0114] 3) 基于簇中心点求解出组间距 $\varphi_k$

$$[0115] \quad \varphi_k = \frac{\sum_{k=1}^K \sqrt{(C_k - C_{k'})^2}}{K}$$

[0116] 式中： $K$ 为簇中心个数， $C_k$ 和 $C_{k'}$ 为簇中心点坐标。

[0117] 4) 基于不同组数 $K$ 下的组内距均值 $\bar{\tau}_k$ 和组间距 $\phi_k$ ，确定最佳聚类簇数 $K_{opi}$ 。具体来说，将比值 $\alpha = \frac{\phi_k}{\bar{\tau}_k}$ 最大时的 $k$ 值默认为 $K_{opi}$ 。

[0118] 步骤3-2，重复步骤3-1，调整相似度阈值 $\gamma$ ，重复求解出不同相似度阈值 $\gamma$ 下的最佳聚类簇数 $K_{opi}$ ，记为 $(\gamma_l, K_{opi}^l)$ ，其中 $\gamma_l$ 表示相似度阈值，该数值取整数；同时各 $K_{opi}^l$ 下含有最优簇数的组内距均值 $\bar{\tau}_k^l$ 和组内距 $\phi_k^l$ 。

[0119] 步骤3-3，基于 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组内均方差 $S_\tau^l$ 和组间均方差 $S_\phi^l$ ，确定最优阈值 $\gamma_{opi}$ 及其最优组数 $K_{opi}$ 。

[0120] 步骤3-3-1，根据 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组内均方差 $S_\tau^l$ ，即：

$$[0121] \quad S_\tau^l = \sqrt{\sum_k^{K_{opi}^l} (\tau_k^l - \bar{\tau}_k^l)^2}$$

[0122] 式中： $\tau_k^l$ 和 $\bar{\tau}_k^l$ 为相似度阈值为 $\gamma_l$ 时，最佳聚类簇数 $K_{opi}$ 下各 $k$ 组的组内距和组内均值，其求解公式如S45，其中 $k$ 为簇数，其 $1 < k \leq K_{opi}$ ， $l$ 为相似度阈值序号。

[0123] 步骤3-3-2，根据 $(\gamma_l, K_{opi}^l)$ 求解出不同相似度阈值 $\gamma_l$ 的组间均方差 $S_\phi^l$ ，即：

$$[0124] \quad S_\phi^l = \sqrt{\sum_k^{K_{opi}^l} (dist(C_k^l) - \phi_k^l)^2}$$

$$[0125] \quad dist(C_k^l) = \sqrt{(C_k^l - C_{k'}^l)^2}$$

[0126] 式中： $dist(C_k^l)$ 表示为相似度阈值为 $\gamma_l$ 时，最佳聚类簇数 $K_{opi}$ 时，各 $k$ 组簇中心之间的距离， $\phi_k^l$ 表示为第 $l$ 相似度阈值时，最佳聚类簇数 $K_{opi}$ 时组内距数值；其中 $k$ 为簇数，其 $1 < k \leq K_{opi}$ ， $l$ 为相似度阈值。

[0127] 步骤3-3-3，基于不同的相似度阈值 $\gamma$ 及其 $K_{opi}$ 下的组内均方差 $S_\tau^l$ 和组间均方差 $S_\phi^l$ ，确定最优阈值 $\gamma_{opi}$ 。具体来说，将比值 $\beta = \frac{S_\phi^l}{S_\tau^l}$ 最大时的 $l$ 值默认为最优阈值。

[0128] 步骤3-4，基于最佳阈值 $\gamma_{opi}$ 和最优聚类簇数 $K_{opi}$ ，确定车辆轨迹的聚类结果，识别出车辆轨迹模式类型，确定正常车辆轨迹和异常车辆轨迹。

[0129] 具体来说，基于步骤3-3确定的最佳阈值 $\gamma_{opi}$ 和最优聚类簇数 $K_{opi}$ 进行车辆轨迹聚类，根据步骤1的路口渠化信息，对聚类的车辆轨迹模式进行分类，将根据路口渠化流向划分以外的车辆轨迹默认为异常车辆轨迹，其余则为正常车辆轨迹，一般情况下，路口的异常车辆轨迹包含非机动车行驶轨迹、车辆逆行行驶轨迹、车辆暂停等异常情况。

[0130] 步骤4，基于步骤3模式分析的车辆轨迹类型提取出单一模式（流向）正常车辆轨迹，进一步对单一模式（流量）车辆轨迹分析，识别出异常轨迹和正常车辆轨迹。

[0131] 步骤4-1，提取出单一模式（流向）正常车辆轨迹，基于车辆轨迹TR特征点数值，以轨迹的轨迹相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 作为特征数值进行层次聚类，划分出异常行驶

轨迹、异常行为轨迹和正常车辆轨迹。

[0132] 步骤4-1-1,对车辆轨迹 $TR_j$ 的轨迹相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 进行求解。具体包括如下步骤:

[0133] 1) 根据车辆轨迹特征确定相似度计算公式,如LCSS算法或DTW算法,进一步求得车辆轨迹相似度 $\lambda$ 。

[0134] 若采用LCSS算法,则基于LCSS算法求得车辆轨迹相似度 $\lambda$ 的计算公式即:

$$[0135] \quad LCSS(TR_1, TR_2) = \begin{cases} 0, & \text{if } TR_1 = \emptyset \text{ or } TR_2 = \emptyset \\ 1 + LCSS(TR_1^{i-1}, TR_2^{j-1}), & \text{if } dist(p_i, q_j) < \gamma \\ \max(LCSS(TR_1^{i-1}, TR_2^j), LCSS(TR_1^i, TR_2^{j-1})), & \text{otherwise} \end{cases}$$

$$[0136] \quad dist(p_i, q_j) = \sqrt{\|p_i - q_j\|^2}$$

$$[0137] \quad D_{LCSS}(TR_1, TR_2) = 1 - \frac{LCSS(TR_1, TR_2)}{\min(len_{TR_1}, len_{TR_2})}$$

$$[0138] \quad \lambda(TR_1) = \frac{\sum (D_{LCSS}(TR_1, TR_s))}{N}$$

[0139] 式中 $TR_1$ 和 $TR_2$ 分别为两条长度为 $m$ 、 $n$ 的轨迹;  $TR_1 = [p_1, p_2, \dots, p_{m-1}, p_m]$ ,  $TR_2 = [q_1, q_2, \dots, q_{n-1}, q_n]$ , 其中 $p_i$ 和

$q_j$ 分别代表轨迹特征点坐标,  $TR_1^i = [p_1, p_2, \dots, p_{i-1}, p_i]$ ,  $TR_1^{i-1} = [p_1, p_2, \dots, p_{i-1}]$ , 同理

$TR_2^j = [q_1, q_2, \dots, q_{j-1}, q_j]$ ,  $TR_2^{j-1} = [q_1, q_2, \dots, q_{j-1}]$ , 其中 $1 \leq i \leq m$ ,  $1 \leq j \leq n$ ,  $\emptyset$ 表示轨迹为空;

$dist(p_i, q_j)$ 表示两个坐标点的欧氏距离,  $\gamma$ 为相似度阈值;  $LCSS(TR_1, TR_2)$ 表示 $TR_1$ 和 $TR_2$ 两条轨迹的最长公共子序列长度;  $D_{LCSS}(TR_1, TR_2)$ 为轨迹 $TR_1$ 和 $TR_2$ 的相似度距离;  $\min(len_{TR_1}, len_{TR_2})$ 表示轨迹 $TR_1$ 长度和 $TR_2$ 长度的较小值;  $D_{LCSS}(TR_1, TR_s)$ 表示 $TR_1$ 和 $TR_s$ 之间的最长公共子序列相似度距离, 其中 $1 \leq s \leq N$ ,  $N$ 表示车辆轨迹数目;  $\lambda(TR_1)$ 表示 $TR_1$ 车辆轨迹的相似度值。

[0140] 若采用DTW相似度,则基于DTW算法求得车辆轨迹相似度 $\lambda$ 的计算公式为:

[0141]  $DTW(TR_1, TR_2) = f(m, n)$

$$[0142] \quad f(i, j) = \|p_i - q_j\| + \min \begin{cases} f(i, j-1) \\ f(i-1, j) \\ f(i-1, j-1) \end{cases}$$

$$[0143] \quad \lambda(TR_1) = \frac{\sum (DTW(TR_1, TR_s))}{N}$$

[0144] 式中 $TR_1$ 和 $TR_2$ 分别为两条长度为 $m$ 、 $n$ 的轨迹;  $TR_1 = [p_1, p_2, \dots, p_{m-1}, p_m]$

[0145]  $TR_2 = [q_1, q_2, \dots, q_{n-1}, q_n]$ , 其中 $p_i$ 和 $q_j$ 分别代表轨迹特征点坐标,

$TR_1^i = [p_1, p_2, \dots, p_{i-1}, p_i]$ ,  $TR_1^{i-1} = [p_1, p_2, \dots, p_{i-1}]$ , 同理  $TR_2^j = [q_1, q_2, \dots, q_{j-1}, q_j]$ ,

$TR_2^{j-1} = [q_1, q_2, \dots, q_{j-1}]$ , 其中 $1 \leq i \leq m$ ,  $1 \leq j \leq n$ ;  $\|p_i - q_j\|$ 表示两个坐标点的欧式距离;  $\lambda$

(TR<sub>s</sub>)表示TR<sub>s</sub>车辆轨迹的相似度值,其中 $1 \leq s \leq N$ ,N代表车辆轨迹数目。

[0146] 一般情况下,LCSS算法和DTW算法的选择由路口特征和轨迹数据量确定。

[0147] 2) 基于车辆轨迹TR的轨迹特征点确定各车辆轨迹的加速度方差 $\alpha^2$ ,即:

$$[0148] \quad \alpha_{i+1} = \frac{\sqrt{\|p_{i+1} - p_i\|^2}}{(f_{i+1} - f_i)^2}$$

$$[0149] \quad \bar{\alpha} = \frac{\alpha_{i+1} + \dots + \alpha_n}{n-1}$$

$$[0150] \quad \alpha^2 = \frac{\sum (\alpha_{i+1} - \bar{\alpha})^2}{n-1}$$

[0151] 式中, $\alpha_{i+1}$ 表示特征点i+1的加速度,其中 $f_i$ 和 $f_{i+1}$ 表示帧数, $\sqrt{\|p_{i+1} - p_i\|^2}$ 表示特征点i与特征点i+1之间的欧式距离, $p_i$ 和 $p_{i+1}$ 表示特征点。

[0152] 3) 基于车辆轨迹TR确定车辆轨迹弧长比 $\sigma$ ,即

$$[0153] \quad \sigma = \frac{\sum_{i=1}^{n-1} \sqrt{\|p_{i+1} - p_i\|^2}}{\sqrt{\|p_n - p_1\|^2}}$$

[0154] 式中, $p_1$ 、 $p_i$ 、 $p_{i+1}$ 、 $p_n$ 均表示车辆轨迹内的特征点; $\sqrt{\|p_{i+1} - p_i\|^2}$ 表示车辆轨迹特征点i和特征点i+1之间的欧式距离; $\sqrt{\|p_n - p_1\|^2}$ 表示车辆轨迹特征点n与特征点1之间的欧式距离。

[0155] 步骤4-1-2,基于上一步骤车辆轨迹TR的相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 数值作为特征数据进行层次聚类(聚类的簇组数K为3),划分为异常车辆轨迹、异常行为轨迹和正常车辆轨迹。

[0156] 具体来说,以相似度 $\lambda$ 、加速度方差 $\alpha^2$ 、弧长比 $\sigma$ 数值作为特征数据进行层次聚类,将车辆轨迹划分为三组数据,根据三组数据的数据量及各组数据离散程度确定数据类型。具体步骤如下:

[0157] 1) 计算三组车辆轨迹的轨迹数量及其与总轨迹数量的比值,若组内轨迹数量最少则那组车辆轨迹数据默认为异常行驶轨迹,转为下一步骤。

[0158] 2) 基于车辆轨迹的特征点分别求剩余两组车辆轨迹数据的离散程度 $\varepsilon$ ,即

$$[0159] \quad \varepsilon = \frac{\sum dist(p_i - \bar{p})^2}{N}$$

$$[0160] \quad dist(p_i - \bar{p}) = \sqrt{\|p_i - \bar{p}\|^2}$$

$$[0161] \quad \bar{p} = \frac{\sum p_i}{N}$$

[0162] 式中,N为组内所有车辆轨迹特征点的总个数; $p_i$ 为特征点, $\bar{p}$ 为簇中心点; $dist(p_i - \bar{p})$ 表示 $p_i$ 点到簇中心点距离。

[0163] 将离散程度 $\varepsilon$ 较大的那组数据默认为异常行为轨迹,较小的那组数据默认为正常

车辆轨迹,同时实现正常车辆轨迹提取。

[0164] 步骤4-2,整合上一步骤的正常车辆轨迹,在DTW算法下对车辆轨迹再次进行层次聚类(聚类的簇组数为2),判别出正常车辆轨迹和离群轨迹。具体步骤如下:

[0165] 1) 提取正常车辆轨迹,基于DTW算法确定两两轨迹之间的相似度,即:

[0166]  $DTW(TR_1, TR_2) = f(m, n)$

$$[0167] \quad f(i, j) = \|p_i - q_j\| + \min \begin{cases} f(i, j-1) \\ f(i-1, j) \\ f(i-1, j-1) \end{cases}$$

[0168] 式中TR1和TR2分别为两条长度为m、n的轨迹;  $TR_1 = [p_1, p_2, \dots, p_{m-1}, p_m]$

[0169]  $TR_2 = [q_1, q_2, \dots, q_{n-1}, q_n]$ , 其中 $p_i$ 和 $q_j$ 分别代表轨迹特征点坐标,

$TR_1^i = [p_1, p_2, \dots, p_{i-1}, p_i]$ ,  $TR_1^{i-1} = [p_1, p_2, \dots, p_{i-1}]$ , 同理  $TR_2^j = [q_1, q_2, \dots, q_{j-1}, q_j]$ ,

$TR_2^{j-1} = [q_1, q_2, \dots, q_{j-1}]$ , 其中  $1 \leq i \leq m, 1 \leq j \leq n$ ;  $\|p_i - q_j\|$  表示两个坐标点的欧式距离。

[0170] 进一步确定两两轨迹之间的相似度矩阵 $S[a][b]$ , 以第i行为例, 其行内数值为DTW轨迹相似度, 即  $\{D_{LCSS}(TR_j, TR_1), D_{LCSS}(TR_j, TR_2), \dots, D_{LCSS}(TR_j, TR_N)\}$ , 式中  $1 \leq j \leq N$ , 表示第j条轨迹; N表示提取的正常车辆轨迹的轨迹总数。

[0171] 2) 利用相似度举证 $S[a][b]$ 内各相似度数值, 通过层次聚类算法将正常车辆轨迹划分为两类。

[0172] 3) 对于两类车辆轨迹分析, 若某一组车辆轨迹数目与车辆轨迹总数目的比值小于异常阈值(一般取30%), 则将该组合数据默认为离群轨迹, 另一组合数据为正常车辆轨迹, 否则转到下一步骤。

[0173] 4) 基于各车辆轨迹的特征点求解出两个组合的车辆轨迹离散程度 $\varepsilon$ , 即:

$$[0174] \quad \varepsilon = \frac{\sum dist(p_i - \bar{p})^2}{N}$$

$$[0175] \quad dist(p_i - \bar{p}) = \sqrt{\|p_i - \bar{p}\|^2}$$

$$[0176] \quad \bar{p} = \frac{\sum p_i}{N}$$

[0177] 式中, N为组内所有车辆轨迹特征点的总个数;  $p_i$ 为特征点,  $\bar{p}$ 为簇中心点;

$dist(p_i - \bar{p})$ 表示特征点 $p_i$ 到簇中心点距离。将离散程度 $\varepsilon$ 大的那组数据默认为离群轨迹, 另一组为正常车辆轨迹。

[0178] 步骤4-3, 重复步骤4-1至4-2, 对路口各模式(流向)下车辆轨迹分析, 划分各流向中异常行驶轨迹、异常行为轨迹、离群轨迹和正常车辆轨迹, 其中将异常行驶轨迹、异常行为轨迹和离群轨迹均默认为异常车辆轨迹。

[0179] 步骤5, 基于步骤3和4中的车辆轨迹识别信息中的异常车辆轨迹, 对路口状况进行分析。

[0180] 步骤5-1, 整合步骤3和4中的异常车辆轨迹, 对异常轨迹发生原因进行分析(如违法车辆行为、躲避行人异常行为、车辆冲突行为), 同时将异常车辆轨迹与异常原因相对应,



建立异常轨迹关联数据库。

[0181] 步骤5-2,基于统计时间段内异常轨迹数目比例及其行为原因对路口渠化组织问题和交通信号控制问题进行分析,识别出不合理路口渠化和不合理路口信号方案。

[0182] 1)若统计时间段内某原因异常轨迹数目大于路口异常状况阈值(一般取30%-40%,根据路口渠化类型确定),则根据其原因分析路口渠化问题是否存在问题。如统计时间段(7天)内某路口均存在车辆躲避行人异常行为,则判定非机动车道与机动车道之间渠化是否合理。

[0183] 2)若同一流向下,统计时间段放行相位阶段内,异常轨迹数目大于路口异常状况阈值(一般取30%-40%之间),则根据原因分析路口信号方案配置是否合理,如某左转相位阶段均在较多异常轨迹数目,则左转相位相序或绿灯时长设置是否合理。

[0184] 以下通过具体示例对本发明提出的车辆轨迹识别分析方法进行说明。

[0185] 示例1为十字路口,选取某十字路口5分钟内球机监控视频,从中提取出车辆ID、帧数、X坐标、Y坐标数值,基于X/Y坐标绘制出车辆原始轨迹点,从中剔除不正确数据和缺失重要特征数据,如图2所示。

[0186] 通过步骤3基于车辆ID对其特征点进行筛选,如提取出ID下的原始数据点,在其6s单位时间中随机筛选出30条数据,绘制出车辆轨迹,具体如图3所示。

[0187] 进而从中将等待红灯等状况(数据点未移动)数据剔除,从5min总时长内筛选出100个特征点,同理对所有车辆ID下的轨迹进行特征点提取(共305条车辆轨迹)。

[0188] 基于LCSS算法求解出305条轨迹之间的相似度关系,其中相似度阈值 $\gamma$ 的初始值为30,列出相似度矩阵 $S[a][b]$ ,将 $TR_1$ 、 $TR_2$ 、 $TR_3$ 、 $TR_4$ 四条数据单独拎出来,其相似度矩阵为:

$$[0189] \begin{bmatrix} 0 & 1 & 1 & 0.6 \\ 1 & 0 & 0.8 & 1 \\ 1 & 0.8 & 0 & 1 \\ 0.6 & 1 & 1 & 0 \end{bmatrix}$$

[0190] 即轨迹 $TR_4$ 与轨迹 $TR_1$ 的相似度为0.6。

[0191] 进一步不断叠加计算确定k组情况下的聚类效果,从而对比各k组情况下聚类的特征点组间距数值和组内距均值,将比值最大的k值设定为最优簇数,即确定 $\gamma = 30$ 情况下的簇数 $K = 44$ 。进一步调整相似度阈值 $\gamma$ 得到不同相似度阈值及其簇数。

[0192] 最终通过组间距均方差和组内距均方差确定最优阈值 $\gamma = 26$ ,其簇数 $K = 42$ 。在特征点轨迹轮廓基础上,根据聚类信息用颜色标注,具体包括左转(东左转、西左转、北左转、南左转)、直行(东直行、西直行、北直行、南直行)、右转(东右转、西右转、北右转、南右转)、掉头(南掉头、北掉头)和其余异常轨迹,具体如图4所示。

[0193] 路口分类情况如图5所示,从中提取出异常车辆轨迹。

[0194] 进一步,对南进口车道左转车辆轨迹进行提取分析,基于LCSS算法求解出各车辆轨迹的相似度,同时求解出各轨迹的弧长比和加速度方差,数据列表图6所示。

[0195] 通过层次聚类将车辆轨迹划分为正常车辆轨迹、异常行为轨迹和异常行驶轨迹,具体分类如图7所示。

[0196] 进一步对其中的正常车辆轨迹提取分析,利用DTW算法求出轨迹之间的相似度矩阵,具体如下:

$$[0197] \quad \begin{bmatrix} 0 & 17.08 & 85 & 219.64 \\ 79.38 & 0 & 47.52 & 181.02 \\ 149.25 & 94.72 & 0 & 110.22 \\ 263.79 & 209.25 & 140.68 & 0 \end{bmatrix}$$

[0198] 从而划分出正常车辆轨迹和离群轨迹,具体如图8所示。

[0199] 对各流向下车辆轨迹分析,整合所有异常车辆轨迹,其数值未超过路口异常状况阈值(异常车辆数目小于轨迹总数的35%)对照监控视频,发现异常轨迹原因,其主要为车辆违法行为(绕路行驶)。

[0200] 示例2为天桥路口,选取某天桥路口5分钟内球机监控视频,从中提取出车辆ID、帧数、X坐标、Y坐标数值,基于X/Y坐标绘制出车辆原始轨迹点,从中剔除不正确数据和缺失重要特征数据,实现原始轨迹清洗和轨迹特征点提取。

[0201] 进一步通过LCSS算法和层次聚类不断迭代找到最优相似度阈值 $\gamma$ 及其簇数 $K$ ( $\gamma=30, K=10$ ),绘制出车辆轨迹信息,确定车辆轨迹类型,具体类别包括左转(入和出)、直行(南向北、北向南)、右转(入和出)及其他异常轨迹,进而分析出异常车辆。所有车辆轨迹如图9所示。

[0202] 进一步对某单一模式下车辆轨迹分析,最终将所有非正常轨迹整合,为交通管理提供支撑依据。

[0203] 以上所述仅为本发明的较佳实施方式,本发明的保护范围并不以上述实施方式为限,但凡本领域普通技术人员根据本发明所揭示内容所作的等效修饰或变化,皆应纳入权利要求书中记载的保护范围内。

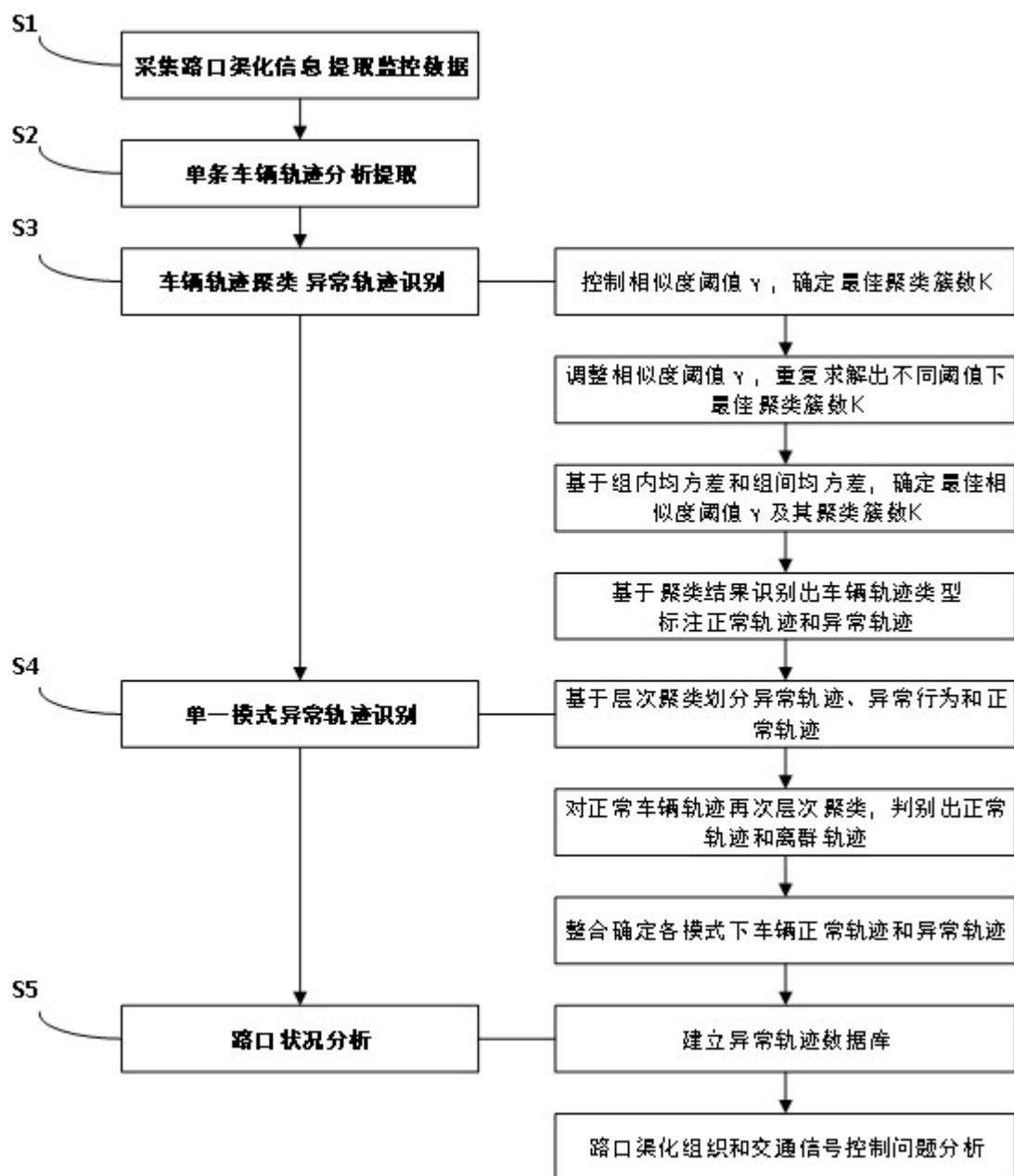


图1

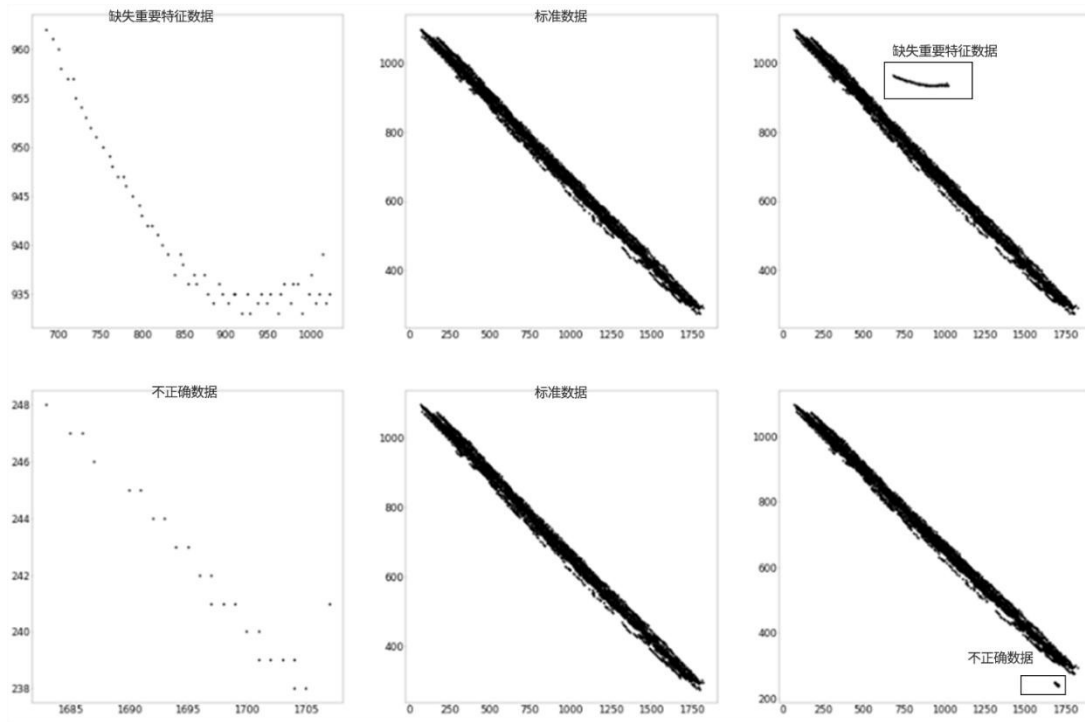


图2

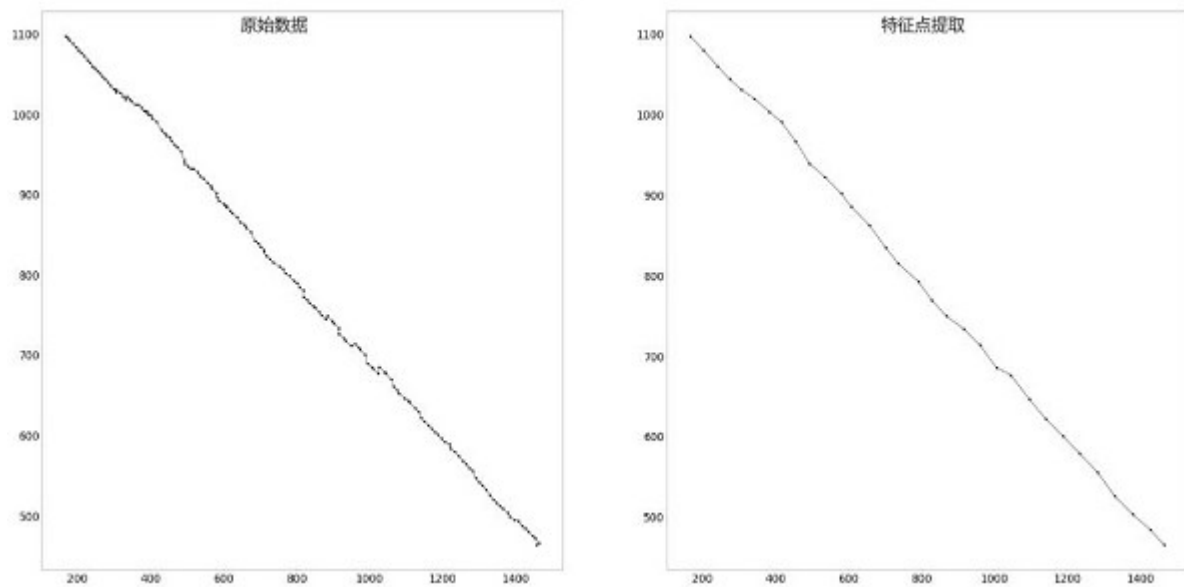


图3



图4

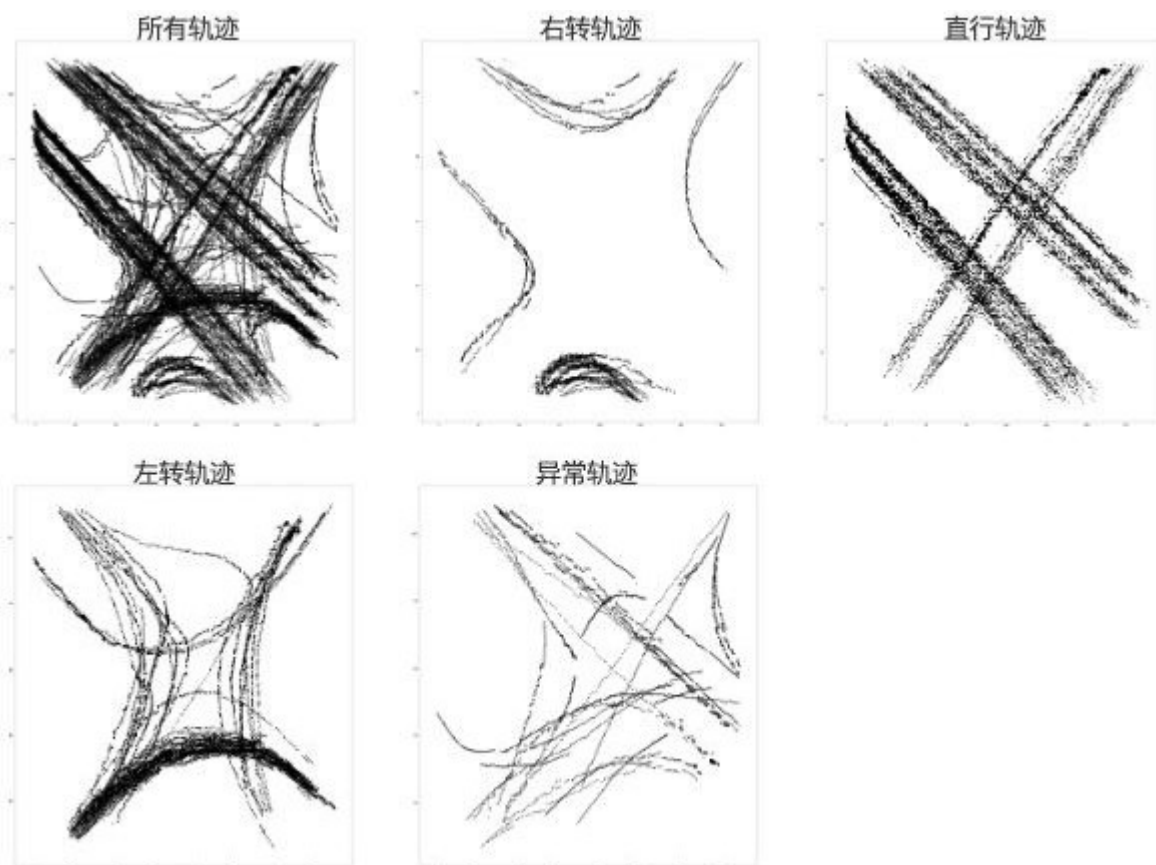


图5

序号	相似度	加速度方差	弧长比
1	0.5	0.068	1.001
2	0.7	0.031	1.003
3	1	0.037	1.034
4	0.5	0.032	1.001
5	0.5	0.027	1.002
6	0.7	0.076	1.000
7	0.6	0.031	1.009
8	0.7	0.089	1.001
9	0.8	0.053	1.000
10	0.8	0.014	1.016

图6

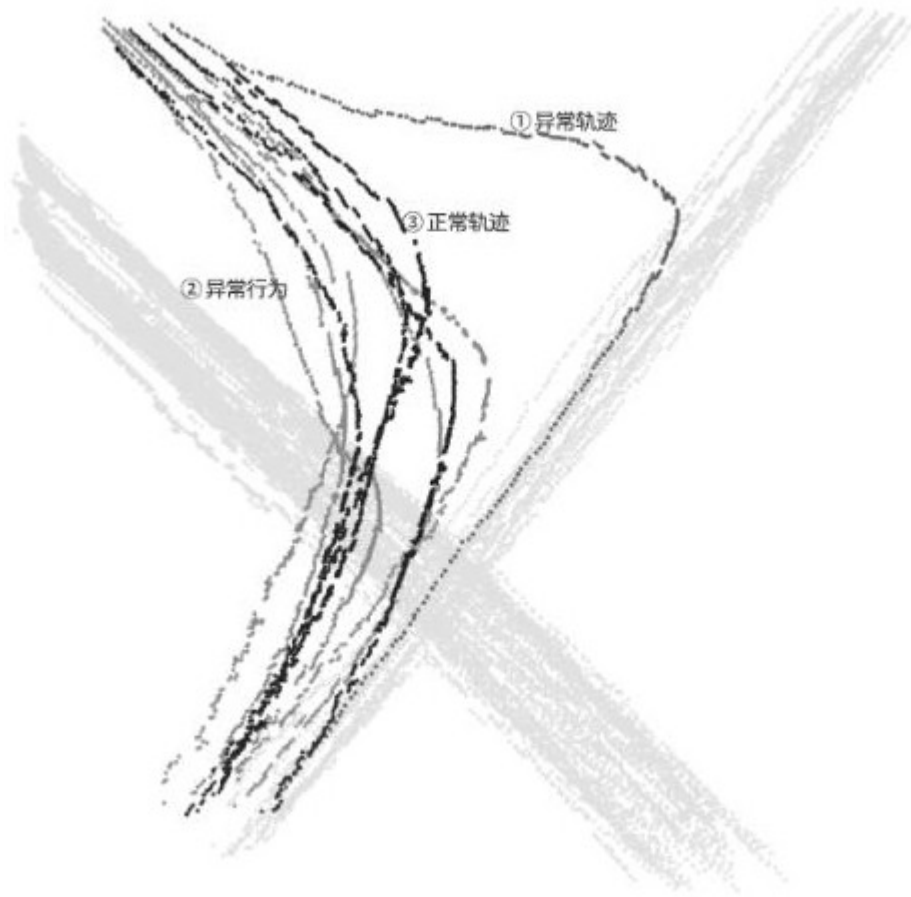


图7

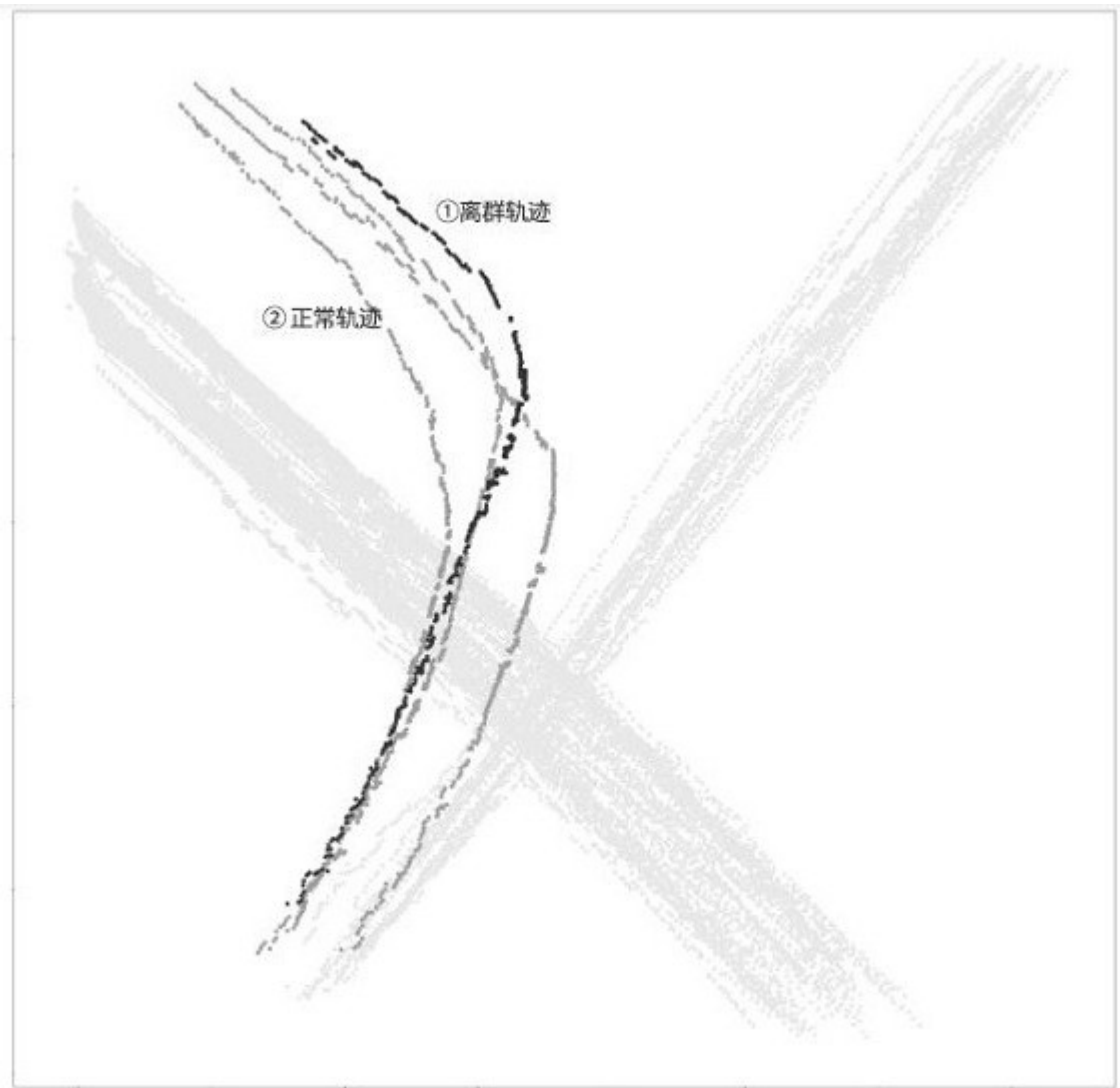


图8





图9