

一种基于车载信号还原机动车道3D地图的大数据方法

黄川, 胡平, 连静
(大连理工大学汽车工程学院, 辽宁大连 116024)

摘 要: 提出一种基于多台行驶中汽车产生数据重建机动车道3D地图的大数据策略. 每台在线汽车上的程序实时上传经过优化的汽车3D坐标信息至服务器. 优化方法为使用最小二乘法结合卡尔曼滤波器, 利用汽车总线信号实时修正汽车的位置, 相比GPS信号, 经纬度和高度误差均降低50%以上. 此外, 还使用遗传算法代替卡尔曼滤波器, 进一步降低卡尔曼滤波器的经纬度误差达16%. 其次, 服务器根据来自多台在线汽车上传的数据建立道路表面的3D点云数据库, 并使用K-聚类算法进行数据挖掘, 可推算出具有多条行车线道路的每条行车线的中心轨迹, 以此建立机动车道3D地图. 所建立的地图可为汽车能耗优化策略提供数据支持, 降低行驶能耗.

关 键 词: 地图重建; 大数据; 卡尔曼滤波器; 遗传算法; K-聚类
中图分类号: P 284.9 **文献标志码:** A **文章编号:** 1005-3026(2020)06-0771-07

A Big Data Method to Rebuild 3D Road Map Based on Vehicle Data

HUANG Chuan, HU Ping, LIAN Jing
(School of Automotive Engineering, Dalian University of Technology, Dalian 116024, China. Corresponding author: LIAN Jing, E-mail: lianjing80@126.com)

Abstract: A big data based 3D map reconstruction method was presented based on multiple vehicles data. On individual vehicle site, in the proposed algorithm, the least square method and Kalman filter were used, and the refined vehicle instant 3D position was uploaded to the server. The original GPS signal error was reduced by over 50%. Next, genetic algorithm was used instead of Kalman filter, resulting in the error being further reduced by over 16%. On the server site, a 3D road surface point cloud database was generated based on the data from multiple vehicles. K-means method was used as the data mining strategy to search for the lane centers from roads with multiple lanes. The reconstructed map can be used by every online vehicles that support the relevant researches for the optimum driving strategy.

Key words: map reconstruction; big data; Kalman filter; genetic algorithm; K-means

随着国家对于汽车工业的排放要求进一步提高,许多研究转向了行驶过程中的优化控制^[1-5]. 这些方法均需要未来行驶道路的坡度数据,目前国内大部分地图数据库并不包含足够精确的坡度信息. 因此应用这些技术需要建立一个具有地形信息的地图数据库. 针对汽车行驶地形的研究工作大致分为计算机视觉^[6-8]、激光探测^[9-10]和加

速度传感器^[11-12]. 计算机视觉与激光探测多应用在对道路坡度和两旁物体进行探测^[6-10]. 加速度传感器多用于探测路面的平整度,尤其对于柏油路面的疲劳损坏等造成行驶过程中颠簸的因素具有良好检测效果^[11]. 然而,上述方法对地形的探测精度并不高,并且需要汽车安装特定设备并由工作人员按照既定的路线行进行数据采集.

本文提出一种后台运行、基于多台配备 GPS 汽车的行驶数据重建机动车道 3D 地图的方法. 首先将 GPS 定位中经纬度坐标转换成平面直角坐标系,并通过总线信号推算汽车行驶轨迹. 其次利用最小二乘法以及卡尔曼滤波器或遗传算法对 GPS 轨迹与总线信号推算轨迹融合,得到更高精度的 3D 位置信息并将这些信息通过移动网络实时上传至服务器. 服务器收到来自多台汽车上传的数据建立数据库并使用 K-聚类算法寻找道路中心线,以此建立机动车道 3D 地图. 该策略可在后台实时运行并上传数据,在用户正常的用车过程中完成高精度机动车道 3D 地图的建立,制图成本低,并且当道路重新建设导致其地形发生变化时可以根据最新数据更新地图数据库.

1 坐标转换及行驶轨迹推算

1.1 坐标转换

GPS 与总线信号推算行驶轨迹使用的坐标系不同,因此需要将其转换为通用横轴墨卡托投影,将经纬度坐标转化成平面直角坐标.

$$x=0.999\,6\times\frac{2\delta+\gamma^2\cdot N\cdot\sin\varphi\cdot\cos^3\varphi\cdot(5-\tan^2\varphi+9\eta^2+4\eta^4)}{2},$$
$$y=0.999\,6\times[\gamma\cdot N\cdot\cos\varphi+\frac{\gamma^3\cdot N}{6}\cos^3\varphi\cdot(1-\tan^2\varphi+\eta^2)+\frac{\gamma^5\cdot N}{120}\cos^5\varphi(5-18\tan^2\varphi+\tan^4\varphi+14\eta^2-58\eta^2\cdot\tan^2\varphi)],N=\frac{a}{\sqrt{1-e^2\sin^2\varphi}},$$
$$\eta=e'^2\cos^2\varphi,e'=\frac{a^2+b^2}{b^2}.$$

(1)

式中:γ 代表经度;φ 代表纬度;a,b,e 分别代表地球近似椭圆长、短半轴及偏心率. 以 GPS 天线正下方的地面为参考点,忽略由于路面颠簸造成天线高度的微小变化. 天线所处位置坐标列为(x₁,y₁,z₁),天线正下方地面位置列为(x,y,z).

$$x=x_1;y=y_1;z=z_1-h/\cos\theta.$$

(2)

式中:h 代表汽车在水平时 GPS 天线距离地面的高度;θ 代表汽车所处地面与水平方向的夹角.

1.2 行驶轨迹推算

实验发现,使用横摆率推算的汽车行驶轨迹准确性要高于前轮转角信号推算的轨迹,因此采用横摆率推算汽车行驶轨迹.

$$\left. \begin{aligned} \alpha &= \int_0^t V_{\text{yaw}} \text{d}t, \\ x'_t &= \int_0^t v_t \cdot \sin\alpha \cdot \cos\theta_t \text{d}t, \\ y'_t &= \int_0^t v_t \cdot \cos\alpha \cdot \cos\theta_t \text{d}t, \\ h'_t &= \int_0^t \sin\theta_t \cdot v_t \text{d}t. \end{aligned} \right\} \quad (3)$$

式中:α 为汽车行驶的方向与初始方向的夹角;x'_t,y'_t,h'_t 为根据车身 CAN 信号推算出的 t 时刻汽车行驶轨迹的三轴坐标;d t 为信号更新的周期;V_{yaw},v_t 分别为车身横摆率和实时车速.

2 数据处理方法

2.1 最小二乘法

总线信号还原的轨迹坐标仅代表 t 时刻车身相对初始时刻的相对位置,并存在系统误差. 采用最小二乘法作为数据初步拟合算法,将总线信号还原的轨迹坐标与同一时刻的 GPS 位置建立对应关系. 以 t₀ 时刻 GPS 坐标作为起始位置.

$$\left. \begin{aligned} [x'',y''] &= T \cdot [x',y'], \\ T &= \begin{bmatrix} \cos\beta & \sin\beta \\ -\sin\beta & \cos\beta \end{bmatrix}. \end{aligned} \right\} \quad (4)$$

式中:T 是方向余弦变换矩阵;β 是旋转角度;[x',y']与[x'',y''] 分别是转换前后的点阵. 程序根据式(5)对 β 进行搜索,n 代表数据中点的个数.

$$E^2 = \sum_{\min} \frac{(x''_i - x)^2 + (y''_i - y)^2}{n}.$$

(5)

2.2 卡尔曼滤波器

GPS 信号短期误差较大,但误差不随着时间增加. 总线信号推算的轨迹在每个测量周期内保持较高精度,但每一步的计算都基于前一步结果的积分,误差随时间而增大. 卡尔曼滤波器^[12] 通过迭代将两个测量结果拟合得到一个更接近真实值的结果. 假设两种测量的误差符合高斯分布,两种测量结果相互独立. 在不存在控制量的情况下,卡尔曼滤波器的初始化公式为

$$\left. \begin{aligned} x_{k|k-1} &= x_{k-1|k-1}, \\ p_{k|k-1} &= p_{k-1|k-1} + Q. \end{aligned} \right\} \quad (6)$$

式中:x_{k|k-1}和x_{k-1|k-1} 分别代表 k 时刻的预测值和 k-1 时刻的修正值并始终相等;而 p_{k|k-1} 和 p_{k-1|k-1} 分别代表 k 时刻的误差协方差的预测值和 k-1 时刻的误差协方差的修正值;Q 是符合高斯分布的过程噪声协方差,即行驶轨迹推算中每一步误差的平方. 通过式(6)初始化得到初始值,

随后的迭代为

$$\left. \begin{aligned} K_k &= \frac{p_{k|k-1}}{p_{k|k-1} + R}, \\ x_{k|k} &= x_{k|k-1} + K_k \cdot (z_k - x_{k|k-1}), \\ p_{k|k} &= (1 - K_k) \cdot p_{k|k-1}. \end{aligned} \right\} \quad (7)$$

式中: K_k 代表 k 时刻的卡尔曼增益; R 是符合高斯分布的测量噪声协方差, 即 GPS 误差值的平方; $x_{k|k}$ 代表 k 时刻的输出修正值; z_k 代表 k 时刻系统的观测值, 即 GPS 测量值。式(6)与式(7)经整理为

$$\left. \begin{aligned} K_k &= \frac{p_{k|k-1}}{p_{k|k-1} + R}, \\ x_{k|k} &= x_{k-1|k-1} + K_k \cdot (z_k - x_{k-1|k-1}), \\ p_{k+1|k} &= (1 - K_k) \cdot p_{k|k-1} + Q. \end{aligned} \right\} \quad (8)$$

2.3 遗传算法

遗传算法中, 以每一步 GPS 坐标为圆心, GPS 误差范围为半径作圆, 随机分布建立原始种群, 以行驶经过这些点后与后续 GPS 点的偏离程度作为适应度函数, 选择偏移度小的个体。适应度函数如下:

$$\left. \begin{aligned} T' &= \begin{bmatrix} \cos\left(\frac{x_1 \cdot \pi}{180}\right) & \sin\left(\frac{x_1 \cdot \pi}{180}\right) \\ -\sin\left(\frac{x_1 \cdot \pi}{180}\right) & \cos\left(\frac{x_1 \cdot \pi}{180}\right) \end{bmatrix}, \\ [(x_f - x_{il})^T, (y_f - y_{il})^T] &= \\ T' \cdot [(x_i - x_{il} + x_2)^T, (y_i - y_{il} + x_3)^T], \\ E &= \min \sum_1^n [(x_f - x_{gn})^2 + (y_f - y_{gn})^2]. \end{aligned} \right\} \quad (9)$$

式中: x_f, y_f 分别代表经过修正后的未来 n 个点的横、纵坐标, 并在下一步被用来计算修正后的方差; x_1, x_2, x_3 分别代表角度修正参数及 x, y 轴修正参数; T' 代表角度修正的变换矩阵; x_i, y_i 分别代表未来 n 个点经过最小二乘法修正后的积分计算结果的横、纵坐标; x_{il}, y_{il} 分别代表被修正点的经过最小二乘法修正后的积分计算结果的横、纵坐标; x_{gn}, y_{gn} 分别代表 GPS 结果中未来 n 个点中每一个点的横、纵坐标; E 是修正结果的最小方差。

2.4 K-聚类

K-聚类^[13]将数据以 K 个类心聚成 K 个聚类。所有数据中随机取 K 个数据作为聚类的初始类心。计算所有数据到类心的距离, 根据数据间的距离来确定它们之间的相异度。并将其归类到对应距离最小的那个类心所在的聚类中。根据聚类结果, 重新计算 K 个聚类各自的中心, 计算方法是取聚类中所有元素各自维度的算术平均数。若

产生的新类心与之前的类不同, 则重复前两步, 直到结果不再变化。算法中 K-聚类是分步进行的, 只针对点云数据库中的部分数据点。选取这些点时首先按照一定距离选取一个点 (x_t, y_t) 作为基准点, k_t 是 t 时刻相对 $(t-1)$ 时刻的移动方向的斜率。过 (x_t, y_t) 点作一条垂直于该时刻移动方向的辅助线, 并在数据库中所有距离参考点距离小于 l , 并且从行驶方向的斜率与 k_t 相差小于 β 的点中, 找到所有距离直线 r 的距离小于 d 的点。

$$\left. \begin{aligned} k_t &= \frac{y_t - y_{t-1}}{x_t - x_{t-1}}, \\ y_r &= \frac{x_r - x_t}{-k_t} + y_t, \\ l &\geq \sqrt{(x_t - x_0)^2 + (y_t - y_0)^2}, \\ \beta &\geq |k_t - k_0|, \\ d &\geq \left| \frac{x_0 + k_t \cdot y_0 - (x_t + k_t \cdot y_t)}{\sqrt{k_t^2 + 1}} \right|. \end{aligned} \right\} \quad (10)$$

式中: x_{t-1}, y_{t-1} 代表汽车在前一时刻的位置; x_r, y_r 表示直线 r 的横、纵坐标; x_0, y_0 代表点云数据库中全部点的坐标。

3 实验环境与结果

实验数据采集于具有设计图纸的内部测试场地, 准确参数可以通过查阅设计图纸获得。测试场地大部分是海拔高度差在 1 m 之内的平路, 同时包含一段上下坡路, 坡路的细节见图 1。测试车采用 BMW 530Le, 测试汽车的 GPS 天线安装在车身顶部, 取车身高度作为天线距离地面的高度为 1.50 m, 信号更新周期为 0.10 s。

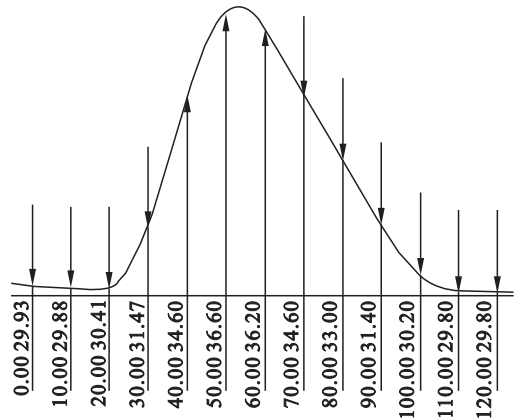


图1 测试场地上下坡局部侧视图(单位:m)
Fig. 1 Test track slope drawing (unit:m)

测试共计 7 圈, 耗时约 900 s, 行驶里程约 17 km, 气温 -10°C , 天气晴朗, GPS 卫星数量持续大于 7。收集的数据包括: 车速、横摆率、车身倾

角、GPS 经纬度及高度. 利用坐标转换以及推算的行驶轨迹分别见图 2 与图 3.

图 2 中, 曲线表示的 GPS 具有长期稳定性, 但是短期会有较大的偏移, 从局部放大图可见, 结果有比较大的波动. 与之相反, 图 3 中曲线代表的总线数据推算的行驶轨迹在短期内可以准确反映汽车运动轨迹, 曲线平滑, 并未出现 GPS 结果中那样的波动. 然而长期来看, 误差会随着累积逐渐增大.

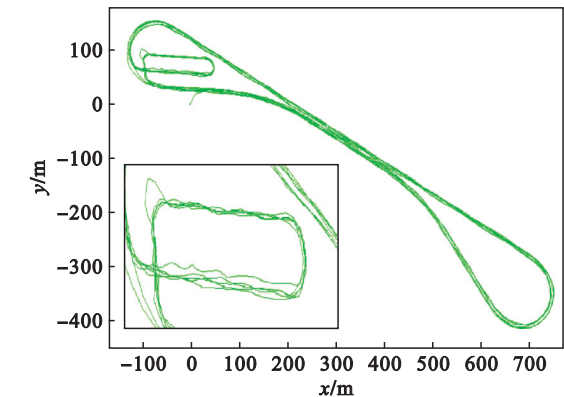


图 2 GPS 测量行驶轨迹
Fig. 2 Driving track based on GPS

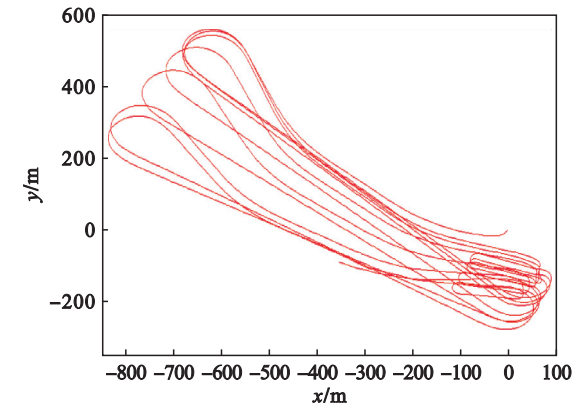


图 3 基于 CAN 的行驶轨迹
Fig. 3 Driving track based on CAN

3.1 最小二乘法

选取测试开始的位置作为初始位置, 通过最小二乘法将两种结果初步拟合, 结果见图 4, 其中 GPS 轨迹与图 2 相比并未发生变化. CAN 相比图 3 离散度明显降低, 而且与 GPS 结果基本重合. 高度是一维变量, 将 GPS 初始高度值作为行驶轨迹推算的高度初始值, 未作其他修正.

3.2 卡尔曼滤波器

卡尔曼滤波器需要使用的一些参数见表 1. 卡尔曼滤波器的经纬度拟合结果见图 4. 结果显示, 卡尔曼滤波器保持了 GPS 与总线信号推算的优点, 相比单纯使用 GPS 的结果, 定位精度大幅提升.

根据设计图纸均匀地选取 72 个点作为标准值, 以不同方法绘制的曲线到这些点的最短距离作为误差值. 结果如图 5 所示, 未经过卡尔曼滤波器的 GPS 平均误差达到 1.65 m, 方差 0.78. 经过卡尔曼滤波器融合后, 平均误差降至 0.72 m, 方差 0.12. 卡尔曼滤波器将误差降低了 56.4%, 描述误差波动程度的方差降低约 84.6%.

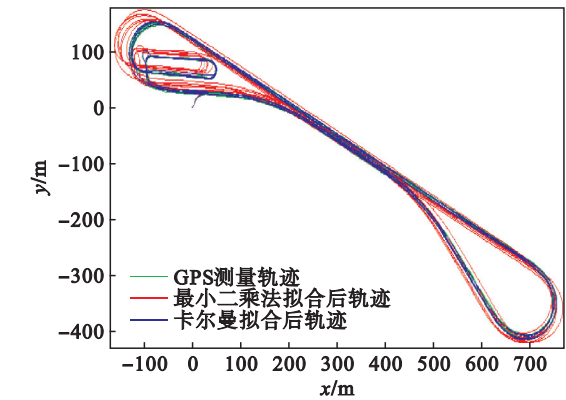


图 4 行驶轨迹拟合结果
Fig. 4 Driving track results

表 1 卡尔曼滤波器设定参数
Table 1 Kalman filter parameters

参数	符号	取值
GPS 精度/m	—	5
CAN 推算精度	—	0.1
初始协方差	—	4.01
GPS 协方差	R	25
CAN 推算协方差	Q	0.01

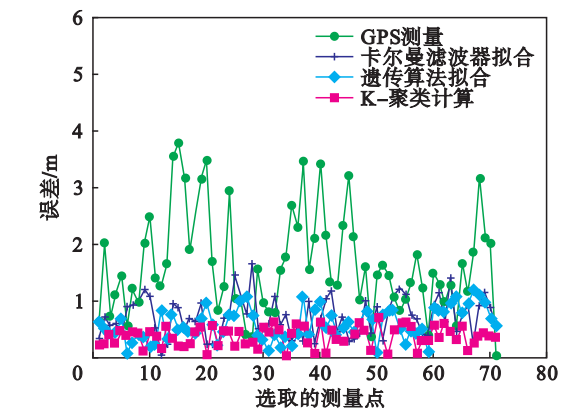


图 5 多种算法结果的误差
Fig. 5 Result errors for corresponding methods

测试车 GPS 高度信号的测量分度值是 1 m, 远高于经纬度转化为 $x-y$ 平面直角坐标后的分度值, 因此, 高度拟合被单独计算并且仅使用卡尔曼滤波器. 测试场地大部分处于水平, 仅在一处有上下坡, 上下坡融合后的结果与图 1 的比较如图 6 所示. 设计图中标注出的参考点共 12 个, 以圆

点表示,曲线描述了卡尔曼滤波器计算后的结果,结果显示卡尔曼滤波器的拟合结果与标准值误差均小于 0.20 m. 其余 72 个参考点的计算高度与设计数据相比,高度误差如图 7 中的曲线所示,平均误差为 0.12 m.

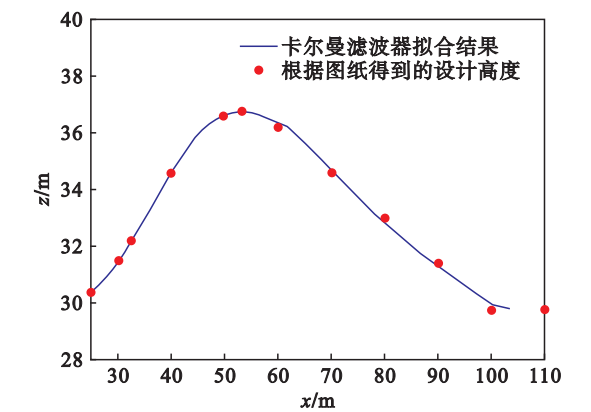


图 6 卡尔曼滤波器上下坡处高度拟合结果
Fig. 6 Kalman filter altitude matching result at the slope

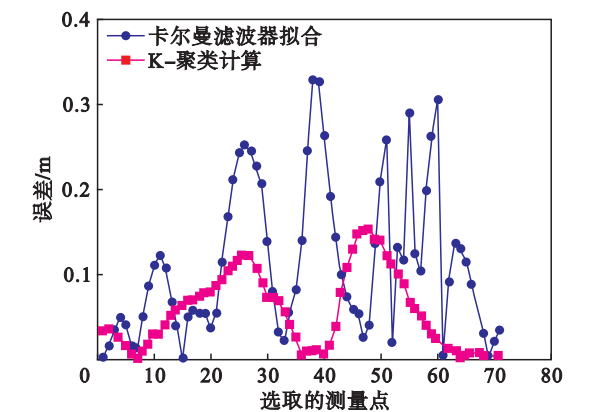


图 7 卡尔曼滤波器全程高度拟合结果误差
Fig. 7 Kalman filter altitude matching result along the track

3.3 遗传算法

$x-y$ 平面上,使用遗传算法进行信号融合可以得到精度高于卡尔曼滤波器的拟合结果. 高度方面,由于 GPS 高度信号分度值过大,遗传算法无法持续稳定收敛至最优结果附近,因此仅使用遗传算法对 $x-y$ 平面数据融合. 遗传算法的相关计算由 Matlab 提供的遗传算法工具箱完成,误差取值与具体的参数设置见表 2,表 3(未列出系统默认值).

72 个采样点使用遗传算法拟合的结果误差由图 5 中的曲线给出,遗传算法的平均误差为 0.60 m,方差 0.08,分别比卡尔曼滤波器的计算结果降低了 16.7% 和 33%. 表明遗传算法在经纬度方面的融合具有比卡尔曼滤波器更小的误差及更好的稳定性.

表 2 遗传算法误差取值 Table 2 Genetic algorithm errors			
参数	符号	单位	取值
X1 误差范围	X1	($^{\circ}$)	$[-1,1]$
X2 误差范围	X2	m	$[-5,5]$
X3 误差范围	X3	m	$[-5,5]$

表 3 遗传算法参数设置 Table 3 Genetic algorithm parameters settings	
项目	设定
变量数量	3
编码类型	双精度
生成函数	约束相关
适应度尺度	等级
依据适应度选择父代	随机均匀分布
突变	约束相关
杂交	随机生成向量

3.4 K-聚类

K-聚类可以将来自不同车道的数据分类并求对应类心. 首先根据式(10)建立运用 K-聚类算法所需要的局部数据,边界条件的定义见表 4. 选用遗传算法的拟合结果,随机选取其中一个点进行数据挖掘处理,计算过程如图 8 所示.

表 4 K-聚类边界条件 Table 4 K-means boundary conditions			
参数	符号	单位	取值
距离参考点最大距离	L	m	10
行驶方向最大偏差	α	($^{\circ}$)	5
距离参考直线最大距离	d	m	3

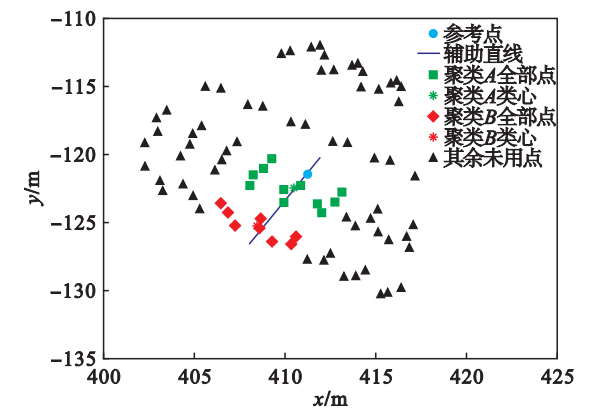


图 8 K-聚类计算过程
Fig. 8 K-means computing process

参考点选自测量数据,在图 8 中以圆点表示. 根据式(10),点云数据库中所有距离参考点 10 m 内的点被列出(图 8 中全部点). 根据每个点的运动方向以及与建立的辅助线的距离选取 K-聚类

的数据点(图 8 中全部非三角形点). 参考点上方的三角形点为反向车道行驶时产生的数据点,行驶方向与参考点完全相反,因而未被采用. K-聚类算法产生了 A、B 两个分类,参考点属于聚类 A. 星号表示聚类的类心,作为行车线的中心建立最终机动车道 3D 地图. 通过以上方法完成对其余参考点的计算. 还原的机动车道 3D 地图在 72 个采样点的横、纵坐标和高度误差分别如图 5 和图 7 中曲线所示. 其中横、纵坐标平均误差降至 0.37 m,方差降为 0.03,相比 GPS 分别降低 78% 和 96%,同时高度平均误差为 0.08 m.

3.5 实 测

取一段真实环境中的山路进行实车测量,测量地点为沈阳市棋盘山,全程约 2 km. 由于无法获得该段道路准确设计参数,只能通过电子地图来评估水平面上大数据建立地图方法的准确性,先对该段道路地图截图进行预处理,将公路位置转化成数据作为比较对象,评估结果见图 9.

通过大数据方法还原的道路表面轨迹与电子地图数据重合,说明二者基本一致. 坡度数据的采集方式为:驾驶测试车沿路行驶并使用角度测量仪测量一次停车所在位置地面的坡度信息,从起点开始每间隔 100 m 进行采集,全程总计采集 20 个数据点,一共采集两次取平均值为测量值,见表 5.

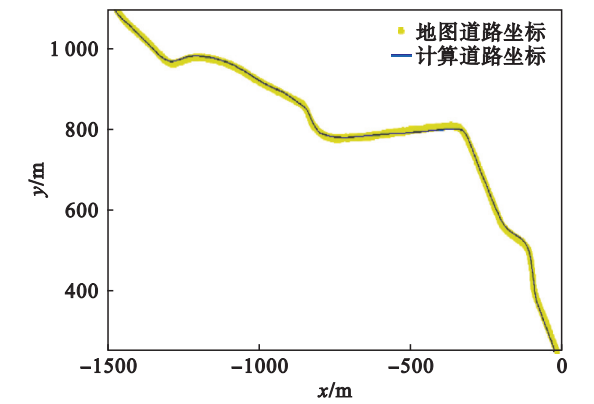


图 9 实测结果
Fig. 9 Test result

表 5 坡度数据采集结果
Table 5 Slope data collection

标记点	1	2	3	4	5	6	7	8	9	10
计算值/%	7.3	5.2	-4.2	-2.5	7.1	6.0	0.0	6.6	5.6	6.1
测量值/%	6.9	4.8	-4.1	-2.2	6.8	6.0	0.0	6.4	5.3	5.6
标记点	11	12	13	14	15	16	17	18	19	20
计算值/%	4.2	-6.1	-16.1	-9.4	4.7	4.1	5.4	-3.4	-7.1	-8.0
测量值/%	4.2	-6.0	-13.5	-7.6	5.0	3.9	5.4	-3.4	-7.0	-8.0

测量值接近于计算结果,进一步计算得出:聚类结果与测量坡度值误差约 0.38%. 结合通过本文使用大数据建立的地图在经纬度上及坡度上的误差考虑,可以认为大数据方法能够有效建立供新能源汽车使用的包含机动车道地形信息的电子地图.

4 结 语

通过利用正常行驶汽车的数据,还原了高精度的汽车实时位置,这个过程中遗传算法相比卡尔曼滤波器精度更高. 实时位置被上传至服务器,服务器利用来自多个汽车上传的数据建立数据库,对数据进行挖掘,绘制机动车道 3D 地图. 实验结果显示该方法所建立的地图水平误差平均为 0.37 m,高度误差平均为 0.08 m. 地图的建立过程由软件自动完成,驾驶员只需正常驾驶汽车,算法可根据最新的数据自动更新 3D 地图数据. 建

立的 3D 地图具有精度高、低成本、实时更新等优点,可以为行驶策略优化提供相关的地形数据支持,降低汽车行驶能耗.

参考文献:

[1] Mahyar V, Nasser L. Ecological adaptive cruise controller for plug-in hybrid electric vehicles using nonlinear model predictive control [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2016, 17(1): 113-122.

[2] Chen Z, Vahidi A. Route preview in energy management of plug-in hybrid vehicles [J]. *IEEE Transactions on Control Systems Technology*, 2012, 20(2): 546-553.

[3] Gong Q, Li Y, Peng Z. Trip-based optimal power management of plug-in hybrid electric vehicles [J]. *IEEE Transactions on Vehicular Technology*, 2008, 57(6): 3393-3401.

[4] Grant M, Ardalan V. An optimal velocity-planning scheme for vehicle energy efficiency through probabilistic prediction of traffic-signal timing [J]. *IEEE Transactions on Intelligent Transportation Systems*, 2014, 15(6): 2516-2523.

[5] Li X, Chen Y, Wang J. In-wheel motor electric ground vehicle energy management for maximizing the travel distance [C]//American Control Conference. Montreal,

- [illegible]

Engineering and Science, 2017, 39 (8): 1403 - 1409.)

[12] 谢彬, 张琨, 张云纯, 等. 基于轨迹相似度的移动目标轨迹预测算法[J]. 计算机工程, 2018, 44 (9): 177 - 183. (Xie Bin, Zhang Kun, Zhang Yun-chun, et al. A trajectory prediction algorithm for moving targets based on trajectory similarity[J]. *Computer Engineering*, 2018, 44 (9): 177 - 183.)

[13] 宋路杰, 孟凡荣, 袁冠. 基于 Markov 模型与轨迹相似度的移动对象位置预测算法[J]. 计算机应用, 2016, 36 (1): 39 - 43. (Song Lu-jie, Meng Fan-rong, Yuan Guan. Location prediction algorithm of moving objects based on Markov model and trajectory similarity[J]. *Computer Application*, 2016, 36 (1): 39 - 43.)

[14] 李晓娟. 在 GPS 数据挖掘中基于向量自回归的 POI 类型转移矩阵规律与预测[J]. 现代计算机, 2019 (1): 33 - 39. (Li Xiao-juan. The rule and prediction of POI type transfer matrix based on vector autoregression in GPS data mining[J]. *Modern Computer*, 2019 (1): 33 - 39.)