

一种基于GPS轨迹的道路拓扑生成方法

谭康 刘建勋 廖祝华

(湖南科技大学计算机科学与工程学院 湘潭 411201)

摘要 复杂路网拓扑的自动生成建立在道路提取和交叉路口识别的基础之上,是智能交通控制和自动导航服务等领域的研究热点之一,基于浮动车或出租车的GPS轨迹可以反映交通路网的拓扑结构。为此,提出了一种基于GPS轨迹的道路拓扑生成方法,即在无道路地图辅助的情况下,该方法基于大规模GPS轨迹,能够快速提取路口交叉点,自动构建具有地理位置信息的拓扑结构和计算相邻路口的网络距离。实验结果表明,该方法能够提取出各个道路交叉点并建立各点之间的拓扑关系。在提取主干道路拓扑实验中,在设置路宽为55米的情况下提取路口交叉点的正确率达到了87.08%,各路口之间的平均网络距离误差率为8.87%,并且能够正确地得到交叉点之间的连通关系。

关键词 道路拓扑,网络距离,路口点,连通关系

中图法分类号 TP315 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.9.008

Topology Generation Method of Road Network Based on GPS Trajectories

TAN Kang LIU Jian-xun LIAO Zhu-hua

(School of Computer Science and Engineering, Hunan University of Science and Technology, Xiangtan 411201, China)

Abstract Automatic generation of complex network topology is based on road extraction and road intersection detecting, which is one of the hot research topics in intelligent traffic control and automatic navigation service fields, and GPS trajectories generated by floating cars or taxis can reflect the road topology. Therefore, this article presented a method to extract road intersection and build topology. It extracts road intersections, builds topology with geographical location information based on large-scale GPS trajectories without auxiliary road map, and calculates network distance between any two adjacent road intersections. The result shows that our method can extract road intersection and build topology effectively. In our experiment, when the road width is set with 55 meters, the accuracy of extracting road intersection is 87.08%, the average error rate between adjacent road intersections is 8.87%, and connected relationship between adjacent intersections can be gotten correctly.

Keywords Road topology, Network distance, Road intersection, Connected relationship

1 引言

复杂路网拓扑的自动生成建立在道路提取和交叉路口识别的基础之上,是智能交通控制和自动导航服务等领域的研究热点之一。基于浮动车或出租车的GPS轨迹不仅能够实时反映特定区域的动态交通信息,而且还反映了交通路网的拓扑结构,这正是导航软件路径规划必需的数据。

本文把连接各路口交叉点之间的道路称为“道路段”。路网拓扑结构就是利用路口交叉点来表达路段与路段之间的连通性。普通的GPS轨迹数据包含经纬度、记录的时间、海拔等信息,所有记录下的GPS点的存储结构都是一样的,这为区分哪些点是路口交叉点带来了困难。

OpenStreetMap^[1]中志愿者可以手动编辑GPS踪迹和添加航摄图像,这种方法可以不用专业的地图绘制车辆,但同一个地方不同志愿者编辑的结果可能不一样;张小国等^[2]在道路段已经给出的情况下通过对所有道路段求交来建立道路拓

扑,但是在确定道路是否相交的问题上阈值不容易把握,过大会造成不相交的道路相交,过小会让原本连接的道路断开;Schroedl等^[3]和Edelkamp等^[4]首先为GPS轨迹聚类找到道路中心线,然后利用中心线确定道路和路口结构;Alireza Fathi等^[5]则利用GPS轨迹相交先找出可能的路口点,然后得到道路线,再利用路口点周围的GPS数据进一步优化路口点位置,但是原始的GPS轨迹数据非常大,并且轨迹相交数目多大才取为路口点的阈值不好设定。

传统的路网提取方法得到的路网信息更新慢、成本高,很多地区的路网整年得不到更新,道路建设等原因导致路径规划经常缺少最新的道路地图信息。出租车每天在城市道路行驶,GPS数据易采集、成本低、更新快,可以很好地弥补传统方法的不足。本文的目的就是在没有任何道路地图辅助的情况下建立一个具有拓扑结构的路网图。根据普通的车载GPS数据的特点和道路的实际情况,本文设计并实现了一种路口交叉点提取并构建拓扑的方法:首先,对原始GPS轨迹

到稿日期:2014-05-22 返修日期:2014-08-16 本文受国家自然科学基金(61272063,61370227),湖南省自然科学基金(13JJB004)资助。

谭康(1988—),男,硕士生,主要研究领域为GIS与移动计算,E-mail: tank_1125@163.com;刘建勋(1970—),男,博士,教授,主要研究领域为服务计算与云计算、工作流;廖祝华(1977—),男,博士,副教授,主要研究领域为网络通信与移动计算。

数据进行处理得到北京市的主干道路网;其次,对道路中因为聚类而缺失的道路点进行增补操作,去掉道路中多余的道路点;然后,提取出路段和道路的交叉点,存储它们之间的连通关系;最后,对结果进行评估。

2 道路拓扑生成方法

2.1 GPS 轨迹聚类

根据 GPS 轨迹进行道路提取的方法已经有很多。本文借鉴欧阳鸿等^[6]的路网提取方法,对北京市五环区域内的主干道路路网进行了提取。由于整个区域内的 GPS 数据达到了 1.57GB,因此将五环分割成 33 个小区域,聚类完毕后数据将大大减少,然后再将各小区域得到的数据放到一起重新聚类一次。聚类的作用是将相同道路的 GPS 轨迹进行合并、简化,从而从大量轨迹中容易提取道路形状,并容易识别交叉路口。文献^[6]使用的是一种滚动式聚类算法:

(1)选取一个未被访问的轨迹点 p_i ;

(2)以 p_i 为圆心、 dm 为半径聚类,计算圆内所有轨迹点经纬度的平均值 p_{ave} ,存储 p_{ave} ;

(3)找出离 p_{ave} 最近的未被访问的轨迹点,重复步骤(2),直到所有轨迹点都聚类完毕。

图 1 是根据 GPS 数据得到的北京市主干道道路网,其全部由 GPS 点构成,聚类半径为 100m。聚类后所有的 GPS 数据包含了经纬度信息,按照存储 GPS 点的顺序将其编号,本文将这些 GPS 点统称为道路点。

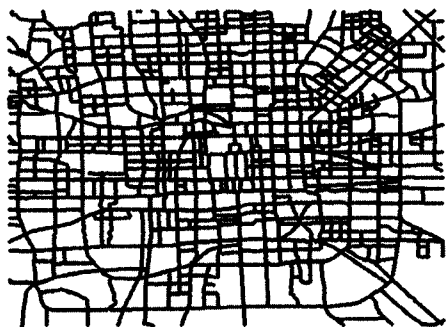


图 1 北京市五环内的主干道路网

2.2 道路点预处理

图 1 中,聚类后的道路点并不是等距的,以 dm 为聚类半径时,大部分道路点都相距 $2dm$ 以内,但有的点即使在同一道路上,经过聚类操作也会相隔超过 $2dm$,这为判断该相邻两点是否有道路连接带来了困难。为了消除这种道路“断开”的现象,必须在两个点之间根据条件增补一个点。

对于直线道路上的道路点,它们的经纬度在聚类后的文件中是按照聚类时的顺序存放的。根据这个特性,当文件中相邻两个点的距离大于 $2d$ 时,从聚类前的原始数据中寻找一条经过这两个相邻点的道路,只要找到了这样的路径,就在这两个点之间补一个点,点的经纬度是这两个点经纬度的平均值;否则就认为这两个道路点没有道路直接连通,不补点。增补的点在文件中要放在两端道路点中间以保持其顺序性。

这种增补道路点的方法有一定的误差,如图 2 所示,当按箭头 1 聚类后又从箭头 2 方向滚动聚类时,“断开”的两个点没有按顺序存储,导致两条相交的道路被认为没有连通。这

些道路点的性质是: $2d$ 范围内除圆心外只包含了不超过一个道路点。找出离 A 最近的另一条道路上的点 B ,利用与上述相同的方法判断是否应该在中间增补一个道路点。补点操作后通常同一道路上相邻的两个点之间相距在 $2dm$ 以内。

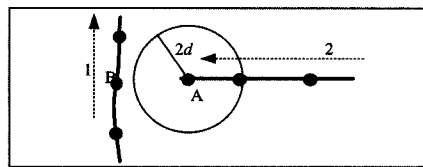


图 2 道路点缺省导致道路“断开”

2.3 路口交叉点提取

要从聚类结果中找出哪些点是道路交叉点,需要充分利用聚类后的道路点特征,其中最具代表性的是 William 等^[7]进行的道路交叉口探测,即利用道路点的度进行道路交叉点识别,该方法的优点是能够快速地探测交叉路口。图 3 是一个小区域内的道路网,聚类半径为 dm 。从图中可以看到,以路口交叉点为圆心,其 $2d$ 范围内至少包含 3 个聚类后的道路点。

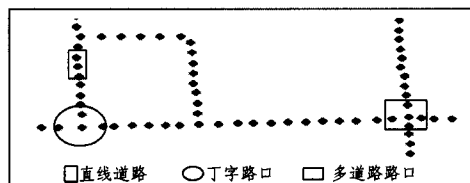


图 3 道路网中道路点的 3 种情况

首先按照这样的规则把所有符合道路交叉点性质的道路点取出来。仅仅按这种规则所选取的路口交叉点是不准确的,可能会导致两种错误:

(1)如图 4 所示,从左向右滚动聚类,当某处道路点比较密集时,得到的结果可能导致相邻的两个点距离在 dm 以内,这时即使是直线道路上的非路口交叉点, $2dm$ 的距离内也可能含有 3 个点。

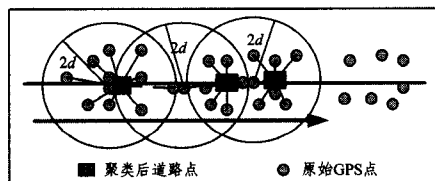
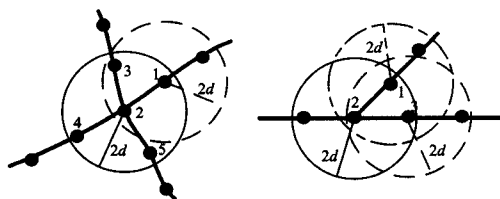


图 4 相邻道路点相距 d 米以内

(2)当道路夹角较小时,如图 5(a)所示,点 1 和点 2 都包含了两个以上的点,但道路点 1 不是路口交叉点。如果不进行路口交叉点正确性判断,道路点 1、2、3 甚至 4、5 都将被取为路口交叉点,并且 1、3 和 4、5 都是成对出现的。



(a)多路口交叉点 (b)丁字路口交叉点

图 5 错误的交叉点

对所有路口交叉点采用如下步骤判断:

(1)如图6所示,如果当前圆内除圆心外还包含3个道路点,则这有可能是一个丁字路口,也可能是图6(b)所示的情况。在聚类结果文件中,当以点2为圆心时,如果是丁字路口,不管以什么顺序聚类,道路点1、2、3、4都不是依次存放的;如果是直线上的道路,道路点将会依次相邻存放。根据这个特性,可以判断出点2是否是正确的路口交叉点。去掉直线上太相邻的点,如图6(b)中的道路点3。圆内必须不包含其他路口交叉点,否则会像图5(b)中,按3→2→1顺序聚类后再进行本次判断,会把路口交叉点2去掉。

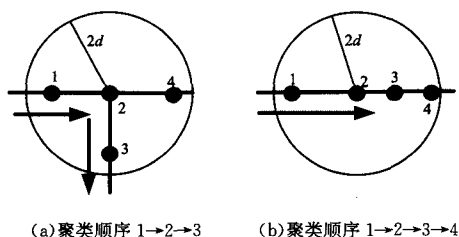


图6 丁字路口修正

(2)如果当前圆内除圆心外还包含了两个以上的成对出现的路口交叉点,则出现了图5的情况。图5(a)中,取包含点数最多的路口交叉点为正确的路口交叉点,这样即使是多交叉路口也能保证取到正确的道路交叉点;图5(b)不能根据聚类后的道路特征判断出哪个是路口交叉点,取最先判断为路口交叉点的点为路口交叉点,这会引入小于 d m的误差。判断后去掉其余点的路口交叉点性质。

(3)通过按存储顺序进行补点的操作后,路口交叉点处往往会产生两个十分靠近的道路点,它们都会被判断为路口交叉点,当这两个路口交叉点小于 d m时,取平均经纬度为路口交叉点,去掉它们的路口交叉点性质。

经过这个操作,就能去掉大部分误判的路口交叉点,保留正确的路口交叉点。

2.4 道路段提取

为了得到两路口交叉点之间的道路段,以当前路口交叉点为起点、 $2d$ m为半径,依次滚动式检索道路点,记录下道路点编号和相邻点的距离,直到找到另一个路口交叉点,将其作为终点,这就是一个完整的道路段,叠加记录下的所有距离长度就是整个路段的网络距离。有时到了一条道路的尽头并不能遇到路口交叉点,存储与道路段连接的单个路口交叉点即可。

经过道路交叉点提取、路段提取后,将道路交叉点、路段上的点分别存储在两个文件中并按顺序编号。

2.5 拓扑结构构建

路网拓扑结构的表示方法有多种,比较常用的方法有邻接矩阵法、前向星型结构(Forward Star Representation)^[8]法。本文利用一个路段表来存储相关数据和拓扑信息,路段表的结构如表1所列,从中可以获知路段相连信息。

表1 路段表格式

路段编号	道路点个数	道路段长度	包含道路点编号
------	-------	-------	---------

路段编号规则是:起始路口交叉点编号/终点路口交叉点编号,如果两个路口交叉点之间有多条道路段相连接,则路段

编号为:起始路口交叉点编号/终点路口交叉点编号/b,依此类推;如果是有尽头的道路,路段编号为:起始路口交叉点编号/0。

3 实验结果

本文实验数据源于微软亚洲研究院从2007年4月到2010年10月在北京收集的167辆出租车的GPS轨迹数据,它还提供了一个北京市shp格式的道路地图,采样频率为5秒一次,包括经度、纬度、采样时间等信息,如表2所列。

表2 实验数据基本描述

纬度	经度	海拔	日期	时间
39.984702	116.318417	492	2008-10-23	02:53:04

所有数据共1.57GB,存放数据的每个plt子文件都存储了出租车行驶的一条行程。实验在Visual Studio 2010+ArcMap10环境下进行。

3.1 选取聚类半径

北京市的主干道比较宽,加之GPS点漂移,如果聚类半径取几十米以内,主干道上会产生多条道路,不利于路口交叉点的提取,如图7所示。为了避免这种情况发生,我们尝试了不同的半径,根据聚类后的道路特征选取聚类半径为100m,如图7(d)所示,从而较好地保持了道路形状。

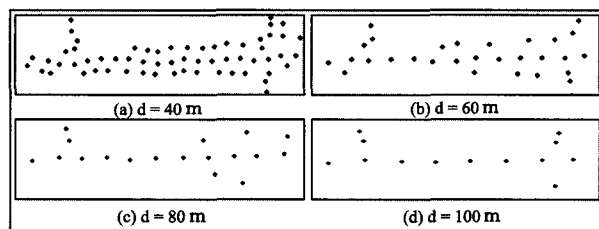


图7 聚类半径不同时的道路特征

3.2 提取路口交叉点结果

图8是提取路口交叉点后形成的主干道路网图,相对较大的点为提取出的路口交叉点。整个五环区域内总共提取出了9781个道路点,包含了804个路口交叉点,形成了1301个道路段,从图中可以看出实验结果基本包含了各个路口交叉点。

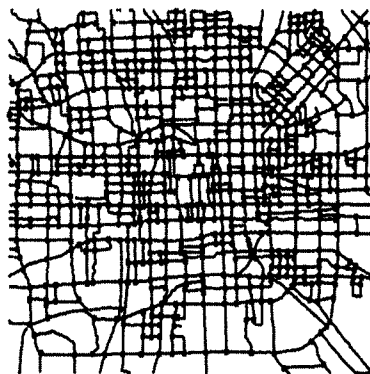


图8 主干道路网路口交叉点提取图

3.3 拓扑构建结果

图9中上半部分为在ArcMap中随机选取的一个区域,下半部分表示这一区域的路段表,从图中可以看出表中的连通性与上半部分的路网连通性是相同的。

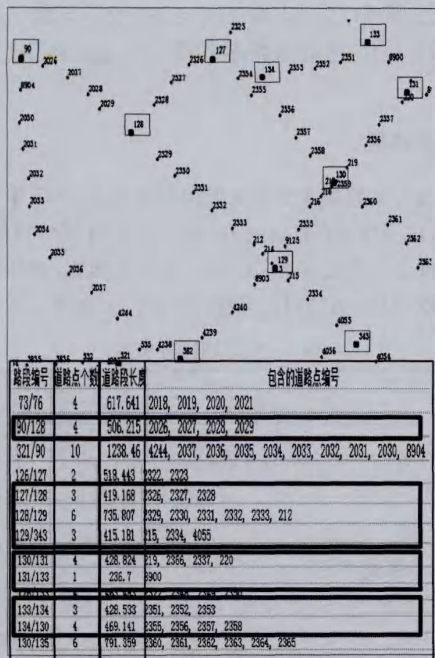


图9 区域主干道路路口拓扑图

3.4 实验评价

由于本文提取的是主干道,没有支路信息,因此本文只对已经选出的道路与商业地图进行比较。为了对路网、路口交叉点进行精度评价,以谷歌、北京 shp 地图为参考路网,分别从定量与定性两个方面进行评价。

3.4.1 定量评估

参考 Zhang 等^[9]提出的缓冲区方法,以北京 shp 地图为缓冲分析对象,获取各缓冲半径内本文提取的路口交叉点与北京 shp 地图路口交叉点相互匹配的路口交叉点个数集合。利用 ArcMap 得到 shp 地图中主干道路路口交叉点,得到的路口交叉点数为 766,以道路交叉点为圆心、缓冲区为半径,只要实验提取的路口交叉点落入到这些缓冲区,就认为路口交叉点提取正确,将这样的路口交叉点称为落入点。城市道路分为快速路、主干路、次干路和支路 4 类;对于人口超过 200 万的大城市,快速路、主干路和次干路宽度分别为:40m~45m、40m~55m、40m~50m^[10]。用覆盖率 α 和误检率 β 来评估本文方法的路口交叉点提取性能^[11],其定义如式(1)、式(2)所示:

$$\alpha = \frac{\sum_{P \in \sigma} N(P)}{\sum_{P \in \varphi} N(P)} \quad (1)$$

$$\beta = 1 - \frac{\sum_{P \in \sigma} N(P)}{\sum_{P \in \varphi} N(P)} \quad (2)$$

其中, φ 表示 shp 地图路口交叉点集合, $\varphi = \{P_1, P_2, \dots, P_n\}$, σ 表示本文提取的结点集合, σ 为落入缓冲区内的路口交叉点集合, $N(P)$ 表示路口交叉点个数。

表 3 是提取的路口交叉点与北京 shp 地图比较的正确率。路网中每个道路点的经纬度是该聚类半径内所有原始轨迹点经纬度之和的平均值,而道路交叉口处 GPS 数量最多,这提高了路口交叉点的提取精度。

表 3 路口交叉点结果正确率

缓冲区半径 (m)	落入点 个数	覆盖率 α	误检率 β	落入点距圆心的平均距离 (m)
40	587	76.63%	27.0%	17.2
45	618	80.68%	23.13%	18.9
50	643	83.94%	20.02%	20.09
55	667	87.08%	17.04%	20.45

从 shp 地图中随机选取 50 个路口交叉点,计算出与其相邻的道路交叉点的网络距离;再找到从实验提取出的相应路口的道路交叉点,计算出与其相邻的道路交叉点的网络距离,与 shp 地图的网络距离进行比较,利用平均距离误差率 P 来判断实验所提取的道路段网络距离的正确性,公式如下:

$$P = \frac{\sum_{i=1}^N |S_i - E_i|}{N S_i} \quad (3)$$

其中, S_i 表示 shp 地图中选定的第 i 个路口与相邻路口的道路段网络距离, E_i 表示实验所提取的相应道路段网络距离, N 表示所有的道路段个数。所提取的 50 个路口交叉点的平均误差率为 8.87%,这说明实验所提取出的道路段网络距离与真实道路的网络距离比较接近。表 4 是随机选取的一个十字路口的网络距离比较结果。

表 4 网络距离比较结果

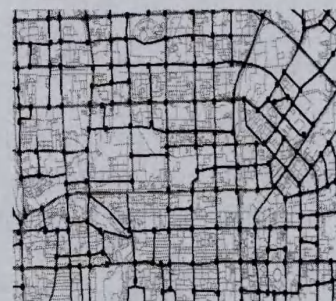
邻接点	1号路口 交叉点 (m)	2号路口 交叉点 (m)	3号路口 交叉点 (m)	4号路口 交叉点 (m)
所选点				
shp 地图 (116.306283, 39.865950)	852	1989	1083	1108
本文结果 (116.410529, 39.939632)	840	1955	1057	985
距离相差率	1.4%	1.7%	2.3%	11.1%
平均误差率	4.13%			

3.4.2 定性评估

利用 GPS 轨迹点自动生成道路中心线,使用 ArcGIS 软件得到坐标系统下的路网信息,将路口交叉点、路网提取结果与参考路网进行图层叠加对比分析。



(a)提取的节点与 Google Earth 的比较结果



(b)提取结果与 shp 地图的比较结果

图 10 实验结果与 Google Earth、shp 地图的比较结果

(下转第 55 页)

- [9] Forlines C, Balakrishnan R. Evaluating tactile feedback and direct vs. indirect stylus input in pointing and crossing selection tasks[C]//Proceedings of the 26th Annual ACM SIGCHI Conference on Human Factors in Computing Systems. Florence, Italy, 2008; 1563-1572
- [10] Forlines C, Vogel D, Kong N, et al. Absolute vs. relative direct pen input[R]. Mitsubishi Electric Research Labs Tech Report, 2006
- [11] Sears A, Shneiderman B. High precision touchscreens: design strategies and comparisons with a mouse[J]. International Journal of Man-Machine Studies, 1991, 34(4): 593-613
- [12] Fitzmaurice G W, Zhai S, Chignell M H. Virtual reality for palmtop computers[J]. ACM Transactions on Information Systems, 1993, 11(3): 197-218
- [13] Partridge K, Chatterjee S, Sazawal, et al. TiltType: accelerometer-supported text entry for very small devices[C]//Proceedings of the 15th Annual ACM Symposium on User Interface Software and Technology. Paris, France, 2002; 201-204
- [14] Wigdor D, Balakrishnan R. TiltText: using tilt for text input to mobile phones[C]//Proceedings of the 16th Annual ACM Symposium on User Interface Software and Technology. Vancouver,

- [15] Rahman M, Gustafson S, Irani P, et al. Tilt techniques: investigating the dexterity of wrist-based input[C]//Proceedings of the 27th Annual ACM SIGCHI Conference on Human Factors in Computing Systems. Boston, USA, 2009; 1943-1952
- [16] Leitner J, Powell J, Brandl P, et al. Flux: a tilting multi-touch and pen based surface[C]//Proceedings of the 27th Annual ACM SIGCHI Conference on Human Factors in Computing Systems. Boston, USA, 2009; 3211-3216
- [17] Tian F, Xu L, Wang H, et al. Tilt menu: using the 3D orientation information of pen devices to extend the selection capability of pen-based user interfaces[C]//Proceedings of the 26th Annual ACM SIGCHI Conference on Human Factors in Computing Systems. Florence, Italy, 2008; 1371-1380
- [18] Oshita M. Pen-to-mime: a pen - based interface for interactive control of a human figure[C]//Proceedings of the 1st EUROGRAPHICS Workshop on Sketch-Based Interfaces and Modeling. Grenoble, France, 2004; 43-52
- [19] Bi X, Moscovich T, Ramos G, et al. An exploration of pen rolling for pen-based interaction[C]//Proceedings of the 21th Annual ACM Symposium on User Interface Software and Technology. Monterey, USA, 2008; 191-200

(上接第 40 页)

图 10 是所得到的路网和路口交叉点的结果中随机选取的一个区域与 Google Earth 卫星地图、shp 地图的比较结果。从整体上看,本文提取的路网、路口交叉点与参考路网基本一致。

结束语 道路拓扑和网络距离计算是路径规划的前提,本文利用出租车 GPS 轨迹数据对北京市城区主干道进行了路网提取,然后提出了一个得到路口交叉点及道路交叉点之间连通性的方法。该方法基于特定对象或人物的出行轨迹,生成适合特定对象或人物出行的道路拓扑结构,从而可以进一步规划和优化其出行路径。实验结果表明,在道路拓扑生成效果方面,本文方法能有效提取各个道路交叉点,并能构建各点之间的拓扑关系,在复杂路网中提取道路交叉点的正确率达到了 87.08%,而且这些正确的交叉点位置与真实道路交叉点中心位置相差不大。

参 考 文 献

- [1] Haklay M, Weber P. OpenStreetMap: User-Generated Street Maps[J]. IEEE Pervasive Computing, 2008, 7(4): 12-18
- [2] 张小国,王庆,王宁,等. 电子地图道路网模型及其自动生成算法研究[J]. 中国图像图形学报, 2001, 6(5): 481-485
Zhang Xiao-guo, Wang Qing, Wang Ning, et al. A Study on Road Network Model in Digital Maps and the Automatic Generation Algorithm of its Database[J]. Journal of Image and Graphics, 2001, 6(5): 481-485
- [3] Stefan S, Kiri W, Seth R. Mining GPS Traces for Map Refinement[J]. Data Mining and Knowledge Discovery, 2004, 9(1): 59-87

- [4] Stefan E, Stefan S. Route Planning and Map Inference with Global Positioning Traces[J]. Computer Science in Perspective, 2003, 2598; 128-151
- [5] Alireza F, John K. Detecting Road Intersections from GPS Traces[J]. GIScience, 2010, 6292; 56-59
- [6] 欧阳鸿,刘建勋,刘毅志,等. 基于步行 GPS 轨迹的路网提取方法[J]. 计算机与现代化, 2014(2): 124-128
Ouyang Hong, Liu Jian-xun, Liu Yi-zhi, et al. An Extraction Method of Road Network Based on Walking GPS trajectories[J]. Computer and Modernization, 2014(2): 124-128
- [7] William A, Gordon A, et al. Automating the Detection and Simplification of Junctions in Road Networks[J]. GeoInformatica, 1999, 3(2): 185-200
- [8] 谢松云,张振中. 脑电信号的若干处理方法研究与评价[J]. 计算机仿真, 2007, 24(2): 326-330
Xie Song-yun, Zhang Zhen-zhong. Research and Evaluation on Some EEG Processing Methods[J]. Computer Simulation, 2007, 24(2): 326-330
- [9] Zhang L, Thiemann F, Sester M. Integration of GPS traces with road map[C]//Proceedings of the Second International Workshop on Computational Transportation Science, 2010. ACM, 2010; 17-22
- [10] 建设部主编. 城市道路交通规划设计规范[S]. 北京: 中国计划出版社, 2012; 7-8
The ministry of construction. Code for Transport Planning on Urban Road[S]. Beijing. China Planning Press, 2012; 7-8
- [11] Liu X, Zhu Y, Wang Y, et al. Road recognition using coarse-grained vehicular traces[R]. Technical Report HPL-2012-26, 2012