

利用地图栅格化的海量浮动车数据 道路匹配快速算法

李宇光¹ 李清泉^{2,1}

1 武汉大学测绘遥感信息工程国家重点实验室,湖北 武汉,430079

2 深圳大学空间信息智能感知与服务深圳市重点实验室,广东 深圳,518060

摘要:提出了在大城市路网环境下快速确定海量浮动车数据匹配路段的方法。首先构建路网道路缓冲区,再对道路缓冲区地图进行栅格化处理,并构建空间位置与道路ID的索引,然后基于每个浮动车数据中的地理位置信息依据索引找出浮动车数据可能的匹配道路,最后对这些道路进行匹配度计算,确定浮动车数据的匹配道路。实验表明,该方法能显著减少每个浮动车数据需要计算匹配度道路的数量,成倍地提高海量浮动车数据道路匹配算法的效率。

关键词:海量浮动车数据;道路匹配;地图栅格化;空间索引

中图法分类号:P208

文献标志码:A

智能交通系统是智慧城市的一个重要组成部分,道路交通信息的采集是智能交通系统的基础之一。浮动车数据指载有GPS设备的车辆在道路上采集的GPS数据,含有车辆行驶路段的交通信息^[1],这种数据能通过无线通讯系统快速发送到数据中心。大量的浮动车所提供的海量浮动车数据能全面、准确、快速、自动、全天时、全天候地提供车辆所在区域道路的交通信息,能精确反映出区域道路网交通信息的时空分布^[2],具有感应线圈、视频摄像等道路信息数据获取无法比拟的优点。目前,国内很多大城市的出租车上都装有GPS和无线通讯模块,浮动车数据已成为城市智能交通系统中道路动态交通信息的主要来源,对海量浮动车数据进行快速处理也成为智能交通系统的基本要求。

浮动车数据中的地理坐标由GPS获取,在收星状态良好的情况下,GPS数据存在20 m左右的平面定位误差^[3-5],而在城市环境中,GPS的收星状态还会受到高大密集建筑物、高架桥、立交桥等因素的影响,使GPS数据的定位误差更大,浮动车数据往往不能直接落在道路范围内,而且城市道路网非常密集、复杂,这就要求在对车载GPS数据应用前必须进行此类数据的地图匹配,

将车辆定位在正确的道路上^[6]。GPS数据的道路匹配研究很多都集中在车载自主导航系统方面,国内外有很多学者在这方面进行了研究,仅国外文献近些年来就提出了超过35种地图匹配的算法^[7],国内也有很多学者提出了多种GPS数据的地图匹配算法^[8-11]。这些算法主要针对自主导航需求或小规模浮动车数据,并且很多算法还包含航位推算数据,多数匹配算法仅对单车车辆的GPS数据进行了道路匹配,很少考虑海量浮动车数据道路匹配的效率问题。面向海量浮动车数据的处理,有学者提出了地图格网划分的算法^[12-13],这种算法与全区域道路匹配算法相比,有效地提高了海量浮动车数据道路匹配的效率,但仍不能满足道路交通信息准确、快速发布的要求。本文所提出的地图栅格化浮动车道路匹配算法的效率比地图格网化的更高,更适合处理海量浮动车数据。

1 浮动车数据道路匹配基本算法

浮动车数据道路匹配的一种基本方法是点到线段的匹配算法^[14],它计算一定范围内所有道路与当前浮动车数据的道路匹配度,是浮动车数据到道路的有效投影距离以及浮动车航向角与道路

收稿日期:2014-01-20

项目来源:深圳市科技研发资金资助项目(ZDSY20121019111146499);深圳市战略性新兴产业发展专项资金资助项目(JCYJ20121019111128765);国家863计划资助项目(2012AA101701)。

第一作者:李宇光,副教授。主要从事交通地理信息系统的理论与方法研究。E-mail:wd_ygli@sina.com

方向夹角的函数,即

$${}^kM_j^i = f({}^k d_j^i, {}^k \alpha_j^i) \quad (1)$$

式中, ${}^kM_j^i$ 表示第 k 个浮动车数据的位置到路网中第 i 条道路中第 j 个路段的匹配度; ${}^k d_j^i$ 表示第 k 个浮动车数据的位置到路网中第 i 条道路中第 j 个路段的有效投影距离; ${}^k \alpha_j^i$ 表示第 k 个浮动车数据位置的航向与路网中第 i 条道路中第 j 个路段的方向的夹角。对于第 k 个浮动车数据来说, 匹配度 ${}^kM_j^i$ 值最小时的道路 i 为该浮动车数据的匹配道路。

由于道路地图更新不够及时以及城市复杂环境(如高楼、高架桥)引起 GPS 数据出现粗差, 使得这种情况下的浮动车数据不能匹配到正确的路段上, 为保证浮动车数据匹配的准确性, 这些数据需要剔除。上述情况下, 这些浮动车数据与最近路段的有效投影距离往往都很大, 明显超过 GPS 数据位置的误差范围, 因此可以设定一个有效投影距离阈值 D_i , 该值的选取与 GPS 的定位精度以及道路的宽度有关。浮动车数据与路段的有效投影距离大于 D_i 时不能满足道路匹配条件, 为避免误匹配, 就不对该浮动车数据进行道路匹配。

如果在整个城市范围内对所有道路进行搜索, 完成浮动车数据道路匹配计算, 则每个浮动车数据需要遍历整个城市的所有道路, 从中计算出匹配度最小的道路作为浮动车数据的匹配道路, 这种方法对于海量浮动车数据处理的计算量非常大, 耗时很长。针对这个问题, 文献[12,13]提出将地图划分为若干个格网, 如 $450\text{ m} \times 250\text{ m}$ [13], 先确定浮动车数据所在的格网, 然后在当前或邻近的格网范围内做道路匹配计算。每个格网的东西、南北范围只有几百 m, 该范围内的道路数量平均只有几十条, 因此浮动车道路匹配计算的速度比在整个城市范围内进行遍历搜索计算的方法有显著提高。

通常, 在浮动车数据位置误差范围内平均只有几条道路, 只需对这几条道路进行道路匹配度计算, 而地图格网化方法平均需要计算几十条道路, 存在很多冗余的计算。对于每天上千万的浮动车数据的处理需求, 计算效率的提高意味着大量时间的节省。针对这一问题, 本文提出了一种道路地图栅格化方法, 该算法只对浮动车数据位置误差范围内的道路进行道路匹配度计算, 实现更高效的海量浮动车数据道路匹配计算。

2 地图栅格化道路匹配算法

地图栅格化道路匹配算法首先根据浮动车数

据位置误差和道路宽度信息构建道路缓冲区; 然后对道路缓冲区图进行栅格化处理, 每个栅格块的位置都对应一个地理坐标范围; 最后建立一对多的栅格索引, 即一个栅格块可以包含多条道路 ID。匹配计算时, 根据每个浮动车数据的位置坐标找到相应的栅格块, 根据其中的索引提取出道路 ID, 最后仅对索引中的道路进行匹配计算, 从而确定出该浮动车数据所匹配的道路。栅格块的范围可以划分的很小, 如 $2.5\text{ m} \times 2.5\text{ m}$, 这样栅格块中的路段就很少, 平均只有几条, 大大减少了待计算道路的数量, 从而能显著提高海量浮动车数据道路匹配计算的效率。

2.1 道路缓冲区

通常, 电子地图中的道路都采用矢量的形式表示 [15], 表示的道路中心线的信息、道路宽度信息作为道路的属性数据。同时, 浮动车数据的位置坐标存在误差, 因此几乎所有浮动车数据都不会位于道路中心线上, 而是落在道路中心线附近。仅对道路矢量地图进行栅格化处理, 可能会使大量栅格内无道路信息, 而位于栅格中的浮动车数据能够匹配到附近的道路, 如图 1 所示。而对道路缓冲区图进行栅格化处理就不会出现上述情况。如图 2 所示, 与浮动车数据可能匹配的道路都在栅格块内, 从而保障了浮动车数据道路匹配的可靠性。

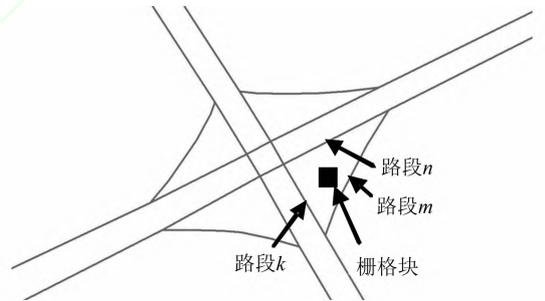


图 1 矢量道路地图

Fig. 1 Road Centre Line

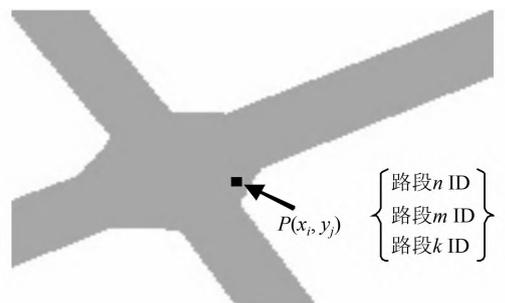


图 2 道路缓冲区栅格图

Fig. 2 Road Buffer Raster Map

道路缓冲区宽度如果设置的小,会导致一些浮动车数据不能匹配到道路上;如果过大,将导致索引数据量庞大,并产生不必要的冗余计算。为了保障浮动车数据道路匹配的可靠性,并尽可能地减少冗余计算,道路缓冲区构建示意图如图3所示,其半宽度 HW_b 等于有效投影距离阈值 D_t ,并可由下列公式算出:

$$HW_b = D_t = E_{\text{pfcd}} + W_{\text{link}}/2 \quad (2)$$

其中, HW_b 表示道路缓冲区的半宽度; E_{pfcd} 表示浮动车数据的平面定位精度; W_{link} 表示道路的宽度,从道路的属性值可以查到。

这样,当某个浮动车数据落在道路缓冲区内时,该条道路才有可能该浮动车数据的匹配道路,而落在道路缓冲区外的浮动车数据由于到这条道路的有效投影距离大于有效投影距离阈值 D_t ,因此不会匹配到这条道路上。

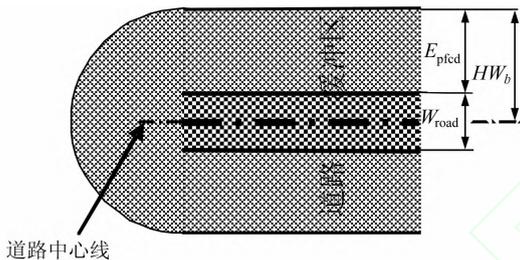


图3 道路缓冲区示意图

Fig. 3 Road Buffer Illustration

2.2 道路缓冲区地图栅格化

对道路缓冲区地图进行栅格化处理就是建立一个城市道路缓冲区的栅格地图,从图形上看与矢量图相似,但该图实际上是一个索引结构,图的像素行列位置与城市道路地理坐标对应,像素内容包含了所覆盖的道路信息。因此,栅格化的结果不是一个简单的道路缓冲区栅格图,而是一个含有有序的栅格行列位置和道路信息的数据结构,该数据结构由行数据和点数据构成。

行数据结构如下:

```
typedef struct _GRIDROW{
    int cnum;
    GRIDPOINT * gp;
}GRIDROW;
```

行数据中有两个变量,整型变量 $cnum$ 表示一行中有道路缓冲区栅格的数量,指针变量 $*gp$ 表示栅格像素数据信息,该变量的数量等于 $cnum$ 中的数值。

栅格像素数据结构如下:

```
typedef struct _GRIDPOINT{
    unsigned short row, col, n;
```

```
    unsigned int * rid;
```

```
}GRIDPOINT;
```

该数据结构中 16 位无符号变量 row 和 col 记录的是栅格所在的行、列信息,表示范围为 $[0, 65535]$; n 表示该像素所覆盖缓冲区道路的数量;无符号整型数组指针变量 $*rid$ 中存储的是道路的 ID。可以看出,道路地图栅格化后是一个行列有序的数据结构,先统计栅格图中每行是否有道路信息,再在每行中统计其中每列的道路信息,有道路信息的存入索引中。从图3可以看出,栅格图中大部分区域没有道路信息,因此只有 $n > 0$,即像素位置覆盖有道路缓冲区时,才能构建栅格像素数据结构,建立索引,以节省内存,因此这里的道路栅格图实际上是一个一对多的空间索引结构,是地理位置与道路 ID 的索引图。在该图中,其行信息的排列是连续、有序的,但由于存在大量无道路信息的栅格,所以列信息的排列是有序的但不连续,这样便于索引信息的快速查询。

2.3 浮动车数据道路匹配计算

有了道路地图栅格化的索引图后,对浮动车数据进行道路匹配。首先根据栅格道路地图的原点坐标对浮动车数据中的位置信息进行处理,计算出与道路地图栅格化后相对应的位置;然后根据浮动车的行位置直接查到栅格地图中对应的行数据,再在这一行数据中通过对半查找算法快速找出浮动车数据所对应的栅格块,如果所对应的栅格块索引信息存在,则从中提取出所有的道路 ID;如果不存在,则表示没有路段与该浮动车数据匹配;最后如果有道路 ID,就仅对这些道路进行浮动车数据的匹配度计算,确定该浮动车数据所匹配的路段,再进行下一个浮动车数据道路匹配计算。

由于是有序的索引结构,对半查找的效率最高,并且每行中有道路 ID 的列很少,而每个栅格块中的道路数量也很少,因此浮动车数据匹配计算的速度就很快。图4给出了一个基于道路地图栅格化方法实现海量浮动车数据道路匹配算法的流程图。

3 实验与分析

为了验证面向海量浮动车数据道路匹配处理的道路地图栅格化算法的有效性,采用了真实的数据进行实验。该路网数据是某公司 2005 年出版的导航地图数据,选取了以武汉市区为中心的一个东西长约 50 km(经度 $30'$)、南北宽约 44 km

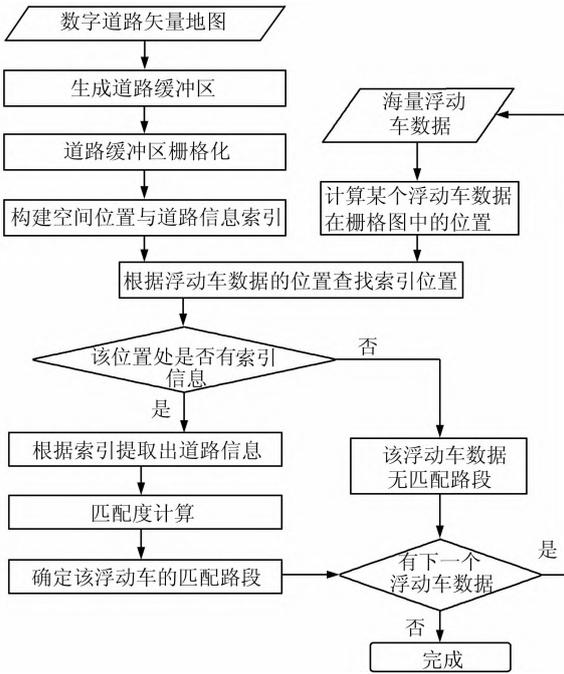


图 4 浮动车数据地图栅格化方法道路匹配算法流程图

Fig. 4 A Flowchart Representing FCD Map-matching Algorithm Based on Road Map to Raster Conversion

(纬度 25') 的约 2 200 km² 的区域。在这个区域的道路导航地图中有 26 436 条道路。浮动车数据采用了 2008 年 12 月 28 日武汉市 11 354 辆出租车全天所记录的浮动车数据, 共有 13 874 689 个。计算机环境的硬件是个人台式计算机, CPU 是英特尔 E6400 双核处理器, 主频为 2.13 GHz, 内存 3 GB; 操作系统是微软 Windows XP SP3, 编程环境为微软 Visual C++ 6.0。

在道路地图栅格化过程中, 浮动车数据的位置误差极限为 20 m, 道路宽度从导航电子地图中道路的属性中获取。道路栅格地图的分辨率为 0.000 025° (约为 2.5 m), 路网栅格化后大小为 20 000 × 16 667。

计算表明, 采用道路地图栅格化算法得出的浮动车数据道路匹配的结果与其他匹配方法的结果完全一致, 没有出现漏匹配和误匹配的情况。该算法完成所有 13 874 689 个浮动车数据道路匹配所消耗的时间为 85 s, 其中地图栅格化用了 32 s, 而匹配计算仅用了 53 s, 平均每个浮动车数据需要判断、计算的道路数量为 2.87 条。表 1 列出了三种海量浮动车数据道路匹配算法的比较情况。

从表 1 中可以看出, 如果处理上千万条浮动车数据, 地图栅格化算法的耗时最短, 索引建立完成后, 平均每 s 可以匹配 261 669.6 个浮动车数

据, 这是由于每个浮动车数据平均只需对 2.87 条道路进行匹配计算, 远少于其他算法需要计算的道路的数量。但栅格化算法构建整个城市路网地图索引所需的时间较长, 适用于连续处理海量浮动车数据; 栅格化索引所需的计算机内存也较大, 但目前 64 位操作系统和编译软件已经成熟, 计算机内存的容量越来越大, 在个人计算机上 8 G、16 G 内存的配置已很普遍, 并支持 32 G 内存, 因此用道路地图栅格化处理任意一个城市的路网数据在个人计算机上就能实现。

表 1 海量浮动车数据不同算法道路匹配耗时比较

Tab. 1 Performance to Map-Matching Huge-Volume FCD

算法	资源消耗		平均涉及的道路数量/条	计算效率/(匹配点数 · s ⁻¹)	
	索引内存	建索引/s			
全区域	0	0	34 173	26 436	405.8
格网划分	1.2 MB	2	273	20.7	50 800
栅格地图	1.7 GB	32	53	2.87	261 669.6

4 结 语

本文提出了一种面向海量浮动车数据的道路地图栅格化的浮动车数据道路匹配算法。实验验证, 在保证匹配结果一致性的情况下, 该方法能够显著地提高匹配算法的效率, 提高了利用海量浮动车数据获取城市道路交通信息的实效性, 对实现智能交通系统有现实意义。

本文采用的道路匹配算法是基于点到线段匹配度最大的方法, 这种方法在城市复杂道路环境下的匹配结果会存在一些误差。后续研究工作将在本算法的基础上对海量浮动车数据道路匹配的准确性、可靠性进行研究。

参 考 文 献

- [1] Brakatsoulas S, Pfoser D, Salas R, et al. On Map-Matching Vehicle Tracking Data[C]. The 31st VLDB Conference, Trondheim, Norway, 2005
- [2] Kuehne R, Schaefer R P, Mikat J, et al. New Approaches for Traffic Management in Metropolitan Areas[C]. IFAC CTS Symposium, Tokyo, Japan, 2003
- [3] US DoD. Global Positioning System Standard Positioning Service Performance Standard[M]. United States: Assistant Secretary of Defense for Command, Control, Communications, and Intelligence, 2001
- [4] Zhang Lina, Yang Zhiqiang, Ou Long, et al. A

- Study of PDA GPS's Location Error[J]. *Chinese Journal of Engineering Geophysics*, 2006, 3(6): 478-483 (张丽娜, 杨志强, 欧龙, 等. 手持式 GPS 定位误差的研究[J]. 工程地球物理学报, 2006, 3(6): 478-483)
- [5] Cheng Xinwen, Chen Xingyi. A Study of PDA GPS's Location Precision[J]. *Bulletin of Surveying and Mapping*, 2004(9): 20-22 (程新文, 陈性义. 手持式 GPS 定位精度研究[J]. 测绘通报, 2004(9): 20-22)
- [6] Greenfeld J S. Matching GPS Observations to Location on a Digital Map[C]. The 81st Annual Meeting of the Transportation Research Board, Washington D C, 2002
- [7] Quddus M A, Ochieng W Y, Noland R B. Current Map-matching Algorithms for Transport Applications; State-of-the Art and Future Research Directions[J]. *Transportation Research Part C*, 2007(15): 312-328
- [8] Su Jie, Zhou Dongfang, Yue Chunsheng. Real-time Map-matching Algorithm in GPS Navigation System for Vehicles[J]. *Acta Geodaetica et Cartographica Sinica*, 2001, 30(3): 252-256 (苏洁, 周东方, 岳春生. GPS 车辆导航中的实时地图匹配算法[J]. 测绘学报, 2001, 30(3): 252-256)
- [9] Qin Ling, Zhang Jianfei, Guo Peng, et al. Study of Key Technologies and Applications of Floating Car Traffic Information Collection and Processing[J]. *Journal of Transportation Systems Engineering and Information Technology*, 2007, 7(1): 39-42 (秦玲, 张剑飞, 郭鹏, 等. 浮动车交通信息采集与处理关键技术及其应用研究[J]. 交通运输系统工程与信息, 2007, 7(1): 39-42)
- [10] Ma Jianjun, Tang Jinjun, Cao Kai. New Intelligent Map-matching Algorithm[J]. *Computer Applications*, 2007, 27(12): 3116-3118 (马建军, 唐进君, 曹凯. 一种新的智能地图匹配算法[J]. 计算机应用, 2007, 27(12): 3116-3118)
- [11] Wang Dongzhu, Dong Jiming, Li Yameng, et al. Map-matching Method Based on Zero-speed Points in Floating Car Data[J]. *Journal of Transport Information and Safety*, 2009, 27(6): 38-42 (王东柱, 董继明, 李亚檬, 等. 浮动车数据中零速度点数据地图匹配方法[J]. 交通信息与安全, 2009, 27(6): 38-42)
- [12] Zhang Wei, Xu Jianmin, Lin Mianfeng. Map Matching Algorithm of Large Scale Probe Vehicle Data [J]. *Journal of Transportation Systems Engineering and Information Technology*, 2007, 7(2): 39-45 (章威, 徐建闽, 林绵峰. 基于大规模浮动车数据的地图匹配算法[J]. 交通运输系统工程与信息, 2007, 7(2): 39-45)
- [13] Wang Zuyun, Du Yong, Wang Gang, et al. A Quick Map-matching Algorithm by Using Grid-based Selecting[C]. 2008 International Workshop on Education Technology and Training & 2008 International Workshop on Geoscience and Remote Sensing, Shanghai, China, 2008
- [14] White C E, Bernstein D, Kornhauser A L. Some Map-matching Algorithms for Personal Navigation Assistants[J]. *Transportation Research Part C*, 2000(8): 91-108
- [15] Li Yuguang, Li Lianying, Li Qingquan, et al. Vector Map Geometric Data Change Detection Based on Grid Method[J]. *Geospatial Information*, 2010, 8(1): 142-145 (李宇光, 李连营, 李清泉, 等. 基于栅格化思想的矢量电子地图几何变化检测[J]. 地理空间信息, 2010, 8(1): 142-145)

A Fast Algorithm for Huge Volume Floating Car Data Map-Matching: A Vector to Raster Map Conversion Approach

LI Yuguang¹ LI Qingquan²

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan 430079, China

2 Shenzhen Key Laboratory of Spatial Smart Sensing and Services, Shenzhen University, Shenzhen 518060, China

Abstract: We propose a fast algorithm for matching huge volume floating car data (FCD) with maps of urban areas. Firstly, a road buffer is constructed. Secondly, all links buffer are converted to raster maps and a spatial index map between the geographical coordinate and link IDs is built. Then, the candidate links are selected based on the coordinates of FCD and the spatial index. Finally, the map-

(下转第 733 页)

Automatic Topology Maintenance Approach for High-level Road Networks

LI Fei^{1,2,3} LUAN Xuechen³ YANG Bisheng^{1,2} LI Qiuping⁴

1 State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing,
Wuhan University, Wuhan 430079, China

2 Engineering Research Center for Spatio-temporal Data Smart Acquisition and Application,
Ministry of Education of China, Wuhan University, Wuhan 430079, China

3 Guangdong Ritu Information Systems Co., Ltd., Foshan 528305, China

4 Center of Integrated Geographic Information Analysis, School of Geography and Planning, Sun Yat-sen University,
Guangzhou 510275, China

Abstract: This paper presents an automatic topology maintenance approach for any extracted high-level road network. Firstly, the minimum spanning tree (MST) is generated on dual graph of road networks for maintaining the whole connectivity relationship. Secondly, the breadth-first search (BFS) and shortest path algorithm are applied to connect the dangling roads and rational roads, which are used to maintain the perceptual connecting relationship of the high-level road networks. An experiment was executed to verify the effectiveness of the approach. The experimental results show that the method achieves good and effective topology maintenance results, reflecting the structural characteristic of a high-level network. Furthermore, a comparison analysis to original high-level roads before and after topology maintenance was also carried out to show the rationalities of navigation path planning.

Key words: road networks; topology maintenance; road hierarchies; navigation route planning

First author: LI Fei, PhD, specializes in navigation, LBS and intelligent transportation system. E-mail: lifei@ritu.cn

Corresponding author: YANG Bisheng, PhD, professor. E-mail: bshyang@whu.edu.cn

Foundation support: The National 863 Program of China, Nos. 2012AA12A211, 2012AA12A204; Guangdong Special Fund for Strategic Emerging Industrial Development (High-end New Electronic Information) Project, No. 2011168036.

(上接第 728 页)

matching degree is calculated between the FCD and the candidate links. The map-matching links are identified in the results. Experiments show that the proposed method reduces the number of links that need be calculated for every FCD, thus resulting in efficiency improvement of the proposed map-matching algorithm.

Key words: huge volume floating car data; map-matching; raster map conversion; spatial index

First author: LI Yuguang, associate professor, specializes in the theories and methods of Geography Information System -Transportation. E-mail: wd_ygli@sina.com

Foundation support: Shenzhen Scientific Research and Development Funding Program, No. ZDSY20121019111146499; Shenzhen Dedicated Funding of Strategic Emerging Industry Development Program, No. JCYJ20121019111128765; The National 863 Program of China, No. 2012AA101701.