

引用格式:傅梦颖,张恒才,王培晓,等.基于移动对象轨迹的室内导航网络构建方法[J].地球信息科学学报,2019,21(5):631-640. [Fu M Y, Zhang H C, Wang P X, et al. A method for constructing indoor navigation network based on moving object trajectory[J]. Journal of Geo-information Science, 2019,21(5):631-640.] DOI:10.12082/dqxxkx.2019.190024

基于移动对象轨迹的室内导航网络构建方法

傅梦颖^{1,2}, 张恒才^{2,3*}, 王培晓^{1,2}, 吴升^{1,2}, 陆锋^{2,3}

1. 福州大学, 数字中国研究院(福建), 福州 350002; 2. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 北京 100101; 3. 海西政务大数据应用协同创新中心, 福州 350002;

A Method for Constructing Indoor Navigation Networks based on Moving Object Trajectory

FU Mengying^{1,2}, ZHANG Hengcai^{2,3*}, WANG Peixiao^{1,2}, WU Sheng^{1,2}, LU Feng^{2,3}

1. The Academy of Digital China, Fuzhou University, Fuzhou 350002, China; 2. State Key Lab of Resources and Environmental Information System, Institute of Geographic Sciences and Natural Resources Research, Chinese Academy of Science, Beijing 100101, China; 3. Fujian Collaborative Innovation Center for Big Data Applications in Governments, Fuzhou 350002, China

Abstract: The indoor navigation network is the basis for pedestrian navigation, information recommendation, and business analysis. The traditional method of manual mapping or semiautomatic extraction of three-dimensional indoor navigation network cannot meet the requirement of high-frequency change of complex indoor space structures. With the continuous development of indoor positioning technology, there is an explosion of trajectory data of indoor moving objects, which provides a possibility for rapid construction and change monitoring of indoor navigation networks. This paper proposes a method of crowdsourcing construction of indoor navigation network based on the trajectory of moving objects. Based on trajectory simplification preprocessing using ST-DBSCAN, an indoor trajectory adaptive rasterization algorithm is proposed to reduce the influence of raster image resolution on the extraction of navigation networks. This approach effectively avoids the failure of navigation networks' topological connection that is caused by the difference of track trajectory density. Moreover, it automatically identifies the connection points between floors by the CFSFDP adaptive clustering algorithm to realize the rapid construction of indoor navigation networks. The experimental data is derived from the real indoor moving object trajectory data provided by Shanghai Palmap Science & Technology Co., Ltd. The experimental results show that, compared with the universal rasterization method, the proposed method improves the accuracy of navigation network construction by an average of 2.43% and improves the accuracy of topology by 12.8%.

Key words: indoor navigation network; trajectory of moving objects; trajectory waypoint; adaptive rasterization; CFSFDP algorithm

*Corresponding author: ZHANG Hengcai, E-mail: zhanghc@lreis.ac.cn

收稿日期:2019-01-16;修回日期:2019-03-11.

基金项目:国家自然科学基金项目(41771436);国家重点研发计划项目(2016YFB0502104、2017YFB0503500);数字福建建设项目(闽发改网数字函[2016]23号)。[**Foundation items:** National Natural Science Foundation of China, No.41771436; National Key Research and Development Program of China, No.2016YFB0502104, 2017YFB0503500; Digital Fujian Program, No.2016-23.]

作者简介:傅梦颖(1993-),女,安徽六安人,硕士生,研究方向为地理信息服务。E-mail: 1429137226@qq.com

*通讯作者:张恒才(1985-),山东济南人,助理研究员,研究方向为移动对象轨迹数据管理与分析。E-mail: zhanghc@lreis.ac.cn

摘要:室内导航网络是行人导航、信息推荐和商业分析的基础。传统人工测绘或半自动提取的室内三维导航网络无法满足复杂室内空间结构高频变化需求。随着室内定位技术的不断发展,室内移动对象轨迹数据爆发式增长,为室内导航网络快速构建与变化监测更新提供了可能。本文提出一种基于移动对象轨迹的室内导航网络构建方法,在基于ST-DBSCAN的轨迹简化预处理基础上,提出了室内轨迹自适应栅格化算法,减弱栅格图像分辨率对导航网络提取的影响,有效避免廊道轨迹密度差异造成的导航网络拓扑连通失效,并通过CFSFDP自适应聚类算法自动识别楼层之间连通点,实现室内导航网络的快速构建。实验数据来源于上海图聚智能科技有限公司提供的某商城真实的室内移动对象轨迹数据,实验结果表明,与普适栅格化方法相比,本文提出的方法将导航网络构建准确率平均提高2.43%,拓扑正确度提高12.8%。

关键词:室内导航网络;移动对象轨迹;轨迹停留点;自适应栅格化;CFSFDP算法

1 引言

室内空间是人类活动的主要空间,例如办公楼、购物中心、医院、机场、地铁站等,据研究表明人类87%左右的时间都在室内空间移动^[1]。室内导航网络是行人导航、个性化信息推荐和商业分析的基础^[2-3]。与交通路网相比,室内空间更为复杂,导航网络提取难度较大,一方面,室内空间属于三维空间,室内三维空间结构复杂,如封闭空间、半封闭空间等,室内三维空间实体众多,如房间、墙壁、门、窗、走廊、电梯、楼梯等,室内三维空间约束多样,如连通约束、障碍约束等,三维拓扑连通关系构建是室内导航网络构建的难点之一;另一方面,室外路网结构相对固定,室内廊道结构狭窄多变,且在一些特定大型购物广场类室内空间,廊道结构变化频率较高。室内导航网络提取多采用人工现场测绘或半人工CAD平面图提取^[4-5]等方法,虽然该类方法可以保证导航路网提取精度,但室内导航网络时效性无法满足要求,进而影响室内导航网络应用的准确性构建。

近年来,随着室内定位技术迅猛发展,如WiFi定位、射频识别(Radio Frequency Identification, RFID)定位、蓝牙或NFC(Near Field Communication)定位、伪卫星定位、ZigBee定位、UWB(Ultra-Wideband)定位、超声波定位、影像匹配与条码定位及地磁定位等,内置定位模块的移动终端如智能手机、平板电脑、PDA等的用户规模不断扩大,移动互联网的不断发展,室内位置服务应用不断增多,如在线导航、基于位置的社交网络、基于位置的广告推送^[9]等,室内空间产生了海量移动对象轨迹数据,为室内导航网络快速构建与变化检测更新提供了一种新的可能。

目前,基于移动对象轨迹数据构建导航网络的研究多集中在室外空间,依据方法模型的不同,分为轨迹聚类法、增量融合法、节点连接法、栅格化法

等。轨迹聚类法^[6-7]是基于同一路段轨迹具有的空间相似性和邻近性,运用聚类算法对轨迹点、线聚类,实现路网提取。增量融合法^[8-10]首先初始化空白地图,逐次将车辆轨迹线添加进空的道路地图上,逐步构造和细化道路图。节点连接法^[11]的关键在于交叉口识别^[12],通常基于车辆轨迹在转弯处产生的速度和方向变化特征识别交叉口位置,通过节点连接形成路网地图。栅格化法^[13-15]首先将轨迹点转换为二值化栅格图像,然后运用图像细化算法提取道路骨架线,最后生成矢量路网地图。

室内移动对象轨迹产生于模糊的自由空间(包括公共廊道、房间等),且室内移动对象具有移动随意性,对基于移动对象轨迹构建室内导航网络提出了很大挑战。鉴于此,本文提出一种室内导航网络众包构建方法,首先基于轨迹速度信息和ST-DBSCAN对移动对象轨迹进行简化,去除轨迹停留点信息;接着,通过改进KNN室内轨迹自适应栅格化算法生成室内轨迹图像,并采用数学形态学优化和细化方法,实现室内二维平面导航网络构建;最后,通过拓扑连通区域识别规则和CFSFDP(Clustering by Fast Search and Find of Density Peaks)自适应聚类算法识别出楼层连通点,实现楼层连接并与平面导航网络融合,实现室内三维导航网络构建。

2 研究方法

基于移动对象轨迹构建室内导航网络的算法流程(图1)为:①室内轨迹预处理,根据轨迹速度信息及ST-DBSCAN算法简化轨迹停留点信息;②室内二维平面导航网络构建,包括室内轨迹去噪、室内轨迹图像生成、室内轨迹图像优化处理、室内平面二维导航网络提取4个步骤;③室内三维导航网络构建,根据室内三维拓扑连通区域识别规则和CFSFDP自适应聚类算法提取出楼层连通点,实现室内三维导航网络自动构建。

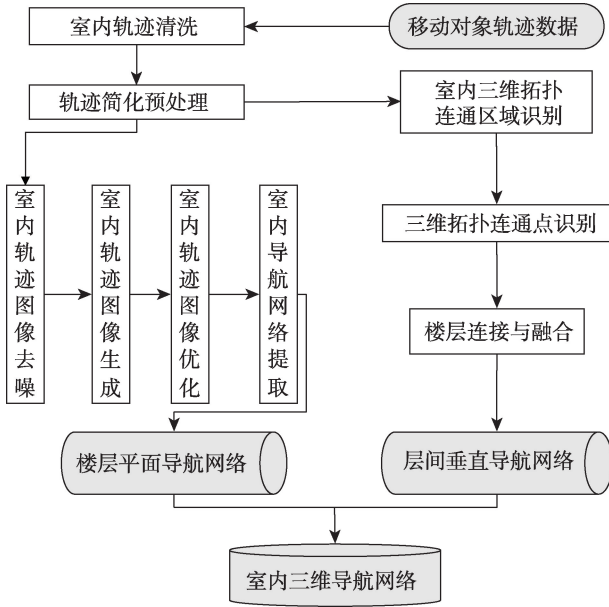


图1 基于移动对象轨迹的室内导航网络构建方法

Fig. 1 Method for constructing indoor navigation networks based on moving object trajectory

2.1 室内轨迹预处理

室内移动对象轨迹包含较多停留点信息,主要分为2种情形^[16]:①移动对象在同一位置停留较长时间,其轨迹会出现大量时间连续且位置不变的轨迹点;②移动对象在一个位置及其周围游荡,其轨迹会出现小范围聚集,形成高密轨迹点簇。本文提出了基于ST-DBSCAN和轨迹速度信息的轨迹简化预处理方法,具体步骤为:首先利用ST-DBSCAN算法^[17]识别出每条轨迹连续时间序列下形成的高密轨迹点簇,并以簇集质心 C_i 取代该簇集中的所有轨迹点 $\{P_i\}_{i=1}^n$ 。如图2(a)所示,轨迹点 $\{p_1, p_2, \dots, p_7\}$ 形成一个停留簇, p_9 虽位置邻近,但超过了时间阈值 ϵ_t , 因此不参与聚类;然后,以质心点 C_i 取代 $\{p_1, p_2, \dots, p_7\}$, 该处理方法在去除轨迹停留点信息的同时避免了因轨迹点去除而导致的坐标和时间

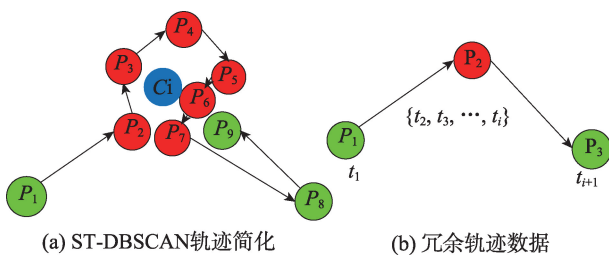


图2 室内轨迹预处理

Fig. 2 Indoor trajectory preprocessing

信息缺失;最后,合并用户轨迹中连续时间下速度为0的轨迹点,只保留一个。如图2(b)所示,某用户在连续时间序列 $\{t_2, t_3, \dots, t_i\}$ 下产生了 i 个坐标一致的 p_2 点,则保留一个 p_2 。

$$C_i(X, Y, T) = \left(\frac{1}{n} \sum_{i=0}^n (x_i), \frac{1}{n} \sum_{i=0}^n (y_i), \frac{1}{n} \sum_{i=0}^n (t_i) \right) \quad (1)$$

2.2 二维平面室内导航网络构建

2.2.1 室内轨迹图像生成算法

经过ST-DBSCAN轨迹简化后轨迹点更集中于廊道,但仍残留不少噪声点,在生成轨迹图像之前首先需要进一步过滤噪声,这里通过栅格过滤方式进一步去噪。

以点 (x_{\min}, y_{\max}) 为栅格图像的原点对栅格图像进行划分,统计所有栅格对应的轨迹点数 N_{ij} , 设置一个统计阈值 rep_{thr} , 将值小于 rep_{thr} 的像元滤除。噪声的过滤程度取决于轨迹栅格图像分辨率的选择,分辨率过大,容易造成廊道点被错误滤除,分辨率过小,容易噪声较多残留。图像分辨率至多应为最小道路宽度的一半^[18], 结合噪声尽可能去除且廊道区域轨迹点尽可能保留的原则,经试验将图像分辨率确定为1 m。接下来只要确定出剩下像元对应的像元值即可得到室内轨迹栅格图像。

本文提出一种改进KNN室内轨迹自适应栅格化算法,根据栅格邻域内点的个数和密度,自动调节栅格化时统计阈值,克服不同廊道区域轨迹密度差异所导致的提取误差,增强轨迹稀疏区域的拓扑连通性,并有效克服噪声的影响。遍历所有栅格像元 (i, j) , 记录所有像元值为0的像元,统计其8邻域像元的像元值,将8像元均值赋给中心像元 (i, j) , 计算过程如式(2)所示。为了克服噪声的影响,防止噪声像元被扩大,则对重新赋值的像元设置限制条件:当中心像元8邻域像元中像元值不为0的像元个数 Np_{thr} 至少为 n ($0 < n < 8$) 个时才对中心像元重新赋值, n 越大,噪声的过滤效果越好,但像元被重新赋值的概率降低,影响道路连通效果。本文将 Np_{thr} 分别设为3和2,得到的实验结果如图3(a)、(b)所示, Np_{thr} 为2时道路连通效果更好(如图中红框区域所示),但一些噪声也不可避免被扩大(如图中绿框区域所示),本文在实验中将 Np_{thr} 设为3,该阈值下产生的少数未完全连道路后续可通过图像优化算法进一步修复。室内轨迹图像生成伪代码如图4所示。

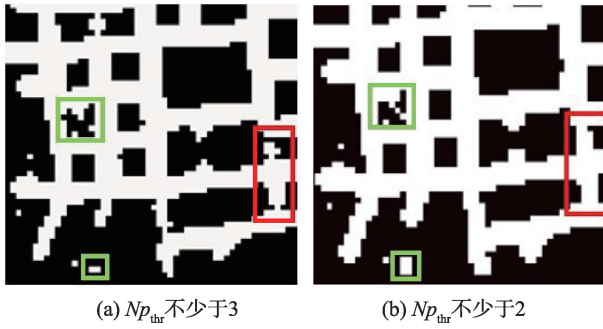


图3 不同阈值下的自适应栅格化结果

Fig. 3 Adaptive rasterization results under different thresholds

算法 室内轨迹图像生成算法

输入: 化简后轨迹 $pts = \{p_i\}$, 分辨率 $resolution$, 像元阈值 rep_{threh} , 邻域 r , 邻域像元个数 n

输出: 中心点坐标矩阵 $centerMat$, 密度矩阵 $densityMat$

```

1: //获取轨迹点的边界范围
2:  $extent = getExtent(pts)$ 
3: // 计算中心点坐标矩阵
4:  $centerMat = computerCenterMat(extent, resolution)$ 
5: // 获取矩阵的行和列
6: for  $i = 0 \rightarrow m$  do
7:   for  $j = 0 \rightarrow n$  do
8:     //计算中心点8邻域内轨迹点密度
9:      $densityList = computerDensity(pts, centerMat[i][j], r)$ 
10:    if  $sum(densityList) < rep_{threh}$  then
11:       $densityMat[i][j] = 0$ 
12:    end if
13:    if  $sum(isZero(densityList)) \geq n$  then
14:       $densityMat[i][j] = average(densityList)$ 
15:    end if
16:  end for
17: end for
18: return  $centerMat, densityMat$ 

```

图4 室内轨迹图像生成算法

Fig. 4 Algorithm for generating indoor trajectory images

$$N'_{ij} = \frac{1}{8} \left(\sum_{\substack{i-1 \leq m \leq i+1; \\ j-1 \leq n \leq j+1}} N_{ij} - N_{ij} \right) \quad (2)$$

式中: N_{ij} 表示像元 (i, j) 的像元值; N'_{ij} 表示自适应栅格化重新赋值后的像元 (i, j) 的像元值。

图5为室内轨迹图像生成过程示意图, 像元 (i, j) 在设定的栅格图像分辨率下的像元值 N_{ij} 为0, 其8邻域像元中有3个像元值不为0, 那么 N_{ij} 由其8邻域像元的值重新赋值。而像元 (m, n) 的8邻域像元中只有一个像元 $(m-1, n)$ 的像元值不为0, 那么将其判定为噪声像元, 不予处理。

邻域像元构成的滑动窗口可根据拓扑不连通

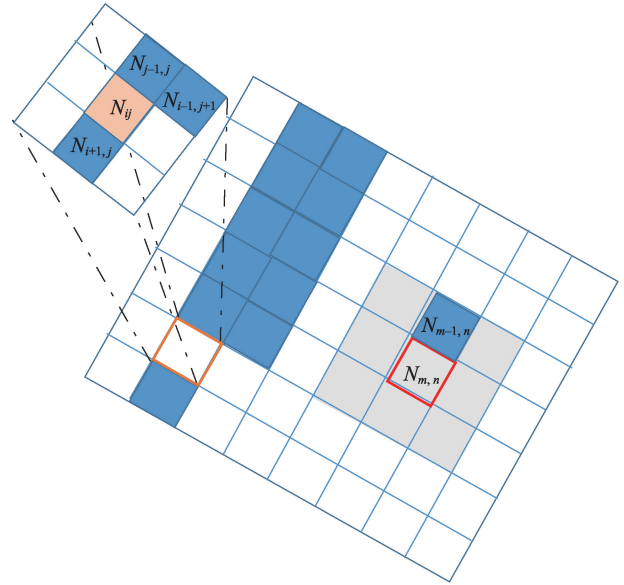


图5 自适应栅格化算法示意

Fig. 5 Schematic diagram of the adaptive rasterization algorithm

区域的面积大小自由调整, 窗口边长不应超过研究区域最大路宽。若噪声多表现为连续的噪声像元, 则可以增加邻域像元中不为0的像元个数, 降低噪声像元被处理的概率。这种栅格化方法能够减弱栅格图像分辨率的选择对路网提取结果的影响, 能够有效避免廊道轨迹密度差异而导致导航网络拓扑不连通的问题, 同时能够较好的克服噪声的影响。

轨迹数据自适应栅格化后需要进行二值化, 得到二值化室内轨迹栅格图像。设定二值化阈值 th_{value} , 若像元 (i, j) 的像元值大于 th_{value} , 像元 (i, j) 的像元值重新赋为1, 反之, 赋为0。 th_{value} 设置的越大, 得到的像元为廊道道路的可靠性越大, 但是也容易导致路面像元过于稀疏。因此, th_{value} 需要根据自适应栅格化重新赋值后的像元值决定。

2.2.2 室内轨迹图像优化算法

改进KNN自适应栅格化得到的室内轨迹图像连通性较好, 但仍存在一些残留噪声、道路轻微裂隙、边缘锯齿等现象, 这些会影响室内导航网络提取精度, 需要对图像进行优化。本文采用数学形态学基本运算对室内轨迹图像进行优化, 包括腐蚀、膨胀、开运算、闭运算等, 在室内轨迹图像优化处理过程中通过结合多种运算以达到理想效果。

(1) 去除噪声。噪声在室内轨迹图像里表现为大小不等的图斑, 有的独立存在, 有的与道路紧密粘连。去噪采用形态学腐蚀操作(式(3))实现, 本文采用图6(a)所示的结构元素, 记为B, 其操作过

程如下:在B中定义一个原点 (a, b) ,以B为滑动窗口对图像逐步扫描,如图6(b)所示,当原点 (a, b) 移动到A的某像元 (i, j) 处时,若B中的元素值与A中的元素值完全相同,则将像元 (i, j) 赋值为1,否则赋为0。图6(c)为图像腐蚀运算后的结果,可见经腐蚀运算后大多数像元的值由1变为了0。结构元素的大小根据最大噪声图斑尺寸进行调整,使得在保证拓扑连通性的情况下噪声尽可能被去除,如图7(a)所示。

$$A \ominus B = \{x, y | (B)_{xy} \subseteq A\} \quad (3)$$

(2)填补裂隙。自适应栅格化得到的室内轨迹图像可能仍存在少量道路裂隙,可采用形态学膨胀运算进行裂隙填补。如式4所示,其原理是当B中原点 (a, b) 移动到A的某像元 (i, j) 处时,若B中的元素值与A中的元素值至少有一个相同时,则将像元 (i, j) 赋值为1,否则赋为0。图6(d)为膨胀运算后的结果,可以看出原本大量值为0的像元变为了1。单纯的膨胀操作会导致噪声也被扩大,因此本文结合了开运算和闭运算同时对图像进行处理,开

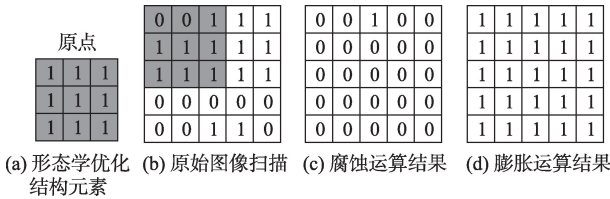


图6 形态学优化过程

Fig. 6 Morphological optimization processes

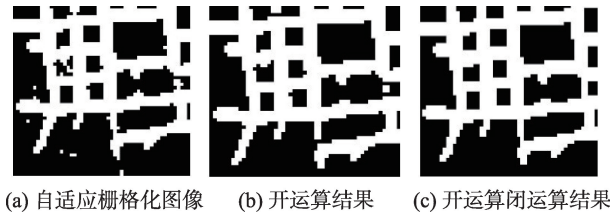


图7 室内轨迹图像优化

Fig. 7 Optimization of indoor trajectory images

0	0	*
0	1	1
*	1	1
B^1		
*	0	0
1	1	0
1	1	*
B^2		
1	1	*
1	1	0
*	0	0
B^3		
*	1	1
0	1	1
0	0	*
B^4		
0	0	0
*	1	*
1	1	1
B^5		
1	*	0
1	1	0
1	*	0
B^6		
1	1	1
*	1	*
0	0	0
B^7		
0	*	1
0	1	1
0	*	1
B^8		

图8 形态学细化采用的结构元素序列

Fig. 8 Sequences of structural elements used in morphological refinement

运算是先对图像执行腐蚀操作然后用同一结构元素膨胀其结果,闭运算反之。由于室内轨迹图像仍存在不少噪声,且部分与道路紧密粘连,因此,本文对室内轨迹图像执行先开运算后闭运算处理,图7(b)是开运算的结果,可以看出,噪声基本被去除但部分道路出现了轻微裂隙,对其结果再作闭运算,结果如图7(c)所示,裂隙基本被填补。

$$A \oplus B = \{x, y | (B)_{xy} \cap A \neq \emptyset\} \quad (4)$$

2.2.3 室内平面导航网络提取算法

室内平面导航网络提取分为2步:①对优化处理后的轨迹图像进行细化,即在保持图像的形状、拓扑结构不变的情况下,提取出图像的中心像元;②将细化结果转换成矢量数据,构成包含路段和结点信息的室内导航网络。

本文采用数学形态学细化算法对室内轨迹图像进行细化。数学形态学细化算法利用击中击不中变换,通过一系列结构元素对图像迭代细化,不断删除被击中的像素,直到结果不再变化,使用结构元素B对图像A细化处理定义为:

$$A \otimes B = A - (A \ominus B) = A \cap (A \ominus B)^c \quad (5)$$

为保证细化结果的对称性,结构元素序列中应包含多种方向的结构元素。本文采用常用的结构元素集合 $B = \{B^1, B^2, \dots, B^n\}$,如图8所示,其中 B^i 是 B^{i-1} 的旋转,结构对序列 $\{B^1, B^2, B^3, B^4\}$ 则用来消去西北、东北、东南和西南4个方向上的像素点,结构对序列 $\{B^5, B^6, B^7, B^8\}$ 用来消去北、东、南、西4方向上的点。“1”表示目标图像上的像素点,“0”表示背景像素点,“*”既可以是目标像素点,也可以是背景像素点。使用序列B对图像A细化处理定义为:

$$A \otimes \{B\} = \left(\left(\left(\left(A \otimes B^1 \right) \otimes B^2 \right) \dots \right) \otimes B^n \right) \quad (6)$$

首先使用 B^1 细化A,再用 B^2 细化上一步的结果,依次进行上述过程直到 B^n 细化A结束,一轮细化完成。不断重复上述过程,直到结果不再发生变化。

最后将室内轨迹图像细化结果转换成矢量数

据就得到了室内导航网络,矢量化过程中需要考虑细化结果残留噪声、细化后的道路细碎分支、矢量道路平滑等,实现噪声和细碎道路过滤、路段平滑等。

2.3 三维室内导航网络快速构建

本文通过移动对象轨迹数据自动识别出楼层连通点,实现楼层自动连接,生成室内三维导航网络。首先基于化简后的用户轨迹识别出潜在三维拓扑连通区域,其次,从潜在连通区域识别出拓扑连通点,最后以连通点连接各楼层。具体步骤如下:

(1) 连通区域识别步骤如下:识别出每条原始轨迹中楼层属性 f_i 不一致且定位时间差小于时间阈值 T_{thr} 的多对轨迹点 $\{\dots, P_i, P_j, \dots | t_j - t_i < T_{thr}, f_i \neq f_j\}$, $P_i = (ID, x, y, t, f)$, $i \in [1, n]$;

如某对轨迹点的位置差在设定的距离阈值 dis_{thr} 范围内,即 $\sqrt{(x_j - x_i)^2 + (y_j - y_i)^2} < dis_{thr}$, 则将该对轨迹点存入楼层连通区域数据集中 $\{P_{ij}\}_{i=1}^n$;

(2) 从识别出的楼层连通区域中识别出 $\{P_{ij}\}_{i=1}^n$ 中的楼层连通点 $\{Q_i\}_{i=1}^m$;

(3) 以位置差为约束,将跨楼层对应连通点连接起来,并建立各楼层连接点与最近道路的连接,生成连接道路。

室内三维拓扑连通点识别伪代码如图9所示。

三维拓扑连通点需要采用聚类算法从步骤(1)识别出的分布于不同区域的三维拓扑连通区域中提取得到。此时,连通区域点密度越大,越接近楼层连通设施的实际位置。常见的基于划分的聚类算法(如K-Means)需事先指定聚类的个数,而楼层连通设施数量是未知的,限制了算法的使用;基于密度的聚类算法(如DBSCAN)在聚类过程中需要指定2个超参数,经验化程度较高,且得到的结果是聚类簇集,需要人为指定连通点的位置,通常取均值位置,均值通常会弱化轨迹点之间的密度差异,而将所有轨迹点同等对待,容易造成连通点位置偏移。

CFSFDP(Clustering by Fast Search and Find of Density Peaks)算法^[19]是一种结合基于密度和基于划分的新颖聚类算法,该算法的核心思想在于聚类中心的刻画,且仅需指定一个参数即截断距离 dc ,即可自动生成聚类中心的个数及位置。聚类中心同时具有两个特点,与本文楼层连通点识别思想不谋而合:

- ① 局部密度最大,被不超过自身密度的邻居点包围;
- ② 与局部密度更大的样本点之间的距离相对较大。

算法 室内三维拓扑连通点识别算法

输入: 群体用户轨迹 $trajArr = \{traj_i\}$, 时间阈值 T_{thr} , 距离阈值 dis_{thr}

输出: 楼层连通点 $Q = \{pt_i\}$

```

1: //获得某个用户的轨迹
2: for  $i = 0 \rightarrow len(trajArr)$  do
3:   //获得某用户轨迹中的轨迹点
4:   for  $j = 0 \rightarrow len(traj)$  do
5:     if  $traj[j].f \neq traj[j+1].f$  then
6:       if  $traj[j+1].t - traj[j].t < T_{thr}$  then
7:         if  $getDis(traj[j], traj[j+1]) < dis_{thr}$  then
8:           //将潜在的楼层连通点添加到集合P中
9:           P.add(traj[j])
10:          P.add(traj[j+1])
11:        end if
12:      end if
13:    end if
14:  end for
15: end for
16: //使用CFSFDP算法识别楼层连通点
17:  $Q = CFSFDP(P)$ 
18: return  $Q$ 

```

图9 室内三维拓扑连通点识别算法

Fig. 9 Algorithm for identifying 3-D indoor topological connections

另外,定位误差的存在会导致道路上而非楼层连通区域的轨迹点也存在层间跳跃的情况,产生异常点。由于CFSFDP算法在识别过程中对密度敏感,因此具有良好的抗噪性能。因此本文采用CFSFDP算法从楼层连通区域中识别出楼层连通点。

对于数据集 $X = \{x_i\}_{i=1}^N$, 其聚类中心取决于两个变量,每一个样本点 x_i 的局部密度 ρ_i (式(7))和距离 δ_i (式(8))。 ρ_i 表示落在以样本点 x_i 为圆心、以 d_c 为半径的圆形区域中样本点个数; δ_i 表示局部密度大于 x_i 的所有样本点中,距离 x_i 最近的样本点与 x_i 之间的距离。

$$\rho_i = \sum_j \chi(d_{ij} - d_c) \quad (7)$$

$$\delta_i = \min_{j: \rho_j > \rho_i} d_{ij} \quad (8)$$

式中: d_{ij} 为样本点 i 和样本点 j 之间的空间距离, d_c 是人为指定的超参数,称为截断距离。 $\chi(x)$ 表示 0~1 函数,当 $x < 0$ 时, $\chi(x) = 1$, 当 $x \geq 0$ 时, $\chi(x) = 0$ 。

根据样本点 x_i 在截断距离 d_c 下的局部密度 ρ_i 和距离 δ_i , 数据集 X 的聚类中心表示如下:

$$\gamma_i = \rho_i \delta_i \quad (9)$$

式中: ρ_i 表示样本点 x_i 的局部密度; δ_i 表示样本点 x_i 的距离。 γ 值越大, 该样本点是聚类中心的可能性越高。

如图 10 所示, 楼层连通区域中每一个点 p_i 在截断距离 d_c 下都有对应的局部密度 ρ_i 和距离 δ , 图中点 O_1 、 O_2 同时具有较大的 ρ 值和 δ 值, 即 γ 值较大。因此, O_1 、 O_2 被确定为数据集 $\{P_{ij}\}_{i=1}^n$ 的 2 个楼层连通点。

3 实验分析

3.1 实验数据

实验数据来自于某商城 2 天的移动对象蓝牙定位数据, 有效轨迹约 4000 条, 总轨迹点逾 400 万个, 数据采样间隔 1~10 s 不等, 如图 11 所示, 采样间隔为 1~2 s 的轨迹点占比 80% 以上, 数据定位精度为 2~3 m, 数据包括用户设备编号、定位时间、 x 坐标、 y 坐标及所在楼层编号 5 个字段(表 1)。

选取 F1、F2 楼层作为试验区域, 原始轨迹通过室内蓝牙定位技术采集得到, 由于手机终端信号不

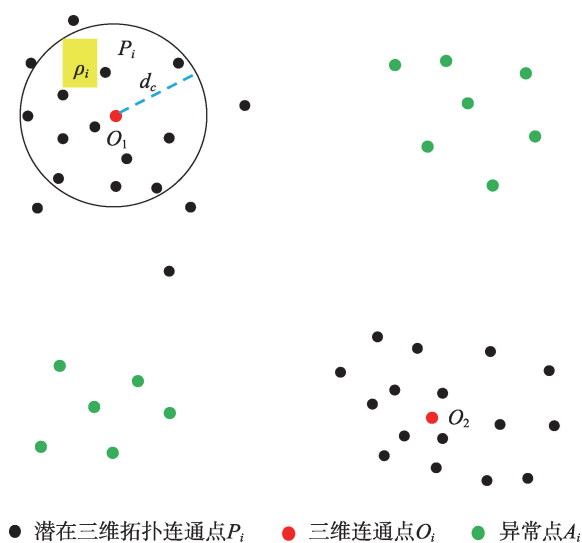


图 10 三维拓扑连通点识别

Fig. 10 Identification of 3-D topological connections

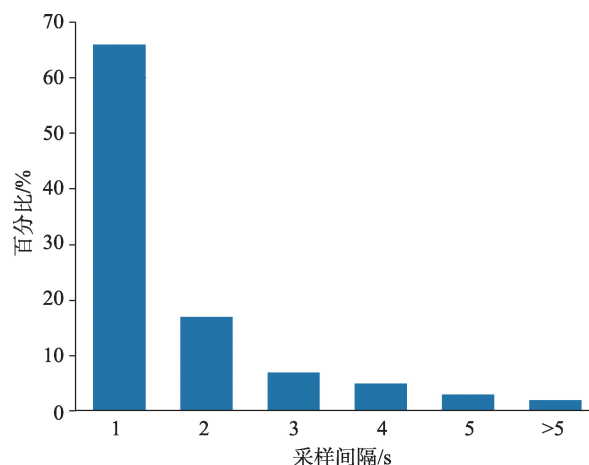


图 11 某商城移动对象轨迹数据采样间隔

Fig. 11 Sampling interval of trajectory data of moving objects in a shopping mall

稳定、人为关闭信号接收、用户停留等原因^[20], 原始轨迹存在异常、漂移、重复等现象, 需要进行数据清洗。主要存在 3 种情形: ① 时间异常轨迹点, 即同一用户在同一时间存在大于一个轨迹点, 只保留一个轨迹点; ② 漂移轨迹点, 即连续时间序列中个别偏移过大的轨迹点, 设置偏移距离阈值, 将偏移距离大于阈值的轨迹点视为离群值, 进行删除; ③ 冗余轨迹点, 同一用户轨迹序列中时间相邻但坐标相同的轨迹点, 进行合并并且只保留一个。

3.2 算法评价方法

采用文献^[21]提出的缓冲区检测方法来评价本文平面导航网络提取精度。对室内原始导航网络分别建立以 0.2、0.5、0.7 m 为半径的缓冲区, 统计落入缓冲区内的路段长度并计算准确率、召回率。同时, 统计实验结果中拓扑正确道路数量与拓扑错误道路数量, 统计百分比, 计算拓扑正确度。

$$P = \frac{\sum_{P \in \sigma} \text{len}(L)}{\sum_{P \in \sigma} \text{len}(L)} \quad (10)$$

$$R = \frac{\sum_{P \in \sigma} \text{len}(L)}{\sum_{P \in \varphi} \text{len}(L)} \quad (11)$$

表 1 移动用户轨迹实例

Tab. 1 Samples of user records

mac	Time	x	y	Floor
000C437***	2017/11/910:00:01	135946***	45097***	1
000C437***	2017/11/910:00:03	135946***	45097***	1
000C437***	2017/11/910:00:04	135946***	45097***	1
.....
000C437***	2017/11/912:20:44	135946***	45097***	2

式中: P 表示准确率; R 表示召回率; ϕ 表示本文提取的导航网络集合; $\phi = \{L_1, L_2, \dots, L_m\}$, φ 表示原有导航网络集合; σ 表示缓冲区内的导航网络集合; $len(L)$ 表示导航网络长度。

Kuntzsch 等^[15]采用基于核密度估计的方法(Kernel Density Estimation Method, KDE)重构路网,即在二维空间中对轨迹数据进行栅格化,构建KDE直方图,根据核密度估计的聚类特性提取出道路网络栅格图像,对栅格图像进行骨架化,计算出代表道路网络几何图形的道路中心线。本文将该方法与自适应栅格化法进行二维导航网络实验对比。通过计算三维拓扑连通点与实际楼层连通设施的距离,作为三维导航网络的评价指标,其中,楼层连通设施位置取门中心位置。

3.3 实验结果与分析

3.3.1 二维平面室内导航网络构建

图12(a)、12(b)是原始轨迹及预处理后的结果,可以看出,由于移动对象停留主要集中于设施区域,因此对所有轨迹进行预处理后产生了明显的区域密度差异,即设施区域轨迹点密度较低,廊道区域轨迹点密度较高。

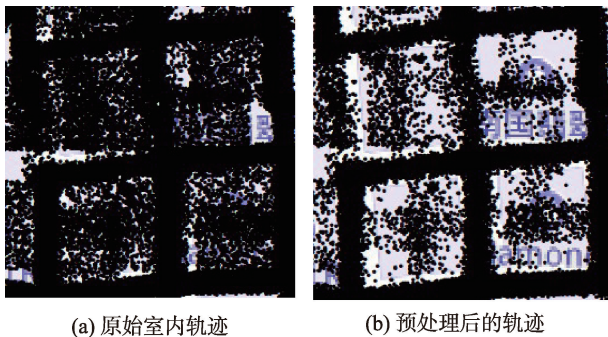


图12 室内轨迹预处理结果

Fig. 12 Pretreatment results of indoor trajectory

图13(a)为8邻域像元中像元值不为0的像元个数 $N_{p_{thr}}$ 不少于3时的自适应栅格化结果,图13(b)是KDE结果。可以看出,本文得到的室内导航网络与KDE栅格化结果相比,道路连通性得到显著增强,如图中红框区域所示。图13(c)是自适应栅格化图像优化结果,可以看出,室内轨迹图像优化后残留噪声基本被去除,道路拓扑连通性进一步加强。图13(d)为采用形态学细化算法提取的室内导航网络栅格骨架线,可以看出,该骨架线基本概括了研究区域的廊道几何结构,道路连通性较好。图14是矢量化后的室内导航网络与商场平面图叠加后的

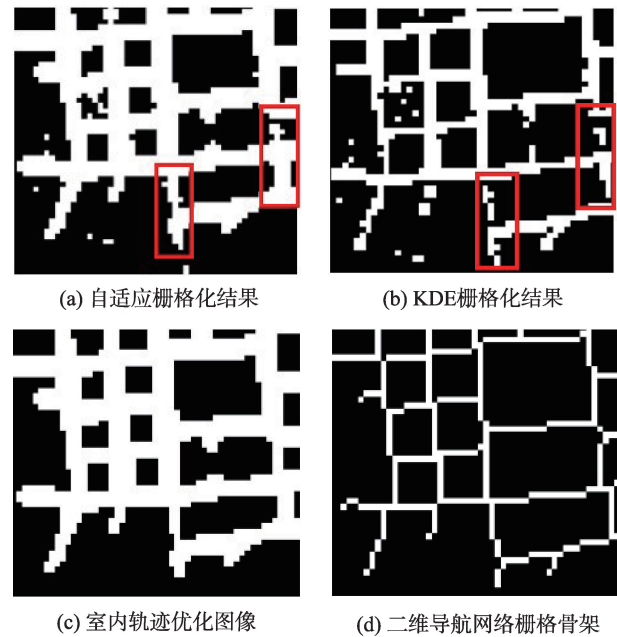


图13 室内二维导航网络提取过程

Fig. 13 Processes of extracting a 2-D indoor navigation network

结果,可以看出,本文方法生成的导航网络基本覆盖了试验区域数据所及之处的所有道路,准确度较高。

采用基于KDE的方法^[15]和本文方法得到的室内平面导航网络结果评价如表2所示,当缓冲区半径分别取值0.2 m、0.5 m、0.7 m时,本文方法导航网络准确率分别达到28.2%、64.4%、80.5%,相比KDE法分别提高1.2%、3.4%、2.7%,准确率平均提高2.43%;本文方法导航网络召回率分别达到25.7%、58.7%、73.3%,相比KDE法分别提高0.9%、2.6%、



图14 室内二维导航网络提取结果

Fig. 14 Extraction result of a 2-D indoor navigation network

1.7%,召回率平均提高1.73%。拓扑正确率提高12.8%。由此可见,本文方法能够有效提高室内导航网络构建的几何精确度和拓扑正确度。

3.3.2 三维室内导航网络构建

图15(a)是按照2.3节方法识别出的楼层连通点集,图中可以看出,高密点集总共有4处,分布于楼层的四周,图15(b)、15(c)是采用CFSFDP算法识别出的三维拓扑点及与平面图叠加的效果,可以看出4个三维连通点均位于电梯位置的附近,4个三维拓扑点与电梯实际位置的距离分别为2.14、

2.93、2.57、2.01 m,准确度较高。以这4个三维拓扑点连接F1、F2楼层,图16是垂直导航网络并与平面导航网络融合后生成的室内三维导航网络结果。由此可见,本文提出的室内导航网络构建方法能够较好地提取该商场的导航网络结构。

4 结论

针对传统室内导航网络构建方法在时效性方面的不足,本文提出一种室内三维导航网络快速众

表2 实验结果评价
Tab. 2 Evaluation of the experiment result

	导航网络缓冲区						拓扑 正确性
	P(0.2 m)	R(0.2 m)	P(0.5 m)	R(0.5 m)	P(0.7 m)	R(0.7 m)	
核密度估计法	27	24.8	61	56.1	77.8	71.6	76.7
本文方法	28.2	25.7	64.4	58.7	80.5	73.3	89.5

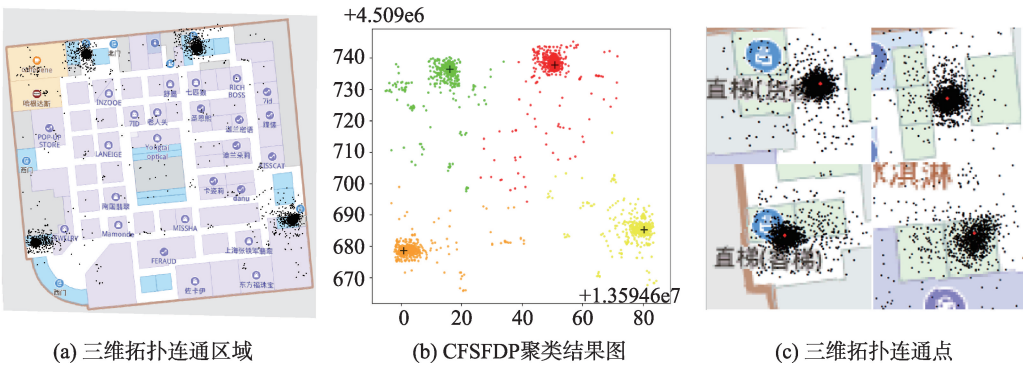


图15 楼层连通点识别及聚类结果
Fig. 15 Identification of floor connections, and clustering result

包构建方法。首先,提出基于ST-DBSCAN的室内轨迹简化预处理方法,去除轨迹停留点信息;接着,通过改进KNN室内轨迹自适应栅格化算法,实现室内二维平面导航网络的提取;最后,通过CFSFDP

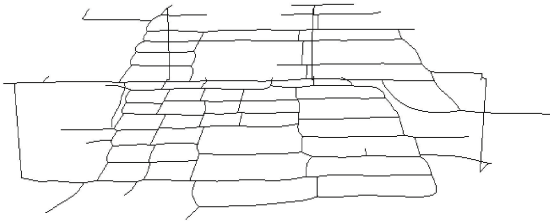


图16 室内三维导航网络提取结果
Fig. 16 Extraction result of a 3-D indoor navigation network

自适应聚类算法识别出楼层三维连通点,实现室内三维导航网络的快速构建。实验结果表明,与基于核密度估计的栅格化方法相比,本文提出的方法将导航网络构建准确率平均提高1.83%,拓扑正确度提高12.8%。能够为传统室内导航网络构建方法提供有效补充,为室内导航网络快速变化检测更新提供支持。室内定位数据误差会影响本研究导航网络提取效果,尤其在轨迹密度较低区域,导航网络提取结果存在较大误差,未来需要进一步提高本研究在稀疏室内轨迹定位数据的适用性。

致谢:上海图聚智能科技股份有限公司为本研究提供了某商城真实室内定位数据支持,在此表示真诚地感谢!

参考文献(References):

- [1] Behar J V. The National Human Activity Pattern Survey (NHAPS): A resource for assessing exposure to environmental pollutants[J]. *Journal of Exposure Analysis and Environmental Epidemiology*, 2000,11(3):231-252.
- [2] 方志祥,徐虹,萧世伦,等.绝对空间定位到相对空间感知的行人导航研究趋势[J].*武汉大学学报·信息科学版*, 2018,43(12):2173-2182. [Fang Z X, Xu H, Xiao S L, et al. Pedestrian navigation research trend: From absolute space to relative space-based approach[J]. *Geomatics and Information Science of Wuhan University*, 2018,43(12):2173-2182.]
- [3] 王培晓,王海波,傅梦颖,等.室内用户语义位置预测研究[J].*地球信息科学学报*,2018,20(12):1689-1698. [Wang P X, Wang H B, Fu M Y, et al. Semantic location prediction of indoor users[J]. *Journal of Geo-information Science*, 2018,20(12):1689-1698.]
- [4] 武恩超,张恒才,吴升.基于中轴变换算法的室内外一体化导航路网自动生成方法[J].*地球信息科学学报*,2018,20(6)730-737. [Wu E C, Zhang H C, Wu S. Automatic generation method of indoor and outdoor integrated navigation network based on mid-axis transformation algorithm [J]. *Journal of Geo-information Science*, 2018,20(6)730-737.]
- [5] 徐战亚,钟赛尚,王媛媛.一种易于更新的室内导航路网构建方法[J].*计算机仿真*,2015,32(12):267-271. [Xu Z Y, Zhong S S, Wang Y W. An easy-to-update method for building indoor navigation network[J]. *Computer simulation*, 2015,32(12):267-271.]
- [6] Wang J, Rui X, Song X, et al. A novel approach for generating routable road maps from vehicle GPS traces[J]. *International Journal of Geographical Information Science*, 2015,29(1):69-91.
- [7] Aronov B, Driemel A, Kreveld M V, et al. Segmentation of trajectories on nonmonotone Criteria[J]. *Acm Transactions on Algorithms*, 2015,12(2):1-28.
- [8] Li J, Qin Q, Han J, et al. Mining trajectory data and geo-tagged data in social media for road map inference[J]. *Transactions in Gis*, 2015,19(1):1-18.
- [9] Tang L, Ren C, Liu Z, et al. A road map refinement method using delaunay triangulation for big trace data[J]. *ISPRS International Journal of Geo-information*, 2017,6(2):45.
- [10] 杨伟,艾廷华.基于车辆轨迹大数据的道路网更新方法研究[J].*计算机研究与发展*,2016,53(12):2681-2693. [Yang W, Ai T H. Research on road network renewal method based on vehicle trajectory big data[J]. *Computer Research and Development*, 2016,53(12):2681-2693.]
- [11] Ahmed M, Karagiorgou S, Pfoser D, et al. A comparison and evaluation of map construction algorithms using vehicle tracking data[J]. *Geoinformatica*, 2015,19(3):601-632.
- [12] Huang J, Deng M, Zhang Y, et al. Complex road intersection modelling based on low-frequency GPS track data [J]. *ISPRS-International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, 2017, XLII-2/W7:23-28.
- [13] Biagioni J, Eriksson J. Map inference in the face of noise and disparity[C]. *International Conference on Advances in Geographic Information Systems*, 2012:79-88.
- [14] Jiang Y, Xiang L I, Xiaojie L I, et al. Geometrical characteristics extraction and accuracy analysis of road network based on vehicle trajectory data[J]. *Journal of Geo-information Science*, 2012,14(2):165-170.
- [15] Kuntzsch C, Sester M, Brenner C. Generative models for road network reconstruction[J]. *International Journal of Geographical Information Science*, 2016,30(5):1012-1039.
- [16] Li Q, Zheng Y, Xie X, et al. Mining user similarity based on location history[C]. *ACM Sigspatial International Conference on Advances in Geographic Information Systems*, 2008:1-10.
- [17] Birant, Derya, Kut. ST-DBSCAN: An algorithm for clustering spatial-temporal data[J]. *Data & Knowledge Engineering*, 2007,60(1):208-221.
- [18] Davies J J, Beresford A R, Hopper A. Scalable, distributed, real-time map generation[J]. *IEEE Pervasive Computing*, 2006,5(4):47-54.
- [19] A R, A L. Machine learning. Clustering by fast search and find of density peaks[J]. *Science*, 2014,344(6191):1492.
- [20] Jin P, Cui T, Wang Q, et al. Effective Similarity search on indoor moving-object trajectories[C]. *International Conference on Database Systems for Advanced Applications*. Springer International Publishing, 2016:181-197.
- [21] Zhang L, Thiemann F, Sester M. Integration of GPS traces with road map[C]. San Jose: *International Workshop on Computational Transportation Science*, 2010:17-22.