基于众包的嵌套流形匹配室内定位方法

周阿鹏[1] 覃锡忠[1] 贾振红[1] NIKOLA Kasabov[2]
(新疆大学信息科学与工程学院 乌鲁木齐 830046)[1]
(奥克兰理工大学知识工程与发现研究所 奥克兰 1020)[2]

摘 要 随着普适应用的兴起,室内定位变得越来越重要。传统的基于指纹的定位方法需要现场勘测,所需时间及工作量巨大,且需实时更新,以适应室内变化,这大大限制了其应用范围。采用众包形式进行室内信息采集,并记录其在室内的大量路径信息,利用嵌套在路径中的低维流形一致性进行地理位置匹配,以建立位置指纹库。通过高斯粒子滤波器对传感器数据进行去噪,进而解决步长差异问题。定位时,根据用户位置的连续性和路径信息筛选出合理的近邻点,继而实现精确定位。在84m^{2}的会议室进行实验,在不需要现场勘测的情况下,所提方法可达到与传统方法可比的定位精度。该方法可以实时适应环境变化,在2周甚至1个月之后,其定位准确性优于传统定位方法。

关键词 室内定位,众包,嵌套流形,高斯粒子滤波,步长差异
中图法分类号 TP391 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2017.08.012

Crowdsourcing-based Indoor Localization via Embedded Manifold Matching

ZHOU A-peng[1] QIN Xi-zhong[1] JIA Zhen-hong[1] NIKOLA Kasabov[2]
(School of Information Science and Engineering, Xinjiang University, Urumqi 830046, China)[1]
(Knowledge Engineering and Discovery Research Institute, Auckland University of Technology, Auckland 1020, New Zealand)[2]

Abstract With the boom of pervasive applications, indoor localization becomes more and more important. The traditional fingerprint based positioning method requires the site survey in which the required time and workload are huge and the real-time updating needs to adapt to the changes in the room. All these factors greatly limit its scope of application. Therefore, the form of crowdsourcing was utilized to collect indoor information and record a large number of path information. The consistency of the low dimensional embedded manifold in the path were used for geographical position matching in order to establish the database of location fingerprints. Gauss particle filter denoising sensor data were used to further solve the problem of pedestrian step difference. According to the continuity of the user's location and the path information, the reasonable nearest neighbor points were selected, and the accurate positioning was realized. The experiments have been carried out in the meeting room of 84m^2, and the experiments can achieve comparable accuracy to the traditional method. The proposed method can adapt to the environmental changes in real time, and the positioning accuracy is better than the traditional positioning method after 2 weeks and even after 1 month.

Keywords Indoor localization, Crowdsourcing, Embedded manifold, Gauss particle filter, Pedestrian step difference

1 引言

无线定位技术进入人们的日常生活要归功于GPS的普及。与GPS类似的系统还有俄罗斯的GLONASS系统、我国的北斗一号区域性卫星导航系统,以及欧盟的伽利略定位系统。这些室外定位系统可以提供很好的定位服务,但是在室内环境中由于建筑物的遮挡,很难接收到卫星信号,因此研究定位的热潮便转移到了室内。如今室内定位研究方兴未艾,尤其是WiFi定位,其流行归功于其作为无线接入技术的流行以及WiFi网络基础设施和终端的普及等[1-2]。

基于WiFi的定位方法主要分为两大类:基于测度的定位和基于指纹的定位。前者利用无线信号的传播模型,根据测得的信号强度计算待定位点到周边AP(至少3个)的距离,进而实现定位[1]。然而由于室内环境复杂,尤其是多径传输等

到稿日期:2016-07-05 返修日期:2016-10-18 本文受中国移动通信集团新疆有限公司研究发展基金项目(XJM2013-01),教育部促进与美大地区科研合作与高层次人才培养项目(2014-2029)资助。
周阿鹏(1990—),男,硕士生,主要研究方向为室内定位,E-mail:1309965165@qq.com;覃锡忠(1964—),男,硕士,副教授,硕士生导师,主要研究方向为信号与信息处理,E-mail:qmqqxz@163.com;贾振红(1964—),男,博士,教授,博士生导师,主要研究方向为光通信技术、信号与信息处理,E-mail:jzhh@xju.edu.cn;NIKOLA Kasabov(1951—),男,博士,教授,主要研究方向为智能信息处理,E-mail:nkasabov@aut.ac.nz。

因素导致在定位过程中产生不可消除的误差，使得定位效果很差，因此其定位精度有限。基于指纹的室内定位方法包括两个过程：离线阶段和在线阶段。在离线阶段，进行现场勘测并规划感兴趣区域，在各个规划参考点离线采集 RSSI 样本，并将其以〈RSSI，参考点坐标〉的形式保存到位置指纹库；在线阶段则根据在待定位点测量到的 RSSI 与位置指纹库进行匹配来实现定位。离线阶段进行现场勘测并建立位置指纹库会消耗巨大的时间、人力，且易受环境动态变化的影响，比如室内状况的变化、AP 变动等，因此该技术的推广应用受到限制[2]。

用户在室内行走时，不能像在室外那样不受约束地从一个位置直线达到另一个位置，而是必须按照特定的路径行走。本文在保证定位精度的前提下，提出利用众包形式在室内收集数据，通过手机内置惯性传感器记录每个行人的详细路径信息，根据内在嵌套的低维路径流形的一致性建立位置指纹库，通过高斯粒子滤波器对数据进行去噪，进而解决步长差异问题。在线阶段充分利用用户位置的连续性和行人的路径信息选择近邻点，采用 Sørensen 距离代替传统的欧氏距离。实验表明所提方法可以达到与传统定位方法可比的定位精度；在经过 1 周、2 周、1 个月后，其结果逐渐优于传统定位方法。

本文解决了以下 4 个问题：

(1)在保证定位误差不是很大的情况下，通过众包形式大大减少了建立位置指纹库所消耗的时间以及人力等；

(2)解决了传统数据库不能适应室内变化的问题，可以根据新采集的众包数据更新指纹库，保证其可适应性；

(3)采用高斯粒子滤波器对传感器数据进行去噪，进而计算已知步数情况下的行走距离，求得各行人步长后对其做归一化处理，以解决步长差异问题；

(4)在线定位时，利用位置的连续性和库内详细路径信息进行近邻点选择，有效排除“伪近邻点”，从而减小定位误差。

本文第 2 节论述了近年来的相关研究进展；第 3 节对本文所要实施的定位过程进行了算法描述；第 4 节给出了实验操作和结果分析；最后总结全文。

2　相关工作

引言部分已提到指纹定位过程包括两个阶段：离线阶段和在线阶段，目前国内外也大多围绕这两个阶段进行研究。Wu 等[3-6]提出用户移动性对定位过程有所助益，将分散的指纹点相互关联起来，通过 MDS 将所收集的数据映射到高维，从而形成指纹空间，对平面地图也进行采样规划映射，并根据其相似性进行匹配。然而，映射到高维以实现数据可视化固然可观，但定然会出现偏转、反射等，使得定位误差增大。Leu 等[7]通过充分利用周边的 AP，提出 RSS“足迹”的概念，并用此来选择近邻点，通过朝向对 RSS 进行滤波以减轻波动性，这些都加大了定位的复杂度，使得定位实时性较差，且不能实时更新指纹库。Wu 等[8]论证了解决室内环境动态变化以及无线地图更新这两个问题的重要性，考虑到用户大部分时间在室内是静止的，因此将其看作是可以移动的参考点并在其静态时采集数据，但此过程的小误差在定位时会被累积。Zhu 等[9]为了消除定期更新 RSSI 地图以保证定位准确度的麻烦并建立细粒度指纹库，提出了时空相似模型，但其解决问题的能力有限，没有实时更新指纹库，且复杂度稍高。

Chang 等[10]和 Wang 等[11]针对随时间变化，动态环境逐渐削弱指纹库稳定性的问题，提出自动确定指纹库更新时间，然后利用定位问题的天然稀疏性，通过 CS 理论来减少数据量和能量消耗，但此过程需计算 RSS 的累计变化，且其服从高斯分布的假设与实际情况有一定的差距。Niu 等[12]以有效利用能量为切入点，通过 ZigBee 接口收集 WiFi 指纹，提出了一种指纹匹配算法来校准成对的指纹；为提高定位精度，设计了 3 种加权距离来进行定位，但其最好距离因对比较少而缺乏推广度。Torres-Sospedra 等[13]对 51 种描述指纹之间的相似性度量函数进行了详细比较，并与传统方法所采用的欧氏距离进行了对比，定位精度可提高 1.7m，由此可见传统方法中选用欧氏距离是有待商榷的。Gao 等[14]描述了利用内置惯性传感器、WiFi 和平面地图约束来侦测转弯，估计运行距离，进而预测其位置，但是该方法需要来自 GPS、邻近设备等的外在参考点，并且没有考虑用户的路径信息。He 等[15]概述了在室内定位领域取得的新进展，说明了运用时空信号模式、用户协作以及运动传感器的重要性，在有效的系统部署方面也阐述了减少离线建立指纹库的工作量、更新指纹库的必要性。Lymberopoulos 等[16]的研究以及近年来的微软室内定位赛事[17]展示了国内外研究机构在该领域的比赛结果，相比其他不用额外设备的定位方法，利用 WiFi 和 IMU(惯性测量单元)完全可以达到很好甚至更优的定位精度。

3　算法描述

3.1　关于流形

在信息科技飞速发展的同时，人们所面对的数据也以指数形式增长，且多表现出高维、小样本特征。目前还没有相应的有效方法对这些数据进行处理，传统的数据分析方法在处理此类数据时往往收效不大甚至失效。

流行学习是近年来才发展起来的一类新的非线性维数约简方法，是指基于局部线性的流形学习方法。流形学习的应用对象是嵌入在高维空间的非线性低维流形上的数据。在认知过程中，人们往往是通过这种非线性低维流形来识别事物的[18-19]。下面给出有关流形的数学定义。

定义 1(拓扑空间)　一个拓扑空间就是一个集对(X,τ)，其中集合 X 为一非空集合，拓扑 τ 是 X 满足以下性质的子集。

1)τ 关于属于它的任意多元素的并运算是封闭的；

2)τ 关于属于它的有限多元素的交运算是封闭的；

3)空集和 X 本身是拓扑 τ 的元素，或者空集和 X 本身属于拓扑 τ。

定义 2(Hausdoff 空间)　如果对于任意集合 X 中的两

个点 X_i 和 X_j，都存在 X_i 的邻域 U 和 X_j 的邻域 V 使得这两个邻域不相交，即 $U \cap V = \emptyset$，则称集合 X 所在的空间为 Hausdoff 空间。

定义 3（流形） 对于位于一个嵌入高维空间的低维流形上的数据，其本身具有低维内在的流形结构。假设 M 是一个 Hausdoff 拓扑空间，若对每一点 $p \in M$ 都有 p 的一个开邻域 U 与 R^d 空间中的一个开子集同胚，则称 M 是 d 维拓扑流形，简称 d 维流形。

用户在室内从一个位置到达另一个位置会形成一定的路径，不能像室外那样可以通过直线到达，这是因为室内存在阻隔，譬如墙壁、家居用品以及实验设备等；经过大量用户的室内运动后会得到室内任意两位置间的路径信息，室内各个位置点之间会形成一个互通网络。当大量用户在室内行走时，这些大量高维数据均匀采样于嵌套在室内空间的低维路径流形上，正是这种潜在的嵌套流形使得通过以众包形式采集到的信息符合一定的规律，通过这种规律性就能建立室内位置指纹库。用户在室内路径上行走时，路径信息通过其所携带手机中内置的惯性传感器进行记录，这里主要通过加速度计来获得路径上各个位置之间的距离信息。当室内各个位置形成一个互通网络时，便得到室内各个位置之间的相互距离。

同时，采集到的数据具有时间标签，按照路径行走就会形成与此路径一致的时间序列数据。当大量众包路径覆盖整个室内区域时，室内各个位置点就通过这些行人路径连接起来，从而在可抵达区域形成一个互通的连通网络。形成互通网络后，两个位置之间往往会有多条路径相连，运用最短路径算法[20]对由路径信息得到的相互点之间的距离进行修正，使得室内相互点之间的距离是最优的，以纠正由于用户行走的随意性以及折返往复等所引起的路径折叠交叉带来的不利影响。室内各个位置都互通时，相互位置点之间的距离是最短的，因此形成的覆盖室内各个位置点之间的相互路径距离是唯一的，而且这与嵌套于室内空间的路径信息流形是一致的。

将室内可到达区域进行网格化处理，在同一个区域（如图 1 所示的 A,B,C,D,E 5 个区域）中，网格点之间采用欧氏距离，不在同一个区域的网格点之间采用实际步行距离代替（这与流形中测地距离类似）。因此，这些网格点之间的相互关系便形成一种类似路径的流形。要从 A 区到达 D 区，由于室内阻隔物的阻挡，必须经过 C 区，按照一定的路径 A→C→D 才能到达；要从 B 区到达 E 区亦然，需要经过路径B→C→E方能到达。

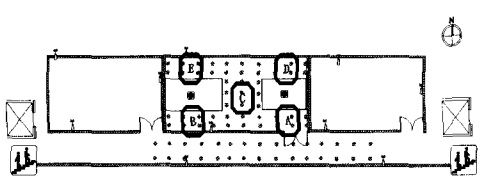


图 1 室内平面网格化结构

根据它们之间嵌套的低维路径流形的一致性，将利用众包形式得到的采集数据映射到室内平面图上，使得采集到的数据便于与地理位置信息相关联，如此便建立了位置指纹数据库。

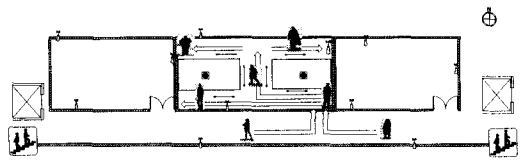


图 2 室内行人移动路径

3.2 嵌套流形的匹配

嵌套流形匹配的步骤如下：

步骤 1 根据 AP 信号强度穿墙跃变的特性（见图 3）确定行人由走廊进入房间的时刻，以此确定图 1 中 A 区的众包数据，该区域内先后次序由采集数据的时间标签确定；

步骤 2 由于从 A 区到 D,E 区以及从 B 区到 D,E 必然经过 C 区（反之亦然），因此 C 区位置的连通性 V 较其他区域高很多（见图 4），以此来确定属于 C 区的众包数据；

步骤 3 对于一条完整路径，其若不经过 C 区，便确定经过 B 区，由此确定属于 B 区的众包数据；

步骤 4 根据强度距离与地理距离的强相关性，在确定了 A,B 区域后，便能确定 D,E 区；

步骤 5 按照 A→B→C→D→E 的顺序将众包数据次序化，利用数据内时间标签以及各点之间记录的传感器数据可以得到指纹点之间的相互距离关联（见图 5）；

步骤 6 按照上述次序得到室内平面网格结构规划图采样点之间的相互距离关系（见图 6），并根据其嵌套路径信息的一致性进行匹配映射；

步骤 7 当用户进行位置询问时，根据其所采集的指纹数据在位置指纹库中进行查找，利用众包路径进行近邻点的筛选（具体过程见 3.3 节），进而确定其位置。

当用户从室外进入室内时，由于墙壁的遮挡，用户携带的移动设备接收到的室内 AP 的信号强度值以及室外（如走廊上安装的 AP）的信号强度值会发生很大的跃变[21]，如图 3 所示。由于移动设备接收到 AP 的信号强度值时其数据是有时间标签的，因此我们可以据此断定用户进入室内的时间及其初始位置，进而，可以断定用户进入了图 1 所示的 A 区，从而实现了步骤 1。

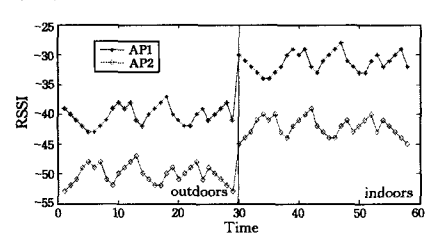


图 3 信号强度经墙体发生跃变

根据匹配步骤 2，C 区的位置点有很大的概率出现在随机选择的两点间的最短路径上。按照众包数据采集的路径次

序，根据式(1)来计算互通网络各个位置点的连通性。

$$B(v)=\sum_{s\neq v\neq t\in V}\frac{\sigma_{st}(v)}{\sigma_{st}} \tag{1}$$

其中，$\sigma_{st}(v)$表示从 s 到 t 经过 v 的最短路径数量，而 σ_{st} 则表示从 s 到 t 的最短路径数量。

感兴趣区域内各个位置的连通性是有很大区分的，位于过渡区域的位置(如图 1 中的 C 区域位置)的连通性远远大于其他区域位置的连通性(见图 4)，并且往往相差好几个数量级，根据这种大的区别可以很容易找到过渡区域 C。而在同一个高连通区域 C 内，连通性相同的位置有很多，但各个位置所接收到的 RSSI 序列(与对应的具体位置相关)是不同的，接收时间也不同(接收数据是有时间标识的)，根据这些信息以及所记录的关联路径信息、进入房间的跃变特性等，将接收到的数据与具体位置相关联，从而建立位置指纹库。

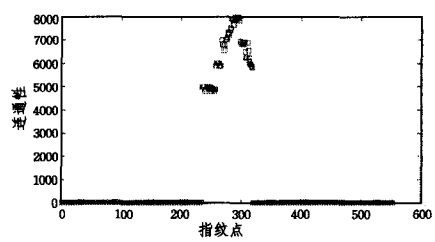


图 4　室内区域的连通性分布

如图 4 所示，有 15.6%的数据的连通性远远大于其他数据的连通性，这部分属于图 1 所示的 C 区，与实际比例吻合，从而实现了步骤 2。

根据强度距离和地理距离的相关性来确定 D，E 区，任意两个位置之间的物理距离与在这两个位置上获取的信号强度向量之间的差异存在着很强的相关性[23]，这是实施指纹定位的理论基础。根据与 A，B 区距离的量度关系可以判断 D，E 区，从而实现了步骤 4。

按照数据时间标签和上述区域规划顺序即 A→B→C→D→E，绘制出众包指纹点之间的相互距离关系，如图 5 所示。

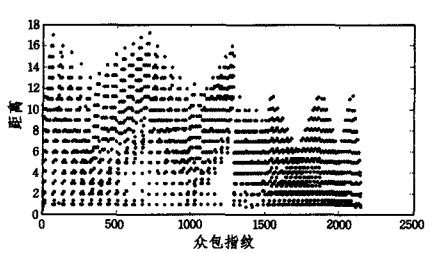


图 5　众包指纹点间的相互距离关系

按上述室内区域顺序进行平面规划，结果如图 6 所示。图中所示的各倾斜平行线以及各水平平行的部分都充分展现了室内路径信息，部分非平行区域则是转弯区。尽管存在部分误差，但路径信息意外地得到保存，且具有很好的相似性。通过两者嵌套的低维路径流形的一致性，将众包形式的指纹点与其地理位置进行映射是可行的，如此便建立了位置指纹库。

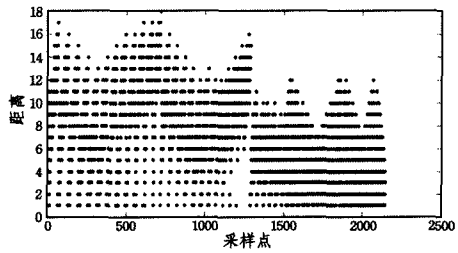


图 6　室内平面采样点间的相互关系

关于惯性传感器，Yuan Gao 和 Edward Y. Chang 等做了大量研究，其采用加速度的周期性来进行计步，以此获得各指纹点之间的行走步数，具体过程可参考文献[14]。

3.3　运用高斯粒子滤波解决步长差异问题

以众包的方式采集数据，指纹库中保存了各个点接收的室内 AP 的 RSSI 序列以及移动设备内置惯性传感器的数据。由于在加速度数据中存在非线性噪声和干扰，导致测量值存在误差，运动较远距离后会累积较大的误差。经研究发现，其误差主要为非高斯误差[21]。在处理这类噪声时，相比其他滤波器，如卡尔曼滤波器(处理线性噪声)、粒子滤波器等，高斯粒子滤波器具有很大优势。

以往方法预先定义一个恒定步长，但行人步长各异，因此用步数或者步数乘以恒定步长来表示距离会造成一定的混乱。

一般同一用户的步长变化相比不同用户间的步长差异要小得多；同时，大量实验数据显示同一用户在不同时刻的步长具有较小的随机差异。对于同一用户，用户行动具有高度的重复性[24]和周期性，其行为模式的变化较小，在经过高斯粒子滤波算法去除噪声和干扰后，少许异常且与其行为模式不符的数据(比如行人起始或停止过程的短时间数据，而此处位置间的距离关联可以由其他用户正常行走时所产生的数据来表示)将被排除，而后采用平均步长作为距离度量单位可以很好地处理这些随机的、小的步长差异，从而可以在一定程度上控制误差，对步长随机差异也具有了一定的抗干扰能力。采用经高斯粒子滤波后的数据来计算特定(已知)步数下的行走距离，以计算各行人步长，并对所得的各步长进行归一化处理，进而利用归一化步长来保持室内全局结构的一致性。

高斯粒子滤波采用粒子滤波方法得到一个高斯密度函数，并以此来近似估计状态的后验概率分布。通常，一个高斯随机变量 x 的密度可表示为：

$$N(x;\bar{x},\Sigma)=(2\pi)^{-m/2}|\Sigma|^{-\frac{1}{2}}e^{[-(x-\bar{x})^{T}\Sigma^{-1}(x-\bar{x})/2]} \tag{2}$$

其中，$\bar{x}$ 是 x 的 m 维向量的均值，Σ 是 x 的协方差阵。

若在 $t-1$ 时刻：

$$p(x_t/y_{t-1})=N(x_t;\overline{x_t},\overline{\Sigma_t}) \tag{3}$$

其中，$\overline{x_t}$，$\overline{\Sigma_t}$ 由先验信息得出。当得到下一个测量值后，进行下一步滤波和概率分布预测。

获得时刻 t 的观测值后，$x_{0:t}=\{x_0,x_1,\cdots,x_t\}$ 的测量更新方程可近似为：

$$p(x_t/y_{0:t})=C_t p(y_t/x_t)p(x_t/y_{0:t-1})$$
$$\approx C_t p(y_t/x_t)N(x_t;\bar{x}_t,\overline{\sum}_t) \quad (4)$$
$$C_t=\left(\int p(x_t/y_{0:t-1})p(y_t/x_t)\mathrm{d}x_t\right)^{-1} \quad (5)$$

高斯粒子滤波更新通过一个高斯分布的 $N(x_t;\hat{x}_t,\overline{\sum}_t)$ 来近似上述滤波概率分布。通常不能精确求解 $p(x_t/y_{0:t})$ 的均值和协方差。对于高斯粒子滤波，首先对重要性函数 $q(x_t/y_{0:t})$ 取样本 $x_t^{(n)}$ 并计算权值 $w_t^{(n)}$，然后获得估计其状态 x_t 的均值和协方差：

$$\hat{x}_t=\sum_{n=1}^{M}w_t^{(n)}x_t^{(n)} \quad (6)$$
$$\sum_t=\sum_{n=1}^{M}w_t^{(n)}(\hat{x}_t-x_t^{(n)})(\hat{x}_t-x_t^{(n)})^H \quad (7)$$

因此，滤波概率分布可近似为 $p(x_t/y_{0:t})\approx N(x_t;\hat{x}_t,\sum_t)$。

此时，高斯粒子滤波的滤波概率分布近似为高斯分布，在 t 时刻对状态转移分布 $p(x_{t+1}/x_t)$ 取样，则可通过新的高斯分布近似预测概率。因此，$x_{0:t}=\{x_0,x_1,\cdots,x_t\}$ 的测量预测方程可近似为：

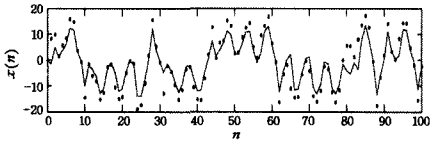
$$p(x_{t+1}/y_{0:t})\approx\int p(x_{t+1}/x_t)N(x_t;\bar{x}_t,\sum_t)\mathrm{d}x_t$$
$$\approx\frac{1}{M}\sum_{n=1}^{M}p(x_{t+1}/x_t^{(n)}) \quad (8)$$

其中，$x_t^{(n)}$ 从 $N(x_t;\bar{x}_t,\sum_t)$ 采样获得，在 t 时刻的观测值中，从 $n(n=1,2,\cdots,M)$ 的状态转移分布 $p(x_{t+1}/x_t^{(n)})$ 抽样可获得 $t+1$ 时刻的状态粒 $x_{t+1}^{(n)}$，其预测概率的均值和协方差为：

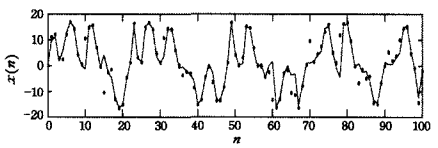
$$\bar{x}_{t+1}=\frac{1}{M}\sum_{n=1}^{M}x_{t+1}^{(n)} \quad (9)$$
$$\overline{\sum}_{t+1}=\frac{1}{M}\sum_{n=1}^{M}(\bar{x}_{t+1}-x_{t+1}^{(n)})(\bar{x}_{t+1}-x_{t+1}^{(n)})^H \quad (10)$$

其预测概率分布可近似为 $p(x_{t+1}/y_{0:t})\approx N(x_{t+1};\bar{x}_{t+1},\overline{\sum}_{t+1})$。

通过进行大量仿真实验来对比粒子滤波器和高斯粒子滤波器的性能。图 7 示出了通过上述两个滤波器滤波后的前 100 个状态和预测值(其中 $M=100$)。



(a)粒子滤波仿真图



(b)高斯粒子滤波仿真图

图 7

仿真表明，相比粒子滤波器，高斯粒子滤波器对非线性数据更具有实时处理能力。

运用高斯粒子滤波器对采集的传感器数据进行去噪处理，利用处理后的数据计算已知步数下的行走距离，进而求得各行人步长，实验结果如表 1 所列。

表 1 运用高斯粒子滤波器计算行人步长/m

行人实际步长	0.53	0.47	0.62
计算步长	0.51	0.44	0.64

3.4 在线定位

在定位阶段，在待定位点采集到该位置处的强度信息时根据最近邻算法在位置指纹库中进行查找定位。在待定位点采集到该处 m 个 AP 的 RSS 数据，表示如下：

$$RSSmean_r=[RSSmean_r^1,RSSmean_r^2,\cdots,RSSmean_r^m]$$

由于用户在室内进行移动时其位置是连续的，从一个位置移动到另一个位置会经过一定的路径，因此根据这种连续性以及路径信息来寻找近邻点，这些近邻点会形成一个小的连通图，去除此连通图内按照式(1)计算的个别跳跃值。通过此方法找到的近邻点可以避免偏离待定位点很远的参考点，对近邻点进行二次选择，筛除围绕待定位点路径以外的“伪近邻点”，从而有效降低定位误差。

选择近邻点时，文献[12]对于距离的选择进行了详细的分析对比，这里采用 sorensen 距离，见式(11)；利用式(12)、式(13)可以计算定位坐标。

$$Distance_{sorensen}(P,Q)=\frac{\sum_{i=1}^{d}|P_i-Q_i|}{\sum_{i=1}^{d}(P_i+Q_i)} \quad (11)$$
$$d_i=\frac{1}{Distance_{sorensen}(RSSmean_i,RSSmean_r)} \quad (12)$$
$$L(x,y)=\sum_{j=1}^{C}\frac{d_j}{\sum_{i=1}^{C}d_i}R(x,y) \quad (13)$$

其中，$L(x,y)$ 是待定位点的坐标，$R(x,y)$ 是最近邻参考点的坐标，C 代表筛选后剩余的参考点数目。

4 实验

4.1 实验设计

实验在面积为 84m² 的会议室(长 12m，宽 7m)进行，如图 1 所示，布置的 AP 采用全向天线 Tenda 300Mbps 无线接入点，型号是 W301A。在室内可抵达区域按 1m×1m 的网格进行采样，获得 69 个采样点。在众包数据收集阶段，由 3 名实验室成员携带手机在室内行走，共形成了 44 条路径，收集到 5520 个指纹数据，其足以覆盖整个室内可抵达区域，采集过程持续 3 个小时。实验中，这些路径的长度以及覆盖区域是互异的。这些众包数据一半用于训练阶段，一半用于在线定位。

4.2 数据分析

在特定点接收到的 RSSI 由于硬件等原因会出现异常值(见图 8)，去除这些异常点，以进行数据缺失处理，t 检验结果表明其属于完全随机缺失；用 SPSS 软件进行缺失值填充处理(见图 9)；然后利用其均值进行后续的定位，数据保存形式如下：

$$RSSmean_i=[RSSmean_i^1,RSSmean_i^2,RSSmean_i^3,\cdots,RSSmean_i^m]$$

定位误差 e 为真实的位置坐标 (x_0,y_0) 与定位估计的位置坐标 (x,y) 之间的距离，即 $e=\sqrt{(x-x_0)^2+(y-y_0)^2}$。

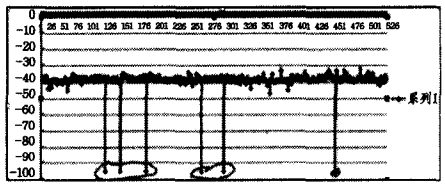


图 8　原始时间序列测量值

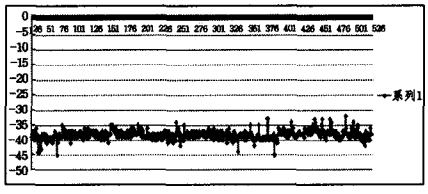


图 9　SPSS 处理后的时间序列测量值

将本文方法与传统的需要现场勘测的 KNN 方法进行对比。选择 KNN 作为对比算法是为了显示通过众包采集数据代替现场勘测所导致的定位精度下降，而不是为了证明本文方法要优于传统方法[3-4]。实验结果如图 10 所示。

随后，经过 1 周、2 周和 1 个月后再次进行定位实验，定位结果如图 11－图 13 所示。结果表明，运用传统采集数据建立指纹库的方法进行定位时，定位准确定性大大降低，而采用众包指纹的方法进行定位时，由于保持了指纹库的更新，其准确性几乎不变。经过 2 周的时间后，室内环境发生了一定的变化，而采用传统方法建立的指纹库就会“过时”，利用其进行定位时误差就会增大。

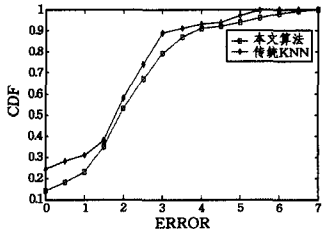


图 10　定位误差累积分布

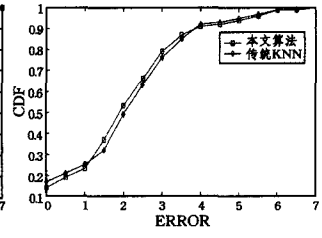


图 11　1 周后定位误差累积分布

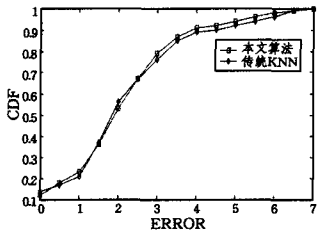


图 12　2 周后定位误差累积分布

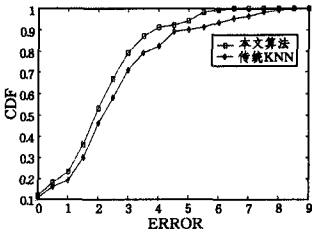


图 13　1 个月后定位误差累积分布

4.3　误差分析

在 A，B，C，D，E 各区都进行定位测试，总体平均定位误差为 1.63m。C 区定位精度最高，平均定位误差为 1.08m；其他区域定位精度与 C 区相比稍低，其中 A 区的平均定位误差为1.35m，B 区为 1.41m，D 区为 1.72m，E 区为 1.57m。在 B，D，E 3 个区域内，行人路径的尽头处（如图 14 中的爆炸区域）的误差要比其他位置大很多，其中 B 区路径尾处的平均定位误差达到了 1.97m，D，E 区路径尾处的平均误差分别为 2.54m和2.38m。这是因为在收集众包数据时，根据行走习惯，人们往往不倾向于走进“死胡同”，行人有时没有走到路径尽头，因此该位置的数据量较其他位置明显偏少，这就导致了该部分的定位误差偏大。

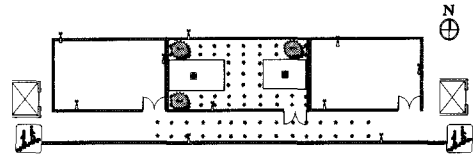


图 14　室内大误差区域的分布

4.4　与 MDS 方法的比较

文献[4]利用多维标度的方法将众包数据映射到高维指纹空间，其房间级别定位误差为 11%，且仅利用 MDS 方法建立位置指纹库进行定位时定位错误较高，于是在此基础上提出运用轨迹匹配的方法来进一步提高定位精度。在将众包数据进行高维可视化映射的过程中，仅利用位置间的相互距离进行映射时会出现偏移等情况，而本文利用嵌套在路径中的低维流形一致性来建立位置指纹库，利用了位置间的距离信息以及众多的路径信息等。

将本文算法与 MDS 映射方法、MDS＋轨迹匹配方法（MDS＋）进行对比，实验结果如图 15 所示。

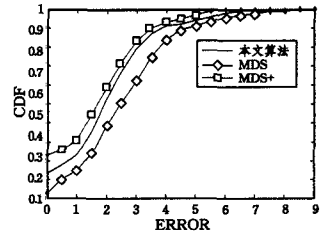


图 15　与 MDS 方法的对比结果

实验表明，在定位精度方面，本文算法优于 MDS 方法，稍逊于 MDS＋轨迹匹配方法；但是轨迹匹配由于是通过最小化积分距离（具体过程参考文献[4]的定位部分）来实现的，因此时间消耗比较大。

结束语　以众包的方式利用所有参与用户的数据，根据其内在低维路径流形的一致性建立指纹库，以此来解除传统现场勘测之苦；提出高斯粒子滤波的方法来解决行人步长差异问题，避免了室内全局结构的混乱，并根据位置的连续性和丰富的路径信息进行近邻点选择。为期 2 个月的实验表明，所提方法与传统方法相比，定位精度下降有限（但省去了现场勘测的过程）；然而，经过一段时间后，传统方法所运用的指纹库由于环境变化而“过时”，其定位准确度会逐渐降低，本文方法则逐渐凸显优势。

接下来的研究重点集中在将该方法用于室内的目标追踪以及室内导航等场景，以更充分利用库内丰富的路径信息。

参 考 文 献

[1]　XI R，LI Y J，HOU M S. The review of indoor positioning methods[J]. Computer Science，2016，43(4)：1-6. (in Chinese)
席瑞，李玉军，侯孟书. 室内定位方法综述[J]. 计算机科学，2016，43(4)：1-6.

[2]　杨峥，吴陈沭，刘云浩. 位置计算：无线网络定位与可定位性[M]. 北京：清华大学出版社，2014：1-3.

[3] WU C, YANG Z, LIU Y. Smartphones based Crowdsourcing for Indoor Localization[J]. IEEE Transactions on Mobile Computing, 2015, 14(2): 444-457.

[4] YANG Z, WU C, LIU Y. Locating in fingerprint space: wireless indoor localization with little human intervention[C]// International Conference on Mobile Computing and Networking. ACM, 2012: 269-280.

[5] WU C, YANG Z, XU Y, et al. Human Mobility Enhances Global Positioning Accuracy for Mobile Phone Localization[J]. IEEE Transactionson Parallel & Distributed Systems, 2015, 26(1): 131-141.

[6] YANG Z, WU C, ZHOU Z, et al. Mobility Increases Localizability: A Survey on Wireless Indoor Localization using Inertial Sensors[J]. ACM Computing Surveys, 2015, 47(3): 1-34.

[7] LEU J S, YU M C, TZENG H J. Improving indoor positioning precision by using received signal strength fingerprint and footprint based on weighted ambient Wi-Fi signals[J]. Computer Networks, 2015, 91: 329-340.

[8] WU C, YANG Z, XIAO C, et al. Static power of mobile devices: Self-updating radio maps for wireless indoor localization[C]// Computer Communications. IEEE, 2015: 2497-2505.

[9] ZHU J Y, ZHENG A X, XU J, et al. Spatio-temporal (S-T) similarity model for constructing WIFI-based RSSI fingerprinting map for indoor localization[C]// International Conference on Indoor Positioning and Indoor Navigation. IEEE, 2014: 678-684.

[10] CHANG L, CHEN X, WANG J, et al. TaLc: Time Adaptive Indoor Localization with Little Cost[C]// ACM MOBICOM Workshop on Challenged Networks. ACM, 2015: 461-490.

[11] WANG J, FANG D, CHEN X, et al. LCS: Compressive sensing based device-free localization for multiple targets in sensor networks[J]. Proceedings-IEEE INFOCOM, 2013, 12(11): 145-149.

[12] NIU J, WANG B, SHU L, et al. ZIL: An Energy-Efficient Indoor Localization System Using ZigBee radio to Detect WiFi Fingerprints[J]. IEEE Journal on Selected Areas in Communications, 2015, 33(7): 1431-1442.

[13] TORRES-SOSPEDRA J, MONTOLIU R, TRILLES S, et al. Comprehensive Analysis of Distance and Similarity Measures for Wi-Fi Fingerprinting Indoor Positioning Systems[J]. Expert Systems with Applications, 2015, 42(23): 9263-9278.

[14] GAO Y, YANG Q, LI G, et al. XINS: the anatomy of an indoor positioning and navigation architecture[C]// Proceedings of the 1st International Workshop on Mobile Location-based Service. ACM, 2011: 41-50.

[15] HE S, CHAN S H G. Wi-Fi Fingerprint-based Indoor Positioning: Recent Advances and Comparisons[J]. IEEE Communications Surveys & Tutorials, 2015, 18(1): 466-490.

[16] LYMBEROPOULOS D, LIU J, YANG X, et al. A realistic evaluation and comparison of indoor location technologies: experiences and lessons learned[C]// The ACM/IEEE Conference on Information Processing in Sensor Networks. 2015: 178-189.

[17] Microsoft Indoor Localization Competition IPSN2015/2016 [EB/OL]. http://research. microsoft. com/en-us/events/indoorlocalizationcompetition 2015/default. aspx.

[18] LUNGA D, PRASAD S, CRAWFORD M M, et al. Manifold-Learning-Based Feature Extraction for Classification of Hyperspectral Data: A Review of Advances in Manifold Learning[J]. IEEE Signal Processing Magazine, 2014, 31(1): 55-66.

[19] ZHANG Z, WANG J, ZHA H. Adaptive manifold learning[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2012, 34(2): 253-265.

[20] FLOYD R W. Algorithm 97: Shortest path[J]. Communications of the ACM, 1962, 5(6): 345.

[21] WU C, YANG Z, LIU Y, et al. WILL: Wireless Indoor Localization without Site Survey[J]. IEEE Transactions on Parallel & Distributed Systems, 2013, 24(4): 839-848.

[22] HARLE R. A Survey of Indoor Inertial Positioning Systems for Pedestrians[J]. Communications Surveys & Tutorials IEEE, 2013, 15(3): 1281-1293.

[23] XIAO Y L, ZHANG S G, WANG J X. An indoor localization algorithm based on multidimensional scaling and region refinement[J]. Chinese Journal of Computers, 2016, 39(67): 6-8. (in Chinese)
肖亚龙，张士庚，王建新. 一种基于多维标度和区域细化的无线室内定位方法[J]. 计算机学报，2016，39(67)：6-8.

[24] GONZÁLEZ M C, HIDALGO C A, BARABÁSI A L. Understanding individual human mobility patterns[J]. Nature, 2008, 453(7196): 779-782.

(上接第 45 页)

[9] CAO S, SALCIZ Z, DING Y T, et al. Temperature-aware task scheduling heuristics on Network-on-Chips[C]// Proceedings of 2016 IEEE International Symposium on Circuits and Systems (ISCAS). Montreal: IEEE, 2016: 2603-2606.

[10] DEB K, PRATAP A, AGARWAL S, et al. A fast and elitist multiobjective genetic algorithm: NSGA-II[J]. IEEE Transactions on Evolutionary Computation, 2002, 6(2): 182-197.

[11] KAUR R, SINGH G. Genetic algorithm solution for scheduling jobs in multiprocessor environment[C]// Proceedings of 2012 Annual IEEE India Conference (INDICON). Kochi: IEEE, 2012: 968-973.

[12] DICK R P, RHODES D L, WOLF W. TGFF: Task Graphs for Free[C]// Proceedings of the Sixth International Workshop on-Hardware/Software Codesign, 1988. Seattle: IEEE, 1988: 97-101.

[13] TOPCUOGLU H, HARIRI S, WU M Y. Performance-effective and low-complexity task scheduling for heterogeneous computing[J]. IEEE Transactions on Parallel and Distributed Systems, 2002, 13(3): 260-274.

[14] WU M Y, GAJSKI D D. Hypertool: A programming aid for message-passing systems[J]. IEEE Transactions on Parallel and Distributed Systems, 1990, 1(3): 330-343.

[15] HUANG W, GHOSH S, VELUSAMY S, et al. HotSpot: a compact thermal modeling methodology for early-stage VLSI design [J]. IEEE Transactions on Very Large Scale Integration Systems, 2006, 14(5): 501-513.