

DOI:10.11918/202212020

基于视觉与 Wi-Fi 双层特征地图的 智能手机室内定位方法

胡钊政^{1,2}, 柳雨婷^{1,2}, 周哲^{1,2}, 黄戈^{1,2}, 孙勋培^{1,2}

(1. 武汉理工大学智能交通系统研究中心, 武汉 430063; 2. 武汉理工大学重庆研究院, 重庆 401120)

摘要:为解决室内环境中 Wi-Fi 定位精度低、视觉定位稳定性差等问题,提出了一种融合视觉与 Wi-Fi 的双层特征地图模型 (vision-CSI map, V-CSI map), 并提出基于隐马尔可夫模型 (Hidden Markov model, HMM) 的智能手机定位方法。V-CSI 地图模型既包含基于栅格的信道状态信息 (channel state information, CSI) 指纹特征, 同时也包含以稀疏安全出口为路标的视觉特征, 并通过参考位置关联完成地图构建。本文将基于 V-CSI 特征地图的定位问题转化为 HMM 问题。首先完成安全出口标志检测与视觉特征匹配, 实现视觉定位, 将定位结果对 HMM 状态进行初始化; 接着, 利用 CSI 指纹匹配完成发射概率建模, 通过高斯模型完成基于运动约束的状态转移概率建模; 最后, 通过前向算法求解 HMM 最优匹配状态, 进而计算位置。在 6 000 平方米办公楼和 3 600 平方米地下停车场对本文算法与模型进行验证。实验结果表明, 两种典型室内场景下, 本文算法平均定位误差约为 1.0 m, 单次定位时间约 170 ms; 相比于单一 CSI 定位, 平均定位误差减少 56% 以上, 说明该算法能够有效提升室内定位的准确性与鲁棒性。

关键词: 室内定位; 智能手机; 双层特征地图; 隐马尔可夫模型; Wi-Fi 定位; 视觉定位

中图分类号: TP242 **文献标志码:** A **文章编号:** 0367-6234(2026)03-0010-10

Smartphone indoor positioning method based on vision and Wi-Fi double-layer feature map

HU Zhaozheng^{1,2}, LIU Yuting^{1,2}, ZHOU Zhe^{1,2}, HUANG Ge^{1,2}, SUN Xunpei^{1,2}

(1. Intelligent Transportation Systems Research Center, Wuhan University of Technology, Wuhan 430063, China;
2. Chongqing Research Institute, Wuhan University of Technology, Chongqing 401120, China)

Abstract: We proposed a smartphone positioning method by formulating the positioning problem as an HMM (hidden markov model, HMM) based on the proposed double-layer feature map consisting of visual and Wi-Fi features (vision-CSI map, V-CSI map) to solve the issue of low accuracy and poor stability in indoor environment. The V-CSI map is modeled by encoding CSI fingerprint features based on grid and visual features of sparse safety exits as well as association locations. The location problem based on the V-CSI feature map is solved as HMM problem in the method. First, the safety exit sign detection and visual feature matching are completed in the visual positioning phase, and the positioning results are employed to initialize and reinitialize the states of HMM. Subsequently, CSI fingerprint features are matched with that of the V-CSI map to complete the emission probability, and the state transition probability is computed by modeling motion constraint with Gaussian model. Finally, the optimal state is derived from the forward algorithm, and the position of the smartphone is readily determined from the weighted average of the closest states. In the experiment, the proposed method is verified in an office building of 6 000 square meters and an underground parking lot of 3 600 square meters respectively. Experimental results show that the average positioning error of the algorithm is about 1.0 m, and the time of a single positioning is about 170 ms in the two typical indoor scenes. Compared with only CSI positioning methods, the average positioning error of our proposed method is reduced by more than 56%. The outstanding performance of experimental results also illustrates that our proposed method can improve the accuracy and robustness of indoor positioning.

Keywords: indoor positioning; smartphone; double-layer feature map; HMM; Wi-Fi positioning; visual positioning

收稿日期: 2022-12-07; 录用日期: 2023-02-14; 网络首发日期: 2024-06-24

网络首发地址: <https://link.cnki.net/urlid/23.1235.t.20240623.1216.004>

基金项目: 国家重点研发计划 (2021YFB2501104); 武汉市科学技术局企业技术创新项目 (2020010601012165, 2020010602011973, 2020010602012003); 武汉理工大学重庆研究院科技创新研发项目 (YF2021-04)

作者简介: 胡钊政 (1979—), 男, 教授, 博士生导师

通信作者: 胡钊政, zzh@whut.edu.cn

随着移动互联网时代的深入发展,位置信息服务(location based service, LBS)已被广泛应用于社会生活的各个领域。全球定位系统(global positioning system, GPS)可以在室外环境中提供精确的位置信息,但人类日常生活中大多时间处于室内,复杂室内环境中的多径效应与非视距干扰会导致GPS无法提供精准的定位信息^[1]。

相较于基于蓝牙和超宽带(ultra wide band, UWB)的智能手机定位方法,基于Wi-Fi的定位方法可利用智能设备普遍搭载的传感器以及场景中广泛部署的Wi-Fi接入点,在不增加额外成本的前提下获取全局坐标位置^[2]。基于Wi-Fi信号的定位技术主要分为两类:三角定位和指纹匹配。三角定位采用信号衰减模型计算接收机与各个Wi-Fi接入点之间的距离,通过三角定位原理获取定位结果,但复杂的室内环境中存在多径效应的影响,很难构建精准的信号衰减模型;指纹匹配方法通过建立接收信号强度指示(received signal strength index, RSSI)与信道状态信息(CSI)指纹数据库,通过指纹相似性匹配得到定位结果。CSI能够提取到比RSSI更细粒度的特征信号,其随时间变化更小,灵敏度更高,能够有效提升定位精度。李华亮等^[3]提出了一种高鲁棒性的RSSI室内指纹定位方法,针对尖峰噪声、接入点(access point, AP)选择及定位系统区域离群坐标问题,分别提出了基于鲁棒主成分分析(robust principal component analysis, RPCA)的环面交集定位方法、AP信息熵加权定位算法与基于凸分解的定位算法。于海涛等^[4]提出一种基于RSS和CSI混合指纹的室内定位方法,通过构建神经网络得到不同位置的权值作为指纹,比较权值差异得到定位结果。但基于指纹匹配的方法依赖于鲁棒的指纹匹配过程,当有动态目标移动时,产生的干扰信号会恶化匹配性能,使得定位精度明显下降。现有的基于CSI指纹的定位研究依赖于高鲁棒性的采集数据和数据降噪等方法,容易因缺少运动约束引发位置跳变,无法满足连续定位需求。

使用单一Wi-Fi传感器定位时,采样点的分布与环境的差异会对定位结果产生较大的影响,需要融合其他传感器以获得更理想的定位效果^[5]。Poulose等^[6]对比分析了基于行人航位推算(pedestrian dead reckoning, PDR)的不同传感器融合技术,筛选

适合航向估计的最佳方法。Tao等^[7]提出了一种基于改进时差载波相位技术的位移估计PDR系统,以变化的位置和航向作为更新信息,通过滑动窗口估计建立PDR系统的步幅和航向偏差。然而,基于智能手机的PDR系统存在累积误差,无法实现长时间定位。Huang等^[8]提出基于多尺度的智能手机融合定位方法,将室内定位分为Wi-Fi粗定位、图像级定位和度量级定位三部分,在节点处进行用户位姿估计,但未加入运动信息且仅限于安全出口标志附近的离散位姿估计。Wang等^[9]提出融合Wi-Fi与图像的室内定位方法,在离线阶段建立基于RSSI均值的指纹数据库,并使用AlexNet对图像进行分割与匹配,在线阶段提取视频帧确定定位区域,再使用基于信号强度的加权平均指纹匹配计算位置,但在用户快速移动时,图像的抖动会影响匹配精度,算法的实时性较差。Zhao等^[10]提出基于相机、Wi-Fi和惯性测量单元(inertial measurement unit, IMU)的多模式室内定位方法,通过分析两张图像间的CSI与IMU数据计算用户移动距离以获取定位结果。但大量观测数据的处理增加了计算的复杂度,且不适用于大型室内场景的连续定位。Sun等^[11]利用AP与参考点之间的几何距离作为指纹,提出基于地图辅助的粒子滤波定位方法,但粒子分布的随机性和大量的粒子计算成本会影响定位的可靠性。HMM能够从预设隐含状态集合中筛选出最优状态,有效规避了预测状态的随机分布^[12]。由此可见,目前基于Wi-Fi与视觉的定位方法由于实验场景路标难以界定、环境差异、设备的异构性以及算法的缺陷,其定位精度与实时性无法满足大众的需求。

针对现有方法的不足,本文提出基于V-CSI双层特征地图的智能手机HMM室内定位方法:将定位场景栅格化,并将栅格定义为隐含状态,通过运动约束建立隐含状态之间的转移关系,利用CSI指纹特征图定义发射概率,结合视觉特征图对隐含状态进行重新初始化,以提升定位精度。

1 本文算法

本文算法整体流程如图1所示,基于V-CSI双层特征地图的HMM智能手机定位方法分为建图和定位两部分。

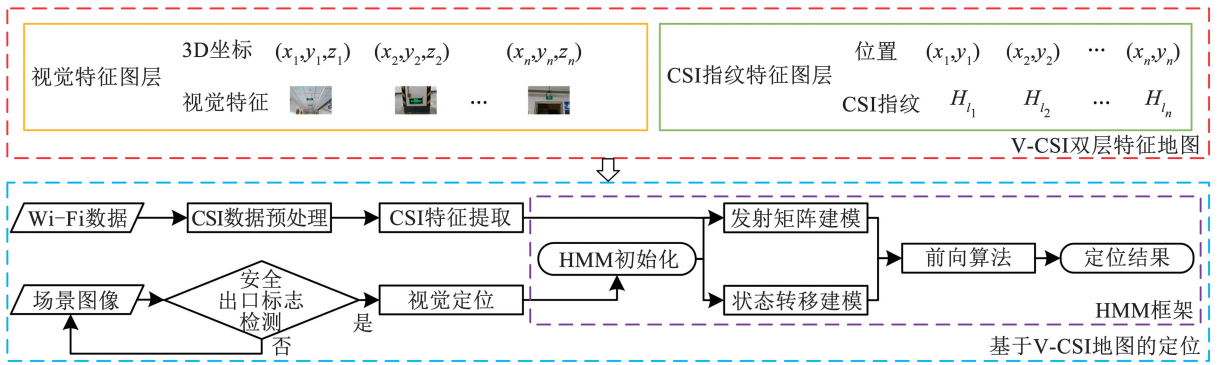


图 1 算法整体流程

Fig. 1 Overall algorithm flow

2 V-CSI 双层特征地图构建

建图阶段,构建整个场景的 V-CSI 双层特征地图,如图 2 所示。构建 CSI 指纹特征图层时,将整个场景划分为 $M \times N$ 个栅格,在每个栅格的中心点采集连续的 Wi-Fi 数据并提取 CSI 信息,建立基于 CSI 幅值的 CSI 指纹特征地图;在 CSI 指纹特征图层的基础上,使用消防安全出口标志作为参考路标,一方面是因为安全出口标志为国家消防安全法规强制要

求安装的引导标识,在室内场景中广泛存在,而且要求安装在出入口的醒目位置,易在视距范围内识别;另一方面是因为消防安全出口标志能够提供固定的参考位置信息。视觉特征图层中的栅格负责存储安全出口标志的视觉场景特征、安全出口标志特征和全局位姿信息。因此,V-CSI 双层特征地图存储的特征信息可以表示为 $W_{V-CSI} = \{M_i, H_i, L_i\}$,其中 M_i 表示视觉特征信息, H_i 表示 CSI 指纹特征信息, L_i 表示位置坐标。

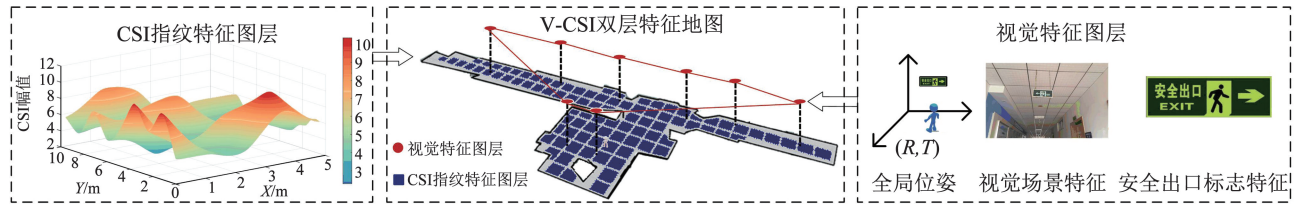


图 2 V-CSI 双层特征地图

Fig. 2 V-CSI double-layer feature map

2.1 CSI 指纹特征图层构建

在构建 CSI 指纹特征图层时,离线采集 Wi-Fi 信息并记录栅格位置坐标,以此作为不同位置的指纹表征。记录采样点位置坐标时,首先建立场景参考坐标系,将场景划分为 $M \times N$ 个栅格,并记录每个栅格中心点在该参考坐标系中的二维坐标,将第 i 个栅格中心点表示为 $l_i = (l_{ix}, l_{iy})$, l_{ix} 和 l_{iy} 分别表示 x 方向和 y 方向的参考坐标。

利用正交频分复用(orthogonal frequency division multiplexing, OFDM)技术将 Wi-Fi 信道划分为若干正交子信道。矢量格式的 CSI 数据可描述通信链路的信道属性,包含更细粒度、更多样化的物理层信息。CSI 信息不仅在时间上更稳定,而且在位置特征方面更具有代表性,因此,更适合作为 Wi-Fi 指纹用于高精度定位。采集细粒度物理层 30 个子载波

的 CSI 数据,每个采样点的 CSI 数据由一个复数矩阵 H_{MIMO} 表示,包含 CSI 特征、子载波和数据包 3 个维度的信息,可以表示为 $H_{MIMO} = \{H_1, H_2, \dots, H_n\}$ 。其中, $H_i = \{H(f_1), H(f_2), \dots, H(f_i) \dots, H(f_{30})\}$ 表示每个数据包中的子载波特征; n 为采集的数据包总数 ($n = 200$); $H(f_i) = |H(f_i)| e^{j\angle H(f_i)}$, $|H(f_i)|$ 与 $\angle H(f_i)$ 分别为幅值和相位。鉴于相位信息的不稳定性,采用幅值信息作为指纹特征。为了提高 CSI 数据的鲁棒性、减小采集过程中人为因素导致的数据异常值影响,在每一个采样点连续采集多个数据包,结合 HAMPPEL 滤波器与高斯滤波去除高频噪声与直流分量;为了减少数据冗余,根据 FIFS 方法^[13]对同一子信道内子载波的信道响应进行平均,将采样栅格中 n 个数据包的所有子信道的功率相加后取平均值作为 CSI 指纹存储在数据库中,公式如下:

$$H_{l_j} = \sum_{i=1}^n \frac{|H_i|}{n} \quad (1)$$

式中: H_{l_j} 为位置 l_j 处采样提取的 CSI 特征值; $|H_i|$ 为在 l_j 处第 i 次采样得到的各个子载波幅值之和。

2.2 视觉特征图层构建

1) 视觉特征提取

构建视觉特征图层时,在含安全出口标志的栅格中存储安全出口标志的视觉特征与参考位置的参考坐标。为了提升定位阶段的运行效率,在离线阶段完成视觉特征提取。为了获取丰富的场景信息,使用 ORB 与 SURF 两种特征描述符来计算安全出口标志的全局特征(全局特征用于描述图像的整体灰度特征,提升匹配速率以获得最接近地图栅格的图像;局部特征用于验证与几何计算)。在提取局部特征时采用 ORB 局部特征描述符;提取全局特征描述符时,首先将图像缩放为 63×63 像素的图像块,以图像块中心点作为特征点,缩减 FAST 特征提取耗时^[14]。相较于传统算法,该方法对光照有较强的鲁棒性。本文采用图像间的差异与给定阈值进行比较生成特征描述符,公式如下:

$$\tau(P; x, y) = \begin{cases} 1: P(x) - P(y) < \gamma \\ 0: P(x) - P(y) \geq \gamma \end{cases} \quad (2)$$

式中: $P(x)$ 是在图像 P 中点 x 的强度值; γ 为阈值。由此可以获得图像块的描述符:

$$f_n(P; x, y) = \sum_{i=1}^n 2^{i-1} \tau(P; x, y) \quad (3)$$

由于 SURF 对视角、光照、尺度变化等情形具有良好的鲁棒性,将图像归一化后的中心点作为 SURF 特征点。首先获取特征点附近圆形区域内的像素 x 与 y 方向的 Haar 小波响应,确定全局 SURF 特征点的方向;再将特征点附近区域划分为 4×4 个子区域,计算每个区域内 Haar 小波响应并赋予高斯系数;对每个子区域内的 dx 与 dy 及其绝对值分别求和,由此得到每个子区域的特征描述符 v :

$$v = (\sum dx, \sum dy, \sum |dx|, \sum |dy|) \quad (4)$$

最后将 16 个子区域的特征描述符组合,形成一个 64 维的特征向量:

$$\mathbf{V} = [v_1, v_2, \dots, v_{15}, v_{16}] \quad (5)$$

式中 $v_i (i=1, 2, \dots, 15, 16)$ 为子区域 i 的特征描述符。

2) 参考位置计算

为了计算 V-CSI 地图中视觉场景在参考系中

的位置,建立整个室内场景的参考坐标系,并标注安全出口标志 4 个顶点在场景参考坐标系中的坐标,基于小孔成像原理建立图像坐标系与参考坐标系之间的转换关系:

$$[u_m, v_m, 1]^T \cong \mathbf{K}[\mathbf{R} \mathbf{t}] [X_w, Y_w, Z_w, 1]^T \quad (6)$$

式中: $[u_m, v_m, 1]^T$ 为安全出口标志顶点在图像坐标系中的坐标; $[X_w, Y_w, Z_w, 1]^T$ 为安全出口标志顶点在场景参考坐标系中的坐标; \mathbf{K} 、 \mathbf{R} 、 \mathbf{t} 分别为相机内参矩阵、旋转矩阵与平移向量。

由于安全出口标志 4 个顶点处于同一物理平面,根据平面之间的单应性关系,设该单应性平面在 $Z=0$ 的平面坐标系中,转换关系如下:

$$[u_m, v_m, 1]^T \cong \mathbf{H} [X_h, Y_h, 1]^T \quad (7)$$

式中: $[u_m, v_m, 1]^T$ 和 $[X_h, Y_h, 1]^T$ 分别表示同一安全出口标志在图像坐标系和平面坐标系中的坐标; \mathbf{H} 为单应性矩阵。场景参考坐标系与平面坐标系之间的关系为

$$[X_w, Y_w, Z_w]^T = \mathbf{R}_p [X_h, Y_h, 0]^T + \mathbf{t}_p \quad (8)$$

式中: \mathbf{R}_p 与 \mathbf{t}_p 分别为旋转矩阵与平移向量。通过上式可将图像点通过单应性矩阵映射至场景参考坐标系,得到相机与安全出口标志之间的几何位姿关系,用于定位过程中的图像特征匹配与位姿计算。

3 基于 V-CSI 双层特征地图的 HMM 定位算法

将基于 V-CSI 双层特征地图的智能手机定位建模为 HMM 问题,通过前向算法获取有效的隐含状态,并与运动约束相融合获取精确、稳定的定位结果,如图 3 所示。具体为:将 V-CSI 地图中的栅格定义为隐含状态,将隐含状态集合记为 $Q = \{q_1, q_2, \dots, q_{n-1}, q_n\}$,其中 n 表示隐含状态的数量;将观测集合定义为 $O = \{o_1, o_2, \dots, o_{m-1}, o_m\}$, m 表示观测数;利用视觉特征定位结果初始化隐含状态,初始化概率记为 $\Pi = \{\pi_1, \pi_2, \dots, \pi_n\}$; $q_i \in Q$ 与 $o_i \in O$ 分别代表 t 时刻的状态与观测。此外,在运动过程中若安全出口标志被观测到,利用视觉特征获取的定位信息重新初始化 HMM,以消除迭代过程中的累积误差;而对于状态转移概率,利用历史定位结果构建基于高斯模型的状态转移模型计算;观测概率由 CSI 指纹匹配结果确定。

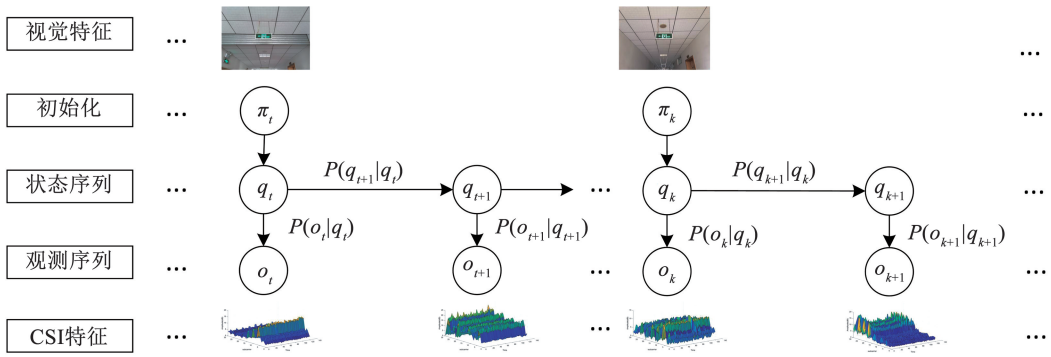


图 3 HMM 算法示意图

Fig. 3 Schematic diagram of HMM algorithm

3.1 HMM 模型初始化

1) 基于视觉特征图层的定位

首先利用 Yolov5 对入口处的安全标志进行检测,当检测到图像中有多个安全标志存在时,结合 CSI 指纹特征的粗定位结果缩小视觉地图匹配的范围,从而确定与当前位置最近的安全标志。为了减少视觉地图匹配耗时,使用 2.2 节中的全局 ORB 和 SURF 描述符来获取与当前图像最近的视觉地图图像。由于两种全局特征来自不同的特征空间,且具有不同的维度,为了消除传统 KNN 算法结果不唯一的问题、提升融合效果,本文从多个特征空间中遍历找到最相似的场景图像,然后使用局部特征匹配验证。

在图像匹配过程中,在全局 SURF 空间采用欧几里得距离作为度量,在全局 ORB 空间采用汉明距离作为度量,如公式(9)所示:

$$\begin{cases} D_o(R^1, R^2) = \sum_{i=1}^{256} \text{XOR}(R_i^1, R_i^2) \\ D_s(F^1, F^2) = \|F^1 - F^2\| = \sqrt{\sum_i (F_i^1 - F_i^2)^2} \end{cases} \quad (9)$$

式中: R^1, R^2 为两个输入的 ORB 全局特征; F^1, F^2 为两个输入 SURF 全局特征;下标 O 和 S 分别代表 ORB 和 SURF 全局空间; R_i^j 表示 ORB 特征的第 i 个元素; F_i^j 表示 SURF 特征 F^j 的第 i 个元素。分别在两个特征空间中对计算得到的 M 个和 N 个近邻赋予权重:

$$\begin{cases} W_o(R^i) = \alpha \frac{D_o(R, R^i)}{\sum_{i=1}^M D_o(R, R^i)} \\ W_s(F^j) = \beta \frac{D_s(F, F^j)}{\sum_{j=1}^N D_s(F, F^j)} \end{cases} \quad (10)$$

式中: $\alpha = \frac{M}{M+N}, \beta = \frac{N}{M+N}$ 分别为两种全局特征空

间的对应权重; $D_o(R, R^i)$ 表示查询图像与全局 ORB 特征空间第 i 张图像的汉明距离; $D_s(F, F^j)$ 表示查询图像与全局 SURF 特征空间第 j 张图像的欧几里得距离。对于从全局 ORB 特征空间得到的 M 个近邻图像,采用 $P = \{p_1, p_2, \dots, p_m, \dots, p_M\}$ 表示近邻图像对应的标签。同理,对于从全局 SURF 特征空间得到的 N 个近邻图像,采用 $G = \{g_1, g_2, \dots, g_n, \dots, g_N\}$ 表示近邻图像对应的标签;定义集合 $X = P \cup G = \{X_1, X_2, \dots, X_i, \dots\}$,遍历集合 X 找到权重最小的图像标签,即拍摄图像与视觉特征地图中某图像位置最近,再使用式(11)计算图像标签 X_i 对应的权重:

$$W(X_i) = \prod_{p_m=X_i} W_o(R^m) \times \prod_{g_n=X_i} W_s(F^n) \quad (11)$$

最后选择权值最小的近邻作为多个特征空间中最相似的场景图像:

$$G^* = \arg \min_X W(X_i) \quad (12)$$

由此得到查询图像与视觉特征地图中最相似的图像,获取对应安全出口标志的三维坐标,用于智能手机绝对位置计算。

在精度要求不高的室内定位场景中,可直接使用标志物的三维坐标作为用户当前位置。但实际上,安全出口标志三维坐标与智能手机位置存在差异,因此使用平面间的映射关系来计算相机与安全出口标志的位姿关系,从而提升定位精度。当前拍摄图像与视觉特征地图图像之间的单应性关系为

$$[u_m, v_m, 1]^T \cong H_{mq} [u_q, v_q, 1]^T \quad (13)$$

式中: $[u_m, v_m, 1]^T$ 为视觉特征地图图像的像素坐标; $[u_q, v_q, 1]^T$ 为查询图像的像素坐标; H_{mq} 为两个坐标系之间的单应性矩阵。通过下式的关系,可以获得当前拍摄图像与场景参考坐标系之间的转换关系:

$$[u_q, v_q, 1]^T \cong H_{mq}^{-1} H [X, Y, 1]^T = H_q [X, Y, 1]^T \quad (14)$$

基于小孔成像模型 $H_q \cong K [r_1, r_2, 1]^T$, (K 为拍摄相机的内参矩阵),计算拍摄图像的位姿:

$$r_1 = \frac{\mathbf{K}^{-1}\mathbf{H}_q^{(1)}}{\|\mathbf{K}^{-1}\mathbf{H}_q^{(1)}\|}, r_2 = \frac{\mathbf{K}^{-1}\mathbf{H}_q^{(2)}}{\|\mathbf{K}^{-1}\mathbf{H}_q^{(2)}\|},$$

$$r_3 = \text{cross}(r_1, r_2), t = \frac{\mathbf{K}^{-1}\mathbf{H}_q^{(3)}}{\|\mathbf{K}^{-1}\mathbf{H}_q^{(3)}\|} \quad (15)$$

然后计算相机在参考坐标系中的位置:

$$\text{pos} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} (-\mathbf{R}_p \mathbf{R}^{-1} \mathbf{t} + \mathbf{t}_p) \quad (16)$$

式中: $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$ 为正交投影矩阵, 通过上式可以获得相机当前的位置坐标, 然后基于欧几里得距离计算各状态概率作为 HMM 初始值。

2) 基于绝对位置的初始概率

为充分利用视觉定位结果, 在得到智能手机的绝对位置后, 建立基于视觉定位结果与状态栅格中心点欧几里得距离的高斯模型, 根据视觉定位结果对当前状态集合赋予不同的概率, 从而确定 HMM 的初始状态值。当前状态对应的初始概率 π_i 为

$$\pi_i = \frac{1}{\sqrt{2\pi}\sigma_\pi} \exp\left(-\frac{F(l_i, l_i^{(j)})^2}{2\sigma_\pi^2}\right) \quad (17)$$

式中: $F(l_i, l_i^{(j)})$ 表示 t 时刻绝对位置定位结果 l_i 与当视觉特征地图中第 j 个栅格中心点的欧几里得距离; σ_π 为根据环境选定的经验值。

3.2 状态转移概率建模

为了增强智能手机定位的连续性和稳定性, 传统方法利用智能手机内部传感器提供运动信息, 再经过 PDR 估计运动状态, 并与 Wi-Fi 信息融合, 但这种方法需要对异构传感器进行精确标定, 而且 PDR 还存在累积误差。为了兼顾模型复杂度与用户运动特征, 本文通过构建短时匀速运动模型, 利用智能手机上一时刻的状态估计当前状态, 如图 4 所示。

具体为, 设 $t-2$ 刻的状态为 $q_k = (l_{kx}, l_{ky})$, $t-1$ 时刻的状态为 $q_i = (l_{ix}, l_{iy})$, 则 t 时刻的状态为

$$\begin{cases} l_{tx} = l_{ix} + v_{t-1}^x \cdot \Delta t \\ l_{ty} = l_{iy} + v_{t-1}^y \cdot \Delta t \end{cases} \quad (18)$$

式中: $v_{t-1}^x = \frac{l_{ix} - l_{kx}}{\Delta t}$, $v_{t-1}^y = \frac{l_{iy} - l_{ky}}{\Delta t}$, 由 $t-1$ 和 $t-2$ 时刻的状态确定; Δt 由系统的定位频率确定。然而在运动过程中不可避免存在噪声, 为了更好地描述 $t-1$ 时刻到 t 时刻的状态转移概率, 构建基于二维高斯分布的状态转移概率模型, t 时刻状态 $q_j = (l_{jx}, l_{jy})$ 状态转移概率为

$$a_{ij} = \frac{1}{2\pi\sigma_a^2} \exp\left(-\frac{(l_{jx} - l_{ix})^2 + (l_{jy} - l_{iy})^2}{2\sigma_a^2}\right),$$

$$\sum_{j=1}^n a_{ij} = 1, \forall i \quad (19)$$

式中: l_{ix}, l_{iy} 由式 (18) 获得; σ_a 为根据用户移动速度设定的经验值, 本文取 $\sigma_a = 0.5$; a_{ij} 为 $t-1$ 时刻状态 i 转移到 t 时刻状态 j 的概率。

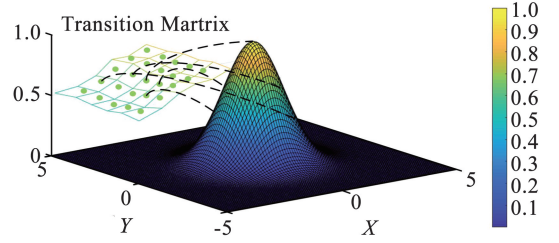


图 4 二维高斯状态转移概率

Fig. 4 Two dimensional Gaussian state transition probability

3.3 发射概率建模

HMM 模型中的观测 o_t 表示在当前状态下观测到的 CSI 特征。使用欧几里得距离描述当前状态下的观测与地图中存储的 CSI 指纹特征之间的相似性。为体现“欧几里得距离越大 (CSI 指纹相似性降低), 发射概率越小”的规律, 使用 0 均值的高斯模型描述当前位置观测与 CSI 指纹地图中不同指纹所在栅格之间的差异。基于 CSI 指纹的发射概率为

$$b_j(o_t) = \frac{1}{\sqrt{2\pi}\sigma_b} \exp\left(-\frac{F(o_t, o^{(j)})^2}{2\sigma_b^2}\right), \sum_{j=1}^n b_j(o_t) = 1 \quad (20)$$

式中: $F(o_t, o^{(j)})$ 表示当前观测 o_t 与地图中对应栅格的指纹特征 $o^{(j)}$ 的欧几里得距离; σ_b 为标准差。

3.4 基于 V-CSI 双层特征地图的位置计算

本文使用前向算法计算 t 时刻各状态的前向概率, 选取最佳的 M 个栅格作为 t 时刻的最优状态, 通过加权平均得到 t 时刻用户在参考坐标系中的位置。为了降低序列匹配的复杂度, 定义 t 时刻的前向算子为

$$\alpha_t(j) = P(q_t = l_j, q_{t-1} = l_i, O_{1:t}) \quad (21)$$

式中: j, i 分别表示 t 时刻、 $t-1$ 时刻的状态序列; l_j, l_i 分别为对应的状态坐标, 则由链式法则展开为

$$\alpha_t(j) = \underbrace{\sum_i P(q_t = l_j | q_{t-1} = l_i)}_{\text{状态转移概率}} \cdot \underbrace{P(O_t | q_t = l_j)}_{\text{观测概率}} \alpha_{t-1}(i); 1 \leq j \leq N, 1 \leq t \leq T \quad (22)$$

对求得的前向概率进行降序排列, 并选取最大的 M 个概率所在的栅格作为加权平均的候选状态, 并对这 M 个最优状态进行归一化:

$$\beta_t(j) = \frac{\alpha_t(j)}{\sum_{j=1}^M \alpha_t(j)} \quad (23)$$

则 t 时刻智能手机的位置为

$$L_t^* = \sum_{j=1}^M \beta_t(j) l_j \quad (24)$$

式中： L_i^* 为加权平均后所获得的智能手机的位置； $\beta_i(j)$ 是归一化后的最优状态系数； l_j 为最优状态对应栅格的中心点坐标。

4 实验结果与分析

4.1 实验设计

为了验证本文提出算法的可行性,在 Android 手机平台上开发定位软件,并在两种不同的室内典型场景进行测试:一种场景是人员流动性较大的武汉市某大学办公室走廊;另一种场景是环境较开阔的武汉市某酒店地下停车场。本实验将定位场景划

分为 $1.2\text{ m} \times 1.2\text{ m}$ 的栅格。测试时,实验人员用不同类型的智能手机(谷歌 Nexus 5;四核 2.3 GHz;三星 SM-T710;四核 1.9 GHz)安装定位软件对算法的可靠性进行评估。图 5(a)、(c)分别为办公楼 4、5 楼(布局一致)与地下停车场的 CAD 平面图,建筑面积分别为 6 000 平方米和 3 600 平方米。办公楼 4 楼有 20 个安全出口标志,5 楼有 22 个安全出口标志,共计 42 个安全出口标志;停车场有 29 个安全出口标志。图 5(b)、(d)分别为通过 YOLOv5 检测安全出口标志与实验人员测试示意图。

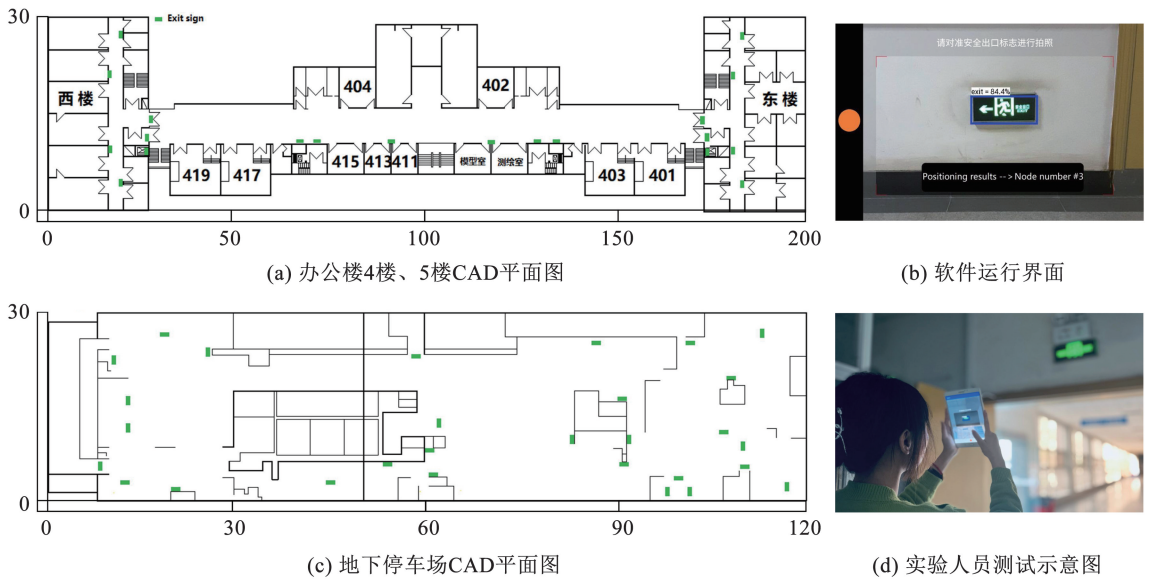


图 5 实验场景与测试示意图

Fig. 5 Schematic diagram of experiment scenarios and tests

建图时,按照本文方法处理所采集的数据,生成 V-CSI 双层特征地图,包括基于 CSI 指纹特征图层的指纹信息和位置坐标、视觉特征图层的图像数据和安全出口标志参考坐标;地图采样点的位置坐标均由激光雷达获取。其中,CSI 数据按 2.1 节方法处理子载波幅值;图像数据包括场景图像、全局特征、局部特征与单应性矩阵。图 6 为实验场景所采集的 CSI 数据和图像数据,采集的图像分辨率为 800×600 。考虑到不同分辨率的栅格会影响定位

的精度(分辨率过小会增加数据量,过大会降低定位精度),结合定位性能综合选择:办公楼场景采用 $1.2\text{ m} \times 1.2\text{ m}$ 栅格,停车场场景采用 $1.0\text{ m} \times 1.0\text{ m}$ 栅格。办公楼场景的特点是人员流动性较强、信号干扰较大,且安全出口标志分布不均(最小间隔 3 m,最大间隔 40 m);停车场场景的特点是光照效果不佳、相似场景较多,且难以提取有效特征点,单一传统的 Wi-Fi 定位或视觉定位均难以达到理想效果。

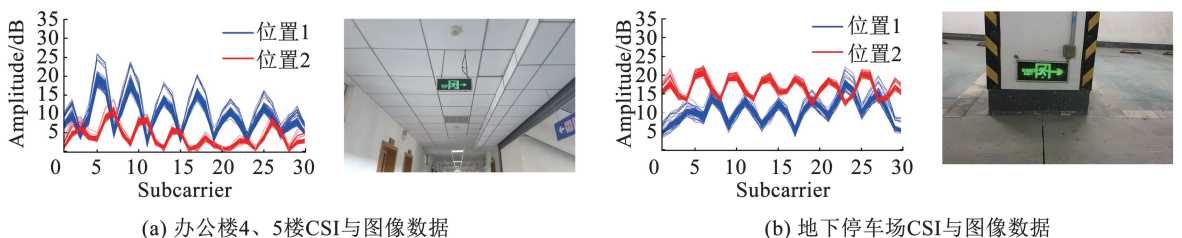


图 6 实验中采集的 CSI 数据与图像数据

Fig. 6 CSI data and image data collected in the experiment

4.2 办公楼场景测试结果

实验人员使用开发的 Android 定位软件进行测试,用户行走轨迹如图 7 所示。其中,蓝线表示规划的真实路线 (Ground Truth), 红线表示本文算法定位结果, 浅蓝色线表示使用基于 CSI 指纹的 KNN ($K=4$) 定位结果, 实验人员行走的路径约 45 m。实验过程:按 3.3 节方法拍摄并识别安全出口标志, 计算用户初始位置;采集当前用户的 Wi-Fi 信息, 提取其 CSI 特征, 按第 3 节构建 HMM 模型得到下一时刻用户的位置;遇到安全出口标志时, 通过智能手机拍摄图像, 按 3.1 节方法进行图像特征匹配, 返回 V-CSI 地图中最接近图像的索引, 计算单应性矩阵并更新 HMM 定位结果。

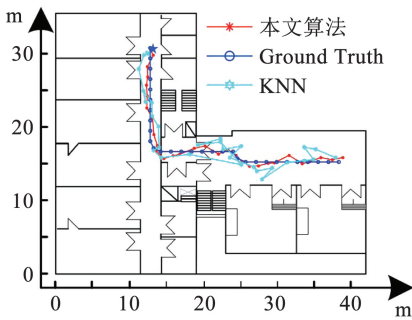


图 7 行走轨迹

Fig. 7 Walking track

为了验证视觉定位的有效性,对场景中安全出口标志进行编号,选取 15 个安全出口标志的识别结果进行统计分析,不同位置的安全出口标志识别准确率如图 8 所示,办公楼的图像识别准确率为 96.7%。虽然《中华人民共和国消防法》规定了安全出口标志的设置要求,但安全出口标志在不同场景下的密度分布并不均衡。在办公楼场景中,安全出口标志的最大间距超过 20 m,最小间距仅为 2.5 m。为了分析安全出口标志密度对定位精度的影响,以“每平方米安全出口标志数”为度量,开展 5 次独立算法验证实验。在办公楼实验场景中,安全出口标志的最小密度为 0.025 个/ m^2 (每 40 平方米 1 个),最大密度为 0.15 个/ m^2 (每 40 平方米 6 个),不同标志密度对定位准确性的影响如图 9 所示。结果表明,随着安全出口标志密度的增大,定位误差减小,定位准确性提升,由此证明准确的视觉定位结果能够有效提升 HMM 序列匹配的精度。

视觉定位结果返回与拍摄图像最相似的视觉特征地图图像,并根据匹配结果计算当前拍摄图像与视觉特征地图之间的单应性矩阵 H_{mq} , 根据式 (14) 计算拍摄图像与场景参考坐标系之间的转换关系

H_q , 得到视觉定位结果。从表 1 可以看出,视觉定位平均定位误差较小,但耗时较长;CSI 指纹定位耗时短,但平均定位误差较大;本文的融合定位方法耗时与 CSI 指纹定位接近,平均误差显著降低,在定位精度和效率方面能够很好地满足实际应用需要。

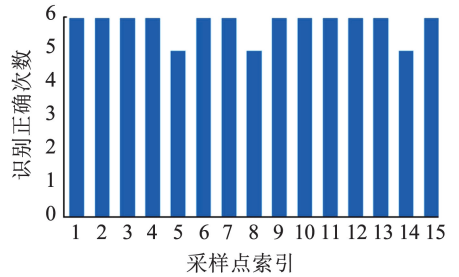


图 8 安全出口标志识别准确率

Fig. 8 Recognition accuracy of exit signs

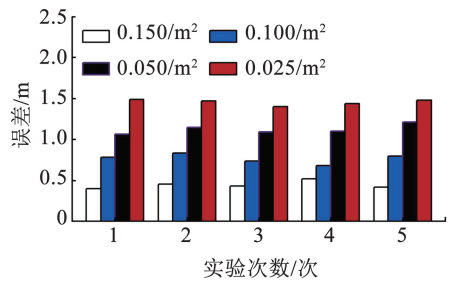


图 9 不同安全出口标志密度定位误差

Fig. 9 Positioning error under different exit sign densities

表 1 三种算法定位结果

Tab. 1 Positioning results of three algorithms

算法	定位耗时/s	平均定位误差/m
视觉定位	0.39	0.40
CSI 指纹定位	0.10	2.50
融合定位	0.15	1.04

表 2 显示了不同融合算法的误差情况。可以看出,在不明显增加定位时间的情况下,相比于文献 [15] 所提出的视觉定位算法、文献 [16] 提出的基于深度学习的指纹定位算法、文献 [17] 提出的基于 CSI 特征融合的定位算法,本文算法定位精度分别提升了 22.4%, 42.9%, 19.4%。

表 2 不同融合算法误差对比

Tab. 2 Error comparison of different fusion algorithms

算法	定位耗时/s	平均定位误差/m
文献 [15]	0.17	1.34
文献 [16]	0.19	1.82
文献 [17]	0.13	1.29
本文算法	0.15	1.04

图 10 为 CSI 指纹定位、CSI + KNN、CSI + HMM

与本文算法的累积误差分布图。结合表 1 中的数据可以看出,在办公楼场景中,由于环境对信号传播的影响,CSI 指纹定位最大定位误差 5.7 m,平均定位误差 2.5 m,误差小于 1.5 m 的仅有 21.7%;本文提出的算法最大定位误差有 3.1 m,平均定位误差 1.04 m,误差小于 1.5 m 的有 78.2%,相较于单一 CSI 指纹定位,平均定位误差减小了 58.4%。

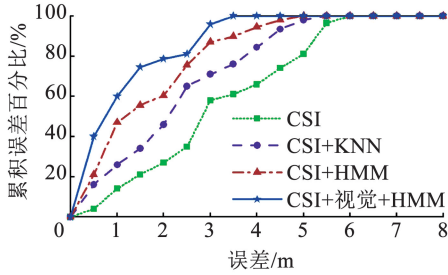


图 10 不同算法误差对比

Fig. 10 Error comparison between different algorithms

4.3 停车场场景测试结果

实验场景环境开阔,且安全出口标志多在高处或主要路径的醒目位置,不易被遮挡,但也存在光照条件较差、标志间距大等问题。实验中测试人员行走轨迹如图 11 所示,实验路径长度约 40 m;场景采样点间距为 1 m,按第 2 节方法建立 CSI 指纹与视觉地图。

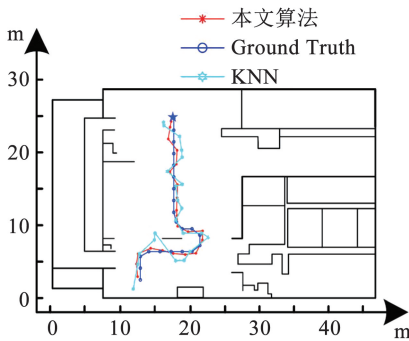


图 11 行走轨迹

Fig. 11 Walking track

为了验证视觉定位的有效性,对停车场场景下的安全出口标志进行编号,选取 15 个安全出口标志统计图像识别准确率,结果如图 12 所示。从图 12 可以看出,停车场的安全出口标志识别准确率为 94.5%。在停车场场景中,安全出口标志间距较大,平均间距为 10 m,最小间距仅为 5 m。为了分析安全出口标志密度对定位精度的影响,同样使用“每平方米安全出口标志数”为度量指标。不同于办公楼实验场景,由于停车场场景开阔,在此场景中统计 50 平方米内安全出口标志的数量。在本场景中,安全出口标志的最小密度为 0.02 个/m²(每 50 平方

米 1 个),最大密度为 0.08 个/m²(每 50 平方米 4 个),不同标志密度下定位准确性如图 13 所示。结果表明,安全出口标志密度越大,平均定位误差越小。从表 3 可知,视觉定位平均定位误差较小,但耗时较长;CSI 指纹定位耗时短,但平均定位误差较大;本文融合定位方法耗时接近 CSI 指纹定位,同时大幅降低了定位误差。

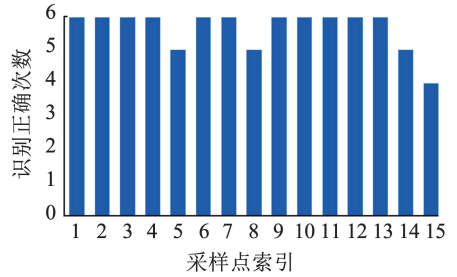


图 12 安全出口标志识别准确率

Fig. 12 Recognition accuracy of exit signs

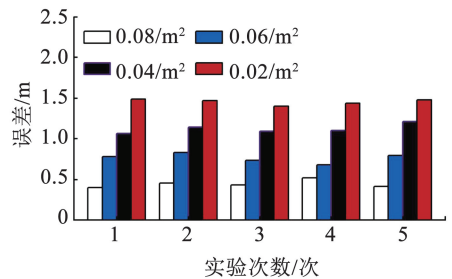


图 13 不同安全出口标志密度定位误差

Fig. 13 Positioning error under different exit sign densities

表 3 视觉定位与 CSI 指纹定位结果

Tab. 3 Analysis of visual and CSI fingerprint positioning results

算法	定位耗时/s	平均定位误差/m
视觉定位	0.43	0.47
CSI 指纹定位	0.11	2.12
融合定位	0.17	0.92

为了验证本文提出算法的有效性,将其他公开文献中的方法应用到本实验场景中,将多次实验的定位结果进行对比,结果如表 4 所示。相较于文献 [15]、文献 [16]、文献 [17] 算法,本文算法定位精度分别提升了 42.5%、26.4%、18.6%,且定位耗时无显著增加。

表 4 不同融合算法误差对比

Tab. 4 Error comparison of different fusion algorithms

算法	定位耗时/s	平均定位误差/m
文献 [15]	0.25	1.60
文献 [16]	0.19	1.25
文献 [17]	0.14	1.13
本文算法	0.17	0.92

图14为CSI指纹定位、CSI+KNN、CSI+HMM与本文方法的累积误差分布图。结合表3中的数据可以看出,相较于办公楼场景,该场景信号干扰较小,但光照对视觉定位影响较大。其中,单一的CSI最大定位误差有5.14 m,平均定位误差2.12 m,误差小于1.5 m的仅有24.4%,而本文方法最大定位误差3.01 m,平均定位误差0.92 m,误差小于1.5 m的有82.7%,相比于单一的CSI指纹定位算法,本文方法平均定位误差减小了56.6%。由此证明,本文方法在没有显著增加定位耗时的情况下,可以有效提升定位精度、降低误差。

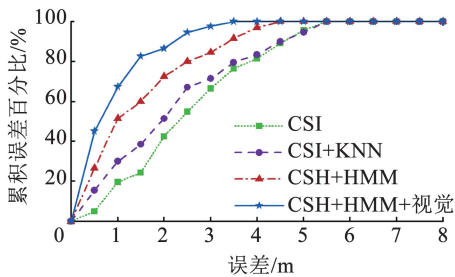


图14 不同算法误差对比

Fig. 14 Error comparison between different algorithms

5 结 语

利用CSI指纹特征和安全出口标志的视觉特征构建V-CSI双层特征地图,提出一种基于V-CSI双层特征地图的HMM智能手机高精度室内定位方法。通过视觉特征层完成HMM模型的初始化与重新初始化,利用CSI指纹特征获取当前状态的发射概率,并运用前向算法获取有效状态,通过有效状态加权平均获得智能手机的位置信息。实验结果表明,本文方法平均定位误差约为1.0 m,相比于单一的CSI指纹定位,平均定位误差降低56%以上,且在信号干扰小、安全出口标志分布密集的场所,本文方法可进一步降低定位误差。同时,本文方法还可与目前主流的定位方法融合,为解决GPS盲区下的室内定位提供新思路。

参 考 文 献

[1] 赵林生,王鸿鹏,刘景泰. 人机共享环境下基于Wi-Fi指纹的室内定位方法[J]. 机器人, 2019, 41(3): 404
ZHAO Linsheng, WANG Hongpeng, LIU Jingtai. An indoor localization method based on Wi-Fi fingerprint in the human-robot shared environment [J]. Robot, 2019, 41(3): 404. DOI: 10.13973/j.cnki.robot.180307

[2] 丁致远. 室内定位及路径规划算法研究[D]. 哈尔滨: 哈尔滨工业大学, 2020
DING Zhiyuan. Research on indoor positioning and path planning algorithms [D]. Harbin: Harbin Institute of Technology, 2020

[3] 李华亮. 高鲁棒性RSSI室内指纹定位算法研究[D]. 吉林: 吉林大学, 2021
LI Hualiang. Research on RSSI indoor fingerprint location algorithm

with high robustness [D]. Jilin: Jilin University, 2021

[4] 于海涛. 基于RSS与CSI混合指纹室内定位研究[D]. 哈尔滨: 哈尔滨工业大学, 2017
YU Haitao. Indoor locating method research based on RSS and CSI fingerprint [D]. Harbin: Harbin Institute of Technology, 2017

[5] 胡判政,刘佳慧,黄刚,等. 融合Wi-Fi、激光雷达与地图的机器人室内定位[J]. 电子与信息学报, 2021, 43(8): 2308
HU Zhaozheng, LIU Jiahui, HUANG Gang, et al. Adaptive affine propagation clustering algorithm for Wi-Fi indoor positioning [J]. Journal of Electronics & Information Technology, 2021, 43(8): 2308. DOI: 10.11999/JEIT200671

[6] ALWIN P, BENAUMEUR S, DONG S H. Performance analysis of sensor fusion techniques for heading estimation using smartphone sensors [J]. IEEE Sensors Journal, 2019, 19(24): 12369. DOI: 10.1109/JSEN.2019.2940071

[7] TAO Xianlu, ZHANG Xiaohong, ZHU Feng, et al. Precise displacement estimation from time-differenced carrier phase to improve PDR performance [J]. IEEE Sensors Journal, 2018, 20(18): 8238. DOI: 10.1109/JSEN.2018.2867225

[8] HUANG Gang, HU Zhaozheng, WU Jie, et al. Wi-Fi and vision integrated fingerprint for smartphone-based self-localization in public indoor scenes [J]. IEEE Internet of Things Journal, 2020, 7(8): 6748. DOI: 10.1109/JIOT.2020.2974928

[9] WANG Zhongshuai, SOKLIEP P, XU Chengpei, et al. Indoor position algorithm based on the fusion of Wi-Fi and image [C]//The 11th International Conference on Advanced Computational Intelligence. Guilin, China: IEEE, 2019: 212. DOI: 10.1109/ICACI.2019.8778542

[10] ZHAO Yanchao, XU Jing, WU Jie, et al. Enhancing camera-based multimodal indoor localization with device-free movement measurement using Wi-Fi [J]. IEEE Internet of Things Journal, 2020, 7(2): 1024. DOI: 10.1109/JIOT.2019.2948605

[11] SUN Meng, WANG Yunjia, LIU Keqiang, et al. Smartphone-based Wi-Fi FTM fingerprinting approach with map-aided particle filter [C]// Proceedings of the 2022 IEEE 12th International Conference on Indoor Positioning and Indoor Navigation (IPIN). Beijing, China, 2022: 1. DOI: 10.1109/IPIN54987.2022.9918110

[12] NASHED S B, ILSTRUP D M, BISWAS J. Localization under topological uncertainty for lane identification of autonomous vehicles [C]//2018 IEEE International Conference on Robotics and Automation (ICRA). Brisbane, Australia, 2018: 6000. DOI: 10.1109/ICRA.2018.8461185

[13] XIAO Jiang, WU Kaishun, YI Youwen, et al. FIFS: Fine-grained indoor fingerprinting system [C]//Proceedings of the 2012 21st International Conference on Computer Communications and Networks (ICCCN) Munich, Germany, 2012. DOI: 10.1109/ICCCN.2012.6289200

[14] 刘国忠,胡判政. 基于SURF和ORB全局特征的快速闭环检测[J]. 机器人, 2017, 39(1): 36
LIU Guozhong, HU Zhaozheng. Fast loop closure detection based on holistic features from SURF and ORB [J]. Robot, 2017, 39(1): 36. DOI: 10.13973/j.cnki.robot.2017.0036

[15] 徐昊玮,廉保旺,邹晓军,等. 基于视觉物体识别的抗差岭估计定位算法[J]. 电子与信息学报, 2018, 40(10): 2453
XU Haowei, LIAN Baowang, ZOU Xiaojun, et al. Visual objects detection based robust ridge regression indoor localization method [J]. Journal of Electronics & Information Technology, 2018, 40(10): 2453. DOI: 10.11999/JEIT170876

[16] WANG Xuyu, GAO Lingjun, MAO Shiwen, et al. DeepFi: Deep learning for indoor fingerprinting using channel state information [C]//Proceedings of the 2015 IEEE Wireless Communications and Networking Conference (WCNC). New Orleans, LA, USA, 2015: 1666. DOI: 10.1109/WCNC.2015.7127718

[17] SANAM T F, GODRICH H. FuseLoc: A CCA based information fusion for indoor localization using CSI phase and amplitude of Wi-Fi signals [C]//Proceedings of the 2019 IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Brighton, UK, 2019: 7565. DOI: 10.1109/ICASSP.2019.8683316