

An Improved Ordinary Kriging Algorithm Fusion of Multidirectional Variability*

LIU Xiaoyu, DENG Ping*

(Key Lab of Information Coding and Transmission, Southwest Jiaotong University, Chengdu 611756, China)

Abstract: In the practice of indoor fingerprint localization technology, the application of Kriging interpolation algorithm can efficiently reduce the workload of fingerprint database establishment. However, the Kriging Interpolation algorithm usually only cover the signal variability in one dimension direction, the variabilities of indoor signals in different dimension direction are not considered. Therefore in this paper, an improved Kriging interpolation algorithm which fuses signal variability in different direction is proposed. The effects of signal variabilities in two dimension indoor environment to the accuracy of interpolation are fully considered. Experiment results show that compared with Kriging interpolation algorithm, Inverse-distance weighting algorithm, the proposed algorithm can get more precisely interpolation data of fingerprint.

Key words: indoor localization; establishment of fingerprint database; Kriging interpolation algorithm; variability
EEACC: 7230 doi: 10.3969/j.issn.1004-1699.2020.07.011

一种融合多方向变异性的改进克里金插值算法*

刘晓宇, 邓平*

(西南交通大学信息编码与传输重点实验室, 成都 611756)

摘 要: 在室内位置指纹定位技术中, 采用克里金插值算法可有效降低指纹数据库构建的工作量。但是, 现有算法通常只考虑了信号在一维方向上的变异性, 没有考虑室内环境中信号在多个方向存在的变异性。为此, 本文提出了一种融合了多个方向信号变异性的改进克里金插值算法, 充分考虑了二维室内环境中信号在不同方向的变异性对插值精度的影响。实验结果表明, 与克里金插值算法、距离反比加权插值算法等算法相比, 本文算法对指纹数据的插值精度更高。

关键词: 室内定位; 指纹数据库构建; 克里金插值算法; 变异性

中图分类号: TP393

文献标识码: A

文章编号: 1004-1699(2020)07-0991-06

目前基于位置指纹的 WIFI 室内定位技术是当下的研究热点之一^[1-2]。然而, 离线阶段中指纹数据库的构建却是一项艰巨的任务, 目前主要使用人工逐点采集的方式进行, 这将耗费大量的人力资源。此外, 在许多情况下, 如办公室或公寓等场所, 数据采集时常受到限制或无法访问。因此, 通过空间插值算法来构建指纹数据库不仅可以减少人力输出, 还可以解决特殊点或特殊环境下指纹数据采集困难等问题。

为解决上述问题, 文献[3]提出线性插值法来构建指纹数据库, 但该方法存在基点处不光滑、插值精度低的问题; 文献[4]提出径向基函数法来确定

周围已知点到插值点的权重, 并考虑在不同的插值形式下采用不同的基函数, 但该方法仅适用于散乱数据。文献[5]提出距离反比加权(Inverse Distance Weighting, IDW)插值法构建指纹库, 该方法实现简单, 但没有考虑数据在空间中的分布情况, 可能会造成估计偏差较大的情况。近年来, 克里金插值算法成为了空间插值算法研究的重点^[6-7]。文献[7-9]将克里金算法作为一种插值方法(包括测量结果)进行了检验, 并显示出插值精度的提高; 文献[10]根据模拟数据评估了克里金插值算法的高效性能。虽然克里金插值算法可有效提升插值精度, 但在二维室内空间环境中, 通常以某一方向下的空间变异

项目来源: 国家自然科学基金项目 61871332)

收稿日期: 2020-04-12 修改日期: 2020-07-07

性来表征信号在不同方向的变异性,这必然导致插值点处的估计指纹数据精度下降。

为此,为提升克里金插值算法的指纹插值精度,本文基于二维室内环境,提出了一种混合多方向变差函数影响的改进克里金插值算法,并通过仿真验证了本文算法的有效性。

1 基于插值的位置指纹数据库构建方法

1.1 距离反比加权插值算法

距离反比加权插值算法是一种典型的加权平均插值算法,目前已被广泛应用于指纹数据库的插值重构中。其实现原理是基于位置距离的相近相似^[11],以待插值点与其邻域内的已知观测点之间的欧式距离的倒数为权重,实现待插值点处的属性值估计。具体实现为:

$$z_m = \sum_{i=1}^n \omega_i z_i \quad (1)$$

$$\omega_i = \frac{1}{d_i^k} \bigg/ \sum_{i=1}^n \frac{1}{d_i^k} \quad (2)$$

$$d_i = \sqrt{(x_i - x_m)^2 + (y_i - y_m)^2} \quad (3)$$

式(1)~式(3)中: z_m 表示待插值点 m 的属性值; z_i 为周围已知观测点的属性值; ω_i 为权重系数; d_i 表示第 i 个观测点与 m 点坐标之间的欧式距离; k 为距离的反比幂指数,它的大小决定了权重随着距离的增减而变化的速率,通常取2。距离反比加权插值算法的优点是计算速度快、易于实现,但该算法仅考虑了待插值点与观测点坐标之间的位置距离关系,忽略了各点在空间上的相对分布,因此距离反比加权插值算法存在一定的应用局限性。

1.2 克里金插值算法

克里金插值算法是由法国科学家乔治·马瑟顿(Georges Matheron)提出的一种最优空间插值算法,目前主要应用于地质统计学领域^[8]。其实现原理可以描述为:以区域化变量为研究对象,依据变差函数理论,探寻空间数据之间的随机性和空间相关性,并以此进行最优线性无偏估计。与距离反比加权插值算法相似,克里金插值算法同样采取加权求平均的插值方式,计算公式为:

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (4)$$

式中: $Z^*(x_0)$ 为待插值点 x_0 处的属性值; $Z(x_i)$ 为已知观测点 x_i 处的属性值; λ_i 为权重系数。不同于距离反比加权插值算法,克里金插值算法的权重系数求解不仅考虑了待插值点与已知观测点之间的位置距离关系,还通过变异理论从整体信号空间变异

结构的角来考量数据点间的相关性^[11],因此较距离反比加权插值算法而言具有更好的鲁棒性。具体地,克里金插值算法的权重系数 λ_i 的求解可以概括为两阶段进行:第一阶段是求取空间区域化变量对应的变差函数;第二阶段是克里金插值。

1.2.1 变差函数的求解

变差函数也称之为半方差函数,主要用来表示区域化变量的空间结构特征。而区域化变量主要指的是区域内所在位置有关的随机变量。具体地,在本文中指的是室内环境中不同位置处的接收信号强度指示(Received Signal Strength Indication, RSSI)值。变差函数的求解是克里金插值算法的核心,主要包含三个步骤:首先利用空间中已知观测点的位置坐标和RSSI信号值计算实验变差函数值;然后根据实验变差函数值的分布情况,选择一种合适的理论变差函数模型;最后利用现有的拟合算法和所选取的理论变差函数模型对实验变差函数值进行数据拟合。

①实验变差函数值

实验变差函数值的求解通常需首先选定空间中的某一方向;然后设定基本滞后距 h 的取值,一般取空间中各观测点的基本间距;最后再计算不同位置距离 ih 下的实验变差函数值,用于后续理论变差函数的拟合。具体计算如下:

$$\gamma'(ih) = \frac{1}{2N(ih)} \sum_{k=1}^{N(ih)} [Z(x_k + ih) - Z(x_k)]^2 \quad (5)$$

式中: $i=1, 2, \dots, m$,其中 m 表示我们设定的最大滞后距间隔的个数; $Z(x_k + ih)$ 和 $Z(x_k)$ 分别对应为空间中坐标点 $x_k + ih$ 和 x_k 处的RSSI信号值; $N(ih)$ 表示被距离 ih 隔开的点对数,点对数值越大,表示求解得到的实验变差函数值越准确,反之误差较大。

②变差函数模型的参数拟合

为了可以对空间中任意点处的RSSI信号值进行估计,需要找到一个合适的理论变差函数模型来拟合我们所求得的一组实验变差函数值 $\gamma'(ih)$ ^[12]。理论变差函数模型主要分为两种:无基台模型和有基台模型。其中,在指纹数据插值过程中使用最为广泛的是有基台的球状模型,具体表达式为:

$$\gamma(h) = \begin{cases} 0 & h=0 \\ C_0 + C \left(\frac{3}{2} \cdot \frac{h}{a} - \frac{1}{2} \cdot \frac{h^3}{a^3} \right) & 0 < h < a \\ C_0 + C & h \geq a \end{cases} \quad (6)$$

式中: C_0 表示块金常数,主要与测量误差有关; C 表示拱高,指空间中变量变异性幅度的大小,而 $C_0 + C$ 就代表着变量在空间上的总变异大小; a 表示变程,

指变量在空间中具有相关性的范围。在变程范围之内,数据之间存在相关性;反之则数据之间互不影响,即变程 a 可以表示为变量在空间中的影响范围。

目前,现有的拟合算法包括加权最小二乘法和加权线性规划法。加权最小二乘法的实现原理是将球状模型的表达式转换为二元线性问题,然后根据加权最小二乘原理,对模型参数 C_0 、 C 以及 a 进行求解。

1.2.2 克里金插值

克里金插值的实现需首先确定用于插值估计的已知观测点,一般以坐标点的欧式距离为衡量标准,选取离待插值点位置距离最近的前 n 个点作为已知观测点;然后再利用已知观测点信息和变差函数确定克里金方程组,进而求得权重系数 λ_i ;最终利用 λ_i 加权求和得到克里金插值结果。其中,克里金方程组的构建是克里金插值部分的关键,具体实现如下。

由于克里金插值算法是一种无偏最优插值,估计无偏性和估计方差最小成为权值 λ_i 的选择标准,因此需要满足以下两个条件:

$$\begin{cases} E[Z^*(x_0) - Z(x_0)] = 0 \\ \text{Var}[Z^*(x_0) - Z(x_0)] = \min \end{cases} \quad (7)$$

区域内研究变量符合二阶平稳假设,由式(7)中的无偏条件可得:

$$\sum_{i=1}^n \lambda_i = 1 \quad (8)$$

同时,根据方差最小特征有:

$$\delta_{\min} = \text{Var}[Z^*(x_0) - Z(x_0)] = E\{[Z^*(x_0) - Z(x_0)]^2\} \quad (9)$$

然后利用拉格朗日乘数法求条件极值我们可以得到:

$$\frac{\partial}{\partial \lambda_i} [\delta_{\min} - 2\mu \sum_{i=1}^n \lambda_i] = 0 \quad (10)$$

最后通过推导可得 $n+1$ 阶线性方程组:

$$\sum_{i=1}^n C(x_i, x_j) \lambda_i - \mu = C(x_0, x_j) \quad (11)$$

式中:式(11)中 $C(x_i, x_j)$ 表示已知观测点之间的协方差, $C(x_0, x_j)$ 表示待插值点与已知观测点之间的协方差; μ 为拉格朗日函数因子。需要注意的是:协方差 $C(x_i, x_j)$ 和 $C(x_0, x_j)$ 无法通过已知观测点所对应的 RSSI 信号值求得,因此需要将协方差转换成空间变量的变异特性,即使用变差函数对协方差进行替换来最终求取权重系数 λ_i 。

当空间中变量满足二阶平稳条件时,变差函数和协方差函数存在一定的数据转换关系。令变量 x_i

和 x_j 对应的变差函数为 $\gamma(x_i, x_j)$,则根据变差函数定义有:

$$\gamma(x_i, x_j) = \frac{E[Z(x_i) - Z(x_j)]^2}{2} \quad (12)$$

$$\gamma(x_i, x_j) = -C(x_i, x_j) \quad (13)$$

再根据式(8)和式(11),可将克里金方程组表示为:

$$\begin{cases} \sum_{i=1}^n \gamma(x_i, x_j) \lambda_i + \mu = \gamma(x_0, x_j) \\ \sum_{i=1}^n \lambda_i = 1 \end{cases} \quad i, j = 1, 2, \dots, n \quad (14)$$

将式(14)写成矩阵形式有:

$$\begin{cases} [K'] [\lambda'] = [M'] \\ [K'] = \begin{bmatrix} \gamma(x_1, x_1) & \cdots & \gamma(x_1, x_n) & 1 \\ \vdots & \cdots & \vdots & \vdots \\ \gamma(x_n, x_1) & \cdots & \gamma(x_n, x_n) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \\ [\lambda'] = [\lambda_1, \lambda_2, \dots, \lambda_n, \mu]^T \\ [M'] = [\gamma(x_1, x_0), \dots, \gamma(x_n, x_0), 1]^T \end{cases} \quad (15)$$

式中:变差函数值 $\gamma(x_i, x_j)$ 可以通过拟合得到的变差函数求得。然后通过求解矩阵方程得到克里金方程组对应的权重系数,代入到式(4)中得到待插值点的 RSSI 信号估计值 $Z^*(x_0)$ 。

2 融合多方向变异性的改进克里金插值算法

为进一步提升基于克里金插值算法的指纹数据库插值精度,本文在克里金插值算法的基础上,提出了一种鲁棒性更好、精度更高的改进克里金插值算法。算法的改进主要体现在以下两个方面:

2.1 引入容限的实验变差函数值求解

当空间中样本点数量较稀疏时,求解实验变差函数值时会出现某一间距 ih (h 为基本滞后距)所对应的点对数 $N(ih)$ 值较小的情况,这时求解得到的实验变差函数值 $\gamma'(ih)$ 的准确性将降低。如果使用了准确性较低实验变差函数值来拟合理论变差函数,不仅会导致理论拟合效果变差,更会直接影响克里金插值估计的准确性,因此确保实验变差函数值求解的准确性十分重要。为解决该问题,本文引入了距离容限 Δh 和角度容限 $\Delta \varphi$,将空间中稀疏的样本点按照区间 $[ih - \Delta h, ih + \Delta h]$ 组成距离组,再将数据按照角度区间 $[\alpha - \Delta \varphi, \alpha + \Delta \varphi]$ 组成角度组。这样,凡是落在角度范围 $\alpha \pm \Delta \varphi$ 以及距离范围 $ih \pm \Delta h$

内的样本点都可以认为是有效点。通过该种方式,可以有效明显地改善点对数 $N(ih)$ 值较小的问题,增大我们所求实验变差函数值的准确性。在本文中,取 $\Delta\varphi = \pi/8$, $\Delta h = h/2$ 。

2.2 融合多方向变异性的克里金插值

在二维室内环境中,对不同方向下的变差函数拟合效果进行分析,发现不同方向下球状模型参数 C_0 、 C 和 a 的求解结果相差较大。因此可以得知:在二维室内空间下,仅用单一方向下的变差函数来进行克里金插值估计存在局限性,因为不能够全面表征信号整体的空间变异特征。针对上述问题,本文提出了一种融合多方向变异性的克里金插值估计方法。

以室内环境作为二维平面,建立空间直角坐标系,然后分别求取 x 方向上和 y 方向上的变差函数,最终可获得 x 方向和 y 方向所对应的变程 a_x 、 a_y 。由上文可知:变程代表数据相关性的范围,所以可以推断如果某方向下的变程越大,则该方向下的数据相关性影响范围就越广,对应的克里金插值结果就越具代表性。所以,为了融合 x 方向和 y 方向的插值估计,用 a_x 、 a_y 计算两方向所对应的估值权重:

$$\begin{aligned}\omega_x &= a_x / (a_x + a_y) \\ \omega_y &= a_y / (a_x + a_y)\end{aligned}\quad (16)$$

同时,用 a_x 和 a_y 确定待插值点的搜索邻域。本文取变程的 $1/8$ 作为搜索区间,即对待插值点 (x_i, x_j) ,只要空间中已知采样点的 x 坐标满足 $[x_i - a_x/8, x_i + a_x/8]$, y 坐标满足 $[x_j - a_y/8, x_j + a_y/8]$,就将该采样点作为已知观测点。然后利用搜索得到的已知观测点信息,根据式(4)和式(15)分别求得 x 、 y 方向针对待插值点的 RSSI 信号估计值。假设 x 方向的估计结果为 est_x , y 方向为 est_y ,则根据式(16)求得的 x 、 y 方向的权重系数融合两方向的克里金估计结果:

$$est = est_x \times \omega_x + est_y \times \omega_y \quad (17)$$

3 实验结果与分析

3.1 实验环境

为有效评估本文算法的插值精度提升效果,分别在真实室内环境下和模拟室内环境下展开了仿真实验。

环境 1 选取某工程实验室大楼二楼作为真实实验环境,该空间内随机部署了 5 个 AP,型号均为 TPLINK-842N,区域示意图如图 1 所示。以图 1 中 $7\text{ m} \times 14\text{ m}$ 的矩形区域作为定位区域,该区域内存在桌子、椅子、柱子等多种类型的障碍物。由于克里金

算法中的变差函数求解需要空间中存在一定量的已知数据点,因此设置采样点间距为 0.5 m ,均匀在空间内设置了 392 个采样点。然后通过电脑上安装的 RSSI 信号采集软件,在每个参考点处对各个 AP 以 1 s/Hz 的频率采集一组 RSSI 信号值,并经过卡尔曼滤波处理后构建得到一个含有 392 条指纹数据的指纹数据库。

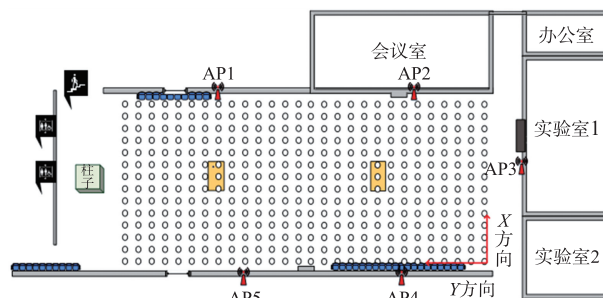


图 1 实验场景 1 示意图

环境 2 利用室内经典的对数路径损耗模型:

$$z_i^j = Z_t - L_0 - 10\beta \lg \|x_i - x^j\| + X_s \quad (18)$$

式中: z_i^j 表示第 i 个参考点处接收到的第 j 个 AP 的 RSSI 值; Z_t 表示发射信号功率; L_0 表示距离 AP 相对较短距离的路径损耗; β 为路径损耗因子,表示路径损耗随距离增加而增大的快慢,与周围的环境有关; x^j 表示第 j 个参考点位置, $\|x_i - x^j\|$ 表示第 i 个参考点与第 j 个 AP 之间的位置距离; X_s 为随机噪声。利用式(18)的室内模型,参照文献[10],产生一个含 5 个 AP 的 $150\text{ m} \times 150\text{ m}$ 的大型模拟室内环境。以 5 m 为间隔距离,将该区域划分为 30×30 的网格,并以每个网格的中心点位置作为采样点位置,生成一个包含有 900 条指纹数据的指纹数据库。其中,各个采样点处的指纹数据均由每个 AP 的 1 000 次 RSSI 测量值求平均得到。具体地,模拟产生的实验环境如图 2 所示。

3.2 插值效果分析

为验证本文算法的稳健性和准确性,选择不掺杂人为因素影响的伯努利随机过程来稀疏化处理环境 1 和环境 2 下所对应的指纹数据库,用于分析插值算法对稀疏化指纹数据库的插值恢复效果。其中用 p 表示剩余指纹数据占指纹数据库中指纹数据总数的百分比,相应地, $1-p$ 就表示删除数据所占的百分比。

为全面分析算法的插值效果,从整体和局部的角度分别对插值算法精度进行分析。用平均估计误差(单位: dBm)分析空间中各个 AP 所对应的平均插值误差;用指纹估计误差百分比(单位: %)分析每条指纹数据的估计误差占整体指纹数据的误差百

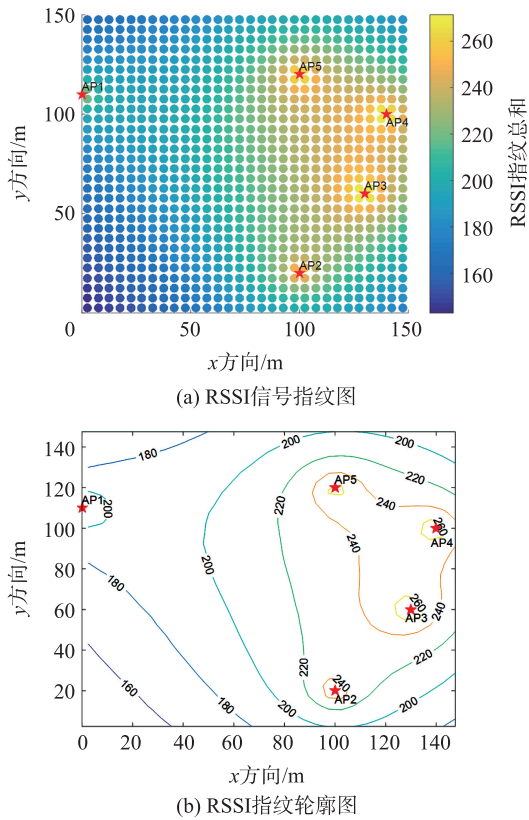


图2 实验场景2示意图

分比。其中,平均 AP 估计误差定义为:

$$\text{err_avg} = \frac{1}{n} \sum_{i=1}^n |AP_i^* - AP_i| \quad (19)$$

式中: n 表示定位环境中 AP 的数量; AP_i^* 表示环境中第 i 个 AP 的估计值;而 AP_i 则表示第 i 个 AP 在空间中的原始信号值。

指纹估计误差百分比则定义为:

$$\text{err_per} = \frac{1}{m} \frac{|R_i^* - R_i|}{R_i} \quad (20)$$

式中: m 表示待插值点的数量; R_i 为第 i 个待插值点处对应的原始指纹数据; R_i^* 则为第 i 个待插值点处的指纹数据估计值。

为了降低伯努利过程随机抽取样本指纹数据对估值效果的影响,将每组对比实验过程重复进行 30 次^[11],以充分反映算法的准确估计效果,具体仿真效果如图 3 和图 4 所示。

通过图 3 和图 4 的对比结果表明,在不同已知指纹百分比 p 值的情况下,本文算法的实际插值误差均小于克里金插值算法和距离反比加权插值算法,证明了本文算法对于插值精度提升的有效性。同时,对比图 3 环境 1 下的采样间距值等于 0.5 m 和图 4 环境 2 下的采样点间距等于 5 m 可知:由于距离反比加权插值算法仅考虑了空间中点与点之间

的位置距离关系,而忽略了点之间的空间变异性,所以在采样点间距增大的情况下,由图 3 中的(b)和图 4 中的(b)可明显看出,距离反比加权插值算法与克里金插值算法的估计效果差距拉大。

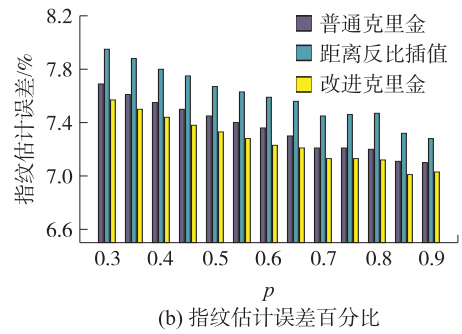
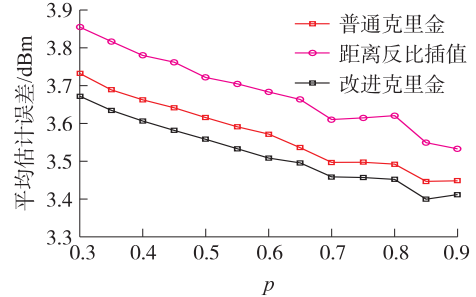


图3 实验环境1对应的插值效果展示

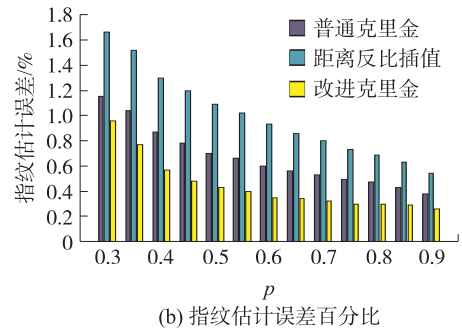
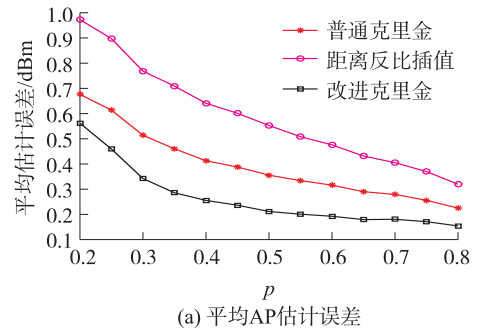


图4 实验环境2对应的插值效果展示

4 结论

为提升克里金插值算法的插值精度,本文提出了一种融合多方向变异性影响的改进克里金插值算法,同时利用概率不同的伯努利随机过程构造了多种稀

疏度的指纹数据库,并在不同的指纹数据库稀疏性水平上,将本文算法与普通克里金算法、距离反比加权算法的插值精度进行了分析对比。仿真结果表明,本文算法估计的 RSSI 指纹更准确,插值精度更高。

参考文献:

- [1] 席瑞,李玉军,侯孟书. 室内定位方法综述[J]. 计算机科学, 2015, 43(4): 1-6.
- [2] 李方敏,张韬,刘凯,等. 基于距离测量和位置指纹的室内定位方法研究[J]. 计算机学报, 2019, 42(2): 111-122.
- [3] 郭红成,罗海勇,尹浩,等. 基于线性插值和动态指纹补偿的分布式定位算法[J]. 传感技术学报, 2009, 22(12): 1795-1801.
- [4] 田增山,代海鹏. 基于曲面拟合的 WiFi 指纹数据库更新[J]. 计算机应用, 2016, 36(5): 1192-1195.
- [5] Li B H, Wang Y F, Lee H K, et al. Method for Yielding a Database of Location Fingerprint in WLAN[J]. IEEE Proceedings Communications, 2005, 152(5): 580-586.
- [6] Zhao H L, Huang B Q, Jia B. Applying Kriging Interpolation for Wifi Fingerprinting Based Indoor Positioning Systems[C]//Proceedings of the 2016 IEEE Wireless Communications and Networking Conference. Doha, Qatar: IEEE, 2016: 1-6.
- [7] Jan S, Yeh S, Liu Y. Received Signal Strength Database Interpolation by Kriging for a Wifi Indoor Positioning System[J]. Sensors, 2015, 15(9): 21377-21393.
- [8] Wang P, Feng Z, Tang Y, et al. A Fingerprint Database Reconstruction Method Based on Ordinary Kriging Algorithm for Indoor Localization[C]//2019 International Conference on Intelligent Transportation, Big Data & Smart City (ICITBS), Changsha, China, 2019: 224-227.
- [9] 李方,王铁成,佟为明. 基于空间变异理论的电子地图构建方法[J]. 哈尔滨工程大学学报, 2012, 33(6): 715-719.
- [10] Liu C, Kiring A, Salman N, et al. A Kriging Algorithm for Location Fingerprinting Based on Received Signal Strength[C]//2015 Sensor Data Fusion: Trends, Solutions, Applications (SDF). IEEE, 2015: 1-6.
- [11] 刘辉元,马金辉,黄琼. 基于改进克里金插值的室内定位位置指纹库构建方法[J]. 重庆邮电大学学报(自然科学版), 2017(6): 45-51.
- [12] 刘志建,关维国,华海亮,等. 基于克里金空间插值的位置指纹数据库建立算法[J]. 计算机应用研究, 2016, 33(10): 3139-3142.



刘晓宇(1994—),女,硕士生,主要研究方向为无线室内定位;



邓平(1964—),男,博士,教授,主要研究方向为无线网络定位技术,现代信号处理,无线传感器网络。