https://doi.org/10.1051/jnwpu/20224040865

基于信道测量实验的 NLOS 误差消除方法对比研究

畅田田[1], 王威[1], 高婧洁[1], 申晓红[2], 姜苏英[1], 谢景丽[1]

(1.长安大学 信息工程学院, 陕西 西安 710064; 2.西北工业大学 航海学院, 陕西 西安 710072)

摘　要: 为了研究不同消除方法对无线电信号由非视距(NLOS)传播而产生的距离估计正偏误差的消除性能,基于信道状态信息(CSI)提取出均值、均方根延迟扩展、偏度、峰度、峰均比特征,并将其与基于到达时间(TOA)的对数估计距离相结合作为特征输入向量,通过建立高斯过程回归(GPR)、最小二乘支持向量机回归(LS-SVMR)与 BP 神经网络训练模型进行实验性能比较。对实际测量的典型室内环境中 2.4~5.4 GHz 的无线传播信道进行误差消除实验,比较不同输入特征、不同带宽和不同频带下的 NLOS 误差消除性能。实验结果表明:GPR 模型表现出最好的 NLOS 误差消除性能,且所提取的 CSI 多特征作为输入向量可以将平均绝对误差(MAE)和均方根误差(RMSE)分别减小 71.12%和 81.36%;随着带宽不断增加,误差消除性能逐渐优化,即可通过增大带宽有效地改善输入特征较少时的 NLOS 定位误差;在多特征输入下,低频带的 NLOS 测距误差与高频带不同,因此将所有可用的频带结合可以比单频带更好地消除 NLOS 定位误差。

关　键　词: 非视距;信道状态信息;到达时间;最小二乘支持向量机回归;高斯过程回归;BP 神经网络

中图分类号: TN925　　**文献标志码:** A　　**文章编号:** 1000-2758(2022)04-0865-10

随着定位技术应用范围的不断扩大,新兴应用对定位的需求已不局限于单纯的室外场景。据调查数据显示,人们在室内的时间是室外的 4 倍,室内定位的需求变得日益迫切。在室外场景中,目前常用的定位系统是全球导航卫星系统(global navigation satellite system,GNSS),包括美国全球定位系统、俄罗斯格洛纳斯、中国北斗导航系统、欧盟伽利略系统。民用 GNSS 设备的定位精度在几米到几十米,使用差分技术的情况下定位精度可以进一步提高。但是卫星定位系统的功率在室内环境下会产生严重衰减,同时室内定位还受到多径效应的影响,导致 GNSS 定位系统在室内会产生较大的定位误差,甚至不能实现定位。另外,室内定位属于短距离小范围定位,定位精度要求较高,仅使用 GNSS 定位无法满足定位精度要求。

目前,常见的室内定位技术主要包括无线传感器网络、声音、红外线、蓝牙和超宽带[1]。其中,超宽带具有能耗低、时间分辨率高、抗多径等优点,被认为是实现高精度室内定位的重要技术之一。然而,在实际情况下,超宽带定位仍存在许多挑战。这些挑战包括信道衰落、低信噪比、多用户干扰、多径效应和非视距(non-line-of-sight,NLOS)传播[2]。其中,NLOS 传播对室内定位的影响最为重要。室内环境下,信号的视距(line-of-sight,LOS)传播经常被障碍物遮挡,即无线发射机和接收机之间的直接路径被阻塞,导致信号形成 NLOS 传播。在这种情况下,信号传输只能通过散射、反射或衍射等路径到达接收端,直接使用基于到达时间(time-of-arrival,

收稿日期: 2021-09-23　　**基金项目:** 国家自然科学基金(61871059,61901057)资助

作者简介: 畅田田(1996—),女,长安大学硕士研究生,主要从事无线电波传播和定位导航研究。

通信作者: 高婧洁(1988—),女,长安大学讲师,主要从事基于无线传感器网络的通信、定位及导航研究。
e-mail:gaojingj@ chd.edu.cn

TOA)[3] 和接收信号强度(received signal strength indication,RSSI)[4] 的测距算法会出现较大偏差,严重影响定位的性能。因此,NLOS 误差消除对提高室内定位精度具有重要的实际意义。

在 NLOS 传播条件下,NLOS 识别和误差消除是提高定位精度的有效方法。最简单的方法是进行 NLOS 识别后,放弃 NLOS 条件下的测量点,直接使用 LOS 条件下的测量点进行测距。但该方法会导致有用信息的丢失。因此,NLOS 识别后,需要进行 NLOS 误差消除,以提高定位精度。Xiao 等[5] 提出了一个融合了假设检验和对数正态传播的模型来估计 NLOS 条件下的收发距离。然而 NLOS 条件下该模型会产生非常大的测距误差。近年来,一些研究人员将机器学习算法引入到 NLOS 定位误差的消除,机器学习算法具有优越的性能。尽管现有许多机器学习算法,但文献主要关注支持向量机(support vector machine,SVM)[6-8] 或高斯过程回归(Gaussian process regression,GPR)[9]。这些方法主要研究了 NLOS 条件下墙壁[10]、人体[11] 等障碍物及不同输入特征[12] 对测距或定位的影响。上述文献没有对同一数据集上不同机器学习模型及神经网络模型的性能进行比较,也没有考虑不同带宽和不同频带对 NLOS 定位误差消除的影响。因此,本研究的主要目的是:①比较不同机器学习算法及神经网络模型的 NLOS 定位误差消除方面的性能;②研究不同带宽和不同频带对 NLOS 定位误差消除的影响。

本文基于信道状态信息(channel state information,CSI)提取出多种特征,建立最小二乘法支持向量机回归(least square-support vector machine regression,LS-SVMR)、GPR 与 BP 神经网络模型,并采用均方根误差(root mean square error,RMSE)、平均绝对误差(mean absolute error, MAE)和累积分布函数(cumulative distribution function,CDF) 3 个分析指标对不同带宽、不同频带下的误差消除性能进行了对比与评价。实验结果表明提取的 CSI 多特征作为输入可以显著提高误差消除性能且 GPR 模型的距离估计误差最小;随着带宽的增大,NLOS 传播引起的测距误差不断减小,因此可通过增大带宽来有效改善输入特征较少时的 NLOS 定位误差;在多特征输入下,低频带的 NLOS 测距精度小于高频带,因此将可用的频带结合可以比单频带更好地消除 NLOS 定位误差。

1 信道模型和多特征提取

1.1 信道模型

在无线通信中,CSI 提供了发射机和接收机之间通信链路的传播信道特征属性[13-14],具体包含每条传输路径的幅值、相位衰落,时延等特征信息。CSI 是信道频率响应(channel frequency response,CFR)的采样值,其描述了频域内每个子载波的幅值和相位

$$H(f_k) = \| H(f_k) \| \mathrm{e}^{\mathrm{i}\sin(\angle H(f_k))} \tag{1}$$

式中:f_k 是第 k 个子载波的频率;$H(f_k)$ 是频率为 f 的子载波的 CSI;$\| H(f_k) \|$ 和 $\angle H(f_k)$ 分别为其幅值和相位。

因此接收端的 CSI 信息可表示为 $\boldsymbol{H} = [H(f_1), H(f_2), \cdots, H(f_k), \cdots H(f_n)]^{\mathrm{T}}$。

1.2 多特征提取

通过快速傅里叶逆变换(inverse fast Fourier transform,IFFT)将 CSI 转换为时域信道冲激响应(channel impulse response,CIR),进而利用信号处理方法可提取出无线信道的诸多特征。

CIR 可表示为

$$h(\tau) = \sum_{i=1}^{N} a_i \mathrm{e}^{-\mathrm{j}\theta_i} \delta(\tau - \tau_i) \tag{2}$$

式中:N 代表多径分量的总数;a_i, θ_i, τ_i 分别表示第 i 条路径的幅值、相位和时延;$\delta(\tau)$ 是狄克拉脉冲函数。

在进行特征提取之前,需要对 CIR 中的噪声进行滤波处理,获得 CIR 中的有效信息。从滤波后的 CIR 中可以提取用来减小 NLOS 测距误差的 5 个关键特征:均值、均方根延迟扩展(root mean square delay spread,RMS-DS)、偏度(skewness)、峰度(kurtosis)和峰均比(peak-to-average ratio,PAR)。

1) 均值 μ:将 CIR 的幅值取样本平均值得到 μ。

2) 均方根延迟扩展:RMS-DS 是功率延迟分布的第二中心矩[15]。RMS-DS 的计算公式为:

$$\sigma_{\mathrm{R}} = \sqrt{\frac{\int_{-\infty}^{\infty} | h(\tau) |^2 \tau^2 \mathrm{d}\tau}{P} - m_{\mathrm{d}}^2} \tag{3}$$

式中:σ_{R} 表示均方根延迟扩展 RMS-DS;P 代表功率;m_{d} 代表平均延迟,即 $m_{\mathrm{d}} = \dfrac{\int_{-\infty}^{\infty} | h(\tau) |^2 \tau \mathrm{d}\tau}{P}$。

3) 偏度 S:定义为样本三阶标准化矩,表达式为

$$S = \frac{\mathrm{E}(|h(\tau)| - \mu_{|h|})^3}{\sigma_{|h|}^3} \tag{4}$$

式中:$|h(\tau)|$ 表示 CIR;$\mu_{|h|}$ 和 $\sigma_{|h|}$ 分别代表 $|h(\tau)|$ 的均值和标准差。

4) 峰度 K:定义为 CIR 幅值的四阶中心矩和二阶中心矩的比值[7]。K 可以表示为

$$K = \frac{\mathrm{E}(|h(\tau) - \mu_{|h|})^4}{\sigma_{|h|}^4} \tag{5}$$

5) 峰均比 PAR:定义为功率最大值与功率平均值的比,可以表示为

$$R_{\mathrm{PA}} = \frac{P_{\max}}{P_{\mathrm{mean}}} \tag{6}$$

式中:R_{PA} 表示峰均比 PAR;$P_{\max}$ 和 P_{mean} 分别表示功率的最大值和平均值。

为了将预测距离与真实距离进行比较,首先在地面标注出发射机和接收机的位置,然后利用 TOPCON@ 全站经纬仪测量收发机的真实坐标,从而获得收发机之间的真实距离。一般情况下,NLOS 传播将导致无法获得原始收发机之间的真实距离,故采用 TOA 算法进行估计。估计距离取对数主要是因为距离和接收功率之间存在线性关系,模型的预测效果好。

对数估计距离:利用 TOA 算法可估计出发射机和接收机间的距离,最后将基于 TOA 估计的发射机和接收机之间的距离转化为其对数值。

$$\hat{d} = \lg(d_{\mathrm{TOA}}) \tag{7}$$

式中:d_{TOA} 表示利用 TOA 算法估计得到的收发机间的距离。

将提取出的多种 CSI 特征与基于 TOA 的 NLOS 对数估计距离 $\hat{d}$ 相结合,可获得 NLOS 误差消除的不同输入向量组合。本文中所有的 NLOS 误差消除方法均是基于公式(8)中所列的特征展开

$$\boldsymbol{X} = [\mu, \sigma_{\mathrm{R}}, S, K, R_{\mathrm{PA}}, \hat{d}]^{\mathrm{T}} \tag{8}$$

2　NLOS 误差消除的方法

为了提高 NLOS 环境下的定位精度,在基于 TOA 的 NLOS 估计距离基础上,综合 CSI 中的多特征提取,分别引入 LS-SVMR、GPR、BP 神经网络对发射机和接收机间的距离进行估计,实现 NLOS 环境下的误差消除,提高定位精度。

2.1　最小二乘支持向量机回归(LS-SVMR)

SVM 是一种监督机器学习算法,可以用于回归问题。LS-SVM 是 SVM 的一种变体,它降低了 SVM 算法的复杂度[16]。在 NLOS 定位误差消除中,输出是发射机和接收机之间的距离 $d(d \in \mathbf{R})$,它依赖于输入的特征向量 $\boldsymbol{x}(\boldsymbol{x} \in \mathbf{R}^n)$。

$$d(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\phi(\boldsymbol{x}) + w_0 \tag{9}$$

式中:$\phi(\cdot)$ 是一个映射函数;$\boldsymbol{w}$ 和 w_0 是该回归器的参数,这些参数值是根据训练集确定的。

LS-SVM 回归通过求解下面的凸优化问题得到

$$\underset{\boldsymbol{w}, w_0, e_k}{\operatorname{argmin}} \frac{1}{2}\|\boldsymbol{w}\|^2 + \frac{1}{2}c\sum_{k=1}^{N} e_k^2$$
$$\text{s.t. } d_k = \boldsymbol{w}^{\mathrm{T}}\phi(\boldsymbol{x}_k) + w_o + e_k, \forall k \tag{10}$$

式中:c 是平衡因子,用来控制训练误差和模型复杂度[17];e_k 为松弛因子。(10) 式可以通过引入拉格朗日乘子和卡鲁什-库恩-塔克条件进行求解。最终,LS-SVM 回归的表达式为

$$d(\boldsymbol{x}) = \sum_{k=1}^{N} \alpha_k k(\boldsymbol{x}, \boldsymbol{x}_k) + w_0 \tag{11}$$

式中:α_k 是拉格朗日乘子;$k(\boldsymbol{x}, \boldsymbol{x}_k)$ 为核函数,本文中选择的核函数为径向基函数。

$$k(\boldsymbol{x}, \boldsymbol{x}_k) = \phi(\boldsymbol{x})^{\mathrm{T}} \cdot \phi(\boldsymbol{x}_k) = \exp\left[-\frac{\|\boldsymbol{x} - \boldsymbol{x}_k\|_2^2}{2\sigma^2}\right] \tag{12}$$

式中:$\boldsymbol{x}_k$ 为核函数中心;σ 为核函数的宽度参数。

2.2　高斯过程回归器(GPR)

高斯过程是指一组服从高斯分布的随机变量集合,它的性质完全由均值函数和协方差函数确定[18]。高斯过程回归模型是将高斯过程用到了回归问题,GPR 模型如下

$$\boldsymbol{d}(\boldsymbol{x}) = \boldsymbol{w}^{\mathrm{T}}\phi(\boldsymbol{x}) + \boldsymbol{n} \tag{13}$$

式中:$\phi(\cdot)$ 是一个映射函数;$\boldsymbol{w}$ 为模型的参数;$\boldsymbol{w} \sim N(0, \sum p)$[19]。$\boldsymbol{n}$ 为高斯噪声且 $\boldsymbol{n} \sim N(0, \sigma_n^2)$。输出向量 $\boldsymbol{d}$ 的先验分布为 $\boldsymbol{d} \sim N(0, \boldsymbol{K} + \sigma_n^2\boldsymbol{I})$,其中 $\boldsymbol{K} = k(\boldsymbol{x}, \boldsymbol{x}_k) = \phi(\boldsymbol{x})^{\mathrm{T}}\sum p\phi(\boldsymbol{x}_k)$ 为协方差函数,$\boldsymbol{K}$ 为单位矩阵。

根据回归模型以及高斯过程的定义,可以得到训练集的输出 $\boldsymbol{d}$ 和测试集的输出 $\boldsymbol{d}_*$ 的联合先验分布

$$\begin{bmatrix} \boldsymbol{d} \\ \boldsymbol{d}_* \end{bmatrix} \sim N\left(0, \begin{bmatrix} \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}_*) + \sigma_n^2\boldsymbol{I} & \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}_*) \\ \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}_*) & \boldsymbol{K}(\boldsymbol{x}, \boldsymbol{x}_*) \end{bmatrix}\right) \tag{14}$$

式中：$(\boldsymbol{x},\boldsymbol{d})$ 和 $(\boldsymbol{x}_*,\boldsymbol{d}_*)$ 分别表示训练集和测试集的输入和输出。对(14)式的联合先验分布取 $\boldsymbol{d}_*$ 的边缘分布可以得到 $\boldsymbol{d}_*$ 最大概率的预测后验分布

$$p(\boldsymbol{d}_* \mid \boldsymbol{x},\boldsymbol{y},\boldsymbol{x}_*,\sigma_n^2) = N[\boldsymbol{d}_* \mid \overline{\boldsymbol{d}_*},\mathrm{cov}(\boldsymbol{d}_*)] \tag{15}$$

式中，估计距离 $\boldsymbol{d}_*$ 的均值和协方差分别为：

$$\overline{\boldsymbol{d}_*} = k(\boldsymbol{x}_*,\boldsymbol{x})(\boldsymbol{K}+\sigma_n^2\boldsymbol{I})^{-1}\boldsymbol{d} \tag{16}$$

$$\mathrm{cov}(\boldsymbol{d}_*) = k(\boldsymbol{x}_*,\boldsymbol{x}_*) - k(\boldsymbol{x}_*,\boldsymbol{x})[k(\boldsymbol{x},\boldsymbol{x}) + (\boldsymbol{K}+\sigma_n^2\boldsymbol{I})]^{-1}k(\boldsymbol{x},\boldsymbol{x}_*) \tag{17}$$

GPR 核函数选择的是平方指数协方差函数

$$k(\boldsymbol{x},\boldsymbol{x}_*) = \phi(\boldsymbol{x})^{\mathrm{T}}\cdot\phi(\boldsymbol{x}_*) = \sigma_f^2\exp\left[-\frac{\|\boldsymbol{x}-\boldsymbol{x}_*\|_2^2}{2\boldsymbol{l}^2}\right] \tag{18}$$

式中：$\boldsymbol{l}$ 和 σ_f^2 均为核函数的超参数。

2.3　反向传播(BP)神经网络

BP 神经网络是一种按照误差反向传播算法训练的多层前馈神经网络[20]。该算法的基本思想是利用梯度下降法，使真实值与网络预测值误差的均方差达到最小[21]。BP 神经网络由输入层、隐藏层和输出层组成，BP 神经网络的拓扑结构如图 1 所示。

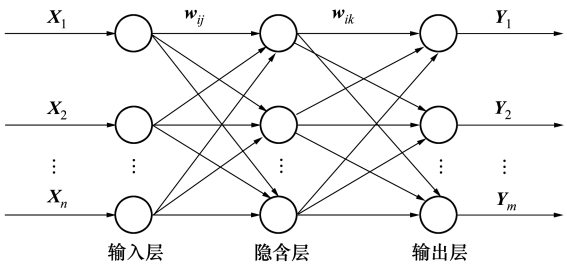


图 1　BP 神经网络的拓扑结构

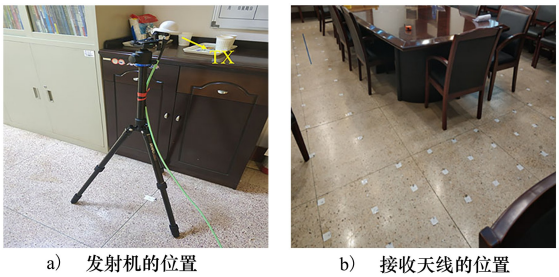
图 1 中，$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_n$ 和 $\boldsymbol{Y}_1,\boldsymbol{Y}_2,\cdots,\boldsymbol{Y}_m$ 分别代表网络的输入值和预测输出值，$\boldsymbol{w}_{ij}$ 表示输入层和隐含层的权值矩阵，$\boldsymbol{w}_{jk}$ 表示隐含层和输出层的权值矩阵。

3　实验数据测量与和结果分析

在本节中，利用 CSI 中提取的多特征向量，结合基于 TOA 的 NLOS 距离估计，分别使用 LS-SVMR、BP 神经网络和 GPR 模型实现 NLOS 定位误差的消除，完成真实环境下的实验测量。根据实验结果对 LS-SVMR、BP 神经网络和 GPR 不同带宽、不同频带和输入特征的 NLOS 误差消除性能进行比较和分析。

3.1　实验环境和测量硬件

为了获得 LOS 和 NLOS 场景下的信道传播特征，在一个 8 m×4 m 的会议室开展了 2.4~5.4 GHz 频段的信道测量活动，环境参数配置如表 1 所示。无线信道测量活动采用了 P5008A 型矢量网络分析仪来测量多组收发机组合的信道传递函数。会议室内设置了 88 个接收天线组成的网络，相邻点间距 0.5 m。此外，在测量活动中使用了 5 个发射天线，其中 3 个位于会议室内，属于 LOS 环境，剩余 2 个位于会议室外的走廊，属于 NLOS 环境。在测量过程中，发射和接收天线的高度均保持在 1.15 m，发射和接收天线的位置由 TOPCON® 全站经纬仪测量，可以根据位置测量得到发射机与接收机之间距离的真实值。图 2 展示了室内信道测量的环境。



a)　发射机的位置　　b)　接收天线的位置

图 2　信道测量环境

表 1　信道测量参数

参数	值
频率/GHz	2.4~5.4
频率采样点	3 201
天线高度/m	1.15
相邻接收点的距离/m	0.5
测量类型	SISO

3.2　数据集的收集

在测量过程中，收集了 405 个样本，每个样本由 3 201个 CSI 频率采样点组成，其中，LOS 传播条件下共获得 264 个样本，剩余的 141 个样本为 NLOS 传播条件。NLOS 条件下的样本被划分为训练集和测试集，其占总样本的比例分别为 70%和 30%。

3.3　实验结果

3.3.1　TOA 测距误差

图 3 给出了 LOS 和 NLOS 条件下基于 TOA 距

离估计误差的 CDF 图。由图可知,LOS 条件下,95%测量结果的测距误差在 0.2 m 范围内;NLOS 条件下,仅 20%测量点的测距误差在 0.2 m 范围内。因此,在 LOS 条件下,TOA 测距可以获得精确的位置信息,但在 NLOS 条件下,TOA 算法会产生较大的测距误差,降低无线定位的精度。NLOS 条件下的测距误差远远大于 LOS 条件是因为发射机和接收机之间的直视路径被障碍物遮挡,导致无线电信号从发射机到达接收机的时延增大。因此,需要消除 NLOS 条件下的测距误差来提高测距精度。

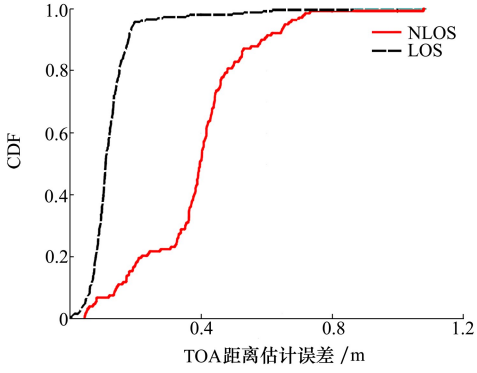


图 3　LOS 和 NLOS 测距误差的 CDF

3.3.2　不同模型下的测距误差

为了对不同模型下的 NLOS 定位误差消除性能进行比较,选取了 RMSE、MAE 和 CDF 3 个评估指标。

1) RMSE:发射机与接收机间距离的真实值和估计值差值的平方与测量次数 n 比值的平方根[22]。RMSE 的计算公式为

$$E_{\mathrm{RMS}} = \sqrt{\frac{1}{n}\sum_{i=1}^{n} d_i^2} \tag{19}$$

式中:d_i 表示第 i 个收发机真实距离与估计距离的差值。

2) MAE:各次测距误差绝对值取平均[23],MAE 可以表示为

$$E_{\mathrm{MA}} = \frac{1}{n}\sum_{i=1}^{n} |\, d_i \,| \tag{20}$$

3) CDF:定位误差在某个精度门限以下的概率[24]。

$$F(x) = P(d_i < x) \tag{21}$$

实验分别针对 LS-SVMR、BP 神经网络、GPR 模型以及基于对数距离路径损耗的 NLOS 距离估计和基于 TOA 的 NLOS 距离估计进行 NLOS 定位误差消除性能比较。

各模型的绝对距离估计误差如图 4 所示。从图中可看出,对数路径损耗模型的距离估计误差是最大的。LS-SVMR 和 BP 神经网络模型的距离估计误差小于 TOA。GPR 模型的距离估计误差最小,因此该实验环境下,不同模型的距离估计误差精度不同,且基于 GPR 模型的误差消除效果最好。

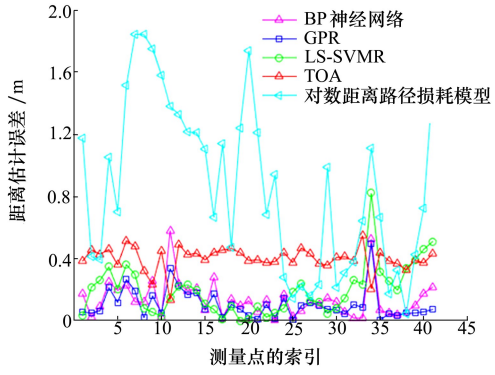


图 4　不同索引点的测距误差

表 2 给出了不同模型的 MAE 和 RMSE 值,进一步对不同模型的距离估计误差进行验证。其中 GPR 模型可将距离估计的 MAE 减小到 0.110 8 m,比 TOA 的距离估计的 MAE 减小 0.146 6 m,误差消除性能高达 71.12%;分别比 LS-SVMR 和 BP 神经网络的 MAE 小 0.087 5 和 0.043 8 m。同样,GPR 的 RMSE 也均小于对数距离路径损耗模型、LS-SVMR 和 BP 神经网络,与 TOA 相比,误差消除性能提高了 81.36%。因此,GPR 模型比 LS-SVMR 和 BP 神经网络有更小的 MAE 和 RMSE 值,即 GPR 模型定位误差消除性能优于 LS-SVMR 和 BP 神经网络模型。

表 2　不同模型的测距误差

模型	MAE/m	RMSE/m
对数距离路径损耗	1.646 0	1.960 4
TOA	0.383 6	0.775 8
LS-SVMR	0.198 3	0.255 3
BP 神经网络	0.154 6	0.202 3
GPR	0.110 8	0.146 6

不同模型距离估计误差的 CDF 如图 5 所示。从图中可以发现,在 NLOS 条件下,基于 GPR 模型 98%的索引点测距误差小于 0.5 m。由于 BP 神经网络精度受训练数据量的影响较大,LS-SVMR 模型

由惩罚参数和核函数参数的选取决定,因此在相同实验环境下,LS-SVMR 和 BP 神经网络模型预测中95%的索引点距离估计误差在 0.5 m 范围内,测距误差偏大。TOA 模型与对数路径损耗模型下的估计误差最大,其中 TOA 模型下仅有 80%的索引点距离估计误差在 0.5 m 范围内。因此,由实验结果可得,对于相同的测试集,GPR 模型的误差消除性能优于神经网络,而且具有容易实现、超参数自适应获取和输出具有概率意义等优点。

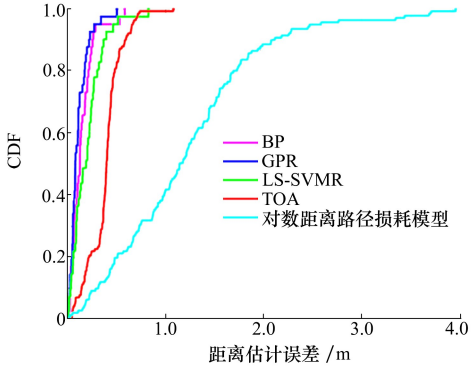


图 5　不同模型距离估计误差的 CDF

3.3.3　不同输入特征的测距误差

表 3 针对不同的输入特征集,结合 GPR 模型获得 NLOS 误差消除,并对误差消除的 MAE 与 RMSE 值进行比较。通过比较可得,本文提出的多特征集 $\boldsymbol{C}_7$ 提供了最佳的性能,即使用更多的输入特征进行 NLOS 定位误差消除是有益的。通过比较特征集 $\boldsymbol{C}_6$ 和 $\boldsymbol{C}_7$ 可以看出特征集中对数估计距离特征 $\hat{d}$ 的存在可以明显提高 GPR 回归模型的精度。

表 3　不同特征向量的测距误差

特征向量	MAE/m	RMSE/m
$\boldsymbol{C}_1 = [\mu, \sigma_R, \hat{d}]$	0.129 9	0.181 6
$\boldsymbol{C}_2 = [\mu, \sigma_R, S]$	0.821 4	1.051 6
$\boldsymbol{C}_3 = [\mu, \sigma_R, S, \hat{d}]$	0.114 1	0.149 5
$\boldsymbol{C}_4 = [\mu, \sigma_R, S, K]$	0.820 3	1.050 9
$\boldsymbol{C}_5 = [\mu, \sigma_R, S, K, \hat{d}]$	0.111 0	0.147 8
$\boldsymbol{C}_6 = [\mu, \sigma_R, S, K, R_{PA}]$	0.802 1	1.032 3
$\boldsymbol{C}_7 = [\mu, \sigma_R, S, K, R_{PA}, \hat{d}]$	0.110 8	0.146 6

3.3.4　不同带宽下的测距误差

为了评估不同带宽对 NLOS 误差消除的影响,利用不同模型对不同带宽下发射机与接收机间的距离进行估计。实验共考虑五种不同的带宽(包括 20 MHz,100 MHz,500 MHz,1 GHz 和 3 GHz),由于对数路径损耗模型下的估计误差最大,实验考虑 4 种不同模型(包括 TOA,LS-SVMR,BP 神经网络和 GPR 模型)。不同带宽下,TOA,LS-SVMR,GPR 和 BP 神模型的 MAE 和 RMSE 值如表 4~7 所示。通过对比可看出 5 种不同的带宽下,TOA 模型的测距误差最大,而 LS-SVMR,BP 神经网络和 GPR 模型都可以用于消除 TOA 模型在 NLOS 条件下的测距误差;随着带宽增大,TOA,LS-SVMR, GPR,BP 神经网络模型的 MAE 和 RMSE 值减小,3 GHz 带宽下,4 种模型的 MAE 和 RMSE 均达到最小值。此外,还可发现不同带宽下 GPR 模型的测距误差最小,误差消除性能最好。

表 4　TOA 模型中不同带宽的测距误差

带宽	MAE/m	RMSE/m
20 MHz	5.557 2	6.792 7
100 MHz	5.060 9	5.820 1
500 MHz	1.200 0	1.390 7
1 GHz	0.674 4	0.875 6
3 GHz	0.383 6	0.775 8

表 5　LS-SVMR 模型中不同带宽的测距误差

带宽	MAE/m	RMSE/m
20 MHz	1.150 2	1.364 5
100 MHz	1.074 6	1.287 2
500 MHz	0.505 9	0.653 6
1 GHz	0.281 4	0.357 0
3 GHz	0.198 3	0.255 3

表 6　GPR 模型中不同带宽的测距误差

带宽	MAE/m	RMSE/m
20 MHz	0.796 0	1.028 5
100 MHz	0.717 2	0.900 7
500 MHz	0.461 3	0.554 0
1 GHz	0.175 2	0.233 1
3 GHz	0.110 8	0.146 6

表 7　BP 模型中不同带宽的测距误差

带宽	MAE/m	RMSE/m
20 MHz	1.006 5	1.176 8
100 MHz	0.843 9	1.032 7
500 MHz	0.455 2	0.621 4
1 GHz	0.269 7	0.266 6
3 GHz	0.154 6	0.202 3

不同带宽测距误差的 CDF 结果如图 6~9 所示。从图可知,随着带宽增加,TOA,LS-SVMR,BP 神经网络和 GPR 模型的 NLOS 测距误差均不断减小。以图 8 中 GPR 模型不同带宽距离估计误差的 CDF 图为例,当带宽从 100 MHz 增加至 500 MHz 时,测距误差显著减小,而当带宽从 1 GHz 增加至 3 GHz 时,NLOS 条件下的测距误差减小不显著。因此,带宽越大,测距误差越小,然而,当带宽足够大时,继续增大带宽不会显著提高测距的精度。

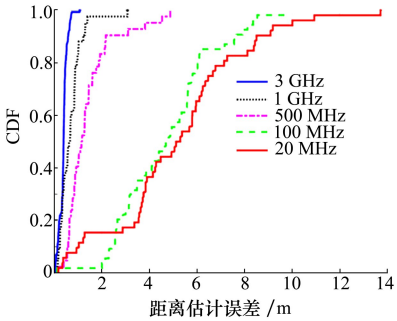


图 6　TOA 模型中不同带宽距离估计误差的 CDF

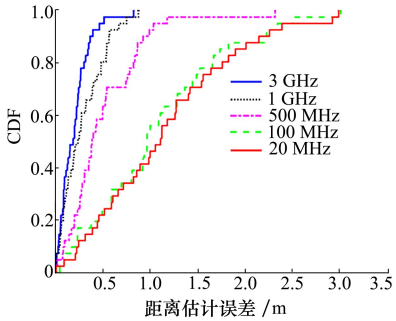


图 7　LS-SVMR 模型中不同带宽距离估计误差的 CDF

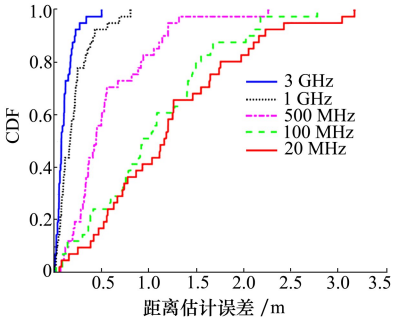


图 8　GPR 模型中不同带宽距离估计误差的 CDF

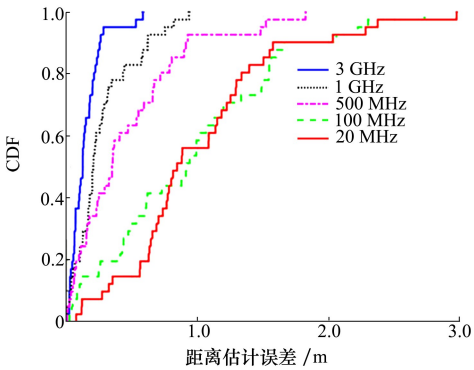


图 9　BP 模型中不同带宽距离估计误差的 CDF

图 10 显示了 GPR 模型在不同带宽和不同输入特征集下 NLOS 距离估计误差的 CDF。图中红色实线表示在带宽为 3 GHz,输入向量为特征集 $\boldsymbol{C}_2$(3 个特征)时的 NLOS 误差消除结果,蓝色虚线表示在带宽为 500 MHz,输入向量为特征集 $\boldsymbol{C}_4$(4 个特征)时的 NLOS 测距误差,绿色虚线表示在带宽为 100 MHz,输入向量为特征集 $\boldsymbol{C}_6$(5 个特征)时的 NLOS 误差消除结果。从图中可以发现大带宽、输入特征少时可以获得比小带宽、多输入特征更好的 NLOS 误差消除结果,即可通过增大带宽来弥补输入特征较少对 NLOS 误差消除产生的不利影响。

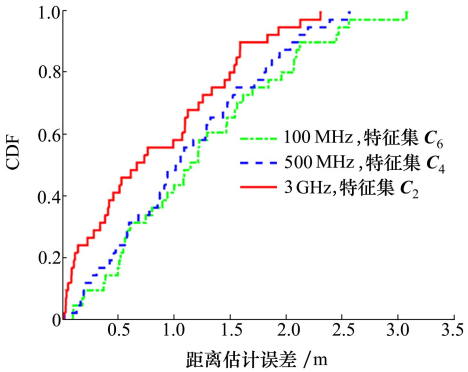


图 10　GPR 模型中不同带宽和不同特征集距离估计误差的 CDF

3.3.5　不同频带下的测距误差

为了比较不同频带的性能,将 3 GHz 带宽划分为 6 个不同频带,即每个频带的带宽为 500 MHz。图 11 给出了详细的频带划分图,从图中可看出,取出 3 个频带,各频带的 MAE 和 RMSE 值如表 8 所示,不同频带测距误差的 CDF 如图 12 所示。首先,由表 8 和图 11 可以看出使用不同频带的测距误差不同。其次,使用频带 6 的测距误差最大,频带 1 的

测距误差最小。因为频带 6 的频率高意味着测距信号的直径分量被高概率阻塞。因此,多特征下低频带可以有效减小 GPR 模型在 NLOS 条件下的定位误差。通过对表 6 和表 8 的比较可以得出另一个重要结论,所有可用的频带结合可以比单频带更好地消除 NLOS 定位误差。

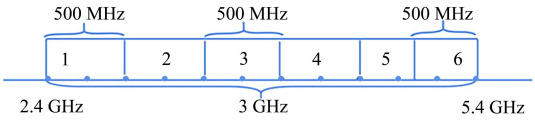


图 11　不同频带的划分

表 8　GPR 模型中不同频带的测距误差

带宽	MAE/m	RMSE/m
频带 1	0.555 2	0.767 6
频带 3	0.727 2	1.003 0
频带 6	1.137 0	1.352 8

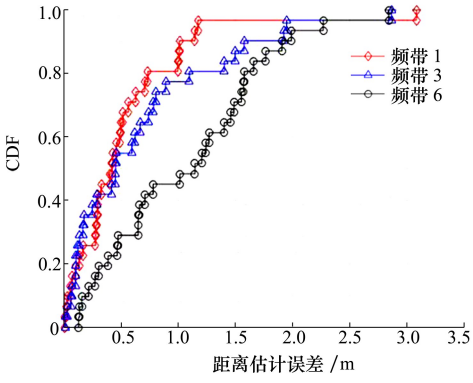


图 12　不同频带距离估计误差的 CDF

图 13 给出了 GPR 模型不同输入特征集和不同频带下 NLOS 误差消除的结果。从图中可看出,频带 1(低频带)在输入向量为特征集 $\boldsymbol{C}_2$ 时的测距误差最小,特征集 $\boldsymbol{C}_2$ 仅包含前 3 个特征。尽管频带 6(高频带)的输入特征包含了 5 个特征,但测距误差是最大的。因此,在输入特征少时,可以通过使用低频带来补偿输入特征较少带来的 NLOS 测距误差。另一方面,较高的输入特征可能会导致不同特征间具有相关性,而且还会增加计算的复杂度,延长程序的运行时间。

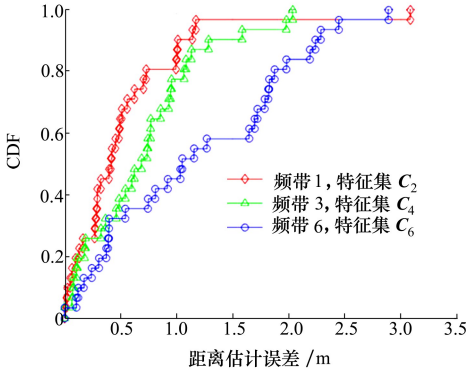


图 13　不同频带和不同特征集距离估计误差的 CDF

4　结 论

本文以实际测量的典型室内办公环境中 2.4~5.4 GHz 频段的信道 CSI 为原始数据,建立了 LS-SVMR、GPR 和 BP 神经网络 3 种不同模型,并利用 CSI 提取出均值、RMS-DS、偏度、峰度、PAR 结合基于 TOA 估计距离等特征进行性能评估实验。实验评估了不同特征、不同带宽和不同频带下的 NLOS 定位误差消除结果,并进行了性能比较。

从实验结果可知:GPR 模型表现出最好的 NLOS 定位误差消除性能且将 CSI 特征与对数估计距离结合形成的特征输入可以显著消除 NLOS 条件下的定位误差;随着带宽的增加,NLOS 误差消除性能逐渐优化,即可通过增大带宽有效地改善输入特征较少时的 NLOS 定位误差;多特征作为输入时,低频带可以有效减小 GPR 模型在 NLOS 条件下的定位误差,因此在输入特征少时,可使用低频带来减小输入特征较少带来的 NLOS 测距误差,且将所有可用的频带结合可以比单频带更好地消除 NLOS 定位误差。

参考文献:

[1]　FERREIRA A G, FERNANDES D , BRANCO S, et al. Feature selection for real-time nlos identification and mitigation for body-mounted uwb transceivers[J]. IEEE Trans on Instrumentation and Measurement, 2021, 70(1): 1-10

[2]　YU K, WEN K, LI Y, et al. A novel NLOS mitigation algorithm for UWB localization in harsh indoor environments[J]. IEEE Trans on Vehicular Technology, 2019, 68(1): 686-699

[3] WU S, ZHANG S, HUANG D. A TOA-based localization algorithm with simultaneous nlos mitigation and synchronization error elimination[J]. IEEE Sensors Letters, 2019, 3(3): 1-4

[4] KATWE M, GHARE P, SHARMA P K, et al. NLOS error mitigation in hybrid RSS-TOA-based localization through semi-definite relaxation[J]. IEEE Communications Letters, 2020, 24(12): 2761-2765

[5] XIAO Z, WEN H, MARKHAM A, et al. Non-line-of-sight Identification and mitigation using received signal strength[J]. IEEE Trans on Wireless Communications, 2015, 14(3): 1689-1702

[6] WU C, HOU H, WANG W, et al. TDOA based indoor positioning with nlos identification by machine learning[C]//2018 10th International Conference on Wireless Communications and Signal Processing, 2018: 1-6

[7] CHITAMBIRA B, ARMOUR S, WALES S, et al. Direct localisation using ray-tracing and least-squares support vector machines [C]//8th International Conference on Localization and GNSS, 2018: 1-5

[8] LI S, SONG B, LUO K. NLOS mitigation for UWB localization based on machine learning fusion method[C]//2019 IEEE Symposium Series on Computational Intelligence, 2019: 1048-1055

[9] YANG X F. NLOS mitigation for UWB localization based on sparse pseudo-input gaussian process[J]. IEEE Sensors Journal, 2018, 18(10): 4311-4316

[10] SILVA B, HANCKE G P. Ranging error mitigation for through-the-wall non-line-of-sight conditions[J]. IEEE Trans on Industrial Informatics, 2020, 16(11): 6903-6911

[11] TIAN Q, WANG K I, SALCIC Z. Human body shadowing effect on UWB-Based ranging system for pedestrian tracking[J]. IEEE Trans on Instrumentation and Measurement, 2019, 68(10): 4028-4037

[12] CHITAMBIRA B, ARMOUR S, WALES S,et al. NLOS identification and mitigation for geolocation using least-squares support vector machines[C]//2017 IEEE Wireless Communications and Networking Conference, San Francisco, 2017: 1-6

[13] 田春元, 余江, 常俊, 等. NWI:基于 CSI 的非视距信号识别方法[J]. 计算机科学, 2020, 47(11): 327-332
TIAN Chunyuan, YU Jiang, CHANG Jun, et al. NWI: CSI-based non-line-of-sight signal identification method[J]. Computer Science, 2020, 47(11): 327-332 (in Chinese)

[14] CHOI J S, LEE W H, LEE J H, et al. Deep learning based NLOS identification with commodity WLAN devices[J]. IEEE Trans on Vehicular Technology, 2018, 67(4): 3295-3303

[15] ZENG H, XIE R, XU R, et al. A novel approach to NLOS identification for UWB positioning based on kernel learning [C]// 2019 IEEE 19th International Conference on Communication Technology, Xi'an, 2019: 451-455

[16] NAM S C, CHOI H B, KO Y B. On mitigation of ranging errors for through-the-body NLOS conditions using convolutional neural networks[C]//2021 23rd International Conference on Advanced Communication Technology, 2021: 141-144

[17] DONG M Y. A low-cost NLOS identification and mitigation method for UWB ranging in static and dynamic environments [J]. IEEE Communications Letters, 2021, 1(1): 2420-2424

[18] CHITAMBIRA B, ARMOUR S, WALES S, et al. NLOS identification and mitigation for geolocation using least-squares support vector machines[C]//IEEE Wireless Communications and Networking Conference, 2017: 1-6

[19] MUSA A, NUGRAHA G D, HAN H, et al. A decision tree-based NLOS detection method for the UWB indoor location tracking accuracy improvement[J]. International Journal of Communication Systems, 2019, 32(13): 1-13

[20] CAO Y, ZHANG L H, Data fusion of heterogeneous network based on BP neural network and improved SEP[C]//2017 9th International Conference on Advanced Infocomm Technology, Chengdu, 2017: 138-142

[21] BREGAR K, MOHORCIC M, Improving indoor localization using convolutional neural networks on computationally restricted devices[J]. IEEE Access, 2018, 6(1): 17429-17441

[22] SILVA B, HANCKE G P. Ranging error mitigation for through-the-wall non-line-of-sight conditions [J]. IEEE Trans on Industrial Informatics, 2020, 16(11): 6903-6911

[23] SRIDHAR B, ALI KHAN M Z. RMSE comparison of path loss models for UHF/VHF bands in India [C]//2014 IEEE Region 10 Symposium, Kuala Lumpur, 2014: 330-335

[24] WANG T, HU K K, LI Z H, et al. A semi-supervised learning approach for UWB ranging error mitigation[J]. IEEE Wireless Communications Letters, 2021, 10(3): 688-691

A comparative study on NLOS error elimination methods based on channel measurement experiment

CHANG Tiantian[1], WANG Wei[1], GAO Jingjie[1], SHEN Xiaohong[2], JIANG Suying[1], XIE Jingli[1]

(1.School of Information Engineering, Chang'an University, Xi'an 710064, China;
2 School of Marine Science and Technology, Northwestern Polytechnical University, Xi'an 710072, China)

Abstract: In order to study the performance of different elimination methods on the distance estimation forward error caused by the non-line-of-sight (NLOS) propagation of radio signals, this paper is based on the mean value, root mean square delay spread, skewness, kurtosis and peak-to-average ratio extracted from the channel state information (CSI), and combine it with the logarithmic estimated distance based on the time of arrival (TOA) as the feature input vector, through the establishment of Gaussian process regression (GPR), least square support vector machine regression (LS-SVMR) and BP neural network training model for experimental performance comparison. Through the actual measurement of the 2.4 to 5.4 GHz wireless propagation channel in the typical indoor environment, the error elimination experiment is carried out to compare the NLOS error elimination performance under different input characteristics, different bandwidths and different frequency bands. The experimental results show that the GPR model has the best NLOS error elimination performance, and the extracted CSI multi-features as the input of the GPR model can reduce the average absolute error and root mean square error by 71.12% and 81.36%, respectively. As the bandwidth continues to increase, the error elimination performance is gradually optimized. By increasing the bandwidth, the NLOS positioning error when the input features are less can be effectively improved. The positioning error of the low frequency band is smaller than that of the high frequency band under the multi-features, so the combination of all available frequency bands can eliminate the NLOS positioning error better than a single frequency band.

Keywords: non-line-of-sight; channel state features; time of arrival; least squares-support vector machine regression; gaussian process regression; BP neural network

引用格式: *畅田田, 王威, 高婧洁, 等. 基于信道测量实验的 NLOS 误差消除方法对比研究[J]. 西北工业大学学报, 2022, 40(4): 865-874*

CHANG Tiantian, WANG Wei, GAO Jingjie, et al. A comparative study on NLOS error elimination methods based on channel measurement experiment[J]. Journal of Northwestern Polytechnical University, 2022, 40(4): 865-874 (in Chinese)

© 2022 Journal of Northwestern Polytechnical University.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.