

基于 Wi-Fi 感知的人员计数方法研究*

周泽仑¹, 戴欢^{1,2}, 黄河², 史文华¹



¹(苏州科技大学 电子与信息工程学院, 江苏 苏州 215000)

²(苏州大学 计算机科学与工程学院, 江苏 苏州 215000)

通讯作者: 戴欢, E-mail: daihuanjob@163.com

摘要: 人员计数是对指定区域内人口数量进行统计或准确估计的一种方法,在许多应用中都发挥了重要作用,例如公共安全、人群控制和营销分析等.传统的基于视频流、电子标签的人员计数方法硬件成本过高,并且基于视频流的人员计数方法在光线不足或有遮挡物的情况下精度低、可靠性差.提出一种基于 Wi-Fi 感知的人员计数方法,该方法对 Wi-Fi 中信道状态信息(channel state information,简称 CSI)进行重构,多子载波的 CSI 有效减少了多径效应的影响,利用解卷相位与线性变换的方法重构 CSI,使得相位信息能以簇的形式集中,避免了原始相位分布范围过大、随机性过高的问题,基于 Hampel 滤波器去除了载波振幅的奇异数据,减少了环境噪声因素对于人员数量特征造成的干扰,保证了利用无线信号进行人员计数的精度和稳定性,最后利用 SVM 分类进行人员计数.实验结果表明,该方法的计数准确度达到了约 95.8%,能够在室内环境下准确地识别出人员的数量.

关键词: CSI;人数统计;相位重构;滤波去噪;支持向量机

中文引用格式: 周泽仑,戴欢,黄河,史文华.基于 Wi-Fi 感知的人员计数方法研究.软件学报,2019,30(Suppl.(1)):62-70. <http://www.jos.org.cn/1000-9825/19007.htm>

英文引用格式: Zhou ZL, Dai H, Huang H, Shi WH. Wi-Fi perception based research on personnel counting method. Ruan Jian Xue Bao/Journal of Software, 2019,30(Suppl.(1)):62-70 (in Chinese). <http://www.jos.org.cn/1000-9825/19007.htm>

Wi-Fi Perception Based Research on Personnel Counting Method

ZHOU Ze-Lun¹, DAI Huan^{1,2}, HUANG He², SHI Wen-Hua¹

¹(School of Electronic and Information Engineering, Suzhou University of Science and Technology, Suzhou 215000, China)

²(School of Computer Science and Engineering, Soochow University, Suzhou 215000, China)

Abstract: Personnel counting are a method of counting or accurately estimating the population in a given area. It plays an important role in many applications, such as public safety, crowd control and marketing analysis. The traditional personnel counting method based on video stream and electronic tag has high hardware cost, and the accuracy and reliability of the personnel counting method based on video stream are low under the condition of insufficient light or occlusion. This paper presents a method of personnel counting based on Wi-Fi perception. This method reconstructs channel state information (CSI) in Wi-Fi. CSI of multi-subcarriers effectively reduces the influence of multipath effect. CSI is reconstructed by deconvolution phase and linear transformation, so that phase information can be centralized in the form of clusters. It avoids the problem of too large range of original phase distribution and too high randomness. Based on Hampel filter, the singular data of carrier amplitude is removed, the interference of environmental noise factors on the number of personnel characteristics is reduced, and the accuracy and stability of personnel counting using wireless signals are guaranteed. Finally, the numbers

* 基金项目: 国家自然科学基金(61702354, 61876121); 苏州科技大学科研项目(XKZ2017004); 江苏省物联网移动互联网技术工程重点实验室开放课题(JSWLW2017004); 研究生科研创新计划(SKSJ18_012, SJCX19_0963)

Foundation item: National Natural Science Foundation of China (61702354, 61876121); Scientific Research Project of Suzhou University of Science and Technology (XKZ2017004); Key Laboratory of Mobile Interconnection Technology Engineering of Jiangsu Province (JSWLW2017004); Graduate Research Innovation Project (SKSJ18_012, SJCX19_0963)

收稿时间: 2019-09-15; 采用时间: 2019-10-24

of people are classified by SVM. The experimental results show that the counting accuracy of the proposed method is about 95.8%, which can accurately identify the number of people in indoor environment.

Key words: channel state information; population statistics; phase reconstruction; filtering and denoising; support vector machine

人员计数在许多应用中都发挥了重要作用,例如一些公共场所需要控制人流数量来保证公共安全,交通管理部门需要利用人流量信息来进行交通建设规划,商城可以利用人流数量分布来设定服务配置.目前人员计数最常用的方法是利用视频流方式实时获取被检测区域的人流图,再利用运动物体检测的方式计数.文献[1]提出基于卡尔曼信道生成人的运动轨迹,然后通过卡尔曼滤波得到的轨迹在进出口方向确定室内人员的数量.文献[2]提出一种基于 SOC 芯片的人员计数系统,该系统利用梯度直方图、支持向量机能够快速通过图像计算出人员的数量.但是基于图像识别的人员计数系统无法在视野不广泛或者光线不充足的情况下起到良好的效果,甚至在更多的情况下,由于遮挡物较多,常常只能得到错误的结果,并且安装额外的视频获取设备需要较高的设备成本.

因此,一些基于射频技术的人员计数方式被提了出来.文献[3]使用脉冲无线电超带宽雷达传感器接收到的信号幅度信息,提出了一种人数统计的算法,该算法利用两台雷达之间的交互信息显著提升了人员计数的精确度.文献[4]提出了两种利用空间和波形特征预测穿过现有 RFID 装置的人数的算法,该算法可以合理、准确地估计通过指定路径的人数.文献[5,6]利用 RFID 标签的方法,为室内每个人配备电子标签,当人员在室内时通过检测电子标签数量的方式来统计人员数量.利用电子标签方式的优点是识别率十分准确,但是为室内每个人配备电子标签需要极高的成本,并且目前电子标签的标准也无法做到完全统一.

随着移动网络和 Wi-Fi 的普及,基于 Wi-Fi 进行人员计数引起了学术界和产业界广泛的兴趣.文献[7]提出采用 RSSI 波动序列的傅里叶谱特征作为机器学习算法的预测因子,在不需要用户携带任何特殊装置的情况下,推导出流量大小,并且利用 HBE-Zigbex-Motes(IEEE 802.15.4)无线传感器网络进行了人员计数实验.文献[8]利用 Wi-Fi 网络来改进建筑物内的定位技术,在工作中利用 Butterworth lowpass 滤波器、Savitzky-Golay 平滑滤波器分析讨论了利用 RSSI 进行人员计数表现.然而,由于多径效应,基于 RSSI 的衰减模型无法在复杂的室内环境中有效地反映信号在多路径传播的过程.文献[9]提出 Trained-once Crowd Counting 系统,该系统利用信道状态信息振幅的特征差数据,通过最小化测试样本与训练样本的 Davies-Bouldin 指数来实现对于室内人员的计数.文献[10]基于信道状态信息提出 FreeCount 系统,使用基于信息论的特征选择方案,选择对人体运动敏感的最具代表性的特征,建立对时间和环境差异具有鲁棒性的分类器,最后通过分类器得到室内人员的数量.

针对上述问题,本文提出一种基于 Wi-Fi 感知的人员计数方法.该方法使用的 CSI 源数据缓解了信号多径效应的影响^[11].使用解卷相位和线性变换的方法重构 CSI,使得 CSI 相位信息能以簇的形式得以集中^[12],避免了原始相位分布范围过大、随机性过高的问题.基于 Hampel 滤波器去除载波振幅中奇异数据,减少环境噪声对于人员数量特征造成的干扰,保证了人员计数的精度和稳定性.最后利用 SVM 分类器识别人员的数量.

1 信道状态信息

复杂的室内环境中,多径效应^[13]会导致无线信号的衰落与相移,影响信号的传输质量,最终导致基于无线信号的应用无法发挥良好的性能.而在现有的正交频分复用(orthogonal frequency division multiplexing,简称 OFDM)系统^[14]中,CSI 可以分别体现折射、散射、能量衰减对于信号的影响,并且描绘出信号从发射端至接收端的传播过程,从而减少多径效应对于信号传播的干扰.在窄带平坦衰落信道中,频域中的 OFDM 系统建模^[15]可以表示为

$$Y = HX + N \quad (1)$$

其中, Y 和 X 分别为接收端和发送端接收到的信号矩阵, H 为 CSI 矩阵, N 为噪声.

假设获取到 N 个子信道的 CSI,其矩阵表示为

$$H = [h_1, h_2, \dots, h_{N_s}] \quad (2)$$

其中,根据 802.11 a/g/n 协议^[18], $N_s=16/56/30/58/114$ 为最大载波数量,每个子信道 h_k 可以表示为

$$h_k = |h_k| e^{j\angle h_k} \tag{3}$$

其中, $|h_k|$ 和 $\angle h_k$ 分别表示子信道载波的振幅与相位.

2 Wi-Fi 感知的人员计数方法

基于 Wi-Fi 感知的人员计数方法分为 4 个阶段:(1) 获取被监测环境的 CSI 数据流;(2) 对 CSI 相位信息进行重构;(3) 使用 Hampel 滤波器去除载波振幅的噪声;(4) 基于有效相位信息与去除噪声后的振幅数据建立人员计数模型,其方法框架如图 1 所示.

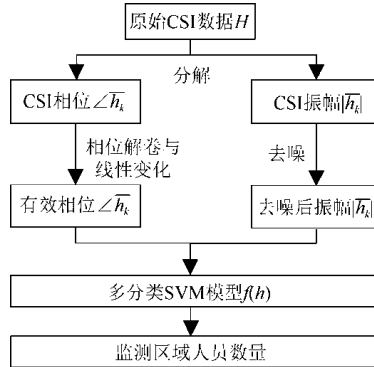


Fig.1 Method framework diagram

图 1 方法框架图

2.1 相位解卷与线性变换

如图 2 与图 3 所示,原始 CSI 在人员静止坐和随机动作时都是规则地环状分布,但是原始相位在室内人数改变时都无序地分布在 $(-\pi, \pi)$ 之间,这就造成了无法利用原始相位作为人员数量的有效特征.为了解决这一问题,文献[16,17]中分别利用相邻载波和天线之间的相位相减方法获取规则的相位差信息.

然而,采用直接相邻载波相位值相减的方法忽略了非相邻载波的关联性和载波频点序号^[18]对于相位变换起到的作用,因此直接利用相位差的方法易造成有效信息的丢失^[12].本文使用一种相位解卷和线性变化的方法对原始 CSI 进行重构,在不丢失有效信息的情况下获取到规则的相位信息.具体如公式(4)所示:

$$\begin{cases} b = \frac{\angle h_{114} - \angle h_1}{f_{114} - f_1}, \\ a = \frac{1}{114} \sum_{k=1}^{114} \angle h_k, \\ \angle \bar{h}_k = \angle h_k - b f_k - a \end{cases} \tag{4}$$

其中, $k=1,2,\dots,114$, 为 114 个载波序号, $f_k=-58, -57, \dots, -3, -2, 2, 3, \dots, 57, 58$ 为 OFDM 副载波的频点序号.该方法将每个载波的相位与所有载波的平均值和频点序号相关联,避免了相位有效信息的丢失.

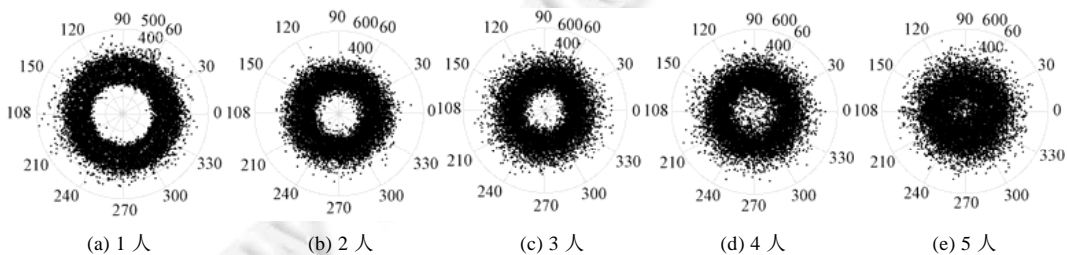


Fig.2 Distribution of CSI in indoor population of 1~5 persons during static seating

图 2 人员在静止坐时室内人数为 1~5 人时的原始 CSI 分布

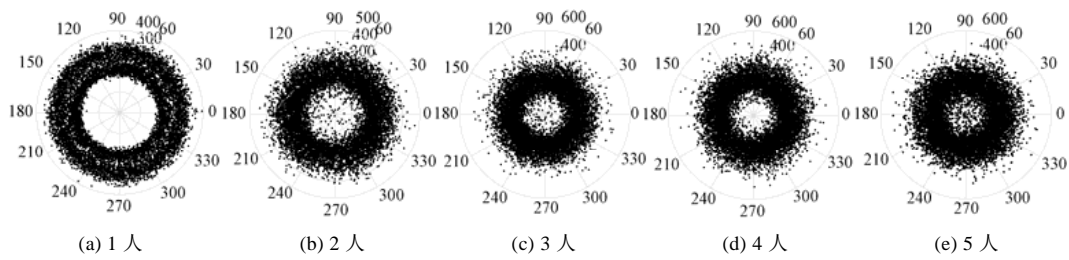


Fig.3 Distribution of CSI in indoor population of 1~5 persons
图 3 人员在随机动作时室内人数为 1~5 人时的原始 CSI 分布

图 4 和图 5 显示了在人员静止以及人员随机动作情况下相位重构之后的 CSI 分布图。

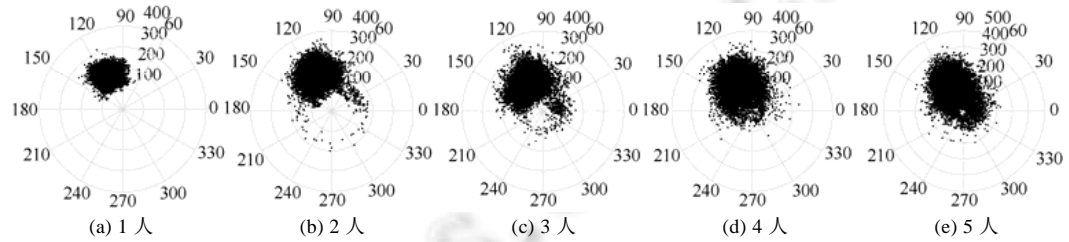


Fig.4 Reconstructed CSI distribution for indoor number of 1~5 persons
图 4 人员在静止坐时室内人数为 1~5 人时重构后的 CSI 分布

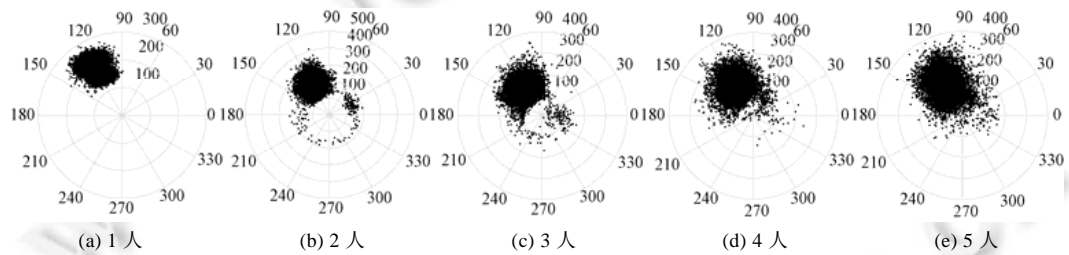


Fig.5 Reconstructed CSI distribution for indoor number of 1~5 persons
图 5 人员在随机动作时室内人数为 1~5 人时重构后的 CSI 分布

从图中可以看出 CSI 在重构之后能以簇的形式规则分布.当人数较少时,重构的相位大都分布在 $(\pi/2, \pi)$,当人数逐渐变多时,分布在 $(-\pi, \pi/2)$ 的相位逐渐变多.由此可见,随着人员数量的增加,无论人员处于静止坐还是随机动作,重构后的 CSI 相位分布都更加广泛.

2.2 振幅去噪

人员静止坐与随机动作状态下 CSI 振幅信息如图 6 和图 7 所示,环境因素引起的噪声主要会使数据产生奇异点数据.本文使用 Hampel 滤波器去除因环境因素产生的噪声,得到去噪后的 CSI 振幅信息.

Hampel 滤波被应用于检测和删除异常值,其基本思想是加权滑动平均.对于 x 的每个样本,如果某个样本与中值相差超过 t 个标准差,则用中值替换该样本.如果 x 是一个矩阵,hampel 将 x 的每一列都看成是独立的通道.其判定数据 x_i 准则如下所示:

$$\begin{cases} |x_i - m| > t \times Std, & \text{判断为奇异数据} \\ |x_i - m| \leq t \times Std, & \text{判断为正常数据} \end{cases} \quad (5)$$

其中, m 表示在数据序列的中值, Std 表示数据序列的标准差, t 表示一种标量阈值.当判断为奇异数据时,则奇异数据的位置用中值代替,其伪代码如算法 1 所示.根据算法 1,可以有效去除环境因素造成的噪声.

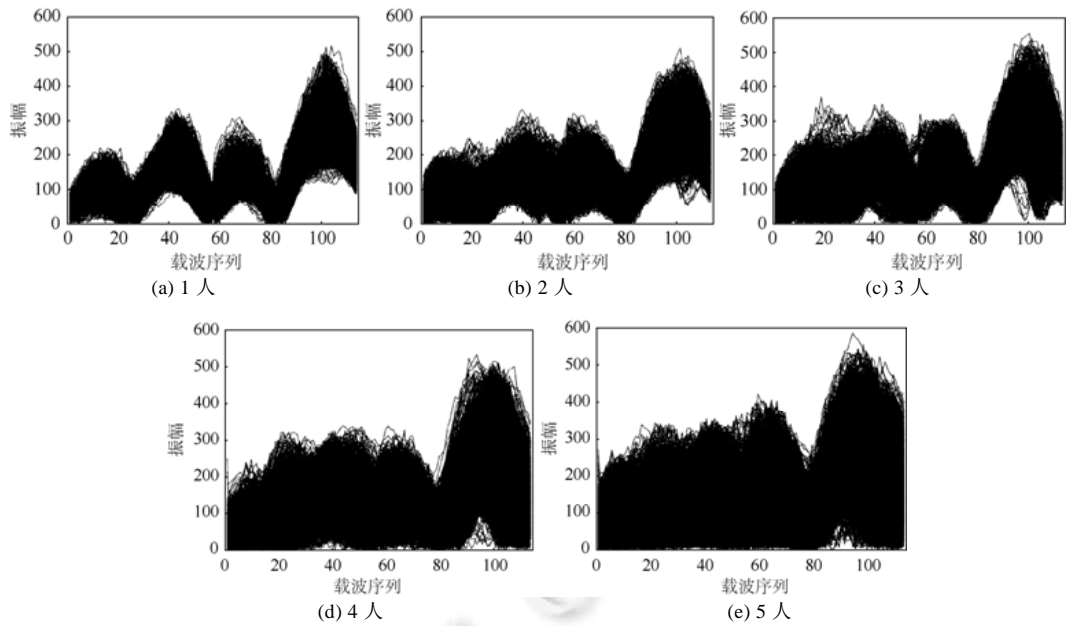


Fig.6 Primary CSI amplitude during static seating

图 6 人员静止坐时原始 CSI 振幅

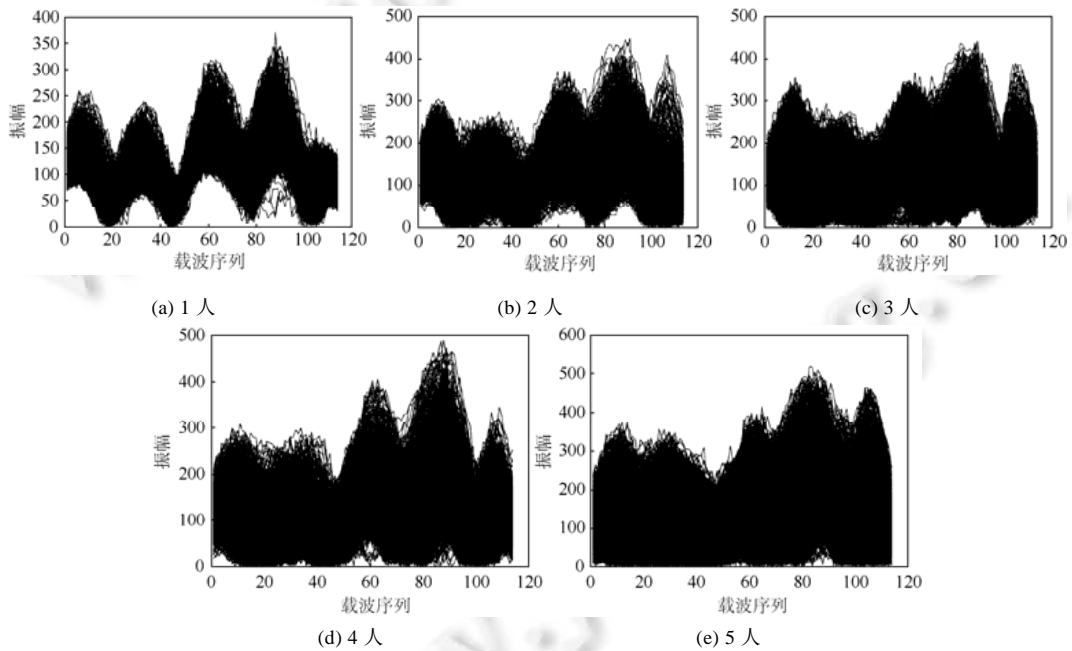


Fig.7 Original CSI amplitude in random action of personnel

图 7 人员随机动作时原始 CSI 振幅

算法 1. 基于 Hampel 滤波的去噪算法.

输入:CSI 的振幅序列 $|H_i|$;

输出:滤除噪声后的振幅序列 $|H'_i|$.

$|H| = [|h_1|, |h_2|, \dots, |h_N|]$, $|H'_i| = [|h'_1|, |h'_2|, \dots, |h'_N|]$

```

| $h'_1| = |h_1|, |h'_2| = |h_2|$  //前两笔数据直接复制
for  $k=3:N$  do
     $m = (|h_k| + |h_{k-1}| + |h_{k-2}|) / 3$  //计算中位数
    if  $||h_k| - m| > t \times Std$ 
then  $|H'_k| = m$  //用中值代替奇异数据
    end if
end for

```

室内人员静止坐(如图 6 所示)和随机动作(如图 7 所示)数据应用该方法检测删除奇异数据去除噪声后,无论在人员静止坐还是在做随机动作时,算法 1 处理过后的振幅数据在人员数量增加时,数据波动性和离散程度都会逐渐变大,这样的分布特性就使得振幅信息也能够有效地作为区分人员数量的特征,也证实了该去噪方法有着较强的鲁棒性.并且,通过图 8 和图 9 对比可知,人员静止坐时去噪效果更加明显,这主要是由于人员在室内的动作对于信号的干扰效果较大.

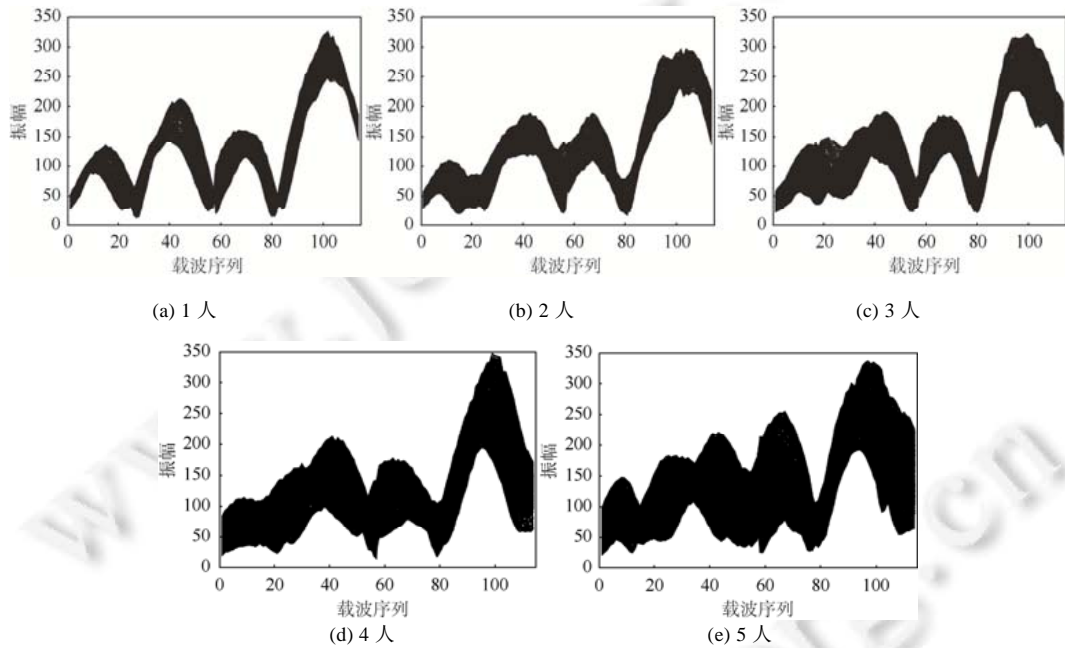


Fig.8 CSI amplitude after denoising during static seating

图 8 人员静止坐时去噪后 CSI 振幅

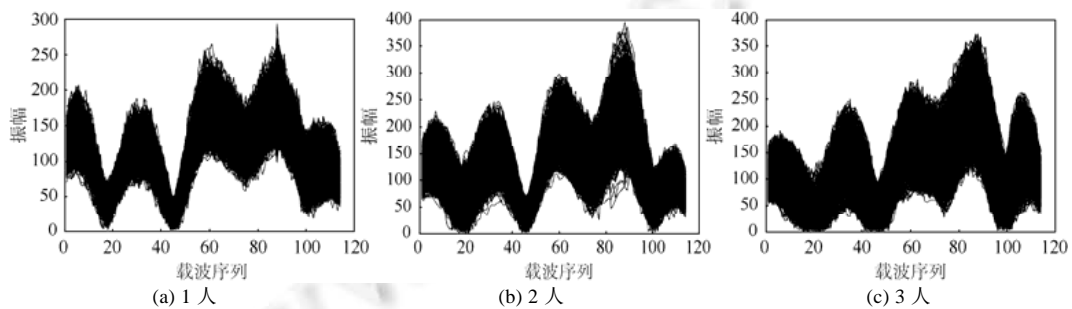


Fig.9 CSI amplitude after denoising during static seating

图 9 人员随机动作时去噪后 CSI 振幅

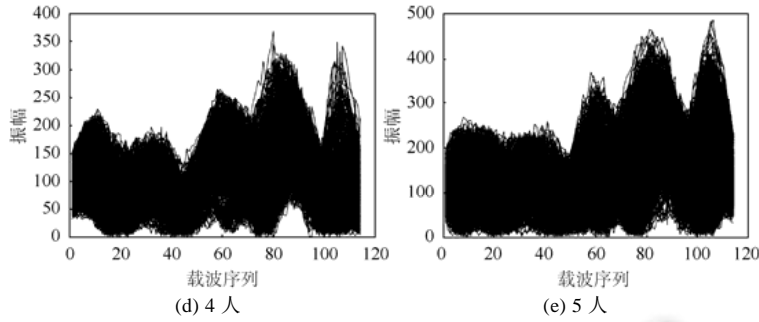


Fig.9 CSI amplitude after denoising during static seating (Continued)

图9 人员随机动作时去噪后 CSI 振幅(续)

2.3 基于多分类SVM人员数量识别

为了最终获得被监测区域人员的数量,本文利用多分类 SVM 算法找出不同人数的判定边界,首先需要构建人员数量和 CSI 组成的样本,即 (\bar{h}_i, y_i) ,其中, \bar{h}_i 由重构的相位 $\angle \bar{h}_i$ 和去噪之后的振幅 $|\bar{h}_i|$ 组成, y_i 表示当前样本人员的数量标记.由于多分类 SVM 的本质是进行多次二分类,因此在每次二分类时,用 $y_i \in \{-1, 1\}$ 作为二分类标记, $y_i=1$ 表示该样本属于该类别, $y_i=-1$ 表示该样本不属于该类别.

在寻找两类样本的划分界面时,两类的划分一般都是非线性的,因此可以将 CSI 通过非线性映射到更高维度的空间,然后在高维度的空间中寻找一个超平面 $wh+b=0$ 将两类分开,其求解方式如公式(6)所示:

$$\min_{w,b} \frac{1}{2} \|w\|^2, \text{ s.t. } y_i(w^T \bar{h}_i + b) \geq 1 \quad (6)$$

其中, \bar{h}_i 为重构之后的相位和去噪之后的振幅组成的 CSI.求解公式(6)可以得到最终分类的决策如公式(7)所示:

$$f(h) = \text{sgn} \left(\sum_{i=1}^N y_i \alpha_i K(\bar{h}_i, h) - b \right) \quad (7)$$

其中, sgn 返回所得结果的正负号, α 为求解公式(6)引入的拉格朗日乘子,满足条件 $0 \leq \alpha \leq C$,其中, C 为惩罚因子常数,设为 1. K 为实现非线性映射的核函数,本文采用的是 RBF 核函数,其映射关系如公式(8)所示:

$$K(\bar{h}_i, \bar{h}_j) = \exp \left(-\frac{\|\bar{h}_i - \bar{h}_j\|^2}{2\delta^2} \right) \quad (8)$$

其中, δ 为 RBF 核函数带宽,设置为 $\sqrt{\frac{N}{2}}$, N 为载波数 114.

3 实验结果与分析

3.1 实验设计

本文实验收集数据所采用的硬件设备为基于 OpenWrt 固件的 TP-link 4900v2 路由器.实验环境布置在 12m×12m 的实验室内.利用自主开发的上位机程序分别收集了不同人数(1~5 人)、4 种状态(静止坐、静止站立、随意走动和所有人随机动作)下的 CSI 数据各 2 万笔,共 40 万笔数据.训练模型时,各取 50% 的数据作为训练集与测试集.收集 114 个载波的信道状态信息时需要利用 OpenWrt 固件将信道带宽设置为 40MHz,信号频率设置为 5G.两台路由器摆在距离 16m 的实验室两端.

3.2 实验度量标准

查准率.对于给定种类 g 的 N_g 个样本,查准率的定义公式为 $N_g T / (N_g T + M_g F)$, $N_g T$ 为本来就为种类 g 的样本被正确识别为种类 g 的数量, $M_g F$ 为不是种类 g 的样本被识别为种类 g 的数量.

查全率.对于给定种类 g 的 N_g 个样本,查全率的定义公式为 $N_g T / N_g$.

3.3 结果分析

为了验证本文提出方法的有效性,实验中我们对比分析了本文方法与文献[9,10]中系统的识别结果.图 10 展示了人员静止坐、静止站立、随意走动、随机动作 4 种情况下人数在 1~5 人时的实验结果.相比于本文方法,图 10(a)中的 Trained-once Crowd Counting 系统在人数为 3 个以上时精确度下滑至 55%,无法有效地分辨出室内人员的数量.图 10(c)中本文方法的识别率比图 10(b)中的 FreeCount 系统识别率整体上升了约 5%.由此可以看出,本文方法相比于其他两种方法,在室内人员计数效果上有着明显的准确度提升.

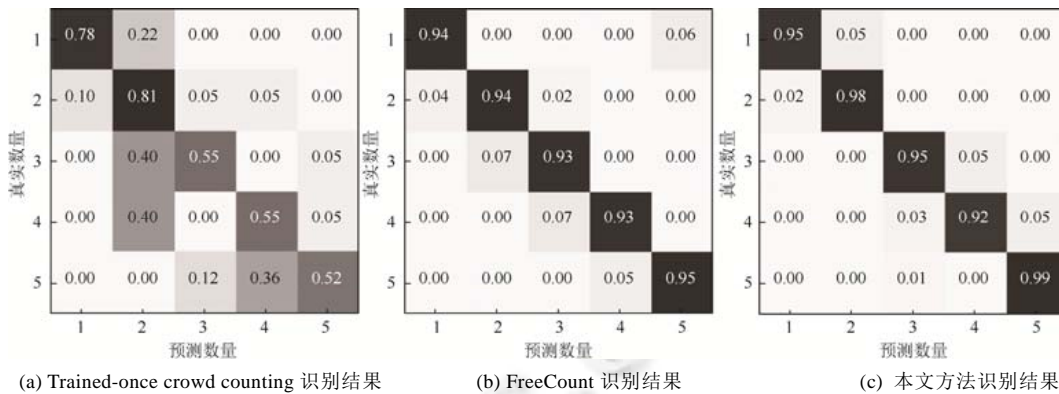


Fig.10 Comparison of recognition results of three methods

图 10 3 种方法识别结果对比

图 11 展示了 5 个人在不同情况下的查准率与查全率.所有动作平均查准率和查全率分别为 96.5%和 96.1%,显示了本文使用的方法能够适应人员在监测区域做不同动作的情况下统计出人员的数量.

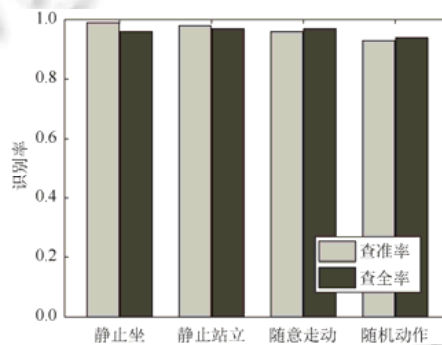


Fig.11 The results of identifying the number of personnel in method of this paper

图 11 本文方法识别人员数量的结果

4 结束语

本文提出的人员计数方法旨在解决目前人员计数硬件成本高、精度低、稳定性差的问题.基于 Wi-Fi 的数据源极大地减少了硬件部署的开销,解决了硬件成本过高的问题.多子载波 CSI 有效减少了多径效应的影响.利用解卷相位与线性变换在不丢失有效信息的情况下获取了规则的相位信息,避免了原始相位分布范围过大、随机性过高的问题.利用 Hampel 滤波器去除噪声后保证了利用无线信号进行人员计数的精度和稳定性.实验结果表明,本文提出的人员计数方法的精度约为 96%,能够准确地识别室内人员的数量.

References:

- [1] Pore SD, Momin BF. Bidirectional people counting system in video surveillance. In: Proc. of the 2016 IEEE Int'l Conf. on Recent Trends in Electronics, Information & Communication Technology (RTEICT). 2016. 724-727.

- [2] Ahmad I, Islam ZUI, *et al.* An FPGA based approach for people counting using image processing techniques. In: Proc. of the Int'l Conf. on Knowledge and Smart Technology (KST). 2019. 148–152.
- [3] Choi JW, Cho SH, Kim YS, *et al.* A counting sensor for inbound and outbound people using IR-UWB radar sensors. In: Proc. of the Sensors Applications Symp. 2016. 1–5.
- [4] Gupta G, Bhoje V, Singh J, Harish AR. Device-free crowd count estimation using passive UHF RFID technology. *IEEE Journal of Radio Frequency Identification*, 2018,3(1):3–13.
- [5] Priyantha NB, Chakraborty A, Balakrishnan H. The cricket location-support system. In: Proc. of the Int'l Conf. on Mobile Computing & Networking. ACM, 2000. 32–43.
- [6] Ji Y, Pandey S, Agrawal P. ARIADNE: A dynamic indoor signal map construction and localization system. In: Proc. of the Int'l Conf. on Mobile Systems. 2006. 1–15.
- [7] Doong SH. Spectral human flow counting with RSSI in wireless sensor networks. In: Proc. of the Int'l Conf. on Distributed Computing in Sensor Systems. IEEE, 2016. 110–112.
- [8] Zourmand A, Sheng NW, Hing ALK, AbdulRehman M. Human counting and indoor positioning system using Wi-Fi technology. In: Proc. of the 2018 IEEE Int'l Conf. on Automatic Control and Intelligent Systems. 2018. 142–147.
- [9] Domenico SD, Sanctis MD, Cianca E, *et al.* A Trained-once crowd counting method using differential Wi-Fi channel state information. 2016. 37–42. [doi: 10.1145/2935651.2935657]
- [10] Zou H, Zhou Y, Yang J, *et al.* FreeCount: Device-free crowd counting with commodity Wi-Fi. In: Proc. of the Globecom IEEE Global Communications Conf. 2018. 1–6.
- [11] Yang Z, Zhou Z, Liu Y. From RSSI to CSI: Indoor localization via channel response. *ACM Computing Surveys*, 2013,46(2):1–32.
- [12] Wang C, Liu J, Chen Y, *et al.* Towards in-baggage suspicious object detection using commodity Wi-Fi. In: Proc. of the 2018 IEEE Conf. on Communications and Network Security (CNS). 2018. 1–9.
- [13] Xie F, Zhang ZC, Zhang P, *et al.* The influence of multipath effect on direction finding error. *Electronic Measurement Technology*, 2010,33(1):29–31 (in Chinese with English abstract).
- [14] Nee RV, Prasad R. OFDM for Wireless Multimedia Communications. 5th ed., Boston: Richard van Nee, 2000.
- [15] Xi W, Zhao J, Li XY, *et al.* Electronic frog eye: Counting crowd using Wi-Fi. In: Proc. of the IEEE INFOCOM 2014—IEEE Conf. on Computer Communications. 2014. 361–369.
- [16] Wang C, Zheng X, Chen Y, *et al.* Locating rogue access point using fine-grained channel information. *IEEE Trans. on Mobile Computing*, 2016,16(9):2560–2573.
- [17] Kumar S, Gil S, Katabi D, *et al.* Accurate indoor localization with zero start-up cost. In: Proc. of the 20th Annual Int'l Conf. on Mobile Computing and Networking—MobiCom. 2014. 483–494.
- [18] I. Std. 802.11n-2009: Enhancements for higher throughput. 2009. <http://www.ieee802.org>

附中文参考文献:

- [13] 谢飞,张忠臣,张鹏,等.多径效应对测向误差的影响.电子测量技术,2010,33(1):29–31.



周泽仑(1995—),男,江苏苏州人,硕士生,CCF 学生会员,主要研究领域为无线传感网络.



黄河(1983—),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为无线传感网络.



戴欢(1983—),男,博士,副教授,CCF 专业会员,主要研究领域为无线传感网络.



史文华(1996—),男,硕士生,CCF 学生会员,主要研究领域为无线传感网络.