

单位代码： 10293 密 级： _____

南京邮电大学
专业学位硕士学位论文



论文题目： 基于 WiFi 的人体行为感知技术研究

学 号	1219043716
姓 名	朱旭
导 师	肖甫
专业学位类别	工程硕士
类 型	全日制
专业（领域）	计算机技术
论文提交日期	二〇二二年六月

Research on Human Behavior Perception Technology Based on WiFi

Thesis Submitted to Nanjing University of Posts and
Telecommunications for the Degree of
Master of Engineering



By

Xu Zhu

Supervisor: Prof. Fu Xiao

June 2022

摘要

在通信与计算机技术飞速发展的背景下，尤其是物联网时代的到来，人类成为信息世界的中心。为实现人与各种设备的交互，需要有效感知人的行为。当前感知技术主要可分为两类：基于传感器的感知和基于射频的感知。本文聚焦射频感知中基于 WiFi 的感知技术，在概述其现有研究工作的基础上，重点介绍我们实现的两个基于信道状态信息（Channel State Information, CSI）的人体行为感知应用方案。

在人机交互过程中，人员身份识别是保证安全可靠交互的前提。传统人员识别通常需要专用的感知设备，其限制了人员识别的应用范围。为实现非设备依赖的人员识别，本文利用 WiFi 信号来收集和识别人员的呼吸特征，提出了一种基于呼吸感知的人员识别系统。首先对 CSI 共轭相乘模型和人员呼吸模型进行理论分析；其次结合中值滤波、经验模态分解（Empirical Mode Decomposition, EMD）算法和基于快速傅里叶变换（Fast Fourier Transform, FFT）的子载波选择策略来分离 CSI 信号中的呼吸分量；然后针对多天线的 CSI 数据流，使用恒虚警（Constant False Alarm Rate, CFAR）寻峰算法来检测呼吸速率；最后通过滑动窗口算法生成数据集，使用卷积神经网络（Convolutional Neural Networks, CNN）完成人员的识别。实验结果表明，本系统检测呼吸速率的平均绝对误差约为 0.5bpm，且系统对三个场景中的人员识别准确率均达到 85%以上，满足人员识别的要求。

在人机交互技术中，高效的手语识别可极大提高交互能力。目前主流的自动手语识别输入设备主要包括基于数据手套的手语识别和基于视觉的手语识别，它们的共同缺点是设备成本高昂且部署不便捷，很难实现聋哑人群的大规模普及。本文基于泛在的 WiFi 信号，设计了一种设备无关的手语识别系统。首先通过 Hampel 滤波器对 CSI 数据进行降噪处理；然后采用基于动态时间归整（Dynamic Time Warping, DTW）的选择策略筛选可用子载波；最后采用基于 CNN 和双向长短期记忆（Bi-directional Long Short-Term Memory, BiLSTM）的深度学习模型来进行手语分类。实验结果显示，对于给定的 6 种代表不同含义的手语，本系统的手语平均识别准确率可达 90%以上。

关键词：行为感知，WiFi，CSI，人员识别，手语识别。

Abstract

Under the background of the rapid development of communication and computer technology, especially the advent of the Internet of Things era, human beings have become the center of the information world. To realize the interaction between people and various devices, it is necessary to effectively perceive human behaviors. Current sensing technologies can be mainly divided into two categories: sensor-based sensing and radio-frequency-based sensing. This thesis focus on the WiFi-based sensing technology in RF sensing. On the basis of an overview of its existing research work, it will focus on two implementations of human behavior sensing applications based on Channel State Information (CSI).

In the process of human-computer interaction, personnel identification is the premise to ensure safe and reliable interaction. Traditional human identification usually requires dedicated sensing equipment, which limits the application scope of human identification. In order to realize non-device-dependent human identification, this thesis uses WiFi signal to collect and identify the breathing characteristics of people, and proposes a human identification system based on breathing perception. Firstly, theoretically analyze the CSI conjugate multiplication model and personnel breathing model; secondly, combine median filtering, Empirical Mode Decomposition (EMD) algorithm and Fast Fourier Transform (FFT)-based subcarrier selection strategy to separate the breathing component in the CSI signal; then, for the multi-antenna CSI data stream, the Constant False Alarm Rate (CFAR) peak-finding algorithm is used to detect the breathing rate; finally, the sliding window algorithm is used to generate the data set, human identification is done by using Convolutional Neural Networks (CNN). The experimental results show that the average absolute error of the system in detecting the breathing rate is about 0.5bpm, and the accuracy rate of human recognition in three scenarios has reached more than 85%, which meets the requirements of human recognition.

In human-computer interaction technology, efficient sign language recognition can greatly improve the interaction ability. The current mainstream automatic sign language recognition input devices mainly include data glove-based sign language recognition and vision-based sign language recognition. Their common shortcomings are high equipment cost and inconvenient deployment, which makes it difficult to achieve large-scale popularization of deaf people. Based on the ubiquitous WiFi signal, this thesis designs a device-independent sign language recognition system. First, the CSI data is denoised by the Hampel filter; then the available subcarriers are screened by a selection

strategy based on Dynamic Time Warping (DTW); finally, a deep learning model based on CNN and Bi-directional Long Short-Term Memory (BiLSTM) is used for sign language classification. The experimental results show that for the given 6 sign languages representing different meanings, the average sign language recognition accuracy of this system can reach more than 90%.

Keywords: Behavior Perception, WiFi, CSI, Human Recognition, Sign Language Recognition.

目录

专用术语注释表	IX
第一章 绪论	1
1.1 课题背景	1
1.2 本文工作	3
1.3 论文组织结构	4
第二章 感知技术研究概述	6
2.1 基于传感器的感知技术	6
2.2 基于射频的感知技术	7
2.2.1 RFID 技术概述	7
2.2.2 UWB 技术概述	7
2.2.3 蓝牙技术概述	8
2.2.4 WiFi 技术概述	9
2.2.5 WiFi 感知技术的研究进展	10
2.3 本章小结	14
第三章 基于信道状态信息呼吸感知的人员识别系统	15
3.1 系统概述	15
3.2 理论模型	17
3.2.1 CSI 共轭相乘	17
3.2.2 菲涅尔域	18
3.3 数据预处理	18
3.3.1 中值滤波	19
3.3.2 EMD 算法	19
3.3.3 子载波选择	21
3.4 呼吸速率检测	22
3.4.1 CFAR 寻峰算法	23
3.4.2 呼吸速率计算	24
3.5 人员识别	24
3.5.1 数据集生成	25
3.5.2 分类模型	25
3.6 实验与评估	25
3.6.1 实验设置	26
3.6.2 表现评估	26
3.7 本章总结	29
第四章 基于 WiFi 的手语识别系统	31
4.1 系统概述	31
4.2 数据采集	33
4.3 数据预处理	33
4.3.1 数据插值	33
4.3.2 Hampel 去除离群值	34
4.3.3 Z-core 标准化	34
4.3.4 子载波选择	35
4.4 手语分类	36
4.5 实验与评估	37
4.5.1 实验设置	37

4.5.2 表现评估 38

4.6 本章总结 42

第五章 总结与展望 43

参考文献 45

附录 1 攻读硕士学位期间申请的专利 48

附录 2 攻读硕士学位期间撰写的论文 49

致谢 50

专用术语注释表

缩略词说明:

CSI	Channel State Information	信道状态信息
EMD	Empirical Mode Decomposition	经验模态分解
FFT	Fast Fourier Transform	快速傅里叶变换
CFAR	Constant False Alarm Rate	恒虚警率
CNN	Convolutional Neural Networks	卷积神经网络
DTW	Dynamic Time Warping	动态时间归整
BiLSTM	Bi-directional Long Short-Term Memory	双向长短时记忆
RFID	Radio Frequency Identification	射频识别
UWB	Ultra Wide Band	超宽带
AIDC	Auto Identification and Data Collection	自动标识与数据采集
WLAN	Wireless Local Area Network	无线局域网
RSSI	Received Signal Strength Indication	接收信号强度指示
MAC	Media Access Control	介质访问控制层
OFDM	Orthogonal Frequency Division Multiplexing	正交频分复用
LOS	Line of Sight	视距路径
LF	Low Frequency	低频
HF	High Frequency	高频
UHF	Ultrahigh Frequency	超高频
DOSM	Dual One Slope Model	双单斜率模型
DOSSOM	Dual One Slope with Second Order Polynomial Model	二阶多项式双单斜率模型
MIMO	Multiple Input Multiple Output	多输入多输出技术
ToF	Time of Flight	飞行时间
TWR	Two-Way Ranging	双向测距
TDofA	Time Difference of Arrival	到达时间差
AoA	Angle-of-Arrival	到达角
LDPL	Log-normal Distance Path Loss	对数距离路径损耗
NIC	Network Interface Controller	网络接口控制器
SVDE	Support Vector Distribution Estimation	支持向量分布估计
STFT	Short-Time Fourier Transform	短时傅里叶变换
DNN	Deep Neural Networks	深度神经网络
CWT	Continuous Wavelet Transform	连续小波变换

SAC	Sparse Approximation based Classification	基于稀疏近似的分类
PCA	Principal Component Analysis	主成分分析
SVM	Support Vector Machine	支持向量机
RBF	Radial Basis Function	径向基函数
LSTM	Long Short-Term Memory	长短时记忆
MLE	Maximum Likelihood Estimator	最大似然估计
DWT	Discrete Wavelet Transformation	离散小波变换
SNR	Signal-to-Noise Ratio	信噪比
CM	Conjugate Multiplication	共轭相乘
BSS	Blind Signal Separation	盲源信号分离
ICA	Independent Component Analysis	独立成分分析
FWPT	Fuzzy Wavelet Packet Transform	模糊小波包变换
HCI	Human-Computer Interaction	人机交互
kNN	k-Nearest Neighbor	k 最邻近分类算法
DBN	Deep Belief Network	深度置信网络
HMM	Hidden Markov Model	隐马尔可夫模型
CFR	Channel Frequency Response	信道频率响应
WD	Wavelet Decomposition	小波分解
IMF	Intrinsic Mode Function	本征模函数
DFT	Discrete Fourier Transform	离散傅里叶变换
MAD	Median Absolute Deviation	中值绝对偏差
ICS	Internal Covariate Shift	内部协变量偏移
RNN	Recurrent Neural Network	循环神经网络

第一章 绪论

1.1 课题背景

基于快速发展的通信与计算机技术，物联网时代各种智能设备实现了爆发式的增长。飞速普及的各类设备，使人类日益成为信息世界的中心。而如何实现人与各种设备交互，甚至是人与虚拟世界的交互，都需要对人的行为进行感知。基于对人体行为进行感知的需求，催生出多种多样的感知手段，涌现了许多受到市场高度认可的应用，覆盖到社会生产生活的各个方面。

目前的行为感知技术，可以从感知手段角度分为主要的两类：基于传感器的感知技术和基于射频的感知技术。其中基于传感器的感知技术依赖于各类传感器设备，例如运动传感器^[1]、光学传感器^[2]和声学传感器^[3]等。光学传感器市场成熟度相当高，不过在行为感知方面，却受到光线条件的限制，感知范围受限，并且存在侵犯用户隐私，泄露用户信息的危险。运动传感器一般常见于手机中的加速度计与陀螺仪，通常作为某类电子设备的一部分发挥作用，所以一般需要用户随身携带此类设备，这不利于行为感知的普适性要求。而其他的声学传感器或红外传感器等，都有着识别率低，易受干扰和作用距离短等劣势。总之，用户需要不依赖设备，精度高，部署成本低，广泛存在的感知手段。

而基于射频的感知技术，主要有射频识别（Radio Frequency Identification, RFID）^{[4][5][6][7]}、超宽带（Ultra Wide Band, UWB）^{[8][9][10]}、蓝牙^{[11][12]}和 WiFi^{[13]~[56]}等。RFID 是自动识别和数据采集（Auto Identification and Data Collection, AIDC）技术中的一种，一个 RFID 系统包含标签、阅读器和主机这些基本组件。RFID 系统利用标签来采集数据，通过标签中的天线，以无线信号的形式将数据传给阅读器，然后阅读器将无线电波转换为可用的数据形式，这些数据被存储在主机中用于进一步的分析。RFID 技术的缺陷有几点：一是标准不统一导致的设备兼容问题，二是 RFID 设备的成本问题，三是 RFID 技术易受信号干扰，安全性较差。超宽带（UWB）是一种快速、安全和低功耗的短距离无线通信协议。它通过无线电波使设备能够相互通信，类似于蓝牙和 WiFi，但更精确、可靠和有效。虽然 UWB 的优点使其有着很好的发展前景，但它的缺点也比较明显，首先探测距离较短，一般只能在一个房间大小的范围内工作。其次现今主流的电子产品对于 UWB 的支持有限，普及度相较于蓝牙和 WiFi 来说很低，这大大限制了它的应用。最后，UWB 依托于设备，这与不依赖设备进行行为感知的愿景相冲突。蓝牙是一种低功耗、用于短距离无线数据传输和语音连接的开放规范。它还可以用于定

位和跟踪人员,在医学、交通和工业生产等领域都有应用。与无线局域网(Wireless Local Area Network, WLAN)相比,蓝牙的总比特率更低,范围更小(通常为 10-15m)。短距离特性使其无法进行密集的本地化部署。

基于 WiFi 的感知技术在无线网络广泛普及的背景下,得到了研究人员越来越多的关注。当前主要基于接收信号强度指示(Received Signal Strength Indication, RSSI)和信道状态信息(Channel State Information, CSI)这两种不同的信号进行研究。来自介质访问控制层(Media Access Control, MAC)层的 RSSI 信息虽然可以从绝大部分的商用 WiFi 设备中获得,但是其粗粒度的特征使其只能检测有限类型的活动,稳定性差,分辨率低。CSI 是衡量信道状况的信息,属于物理层,来自正交频分复用(Orthogonal Frequency Division Multiplexing, OFDM)系统下解码的子载波。与 RSSI 相比,对于多个 OFDM 子载波,CSI 的初始测量值包含一组复值,包括幅值和相位信息。每个中心频率不同的子载波经历不同的多径衰落效应,所有子载波共同细粒度的表示无线信道的特征。目前,基于 WiFi 的感知技术在传感、识别和检测领域都有诸多应用,如人员识别、生命体征检测、活动识别和手势识别等。

人员识别是各种智能应用的重要功能。这些应用程序通常需要人员识别功能来控制访问权限和提供服务。当前最先进的人员识别技术依赖于人类的生理和行为特征,称为生物识别技术。生物特征识别是指根据个体的人类生物学和行为特征自动识别个体。用于人类识别的生物特征包括指纹、面部、虹膜和语音。由于生物特征是个人固有的,并且独一无二,因此生物特征被广泛用于人员识别。此外,由于生物特征伪造的难度较高,基于生物特征的技术能够应对日益增长的安全威胁,并在个性化和便利性方面比密码等传统安全方法具有明显的优势。但即使如此,这些基于生物特征的人员识别方法通常也有劣势,例如,相机可用于面部识别,但收集用户的照片时需要设置特定方向和范围;基于相机的人员识别方法也会受较差的光线条件限制,或遭到来自 AI 的攻击^[57]。另一方面,日常生活中的 WiFi 信号无处不在,通过提取人类造成的信号变化与其行为之间的相关性,利用 WiFi 信号进行面向人的感知是一种有前景的传感技术。这种无设备解决方案利用广泛存在的 WiFi 基础设施,为用户提供非侵入式识别,显著改善了用户体验,扩大了应用范围。WiFi 感知的兴起为人员识别提供了新的解决方案。

识别用户通常取决于用户的生物特征,例如指纹。在本文中,则关注呼吸这种独特的人类生物学特征,它能满足人员识别的要求。人的呼吸通常由人的身高、体重和肺的结构决定。人体的呼吸会导致胸腔移动(5-12mm),从而改变 WiFi 信号在空间中的传播路径。通过分析 WiFi 传播路径变化,可以实现对人员呼吸的细粒度感知。在通过感知呼吸进行人体识别的场景中,用户呼吸时胸部的移动速度和距离不同,对 WiFi 传播路径的影响也不同。接收端分析

WiFi 信号变化,然后在用户数据库中查找匹配记录,实现对人的识别,这对许多面向用户的应用程序很有帮助。所以本文在人员识别和生命体征监测的研究基础上,提出了基于信道状态信息呼吸感知的人员识别系统。

手势识别是计算机科学的一个不断发展的领域,有着相当多的研究与应用。手势识别是人机交互的重要组成部分,例如 VR 设备的使用等。本文中的手语识别,也可以视为手势识别的细分领域。手语是一种使用手部动作、肢体动作和面部表情进行交流的语言,主要由聋哑人和部分听力正常的人使用。手语有着和口语和书面语一样的复杂性和表达力,有着自己独特的语法结构。不同地区的手语是分化的,这使得不同地区的聋哑人群交流时,也会像正常人一样遇到语言障碍。种种因素表明,普通人学习手语的成本并不低,然而手语使用者在任何国家都占总人口的 0.1%以上,因此,为这些群体开发更多交流手段,能够提高他们在工作生活中的便利性,帮助他们更好的融入社会。

从现有的手语输入设备来看,手语识别系统主要分为基于数据手套的手语识别系统和基于视觉(图像)的手语识别系统。基于数据手套的手语识别系统,是利用数据手套和位置跟踪器测量手势在空间运动的轨迹和时序信息。这种基于传感器的方法的优点是系统的识别率高,缺点是执行手语的人要穿戴复杂的数据手套和位置跟踪器,输入设备成本较高,不利于大规模的推广。基于视觉的手势识别是利用摄像机采集手势信息,并进行识别。该方法的优点是输入设备较为常见,相对于数据手套进一步降低了成本,但识别率比较低,实时性较差,受光照等条件束缚,并且有泄露用户个人信息的危险。所以本文提出了基于 WiFi 的手语识别系统,相对而言是一种低成本,适用范围广,能降低隐私泄露风险的解决方案。

1.2 本文工作

本文首先介绍了行为感知的主要技术方向,即基于传感器的感知和基于射频的感知,其中着重介绍了 WiFi 感知技术的发展,比较并分析了 RSSI 和 CSI 的优缺点,并列举了 WiFi 感知技术在人员识别和生命体征监测等多个领域中的研究进展。基于商用 WiFi 设备,在第三章和第四章,分别设计并实现了两个人体行为感知领域的应用:基于信道状态信息呼吸感知的人员识别系统和基于 WiFi 的手语识别系统。本文的主要工作如下:

(1) 介绍了当前行为感知技术的主流技术,大致分为基于传感器的感知和基于射频的感知,其中用于感知的传感器可分为动作传感器、视觉传感器和声学传感器等,基于射频的感知技术则有 RFID、UWB、蓝牙和 WiFi 等不同的门类。

(2) 比较并分析了 RSSI 和 CSI 的优缺点,相较粗粒度、不稳定且具有时间波动性的 RSSI

信号, CSI 稳定性高, 能够让所有子载波一起以一种细粒度的方式描绘无线信道, 进一步拓展了应用场景。

(3) 考虑到传统的人员识别方法通常对用户有特殊要求且硬件成本较高。第三章提出了一个基于信道状态信息呼吸感知的人员识别系统。首先对 CSI 的共轭相乘模型和菲涅尔域模型进行理论分析。然后结合中值滤波、经验模态分解 (Empirical Mode Decomposition, EMD) 算法和基于快速傅里叶变换 (Fast Fourier Transform, FFT) 的子载波选择策略来分离 CSI 信号中的呼吸分量。基于多天线的 CSI 数据流, 使用恒虚警 (Constant False Alarm Rate, CFAR) 寻峰算法来检测呼吸速率。通过滑窗算法生成数据集, 使用卷积神经网络 (Convolutional Neural Networks, CNN) 完成人员识别的任务。针对不同的因素, 本文对人员识别系统进行了多次实验。实验结果表明, 本系统检测呼吸速率的平均绝对误差接近 0.5bpm, 人员识别准确率达到了 85% 以上, 可以满足人员识别的要求。

(4) 为了增加聋哑人士的交互手段, 改善他们的交流便利性。第四章提出了一个基于 WiFi 的手语识别系统。本系统基于商用设备实现, 不需要用户佩戴其他设备, 成本低廉, 易于部署。系统采用 Hampel 滤波器滤除原始数据的离群值, 去除环境噪声, 并通过一维线性插值和归一化的手段保证数据样本的一致性。采用基于动态时间归整 (Dynamic Time Warping, DTW) 的选择策略筛选有效的子载波。识别手语时采用基于 CNN 和双向长短期记忆 (Bi-directional Long Short-Term Memory, BiLSTM) 的深度学习模型, 识别准确率优于主流机器学习模型。

1.3 论文组织结构

本文的组织结构如下:

第一章: 绪论。介绍了课题背景、迄今国内外研究者的在行为感知领域的研究现状, 简要概括了本文所做的研究工作和论文的基本组织结构。

第二章: 感知技术研究概述。人体行为感知研究在技术上可分为基于传感器的感知和基于射频的感知, 本章第一节举例介绍了用于感知的各类传感器, 第二节介绍了射频感知的各个技术分支, 如 RFID、UWB、蓝牙和 WiFi 等。其中着重阐述了 WiFi 感知中 RSSI 信号与 CSI 信号的差异, 详细介绍了它们在多个领域中的研究成果。

第三章: 基于信道状态信息呼吸感知的人员识别系统。本章共五节。第一节系统概述描述了实现系统时面临的挑战和本系统为解决挑战所做的贡献。第二节为先验知识, 介绍了 CSI 共轭相乘模型和基于菲涅尔域的人员呼吸模型, 论证了 CSI 共轭相乘模型用于实验的可行性。第三节到第五节详细介绍了人员识别系统的实施流程。第六节介绍了实验的设置, 并评估了

系统在各种场景下的表现。最后一节是本章总结，回顾人员识别系统并提出可以改进的地方。

第四章：基于 WiFi 的手语识别系统。本章共六节。首先介绍了实验背景，系统实施遇到的挑战和所做的贡献。第一节系统概述给出了手语识别系统的大致流程。第二节描述了采集数据的具体细节，第三节具体介绍了数据预处理的各项步骤，第四节介绍了手语分类模型。第五节是实验设置和表现评估。最后一节对手语识别系统作出总结。

第二章 感知技术研究概述

人类行为感知近些年来受到了广泛的关注并成为一个重要的研究领域。本文针对现有感知系统的基本原理、技术、系统结构和典型应用进行概述。基于感知技术将现有感知手段分为两类：基于传感器的感知技术和基于射频的感知技术，并详细阐述两大类中的具体细分技术。着重介绍了基于 WiFi 信号衍生出的四类应用，即人员识别、生命体征监测、活动识别和手势识别。讨论利用无线信号进行人类行为感知的未来研究方向和局限性。

2.1 基于传感器的感知技术

研究者为了对人员进行有效的感知，部署了各类传感器，对人员的行为活动进行跟踪、识别和检测，其中包括运动传感器^[1]，基于视觉的传感器^[2]和基于声学的传感器^[3]等。本节将简要介绍以上提到的三种传感器，以及它们的局限性。

(1) 基于运动传感器的方法通常需要用户佩戴专门的设备来跟踪身体运动，这些设备内置了诸如加速计、陀螺仪和磁力计等用于记录各类数据的传感器。而传感器记录的数据，则可用于行为识别，定位和轨迹追踪。基于运动传感器的感知已经是一个相当成熟的技术领域，有着广泛的应用。然而，在实际的生产生活中，要求用户一直佩戴相关设备是一件相当麻烦的事，所以难以实现普遍的人员监测，因此这类感知手段的普适性并不高。

(2) 传统基于视觉的方法可以用于被动的感知、识别和检测，这些方法硬件上依赖于摄像机或可见光传感器。研究者利用这些硬件设备采集人的视频或图像，然后使用计算机视觉、机器学习或深度学习技术来识别人类活动。基于视觉的方法存在一些局限性，包括视距路径（Line of Sight, LOS）要求，这意味着障碍物（墙壁、门等）的存在将妨碍设备的准确检测。其次是光照要求，这意味着在低光照条件、烟雾或不透明障碍物的干扰下，很难进行感知，因此只能在特定的光照条件下工作。最后则是关于侵犯个人隐私的担忧，即视频和图像可能包含大量个人隐私信息，这类传感器不能用于许多具有私密性的位置地点，如浴室和卧室等。

(3) 基于声学的方法的稳定性容易受到环境噪声和周围物体的干扰，而且由于声信号的快速衰减，测量范围也受到限制。

总的来说，此类基于传感器的感知技术，发展相当成熟，已经得到了广泛的应用。然而利用传感器系统进行感知，往往需要部署和配置复杂的软硬件设施，而且单一传感器能够感知的信息有限，往往需要多种类的传感器相互配合才能应对更复杂的场景。此外传感器的软

硬件设施一般成本高昂，对于普通大众来说难以承受。在涉及隐私的场景中，传感器侵犯隐私的可能性也会引发用户的担忧。最后，随着用户对于设备无关的感知需求越来越高，基于传感器的感知技术也越发不具有普适性。

2.2 基于射频的感知技术

非接触和非视距的特性使得基于射频的方法成为更加理想的感知手段。在基于射频的系统中，从发射机发射的无线信号在到达接收机之前，受到人类和环境相应活动的影响。通过分析接收到的无线信号，能够实现对人的感知、识别和检测。在近些年，基于射频的技术得到了广泛的发展，包括 RFID、蓝牙、UWB 和 WiFi 等。

2.2.1 RFID 技术概述

RFID 是一种利用无线射频方式进行非接触式双向通信交换数据以达到自动识别的技术。一个 RFID 系统包含标签、阅读器和主机这些基本组件。RFID 系统一般有三种类型：低频（Low Frequency, LF）、高频（High Frequency, HF）和超高频（Ultrahigh Frequency, UHF）。LF 系统的范围在 30KHz 到 300KHz 之间，读取范围达到 10cm，常用于访问控制和人群监控。HF 系统的范围在 3MHz 到 30MHz 之间，读取范围从 10cm 到 1m，通常用于电子票据、支付或用户体验应用程序。UHF 系统的范围在 300MHz 到 3GHz 之间，读取范围高达 12m，常用于零售库存跟踪、门禁和资产管理。

研究人员利用 RFID 开发了许多室内定位^[4]、轨迹跟踪^[5]和活动监测^[6]系统。例如，Hatem^[4]等人提出了基于有源 UHF-RFID 标签的两个新经验路径损耗模型：双单斜率模型（Dual One Slope Model, DOSM）和二阶多项式双单斜率模型（Dual One Slope with Second Order Polynomial Model, DOSSOM），使用这两种新模型的距离误差与实际测量值非常匹配，体现了这两种新的室内传播模型的稳定性。Yang 等人^[7]提出了一种基于 RFID 的驾驶环境呼吸监测系统。他们将多个 RFID 标签连接在安全带上，系统根据从标签采样的相位值估计驾驶员的呼吸速率，同时利用标签的多样性来削弱驾驶环境中强噪声的影响。并采用张量补全和张量典范多元分解方法处理相位值，以克服跳频、采样率不稳定、车辆振动和其他环境运动的影响。

2.2.2 UWB 技术概述

UWB 是一种快速、安全和低功耗的短距离无线通信协议。它通过无线电波使设备能够相

互通信，与蓝牙和 WiFi 类似，超宽带是一种使用无线电波的无线通信协议，但更精确、可靠和有效。UWB 发射器在宽频谱频率上发送数十亿个无线电脉冲，然后 UWB 接收器将脉冲转换为数据。UWB 脉冲可用于感知两个发射器之间的距离。脉冲的持续时间越短，距离测量就越精确。UWB 能够实时确认精度，因为它每秒发送多达 10 亿个脉冲（约每纳秒 1 个）。UWB 通过脉冲发送模式对信息进行编码。此外为了增加 UWB 的范围和接收可靠性，研究人员已将多输入多输出（Multiple Input Multiple Output, MIMO）技术添加到协议标准中。

UWB 一般有着低功耗、高精度和信号干扰小的优点。UWB 是低功耗系统，适合于嵌入智能手机这种受限于电池的便携设备。UWB 不依赖于信号强度，而是使用飞行时间（Time of Flight, ToF）、双向测距（Two-Way Ranging, TWR）、到达时间差（Time Difference of Arrival, TDoA）、到达角（Angle-of-Arrival, AoA）等技术来确定与另一个设备的距离。精确的角度与精确的距离相结合，意味 UWB 在精确定位等应用领域的精度远超其他感知技术。UWB 使用 3.1~10.6GHz 频率，因此任何频段的信号干扰的可能性都比较有限，这通常是蓝牙和 WiFi 的缺点。但 UWB 的缺点也比较明显，首先探测距离较短，一般只能在一个房间大小的范围内工作。其次现今主流的电子产品对于 UWB 的支持有限，普及度相较于蓝牙和 WiFi 来说很低，这大大限制了它的应用。最后，UWB 依托于具体的设备，这与不依赖设备进行行为感知的初衷相冲突。

UWB 是一种很有前途的感知技术，在手势检测、人员识别，呼吸监测方面都有着应用，例如 Junbum Park 等^[8]提出的基于脉冲超宽带（Impulse Radio-Ultra Wide Bandwidth, IR-UWB）雷达的手势识别算法，识别准确率高达 99.69%。Assawaroongsakul 等人^[9]利用 UWB 脉冲具有高空间分辨率并能够穿透介电材料的特点，在 3GHz 频率范围内穿过墙壁收集人类活动的信号，将信号数据与深度神经网络模型结合，对站立、行走、坐、躺和无人状态 5 类活动进行分类，精度高达 96.94%。T J Daim 等人^[10]基于 UWB 雷达传感器，以非接触的方式确定儿童的各种呼吸状态。

2.2.3 蓝牙技术概述

蓝牙是一种低功耗、用于短距离无线数据传输和语音连接的开放规范。它还可以用于定位和跟踪人员，在医学、交通和工业生产等领域都有应用。作为小型收发器的蓝牙标签具有唯一的 ID，可用于定位。基于蓝牙的定位系统利用接收信号强度和距离之间的关系来计算米级精度的距离^{[11][12]}。与 WLAN 相比，蓝牙的总比特率更低，范围更小（通常为 10-15m），短距离特性使其无法进行密集的本地化部署。

Iglesias 等人^[11]研究将蓝牙信号作为定位系统中的信息源,分析其功能,并开发了一组算法来转换蓝牙数据,以改进定位过程。他们在场景中布置节点网络,节点由蓝牙设备和定向天线组成,系统通过查询所有可见蓝牙设备的 RSSI 值来工作,采用指纹技术为每个设备确定最可能的位置网格单元。Hou 等人^[12]基于蒙特卡罗方法,利用加速度计、指南针和低功耗蓝牙信标,来定位在室内使用移动设备的人,对用户坐标的计算误差小于 1m。

2.2.4 WiFi 技术概述

基于广泛存在的商用 WiFi 设备, WiFi 信号已被广泛用于各种无线感知应用。当前主要基于 MAC 层的 RSSI 和物理层的 CSI 这两种不同的信号进行研究。

RSSI 的中文含义是接收信号强度指示。大多数 WiFi 设备都能采集 RSSI 信息,它表示无线信号在一定距离内的路径损耗,可以根据对数正态距离路径损耗(Log-normal Distance Path Loss, LDPL)模型推导出:

$$P(d) = P(d_0) + 10\gamma \log \frac{d}{d_0} + X_\delta \quad (2.1)$$

$P(d)$ 表示 RSSI 测量值,表示距离 d 处的路径损耗,单位为分贝(dB), $P(d_0)$ 是参照距离 d_0 处的路径损耗, γ 是路径损耗指数, X_δ 是由平坦衰落引起的正常零均值噪声。

然而研究人员注意到,无线传感区域内人体的存在会导致信号衰减,从而导致 RSSI 测量值的变化。因此, RSSI 近年来已广泛应用于人类活动传感,例如无设备室内定位、房间内人群密度估计和呼吸频率监测。尽管 RSSI 信息很容易在商用 WiFi 设备中获得,无需额外的硬件,但由于其粗粒度的特性,使得它只能检测有限类型的人员活动。此外,众多研究已经表明,即使在静态的室内环境中, RSSI 的稳定性也无法保证,这让基于 RSSI 的系统在许多应用场景中都不能可靠工作。

为了实现准确可靠的人员活动传感,必须利用更细粒度的 CSI,它代表了信号散射、衰落和随距离衰减的功率的综合效应。由于无线信号几乎可以穿过室内环境中的所有角落,因此人体的存在或移动会改变无线信号的传播,都会导致多条反射信号的微小变化。所有这些多径信号都有助于 CSI 测量值用于检测和跟踪人体运动。与 RSSI 相比, CSI 由一组复数值组成,包括幅度和相位信息,包含了 OFDM 子载波。每个子载波的中心频率略有不同,会经历不同的多径衰落效应,并且所有子载波一起以一种细粒度的方式描绘了无线信道。IEEE 802.11n 标准可以为 52 和 128 个子载波分别提供 20MHz 和 40MHz 带宽,新兴的 802.11ac 标准支持更宽的带宽。CSI 本质上允许细粒度的信道估计,可以表示为:

$$H = (H_1, H_2, \dots, H_i, \dots, H_N)^T, i \in [1, N] \quad (2.2)$$

公式中 N 代表子载波的数目，而 H_i 可以表示为：

$$H_i = |H_i|e^{j\sin(\angle H_i)} \quad (2.3)$$

其中 $|H_i|$ 和 $\angle H_i$ 分别是第 i 个子载波的幅度与相位。与 RSSI 类似，CSI 测量值可以在任何具有现成 WiFi 网络接口控制器（Network Interface Controller, NIC）的设备上利用修改的驱动程序获得，例如 Intel 5300 网卡和 Atheros 9580 网卡。

2.2.5 WiFi 感知技术的研究进展

现在越来越多的研究人员广泛采用 CSI 来进行人体活动感知，例如人员识别，生命体征监测，活动识别和手势识别等。下面将详细介绍 WiFi 在不同应用方向上的研究进展。

（1）基于 WiFi 的人员识别

由于不同人之间存在固有的行为和生理差异，研究人员已经证明了通过受人类活动影响的无线信号来执行用户身份验证的可能性。这种无设备方法利用物联网设备（即智能冰箱、智能电视和恒温器等）提供的 WiFi 信号，成本低且易于部署，并且还保护了用户的隐私。在本节中，将简述使用 WiFi 信号进行人员识别的相关工作。

目前，利用 WiFi 信号进行人员识别的方法主要有两种。第一种是基于人体动作的识别。由于每个人的行为习惯不同，在做同样的动作时，对无线信号的干扰也不同。WiID^[13]使用短时傅立叶变换（Short-Time Fourier Transform, STFT）提取人员动作的时频特征，并结合支持向量分布估计（Support Vector Distribution Estimation, SVDE）为不同的人构建运动模型。WiID 的 7 人识别准确率可达 85%以上。WiHF^{[14][15]}提出了一种基于 seam-carving 算法的跨域动作识别新方法，将获得的动作特征与深度神经网络（Deep Neural Networks, DNN）结合实现跨域人员识别，6 人识别准确率可达 96.74%。第二种是基于步态的识别。与动作识别相比，这种方法与人体生物特征有关，同时不易受环境干扰，鲁棒性更强。WiFi-ID^[16]使用连续小波变换（Continuous Wavelet Transform, CWT）从 CSI 幅度数据中获取频谱特征，并通过 ReliefF 算法实现特征权重的划分，最后结合基于稀疏近似的分类（Sparse Approximation based Classification, SAC）实现基于步态的行人识别，对于 6 位实验者，实现了 77%的识别准确率。Wii^[17]使用主成分分析（Principal Component Analysis, PCA）从 CSI 数据中提取用户步态特征，并使用支持向量机（Support Vector Machine, SVM）和径向基函数（Radial Basis Function, RBF）内核来完成人员识别。随着深度学习的不断发展，最新的工作开始使用深度网络来实现高精度的人体识别。Deep-WiID^[18]结合门控循环单元和平均池化提取步态特征，然后基于深度学习实现人员识别，对 15 人的最大检测精度为 92.5%。CSIID^[19]和 HumanFi^[20]都使用长短

时记忆网络 (Long Short-Term Memory, LSTM) 从 CSI 信息中提取用户的步态特征来实现人员识别。CSIID 可以最多对 6 人达到 94.8% 的检测准确率。HumanFi 可以识别 24 个人, 准确率为 96%。GateWay^[21] 提取步态周期, 估计每一步的速度, 然后将其用作人的唯一标识符。Gate-ID^[22] 提出了一种新颖的启发式算法来识别用户的步行方向, 然后使用基于注意力的深度学习模型来识别用户。Gate-ID 对 6~20 人的识别准确率可达 90.7%~75.7%。在步态识别中, 如果新用户模仿合法授权用户的步态频率, 会混淆检测结果, 进而破坏系统的安全性。呼吸作为人体特有的生物特性, 与使用者的身高和体重有关。因此, 基于呼吸感知的人员识别是一种更加可靠的技术。

(2) 基于 WiFi 的生命体征监测

生命体征 (呼吸和心率等) 和生物统计数据是评估一个人的睡眠质量、压力水平和健康状况的重要指标。传统基于传感器的方法可以准确地跟踪生命体征。然而, 这些方法要么需要在良好的照明条件下工作, 要么需要复杂的安装和维护工作。不同的是, 基于射频的方法由于其低成本、无接触、易于部署的特性而变得更具吸引力。以下将主要回顾基于 RSSI 和基于 CSI 的生命体征监测研究。

许多研究观察到, 与呼吸和心跳相关的微小运动也会影响无线信道, 从而导致 RSSI 读数波动。受这种观察的启发, Kaltiokallio 等人^[23] 基于 IEEE 802.15.4 无线传感网络中的 16 个频道来测量 RSSI, 以检测用户的呼吸频率。考虑其他身体运动的干扰, Patwari 等人^[24] 定义“断点”来标识由于用户的运动干扰而引起的 RSSI 信号的突然变化, 并应用适当的均值去除手段, 以确保更加鲁棒的呼吸频率估计。然而上述方法通常需要额外部署无线网络设备, 高密度放置传感节点。BreathTaking^[25] 将呼吸信号建模为正弦波形, 并根据在大约 20 个无线链路上收集到的 RSSI 测量值, 应用最大似然估计 (Maximum Likelihood Estimator, MLE) 来估计呼吸速率。此外, UbiBreath^[26] 可以准确估计用户的呼吸频率, 误差小于 1bpm, 并且检测呼吸暂停的准确率超过 96%。

当一个人呼吸时, 他的胸腔会在很短的距离内移动。这种移动改变了 WiFi 信号的传播路径。基于这一原理, 研究人员开始利用无线信号的 CSI 来感知人体呼吸。Wi-Sleep^[27] 首先开始使用无线信号进行呼吸感知。它通过从 CSI 中提取与呼吸相关的节奏模式来检测人们的睡眠。在后续研究^[28] 中, 提出了一种新的节奏估计算法, 可以实现不同睡眠姿势的呼吸频率估计。TVS^[29] 利用时域和频域的 CSI 信息来准确感知睡眠者的呼吸频率和心率。由于 CSI 相位信息的敏感度很高, 部分工作利用相位信息来实现精确的呼吸感知。TensorBeat^[30] 使用 CSI 的相位差信息, 结合 CP 分解和峰值检测方法来计算多人的呼吸频率。估计单人呼吸频率的最大误差小于 0.9bpm。PhaseBeat^[31] 改进了 TensorBeat, 并使用离散小波变换 (Discrete Wavelet

Transformation, DWT) 分析相位差信息来估计人的呼吸频率和心率。在最新工作中, PhaseBeat^[32]在 Monitor 模式下收集数据, 呼吸频率的中值误差为 0.25bpm。TR-Breath^[33]使用 Root-MUSIC 算法实现可支持 9 人的高精度呼吸频率检测。Wang 等人^[34]设计了一种自适应子载波组合算法来增强呼吸信号的信噪比 (Signal-to-Noise Ratio, SNR)。最后, 他们结合迭代动态规划和轨迹连接算法来跟踪多人的呼吸路径, 人群计数准确率可达 97.9%。许多相关工作侧重于建立无线通信过程的理论模型以实现人员呼吸感知。Wang 等人在^[35]中详细描述了 WiFi 信号中的菲涅尔域模型, 建立了人体呼吸与无线信号变化的理论模型。Yang 等人^[36]结合菲涅尔域模型, 将人体睡眠时的呼吸运动建模为高斯模型, 将姿势变化转化为垂直位移, 优化收发器部署, 实现多人睡眠呼吸检测, 平均绝对误差为 0.5-1bpm。FullBreathe^[37]利用双天线 CSI 的共轭相乘 (Conjugate Multiplication, CM) 来消除相位偏移, 从而解决了菲涅尔域模型在呼吸感知中的“盲点”问题。FarSense^[38]为 CSI 的理论模型提出了 CSI Ratio 模型, 结合 CSI Ratio 模型的幅度和相位信息, 呼吸检测的范围可以扩展到 8 米左右。针对多人检测问题, MultiSense^[39]依据盲源信号分离 (Blind Signal Separation, BSS) 模型采用独立成分分析 (Independent Component Analysis, ICA) 方法实现多人呼吸信号的分离。Liu 等人^[40]首先通过对 CSI 进行 EMD 信号分解来实现基于呼吸特征的用户识别, 然后从 CSI 中提取形态特征和模糊小波包变换 (Fuzzy Wavelet Packet Transform, FWPT) 特征。最后, 结合 DNN 分类器实现用户的身份认证。但是这项工作仅适用于原始 CSI 数据, 需要控制用户与计算机终端的距离, 限制了其使用范围。基于 WiFi 的呼吸感知应用主要用于人群计数和睡眠检测。

(3) 基于 WiFi 的活动识别

活动识别是许多生产生活应用的关键支撑技术, 包括人机交互 (Human-Computer Interaction, HCI)、老年护理、生活管理和安全监控等。一般来说, 人类的活动可以分为两大类, 常规活动 (例如日常活动和异常人体运动) 和手势 (例如手或手指的运动)。本节将简述活动识别的相关研究。

无线信号很容易受到与人类活动相关的身体运动的影响, 导致 RSSI 出现特殊的波动模式。Sigg 等人^{[41][42]}利用人体运动引起的 WiFi 信号的 RSSI 波动, 提出了一种无设备人体活动识别系统。具体来说, Sigg 等人^[41]从 RSSI 信号中提取 17 个经验特征 (例如最高信号峰值和信号强度中值等), 并利用 k 最近邻 (k-Nearest Neighbor, kNN) 分类器识别四种常规活动。为了进一步提高基于 RSSI 的系统的环境感知能力, Sigg 等人^[42]通过使用主动或被动系统检测单个个体的静态和动态活动, 并进一步识别这四种常规活动。然而由于 RSSI 的低分辨率和有限的感知能力, 难以实现细粒度的活动识别, 所以最近的研究倾向于利用 CSI 来获得更好的识别性能。Wang 等人^[43]提出了第一个成果 E-eyes, 探索使用细粒度 CSI 来识别日常活动。

E-eyes 试图利用位置和活动特征之间的关系, 开发一个面向位置的活动识别系统, 仅使用一对 WiFi 接入点来区分一组就地活动 (例如做饭、学习、吃饭和睡觉) 和步行活动 (例如从一个房间走到另一个房间)。然而, 上述系统是基于实证研究的, 缺乏对 CSI 测量值与人类活动之间关系的理论解释。因此, Wang 等人^{[44][45]}提出了一种名为 CARM 的人体活动识别系统, 它建立了 CSI-speed 模型和 CSI-activity 模型来量化人体不同部位的运动速度与特定人体活动之间的相关性。所提出的系统可以在经过训练和未经训练的环境中工作, 在这些环境中评估大量的日常活动 (例如步行、跑步、跌倒和拳击)。

(4) 基于 WiFi 的手势识别

早期的手势识别系统主要依靠从现成设备中提取的 RSSI 来识别不同的手势。Sigg^[46]通过手机检测 RSSI 的波动来识别 11 种不同的手势, 但识别准确率低至 51.0%。为了消除 RSSI 的环境影响, Melgarejo 等人^[47]利用定向天线和短距离无线传播特性, 对 25 个手势实现了更高的识别精度, 所提出的手势识别系统已得到成功应用, 例如基于手势的电子轮椅激活和基于手势的汽车信息娱乐系统控制。Abdelnasser 等人^[48]设计了 WiGest, 进一步将 8 种手势的识别准确率提高到 96%。采用小波降噪消除 RSSI 测量值中的环境噪声。此外, WiGest 不需要训练步骤, 并且在非视距场景中运行良好。然而, 由于 RSSI 的粗粒度和对环境变化的高度敏感性, 基于 RSSI 的手势识别系统无法捕获更细粒度的手势 (例如手指手势和击键)。

为了进一步提高识别精度并捕捉更细微的手势运动, 研究人员在手势识别中开始普遍使用细粒度的 CSI。Nandakumar 等人^[49]提出利用 RSSI 和 CSI 信息来识别手臂运动, 并且对于四个手臂手势 (右、左、推、拉) 可以达到 91.0% 的识别准确率。相比之下, WiG^[50]是一个仅依赖 CSI 的手臂手势识别系统。WiG 的识别准确率在视距场景下高达 92%, 在非视距场景下平均识别准确率为 88%。然而, 上述两个系统只有训练和测试阶段在设置 (即位置和方向相同) 一致的情况下才能有效地工作。为了克服这种限制, Virmani 等人^[51]提出 WiAG 来识别用户不同位置和方向的手臂手势, 很大程度上提高了实用性。WiAG 背后的关键思想是通过提出的手势转换功能将训练样本转换为所有可能配置中所有手势的虚拟样本。除了常规的手势识别外, WiDraw^[52]还可以通过 AoA 估计来连续跟踪手的轨迹以实现空中绘图。

Li 等人^[53]提出 WiFinger 系统, 用于识别美国手语的 9 个手指手势, 而不是跟踪整只手, 准确率高达 90.4%。WiFinger 可以在现成的 WiFi 环境中连续输入文本, 以方便人机交互。SignFi^[54]进一步利用 CSI 来识别涉及头部、手臂、手和手指手势的手语。该系统通过使用 CNN 将识别能力扩展到 276 个手势。此外, Z.Tian 等人提出 WiCatch^[55], 通过在时域中基于 CSIM 样本构建虚拟天线阵列来检测双手手势。WiSign^[56]可以基于功率谱密度自动分割原始 CSI 数据, 构造了一个五层的深度置信网络 (Deep Belief Network, DBN) 来自动提取孤立片段的特

征，使用隐马尔可夫模型（Hidden Markov Model, HMM）、高斯混合算法和前向-后向算法来识别 40 个手语动作，准确率达到了 92%。

2.3 本章小结

本章对人体行为感知现有的技术方向作了简短的介绍。目前感知技术主要可以分为基于传感器的感知技术和基于射频的感知技术。目前的传感器种类繁多，技术发展成熟，得到了市场的广泛应用与认可，然而基于传感器的感知技术，对于软硬件设施有着较高要求，且适用范围窄，存在侵犯隐私的可能，这都限制了传感器推广到更多的感知应用场景。而基于射频的感知技术，主要可以分为 RFID、UWB、蓝牙和 WiFi 等，其中基于 WiFi 的感知技术，研究人员早期利用来自 MAC 层的 RSSI，但由于粗粒度的 RSSI 信息不适用于复杂、有着高精度需求的应用场景，后续研究转而利用来自物理层的 CSI，其细粒度的信号特性使其在生命体征监测、活动与手势识别、轨迹定位与追踪和人员识别等方面都有着广泛的应用。下面的两个章节，将详细介绍两个基于 WiFi 的感知技术研究成果：基于信道状态信息呼吸感知的人员识别系统和基于 WiFi 的手语识别系统。

第三章 基于信道状态信息呼吸感知的人员识别系统

人员识别是保证各种智能应用安全使用的基础。这些应用程序通常需要人员识别功能来控制访问权限和提供服务。当前最先进的人员识别技术依赖于人类的生理和行为特征,称为生物识别技术。生物特征识别是指根据个体的人类生物学和行为特征自动识别个体。用于人类识别的生物特征包括指纹、面部、虹膜和语音。由于生物特征是个人固有的,并且独一无二,因此生物特征被广泛用于人员识别。此外,由于生物特征伪造的难度较高,基于生物特征的技术能够应对日益增长的安全威胁,并在个性化和便利性方面比密码等传统安全方法具有明显的优势。然而这些传统的人员识别方法通常对用户有特殊要求。例如,相机可用于面部识别,但收集用户的照片时需要设置特定方向和范围。基于相机的人员识别方法也会受较差的光线条件限制,或遭到来自 AI 的攻击^[57]。另一方面,日常生活中的 WiFi 信号无处不在,通过提取人类造成的信号变化与其信息之间的相关性,利用 WiFi 信号进行面向人的感知是一种有前景的传感技术。这种无设备解决方案利用广泛存在的 WiFi 基础设施,为用户提供非侵入式识别,显著改善了用户体验,扩大了应用范围。WiFi 感知技术的兴起为人员识别提供了新的解决方案。

本文研究呼吸这种独特的人类生物学特征。人的呼吸通常由人的身高、体重和肺的结构决定。人体的呼吸会导致胸腔移动(5~12mm),从而改变 WiFi 信号在空间中的传播路径。通过分析 WiFi 传播路径变化,可以实现对人员呼吸的细粒度感知。在通过感知呼吸进行人体识别的场景中,用户呼吸时的胸部的移动速度和距离不同,对 WiFi 传播路径的影响也不同。接收端分析 WiFi 信号变化,然后在用户数据库中查找匹配记录,实现对人的识别,这对许多面向用户的应用程序很有帮助。所以本文在人员识别和生命体征监测的研究基础上,提出了基于信道状态信息呼吸感知的人员识别系统。

3.1 系统概述

本系统在基于 WiFi 的生命体征监测和人员识别的研究基础上提出,包含数据采集、数据预处理、呼吸速率检测和人员识别这四个主要模块,系统流程如图 3.1 所示。

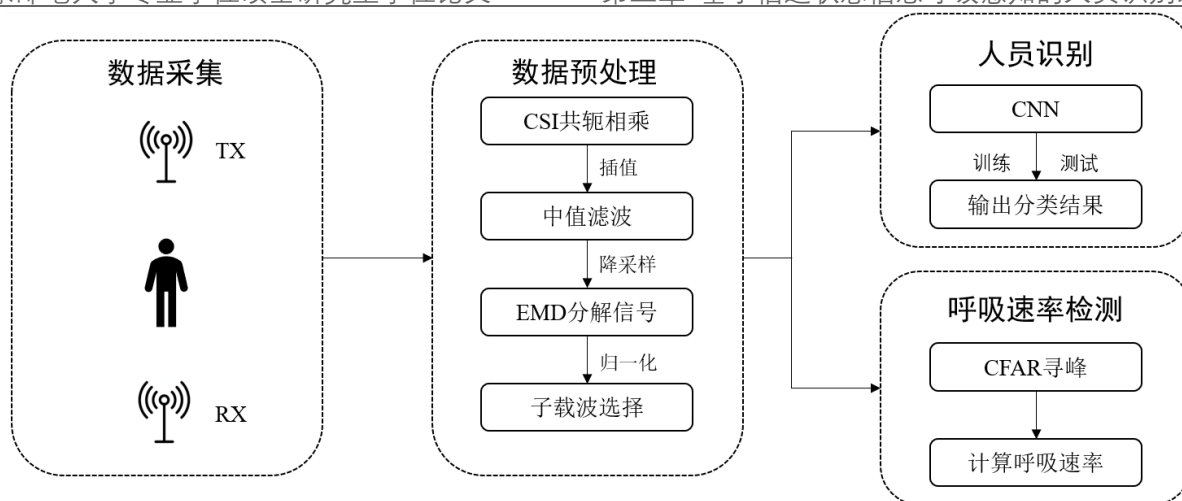


图 3.1 人员识别系统流程图

本章提出的基于信道状态信息呼吸感知的人员识别系统的主要贡献如下：

(1) 从原始 CSI 数据中分离出与呼吸相关的动态分量，会受到环境噪声的干扰。目前提取呼吸成分的理论模型主要有 TVS^[29]提出的时频分析、TensorBeat^[30]使用的相位差、FarSense^[38]提出的相位比和 FullBreath^[37]提出的共轭相乘。本文在采用 CSI 共轭相乘理论模型的基础上，结合 MIMO 技术，能够消除部分相位偏移，增强数据的可用性。

(2) 为了从 CSI 数据中分离出具有高周期性的呼吸曲线，本文采用了 EMD 算法，而先前的 PhaseBeat^[32]对 CSI 相位差使用离散小波变换来提取低频分量。这两种方法都可以用于重建呼吸信号曲线，相较于 DWT 方法，EMD 能够在没有预选基函数的情况下分解原始信号，只依赖于信号本身。此外本文还采用了基于 FFT 的子载波选择策略，筛选周期波动明显的子载波，提高呼吸数据的可靠性。

(3) 为了精确计算用户的呼吸频率，需要对筛选出的呼吸曲线进行寻峰操作。本文采用了 CFAR 算法寻找峰值，与传统的寻峰算法相比，CFAR 寻峰算法能较好的分辨假峰和重峰，在保证检测概率的前提下，能实时改变门限阈值。基于 CSI 共轭相乘的幅度和相位信息，结合基于 FFT 的子载波选择策略和 CFAR 寻峰算法，消除了呼吸感知中的盲点区域，有效提高了呼吸频率的检测精度。

(4) 在人员检测方面，本文利用滑动窗口算法，将采集到的用户的 CSI 呼吸序列分割，生成用于分类任务的数据集。然后将数据集送入卷积神经网络中进行分类。实验结果表明，对于 3 个场景中的 11 名用户的识别准确率都达到了 85%以上。

本文接下来的部分安排如下。首先在第二节中介绍实验的理论模型，然后在第三节中介绍人员识别系统数据预处理的具体流程，在第四节中介绍呼吸速率检测的算法细节，在第五节中介绍如何实现人员识别，第六节给出人员识别系统的实验布置和表现评估，最后在第七节中对本章进行总结。

3.2 理论模型

本节将介绍 CSI 共轭相乘模型和基于菲涅尔域的人员呼吸模型。

3.2.1 CSI 共轭相乘

根据 WiDFS^[58]的研究, 在 WiFi 收发设备工作的过程中, 由于 NIC 处理信号的缺陷和硬件的多样性, 会导致相位的偏移, 与这两种问题相关的信道频率响应 (Channel Frequency Response, CFR) 可以记为:

$$\begin{cases} H_{n,k}^e = \rho_k^{agc} e^{-j(\varphi_{n,k}^{FO} + \varphi_{n,k}^{TO})} \\ H_m^h = \rho_m^h e^{-j\varphi_m^h} \end{cases} \quad (3.1)$$

其中 $H_{n,k}^e$ 是第 n 个子载波, 第 k 个采样点上的 CFR, H_m^h 是第 m 个接收天线的 CFR。 ρ_k^{agc} 表示接收端自动增益控制器 (Automatic Gain Controller, AGC) 设置的增益, 它随不同的信道特性动态变化。 $\varphi_{n,k}^{FO}$ 和 $\varphi_{n,k}^{TO}$ 分别是与频率偏移和时间偏移有关的相位偏移。 ρ_m^h 和 φ_m^h 是第 i 个接收天线处与 WiFi 硬件相关的衰减和相移。

那么第 m 个接收天线, 第 n 个子载波, 第 k 个采样点的 CSI 数据可以表示为:

$$\begin{aligned} CSI_{m,n,k} &= H_m^h H_{n,k}^e \sum_{l=1}^L H_{m,n,k}(l) \\ &= H_m^h H_{n,k}^e (H_{m,n}^S + H_{m,n,k}^D) \end{aligned} \quad (3.2)$$

其中 L 是信号传播路径的总数, $H_{m,n,k}(l)$ 代表第 l 条传播路径的 CFR, $H_{m,n}^S$ 和 $H_{m,n,k}^D$ 分别代表 CSI 的静态分量与动态分量。

以接收天线 1 和接受天线 2 为例, 它们之间的 CSI 共轭相乘可以用公式表示为:

$$\begin{aligned} CSI_{1,n,k} \overline{CSI}_{2,n,k} &= (H_{n,k}^e H_1^h) (H_{1,n}^S + H_{1,n,k}^D) (\overline{H}_{n,k}^e \overline{H}_2^h) (\overline{H}_{2,n}^S + \overline{H}_{2,n,k}^D) \\ &= \|H_{n,k}^e\|^2 H_1^h \overline{H}_2^h H_{1,n}^S \overline{H}_{2,n}^S + \|H_{n,k}^e\|^2 H_1^h \overline{H}_2^h (H_{1,n}^S \overline{H}_{2,n,k}^D + H_{1,n,k}^D \overline{H}_{2,n}^S + H_{1,n,k}^D \overline{H}_{2,n,k}^D) \end{aligned} \quad (3.3)$$

其中 $\|\cdot\|$ 代表计算幅度, 而且 $\|H_{n,k}^e\| = \rho_k^{agc}$ 。由上式可知, 共轭相乘后, $H_{n,k}^e$ 中的 $\varphi_{n,k}^{FO}$ 和 $\varphi_{n,k}^{TO}$ 被消除。所以可以得出结论, CSI 共轭相乘起到了一定消除相位偏移的作用。

3.2.2 菲涅尔域

在 WiFi 信号的传播模型中, 菲涅尔域是以收发装置为焦点的一系列同心椭圆。第 n 个菲涅尔域的定义为: 从发射端到菲涅尔域椭圆边界的任何一点, 再到接收端的反射路径, 减去直射路径, 等于 $\frac{n\lambda}{2}$, 其中 λ 为波长。当物体越过菲涅尔域的边界时, 接收到的信号会出现波峰或波谷, 但是如果物体沿椭圆移动, 则反射信号路径的长度不会改变, 因此接收到的信号保持不变。如图 3.2 所示, 用户站在视距路径上, 方向垂直于视距路径。当用户呼吸时, 其对应的胸部位移距离在 5mm~12mm 之间, 跨越了菲涅尔域的边界, 所以接收到的信号会产生相应的峰值。

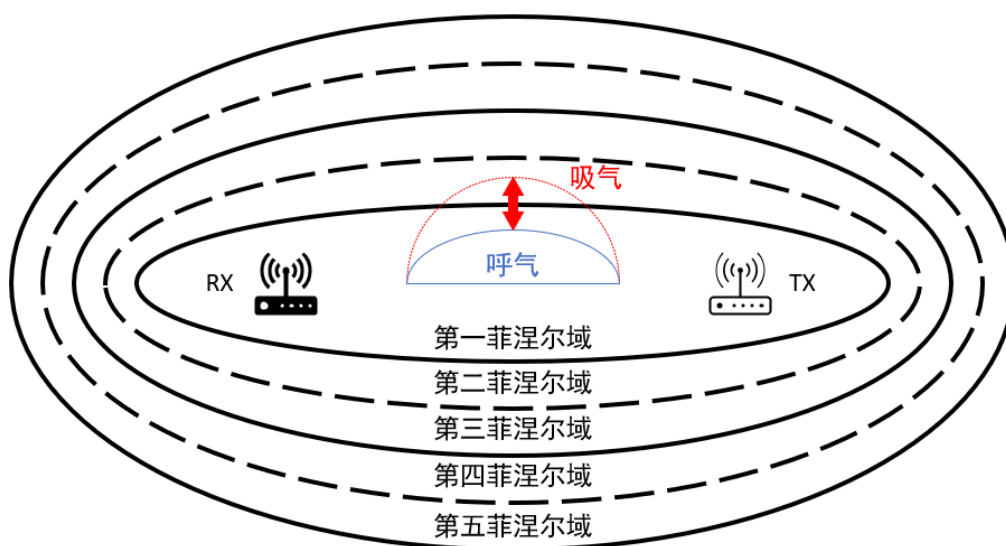


图 3.2 基于菲涅尔域的人员呼吸模型

3.3 数据预处理

在人员识别系统中, 接收端使用 Halperin 开发的 CSI tool^[59]来收集 CSI 数据。本系统的接收机和发射机终端分别配备 3 个全向天线。所有终端的无线网络接口控制器都是 Intel 5300 网卡, 可以提供 30 个子载波的 CSI 数据。人员识别系统的采样率为 200Hz, 由于人的呼吸频率通常远低于 1Hz, 因此该采样率足以用于人的呼吸感知。为了保证通信数据的可控性, 系统发送和接收数据时使用 Monitor 模式来代替传统的 AP 模式。

通过 Monitor 模式采集的 CSI 数据可能会丢包, 丢包率约为 0.1%-0.05%, 会影响原始数据的准确性。为了使采集的数据更加准确, 首先对原始数据进行线性插值, 插值长度为采样率*时间长度。然后使用中值滤波器滤除 CSI 数据中的部分高频噪声与脉冲噪声。由于人的呼吸频率小于 1Hz (人的呼吸频率通常在 10bpm 到 37bpm 之间), 为了减少计算数据量, 本文

对去噪后的数据进行了降采样，新的采样频率为 20Hz。

经过数据预处理后，将原始数据处理成频率为 20Hz 的低频数据。然而，人体呼吸频率通常为 10bpm~37bpm，相应的信号频率为 0.1667Hz~0.6167Hz。因此，有必要进一步从处理后的数据中提取低频信息。目前常见的信号分析手段主要有小波分解（Wavelet Decomposition, WD）和 EMD，本系统选择 EMD 算法，进一步分离 CSI 的低频分量。

最后由于不同子载波对于呼吸的感知粒度不同，本文利用快速傅里叶变换设定选择指标，对子载波进行筛选，增强 CSI 呼吸数据的可靠性。

3.3.1 中值滤波

中值滤波是一种常见的去噪方法，广泛用于涉及语音、信号和图像处理的应用中。标准中值滤波器是一种简单的非线性平滑滤波器，能够有效替代线性滤波，因为它可以抑制噪声，同时保持信号值的边缘信息，在降低脉冲型噪声方面特别有效。标准中值滤波器在某点的输出是以该点为中心的滑动窗口内所有输入数据的中值，假设 $\{m(k)|1 \leq k \leq L\}$ 和 $\{n(k)|1 \leq k \leq L\}$ 分别是标准中值滤波器的输入与输出，滑动窗口长度为 $2N + 1$ ，其中 L 为输入序列的长度，那么输出 $n(k)$ 可以用如下公式计算：

$$n(k) = \text{med}\{m(k - N), \dots, m(k - 1), m(k), m(k + 1), \dots, m(k + N)\} \quad (3.4)$$

考虑到序列开始和结束时的处理， $m(1)$ 和 $m(L)$ 分别在输入的开头和结尾重复 N 次，这是因为序列开始前和结尾后缺少元素，不足以填充窗口。这就会带来中值滤波的边界问题，一般应对方法是避免处理边界，即裁切信号边界，或缩小信号序列边界处的滑动窗口的大小。图 3.3 展示了中值滤波前后某条子载波序列的变化。

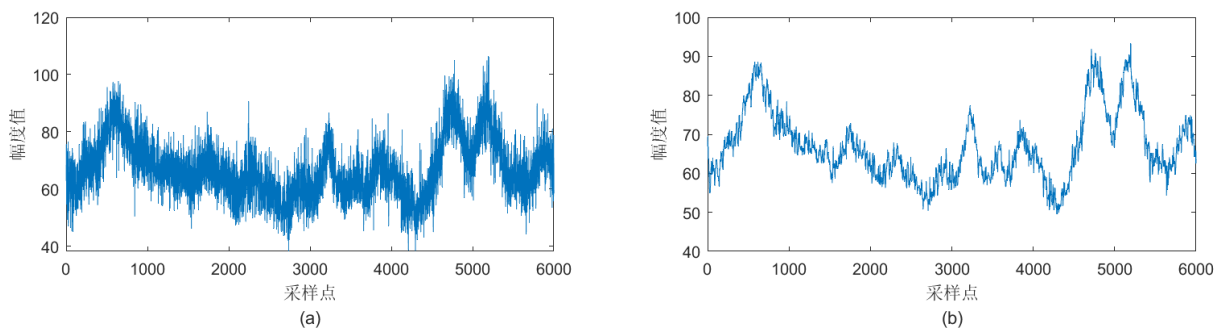


图 3.3 中值滤波效果图：（a）中值滤波前的 CSI 序列；（b）中值滤波后的 CSI 序列

3.3.2 EMD 算法

EMD 是一种基于时域进行信号分解的手段，与傅里叶变换和小波分解等其他分析方法类

似, 适用于分解非线性和非平稳的信号序列。与传统小波分解相比, EMD 没有基函数并且只依赖于原始信号本身。EMD 的整个分解过程是基于信号的最大最小值、局部平均值和时间尺度特征自动进行的, 将信号分解为从高频到低频的一系列本征模函数(Intrinsic Mode Function, IMF), 每个模函数都包含有关原始信号频率如何随时间变化的信息^[60]。

IMF 必须满足以下两个条件: 首先整个数据序列之内, 极值点数与零点数之间的差值不能超过 1。其次由局部最大值和局部最小值定义的上下包络线, 任何时候其均值必须为 0。

EMD 将目标信号分解成 IMF 的过程被称为筛选过程, 其具体步骤如下:

(a) 获取信号 $x(t)$ 的所有局部最大值点和局部最小值点。

(b) 通过使用三次样条函数(CSF)插值, 连接所有局部最大值得到上包络线 $e_1(t)$, 连接所有局部最小值得到下包络线 $e_2(t)$ 。

(c) 提取 $e_1(t)$ 和 $e_2(t)$ 的均值: $m(t) = [(e_1(t) + e_2(t))/2]$ 。从原始数据序列中减去平均信号得到 $h_1(t) = x(t) - m(t)$ 。若 $h_1(t)$ 满足 IMF 条件, 则 $h_1(t)$ 为信号的第一个 IMF, 否则将 $h_1(t)$ 将为新的输入信号, 计算 $h_1(t)$ 上下包络线均值 $m_1(t)$, 重复该过程 k 次, 直到 $h_k(t) = [h_{k-1}(t) - m_{k-1}(t)]$ 满足 IMF 条件, 记为 $c_1(t)$, 也就是 IMF1。

(d) $r(t) = x(t) - c_1(t)$, 作为新的输入信号 $x(t)$, 然后重复步骤(a)~(d), 找到 IMF2, 记为 $c_2(t)$, $r_2(t) = r_1(t) - c_1(t)$ 。

当 $r_n(t)$ 变得足够小或无法从 IMF 分量中提取单调函数时, EMD 的筛选过程停止。初始的 $x(t)$ 可以表示为有限数量的 IMF 和最终残差的总和, 如下方的公式所示:

$$x(t) = \sum_{n=1}^N IMF_n + r_n(t) \quad (3.5)$$

由于 EMD 分解出的 IMF 依次从高频到低频排列, 所以本系统选择去除最开始的两个 IMF, 将剩下的 IMF 相加来重构呼吸信号。重构的呼吸信号周期性更强, 信号波形更加平滑, 一定程度上也减少了高频噪声的影响。EMD 算法分解和重构信号的示例如图 3.4 所示。

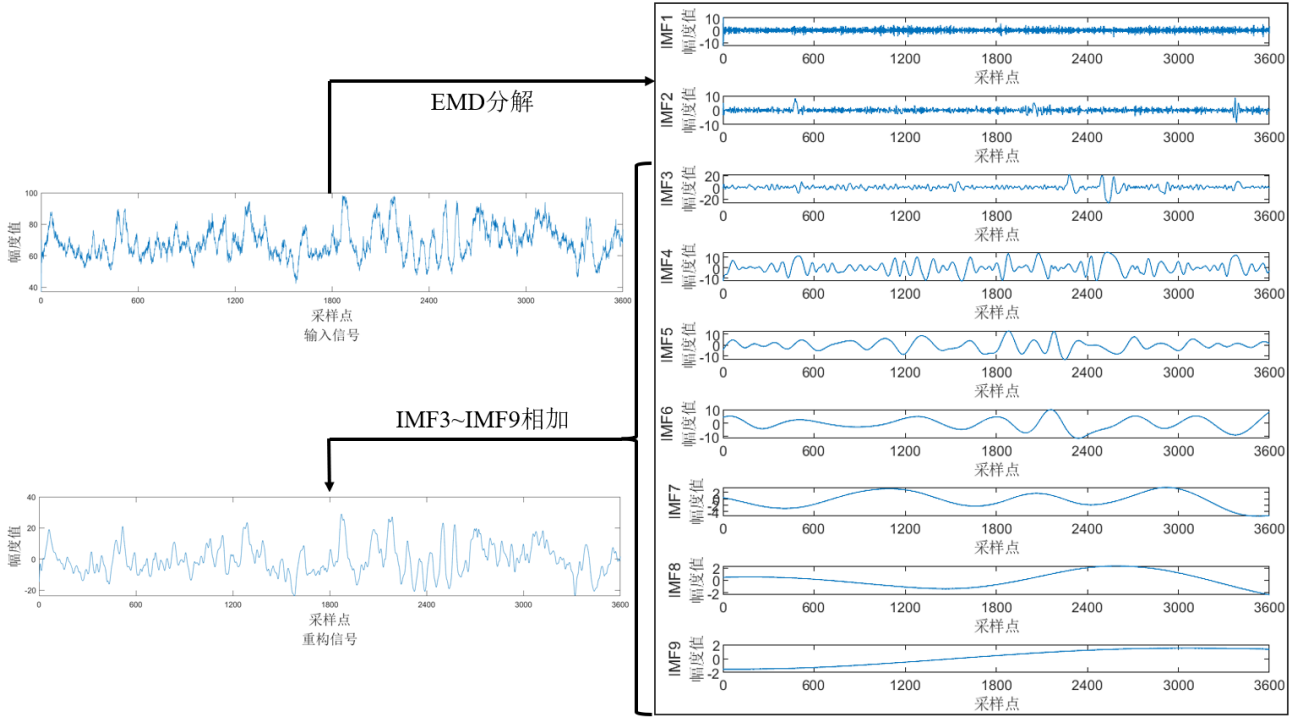


图 3.4 EMD 算法分解和重构信号

3.3.3 子载波选择

FFT 是一种计算序列的离散傅里叶变换 (Discrete Fourier Transform, DFT) 或其逆变换的算法。它能将信号从初始的时域转换为频域表示。在信号处理中, FFT 构成了频域分析 (频谱分析) 的基础, 常用于信号滤波、频谱估计、数据压缩等应用。

由于子载波的反射路径不同, 因此其对于呼吸的感知粒度也不同。所以本章的人员识别系统对子载波进行快速傅里叶变换, 设计了基于呼吸频率范围的频谱指标, 基于此指标选择具有与呼吸信号高度相关的子载波, 用于进行更准确的呼吸速率检测和人员识别。

首先对 30 个子载波组成的矩阵 $A = (A_1, A_2, \dots, A_{30})^T$ 按列进行快速傅里叶变换, 得到一个复数矩阵 $C_{complex}$, 然后对该矩阵作求模运算, 那么就可以得到由 30 个子载波频谱信息组成的实数矩阵 C_{fft} , 计算过程如以下公式所示:

$$\begin{cases} C_{complex} = FFT(A) \\ C_{fft} = abs(C_{complex}) \end{cases} \quad (3.6)$$

人正常情况下的呼吸频率范围为 $[0.1667\text{Hz}, 0.6167\text{Hz}]$, 系统在这里选取的频率范围则是 $[0.15\text{Hz}, 0.5\text{Hz}]$, 足以满足要求。在这里需要先确定 C_{fft} 中子载波序列的第 n 个数, 与频率 f 、采样频率 f_s 和序列总长 N 之间的关系, 公式如下:

$$f = \frac{n * f_s}{N} \quad (3.7)$$

假设呼吸最低频率 0.15Hz 和最高频率 0.5Hz 分别对应序列号 n_1 和 n_2 ，所以第 i 个子载波基于 FFT 的筛选指标的计算公式如下：

$$a(i) = \frac{\sum_{n=n_1}^{n_2} C_{fft}(n, i)}{\sum_{n=1}^N C_{fft}(n, i)} \quad (3.8)$$

最终从筛选指标中选择值最大的前 10 个子载波，作为结果输出。图 3.5 展示了一个基于 FFT 的子载波选择的示例。

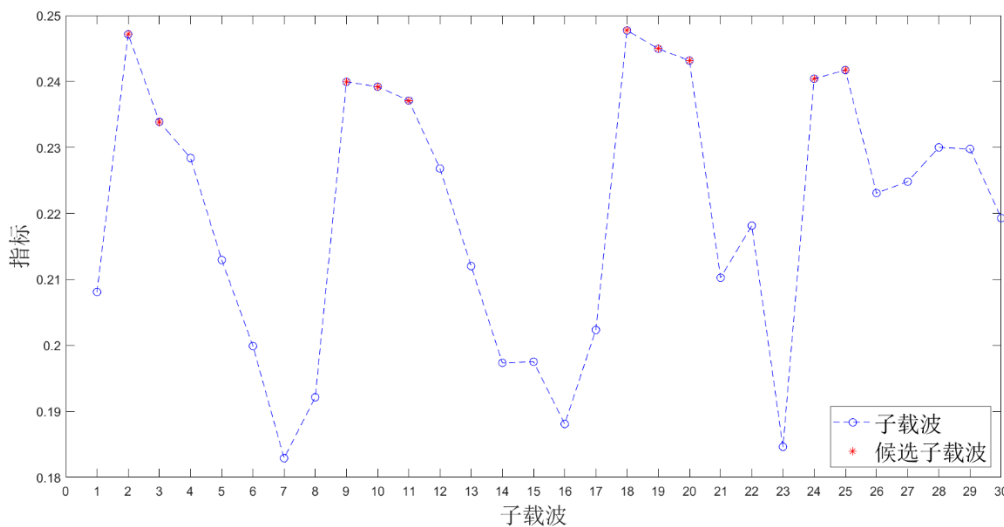


图 3.5 基于 FFT 的子载波选择

3.4 呼吸速率检测

呼吸速率检测的核心步骤，主要通过对筛选出的子载波进行寻峰操作来实现。一般可以将两个相邻峰值之间的数据片段视为一个完整的呼吸片段，所以寻峰算法寻找的峰值点是否准确，将会很大影响呼吸速率的检测精度。目前主流的寻峰算法主要有简单比较法、导数法、CFAR 寻峰算法、对称零面积法和线性拟合寻峰法等。

简单比较法对平滑后的曲线进行基于最大值的比较，只适合寻找强单峰，而且很容易受噪声影响。导数法是对平滑后的曲线，进行一阶、二阶和三阶求导，根据求取数据的均值设置门限值，再进行峰值筛选，缺点是对于弱峰和强峰的分辨效果都不太理想。对称零面积法选择对称零面积简单函数作为变化函数，然后按照对称零面积设置窗函数，计算峰的净面积和总面积的标准偏差，通过阈值判断该峰是否为真。线性拟合寻峰法根据最小二乘法得出线性函数系数，然后根据拟合函数得出拟合曲线，最后计算相应峰的高度和宽度。对称零面积法和线性拟合寻峰法缺点都是函数计算复杂，算法计算量大。

所以本章的人员识别系统采用了基于 CFAR 的寻峰算法，CFAR 是雷达系统中使用的一种自适应算法，用于在噪声和杂波背景下检测目标回波，超过其阈值的任何返回信号或回波都可以被认为来自目标^[61]。

3.4.1 CFAR 寻峰算法

CFAR 算法寻找峰值的流程大致分为三个步骤：首先在数据段的左右两侧设置两个移动平均函数，对训练单元内的数据进行平滑处理，获得优化后的曲线，然后根据选定指标求取数据的平均值，并根据平均值设立门限阈值，最后按照阈值寻找数据段测试单元中的峰值。

图 3.6 展示了数据段的结构，根据测试单元中的值是否大于阈值，来决定其是否被检测。而阈值水平根据训练单元的估计噪声功率来计算。紧邻测试单元的是两个保护单元，它们的作用主要是降低测试单元中数据对训练单元的影响，其数量由用户定义。

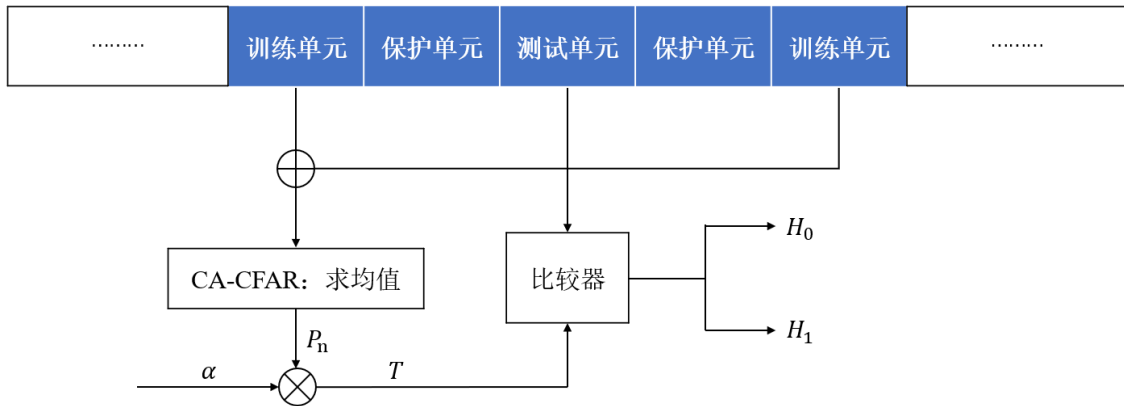


图 3.6 CFAR 算法的工作流程

假设 x_j 是测试单元中的第 j 个值，如果 x_j 大于测试单元内的其他值，并且 $x_j > T$ ，那么 x_j 可以视为测试单元内的唯一峰值。这里的 T 为阈值，其计算公式如下：

$$T = \alpha P_n \quad (3.9)$$

$$\alpha = N \left(P_{fa}^{-\frac{1}{N}} - 1 \right) \quad (3.10)$$

$$P_n = \frac{1}{N} \sum_{m=1}^N x_m \quad (3.11)$$

其中 α 是预定义的阈值因子， P_n 代表着训练单元的估计噪声功率水平， x_m 是训练单元中的元素， N 为训练单元中元素的个数， P_{fa} 则代表误报率。由上述公式可以看出，这是一个典型的均值类 CFAR 模型。

总的来说，CFAR 算法能较好地分辨假峰和重峰，对杂波的抑制能力较强，并且能够在保

证检测概率的前提下，实时的改变门限阈值。图 3.7 展示了 CFAR 寻峰算法在一段 CSI 呼吸数据上的寻峰效果。

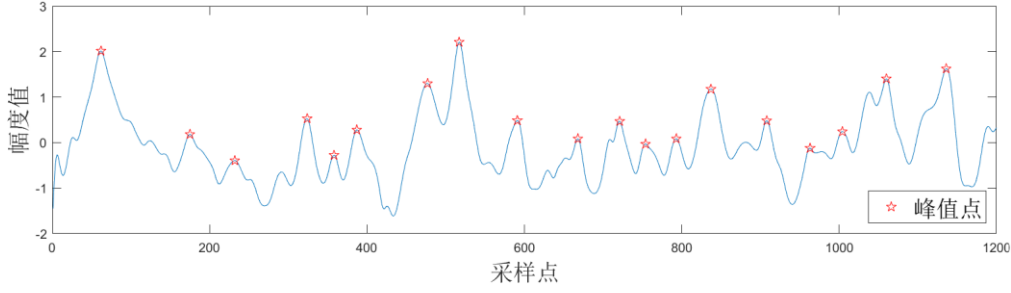


图 3.7 CFAR 算法寻峰效果图

3.4.2 呼吸速率计算

为了精确计算用户的呼吸速率，减少计算误差，增强检测的鲁棒性，本系统基于 FFT 的选择策略，进行子载波筛选，结合子载波的幅度和相位信息来进行呼吸速率检测，根据 FullBreath^[37]的推导，这种做法能够消除 WiFi 呼吸感知中的“盲点”问题。CFAR 寻峰算法将会输出每条子载波幅度和相位曲线的峰值点坐标的集合，单条呼吸曲线的呼吸速率计算公式如下：

$$BPM = \frac{(k-1) * resample_{rate}}{loc_2 - loc_1} \quad (3.12)$$

其中 loc_1 是第一个峰值对应的位置， loc_2 是最后一个峰值对应的位置， k 是这条呼吸曲线寻峰得出的峰值总数， $resample_{rate}$ 是数据重采样率。为了提高检测结果的准确性，本系统采用了 MIMO 多天线技术，收发设备都配有 3 根天线，每个天线对选择 10 个子载波来提取呼吸频率，所以最终的呼吸频率如下：

$$\overline{BPM} = \frac{1}{N_T * N_R * 2K} \sum_{t=1}^{N_T} \sum_{r=1}^{N_R} \sum_{k=1}^K (BPM_A + BPM_P) \quad (3.13)$$

其中， K 是子载波选择后的子载波个数，其值为 10，而 N_T 和 N_R 分别是发射天线和接收天线的数目。 BPM_A 和 BPM_P 则是根据每条子载波的幅度曲线和相位曲线分别求出的呼吸频率。

3.5 人员识别

人员识别必须先保留用户的呼吸样本作为数据库的基础，然后在识别时重新收集新的呼吸样本以进行匹配。本文将使用滑动窗口算法将采集到的 CSI 序列切割成大小一致的样本，并将其送入卷积神经网络中进行分类。

3.5.1 数据集生成

由于实验采集数据时，每个人采集到的是一段长达 3 分钟的呼吸数据，所以每个人对应的 CSI 序列是一个列数为 $N_{resample} * t$ ，行数为 $N_T * N_R * K$ 的矩阵，其中 $N_{resample}$ 是重采样率， t 是采样时长， N_T 和 N_R 分别是发送和接收天线的数目， K 为子载波选择的数目。因此本文采用滑动窗口算法分割 CSI 序列，窗口大小设置为 200，步长设置为 40。这样每个人的 CSI 序列将产生 86 个样本。

3.5.2 分类模型

在使用滑动窗口算法生成数据集后，将其送入卷积神经网络中进行分类，模型的简化结构如图 3.8 所示，包含了卷积层，平均池化层，dropout 层，全连接层和分类输出层。

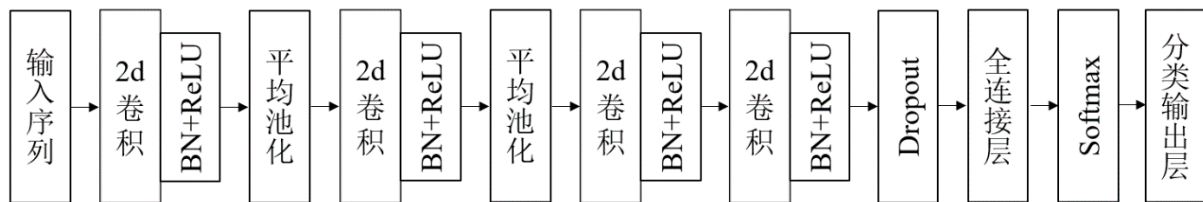


图 3.8 CNN 模型的架构图

卷积层：卷积层可以提取输入数据的局部或全局特征，从而获得对输入数据的有效描述。每个卷积层后都紧接着一个批处理归一化层（Batch Normalization，BN），使用线性整流函数（ReLU）作为激活函数。

平均池化层：池化层通过总结特征的存在来对特征进行下采样，降低计算复杂度。常见池化方法有平均池化和最大池化，分别总结平均特征和活跃特征的存在。

Dropout 层：在训练期间从神经网络中随机忽略单元，防止单元过度共同适应，能有效降低过拟合的概率。

全连接层：作用是把之前提取的特征映射到样本空间中用作分类。

Softmax 函数：返回每个类的概率，输出总和为 1 的概率分布。

3.6 实验与评估

本节主要介绍实验的设备配置和场景安排的情况，并给出对基于 CSI 呼吸感知的人员识别系统的表现评估：包括人员识别的准确率，呼吸速率的计算误差，以及不同因素对于人员

识别的影响。

3.6.1 实验设置

在人员识别系统中,采用 170mm*170mm 的 mini PC 终端和 ThinkPad 笔记本作为收发设备,均搭载 Intel 5300 无线网卡。为获取高精度数据,数据采集设备采用 Monitor 模式进行无线通信,此设备也可以工作在 AP 模式。数据采集设备的工作频率为 5.8GHz,发包速率为 200Hz。收发终端均配备 3 根全向天线,天线间距 3cm (小于无线信号波长 4cm),天线高度 1.2m,设备间距 2m。接收端安装 Linux CSI Tool 来采集数据,使用 MATLAB R2021 作为实现整个人员识别系统的开发软件。

图 3.9 (a) 和 3.9 (b) 显示了实验场景的布局。本文在实验室和会议室这两个不同的房间测试了人员识别系统的性能。实验包含 3 个测试场景,场景 1、场景 2 和场景 3 的实验设置都相同。在每个场景中,一共收集了 11 名不同身高和体重的志愿者(8 名男性和 3 名女性)的呼吸数据,志愿者的年龄在 22~30 岁之间。本文预先收集了志愿者的呼吸数据并将其存储在数据库中。然后在识别用户时生成新的呼吸数据,人员识别系统根据前文的方法识别用户。基于场景 1,本节也讨论了收发设备距离和用户方向对于实验结果的影响。

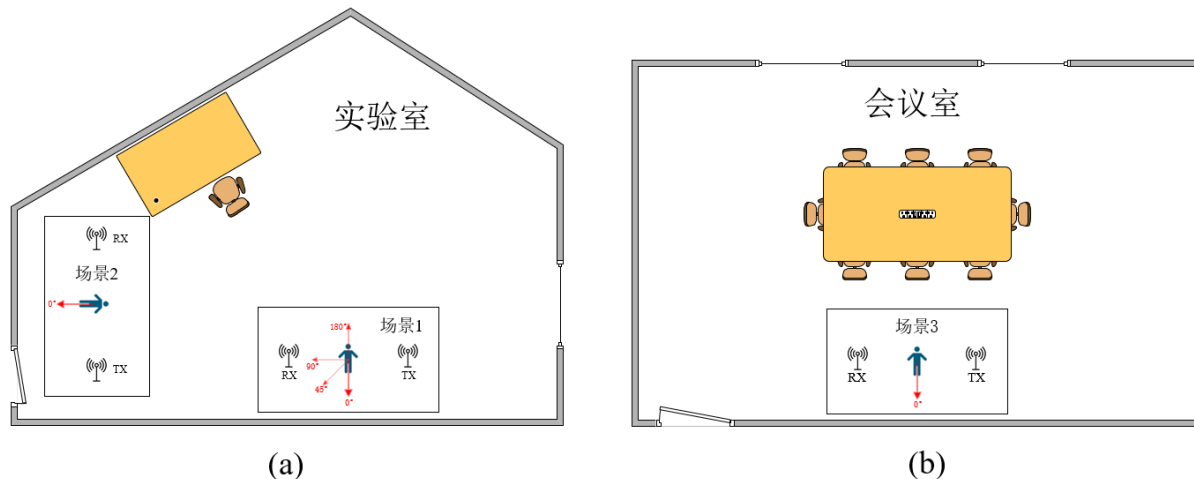


图 3.9 实验场景示意图: (a) 实验室; (b) 会议室

3.6.2 表现评估

呼吸频率的估计误差: 图 3.10 显示了在检测同一呼吸频率时,本系统和其他呼吸检测算法的比较结果。TVS^[29]使用频域信息这种粗粒度特征来估计呼吸频率,其估计呼吸的平均绝对误差约为 1bpm。TensorBeat^[30]使用 CSI 相位差数据来计算细粒度呼吸频率,由于无线感知领域存在“盲点”,仅使用相位信息会造成一定程度的估计误差。为了解决 CSI 感知的“盲点”

南京邮电大学专业学位硕士研究生学位论文 第三章 基于信道状态信息呼吸感知的人员识别系统

问题，FullBreathe^[37]采用 CSI 共轭相乘的方法提取信号特征，而本系统的平均绝对误差约为 0.5bpm，接近 FullBreathe 系统。

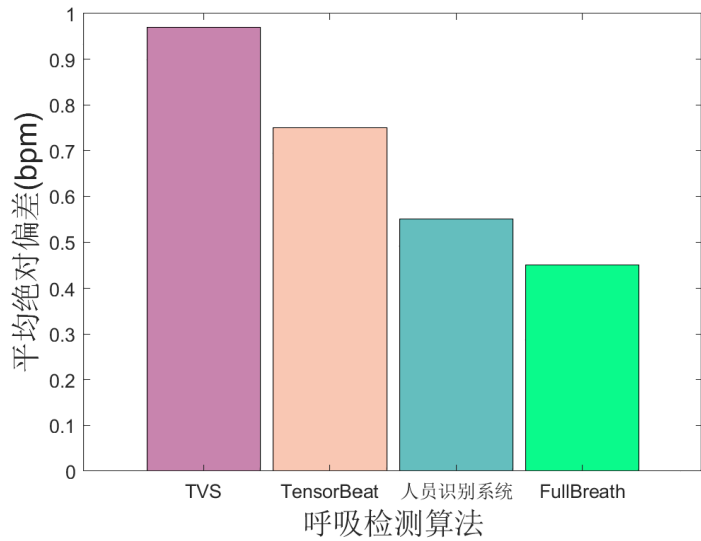


图 3.10 不同呼吸检测算法的表现对比

人员识别准确率：场景 1 作为实验的基准场景，其人员识别准确性的混淆矩阵如图 3.11 所示，可见其总体识别准确率达到了 87.76%，其中第 3、6、7、9、10、11 用户的识别准确率较低，在 70%~90%之间，其余用户的识别准确率则达到了 90%以上。实验结果证明了人员识别系统能够准确的识别用户。

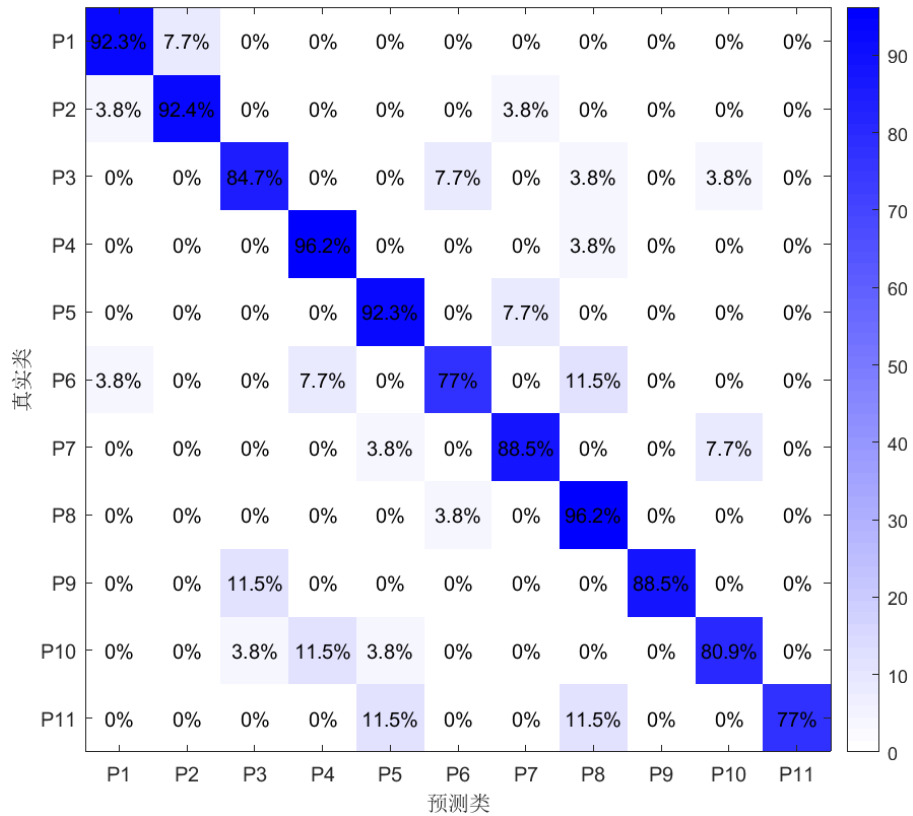


图 3.11 场景 1 人员识别的混淆矩阵

不同场景的影响：本文在实验室和会议室中选择了 3 个场景进行实验，如图 3.9 所示。图 3.12 展示了三个场景中的人员识别准确率。实验结果表明，实验室中的场景 1 和场景 2 的识别准确率都达到了 86%以上，由于会议室狭小的空间中桌椅等影响无线信号传播的物体较多，所以会议室中场景 3 的识别准确率降低到了 85.67%。

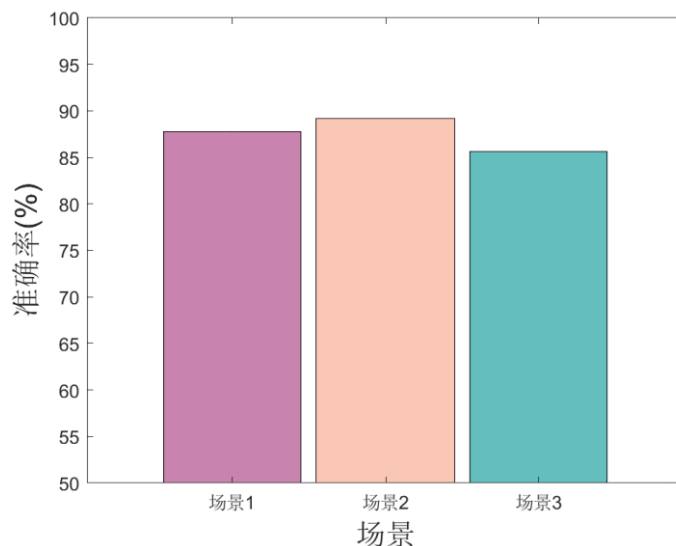


图 3.12 不同场景中的人员识别准确率

收发设备不同间距的影响：在场景 1 中，将发送端和接收端设备的间距分别调整为 1.5m、2m 和 2.5m，并采集 CSI 呼吸数据进行实验，结果如图 3.13 所示。实验结果表明，本系统的人员识别准确率会随收发设备距离的增加而略有降低，但仍保持着 85%以上的准确率。

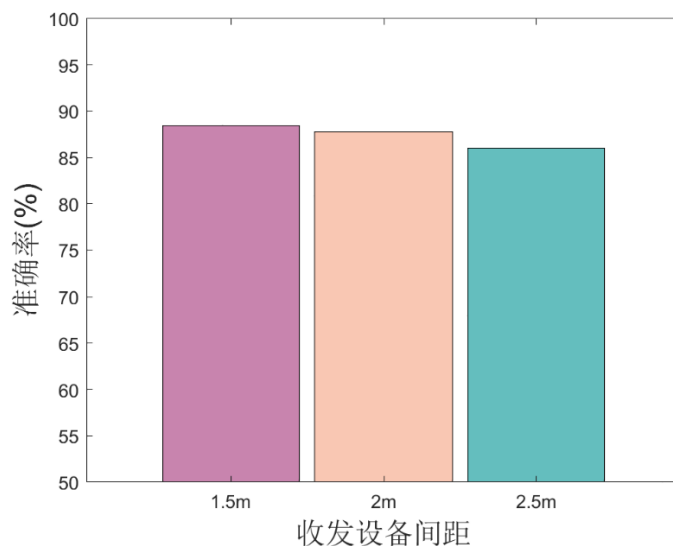


图 3.13 收发设备不同距离时的人员识别准确率

人员不同朝向的影响：在场景 1 中的 4 个人员朝向上进行人员识别。图 3.14 的实验结果显示，90°时的检测结果最差，0°和 180°时识别准确率可达 86%以上。在 0°和 180°时，用户的呼吸方向垂直于菲涅尔域的切线，这会导致反射路径发生显著的变化。在 90°时，用户的呼吸方向与切线方向平行，所以很难检测到呼吸引起的信号传播路径的变化，此时人员识别准

确率跌至 50%以下。

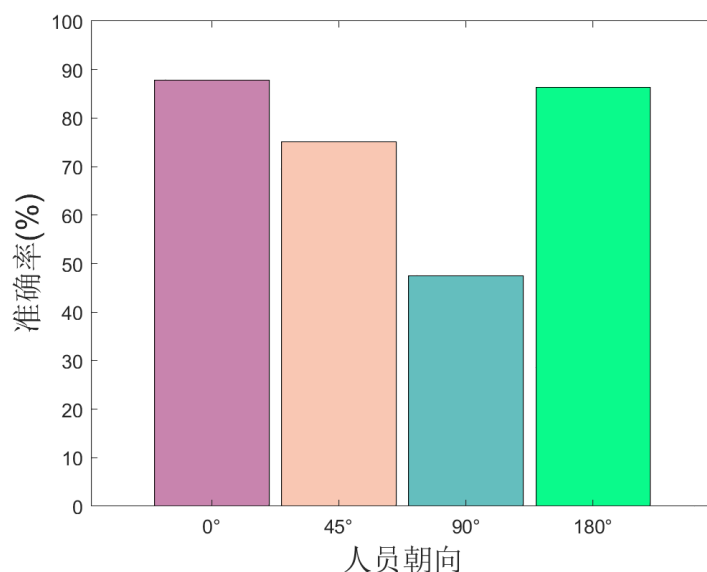


图 3.14 人员不同朝向的识别准确率

不同用户数目的影响：为了探究目标用户的数量对于识别准确率的影响，实验分别基于 5 人、8 人和 11 人的呼吸数据对人员识别系统进行了验证。实验结果图 3.15 表明，随着用户数目的增加，人员识别的精度会逐渐降低，从 90%降低到 87%。

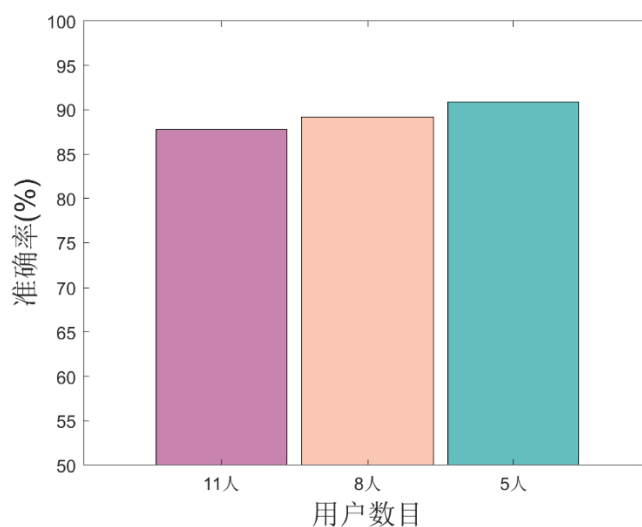


图 3.15 识别不同数目用户时的准确率

3.7 本章总结

本章提出了一种使用商用 WiFi 设备的高精度人员识别系统。它对用户的呼吸进行感知，然后识别用户。本文首先对 CSI 的共轭相乘模型和菲涅尔域模型进行理论分析。然后结合中值滤波、EMD 算法和基于 FFT 的子载波选择策略来分离 CSI 信号中的呼吸分量。基于多天线的 CSI 数据流，使用 CFAR 寻峰算法来检测呼吸速率。通过滑动窗口算法生成数据集，使

用 CNN 完成人员识别的任务。针对不同的因素,本文对人员识别系统进行了多次实验。实验结果表明,本系统检测呼吸速率的平均绝对误差约为 0.5bpm。本系统对 3 个场景中的人员识别准确率都达到了 85%以上,能够进行较高准确率的人员识别。

人员识别系统还有很多需要改进的地方。第一、基于 EMD 的信号分解算法,如何选择 IMF 来重构呼吸信号,可以参考 Liu 等人^[40]提出的互信息的概念,动态确定用于重构信号的 IMF 数量。第二、均值类 CFAR 寻峰算法可以继续改进,例如 CFAR 算法中的自适应类,通过分析训练单元数据选择不同的决策方法。第三、可以进一步改进分类模型,优化分类器的性能。

第四章 基于 WiFi 的手语识别系统

手语是一种使用手部动作、肢体动作和面部表情进行交流的语言，主要由聋哑人和部分听力正常的人使用，包括聋哑人的家属和口译员等。手语是聋哑人最有效的交流手段，它在相同时间输出信息的密度，远高于其他手段。手语有着和口语和书面语一样的复杂性和表达力，但手语不是口语和书面语的另一种表现形式，有着自己独特的语法结构。不同地区的手语是分化的，手语中也有语系，虽然存在一种通用的手语，但其本身并不是一种独立的手语语言，没有固定的语法和词库。种种因素表明，普通人学习手语的成本并不低，然而手语使用者在任何国家都占总人口的 0.1% 以上，因此，为这些群体开发更多交流手段，能够提高他们在工作和生活中的便利性，帮助他们更好的融入社会。所以本章提出了一种基于 WiFi 的手语识别系统，用户在接收设备和发射设备的视距路径上执行手语动作，系统处理收集到的 CSI 数据并识别其对应的手语含义。

从现有的手语输入设备来看,手语识别系统主要分为基于数据手套的手语识别系统和基于视觉的手语识别系统。基于数据手套的手语识别系统，是利用数据手套和位置跟踪器测量手势在空间运动的轨迹和时序信息。这种方法的优点是系统的识别率高。缺点是打手语的人要穿戴复杂的数据手套和位置跟踪器，输入设备成本较高，不利于大规模的推广。基于视觉的手势识别是利用摄像机采集手势信息，并进行识别。该方法的优点是输入设备较为常见，相对于数据手套进一步降低了成本，但识别率比较低，实时性较差，受光照等条件束缚，并且有泄露用户个人信息的危险。而基于 CSI 的基本手语识别技术，相对而言是一种低成本，适用范围广，能降低隐私泄露风险的解决方案，而且这种方案，可以推广到智能交互等更广阔的应用场景中去。

4.1 系统概述

本手语识别系统用于数据采集的收发设备分别为搭载了 Intel 5300 无线网卡的笔记本电脑和 mini PC 机，均为商用设备。从作为国家语言文字规范的《国家通用手语常用词表》中挑选出用于识别的手语。对采集到的原始 CSI 数据预处理后，送入深度学习模型中进行分类识别。本章的主要挑战与贡献如下：

(1) 相较于市面上已有的商用手语识别系统，本系统基于商用设备实现，不需要用户佩戴其他设备，成本低廉，易于部署。

(2) 由于 WiFi 信号传播过程中周围环境的变化，无线信号传播路径也会发生变化，从而导致接收到的 CSI 数据产生波动。因此本文采用 Hampel 滤波器滤除原始数据的离群值和去除环境噪声，并通过一维线性插值和归一化的手段保证数据样本的一致性。

(3) 由于不同子载波对于环境中人体行为变化的敏感度不同，因此有必要对子载波进行筛选。本系统采用基于 DTW 的选择策略筛选有效的子载波。

(4) 本系统识别手语时采用基于 CNN 和 BiLSTM 的深度学习模型，识别准确率优于现有的主流机器学习模型和 CNN 模型，平均识别准确率可达到 97.8%。

基于 WiFi 的手语识别系统的总体结构如图 4.1 所示，通过搭载 Intel 5300 无线网卡的笔记本电脑和 mini PC 机作为收发设备，采集执行手语动作时的原始 CSI 数据，然后将原始数据进行预处理，包括去噪，归一化和子载波筛选，最后将处理后的数据送入 CNN 结合 BiLSTM 的深度学习模型中进行识别。系统执行的流程如下所示：

(1) 数据采集阶段：本手语识别系统的收发设备分别为搭载了 Intel 5300 无线网卡的笔记本电脑和 mini PC 机，各自配备 3 根外接的全向天线，设备固定在实验场地中，当发送端开始向接收端发送数据时，支援者同时开始执行手语动作，当数据发送完毕后，一个手语样本也采集完毕。

(2) 数据预处理阶段：因为用于数据传输的无线链路有时会出现不稳定，导致一些数据样本中的数据包丢失，因此为了保证样本的一致性，在对数据进行降噪处理之前，对数据进行插值。然后采用 Hampel 滤波器滤除原始数据中存在的离群值，对数据进行降噪处理。由于不同子载波对于环境中人体行为变化的敏感度不同，因此有必要对子载波进行筛选，而且子载波选择也在一定程度上起到了数据降维的作用，能够降低分类时的复杂度，本章的子载波选择策略基于 DTW。最后，为了保证数据样本的一致性，对数据进行归一化处理。

(3) 手语分类阶段：为了达到预期的识别准确率，本系统的识别模块采用了基于 CNN 和 BiLSTM 的深度学习模型。

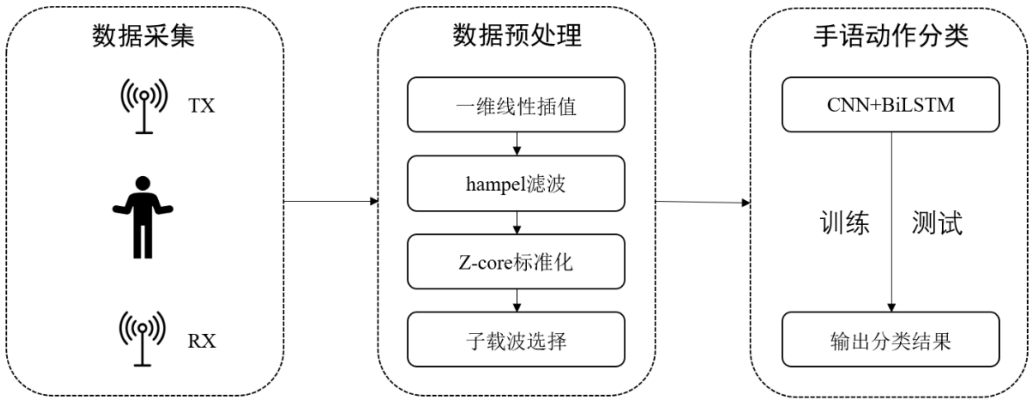


图 4.1 手语识别系统流程图

4.2 数据采集

在此手语识别系统中,接收端使用 CSI tool 来收集 CSI 数据。系统的接收机和发射机终端分别配备 3 根外接的全向天线。所有终端的无线网络接口控制器都是 Intel 5300 网卡,可以提供 30 个子载波的 CSI 数据。系统的采样率设置为 100Hz,数据单次发送时长设置为 5s,高采样率可以确保细粒度识别的要求,时长也足以让志愿者执行完一次手语动作。同时为了保证通信数据的可控性,本系统发送和接收数据时使用 Monitor 模式来代替传统的 AP 模式。AP 模式的缺点有多个方面,首先 ping 命令速度比较慢,其次数据发送有一定随机性,不能精确控制发送的参数,例如精确控制发送时间和单次发送的数据包数目等。而当工作在 Monitor 模式下时,可以直接通过终端命令自定义采样率、发包总数、发送天线个数和频宽模式。

4.3 数据预处理

由于 WiFi 信号传播过程中周围环境的干扰,无线信号传播路径也会发生变化,加上实验硬件自身的缺陷,以及无线通信链路本身存在的不稳定性,都会导致接收到的 CSI 数据不可避免的受到噪声干扰,与真实数据相去甚远,因此需要多种信号处理手段使原始实验数据变得可用。

数据预处理模块主要有四个部分,分别为数据插值, Hampel 去除离群值, Z-core 标准化和子载波选择。下面将详细描述这些步骤。

4.3.1 数据插值

插值是在一组已知数据点的范围内添加新数据点的技术。在数学中,线性插值是一种使用线性多项式在已知数据点的离散集合范围内构造新数据点的曲线拟合方法,适用于一维数据。

即使采集数据时采用了参数自定义,不易丢包的 Monitor 模式,实验时,由于无线信道自身的不稳定性,实际接收到的数据也会存在一定程度丢包,丢包程度一般在 10%以内,所以,本章采用一维线性插值方法对数据进行插值,以获取间隔均匀,长度一致的 CSI 序列。

4.3.2 Hampel 去除离群值

观察读取得初始 CSI 值, 会发现幅度值的序列中存在少量与其他数值差异较大的点, 这些离群值可以使用 Hampel 滤波器进行去除。

Hampel 滤波器使用 Hampel 标识符检测并去除输入信号中的离群值。Hampel 标识符基于中值和中值绝对偏差 (Median Absolute Deviation, MAD) 对离群值的位置和散布进行鲁棒的估计。对于输入数据的每个 CSI 样本, 都包含了 30 个子载波, 以一个子载波对应的, 长度为 n 的数据序列 $\{x_1, x_2, \dots, x_n\}$ 为例, 假设滑动窗口的大小为 $2k + 1$, 那么窗口序列的中值 m_i 为:

$$m_i = \text{median}(x_{i-k}, x_{i-k+1}, \dots, x_i, \dots, x_{i+k-1}, x_{i+k}) \quad (4.1)$$

中值绝对偏差 S_i 的公式如下:

$$S_i = \text{median}(|x_{i-k} - m_i|, \dots, |x_{i+k} - m_i|) \quad (4.2)$$

那么, 落在封闭区间 $[m_i - t * S_i, m_i + t * S_i]$ 以外的点可以视为离群值, 其中 t 是人为设置的参数。将 Hampel 滤波应用于所有样本的所有子载波。图 4.2 展示了一个子载波序列滤除离群值前后的对比。

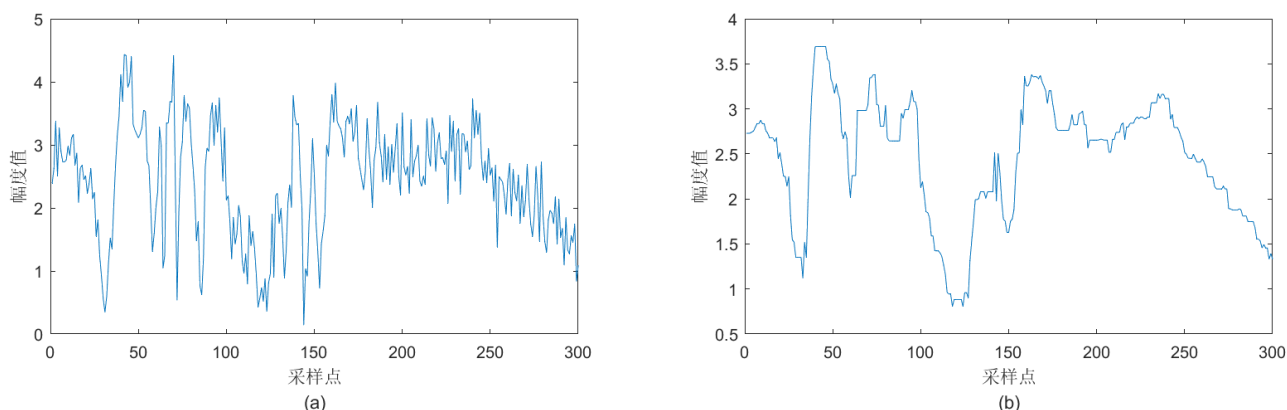


图 4.2 Hampel 滤波效果图: (a) Hampel 滤波前的子载波序列; (b) Hampel 滤波后的子载波序列

4.3.3 Z-core 标准化

数据标准化, 指把按比例缩放的数据置于一个小的特定区间之内。数据标准化一般是进行数据分析的常见准备工作, 可以消除评价指标量纲和取值规模的影响, 有利于不同量级或单位的指标间的比较或加权计算等。同理在深度学习中, 标准化可以使特征的权重保持一致, 这样网络就不会偏向于那些数值较大的特征, 减少了内部协变量偏移 (Internal Covariate Shift, ICS) 并加速了深度神经网络的训练, 减少了梯度对参数规模或其初始值的依赖性, 使优化更快, 并加速了随机梯度下降的收敛。

目前常见的数据标准化方式主要有 min-max 标准化和 Z-core 标准化等。本章采用 Z-core

标准化方法, Z-score 标准化使输入的数据序列 $\{x_1, x_2, \dots, x_n\}$ 均值为 0, 标准差为 1, 它的计算公式为:

$$x^* = \frac{x - \mu}{\delta} \quad (4.3)$$

其中为 x 数据序列中的值, x^* 是数据点标准化之后的结果, μ 为序列的总体平均值, δ 为序列的总体标准差。

4.3.4 子载波选择

由于子载波的反射路径不同, 因此其对于手语动作的感知粒度也不同。在本章的系统中, 通过筛选不同天线对之间的子载波, 并选择具有高度相关性的子载波, 来用于更高精度的手势分类。主要思想就是: 在接收到每个样本的 30 个子载波里, 受环境噪声干扰较小的子载波, 往往呈现出更相似的波形, 而那些受干扰较大的子载波, 其波形会有更多的畸变, 与正常的子载波相差较大, 并且在受干扰较大的子载波之间, 也较少出现波形相似的情况, 这与受干扰较小的子载波的情况不同。

在此步骤本章选择了 DTW 作为子载波选择的指标。DTW 一般用于测量两个时间序列之间的相似性, 适用于视频、音频或图形数据等任何可以转换为线性序列的数据。

首先计算数据样本中 30 个子载波 $A = (A_1, A_2, \dots, A_{30})^T$ 彼此之间的 DTW 距离, 可以得到一个 30*30 的系数矩阵 C :

$$\begin{cases} C = \begin{bmatrix} c_{1,1} & \cdots & c_{1,30} \\ \vdots & \ddots & \vdots \\ c_{30,1} & \cdots & c_{30,30} \end{bmatrix} \\ c_{m,n} = dtw(A_m, A_n) \quad m, n \in [1, 30] \end{cases} \quad (4.4)$$

在得到 30 个子载波的系数矩阵之后, 按行进行求和, 就可以得到基于 DTW 的筛选指标:

$$K_{DTW}(i) = \sum_{j=1}^{30} C_{i,j} \quad (4.5)$$

从中找出筛选指标最小的 10 个子载波作为候选子载波, 示例如图 4.3 所示。

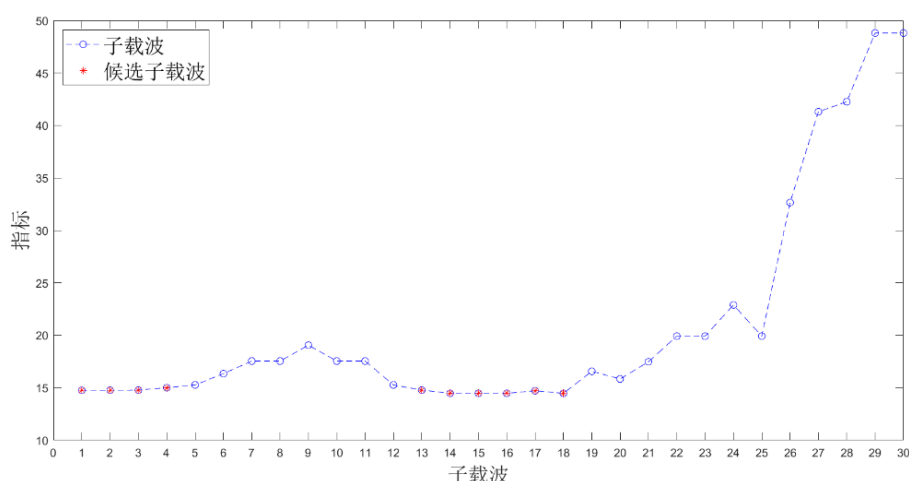


图 4.3 基于 DTW 的子载波选择

4.4 手语分类

为了对实验选取的多个手语手势进行识别，本章采用一个基于 CNN 和 BiLSTM 的深度学习模型。模型的架构如图 4.4 所示，包含了卷积层，池化层，BiLSTM 层，全连接层，softmax 层和分类输出层。

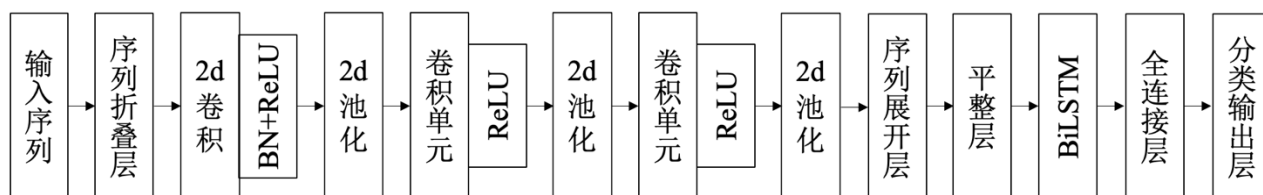


图 4.4 分类模型架构图

首先，预处理后的 CSI 数据先进行折叠，然后送入 3 层二维卷积层中，通过在不同的卷积层中使用不同大小的卷积核，可以提取输入数据的局部或全局特征，从而获得对输入数据的有效描述。然后，通过池化操作，降低特征维度以避免过拟合，同时保持特征尺度不变。每个卷积层后都紧接着一个批处理归一化层，使用 ReLU 作为激活函数。第四层是最大池化层，其主要作用为减少模型中参数的数量，降低模型过拟合的概率。然后，序列展开层恢复输入数据的序列结构。

接下来的 Flatten 层，把输入的多维数据转变为一维数据后送入 BiLSTM 层中，BiLSTM 由两个 LSTM 组成，一个接收正向输入，一个接收反向输入。LSTM 作为循环神经网络（Recurrent Neural Network, RNN）的一种，非常适合对时序序列进行建模，而在 LSTM 基础上改进的 BiLSTM，解决了 LSTM 无法编码从后向前信息的问题，适用于更细粒度的分类任务。

BiLSTM 层后的全连接层，作用是把之前提取的特征映射到样本空间中用作分类，其输出一个 $1 * K$ 的矩阵，其中 K 为手语动作的种类个数。Softmax 函数属于多类逻辑回归，是逻辑回归的推广，一般用于多分类的任务，它接收全连接层的输出，将其转化为归一化的概率分布，其公式如下：

$$\sigma(\vec{Z})_i = \frac{e^{Z_i}}{\sum_{j=1}^K e^{Z_j}} \quad (4.6)$$

其中 \vec{Z} 为 Softmax 的输入向量，由 (Z_1, \dots, Z_K) 组成。 Z_i 都是 Softmax 函数的输入向量的元素，可以取任何实数值，所以这不是一个有效的概率分布。标准指数函数 e^{Z_i} 应用于输入向量的每个元素，将其转化为一个大于 0 的正值。公式底部的项 $\sum_{j=1}^K e^{Z_j}$ 是归一化项，它确保函数的所有输出值总和为 1，并且每个值都在 $(0,1)$ 范围内，从而构成有效的概率分布 $P = [P_1, P_2, \dots, P_K]$ ，其中 P_i 为每个手语动作的相对概率值。

最后，分类输出层输出 Softmax 结果 P 中概率最高的手语动作作为分类结果。

4.5 实验与评估

本节中将主要介绍实验场地、设备和人员的布置情况，对基于 WiFi 的手语识别系统的性能表现进行评估，并探究各种因素对于手语动作分类准确率的影响。

4.5.1 实验设置

在手语识别系统中，采用 170mm*170mm 的 mini PC 终端和 ThinkPad 笔记本作为收发设备，均搭载 Intel 5300 无线网卡。手语识别采用 Monitor 模式进行无线通信，工作频率为 5.8GHz，发包速率 100Hz。收发终端配备 3 根全向天线，天线间距 3cm（小于无线信号波长 4cm），天线高度 1.2m，设备间距 1.2m。接收端安装 Linux CSI Tool 采集 CSI 数据，使用 MATLAB R2021 作为实现整个手语识别系统的开发软件。

图 4.5（a）和图 4.5（b）显示了实验场景的布局。本文在实验室和会议室这两个不同的房间测试了手语识别系统的性能，志愿者站在收发设备视距路径的中心执行手语动作。在两个房间中，分别收集了 3 名志愿者执行的 6 个手语手势，每个手语手势执行 50 次。从作为国家语言文字规范的《国家通用手语常用词表》中挑选出用于识别的 6 个手语，都是日常生活中的常用词语，如图 4.6 所示。手语识别系统根据第 4.4 节的方法识别手语。基于实验室内采集到的数据，本文 4.5.2 节也讨论了不同天线距离、视距路径和不同分类器对于实验结果的影响。

响。

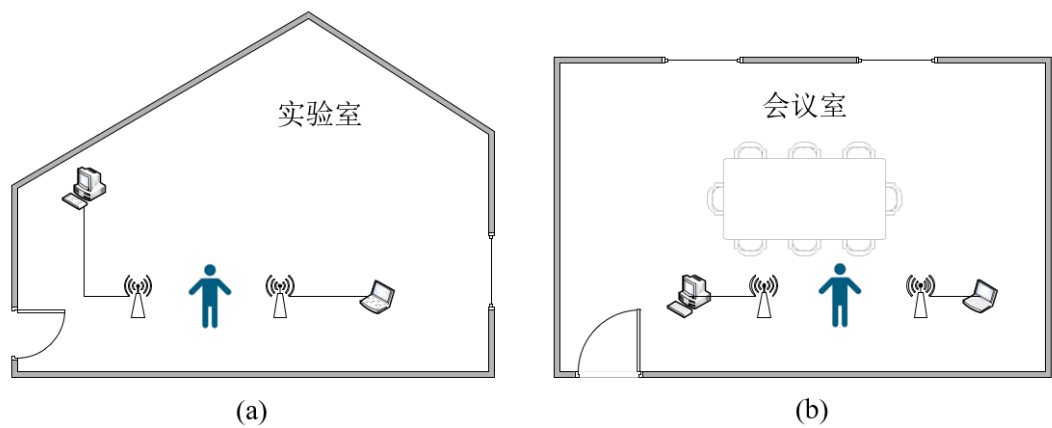


图 4.5 实验场景示意图：(a) 实验室场景布局；(b) 会议室场景布局

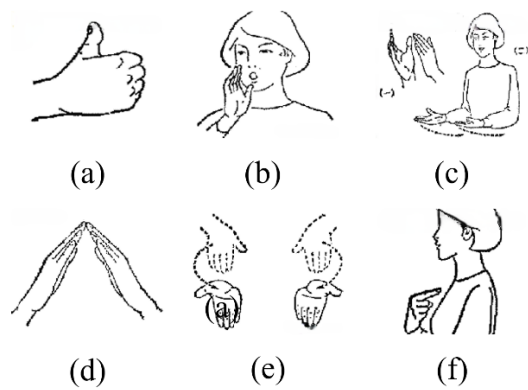


图 4.6 实验选择的六种手语含义：(a) 好；(b) 叫；(c) 欢迎；(d) 家；(e) 什么；(f) 我

4.5.2 表现评估

手语识别系统识别志愿者执行的 6 种代表不同含义的手语，通过数据采集、信号预处理和手语分类这 3 个模块分别收集、处理和分析 CSI 数据，得出正确的分类结果。6 种手语动作的平均识别准确率达到了 97.78%，图 4.7 展示了 6 种手语动作的混淆矩阵图，总的来说，手语动作 2~4 的识别准确率都达到了 100%，只有动作 1 和动作 6 的识别准确率略低，为 93.33%。实验结果表明，手语识别系统可以有效识别出 6 种代表不同含义的手语动作。

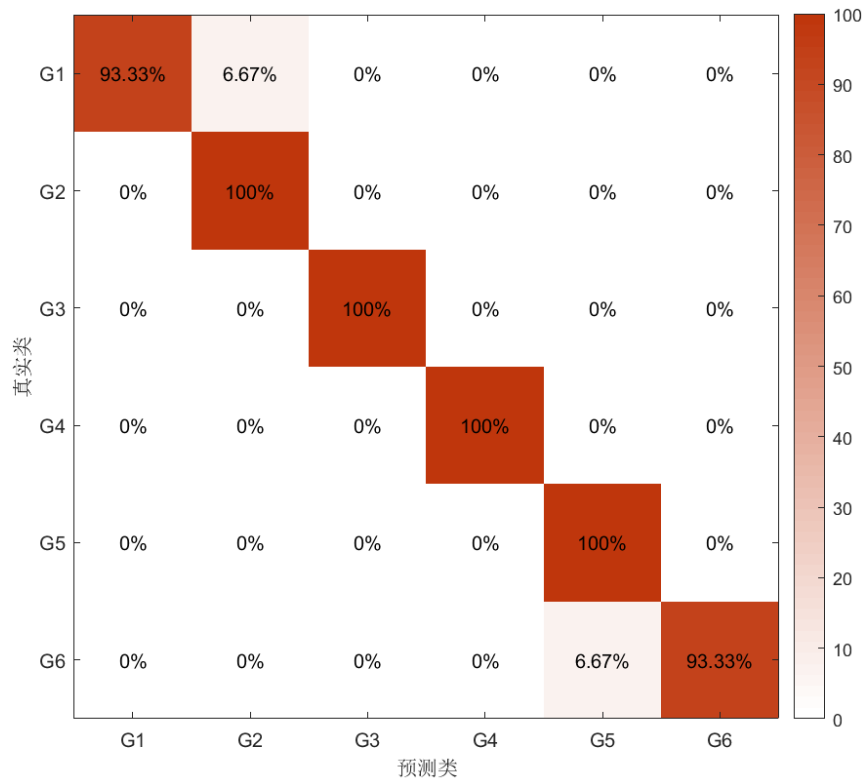


图 4.7 手势识别的混淆矩阵图

不同环境的影响：为了测试手语识别系统在不同的环境中能否正常的工作，安排志愿者在会议室中进行实验，按照与之前实验室相同的参数配置设备：Monitor 模式，发包速率 100Hz，天线间距 3cm，天线高度 1.2m，设备间距 1.2m。两个不同场景下的手语动作识别准确率对比如图 4.8 所示，可以看出，在不同的环境中，手语识别系统仍能保持较高的识别准确率。

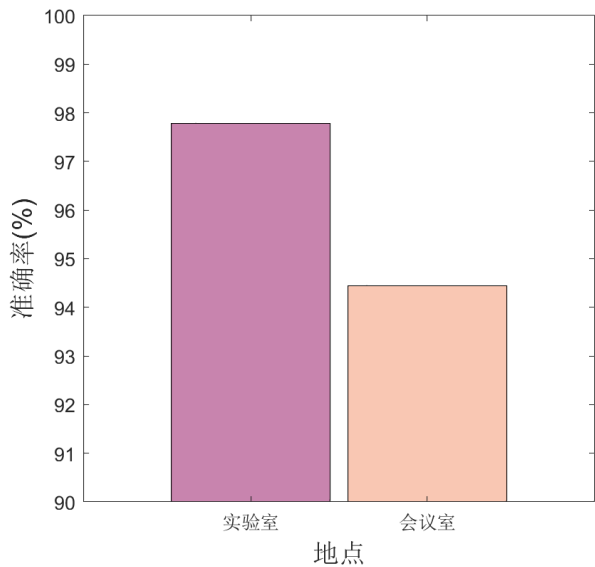


图 4.8 不同环境中手语识别系统的识别准确率

收发设备不同间距的影响：将接收端和发送端的设备间距分别设置为 1.2 米，2.4 米和 3.6 米，其他的因素设置都相同，测试不同距离下手语识别系统的性能表现。结果如图 4.9 所示。总体来说，手语动作的识别准确率随收发设备距离的增大而降低，但即使在 3.6 米的距离下，

依然保持着高于 90% 的识别准确率，证明手语识别系统的部署范围具有较大的灵活性。

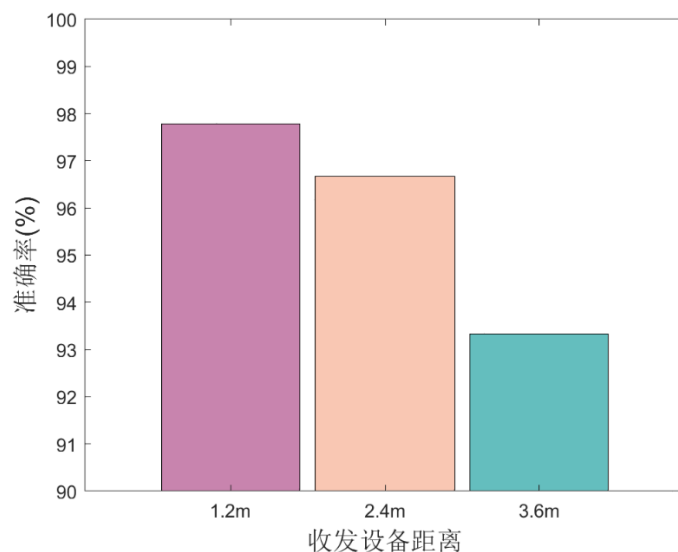


图 4.9 收发终端不同间距下的识别准确率

是否视距路径的影响：与基准实验中志愿者站在视距路径的中心执行手势动作不同，此对比实验要求志愿者的位置距离视距路径中心 1m 远，面向视距路径执行手语。识别准确率如图 4.10 所示，可以发现非视距路径下手语动作的识别准确率下降较为严重。

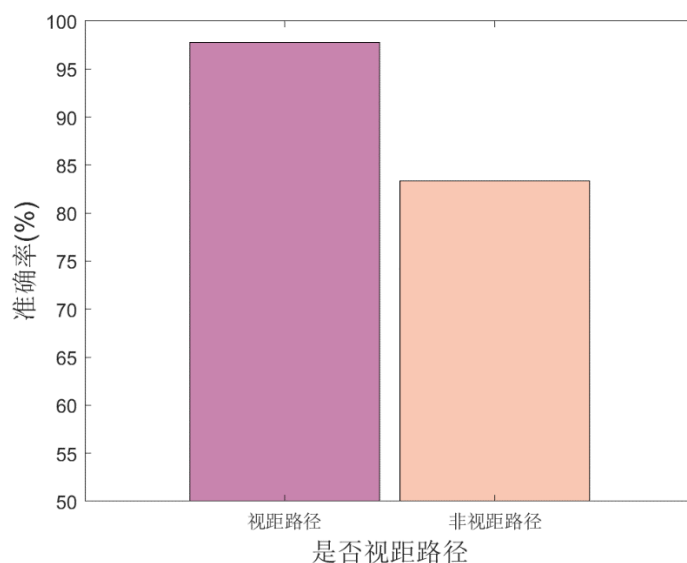


图 4.10 视距路径与非视距路径下的识别准确率

不同分类器的性能：在这里选择了 kNN 算法、SVM 算法和 CNN 算法与本文采用的分类模型进行对比。图 4.11 证明，基于 CNN 和 BiLSTM 的深度学习模型在手语识别的任务上，优于这几个主流的分类算法。

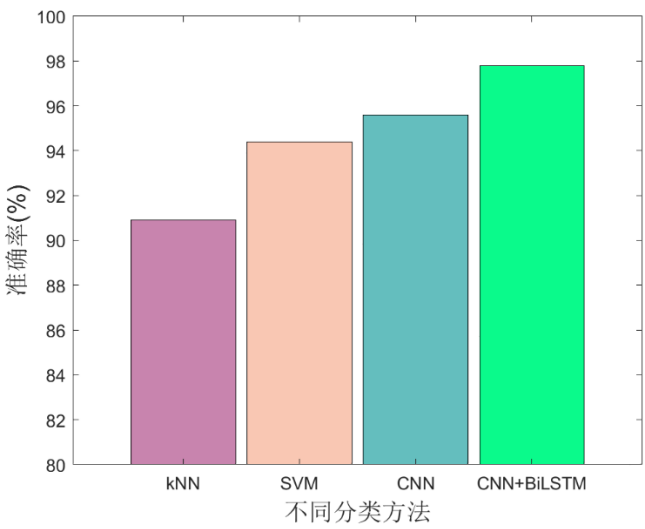


图 4.11 不同分类方法的表现

用户差异性对手语识别的影响：为了验证用户差异性对于手语识别的影响，分别基于 3 名不同体态的志愿者采集的手语数据进行实验。结果图 4.12 表明，用户的差异性对于手语识别的影响不大，总体识别率的差距都在 1%以内。

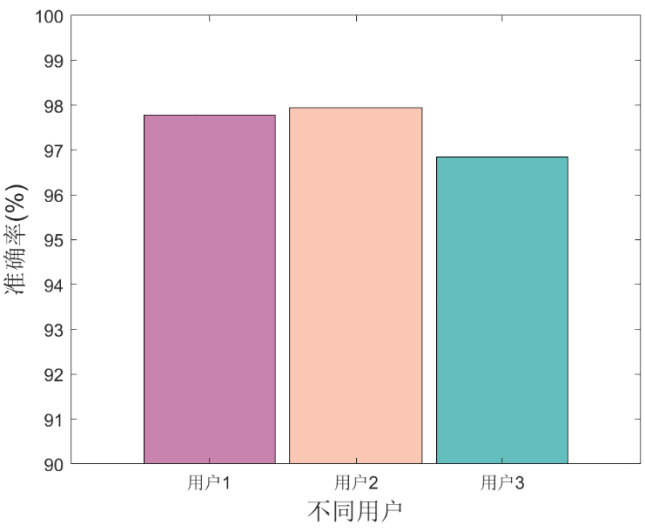


图 4.12 用户差异性对手语识别的影响

现有手语/手势研究的比较：本文选取 WiFinger^[53]、WiCatch^[55]和 WiSign^[56]。这三个研究工作与手语识别系统进行比较。对比结果如表 4.1 所示。

表 4.1 现有手语/手势研究的比较

	动作类型	动作数量	分类模型	识别准确率
WiFinger ^[53]	手语	9	DWT	90.4%
WiCatch ^[55]	双手手势	9	SVM	95%
WiSign ^[56]	手语	40	HMM	92%
手语识别系统	手语	6	CNN+BiLSTM	97.8%

4.6 本章总结

本章提出了一个基于 WiFi 的手语识别系统，介绍了该系统的研究背景，给出了系统实施的详细步骤，最后评估了系统在不同环境，不同因素影响下的表现。手语识别系统主要分为数据采集、数据预处理和手语分类这三大模块。系统在商用 Intel 5300 网卡的 Monitor 模式工作，收集来自于物理层的 CSI。同时为了获得可用的实验数据，采用了一系列信号处理的手段，先是对原始 CSI 数据进行一维线性插值。采用 Hampel 滤波器去除数据中的离群值，平滑曲线，去除环境噪声。对数据进行 Z-core 标准化。并基于 DTW 计算子载波之间的相关性，进行子载波选择。手语分类则采用了基于 CNN 和 BiLSTM 的深度学习模型。实验结果表明，本手语识别系统在各种环境中能够精确识别各类手语动作，系统采用的深度学习模型优于大部分分类器的表现。

第五章 总结与展望

随着通信与计算机技术飞速发展,尤其是物联网时代的带来,基于传感器和基于射频的行为感知技术得到了越来越多研究者的重视,在活动与手势识别、身份认证、人员识别和生命体征等领域进行了大量的研究。

在此基础上,本文首先介绍了行为感知的主要技术方向,着重介绍了其中 WiFi 感知技术的发展,并列举了 WiFi 感知技术在多个领域中的研究进展。分别提出了两个基于商用 WiFi 设备的行为感知应用:基于信道状态信息呼吸感知的人员识别系统和基于 WiFi 的手语识别系统。本文的主要工作如下:

(1)介绍了当前行为感知技术的主流技术,大致分为基于传感器的感知和基于射频的感知,其中用于感知的传感器可分为动作传感器、视觉传感器和声学传感器等,基于射频的感知技术则有 RFID、UWB、蓝牙和 WiFi 等不同的门类。

(2)比较并分析了 RSSI 和 CSI 的优缺点,相较粗粒度、不稳定且具有时间波动性的 RSSI 信号,CSI 稳定性高,能够让所有子载波一起以一种细粒度的方式描绘无线信道,进一步拓展了应用场景。

(3)考虑到传统的人员识别方法通常对用户有特殊要求且硬件成本较高。第三章提出了一个基于信道状态信息呼吸感知的人员识别系统。首先对 CSI 共轭相乘模型和菲涅尔域模型进行理论分析。然后结合中值滤波、EMD 算法和基于 FFT 的子载波选择策略来分离 CSI 信号中的呼吸分量。基于多天线的 CSI 数据流,使用 CFAR 寻峰算法来检测呼吸速率。通过滑动窗口算法生成数据集,使用 CNN 完成人员识别的任务。实验结果表明,本系统检测呼吸速率的平均绝对误差约为 0.5bpm,对 3 个场景中的人员识别准确率都达到了 85%以上,对不同人员能够进行有效的识别。

(4)为了增加聋哑人士的交互手段,改善他们的交流便利性。第四章提出了一个基于 WiFi 的手语识别系统。本系统基于商用设备实现,不需要用户佩戴其他设备,成本低廉,易于部署。系统采用 Hampel 滤波器滤除离群值,去除环境噪声,并通过一维线性插值和归一化的手段保证数据样本的一致性。采用基于 DTW 的选择策略筛选有效的子载波。识别手语时采用了基于 CNN 和 BiLSTM 的深度学习模型,识别准确率达到了 97.8%,能满足手语识别任务的要求。

本文的工作存在一些改进之处。人员识别系统可以基于互信息的概念动态确定用于重构信号的 IMF 数量。使用的均值类 CFAR 寻峰算法可以继续改进,例如选择 CFAR 算法中的自

适应类，通过分析训练单元数据选择不同的决策方法。可以进一步改进分类模型，优化分类器的性能。而手语识别系统可以展开对于连续手语动作识别的研究，采用深度学习方法而不是以方差作为阈值来分割连续的手语动作。

参考文献

- [1]. Mukhopadhyay S C. Wearable sensors for human activity monitoring: A review[J]. IEEE Sensors Journal, 2014, 15(3): 1321-1330.
- [2]. Li W, Pu H, Wang R. Sign Language Recognition Based on Computer Vision[C]. IEEE International Conference on Artificial Intelligence and Computer Applications, 2021: 919-922.
- [3]. Lan J, Zhang Z, Xiong S. Acoustic detection for vehicle targets and recognition by data fusion[C]. IEEE Instrumentation and Measurement Technology Conference Proceedings, 2005, 1: 551-553.
- [4]. Hatem E, Colin E, Abou-Chakra S, et al. New empirical indoor path loss model using active uhf-rfid tags for localization purposes[C]. IEEE International Conference on RFID Technology and Application, 2018: 1-6.
- [5]. Raptopoulos A, Yioultsis T, Dimitriou A G. Particle filter object tracking by a handheld UHF RFID reader[C]. IEEE International Conference on RFID Technology and Applications, 2019: 342-347.
- [6]. Li X, Yao D, Pan X, et al. Activity recognition for medical teamwork based on passive RFID[C]. IEEE International Conference on RFID, 2016: 1-9.
- [7]. Yang C, Wang X, Mao S. Respiration monitoring with RFID in driving environments[J]. IEEE Journal on Selected Areas in Communications, 2020, 39(2): 500-512.
- [8]. Park J, Cho S H. IR-UWB radar sensor for human gesture recognition by using machine learning[C]. IEEE International Conference on High Performance Computing and Communications, 2016: 1246-1249.
- [9]. Assawaroongsakul P, Khumdee M, Phasukkit P, et al. Deep Learning-Based Human Recognition Through the Wall using UWB radar[C]. International Joint Symposium on Artificial Intelligence and Natural Language Processing, 2021: 1-4.
- [10]. Daim T J, Lee R M A. Child Respiration States Determination by Using IR-UWB Radar Sensor-Based Human Motion Detection Method[C]. IEEE International Colloquium on Signal Processing and Its Applications, 2021: 133-137.
- [11]. Iglesias H J P, Barral V, Escudero C J. Indoor person localization system through RSSI Bluetooth fingerprinting[C]. International Conference on Systems, Signals and Image Processing, 2012: 40-43.
- [12]. Hou X, Arslan T. Monte Carlo localization algorithm for indoor positioning using Bluetooth low energy devices[C]. International Conference on Localization and GNSS, 2017: 1-6.
- [13]. Shahzad M, Zhang S. Augmenting user identification with WiFi based gesture recognition[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(3): 1-27.
- [14]. Li C, Liu M, Cao Z. WiHF: Enable user identified gesture recognition with WiFi[C]. IEEE Conference on Computer Communications, 2020: 586-595.
- [15]. Li C, Liu M, Cao Z. WiHF: Gesture and user recognition with WiFi[J]. IEEE Transactions on Mobile Computing, 2020.
- [16]. Zhang J, Wei B, Hu W, et al. WiFi-id: Human identification using wifi signal[C]. International Conference on Distributed Computing in Sensor Systems, 2016: 75-82.
- [17]. Lv J, Yang W, Man D, et al. Wii: Device-free passive identity identification via WiFi signals[C]. IEEE Global Communications Conference, 2017: 1-6.
- [18]. Zhou Z, Liu C, Yu X, et al. Deep-WiID: WiFi-based contactless human identification via deep learning[C]. IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, 2019: 877-884.
- [19]. Wang D, Zhou Z, Yu X, et al. CSIID: WiFi-based human identification via deep learning[C]. International Conference on Computer Science and Education, 2019: 326-330.

- [20].Ming X, Feng H, Bu Q, et al. HumanFi: WiFi-based human identification using recurrent neural network[C]. IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation, 2019: 640-647.
- [21].Wu C, Zhang F, Hu Y, et al. GaitWay: Monitoring and recognizing gait speed through the walls[J]. IEEE Transactions on Mobile Computing, 2020, 20(6): 2186-2199.
- [22].Zhang J, Wei B, Wu F, et al. Gate-ID: WiFi-based human identification irrespective of walking directions in smart home[J]. IEEE Internet of Things Journal, 2020, 8(9): 7610-7624.
- [23].Kaltioikallio O, Yigitler H, Jantti R, et al. Non-invasive respiration rate monitoring using a single COTS TX-RX pair[C]. IEEE International Symposium on Information Processing in Sensor Networks, 2014: 59-69.
- [24].Patwari N, Brewer L, Tate Q, et al. Breathfinding: A wireless network that monitors and locates breathing in a home[J]. IEEE Journal of Selected Topics in Signal Processing, 2013, 8(1): 30-42.
- [25].Patwari N, Wilson J, Ananthanarayanan S, et al. Monitoring breathing via signal strength in wireless networks[J]. IEEE Transactions on Mobile Computing, 2013, 13(8): 1774-1786.
- [26].Abdelnasser H, Harras K A, Youssef M. UbiBreathe: A ubiquitous non-invasive WiFi-based breathing estimator[C]. ACM International Symposium on Mobile Ad Hoc Networking and Computing, 2015: 277-286.
- [27].Liu X, Cao J, Tang S, et al. Wi-Sleep: Contactless sleep monitoring via WiFi signals[C]. IEEE Real-Time Systems Symposium, 2014: 346-355.
- [28].Liu X, Cao J, Tang S, et al. Contactless respiration monitoring via off-the-shelf WiFi devices[J]. IEEE Transactions on Mobile Computing, 2015, 15(10): 2466-2479.
- [29].Liu J, Wang Y, Chen Y, et al. Tracking vital signs during sleep leveraging off-the-shelf wifi[C]. ACM International Symposium on Mobile Ad Hoc Networking and Computing. 2015: 267-276.
- [30].Wang X, Yang C, Mao S. TensorBeat: Tensor decomposition for monitoring multiperson breathing beats with commodity WiFi[J]. ACM Transactions on Intelligent Systems and Technology, 2017, 9(1): 1-27.
- [31].Wang X, Yang C, Mao S. PhaseBeat: Exploiting CSI phase data for vital sign monitoring with commodity WiFi devices[C]. IEEE International Conference on Distributed Computing Systems, 2017: 1230-1239.
- [32].Wang X, Yang C, Mao S. On CSI-based vital sign monitoring using commodity WiFi[J]. ACM Transactions on Computing for Healthcare, 2020, 1(3): 1-27.
- [33].Chen C, Han Y, Chen Y, et al. TR-BREATH: Time-reversal breathing rate estimation and detection[J]. IEEE Transactions on Biomedical Engineering, 2017, 65(3): 489-501.
- [34].Wang F, Zhang F, Wu C, et al. Respiration tracking for people counting and recognition[J]. IEEE Internet of Things Journal, 2020, 7(6): 5233-5245.
- [35].Wang H, Zhang D, Ma J, et al. Human respiration detection with commodity wifi devices: do user location and body orientation matter?[C]. ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016: 25-36.
- [36].Yang Y, Cao J, Liu X, et al. Multi-person sleeping respiration monitoring with COTS WiFi devices[C]. IEEE International Conference on Mobile Ad Hoc and Sensor Systems, 2018: 37-45.
- [37].Zeng Y, Wu D, Gao R, et al. FullBreathe: Full human respiration detection exploiting complementarity of CSI phase and amplitude of WiFi signals[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(3): 1-19.
- [38].Zeng Y, Wu D, Xiong J, et al. FarSense: Pushing the range limit of WiFi-based respiration sensing with CSI ratio of two antennas[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2019, 3(3): 1-26.
- [39].Zeng Y, Wu D, Xiong J, et al. MultiSense: Enabling multi-person respiration sensing with commodity wifi[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(3): 1-29.
- [40].Liu J, Chen Y, Dong Y, et al. Continuous user verification via respiratory biometrics[C]. IEEE Conference on Computer Communications, 2020: 1-10.

- [41]. Sigg S, Shi S, Buesching F, et al. Leveraging RF-channel fluctuation for activity recognition: Active and passive systems, continuous and RSSI-based signal features[C]. Proceedings of International Conference on Advances in Mobile Computing and Multimedia, 2013: 43-52.
- [42]. Sigg S, Scholz M, Shi S, et al. RF-sensing of activities from non-cooperative subjects in device-free recognition systems using ambient and local signals[J]. IEEE Transactions on Mobile Computing, 2013, 13(4): 907-920.
- [43]. Wang Y, Liu J, Chen Y, et al. E-eyes: device-free location-oriented activity identification using fine-grained wifi signatures[C]. Proceedings of the Annual International Conference on Mobile Computing and Networking, 2014: 617-628.
- [44]. Wang W, Liu A X, Shahzad M, et al. Understanding and modeling of wifi signal based human activity recognition[C]. Proceedings of the Annual International Conference on Mobile Computing and Networking, 2015: 65-76.
- [45]. Wang W, Liu A X, Shahzad M, et al. Device-free human activity recognition using commercial WiFi devices[J]. IEEE Journal on Selected Areas in Communications, 2017, 35(5): 1118-1131.
- [46]. Sigg S, Blanke U, Tröster G. The telepathic phone: Frictionless activity recognition from wifi-rssi[C]. IEEE International Conference on Pervasive Computing and Communications, 2014: 148-155.
- [47]. Melgarejo P, Zhang X, Ramanathan P, et al. Leveraging directional antenna capabilities for fine-grained gesture recognition[C]. Proceedings of the ACM International Joint Conference on pervasive and ubiquitous computing, 2014: 541-551.
- [48]. Abdelnasser H, Youssef M, Harras K A. Wigest: A ubiquitous wifi-based gesture recognition system[C]. IEEE Conference on Computer Communications, 2015: 1472-1480.
- [49]. Nandakumar R, Kellogg B, Gollakota S. Wi-fi gesture recognition on existing devices[J]. arXiv preprint arXiv, 2014: 1411.5394.
- [50]. He W, Wu K, Zou Y, et al. WiG: WiFi-based gesture recognition system[C]. IEEE International Conference on Computer Communication and Networks, 2015: 1-7.
- [51]. Virmani A, Shahzad M. Position and orientation agnostic gesture recognition using wifi[C]. Proceedings of the Annual International Conference on Mobile Systems, Applications, and Services, 2017: 252-264.
- [52]. Sun L, Sen S, Koutsonikolas D, et al. Widraw: Enabling hands-free drawing in the air on commodity wifi devices[C]. Proceedings of the Annual International Conference on Mobile Computing and Networking, 2015: 77-89.
- [53]. Li H, Yang W, Wang J, et al. WiFinger: Talk to your smart devices with finger-grained gesture[C]. Proceedings of the ACM International Joint Conference on Pervasive and Ubiquitous Computing, 2016: 250-261.
- [54]. Ma Y, Zhou G, Wang S, et al. Signfi: Sign language recognition using wifi[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(1): 1-21.
- [55]. Tian Z, Wang J, Yang X, et al. WiCatch: A Wi-Fi based hand gesture recognition system[J]. IEEE Access, 2018, 6: 16911-16923.
- [56]. Zhang L, Zhang Y, Zheng X. WiSign: Ubiquitous american sign language recognition using commercial wi-fi devices[J]. ACM Transactions on Intelligent Systems and Technology, 2020, 11(3): 1-24.
- [57]. Jafar M T, Ababneh M, Al-Zoube M, et al. Forensics and analysis of deepfake videos[C]. IEEE International Conference on Information and Communication Systems, 2020: 053-058.
- [58]. Wang Z, Zhang J A, Xu M, et al. Single-Target Real-Time Passive WiFi Tracking[J]. IEEE Transactions on Mobile Computing, 2022.
- [59]. Halperin D, Hu W, Sheth A, et al. Tool release: Gathering 802.11 n traces with channel state information[J]. ACM SIGCOMM Computer Communication Review, 2011, 41(1): 53-53.
- [60]. 王唯嘉, 肖明清, 张磊, 陈茂才. 多小波包下经验模态分解去噪研究[J]. 计算机工程, 2015, 41(12): 125-129.
- [61]. 高亚军, 张冠杰, 陈矛. 基于 FPGA 的 CFAR 设计与实现[J]. 火控雷达技术, 2006(03): 64-68.

附录 1 攻读硕士学位期间申请的专利

[1] 肖甫, 朱旭, 郭政鑫, 盛碧云, 桂林卿, 周剑. 一种基于信道状态信息呼吸感知的人员身份识别方法, 202111181521.8, 2021.10.11。

附录 2 攻读硕士学位期间撰写的论文

- [1] Zhengxin Guo, **Xu Zhu**, Linqing Gui, Biyun Sheng, Fu Xiao. “BreathID: Respiration Sensing for Human Identification Using Commodity WiFi”, IEEE Systems Journal. (在投)

致谢

三年的研究生阶段即将结束，我也将结束学生生涯迈入社会。在南京邮电大学学习的七年是我最宝贵的经历与回忆。回首读研的三年，向给予我无私帮助的各位师长和同学致以衷心的感谢和美好的祝愿。

本研究在导师肖甫老师和盛碧云老师的指导下进行，从研究方向的确定，到开题报告、实验和最终的论文定稿，都离不开两位老师的指导。两位老师严谨的科研态度和深厚的专业功底，是我学习的目标与榜样。

学习和生活中同样离不开同学与朋友的关怀，在这里尤其感谢郭政鑫师兄一直以来的无私帮助。

特别感谢我的家人们的默默付出，他们是我求学路上的坚实后盾。

最后，感谢在座的各位专家和老师，感谢你们对本文的批评指正。

谢谢大家！