

面向智能家居的 Wi-Fi 手势识别交互系统

李瑾慧, 李莉

(北京邮电大学信息与通信工程学院, 北京 100876)

5 **摘要:** 基于 Wi-Fi 信道状态信息 (Channel State Information, CSI) 非接触式手势识别技术因其无需穿戴设备、保护用户隐私等优势, 在智能家居人机交互领域展现出广阔的应用前景。然而, 现有研究普遍面临两方面挑战: 跨域感知模型难以适应不同用户、环境与位置的变化, 以及高精度模型参数量大、无法在边缘设备上实时部署。针对上述问题, 本文设计并实现了面向智能家居场景的 Wi-Fi 手势识别实时交互系统 Wi-MotionSense。系统采用端到端流水线
10 架构, 涵盖 CSI 数据采集、实时预处理、动作分类与用户交互四个功能模块, 基于商用路由器与普通笔记本电脑即可完成部署。动作分类模块采用参数量仅 0.024M 的轻量化跨域识别模型, 具备良好的跨用户、跨环境泛化能力。系统在真实家庭环境 (5×3.5m 客厅) 中对 6 名未参与训练的新用户进行了完整测试, 整体手势识别准确率为 88.61%, 端到端延迟仅为 350ms, 跨用户泛化准确率达 85.8%~90.8%, 验证了系统在真实非受控场景中的实用性与鲁棒性。

15 **关键词:** 计算机应用; Wi-Fi 感知; 手势识别; 智能家居

中图分类号: TP393.17

A Interactive System for Wi-Fi-Based Gesture Recognition in Smart Homes

20

Li Jinhui, Li Li

(School of Information and Communication Engineering, Beijing University of Posts and Telecommunications, Beijing 100876)

25 **Abstract:** Contactless gesture recognition technology based on Wi-Fi Channel State Information (CSI) offers significant advantages in smart home human-computer interaction, such as eliminating the need for wearable devices and preserving user privacy. However, existing research generally faces two challenges: cross-domain perception models struggle to adapt to variations in different users, environments, and locations; and high-precision models often involve a large number of parameters, making real-time deployment on edge devices difficult. To address these issues, this paper designs and
30 implements Wi-MotionSense, a real-time interactive system for Wi-Fi gesture recognition tailored for smart home scenarios. The system adopts an end-to-end pipeline architecture, encompassing four functional modules: CSI data collection, real-time preprocessing, action classification, and user interaction. It can be deployed using commercial routers and standard laptop computers. The action classification module employs a lightweight cross-domain recognition model with only 0.024M
35 parameters, demonstrating strong cross-user and cross-environment generalization capabilities. The system was comprehensively tested in a real home environment (a 5×3.5m living room) with six new users who were not involved in the training. The overall gesture recognition accuracy achieved was 88.61%, with an end-to-end latency of only 350ms. The cross-user generalization accuracy ranged from 85.8% to 90.8%, validating the system's practicality and robustness in real-world, uncontrolled
40 scenarios.

Key words: Computer Applications; Wi-Fi Sensing; Gesture Recognition; Smart Home

0 引言

随着物联网与智能家居发展, 非接触式人机交互需求日益迫切。Wi-Fi 信号凭借覆盖广、

作者简介: 李瑾慧 (2001-), 女, 主要研究方向: WiFi 感知

通信联系人: 李莉 (1978-), 女, 副教授, 主要研究方向: 未来网络理论与技术、物联网、软件定义网络等. E-mail: lili66@bupt.edu.cn

45 无需穿戴、保护隐私等优势，以 CSI 为核心的 Wi-Fi 感知技术能捕捉人体运动引起的信道细
粒度变化，已在手势识别、跌倒检测等场景展现出广阔前景。

在实时感知系统研究方面，早期 WiSeac^[1]、WiGest^[2]、E-eyes^[3]等工作验证了 Wi-Fi 手
势识别的可行性，但受限于信号粒度或跨环境适应性。CSI 技术引入后，CARM^[4]建立了 CSI
速度谱与人体活动的关联模型，成为里程碑工作；WiHear^[5]、RT-Fall^[6]等系统进一步拓展了
50 感知场景；SignFi^[7]实现了复杂手语识别，WiFine^[8]在 Raspberry Pi 4B 上以 190ms 延迟达到
96% 准确率，推动了边缘化部署。然而，实时系统普遍存在跨域泛化能力不足的问题。在跨
域算法研究方面，多项研究专注于用 CSI 信号的幅度和相位差异生成高级动作指标^[9-11]，
Widar3.0^[12]提出了体速度分布域无关特征，EI^[13]、Wi-AM^[14]等研究通过对抗网络机制利用
目标域数据辅助训练。但是上述方法普遍参数量大、推理开销高，难以部署于资源受限的边
55 缘设备，且大多停留在离线验证阶段，缺乏完整的实时系统实现。

现有研究在两个方向存在割裂：实时系统缺乏跨域泛化能力，跨域算法难以落地为可部
署的完整系统。如何将轻量化跨域识别与端到端实时交互结合，是当前亟待解决的问题。为
此，本文设计并实现了面向智能家居场景的 Wi-Fi 手势识别实时交互系统 Wi-MotionSense。
系统采用端到端架构，涵盖 CSI 数据采集、实时预处理、动作分类与用户交互四个模块，基
60 于商用路由器与普通笔记本即可部署。动作分类模块中，本文使用双模态教师蒸馏：以 BVP
教师提供域无关跨域监督，以 CSI 教师提供同模态语义对齐监督，联合指导轻量化学生模型
训练。系统在真实家庭环境中验证，端到端延迟仅 350ms，对 6 名用户的识别准确率达
85.8%~90.8%。

本文主要贡献如下：①设计并实现完整的端到端实时系统 Wi-MotionSense，整体准确
65 率 88.61%；②提出双模态教师协同蒸馏框架，将跨域感知能力迁移至仅 0.024M 参数的轻量
化模型，适合 IoT 设备部署；③在真实家庭环境中完成系统级综合验证，从准确率、响应延
迟与用户泛化性三维度验证了系统的实用性与鲁棒性。一级标题

1 Wi-MotionSense 系统总体设计

本章从系统层面对 Wi-MotionSense 进行整体介绍，包括系统的设计目标、整体架构以及
70 端到端工作流程，为后续各模块的详细实现描述提供全局视角。Wi-MotionSense 以商用 Wi-Fi
设备为硬件基础，通过四个功能模块的协同配合，构建从原始无线信号采集到用户交互指令
输出的完整端到端闭环，兼顾跨域感知能力与实时交互体验，面向普通家庭环境的实际部署
需求。

1.1 设计目标

75 Wi-MotionSense 实时人体动作识别系统以智能家居场景为核心应用背景，旨在构建一
套基于 Wi-Fi 信道状态信息 (CSI) 的非接触式、轻量化人体动作识别系统。系统的整体设
计目标围绕感知能力、实时性与易用性三个核心维度展开，力求在有限的硬件资源条件下，
实现准确、高效、易部署的动作识别能力。

在感知与识别能力方面，系统须保障信号采集的稳定性与数据质量，收发链路维持不低

80 于 200 包每秒的采集速率，并集成实时预处理机制对原始 CSI 数据进行清洗与规整。在此基础上，系统采用“检测—分类”级联架构，先实现对静止状态与有效动作的鲁棒区分，再对触发后截取的 2 秒 CSI 窗口执行语义级动作分类。在典型家居多径干扰环境下，系统对四类预定义动作的分类准确率设计目标不低于 80%，从动作检测触发至识别结果输出的端到端延迟须严格控制在 2 秒以内。

85 在轻量化与易用性方面，系统仅依托标准商用 Wi-Fi 路由器与普通计算设备即可完成部署，核心识别模型采用轻量化深度学习架构，支持在边缘计算或资源受限平台上稳定运行，无需引入专用硬件。用户交互层面则提供基于 Web 的可视化界面，支持 CSI 波形实时展示与识别结果呈现，用户仅需通过浏览器即可便捷访问与操作，显著降低系统的部署门槛与维护成本。

90 **1.2 系统整体架构设计**

基于上述需求分析，Wi-MotionSense 系统采用分层架构设计，将整个系统划分为四个功能模块：CSI 信号采集模块、数据预处理模块、动作分类模块与用户交互模块，如图 1 所示。各模块之间通过标准化接口协议进行数据交互，职责边界清晰，耦合度低，共同构成完整的端到端处理链路。这一分层设计不仅有助于系统各部分的独立开发与测试，也为后续功能扩展与性能优化提供了良好的结构基础。

95

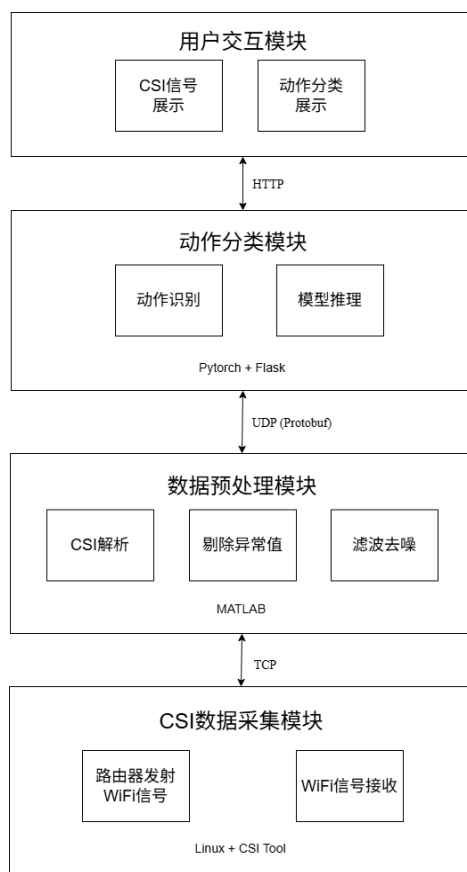


图 1 系统架构

100 CSI 信号采集模块作为系统的感知入口，承担着从物理环境中获取原始无线信道信息的核心职责。该模块由 Wi-Fi 路由器与搭载 Linux 内核及 Intel 5300 网卡的接收端设备协同构成。路由器采用支持 IEEE 802.11be 协议的标准商用设备，工作于 2.4GHz 频段，通过多天线配置实现 MIMO 传输，持续向监测区域广播无线信号。接收端运行 CSI Tool 软件，

从物理层实时提取 3×3 MIMO 配置下 30 个有效子载波的复数信道响应矩阵，并将原始信道数据封装为标准格式后向上层传输，该模块仅依赖标准商用硬件即可完成部署，无需专用采集设备，有效降低了系统的硬件成本与部署门槛。

105 数据预处理模块基于 MATLAB 平台构建，承担对原始 CSI 数据进行清洗、增强与标准化转换的关键职责，旨在为后续识别模型提供规整、纯净的输入特征。该模块通过解析来自采集层的 TCP 数据流提取复数信道矩阵，随后依次执行异常值剔除、滤波去噪与数据规整等操作。为保障与下游模块的高效跨平台交互，预处理完成的数据经由 Protobuf 进行序列化封装，转换为结构紧凑、语言中立的标准格式，并通过 UDP 协议执行分块传输，确保大
110 数据量下特征传输的可靠性。

动作分类模块是系统的智能决策中枢，整体基于 PyTorch 深度学习框架构建，采用“检测—识别”级联架构，由动作检测与动作识别两个核心子模块协同实现从信号感知到语义输出的完整推理流程。动作检测子模块对预处理后的实时 CSI 数据流进行持续监测，通过提取时域能量、方差等特征并与自适应阈值进行比对，实现对静止状态与有效动作事件的鲁棒区分，并在判定动作发生时立即触发后续处理流程。动作识别子模块则基于双模态教师协同蒸馏框架训练的轻量化学生模型，对触发截取的标准化信号窗口执行端到端推理，在低算力约束下实现对四种预定动作的高精度分类，并将识别结果经实时推送至前端。

115 用户交互模块作为系统与用户之间的连接，基于 Web 技术构建，集成了实时信号监控与识别结果展示两大功能。该模块利用 ECharts 图表库将底层 CSI 数据转化为动态可视化的多天线波形图，使用户能够直观观察信号随人体动作产生的实时变化。与此同时，模块通过 RESTful API 接口为外部应用提供标准化的动作数据接入能力，支持智能家居设备的联动控制，有效拓展了系统的应用场景与实用价值。

1.3 系统工作流程

125 Wi-MotionSense 系统的完整工作流程可划分为信号采集、数据预处理、动作分类与结果交互四个依次推进的阶段，各阶段紧密衔接，共同保障系统从原始物理信号到最终动作语义输出的全链路处理能力。

系统启动后，Wi-Fi 路由器持续向监测区域广播无线信号。当人体在监测范围内产生动作时，其躯体对无线信号传播路径的干扰将引发信道特性的相应变化，并被实时记录于 CSI 数据之中。接收端通过 CSI Tool 软件持续采集信道响应数据，并经 TCP 协议上传至预处理模块。预处理模块随后对原始数据依次执行异常值剔除、滤波去噪与归一化处理，生成格式统一的时序样本，并经 Protobuf 序列化封装后通过 UDP 协议实时传输至动作分类模块。
130 动作分类模块收到数据后，首先由检测子模块对实时 CSI 数据流进行持续监测，当提取的时域特征变化量超过自适应阈值时，判定有效动作发生并触发截取机制，锁定 2 秒长度的标准化信号窗口。随后，识别子模块将该窗口数据送入轻量化学生模型完成端到端推理，输出动作类别与置信度。最终，识别结果经由 RESTful API 实时推送至前端交互界面，用户可通过 Web 界面直观查看 CSI 波形变化与识别结果，外部应用亦可通过标准化接口接收动作数据，实现智能家居设备的联动控制。

2 Wi-MotionSense 系统实现

本节将对 Wi-MotionSense 系统各模块的具体实现细节进行详细介绍。系统整体基于标

140 准商用硬件设备搭建，主要使用华为 XIHE-BE70 路由器、中兴 MU5002 5G 移动热点设备以及装配了 Intel 5300 网卡的联想 ThinkPad X201i 笔记本电脑，编程语言涵盖 C 语言、Python 与 MATLAB。系统实现了两种创新交互模式：一是基于 Web 的实时 CSI 信号可视化展示界面，当系统识别到特定动作时，对应图标会发出橙色光效并产生跳动动画，为用户提供直观反馈；二是集成了 KM（Keyboard Mapper）模块，将识别结果映射至预设键盘按键，实现基于人体动作的游戏角色实时控制，有效拓展了系统的应用场景，Wi-Fi 感知知识识别实时系统流程图如图 2 所示。

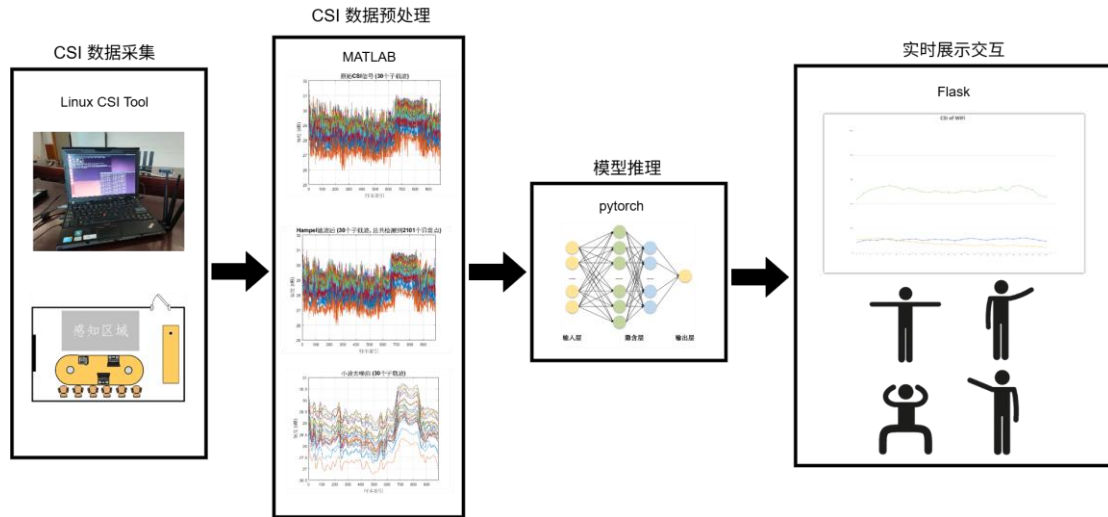


图 2 Wi-Fi 感知知识识别实时系统流程图

2.1 CSI 数据采集模块

150 CSI 数据采集模块为上层算法提供高质量的原始输入，在信号发射端，系统采用华为 XIHE-BE70 路由器作为主信号源，该路由器支持 IEEE 802.11be 标准，具备 2.4GHz 与 5GHz 双频段发射能力，在 AP 模式下可提供稳定的多天线信号覆盖。与此同时，系统引入中兴 MU5002 5G 移动热点设备作为基础网络信号源，MU5002 支持 Sub-6GHz 频段，下行峰值速率可达 3.8Gbps，内置 4500mAh 锂电池，可满足长时间连续工作的续航需求，充分保障了系统在实验与部署场景中的稳定运行。在信号接收端，系统采用联想 ThinkPad X201i 笔记本电脑作为专用 CSI 数据采集设备。该设备装配 Intel 5300 网卡，能够提供幅度与相位信息完整的复数信道矩阵。

160 在数据采集机制上，系统的探测信号并非专用帧，而是由接收端周期性发送 ICMP 回显请求至路由器，路由器协议栈自动生成并回复 ICMP 应答帧，该帧经由射频前端承载于 2.4GHz 频段的 OFDM 子载波上发射，从而完成探测信号的激发与广播。将 Intel 5300 网卡配置为特殊监听模式，通过 CSI Tool 软件以 500 毫秒为间隔实时监听并捕获来自路由器的数据包。每当捕获到一个 ICMP 应答帧时，系统即从网卡硬件寄存器中提取对应的 CSI 矩阵，以复数形式记录该帧在多天线、多子载波上的信道响应。最终，数据采集程序通过 TCP 连接将 CSI 数据实时传输至推理主机的 8090 端口，完成从物理层感知到数字信号获取的完整链路。

2.2 数据预处理模块

165 数据预处理模块基于 MATLAB 平台构建，采用 TCP 服务器架构，在 8090 端口建立与采集层的稳定连接，使用 read_bfee 函数对原始数据流进行解析。每个数据单元 (bfee)

170 由 field_len、code 与 field 三部分构成,当 code 值为 187 时,field 中携带有效的 CSI 信息,系统仅对满足此条件的数据帧进行后续解析,从而提取出包含天线编号、子载波索引与复数信道响应的完整信道矩阵。

175 在数据缓存与管理层面,系统构建了基于滑动窗口的时序数据管理机制,在内存中维护一个容纳最近 3 秒时间跨度的数据缓存区。新数据到达时,系统采用先进先出策略更新窗口内容,在尾部添加最新帧的同时移除头部过时数据,从而确保窗口内数据的连续性与时效性。系统通过解析 Protobuf 消息体中的 Matrix 域与 PRBITEM 字段,精确定位各数据单元对应的天线源和频率位置,并依据时间戳、天线编号、载波索引等关键标识符将分散的 CSI 信息整合为规范化的三维数据矩阵。

180 模块采用两阶段处理策略。第一阶段执行异常值检测与修正,应用 Hampel 滤波算法对 CSI 幅度数据进行处理:该算法通过计算滑动窗口内数据的中位数与中位数绝对偏差(MAD)来识别异常值,当数据点偏离中位数超过 3 倍 MAD 时,将其标记为异常并以窗口中位数替代,滑动窗口大小设置为 10 个数据点,有效消除了因硬件故障或环境突变引入的脉冲噪声。第二阶段执行基于小波变换的多尺度降噪处理:系统选取 Daubechies4 小波作为基函数,对 Hampel 滤波后的 CSI 信号进行 4 层小波分解,基于自适应阈值策略对高频细节分量实施噪声抑制,最终通过逆变换重构得到高品质的净化 CSI 数据序列。经过上述两阶段处理,信号信噪比得到显著提升,为后续动作检测提供了高质量的输入数据。185 预处理完成后,数据经 Protobuf 进行序列化封装,转换为结构紧凑、语言中立的标准格式,并通过 UDP 协议经由 9090 端口实时传输至动作分类模块,在保障传输实时性的同时确保数据的完整性与跨平台兼容性。

2.3 动作分类模块

190 动作分类模块是系统的智能决策中枢,基于 Python 3.10.14 平台搭建,以 PyTorch 2.3.1 作为深度学习推理引擎,同时集成 Flask 3.1.0 Web 框架以对外提供标准化数据接口。在架构设计上,模块采用“检测—识别”两阶段级联方式,由动作检测与时序分割子模块和深度学习推理子模块协同配合,共同构成从信号感知到语义输出的完整推理链路。

195 在动作检测与时序分割阶段,模块基于累积阈值检测算法对预处理后的实时 CSI 数据流展开持续监控,通过量化各子载波信号偏离环境静态基线的累积程度来判别有效动作的发生。系统启动后,首先采集前 20 秒的 CSI 数据并计算各子载波幅度均值,以此建立反映当前环境信道静态特征的基准参考。进入实时监测阶段后,系统对逐帧到达的 CSI 数据计算各子载波幅度与基准值的绝对差值,并在 5 帧(0.5 秒)的累积窗口内进行时间维度上的叠加累计。当某一子载波的累积差值突破基准值 1.5 倍的预设阈值时,该子载波被判定为“活跃状态”;一旦活跃子载波的数量超过总数的半数,系统便认定当前存在显著的信道扰动,随即触发动作数据段的提取流程。提取过程采用固定时间窗口策略:以触发时刻为起点,向前截取 0.5 秒数据作为动作前期段,向后截取 1.5 秒数据作为动作执行段,两者合并形成时长为 2 秒的完整动作窗口。这一窗口划分方式契合人体动作的生理时序规律,可完整涵盖动作的准备、执行与恢复三个阶段。

205 在深度学习推理阶段,系统将截取的 2 秒 CSI 数据窗口转化为张量格式,对复数 CSI 信号实施幅度与相位的分离提取及归一化处理,进而构建高表达性的特征向量。该特征向量被输入至基于双模态教师协同蒸馏框架训练得到的轻量化学生模型,完成针对四种预定义动作的端到端分类推理,并依据模型输出的概率分布确定最终的动作类别与置信度。识别结果

210 随后经由 Flask 封装的 RESTful API 实时推送至前端交互界面，完成整个推理流程的闭环输出。

2.4 用户交互模块

交互模块作为系统面向用户的核心接口层，基于 HTML5、CSS3 与 JavaScript 技术栈构建前端界面，并借助 Apache ECharts 5.4 图表库实现 CSI 信号的实时可视化呈现。模块通过 RESTful API 与 Flask 后端建立异步通信机制，以 80 毫秒为间隔持续向后端轮询数据，保障界面信息的实时刷新与流畅展示。界面主体由信号展示区与动作展示区两部分构成。信号展示区利用 ECharts 线图组件同步绘制三路天线的 CSI 幅度变化曲线，横轴对应 30 个子载波编号，纵轴覆盖 0 至 100 的幅度区间。动作展示区设有四个与预定义动作一一对应的状态图标，当识别模块输出特定动作类别时，前端通过 updateGesture 函数动态更新对应图标的 CSS 样式，触发橙色光效与跳动动画，以直观的视觉反馈向用户呈现当前识别结果。前端交互页面如图 3 所示：

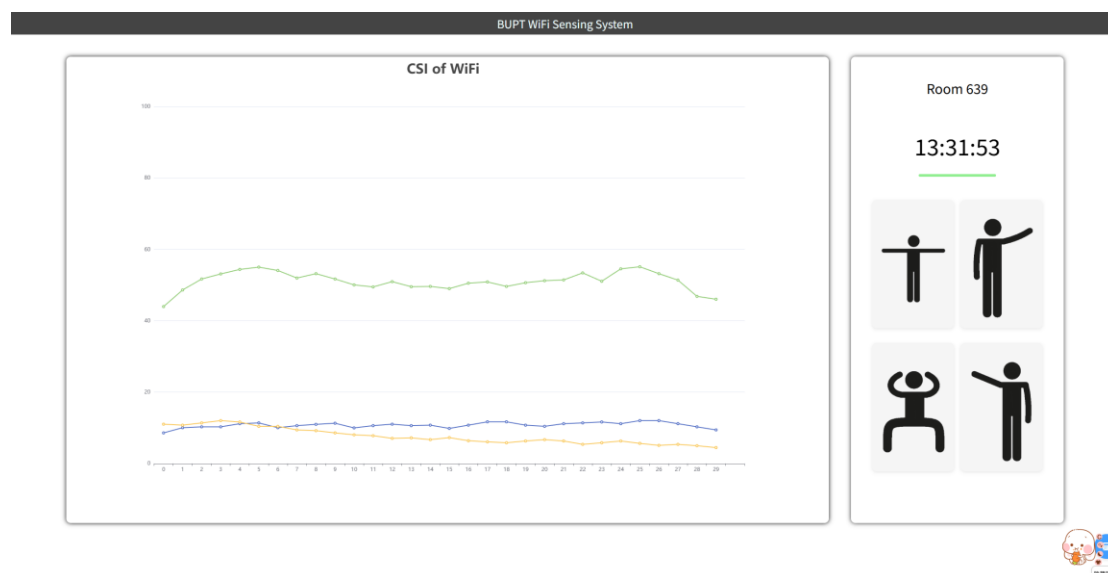


图 3 用户交互页面

在扩展交互能力方面，系统集成 miniKM 游戏控制模块，构建了从动作识别结果到键盘指令的完整映射链路。该模块分为控制端与被控端两部分：控制端通过 win32api 库捕获当前动作识别结果，并经 UDP 协议将对应按键指令发送至目标设备；被控端接收指令后，调用 Windows API 完成虚拟键盘事件的模拟触发。系统预设了四种动作与方向键的对应关系，up、left、right、down 四类动作分别映射至键盘的上、左、右、下方向键，用户可通过肢体动作直接操控游戏角色移动。这一“动作—指令”映射框架具备良好的可扩展性，其底层逻辑同样适用于智能家居设备控制、多媒体指令触发等更广泛的应用场景，为系统向实用化方向延伸提供了技术支撑。

3 实验验证

3.1 实验设置

为全面评估 Wi-MotionSense 系统在实际应用场景下的功能完整性与性能可靠性，本文在典型家庭环境中搭建了系统测试平台并制定了规范化的测试方案。

235 在测试环境方面,实验选取一间面积为 5m×3.5m 的家庭客厅作为测试场地。该场景配置有冰箱、桌椅、盆栽等常见家居物品,墙面悬挂画框,地面铺设瓷砖,还原了真实居住环境的空间布局与多径干扰条件,能够有效模拟系统在实际部署中所面临的复杂电磁环境。

240 在硬件部署方面,系统采用三点式布局进行设备配置。左侧部署 CSI 数据采集终端,由搭载 Intel 5300 网卡的联想 ThinkPad X201i 笔记本电脑担任,运行 CSI Tool 软件负责实时捕获无线信道状态信息;右侧放置 Wi-Fi 信号发射设备,采用华为 XIHE-BE70 路由器作为系统射频信号源;中央位置设置由 mini 主机与显示器构成的信号处理与显示单元,负责接收采集终端传输的 CSI 数据,并执行实时预处理、特征提取与动作识别,同时通过 Web 界面呈现识别结果与信号可视化图表。所有关键设备均通过专业可调节三脚架固定,设备高度统一标定为 100cm,以保障测试过程中空间几何配置的一致性与实验结果的可重现性,测试场景如图 4 所示

245



图 4 测试场景

250 实验共招募六名测试人员参与,覆盖不同年龄段与体型,具备一定的人员多样性。测试过程中,各测试人员在上述生活化场景中依次完成系统预定义的四类动作,每类动作进行重复测试,以系统评估模型在真实用户条件下的分类准确性与鲁棒性。

3.2 测试结果及分析

为全面评估 Wi-MotionSense 系统在实际部署环境中的综合表现,本节从功能性测试与非功能性测试两个维度对测试结果进行系统梳理与分析,重点围绕信号采集稳定性、动作识别准确性、交互响应实时性以及用户泛化能力四个方面展开。

3.2.1 信号采集测试结果分析

255

260 信号采集模块作为整个系统数据流链路的起点,其稳定性与可靠性直接决定了后续动作识别的数据质量。在连续 30 分钟的采集测试中,系统 CSI 数据接收成功率达到 99%,仅有约 1% 的数据包因网络波动或短暂信号遮挡而丢失,满足预期性能要求。MATLAB 可视化界面实时显示的多天线 CSI 幅度波形清晰连续,当测试人员在监测区域内执行动作时,对应子载波的信号幅度出现明显波动,直观验证了人体动作对无线信道状态的干扰机制,表明采集链路能够有效捕获由人体运动引发的信道变化特征,为后续识别算法提供了可靠的物理信号基础。总体而言,信号采集模块在典型家庭多径干扰环境下表现出良好的稳定性,具备支撑实时动作识别任务的数据采集能力。

3.2.2 动作识别功能测试结果分析

265 动作识别功能测试旨在验证系统对四种预定义动作——挥手、拍手、滑动与画 Z——的
端到端识别能力。测试采用随机序列混合执行的方式，对四类动作进行组合测试，评估系统
在动作快速切换场景下的识别稳定性。测试结果表明，Wi-MotionSense 系统能够成功识别
全部四种预定义动作，当测试人员执行特定动作时，系统可正确判定动作类别，并在 Web 界
270 面上触发对应图标的橙色高亮与跳动动画效果，实现了从底层信号采集到上层语义输出的完
整功能闭环。动作识别功能测试如图 5 所示，测试过程中，动作检测模块能够正常触发截取
流程，系统在动作间快速切换时亦未出现明显的响应滞后或误触发现象，验证了“检测—识
别”级联架构在功能层面的有效性与完整性。

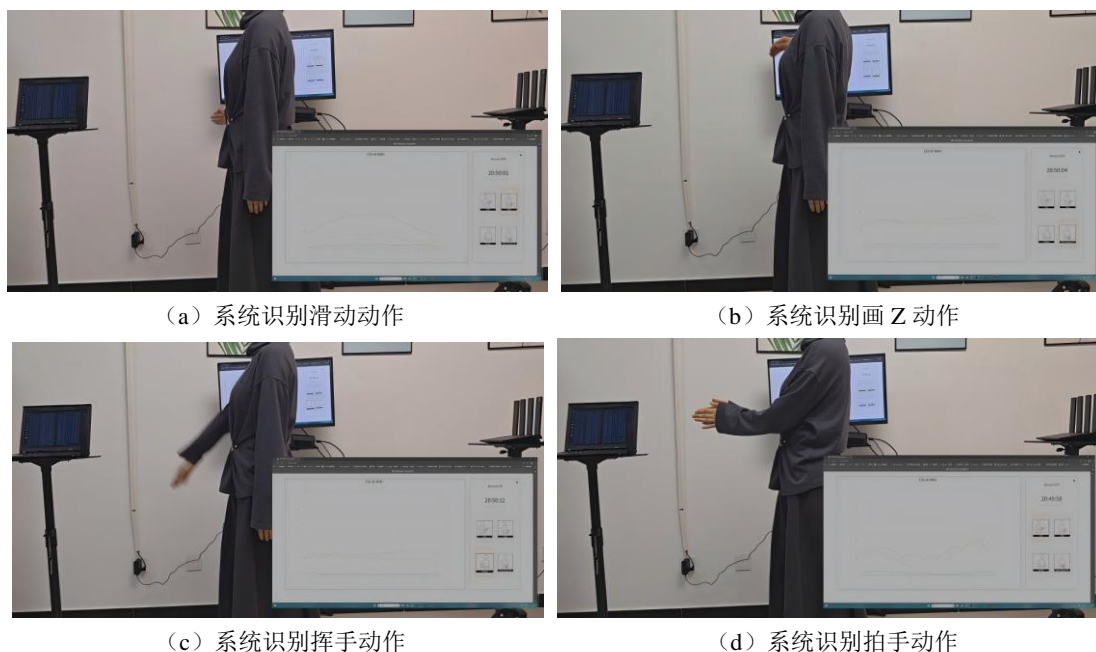


图 5 动作识别功能测试

3.2.3 用户交互功能测试结果分析

280 用户交互功能测试在动作识别的基础上，进一步验证了系统将识别结果转化为实际设备
控制指令的能力。测试借助 miniKM 键盘映射模块，将挥手、拍手、滑动与画 Z 四种动作
分别绑定至键盘的下、上、右、左方向键，通过观察游戏角色是否按预设映射关系作出正确
响应来评估映射逻辑的准确性。键盘映射控制功能测试如图 6 所示。测试结果显示，四种动
285 作与方向控制指令之间的映射关系均能正确执行，未出现映射错误或指令丢失的情况。当测
试人员执行挥手动作时，游戏角色向下移动；拍手对应向上移动；滑动对应向右移动；画 Z
对应向左移动，所有映射指令均得到准确传递与响应。这一结果验证了 miniKM 模块的功
能完整性与映射逻辑的正确性，同时也证明了基于 Wi-Fi 感知的动作识别结果可作为可靠
的控制信号源，为无接触式人机交互应用提供有效支撑。

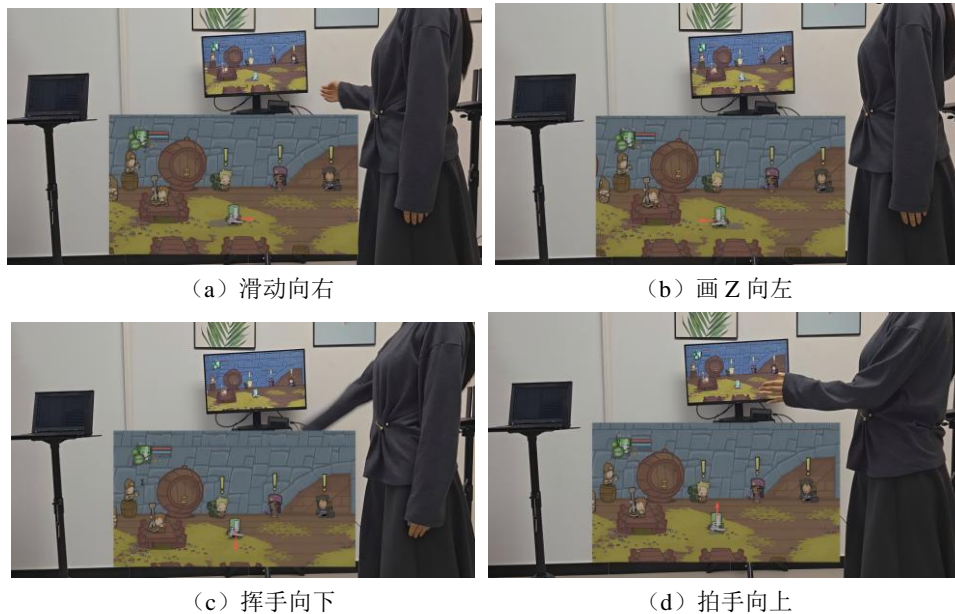


图6 键盘映射控制功能测试

3.2.4 系统准确性测试结果分析

系统准确性测试由六名测试人员参与,每人对每种动作重复执行 30 次,共产生 720 个有效测试样本。测试过程严格控制环境变量,保持设备位置、距离及环境条件一致,以确保结果的可比性。统计结果显示,720 个样本中共有 82 个识别错误,系统整体识别准确率约为 88.61%,表明系统在典型家庭多径干扰环境下具备较为可靠的识别能力。

从各动作类别的识别表现来看,挥手、拍手与滑动三类动作的识别准确率相对较高,而画 Z 动作的识别准确率显著低于其他三类。分析其原因,画 Z 动作在时域上具有较为复杂的轨迹结构,其信道扰动模式与其他动作存在一定程度的特征重叠,在多径干扰较强的实际环境中更容易引发误分类。这一结果也在一定程度上揭示了系统在复杂轨迹类动作识别上仍存在的提升空间,后续可通过增强训练数据多样性或引入更强特征表达能力的模型结构加以改善。

3.2.5 系统实时性测试结果分析

系统实时性测试采用视频同步分析方法进行延迟测量,使用 60fps 摄像设备同步录制用户动作执行过程与系统响应画面,通过逐帧分析计算动作起始时刻至游戏角色作出响应移动之间的时间差,以获取可靠的端到端延迟数据。测试结果表明,系统在标准测试环境下的端到端平均延迟为 350ms,最小延迟为 250ms,最大延迟为 580ms。整体而言,系统响应性能满足实时人机交互的要求,绝大多数情况下延迟控制在 500ms 以内,用户可获得较为流畅的交互体验。最大延迟 580ms 的出现,主要与动作检测阶段的触发时机判定及网络传输抖动有关,在环境干扰较强或动作执行不够标准时,检测模块可能需要更多帧数的累积才能触发截取流程,从而带来额外的响应延迟。总体来看,350ms 的平均端到端延迟验证了基于 Wi-Fi 感知的实时人机交互系统在响应速度方面的技术可行性,能够满足智能家居控制、游戏娱乐等对时延有一定要求的典型应用场景。

3.2.6 用户泛化测试结果分析

用户泛化测试邀请 6 名具有不同身体特征的被试者参与,身高范围覆盖 165cm 至

190cm, 体重分布在 55kg 至 80kg 区间, 年龄跨度为 17 至 52 岁, 构成了具有一定代表性的多样化用户样本。每名被试者按照统一规范完成四种动作的测试, 通过对比个体间的识别效果评估系统的泛化能力。

320 测试结果显示, 6 名被试者的动作识别准确率分布在 85.8% 至 90.8% 之间, 整体波动幅度处于合理范围。这表明系统对不同身体特征用户的动作识别具有较为稳定的适应能力, 用户个体差异对系统识别性能的影响在可接受范围之内。分析个体间差异的成因, 身高与体型的不同会改变人体对无线信号的散射特性, 进而影响 CSI 信号中动作相关特征的分布规律; 年龄差异则可能带来动作执行速度与幅度的差异, 对检测模块的触发阈值判定产生一定干扰。尽管如此, 所有被试者的识别准确率均维持在 85% 以上, 验证了基于 Wi-Fi CSI 的动作识别方法对多样化用户群体具备良好的泛化能力, 为系统在实际多用户环境中的稳定部署提供了可靠的性能保障。

325 综合上述六项测试结果可以看出, Wi-MotionSense 系统在功能完整性、识别准确性、响应实时性与用户泛化能力等方面均达到或超过了设计目标, 初步验证了系统在真实家庭环境中的工程可行性与实用价值。

330

4 结论

本文设计并实现了面向智能家居场景的 Wi-Fi 手势识别实时交互系统 Wi-MotionSense, 系统涵盖 CSI 数据采集、实时预处理、动作分类与用户交互四个功能模块, 构建了从原始无线信号到用户交互指令的完整端到端处理链路。在动作分类模块中, 提出双模态教师协同蒸馏框架, 通过 BVP 教师与 CSI 教师的协同监督, 将跨域泛化能力迁移至参数量仅 0.024M 的轻量化学生模型, 在保证跨域识别性能的同时满足了边缘实时推理的资源约束。系统在真实家庭环境中进行了全面的系统级测试, 对 720 次手势测试实现了 88.61% 的整体识别准确率, 端到端延迟仅为 350ms, 对 6 名用户的泛化准确率达 85.8%~90.8%, 验证了系统在真实非受控场景中的实用性与鲁棒性。

335

340 当前系统仍存在一定局限性: 一方面, 系统目前仅支持 4 种手势类别, 无法满足更丰富的交互指令需求; 另一方面, 动作检测算法基于固定阈值设计, 在信号环境发生较大变化时可能出现漏检; 此外, 系统尚未在多人同时运动的复杂场景下进行验证。

345

未来工作将从以下几个方向展开: 扩展可识别手势集合, 通过增加训练类别和优化模型结构支持更多样的交互指令; 研究自适应动作检测机制, 提升系统在动态复杂环境中的鲁棒性; 探索多用户感知场景下的信号分离与个体识别技术; 并进一步推动系统与智能家居 IoT 设备生态的集成, 向更广泛的实际部署应用迈进。

345

[参考文献] (References)

- 350 [1] Abdelnasser H, Harras K A, Youssef M. WiGest demo: A ubiquitous WiFi-based gesture recognition system[C]//2015 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS). IEEE, 2015: 17-18.
- [2] Pu Q, Gupta S, Gollakota S, et al. Whole-home gesture recognition using wireless signals[C]//Proceedings of the 19th annual international conference on Mobile computing & networking. 2013: 27-38.
- 355 [3] 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 20th annual international conference on Mobile computing and networking. 2014: 617-628.
- [4] Wang W, Liu A X, Shahzad M, et al. Understanding and modeling of wifi signal based human activity recognition[C]//Proceedings of the 21st annual international conference on mobile computing and networking. 360 2015: 65-76.
- [5] Wang G, Zou Y, Zhou Z, et al. We can hear you with Wi-Fi![C]//Proceedings of the 20th annual international conference on Mobile computing and networking. 2014: 593-604.
- [6] Wang H, Zhang D, Wang Y, et al. RT-Fall: A real-time and contactless fall detection system with commodity WiFi devices[J]. IEEE Transactions on Mobile Computing, 2016, 16(2): 511-526.
- 365 [7] 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.
- [8] Xing T, Yang Q, Jiang Z, et al. WiFine: Real-time gesture recognition using Wi-Fi with edge intelligence[J]. ACM Transactions on Sensor Networks, 2022, 19(1): 1-24.
- [9] Abuhoureyah F, Sim K S, Wong Y C. Multi-user human activity recognition through adaptive location-independent WiFi signal characteristics[J]. IEEE Access, 2024, 12: 112008-112024.
- 370 [10] Meng W, Chen X, Cui W, et al. WiHGR: A robust WiFi-based human gesture recognition system via sparse recovery and modified attention-based BGRU[J]. IEEE Internet of Things Journal, 2021, 9(12): 10272-10282.
- [11] Ding X, Yu X, Zhong Y, et al. Robust gesture recognition method toward intelligent environment using Wi-Fi signals[J]. Measurement, 2024, 231: 114525.
- 375 [12] Zhang Y, Zheng Y, Qian K, et al. Widar3. 0: Zero-effort cross-domain gesture recognition with Wi-Fi[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(11): 8671-8688.
- [13] Jiang W, Miao C, Ma F, et al. Towards environment independent device free human activity recognition[C]//Proceedings of the 24th annual international conference on mobile computing and networking. 2018: 289-304.
- 380 [14] Xie J, Li Z, Feng C, et al. Wi-AM: Enabling cross-domain gesture recognition with commodity Wi-Fi[J]. Sensors, 2024, 24(5): 1354.