

AIoT时代的智能无线感知： 特征、算法、数据集

杨 铮¹ 郑 月¹ 吴陈沐²

¹ 清华大学

² 马里兰大学

关键词：无线感知 信号特征 识别算法 数据集

背景

人类对物理世界的感知经历了从主观感受到传感器再到传感网的发展阶段，如图1所示。在传感器变得越来越小、数据采集变得越来越普及的同时，有一个问题也日益突出——传感系统的部署成本太高。特别是随着感知范围和规模不断增大，部署和维护长期稳定运行的大规模传感系统的难度和成本也越来越高。以室内人员定位为例，虽然我们可以在特定的房间、工厂里部署定位设施和系统，但是这些特定的区域与人类活动所处的整个物理世界相比无异于九牛一毛。当我们面对全世界范围内的住宅、学校、商场、机场、写字楼时，这种部署方式就显得无能为力了。那么有没有可能突破传统思维，在不部署任何专用传感器的情况下也能感知各种环境信息呢？已有研究工作利用电视广播信号来获取能量，为低功耗物联网设备供电，受此启发，我们是否能够实现非传感器感知，利用环境中已经存在的无线射频信号（例如Wi-Fi、RFID、蓝牙、Zig-Bee以及广播电视信号等）来感知人的动作行为？如果可行，我们就无须在环境中部署专用的传感器，也不需要人员携带传感器。目前无线网络在全球范围内已经广泛普及，这种感知物理世界的方式将显著降低部署成本，在易用性、普适性等方面取得重

要突破。

环境中已有的无线信号（声、光、射频信号等）在完成本职任务（照明、通信等）的同时，还可以“额外”用来感知环境。以射频信号为例，信号发射机产生的无线电波在传播过程中会发生直射、反射、散射等物理现象，从而形成多条传播路径。这样一来，在信号接收机处形成的多径叠加信号就携带了反映信号传播空间的信息。无线感知技术（或称为非传感器场景感知技术）通过分析无线信号在传播过程中的变化，获得信号传播空间（信道）的特性，以实现场景的感知。这里的场景既包括人



图1 感知手段的发展

的因素（是否有人以及人的位置、姿势、动作等），也包括其他外物的因素。非传感器感知提供了一种全新的物理世界感知方式，即无须部署传感器，只“复用”环境中已有的无线信号即可实现场景感知。非传感器场景感知将人类对物理世界的感知方式推动到了一个新的阶段——从以前单纯依赖部署专用传感器的方式升级为“专用”与“复用”相结合的方式。

无线感知技术将感知与通信合二为一，具有三个鲜明特点——“三无”：(1) 无传感器 (sensorless)，感知人和环境不再需要部署专门的传感器，这有别于无线传感网中由传感器负责感知而无线信号负责通信；(2) 无线 (wireless)，无须为通信及传感器部署有线线路；(3) 无接触 (contactless)，相较于现在市场上的各种可穿戴式智能设备，无线感知更向前迈了一步，无须用户佩戴任何设备。

无线感知技术的感知对象包括环境、物品和人，潜在应用十分丰富。以感知人为例，无线感知技术可以用于被动式人员感知。“被动式”在这里指的是人员不需要携带任何电子设备，用以区别传统无线定位系统中，通过定位人所携带的电子设备来定位人员，这样的方式也被称作设备非绑定的 (device-free) 或者非侵入式的 (non-invasive)。被动式人员检测可广泛使用于各种普适计算的应用中，提供更好的基于用户位置的服务。例如，博物馆中参观者接近某个展品时自动播放展品说明，超市统计近期最受关注的商品，或者在电梯及车厢中统计乘客数量等。非传感器感知还可以作为一种新型人机交互方式，通过识别人的行为（姿势、动作以及手势等微小运动）来遥控电子设备（计算机、游戏机、智能硬件等），完成特定的功能或提供交互式体感游戏；也可以用于智慧医疗监护，检测人员的睡眠质量以及老年人的意外跌倒等。被动感知的模式还契合安全保卫应用的需求。在涉密区域监控、人员入侵检测、灾难应急响应、重要物品保护等与安全相关的应用中，需要及时发现未携带任何无线通信设备的人员（工作人员或入侵者）是否在敏感区域出现并监测他们的活动。传统安防传感器中的红外传感器或者摄

像头都受到可视角度的限制，只能在一个很有限的角度内监测目标，而且不能应对烟雾、遮挡乃至视觉欺骗的情况。在电影和生活中已经出现了针对传统安防传感器局限性的人员入侵方法，而基于无线信号的非传感器感知可以有效克服此类漏洞。

近年来，无线感知技术引起了学术界的广泛关注。在无线网络与移动计算的著名会议与期刊（例如 ACM SIGCOMM、ACM MobiCom、ACM MobiSys、IEEE INFOCOM、USENIX NSDI、*IEEE/ACM ToN*、*IEEE JSAC*、*IEEE TMC*）上，许多论文将无线感知从梦想带入现实，在提升感知精度、提高鲁棒性、拓展应用场景等方面做出了重要贡献。除此之外，工业界也在探索非传感器感知的产品化，在智能家居、安防监控、健康监护等方面形成了实用的产品。

问题与挑战

无线感知蓬勃发展的另一面是其在实际应用中的局限性，具体体现在**特征、模型、数据集**三个方面：

1. 有效特征湮没：信号特征与背景环境相关，导致感知结果依赖部署环境，泛化能力差、学习训练成本高。造成环境依赖性强的根本原因在于，已有研究工作提取的信号特征（例如信号幅度、相位等）严重依赖于系统部署的具体环境。因此，不同的使用环境、不同的用户，甚至同一用户的不同位置、不同朝向等都会降低感知准确性，导致完全无法实现感知。针对新环境，需要重新采集数据进行训练，造成无线感知普适性差并且学习训练成本高。但令人遗憾的是，已有的工作绝大部分都是基于环境相关的信号特征来实现的。

2. 识别模型粗陋：缺乏在无线信号空间对人的行为活动的精细时空建模，导致感知精度低、鲁棒性差。人的姿势、动作、行为模式具有特定的时空特征。计算机视觉方向的研究工作针对图像和视频数据对人体运动特征进行建模并识别，取得了较大突破。然而在无线信号空间，缺乏对人

行为活动的精细时空建模。因此,已有的工作通常简单地使用计算机视觉领域的方法,直接在原始信号层次上进行识别,或者仅仅使用时间相关性或空间相关性,缺乏同时整合利用二者的有效手段,造成相似动作难以准确判别,在实际环境中感知精度低、鲁棒性差。

3. 数据集缺失: 高质量公开数据集的缺失造成性能比较不客观、实验结果难复现、技术进步难积累。越来越多的研究人员认识到,高质量的数据集在提取有效特征、训练精确模型以及提升跨场景学习能力等方面会起到至关重要的作用。特别是深度学习算法的广泛使用,对数据集的规模和质量提出了更高的要求。然而目前公开的数据集非常少,且数据量不足、应用场景少,难以满足深度学习算法的需求,往往导致识别模型欠拟合、泛化能力差。高质量公开数据集的缺失造成许多研究工作陷入“自说自话”的境地——性能比较不客观、实验结果难复现、技术进步难积累。

特征

既然无线信号在传播过程中“调制”了环境信息,那么如何从接收信号中“解调”这些环境信息呢?例如:如果手机接收到的 Wi-Fi 信号较弱,可能是由于手机距离无线路由器较远;而如果手机接收到的 Wi-Fi 信号强度骤降,很可能是因为手机进入了某些特定的封闭空间如电梯等。在这个例子中,我们以接收信号强度 (Received Signal Strength Indicator, RSSI) 为特征,推断手机的位置和所处的环境。信号特征的选取对于感知精度、可靠性、模型泛化能力等具有关键影响。

在无线感知领域, RSSI 的普适性使其被广泛用于无线室内定位、被动式人员检测等应用中。理论上,可将 RSSI 代入无线信号传播模型估算信号传播距离,也可把 RSSI 作为特定地点的无线信号特征“指纹”,还能通过 RSSI 的波动推断是否有人或其他障碍物阻挡无线链路。但在室内环境中, RSSI 会因信号多径传播引起的小尺度阴影衰落而不再随

传播距离增加单调递减,从而限制了测距精度。另外,多径传播也会引起信号强度波动,从而导致定位时无线信号指纹错误匹配。

为刻画多径传播,无线信道通常可用信道冲激响应 (Channel Impulse Response, CIR) 或频率响应 (Channel Frequency Response, CFR) 来建模^[1-3]。以前精确测量 CIR/CFR 通常需要专业信道测量仪器,例如网络矢量分析仪。而现在,在普通 Wi-Fi 设备上,我们也能以信道状态信息 (Channel State Information, CSI) 的形式获取一个采样版本的 CFR。RSSI 只反映了多径叠加的总幅度,而 CSI 呈现了不同频率(对应不同子载波)下多径传播的幅度和相位,从而更加精确地刻画了具有频率选择性衰落特性的信道^[1]。

相位信息是 CSI 区别于 RSSI 的一大特征。然而,受载波频率误差、采样频率偏差等因素的影响,CSI 相位通常包含较多噪声。因此,CSI 相位信息必须先处理再使用^[4]。

波达角 (Arrival of Angle, AoA) 与 **飞行时间** (Time of Flight, ToF) 是雷达信号与声音信号定位、追踪目标时常用的两种特征。然而受多径效应与信号带宽的限制,在室内环境下准确估计经由目标反射而来的信号的 AoA 与 ToF 具有一定挑战性。已有工作^[5]观察到在发射机、接收机或目标只有微小位移时,经由目标一次反射而来的信号的 AoA 比其他多次反射信号的 AoA 更加稳定,而且多次反射信号可能只在某些接收机的 AoA 谱上出现。因此可通过多径压缩的算法,滤除多次反射信号,获得定位目标的方位信息。此外,在多天线的基礎上,可利用子载波多样性进行空间平滑,从而利用多个数据包提升 AoA 估计精度^[6]。为了克服直接测量 AoA 与 ToF 的不精确性,最新的研究工作 RIM^[3]通过对天线阵列收到的 CSI 信息进行处理,测量移动目标的相对位移(位移方向、大小、速度等),创新性地利用无线射频信号实现了类似惯性传感器的功能,使得远距离目标追踪精度达到了 8.4 厘米。

多普勒频偏 (Doppler Frequency Shift, DFS) 也是一种常见的用于刻画人员活动与行为的特征,这是因为目标移动将导致经由目标反射的信号的路径

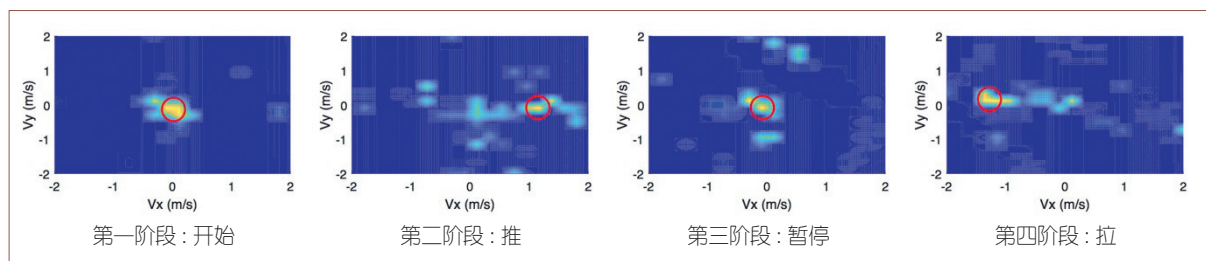


图2 动作“推与拉”的BVP

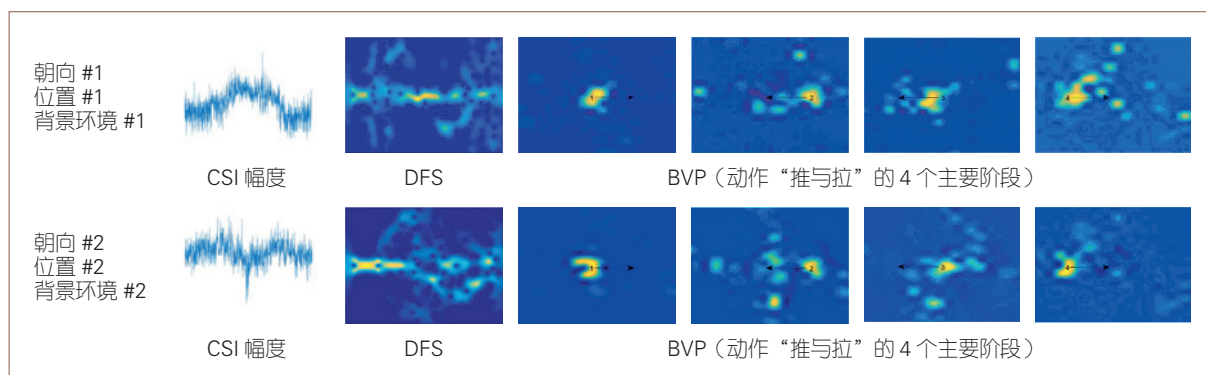


图3 同一动作在不同场景下CSI幅度、DFS与BVP的对比

长度产生变化,使得观测到的信号频率发生一定的偏移。对CSI功率进行时频分析(如短时傅里叶变换和小波变换),可从功率的动态变化中提取人体运动导致的DFS^[7]。一些工作进一步观察到:只有目标的径向速度分量对应着反射路径长度的变化速率,提取到的DFS和运动速度方向、人员位置存在着一定的几何约束关系。因此,通过添加更多链路可消除速度求解的歧义性,获得完整的人员速度信息,并应用于定位、追踪、活动识别等多种场景^[8]。

多维度信号特征的融合通常可以取得比使用单一信号特征更好的感知效果。对于许多环境,多链路条件可能过于严格,比如人们通常只在家里配置一个路由器。在只有一对收发设备的情况下,需要提取出尽可能多的参数,才能对目标进行细粒度的感知与定位。已有工作^[9]将CSI建模成关于幅度衰减、DFS、ToF、AoA的函数,使用基于期望最大化的算法对多径信道的多参数进行联合估计,并利用图匹配技术得到经由目标反射路径的长度与目标的方位信息的最优解,最终实现了单链路下的分米级被动式人员追踪系统。通过配置更多的发射

天线,可对反射路径的DFS、ToF、AoA和出发角(Angle of Departure, AoD)进行联合估计,进一步提升单链路下的感知精度^[10]。

上述的信号特征都是**环境依赖的**——在反映人的活动的同时,也反映环境情况。然而,如果不能有效地区分人和环境的信息,场景感知的效果必将严重依赖于系统的部署环境,使得不同的使用环境、不同的用户,甚至同一用户的不同位置、不同朝向等都会降低感知准确性,还会造成感知普适性差、学习训练成本高等问题。

最新的解决方案是通过提取**环境无依赖的信号特征**,在信号特征层面上提升感知模型的泛化能力。环境无依赖的信号特征不受人员位置、朝向与背景环境等因素的影响,只反映人员活动本身的信息。通过对信号的精细建模,我们提出一种环境无依赖的信号特征——人体坐标系下速度谱(Body-coordinate Velocity Profile, BVP)^[2]。如图2所示,BVP是一个三维特征,它刻画了人员在进行特定活动时信号能量在不同速度(对应产生运动的不同身体部位)下分布的变化趋势。与单纯的躯干速

度 (velocity) 信息相比, BVP 传达的信息更为丰富。而且由于其构建在以人员位置为原点、人员朝向为 x 轴正方向的人体坐标系中, 避免了在地球坐标系下不同位置与朝向会对同一活动的速度分布产生不同影响的问题。

图 3 展示了一位用户在不同环境、不同位置、不同朝向下完成同一手势时, 感知系统采集得到的三个信号特征: CSI 幅度、DFS 以及 BVP。由于 CSI 主要刻画了信号传播空间的特性, 受背景环境的影响, 即使是同一手势对应的 CSI 也可能完全不同。DFS 虽然在原理上体现了目标移动的特性, 然而由于其只能反映目标的径向运动速度信息, 不同位置或朝向依然有可能影响 DFS 的大小。BVP 有效克服了 CSI 与 DFS 对环境的依赖性: 同一手势在不同环境、不同位置、不同朝向下对应的 BVP 保持一致的变化趋势。因此, 以 BVP 为输入的感知分类器无须重复训练、收集额外数据等操作, 即可实现跨位置、跨朝向、跨环境的活动识别。

识别算法

在获得人体运动特征后, 机器学习方法可以解决人员行为识别问题。已有工作^[4]尝试利用 CSI 幅度和相位信息的变化特征作为输入, 使用 SVM 算法对环境中的是否存在运动物体进行分类预测, 从而实现运动物体的被动式检测。对于有明显时序特性的行为, 可以利用隐马尔科夫模型刻画其时序特性^[7]。然而, 传统的机器学习方法往往无法处理维度过高、结构过于复杂的特征, 具有一定的局限性。

随着深度学习浪潮的兴起, 无线感知领域对深度学习进行了大量的尝试。深度学习方法在计算机视觉领域取得了巨大的成功, 产生了许多精确可靠的识别模型。为了借用这些模型, 绝大部分已有无线感知工作不得不将原始 CSI 或从 CSI 提取到的特征以热力图或类似的形式转化为图片, 作为识别模型的输入, 并将识别模型作为黑盒来对待, 最终获得识别结果。然而, 这种方式忽视了视觉感知与无线感知在信号层面上的区别 (电磁波频率、成

像方式等), 缺乏在无线信号空间对行为活动的精细时空建模, 造成识别模型性能欠佳。Widar3.0^[2]利用卷积神经网络与循环神经网络分别挖掘输入特征 BVP 在空间维度与时间维度的特性, 对 6 种常见的人机交互手势 (推拉、横扫、拍手、滑动、画圆、画之字) 进行判别, 准确率高于 90%。STFNets^[11]没有使用常规的卷积神经网络和循环神经网络, 而是提出了一种新的基础神经网络构建模块——短时傅里叶神经网络, 来直接学习多种传感输入在频域中的特征, 并发现将信号转换到频域会极大地简化学习过程。

泛化能力差是当前识别算法的另一个主要问题。泛化能力是指模型对未知数据的适应能力。对于无线感知来说, 泛化能力至关重要——识别模型一旦训练好之后, 就应该可以应用于千家万户, 而不是到千家万户后还要用户重新采集大量数据并自己训练。为了解决跨域的识别问题, Widar3.0 通过挖掘环境无依赖的信号特征来提升泛化能力。与更专注在信号层面的 Widar3.0 不同, EI^[12]在模型设计上进行了探索——引入对抗学习的思想, 通过设计更加复杂的网络结构, 定义新的损失函数, 在充分利用无标签数据的基础上, 直接利用新的模型, 学习信号特征在不同环境中的共性表达。

数据集

高质量公开数据集对科研的推动作用是巨大的。在计算机视觉领域, ImageNet^[13]等公开数据集为深度学习算法的成功奠定了数据基础。无线感知领域对高质量公开数据集的需求更加迫切, 原因是无线射频信号的传播特性导致实验结果对设备的部署情况依赖性更高、对环境条件更加敏感。多伦多大学与斯坦福大学于 2017 年公开了一个包含 4GB 原始 CSI 信息的数据集^[14], 记录了 6 位用户的 6 种活动 (躺、摔倒、走路、跑步、坐下、起立)。威廉玛丽学院的研究团队于 2018 年公开了两种实验场景 (办公室与住宅) 下, 6 位用户用 276 种手语表达时对应的原始 CSI 数据^[15]。清华大学的研究团

队于2018年公开了1位用户在三个场景的80条运动轨迹数据,可用于被动式定位与追踪的研究^[8,9]。然而,已有的公开数据集大都存在数据量不足、应用场景少等问题。

作者团队在这方面做了进一步探索,公开了Widar3.0的手势识别数据集¹。该数据集包含原始CSI数据以及提取出的信号特征(DFS和BVP),包括75个不同场景下(包括不同位置、朝向与环境)采集的约26万组动作实例,总时长超过144小时,数据规模约为325GB。该数据集自公布一个月之内,已有200余位来自中国、美国、日本、新加坡、澳大利亚等多个国家的用户进行访问和下载。希望该数据集的推出能让无线感知领域的研究工作变得更“容易”——站在研究最前沿并做出可信的创新结果。

展望

人类对物理世界的感知进入了泛在智能的新阶段,物联网与人工智能技术共同推动人类社会从万物互联走向万物智联。2019年风靡一时的词汇“人工智能物联网”(AIoT)就是来自“人工智能”(AI)与“物联网”(IoT)两个词的结合。不同于以往,这次工业界率先预见了这一变革。华为确定了“构建万物互联的智能世界”的新愿景,小米把“AIoT”提升为核心战略,BAT等互联网公司也都提出了相似的发展战略。作为物联网与人工智能的交叉领域,无线感知正反映了这一趋势,成为当前学术界研究和工业界追寻的热点。在本领域,工业界有产品化优势,擅于将新技术融合进智能产品;学术界有创新探索的驱动力,擅于研发新技术。

其实,早在两千多年前,荀子就讨论了人类感知与智能的关系,在《荀子·正名篇》中郑重写下了“知之在人者谓之知,知有所合者谓之智”。但是,荀子肯定没能预料到“知”和“智”一旦结合起来,并且从“人”延伸到“物”,是一个怎样的泛在智能的时代。两千多年后的我们,有能力预料吗? ■



杨 铮

CCF 专业会员。清华大学副教授、博士生导师。主要研究方向为物联网,包括传感网、工业互联网、RFID、定位导航、群智感知、移动数据挖掘等。
yangzheng@tsinghua.edu.cn



郑 月

清华大学博士生。主要研究方向为无线感知。zhengyue15@mails.tsinghua.edu.cn



吴陈沐

马里兰大学助理研究员。主要研究方向为物联网,无线感知和定位。
cswu@umd.edu

参考文献

- [1] Yang Z, Zhou Z, Liu Y. From RSSI to CSI: Indoor Localization via Channel Response[J]. *ACM Computing Surveys*, 2013, 46(2):1-32.
- [2] Zheng Y, Zhang Y, Qian K et al. Zero-Effort Cross-Domain Gesture Recognition with Wi-Fi[C]// *Proceedings of the 17th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys)*. ACM, 2019.
- [3] Wu C S, Zhang F, Fan Y, and et al. RF-based inertial measurement[C]// *Proceedings of the ACM Special Interest Group on Data Communication (SIGCOMM)*. ACM, 2019.
- [4] Qian K, Wu C, Yang Z, et al. PADS: Passive Detection of Moving Targets with Dynamic Speed using PHY Layer Information[C]// *Proceedings of the 20th IEEE International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2014.
- [5] Xiong J, Jamieson K. ArrayTrack: A Fine-Grained Indoor Location System[C]// *Proceedings of the 10th USENIX conference on Networked Systems Design and Implementation*. USENIX Association, 2013.

更多参考文献: <http://dl.ccf.org.cn/cccf/list>

¹ Widar3.0 数据集下载地址: <http://tns.thss.tsinghua.edu.cn/widar3.0/>。

- [6] Kotaru M, Joshi K, Bharadia D, et al. SpotFi: Decimeter Level Localization Using Wi-Fi[J]. ACM Sigcomm Computer Communication Review, 2015, 45(4):269-282.
- [7] Wang W, Liu A X, Shahzad M, et al. Understanding and Modeling of Wi-Fi Signal Based Human Activity Recognition[C]// ACM MobiCom. ACM, 2015.
- [8] Qian K, Wu C S, Yang Z and et al. Widar: Decimeter-level passive tracking via velocity monitoring with commodity Wi-Fi[C]//Proceedings of the 18th ACM International Symposium on Mobile Ad Hoc Networking and Computing(MobiHoc). ACM, 2017.
- [9] Qian K, et al. Widar2. 0: Passive human tracking with a single wi-fi link[C]//Proceedings of the 16th Annual International Conference on Mobile Systems, Applications, and Services (MobiSys). ACM, 2018.
- [10] Xie Y X, Xiong J, Li M and et al. mD-Track: Leveraging multi-dimensionality for passive indoor Wi-Fi tracking[C]//The 25th Annual International Conference on Mobile Computing and Networking (MobiCom). ACM, 2019.
- [11] Yao S, Piao A, Jiang W, et al. STFNet: Learning Sensing Signals from the Time-Frequency Perspective with Short-Time Fourier Neural Networks[J]. ACM. 2019.
- [12] Jiang W J, et al. Towards environment independent device free human activity recognition[C]//Proceedings of the 24th Annual International Conference on Mobile Computing and Networking (MobiCom). ACM, 2018.
- [13] Deng J, Dong W, Socher R, et al. ImageNet: a Large-Scale Hierarchical Image Database[C]// 2009 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2009), Miami, Florida, USA. IEEE, 2009.
- [14] Siamak Y, Hirokazu N, Sankalp D and et al. A survey on behavior recognition using WiFi channel state information[J]. IEEE Communications Magazine. 2017, 55(10): 98-104.
- [15] Ma Y S, Zhou G, Wang S Q and et al. SignFi: Sign language recognition using WiFi[C]//Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies. 2018.