

基于双流融合网络的非接触式IR-UWB人体动作识别方法

张传宗, 王冬子, 郭政鑫, 桂林卿, 肖甫

引用本文

张传宗, 王冬子, 郭政鑫, 桂林卿, 肖甫. 基于双流融合网络的非接触式IR-UWB人体动作识别方法[J]. 计算机科学, 2025, 52(1): 221-231.

ZHANG Chuanzong, WANG Dongzi, GUO Zhengxin, GUI Linqing, XIAO Fu. [Contact-free IR-UWB Human Motion Recognition Based on Dual-stream Fusion Network](#) [J]. Computer Science, 2025, 52(1): 221-231.

相似文章推荐 (请使用火狐或 IE 浏览器查看文章)

Similar articles recommended (Please use Firefox or IE to view the article)

[基于层次化视觉注意力的富语义视频对话生成](#)

Generation of Enrich Semantic Video Dialogue Based on Hierarchical Visual Attention

计算机科学, 2025, 52(1): 315-322. <https://doi.org/10.11896/jsjcx.231100107>

[基于骨架特征的瓶颈层多尺度图卷积动作识别方法](#)

Bottleneck Multi-scale Graph Convolutional Network for Skeleton-based Action Recognition

计算机科学, 2024, 51(11A): 231000073-5. <https://doi.org/10.11896/jsjcx.231000073>

[基于视频理解的教学过程感知与分析](#)

Perception and Analysis of Teaching Process Based on Video Understanding

计算机科学, 2024, 51(10): 56-66. <https://doi.org/10.11896/jsjcx.240400109>

[基于可穿戴设备的开放集动作识别技术研究](#)

Study on Open Set Activity Recognition Technology Based on Wearable Devices

计算机科学, 2024, 51(4): 291-298. <https://doi.org/10.11896/jsjcx.230300158>

[基于视觉的神经网络三维动态手势识别方法综述](#)

Review of Vision-based Neural Network 3D Dynamic Gesture Recognition Methods

计算机科学, 2024, 51(4): 193-208. <https://doi.org/10.11896/jsjcx.230200205>

基于双流融合网络的非接触式 IR-UWB 人体动作识别方法

张传宗 王冬子 郭政鑫 桂林卿 肖甫

南京邮电大学计算机学院、软件学院、网络空间安全学院 南京 210023

江苏省无线传感网高技术重点实验室 南京 210023

(zhangcz@njupt.edu.cn)

摘要 随着智能感知技术的飞速发展,人机交互(Human Computer Interaction, HCI)领域迎来了全新的发展态势。传统的人机交互方法主要依赖可穿戴设备或者摄像头采集用户的行为数据,虽然识别精准,却存在不小的局限性。具体而言,可穿戴设备会给用户带来额外的使用负担,而基于摄像头的方案不仅会受到环境光线的影响,还会涉及用户隐私的泄露,这些因素均限制了其在日常生活中的广泛应用。为了突破这些限制,实现精确的、非接触式人机交互应用,利用无线射频(Radio Frequency, RF)领域中脉冲超宽带(Impulse Radio Ultra-Wideband, IR-UWB)所具有的高灵敏度和精细空间分辨率等优势,提出了一种基于双流融合网络的非接触式人体动作识别方法。该方法捕获目标运动所导致的时域信号变化,并通过时对域特征进行多普勒频移变化,提取到对应的频域特征。在此基础上,构建了一个融合多维卷积神经网络(Convolutional Neural Networks, CNNs)和 GoogLeNet 模块的双流网络模型,以实现高精度的动作识别。通过广泛的实验测试,结果表明所提方法对 8 种常见人体动作的平均识别准确率达到 94.89%,并且在不同的测试条件下均能保持超过 90% 的识别准确率,进一步验证了所提方法的鲁棒性。

关键词: 人机交互;无线感知;脉冲超宽带;动作识别

中图分类号 TP391

Contact-free IR-UWB Human Motion Recognition Based on Dual-stream Fusion Network

ZHANG Chuanzong, WANG Dongzi, GUO Zhengxin, GUI Linqing and XIAO Fu

School of Computer Science, Nanjing University of Posts and Telecommunications, Nanjing 210023, China

Jiangsu High Technology Research Key Laboratory for Wireless Sensor Networks, Nanjing 210023, China

Abstract With the rapid development of intelligent sensing technology, the field of human computer interaction(HCI) has entered a new era. Traditional HCI methods, predominantly reliant on wearable devices and cameras to collect user behavior data, have significant limitations despite their precise recognition capabilities. Wearable devices, for instance, impose additional burden on users, whereas camera-based solutions are susceptible to ambient lighting conditions and pose significant privacy concerns. These challenges considerably restrict their applicability in daily life. To solve these challenges, we utilize the exceptional sensitivity and spatial resolution of impulse radio ultra-wideband(IR-UWB) in the field of radio frequency(RF) to propose a novel and contact-free method for human motion recognition based on a dual-stream fusion network. This method adeptly captures the temporal signal variations caused by target movements and extracts the corresponding frequency-domain features by analyzing Doppler frequency shift(DFS) changes on the time-domain signals. Subsequently, a sophisticated dual-stream network model, integrating multi-dimensional convolutional neural networks(CNNs) and GoogLeNet modules, is developed to facilitate precise action recognition. Through extensive experimental tests, the results show that the proposed method achieves an average accuracy of 94.89% for eight common daily human actions and maintains an accuracy of over 90% under varying test conditions, thereby validating the robustness of the proposed method.

Keywords Human computer interaction, Wireless sensing, Impulse radio ultra-wideband(IR-UWB), Motion recognition

到稿日期:2024-04-15 返修日期:2024-07-02

基金项目:国家杰出青年科学基金(62125203);国家自然科学基金(61932013);江苏省重点研发计划重大科技示范项目(BE2022798);南京邮电大学引进人才科研启动基金(NY223041)

This work was supported by the National Science Fund for Distinguished Young Scholars of China(62125203), National Natural Science Foundation of China(61932013), Key Research and Development Program of Jiangsu Province(BE2022798) and Natural Science Research Start-up Foundation of Recruiting Talents of Nanjing University of Posts and Telecommunications(NY223041).

通信作者:郭政鑫(guozx@njupt.edu.cn)

1 引言

随着科技的飞速进步,人体动作识别和行为感知技术在各个领域中都扮演着重要的角色,如智能家居、安防监控、医疗护理和人机交互。这些技术意义重大,有着广泛的应用前景。目前,智能家居在实际运用中部分功能仍需通过手机、平板等设备上的互联网页面控制来实现,并未达到真正智能化的水平。将人体动作识别和行为感知等技术融入智能家居系统,以实现系统的智能化与高效化,仍是一个需要深入研究的课题。

常见的人体动作识别和行为感知技术方案可分为3类:基于视觉的方法、基于可穿戴传感器的方法以及基于无线信号的方法。基于计算机视觉^[1]的系统通常利用分段摄像机将视频捕获到图像序列中,并采用机器学习算法来识别行为。尽管这些方法在准确度方面具有明显优势,并已被广泛应用于许多公共场所,但这类传感器对存储空间和信息处理能力要求较高,同时存在视野盲区、对光照条件敏感以及个人隐私易泄露等缺点。另外,这些设备成本昂贵,部署在公共区域容易遭到破坏。基于可穿戴传感器的系统利用多个人体运动传感器(加速度计、陀螺仪等)感应身体的运动^[2]。穿戴式传感方法善于感知微妙的动作,因此可以实现对动作的精细识别。然而,用户被迫佩戴各种笨重且昂贵的传感器,这极大地限制了他们的自由度并降低了可扩展性,且硬件存在便携性差、容易损坏、误报率高和识别精度低等问题^[3]。基于无线信号的方法利用 RFID(Radio Frequency Identification)、蓝牙、Wi-Fi 以及雷达等技术完成用户无传感器的感知。其中,RFID 能够利用射频信息来自主地辨认出目标物体,同时也能收集到所有的相应数据^[4],这种辨认方式不需要任何人为介入,但需要专用的阅读器;蓝牙信号便捷但覆盖范围小(通常为 10~15m),这些约束的存在^[5]限制了其应用。相比之下,Wi-Fi 在我们的日常生活中无处不在,其通过商用设备就可以获取,成本可控,因此成为人们的关注热点。基于 WiFi 的动作识别可分为两类,即基于 RSSI^[6]和 CSI^[7]。该方案虽然摆脱了穿戴问题,但是容易受到周围环境的影响,稳定性较差。

与以上技术相比,超宽带(Ultra-Wideband, UWB)雷达通过发射低频段的电磁波来实现人体目标行为识别(人体目标的肢体动作信息可从雷达回波中的幅值、相位等参数中进行提取)。基于超宽带(UWB)雷达的人体动作识别技术的原理是雷达回波信号中包含了人体目标的运动信息,经过相关处理和特征提取后可获得人体动作的特征信息,再通过识别算法对人体动作进行分类识别。UWB 雷达具有低功耗、不侵犯隐私、不受视距环境影响、探测距离远、分辨率高等特点,符合人体动作识别的要求,因此研究基于 UWB 雷达的人体动作识别技术不仅在理论上具有研究意义,而且在实际应用中也具有重要价值。

综上所述,常见人体行为识别方法存在着高度依赖环境、辨别特征能力弱等问题。针对上述问题,本文基于双流融合网络的 IR-UWB 非接触式人体动作识别系统,分别从时域和频域方面展开分析并提取特征,通过双流网络模型进行训练,

最后通过 softmax 分类网络层实现人体动作识别。首先,通过 IR-UWB 雷达获取人体动作信号数据,并通过静态杂波抑制、动态噪声滤波等方法对信号数据进行预处理;然后通过经验性实验分析不同的人体运动导致信号在时域和频域中呈现出的不同特性,从而提取人体动作的时域、频域特征;再通过双流网络进行训练,经过多策略融合,得到适用于 IR-UWB 人体动作识别的融合特征,并通过 softmax 分类器对人体动作进行识别分类;最后综合多因素对实验结果进行分析。大量的实验测试表明,本文提出的方法对 8 种常见人体动作的平均识别准确率超过 94%。

本文第 2 章介绍了 IR-UWB 动作识别的相关工作;第 3 章介绍了基于 IR-UWB 的动作识别系统原理,给出了 IR-UWB 雷达信号建模过程,并对信号进行时、频域特征分析;第 4 章详细说明了本文提出的基于双流融合网络的 IR-UWB 非接触式人体动作识别方法的系统实现过程;第 5 章对本文提出的动作识别系统开展大量的实验测试,并分析系统性能;最后总结全文。

2 基于 IR-UWB 动作识别的相关工作

非接触式动作识别技术指基于无线射频设备、雷达传感器等无需和人体直接接触的设备来获取人体目标的运动数据。根据所采用的不同射频信号,目前非接触式动作识别的相关工作可以分为基于 Wi-Fi、基于 RFID、基于 mmWave 和基于 IR-UWB 的动作识别工作。

2.1 基于无线射频信号的动作识别

利用无线射频信号进行人体动作识别具有不受光照影响和尊重个人隐私等独特优势,这些优势是其他识别技术无法比拟的,在非接触式人体动作识别领域具有巨大的发展潜力。现有的无线射频信号动作识别技术主要有 Wi-Fi 技术、RFID 技术、mmWave 技术和 IR-UWB 技术等。

Wi-Fi 技术主要通过分析用户日常活动中的不同动作对空间中 Wi-Fi 信号的扰动,实现动作识别任务。Yan 等^[8]提出的 WiAct 通过提取 CSI 数据的方差特征完成人体动作感知任务。为了识别信号中的动作部分,他们提出了一种全新的自适应活动切割算法,并从多根 Wi-Fi 天线中提取了与活动相关的多普勒频移数值作为输入,送入极限学习机分类器进行处理,以识别 10 种不同活动的情况。该算法的整体识别准确率达到 94.20%。Zhou 等^[9]提出了 LT-WioB 感知系统,其中的轻量级卷积模块能用少量数据样本进行训练和完成识别任务,整体平均识别准确率达到 96.1%,并很大程度上减少了计算开支。Xiao 等^[10]提出了一种基于 CSI 的跨区域人类活动识别框架,该框架设计了一种自适应标签平滑方法用于提高识别模型的泛化能力,提高了 Wi-Fi 在跨场景应用下的感知能力。

Xia 等^[11]提出了一种基于 RFID 技术的非接触式姿势识别解决方案 RF-Care,通过分析 RFID 标签阵列中 RSSI 值的变化来捕获环境中的活动信息,该系统使用 9 个标签和 1 个阅读器,动作识别准确率达到 98%。Oguntala 等^[12]提出一个两层的完全连接的长短期记忆递归神经网络模型,通过接收

无源 RFID 标签的 RSSI 信息获取详细的活动信息,准确率达到 98.18%。Li 等^[13]提出了一种 R&P 动作识别方案,通过提取部署在环境中的 RFID 标签的相位和 RSSI 值进行非接触式的活动识别,该方案可以实现平均 87.9%的准确率。

Li 等^[14]利用 mmWave 雷达对采集到的人体运动时域数据作短时傅里叶变换得到频域信息,并使用双向二维 PCA 在频域提取微多普勒特征,使用 SVM 分类器最终实现了 91.9%的动作识别效果。Kim 等^[15]使用原始的微多普勒频谱图作为 DCNN 的数据输入进行训练和分类,在 7 种不同人类动作中实现了 90.9%的分类准确率。之后, Kim 等^[16]又利用雷达反射回波形成的点云信号来对人类活动进行分类并结合二维深度卷积递归神经网络捕捉点云在不同时刻的特征,其分类准确率超过 97%。Abedi 等^[17]提出了一个将 mm-Wave 雷达用于家庭环境中实时监测的非接触式身体活动识别系统,选择门控循环网络模型进行实时部署,其对家庭体育活动的分类总体准确率为 93%。

现有的 Wi-Fi 设备具有良好的普适性和通用性,但是目前基于 Wi-Fi 的 CSI 信号需要利用特定的网卡才能收集,对部署产生了一定的限制。RFID 技术射频标签非常微小且轻量化,但是其贴附式的使用方法使其并没有完全摆脱接触式的感知范围。目前基于 mmWave 雷达的感知技术可以取得精度最高的细粒度人员感知,然而现有的商用 mmWave 雷达系统的制造成本和功耗较高,不适合大规模部署。与上述系统相比,商用 IR-UWB 雷达具有精度高、无累计误差、成本低、功耗低和抗干扰能力强的优势,已被应用于智能手机、笔记本电脑等产品当中,适合在室内场景中大规模部署,以实现高精度的目标人体动作识别等工作。

2.2 基于 IR-UWB 雷达传感器的动作识别

由于其具有高分辨率和强抗多径干扰能力等优势,IR-UWB 雷达被提出并应用于人体动作识别技术研究。2012 年 Bryan 等^[18]首次将 UWB 雷达用于人体动作识别领域,采用主成分分析算法提取 UWB 雷达信号的主要分量部分作为特征,并结合 SVM 对 8 种典型的人体动作进行分类识别,平均准确率超过 85%。Ahmed 等^[19]提出将 UWB 雷达信号转换为三维图像模式的方法,使用 GoogleNet 中基于初始模块的变体来分析图像中的模式特征,针对 8 种不同的手势,实现了 95%的分类准确率。Ding 等^[20]针对人体动作的原地姿态,利用集成学习方法对 UWB 雷达信号数据进行分层处理,所采用的 S-K 近邻模型对 6 种人体姿态的识别率达到 94.4%。Jiang 等^[21]利用小波变换分解 UWB 雷达快慢时间矩阵,使其呈现更多的局部细节信息,进一步综合离散反射图的特征,实现对人体动作的识别。Jiang^[22]利用 WPD 提取动作的 UWB 雷达信号能量分布特征,并用改进的混沌自适应遗传算法优化 SVM 分类模型,实现了超过 97%的人体动作识别效果。本文基于以上相关研究工作,从 IR-UWB 雷达信号中提取时域、频域的动作特征信息,构建双流网络模型,根据不同策略进行特征融合,开展人体动作识别工作,设计了高效的分类算法,提高了 IR-UWB 动作识别的准确度。

3 基于 IR-UWB 的动作识别系统

本章首先介绍 IR-UWB 雷达信号的建模过程,然后分析不同的人体运动导致信号在时域和频域中呈现出不同特征。

3.1 IR-UWB 雷达信号建模

在室内环境中,电磁波在传播过程中会在多种障碍物如墙壁、家具和办公用品等表面发生反射和散射,从而使得接收的信号是所有路径信号的叠加。当感知范围内存在执行不同动作的人员时,其躯干和肢体的运动会导致 IR-UWB 雷达信号的传播路径长度发生改变,从而使得回波信号中包含相应的运动特征信息。首先对 IR-UWB 雷达信号进行建模,分析多径的传播时延等信息,以实现空间中人体动作的感知。

IR-UWB 雷达发射的基带信号 $s(t)$ 为高斯脉冲。在使用同相单载波进行上变频操作后,发射信号可以表达为:

$$x_n(t) = s(t - nT_q) \cdot \cos(2\pi f_\lambda(t - nT_q)) \quad (1)$$

其中, f_λ 表示载波中心频率,是电磁波传播速度 c 和载波波长 λ 的比值,中心频率越低,波长越长,信号的穿透能力越强; f_q 表示脉冲的重复频率; T_q 表示脉冲重复间隔; n 表示发射脉冲的序号; q 表示多径传播路径的序号。

假设脉冲信号在传播过程中存在包括直射路径和反射路径在内的共 Q 条不同的路径,则无线信道对第 n 个脉冲的冲击响应可表示为:

$$h_n(t) = \sum_{q=1}^Q a_q \sigma(t - \tau_q - \tau_q^D(nT_q)) \quad (2)$$

其中, a_q 表示第 q 条多径的信道增益, τ_q 是第 q 条多径的传播时延, $\tau_q^D(nT_q)$ 表示由第 q 条多径的多普勒频移引起的时延。目标距离雷达为 D_q , 运动速度为 v_q 时,存在关系: $D_q = 2 \cdot c \cdot \tau_q$, $v_q = c \cdot \tau_q^D(nT_q) / 2nT_q$, 其中 c 表示光速。目标距离 D_q 是信号传播路径长度 $c \cdot \tau_q$ 的两倍。当感知区域内存在多个感知目标时,只有当雷达到两个目标的距离差大于距离分辨率 ΔD 时才能够被分辨出来。对于一个发射脉冲带宽为 B 的雷达,其距离分辨率为 $\Delta D = c / 2B$ 。雷达的距离分辨率仅取决于信号的带宽,带宽越大,距离分辨率就越高,更有助于感知人体的微小运动。

接收信号 $r_n(t)$ 是发射信号与信道冲激响应的卷积。

$$r_n(t) = h_n(t) * x_n(t) \\ = \sum_{q=1}^Q a_q s(t - nT_q - \tau_q - \tau_q^D(nT_q)) \cdot \cos(2\pi f_\lambda(t - nT_q - \tau_q - \tau_q^D(nT_q))) + \omega(t) \quad (3)$$

其中, $\omega(t)$ 表示高斯噪声。

在经过 IQ 下变频操作后,接收端的基带输出信号为:

$$r_n^b(t) = \sum_{q=1}^Q a_q \cdot e^{j2\pi f_\lambda(\tau_q + \tau_q^D(nT_q))} \cdot s(t - nT_q - \tau_q - \tau_q^D(nT_q)) + \omega(t) \quad (4)$$

当对基带输出信号以 T_s 为间隔进行采样后,第 k 个采样点可表示为:

$$r_n^b[k] = \sum_{q=1}^Q a_q \cdot e^{j2\pi f_\lambda(\tau_q + \tau_q^D(nT_q))} \cdot s[kT_s - nT_q - \tau_q - \tau_q^D(nT_q)] + W[k] \quad (5)$$

IR-UWB 雷达将时序的基带采样信号转换成矩阵形式,其发射的信号为周期性脉冲序列。同一个脉冲序列的接收

回波被存储在相同的列中,不同脉冲序列的接收回波按行堆叠。因此,将从行的方向看过去的维度定义为慢时间维度。由于列与列之间的数据采样间隔通常大于脉冲持续时间,因此将从列的方向看过去的维度定义为快时间维度。对 n 个脉冲回波的基带信号进行数字采样后可得到快慢时间矩阵。

$$R = \begin{bmatrix} r_1^b[1] & \cdots & r_n^b[1] \\ \vdots & \ddots & \vdots \\ r_1^b[k] & \cdots & r_n^b[k] \end{bmatrix} \quad (6)$$

其中, n 表示脉冲序号, k 表示采样点序号。如图 1 所示,快慢时间矩阵是对环境中的多径信息的整体刻画,反映了环境中物体距离随时间的变化关系。

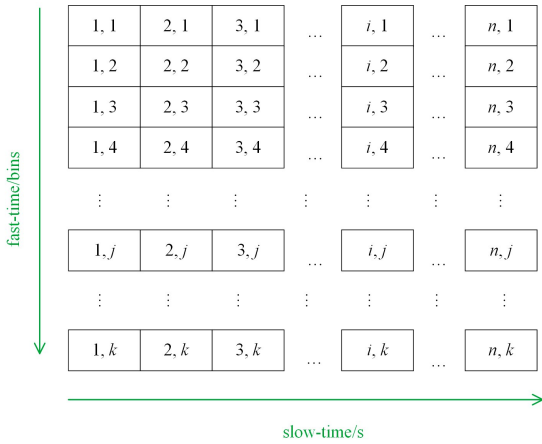


图 1 接收信号谱(接收信号矩阵)

Fig. 1 Received signal spectrum(received signal matrix)

3.2 IR-UWB 信号时域和频域特征

IR-UWB 的快慢时间矩阵反映了空间中物体与雷达设备的距离信息。由于人体动作会导致持续性的回波信号变化,因此可以使用时频分析方法提取时域中与人体动作相关的

特征。此外,人体动作导致躯干和肢体以不同的速度进行运动,通过对快慢时间矩阵的多普勒频移估计,可以提取与人体动作相关的频域特征。

1)时域特征。利用 IR-UWB 雷达技术可以将人体分解成许多离散的点,当人体运动时,四肢和躯干的散射点与雷达的径向距离和空间位置会随着时间而变化。图 2 展示了人体在某一距离点的 6 种不同动作(打拳、踏步、转身、推、拉和下蹲)的人体散射点信号幅值随时间的变化趋势。在该距离点,不同类别的人体动作在信号幅值变化的趋势上存在一定差异,而相同类别的动作则呈现出相似的变化规律。在打拳过程中,2 次快速出拳的动作对应有 2 个时间上人的手部肢体到达该距离点位置,对应图 2(a)中两处峰值点;在踏步过程中,人的腿部、胳膊肢体交替摆动多次,多次到达该距离点形成多个峰值;在转身过程中,基于人的形体构造,当人正对方向与雷达两次垂直时,形成幅度峰值;在推和拉的过程中,信号幅值的变化分别体现了手部肢体靠近、远离该距离点的变化;在下蹲的过程中,人体会有前倾的现象,之后维持蹲下的状态,因此信号幅值中体现了这一变化过程。

2)频域特征。由于多普勒效应的存在,即当物体和雷达之间存在相对运动时,雷达接收到的脉冲信号频率和发射的信号频率存在特定的偏移,因此可以通过多普勒效应来捕捉运动目标的多普勒频率变化,从而计算出其运动速度,借此推测运动状态。图 3 展示了打拳、踏步、转身和下蹲这 4 种动作的动作信号频谱,不同的动作会对到达信号产生影响,从而导致不同的频率变化。理想室内空旷场景中,当人员静止时,几乎不存在信号频率变化;当人体运动时,信号频谱中产生一定宽度的“亮带”,对应于不同的人体动作,动作幅度越大、速度越快,接收信号的频率变化就越大,这条“亮带”的宽度也就越宽、能量越强。

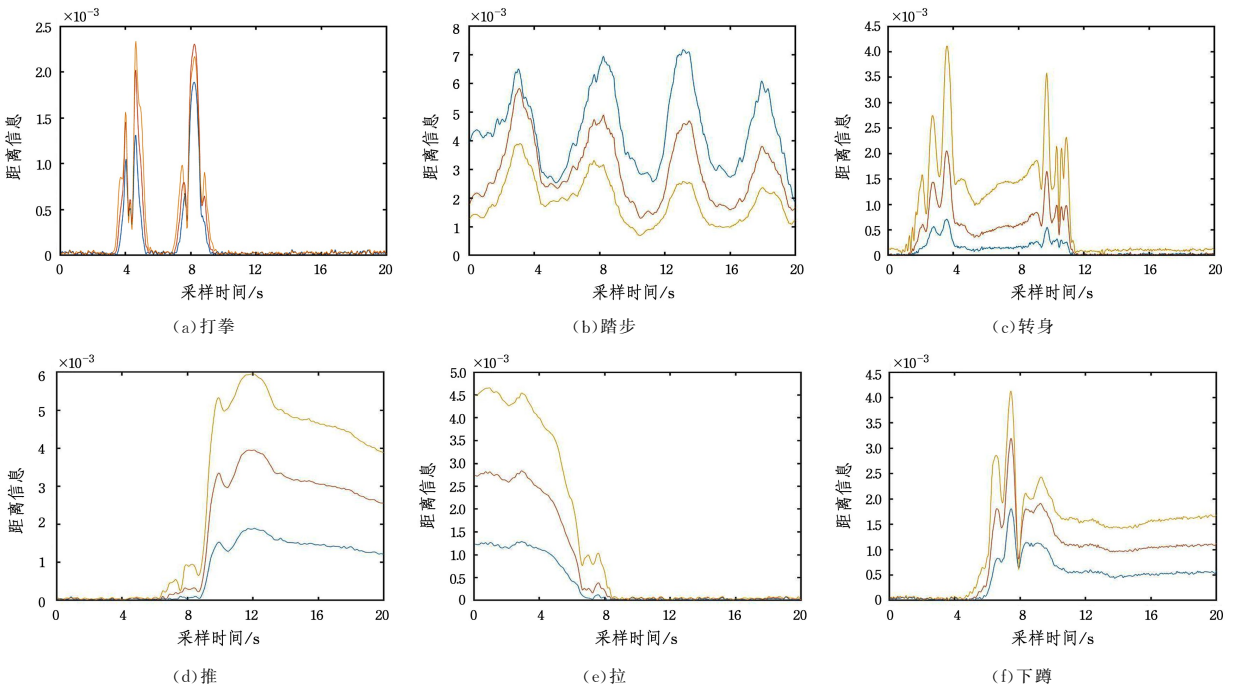


图 2 人体动作时间-信号幅度特征

Fig. 2 Time-Signal amplitude features of human motions

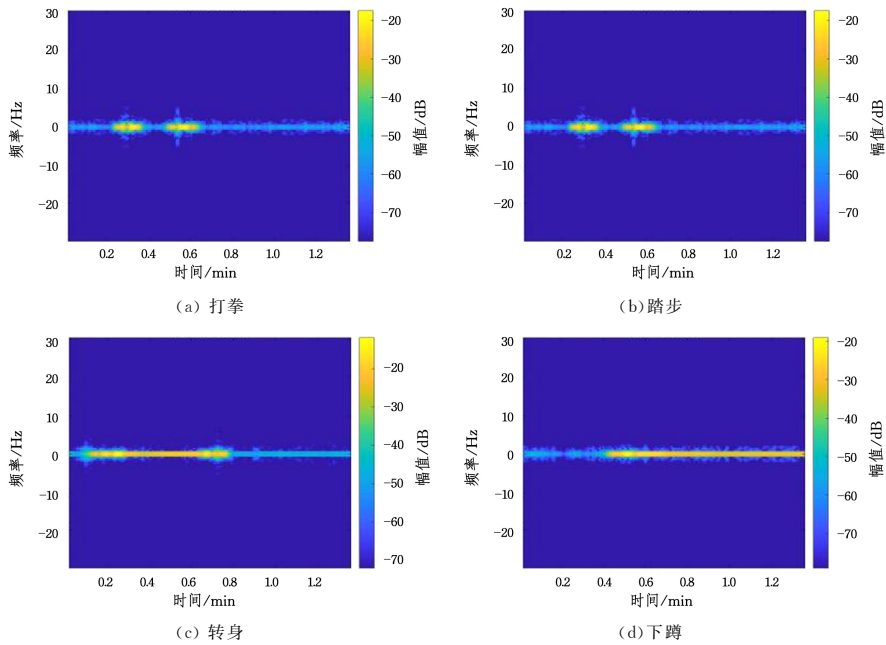


图3 人体动作脉冲信号频谱图

Fig. 3 Spectrum diagram of human motion pulse signal

4 系统实现

IR-UWB 非接触式人体动作识别的工作中,首先需要通过 IR-UWB 雷达获取人体动作脉冲回波数据,并通过静态杂波抑制、动态噪声滤波等方法对数据进行预处理,消除硬件环境和人体运动导致的噪声干扰;然后,通过分析人体行为动作导致脉冲信号产生的时间延迟和频率偏移,提取回波信号中的时域和频域特征,并进行归一化处理;最后,融合多维卷积神经网络和 GoogLeNet 模块^[23]构建双流网络模型,实现基于 IR-UWB 的非接触式人体动作识别。由于人体不同的行为动作会导致脉冲信号的传播路径长度以及多普勒频率偏移发生相应的变化,因此,通过提取回波信号的时域和频域特征,并使用双流融合网络模型来识别人体动作是本文工作的核心。

为了实现上述人体动作识别工作,本文开发了基于双流融合网络的 IR-UWB 非接触式人体动作识别系统,系统框架如图 4 所示。该系统主要由 4 个模块构成,分别是信号预处理模块、行为特征提取模块、模型训练模块和行为识别模块。首先是信号预处理模块,本系统通过单个 IR-UWB 雷达获取空间中人体运动的反射信号,通过静态杂波抑制对信号中由硬件缺陷和环境噪声导致的直流分量和静态杂波进行消除,并采用 Hampel 和 Savitzky-Golay 滤波消除人体运动过程中产生的动态噪声;然后,通过解析不同人体运动导致的信号传播路径长度变化和 Doppler 频率偏移,提取出反射信号中的时域和频域特征;最后,融合多维卷积神经网络和 GoogLeNet 模块构建双流网络,对时域、频域特征进行映射,通过 softmax 算法分类,最终实现人体动作识别。

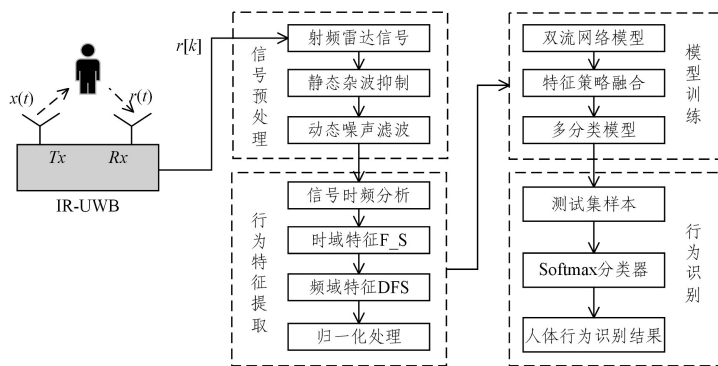


图4 基于双流融合网络的 IR-UWB 非接触式人体动作识别系统框架

Fig. 4 Framework of IR-UWB contact-free human motion recognition system based on dual-stream fusion network

4.1 IR-UWB 信号预处理

由于存在设备缺陷和环境噪声,IR-UWB 的回波信号中不可避免地存在大量干扰。同时,由于 IR-UWB 雷达的收发天线集成在同一张芯片上,不可避免地会存在反射泄露和近距离回波进而导致回波信号中存在直流分量。从

图 5 中可以看到,在慢时间维度接近 0 的时间序列中能量峰值普遍较高,而这部分直流分量会完全淹没其中的反射信号。与此同时,由于多径效应的存在,复杂的空间环境中的墙壁、桌椅和办公用品等障碍物会使得脉冲信号发生反射、折射和衍射,且这部分信号分量叠加在快慢时间

矩阵中,从而导致人体运动反射的信号失真。这两种信号干扰都属于静态杂波。

为了消除设备缺陷和环境噪声导致的干扰,从而准确地提取出人体运动反射的信号,本节首先对 IR-UWB 快慢时间矩阵进行静态杂波抑制。设备缺陷导致的直流分量在每次雷达启动时都保持相对稳定。同时,本文实验中设定人体目标到雷达设备的距离是固定的,这样每一条脉冲回波上人体目标的时延就是不变的,对所有的脉冲回波计算均值就可以得到环境参考的接收脉冲,然后将接收到的每个脉冲信号与参考信号相减以获得经过静态杂波抑制处理后的回波信号。此操作的核心思想在于进行平均处理后再相减。具体操作步骤为:首先对所有接收到的脉冲信号进行平均处理以得到参考信号,然后将接收到的每个脉冲信号减去参考信号就可以得到 t 时间的目标回波信号。

$$\overline{R(f,t)} = R(f,t) - \frac{1}{T} * \sum_{i=0}^T R(f,t) \quad (7)$$

其中, T 是信号脉冲总数,这里 $T=96$,由硬件设备决定; f 是载波频率。

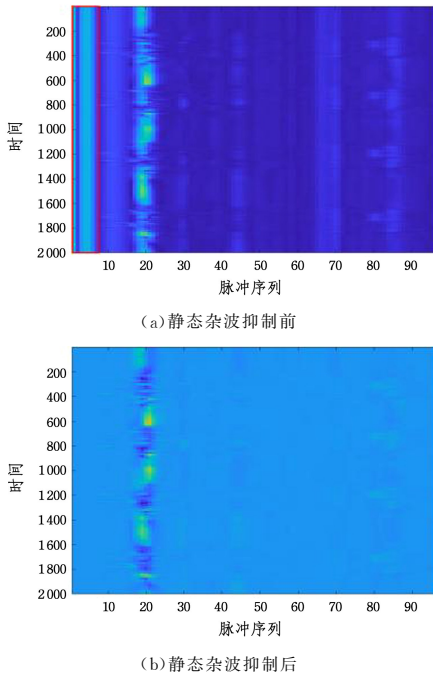


图5 静态杂波抑制效果图

Fig. 5 Static clutter suppression renderings

环境中的噪声除了静态和直流分量导致的信号反射成分外,还有空气震动、人体抖动以及环境晃动导致的动态噪声。而且信道本身的不稳定性也会随机产生相应的热噪声,从而导致采集到的信号中存在异常值。为了消除序列中的异常值,首先采用 Hampel 滤波器来检测和消除序列中的异常值,然后采用 Savitzky-Golay 滤波器来使整个完整的时序序列更加平滑,从而有利于提取人体运动信号。

Hampel 滤波是一种基于中位数的异常值检测和替换方法,通过对信号进行中位数滤波,并计算每个数据点与中位数之间的距离来识别异常点。如果距离大于某个阈值,则该数据点被视为异常点,并用中位数替换。具体 Hampel 滤波方法为:

$$\varphi(n) = \begin{cases} \varphi(n), & \varphi(n) - \varphi_m(n) \leq \sqrt{\frac{1}{n} \sum_{i=n-3}^{n+3} (\varphi(i) - \varphi_m(i))^2} \\ \varphi_m(n), & \varphi(n) - \varphi_m(n) \geq \sqrt{\frac{1}{n} \sum_{i=n-3}^{n+3} (\varphi(i) - \varphi_m(i))^2} \end{cases} \quad (8)$$

其中, n 为观察时间内的帧数, φ_m 为 Hampel 滤波器窗口中的相位中值。

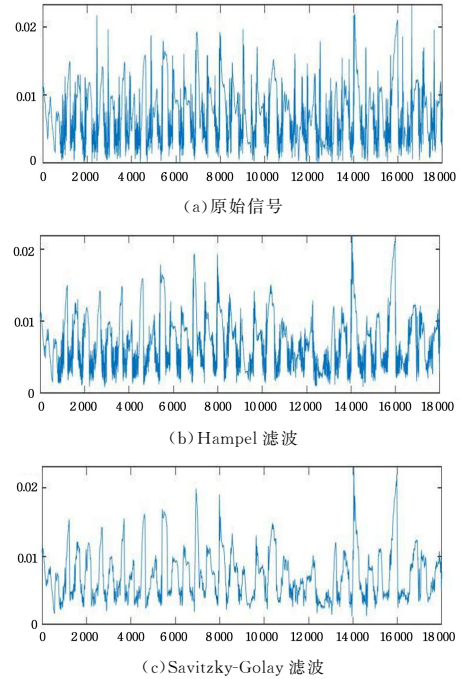


图6 信号数据异常值消除和数据平滑

Fig. 6 Signal data outlier removal and data smoothing

Savitzky-Golay 滤波是一种用于数据平滑的去噪技术,其原理是对曲线局部特征进行多项式拟合。通过最小二乘法来确定权重系数,从而实现移动窗口的加权平均操作。经过滤波处理后的数据能够有效保留局部特征。具体的 SG 滤波方法为:

$$r_{k,SG} = \bar{r}_k = \frac{1}{P} \sum_{i=-w}^{+w} r_{k+i} p_i \quad (9)$$

其中, p_i/P 为平滑系数,由原始 IR-UWB 信号脉冲序列经过最小二乘拟合求得; w 为 SG 滤波的滑动窗口长度,本文中的 IR-UWB 雷达采样率设定为 60 Hz,对应的滑窗长度为 60。图 6(b) 和图 6(c) 是单一脉冲信号经过 Hampel 滤波和 Savitzky-Golay 滤波后的信号时间-幅值序列,明显可以看出,与原始信号序列相比,其既消除了异常值,同时也得到了平滑的数据序列。

4.2 动作特征提取

4.2.1 时域特征提取

IR-UWB 雷达回波信号反映了感知区域内的所有反射信号,而人体动作信号往往局限在人员位置附近,如图 7(a) 所示,运动信号聚集在“山峰”附近。因此在进行特征提取前,需要先定位动作主体的所在位置。相较于人体运动信号,感知空间中的背景环境导致的回波幅值变化较小,导致 IR-UWB 回波信号在运动主体附近的快时间维度中出现较大的幅值波动,因此设置动态阈值 A_{th} 定位动作主体。具体来说,首先

设置动态滑动窗口,然后对窗口内的幅值波动差分求和 A_{sum} 。当前时间窗口内存在动作主体时,IR-UWB 接收信号的强度会出现较大波动。当波动的幅度超过预先定义的阈值时,即 $A_{\text{th}} > A_{\text{sum}}$,记录当前时间窗口作为检测到的动作起点;然后持续记录后续时间窗口内的信号变化,当波动的幅度小于预先定义的阈值时,则视为动作停止,并记录当前时间窗口作为动作的结束时间。

为了准确获取人体的反射信号,针对去除噪声后的快慢时间矩阵,首先采用图像算法中的 Hot-Zone^[24]算法来获取用户的反射区域,图 7(b)为采用 Hot-Zone 算法提取到的人体反射区域图。Hot-Zone^[24]算法主要根据贝叶斯空间误差模型对反射区域图中的高频反射边界进行分离,进而获取目标人体的反射信号分量 $r(t)$ 。由于不同的动作持续时间和动作幅度覆盖范围有所差异,因此分离后动作的快慢时间矩阵中的快慢时间维度存在差异,这会对后续模型的输入工作造成困扰,需要提前调整样本矩阵维度。本文采用 MATLAB 函数 `imresize` 对样本矩阵进行维度归一化。

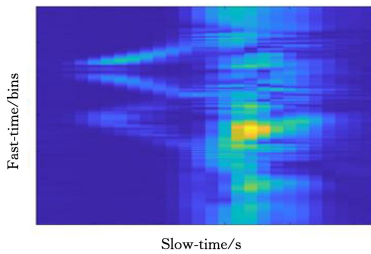
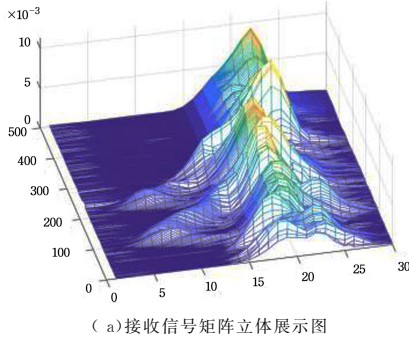


图 7 接收信号快慢时间矩阵图

Fig. 7 Diagram of fast-slow time matrix of received signals

4.2.2 频域特征提取

当雷达和人体存在相对运动时,由于多普勒效应的存在,雷达接收到的信号频率与信号发射频率之间存在频率偏移。计算多普勒频率偏移并将其作为该动作的频域特征。首先提取雷达回波信号中当前动作的慢时间序列 $x(t)$ 。

$$x(t) = s(t - nT_p) \cdot \cos 2\pi f_0(t - nT_p) + \Phi \quad (10)$$

其中, f_0 是发射频率, Φ 为初始相位。

雷达接收到的回波信号 $x_r(t)$ 为:

$$x_r(t) = s(t - nT_p) \cdot \cos(2\pi f_0(t - nT_p - t_r) + \Phi) \quad (11)$$

其中, $t_r = 2R(t)/c$ 为回波信号相对于发射信号的延迟, $R(t)$ 为雷达与目标之间的距离, c 为光速。

如果目标静止,则 $R(t)$ 为常数,且令 $R(t) = R_0$ 。脉冲回波与发射信号之间的固定相位差为 $2\pi f_0(t_r) = 2\pi f_0 \cdot 2R_0/c = 4\pi R_0/\lambda$,其中 λ 为波长。

如果目标与雷达之间存在相对运动,距离 $R(t)$ 随时间变化,设目标相对雷达的径向运动速度为 v_r ,则 $R(t) = R_0 - v_r t$ 。故回波延迟为:

$$t_r = \frac{2R(t)}{c} = \frac{2(R_0 - v_r t)}{c} \quad (12)$$

可得回波信号与发射信号之间的相位差为:

$$\Delta\Phi = -2\pi f_0 \frac{2(R_0 - v_r t)}{c} \quad (13)$$

对应产生的频率差为:

$$f_d = \frac{1}{2\pi} \frac{d\Delta\Phi}{dt} \quad (14)$$

其中, f_d 为多普勒频移。图 8 为人体动作信号提取的 DFS 谱图。当目标朝向雷达运动时,多普勒频移呈正数,接收信号频率高于发射信号;当目标远离雷达运动时,多普勒频移为负数,接收信号频率低于发射信号。不同人体动作的信号 DFS 谱图具有显著的特征差异。

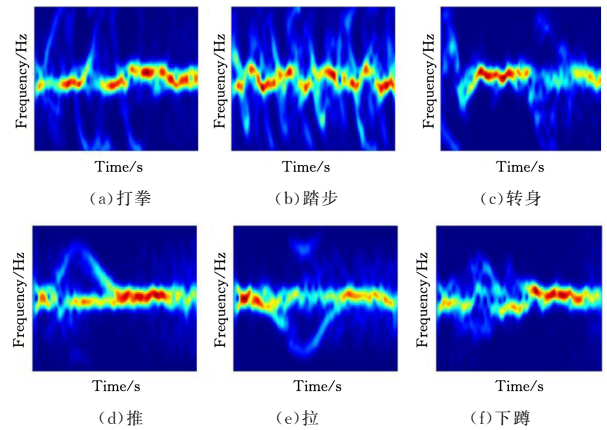


图 8 人体动作信号 DFS 谱图

Fig. 8 DFS spectrum of human motion signals

4.3 双流融合网络模型

基于人体动作 IR-UWB 信号的时频特征分析,本文提出一种融合时域、频域特征的双流网络模型,该模型包括时域流和频域流。时域流采用常见的多维卷积神经网络 CNNs,频域流采用多个 GoogLeNet 模块的新组合。通过两个分流网络对时域、频域特征进行进一步提取,并采用多策略融合方式,得到人体动作的融合特征,配合 softmax 决策函数完成动作标签的分类。

4.3.1 时域流网络

本文时域流网络中使用的卷积神经网络 CNNs 结构包括卷积、池化和局部响应归一化(LRN)等层次,其输入为 IR-UWB 接收信号快慢时间矩阵。在时域流 CNNs 中,卷积层通过多个卷积核提取时域特征,生成特征列组;池化层负责下采样操作,以合并特征并减少维度;LRN 层被用来创建竞争机制,对局部神经元的活动进行调节,强调高响应值,抑制低反馈值,同时增强泛化能力。

本文设置的时域流 CNNs 输入层尺寸为 224×224 ,一共有 6 个卷积层,3 个池化层。第 1 个卷积层的尺寸为 9×9 ,步长为 3;第 2 个卷积层的尺寸为 7×7 ,步长为 3;第 3-5 个卷积层的尺寸为 5×5 ,步长为 2;第 6 个卷积层的尺寸为 3×3 ,步长为 1;池化层的尺寸都为 3×3 ,步长为 2。为了保证网络

结构的非线性,在所有卷积层的后面加入一个非线性激活函数 Relu,同时,在两个池化层后连接一个 LRN 层。

4.3.2 频域流网络

为了实现更好的提取特征,经典 CNN 架构一般会增加卷积层和池化层,然而层数的线性增加会增加网络深度,一定程度上会导致过度拟合,且由于计算复杂度的增加和梯度色散的问题,模型难以进行优化。本文在频域流网络中加入 GoogLeNet 模块,从宽度上优化网络,实现信息跨通道连接的效果。优化后的网络在同等条件下能够更好地提取特征,从而达到提升训练结果的目的。

Inception 是 GoogLeNet 网络中最为关键的结构,如图 9 所示,它将单一的卷积核更改为与池化层配合使用的卷积核组合模式,包括 3 组并行的卷积核(1×1,3×3,5×5),而池化层仅使用单个池化核。使用这种结构具有两个优点:一方面,

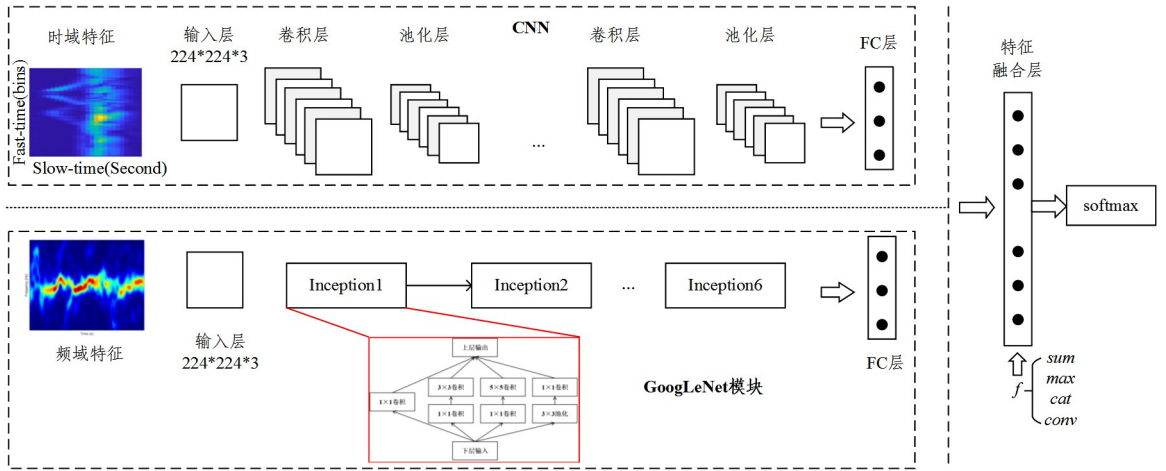


图 9 双流融合网络模型

Fig. 9 Dual-stream fusion network model

4.3.3 特征融合

本文中的双流网络模型^[25],FC 层输出特征尺寸相同,通过在双流网络 FC 层之后添加一个汇聚层的方式进行特征融合。分类是用 softmax 决策函数来执行的,具有最大概率的类被认为是人体动作的预测类。如图 9 所示,采取多种融合方式进行测评。

设定特征融合函数 f , $x_t^a \in (R^{H \times W \times D})$, $y_t^b \in (R^{H' \times W' \times D'})$ 分别为 t 时刻待融合的时域、频域特征,特征融合结果为 $y_t \in R^{H'' \times W'' \times D''}$,其中 H, W, D 分别为特征的高、宽、通道数,这里 $H = H' = H'', W = W' = W'', D = D' = D''$ 。

1) 加和融合

加和融合函数定义为 f^{sum} ,其方式是计算两个特征的加和。特征融合结果 $y^{\text{sum}} = f^{\text{sum}}(x^a, x^b)$,即计算两个特征在同一空间点 (i, j, d) 的和。

$$y_{i,j,d}^{\text{sum}} = x_{i,j,d}^a + x_{i,j,d}^b \quad (15)$$

其中, $1 \leq i \leq H, 1 \leq j \leq W, 1 \leq d \leq D$ 。

2) 最值融合

最值融合函数定义为 f^{max} ,采用两个特征图的最大值。融合的结果表示为 $y^{\text{max}} = f^{\text{max}}(x^a, x^b)$,即将两个特征在同一空间点 (i, j, d) 的最大值作为融合后该点的值。

$$y_{i,j,d}^{\text{sum}} = \max\{x_{i,j,d}^a, x_{i,j,d}^b\} \quad (16)$$

提升了图像尺度适应能力,多卷积核并行的方式让网络能够应用于多种尺寸的图像;另一方面,计算能力得到了一定的提升,通过 3 组卷积核构成的稀疏矩阵可以实现计算效率的提升,这样一来,在同等算力的情况下网络能提取更多的特征。输出量随着卷积核数量的增加而增加,同时稀疏矩阵的产生也避免了可能出现的计算爆炸。因此每次卷积操作之前或者池化操作之后,可以连接一个 1×1 的卷积操作来对图像进行降维,以减少图像的通道数,从而降低图像的计算量。

本文使用 6 个 Inception 模块组合成频域流网络,子模块中每个卷积层的输出被连接在一起以形成对应模块的总输出;保留 CNN 的隐藏层,数据输入尺寸为 224×224 。如图 9 所示,将时域、频域特征分别输入双流网络模型进行训练,设置网络的输出尺寸相同,训练后的输出结果通过全连接 FC 层送往下层进行特征融合。

其中, $1 \leq i \leq H, 1 \leq j \leq W, 1 \leq d \leq D$ 。

3) 连接融合

连接融合函数定义为 f^{cat} ,将两个特征图在通道上串联。融合的结果表示为 $y^{\text{cat}} = f^{\text{cat}}(x^a, x^b)$,即将两个特征通道 d 在相同的空间位置 (i, j) 连接,融合的特征的通道数变为 $2d$ 。

$$y_{i,j,2d}^{\text{cat}} = x_{i,j,d}^a \quad (17)$$

$$y_{i,j,2d-1}^{\text{cat}} = x_{i,j,d}^b$$

其中, $1 \leq i \leq H, 1 \leq j \leq W, 1 \leq d \leq D, y \in R^{H \times W \times 2D}$ 。

4) 卷积融合

卷积融合函数定义为 f^{conv} ,首先在通道 d 的同一空间点 (i, j) 上将两个特征图堆叠起来,并使用上述连接融合函数对它们进行融合,然后对堆叠的数据进行卷积。融合的结果表示为 $y^{\text{conv}} = f^{\text{conv}}(x^a, x^b)$,即输出通道的数量为 D ,卷积核 f 的大小为 $1 \times 1 \times 2D$,用于将维度减少为原来的一半。

$$y^{\text{conv}} = y^{\text{cat}} * f + b \quad (18)$$

卷积融合将连接融合之后得到特征的通道数降低至 D 。

5 实验与分析

5.1 实验设置与场景布置

为了对本文提出的基于双流融合网络的 IR-UWB 非接触式人体动作识别方法的性能进行综合测试,我们进行了

以下多组对比实验。实验设备如图 10 所示,本文使用一台商用的 IR-UWB 无线射频设备获取人员的动作信号,该设备采用的是 Novelda 公司的商用低功耗 X4M05 芯片,设备工作频率为 7.3GHz 或者 8.7GHz,中心带宽为 1.5GHz。本文使用一个 CPU 为 AMD Ryzen 7 5800H,内存为 16GB,GPU 为 GeForce RTX 3060 Ti 的笔记本电脑来进行数据处理和网络模型的训练。

本文的实验场景是一个面积为 $4\text{ m} \times 4\text{ m}$ 的空旷区域,感知区域范围已满足日常活动检测的需求。现场部署了 1 个 IR-UWB 设备,支架高度为 1.2 m,正对人体,以便接收人体反射的脉冲无线信号,每次动作识别仅有一个用户参与实验。



图 10 实验设备和实验场景

Fig. 10 Experimental installation and experimental scene

图 11 展示了实验中使用的 8 个人体日常行为动作,分别是站立、打拳、踏步、转身、挥手、推、拉和下蹲。此外,邀请了 6 名用户来参与收集实验数据,并考虑了身高、体型、性别等的差异性。总共为每个动作收集超过 240 组样本,其中 80% 的样本用于模型训练,20% 用于测试评估。

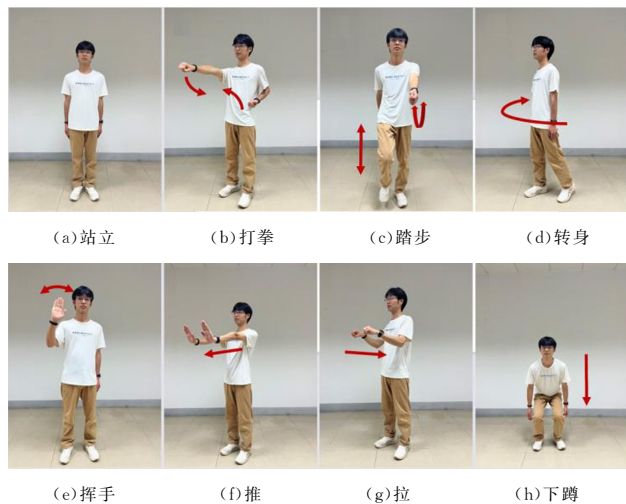


图 11 人体动作示意图

Fig. 11 Diagram of human motions

5.2 系统性能分析

为了对本文提出的方法进行综合性能测试,将人体动作识别准确率作为主要性能指标来评估系统性能。动作识别准确率主要取决于动作标签预测结果与实际动作类别是否相吻合。

1) 总体识别效果

本文动作识别实验共采集了 2145 组数据,其中 80% 的数据作为训练集,20% 的数据作为测试集,使用 softmax 算法

对 8 种不同的人体行为进行识别。图 12 为人体动作识别混淆矩阵,可以看到,除了转身动作的识别准确率为 87% 外,其他人体动作识别准确率均超过 92%。8 种人体动作中,站立和转身、打拳和挥手、推和拉、踏步和下蹲分别为相似的人体动作。实验结果显示,即使是相似的人体动作类型,本文方法仍能够进行有效区分。

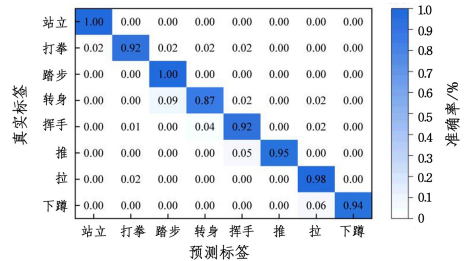


图 12 人体动作识别混淆矩阵

Fig. 12 Confusion matrix of human motion recognition

2) 不同滤波方式的影响

为了验证本文信号预处理阶段中滤波处理方式对系统性能的影响,在对原始信号进行静态杂波抑制处理的基础上,分别在无滤波、仅 Hampel 滤波、仅 Savitzky-Golay 滤波和双重滤波情况下对信号数据提取时域、频域特征,分别送入单流网络和双流融合网络模型进行训练分类。实验结果如图 13 所示。由于动态环境中信号存在数值异常、序列颠簸等因素,相较于不采用滤波的情况,对信号数据进行 Hampel 滤波、S-G 滤波可以消除异常值,平滑数据序列,因此动作识别准确率有所提升,而同时采用两种滤波方式的识别效果更佳。

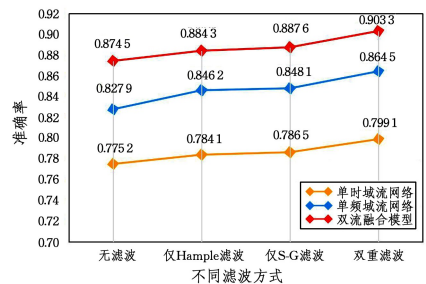


图 13 不同滤波方式识别效果对比

Fig. 13 Comparison of recognition effects of different filtering methods

3) 不同融合策略的影响

为了验证本文提出的双流融合网络模型对系统性能的影响,我们设计了不同的融合策略:时域流和频域流特征输入融合层的比重为 1:1,1:2,1:3,1:5 的加和融合策略,以及最值、连接和卷积融合策略。另外设置对照组:单时域流网络(CNNs)和单频域流网络(GoogLeNet 模块)。具体实验结果如图 14 所示,可以看到,不采用双流融合网络模型时,两种单流网络的动作识别准确率均较低,不到 90%;采用加和融合策略后,不同加权权重对双流网络模型的动作识别准确率也有所影响,由于频域特征反映更加细腻的空间动作移动,因此增大频域流特征的输入比重会使动作识别准确率得到提升。对比可知,卷积融合的方式表现最好,取得了 94.89% 的准确率;其次为连接融合,准确率为 92.63%。

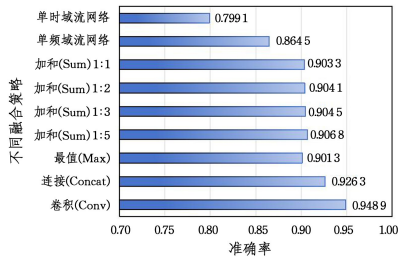


图 14 不同融合策略识别效果对比

Fig. 14 Comparison of recognition effects of different fusion strategies

4) 人员差异的影响

为了测试本文提出的人体动作识别方法对不同人员的动作识别结果,分别对性别、年龄、体态不同的 6 名用户进行测试。用户在 $4\text{ m} \times 4\text{ m}$ 大小的空旷室内环境中做 8 种规定动作,设备发射端与用户距离固定为 1.2 m 。具体的测试结果如图 15 所示。由于每个人的体型和动作幅度有所差异,6 名实验者的动作识别准确度存在一些差异,但总体准确率均保持在 90% 以上,这表明本文提出的动作识别方法对不同人员具有鲁棒性。

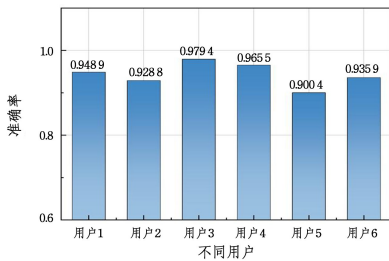


图 15 不同用户识别效果对比

Fig. 15 Comparison of recognition effects of different users

5) 距离差异的影响

为了评估人员与雷达设备之间的距离对系统性能的影响,设置雷达与用户距离为 1.2 m , 2.4 m , 和 3.6 m 。对 6 名用户分别进行了多组实验,计算在每个距离下的动作识别准确度,具体测试结果如图 16 所示。可以看到,人员距离雷达设备 1.2 m 时,识别准确率能达到 96.82% ;随着距离的增大,回波信号受到周围环境干扰的影响增加,系统性能也有所下降,但即使在 3.6 m 的距离下,准确率也超过 90% 。

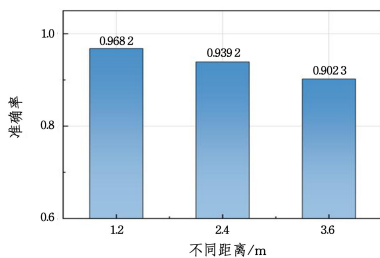


图 16 不同距离识别效果对比

Fig. 16 Comparison of recognition effects at different distances

6) 相对方位的影响

为了测试相对方位对动作识别准确性的影响,分别针对 45° , 30° , 0° , -30° 和 -45° 这 5 个角度下的动作识别进行了测试,具体的测试结果如图 17 所示。当用户正对雷达设备时,

即相对方位角度为 0° 时,动作识别效果最好,准确率为 95.22% 。随着用户与雷达设备的相对方位角度增加,动作识别准确率下降。这是由于人体自身对四肢动作产生遮挡,信号传播遇到人体后发生反射、折射和散射,因此不同的动作方向会导致信号传输路径的改变和信号强度的波动,进而影响动作识别的准确性。

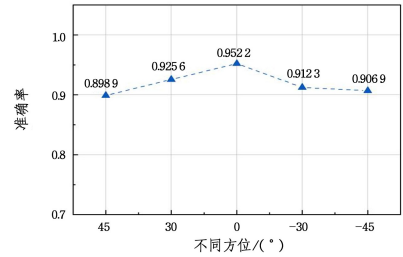


图 17 不同方位识别效果对比

Fig. 17 Comparison of recognition effects in different directions

7) 与其他行为识别算法的比较

将本文方法与其他几种行为识别算法进行了比较,结果如图 18 所示。在本文数据集上对具有 7 层网络结构的 3D-CNN 分类器和由 9 个初始模块组成的原始 GoogLeNet 分类器进行了训练和评估,两个分类器的准确率分别为 89.62% 和 90.25% 。作为对比方案, Kim 等^[26] 和 Skaria 等^[27] 都使用深度卷积神经网络(DCNN)进行特征训练,前者仅使用微多普勒特征作为输入,后者将双接收天线雷达的输出映射到三输入信道 DCNN。结果表明,相较于其他分类算法,本文方法实现了最高的人体动作识别准确率,为 94.89% 。

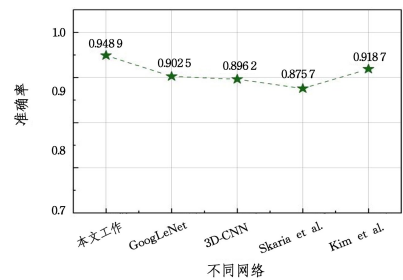


图 18 不同算法的识别效果对比

Fig. 18 Comparison of recognition effects of different algorithms

结束语 本文提出了一种基于双流融合网络的 IR-UWB 非接触式人体动作识别方法。该方法使用 IR-UWB 捕获目标运动所导致的时间延迟和频率偏移,并提取时域和频域特征。在此基础上,构建一个融合多维卷积神经网络和 GoogLeNet 模块的双流网络模型,通过综合多种特征融合策略,成功实现了 IR-UWB 技术下的非接触式人体动作识别功能。大量对比实验验证了本文提出的方法能够实现人体动作识别高精度的效果。在未来的工作中,将进一步考虑运动人体连续动作的识别,以及尝试开展空间内多目标动作识别研究工作。

参考文献

- [1] YANG X, TIAN Y L. Effective 3d action recognition using eigenjoints[J]. Journal of Visual Communication and Image Representation, 2014, 25(1): 2-11.

- [2] XU Y, CHEN J, YANG Q, et al. Human posture recognition and fall detection using Kinect V2 camera[C]//2019 Chinese Control Conference(CCC). IEEE, 2019; 8488-8493.
- [3] LI X, WANG Y, ZHANG B, et al. PSDRNN: An efficient and effective HAR scheme based on feature extraction and deep learning[J]. IEEE Transactions on Industrial Informatics, 2020, 16(10): 6703-6713.
- [4] HE X, ZHU J, SU W, et al. RFID based non-contact human activity detection exploiting cross polarization[J]. IEEE Access, 2020, 8: 46585-46595.
- [5] CASTILLO-CARA M, LOVÓN-MELGAREJO J, BRAVORO-CCA G, et al. An empirical study of the transmission power setting for bluetooth-based indoor localization mechanisms[J]. Sensors, 2017, 17(6): 1318.
- [6] KJARGAARD M B, WIRZ M, ROGGEN D, et al. Mobile sensing of pedestrian flocks in indoor environments using wifi signals[C]//2012 IEEE International Conference on Pervasive Computing and Communications. IEEE, 2012; 95-102.
- [7] 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.
- [8] YAN H, ZHANG Y, WANG Y, et al. WiAct: A passive WiFi-based human activity recognition system[J]. IEEE Sensors Journal, 2019, 20(1): 296-305.
- [9] ZHOU Q, YANG Q, XING J. Enabling efficient WiFi-based occupant behavior recognition using insufficient samples[J]. Building and Environment, 2022, 212: 108806.
- [10] XIAO C, LEI Y, LIU C, et al. Mean teacher-based cross-domain activity recognition using WiFi signals[J]. IEEE Internet of Things Journal, 2023, 10(14): 12787-12797.
- [11] XIA Z, LIU J, GUO S. Rf-care: Rfid-based human pose estimation for nursing-care applications[C]//2021 IEEE International Conference on Robotics and Biomimetics(ROBIO). IEEE, 2021: 1384-1389.
- [12] OGUNTALA G A, HU Y F, ALABDULLAH A A S, et al. Passive RFID module with LSTM recurrent neural network activity classification algorithm for ambient-assisted living[J]. IEEE Internet of Things Journal, 2021, 8(13): 10953-10962.
- [13] LI L, BAI R, XIE B, et al. R&P: a low-cost device-free activity recognition for E-health[J]. IEEE Access, 2017, 6: 81-90.
- [14] LI J, PHUNG S L, TIVIVE F H C, et al. Automatic classification of human motions using Doppler radar[C]//The 2012 International Joint Conference on Neural Networks(IJCNN). IEEE, 2012: 1-6.
- [15] KIM Y, MOON T. Human detection and activity classification based on micro-Doppler signatures using deep convolutional neural networks[J]. IEEE Geoscience and Remote Sensing Letters, 2015, 13(1): 8-12.
- [16] KIM Y, ALNUJAIM I, OH D. Human activity classification based on point clouds measured by millimeter wave MIMO radar with deep recurrent neural networks[J]. IEEE Sensors Journal, 2021, 21(12): 13522-13529.
- [17] ABEDI H, ANSARIYAN A, MORITA P P, et al. AI-powered non-contact in-home gait monitoring and activity recognition system based on mm-wave FMCW radar and cloud computing[J]. arXiv:2208.05905, 2022.
- [18] BRYAN J D, KWON J, LEE N, et al. Application of ultra-wide band radar for classification of human activities[J]. IET Radar, Sonar & Navigation, 2012, 6(3): 172-179.
- [19] AHMED S, CHO S H. Hand gesture recognition using an IR-UWB radar with an inception module-based classifier[J]. Sensors, 2020, 20(2): 564.
- [20] DING C, ZHANG L, GU C, et al. Non-contact human motion recognition based on UWB radar[J]. IEEE Journal on Emerging and Selected Topics in Circuits and Systems, 2018, 8(2): 306-315.
- [21] JIANG L B, LI C, CHE L. Using two-dimensional wavelet packet decomposition for human action recognition ultra-wideband radar[J]. Journal of electronic measurement and instrument, 2018, 32(8): 69-75.
- [22] JIANG J. Human posture recognition based on UWB and support Vector Machine [D]. Beijing: Beijing University of Posts and Telecommunications, 2015.
- [23] AL-QIZWINI M, BARJASTEH I, AL-QASSAB H, et al. Deep learning algorithm for autonomous driving using googlenet [C]//2017 IEEE Intelligent Vehicles Symposium (IV). IEEE, 2017: 89-96.
- [24] LEE J, ABDEL-ATY M, CHOI K, et al. Multi-level hot zone identification for pedestrian safety[J]. Accident Analysis & Prevention, 2015, 76: 64-73.
- [25] MAO M, ZHANG R, ZHENG H, et al. Dual-stream network for visual recognition[J]. Advances in Neural Information Processing Systems, 2021, 34: 25346-25358.
- [26] KIM Y, TOOMAJIAN B. Hand gesture recognition using micro-Doppler signatures with convolutional neural network[J]. IEEE Access, 2016, 4: 7125-7130.
- [27] SKARIA S, AL-HOURANI A, LECH, M, et al. Hand-Gesture Recognition Using Two-Antenna Doppler Radar with Deep Convolutional Neural Networks[J]. IEEE Sensors Journal, 2019, 19: 3041-3048.



ZHANG Chuanzong, born in 1997, post-graduate. His main research interests include IR-UWB and mobile computing.



GUO Zhengxin, born in 1993, Ph.D, lecturer. His main research interests include wireless sensing, mobile computing, deep learning and Internet of Things.