



面向智慧博物馆的基于毫米波雷达稳健的手语识别

赵学荣¹, 王旋¹, 刘彤¹, 郑霞², 江翼成²

(1. 西北大学信息科学与技术学院, 陕西 西安 710127;

2. 浙江大学艺术与考古学院, 浙江 杭州 310028)

摘 要: 智慧博物馆是利用物联网、人工智能等设备或技术, 构建人、物、空间信息交互通道的博物馆新形态。手语识别技术既能让听障语障观众无障碍参观博物馆, 也有助于解析观众自然状态下的手势互动。然而, 基于摄像头或可穿戴设备的方法在博物馆中可能有隐私安全或使用不便等问题。提出一种基于毫米波雷达稳健的手语识别方法, 首先提取不同手势相对于雷达距离和速度随时间变化的特征, 其次采用基于物理意义的增强处理, 最后设计残差网络进一步剔除两种特征预处理后的与环境相关信息, 对其进行特征融合并实现分类。实验表明, 该方法可以有效识别手语, 在测试环境和用户位置改变时也能达到平均 90%以上的精度, 为智慧博物馆的手语手势识别提供了一种新方法。

关键词: 手语识别; 毫米波雷达; 残差网络; 智慧博物馆

中图分类号: TP393

文献标志码: A

doi: 10.11959/j.issn.1000-0801.2023144

mmWave radar based robust sign language recognition for the smart museum

ZHAO Xuerong¹, WANG Xuan¹, LIU Tong¹, ZHENG Xia², JIANG Yicheng²

1. School of Information Science and Technology, Northwest University, Xi'an 710127, China

2. School of Art and Archaeology, Zhejiang University, Hangzhou 310028, China

Abstract: A smart museum is a new form of a museum, which uses devices or technologies including the Internet of things (IoT) and artificial intelligence (AI) to build the information interaction channels between people, things, and space. Sign language recognition not only assists the visitors who have hearing or speech impairment to visit the museum without barriers but also helps study the visitors' natural gesture interaction. However, the methods based on cameras and wearable devices may have issues like privacy or usability when applied to museum spaces. Therefore, a robust sign language recognition method based on millimeter-wave radar was proposed. Different features of distance and velocity changes relative to the radar device were firstly extracted in this method, then a physical data enhancement method was adopted to expand the training data. Finally, a ResNet based on the pre-processed distance time features and Doppler time features was designed to further remove the environment-related information and perform

收稿日期: 2023-04-08; 修回日期: 2023-07-02

通信作者: 郑霞, zhengxia@zju.edu.cn

基金项目: 国家重点研发计划项目 (No.2019YFC1521105)

Foundation Item: The National Key Research & Development Program of China (No.2019YFC1521105)



feature fusion for classification. Experimental results show that this method can effectively recognize sign language and achieve an averaged recognition accuracy of over 90% when the testing environment and the user's location change, providing a new method for smart museum wireless sign language and gesture recognition.

Key words: sign language recognition, millimeter-wave radar, ResNet, smart museum

0 引言

对于听障语障人群来说,由于无法通过语音进行交流,手语成为他们表达和沟通的重要方式。然而,对于没有经过手语训练的人来说,理解手语并且与听障语障群体交流是困难的。因此,开发一个手语识别系统来实现听障语障人群与其他人的沟通显得尤为重要。随着智慧家居、智慧楼宇等智慧空间技术的发展,手语识别系统也有潜力用于人和空间中设备的智能互动,不仅能优化无障碍服务,也能惠及更多公众,提升用户的手势交互和沉浸感体验。

近年来,“智慧博物馆”成为我国博物馆建设的发展趋势。智慧博物馆要求利用物联网、人工智能等设备或技术建立“人-物-空间”三者之间的双向信息通道,实现对博物馆各要素的透彻感知和泛在互联,满足信息时代下观众对博物馆的各种需求^[1]。手语识别系统在智慧博物馆的背景下发展出新的内涵和外延。一方面,手语智能识别有利于实现针对听障语障人群的智慧导览,使观众通过手语无障碍地获得观展引导;另一方面,手语之外的手势识别也可以用于交互装置的互动,实现智慧展示;此外,手势作为一个很重要的非语言符号,在观众与展品、周边的人、环境互动的过程中都扮演着重要角色^[2-3]。如果能对观众自然状态下的手势进行监测,策展方就能深入解析观众言语外的行为数据,从而优化展示和服务的效果。

随着现代科技的迅猛发展,各种手势识别方法不断涌现。现有的手语识别系统可分为基于计算机视觉、基于可穿戴设备和基于无线信号等,目前这几类手语识别系统均缺乏在博物馆空间中

较为普遍的技术应用案例。其中,基于计算机视觉的手语识别系统^[4-6]需要采集人的身体全部信息,存在隐私泄露问题,在跟踪调查中也较难征得对隐私敏感观众的预先同意^[7]。此外,在展览环境下,空间布局较为灵活,不仅光照条件多变,也可能存在展品遮挡,识别效率也会明显下降。基于可穿戴设备的手语识别系统^[8-10]需要执行者佩戴特殊定制的手套或手表等设备来获取手臂的移动信息。然而,尽管该设备能够实现隐私不敏感的监测,提供观众定制化服务,但需要手语执行者长时间的全程佩戴,影响用户的舒适性和参观体验。相比于前两种方法,由于无线信号(如Wi-Fi、射频识别(radio frequency identification, RFID)、音频和毫米波)具有许多优点,例如不受光线影响、无须与用户接触、成本低且部署灵活,因此它们更适用于博物馆展览等手语识别应用环境^[11-16]。

使用Wi-Fi和RFID信号实现的手语识别系统通常具有低成本和易于部署的优点。然而,这些系统容易受环境变化、位置变化和其他干扰的影响,因此在识别准确率和稳定性方面存在局限性。而基于音频信号实现的手语识别系统也被广泛研究,但由于声波在空气中衰减很快,因此音频设备和手之间的距离必须非常短(通常小于1 m),导致其识别范围也有很大的局限性。

相比于Wi-Fi、RFID和音频信号,毫米波信号具有更短的波长,因此可以捕获更细粒度的手势信息。毫米波信号也具有更大的带宽,因此具有更大的感知范围。mmGesture^[17]通过实验证明了距离多普勒图像(RDI)、距离角度图像(RAI)、多普勒(Doppler)角度图像(DAI)以及微多普勒频谱图这4种图像在手势识别中的有效性。一些研

究工作^[18-19]仅基于单个 Doppler 频谱图进行手势识别,而文献[20-21]的方法则基于 RDI 图实现。与现有的解决方案不同,本文首先提取 2D 距离时间特征和 2D 多普勒时间特征,并分别用神经提取深层特征后再进一步融合,实现了两个特征相互补充的效果,提升了系统的识别精度和稳健性,同时也避免使用 3D 卷积,降低了计算量。此外,本文采用基于物理意义的数据增强方法,对采集的数据集进行速度、距离、时间维度的平移和拉伸,以扩充训练数据集,使其包含用户不同的执行手势速度和执行位置的数据,该方法进一步增强了识别方法在环境、位置和手势执行改变时的稳健性。本文提出的基于毫米波信号稳健的手语识别方法有潜力应用于博物馆展览等实际应用场景。

1 系统结构

本文的手语识别系统如图 1 所示,首先通过毫米波雷达设备采集手势信号,通过对距离和速度特征进行处理,消除静态反射物的干扰和环境、位置等因素的影响,通过基于残差网络的 Range 特征提取和 Doppler 特征提取,并将它们提取的特征进行特征融合,再通过分类网络进行识别,最终得到精准的手语识别结果^[22]。相比于无线信号识别方法的常见问题,如受环境、用户位置变化和不同用户动作偏差等影响导致识别精度下降,本文方法具有更高的识别精度和稳定性。

2 毫米波信号处理

由于目标肢体的运动,毫米波雷达接收的中频

信号在时域和频域中都会发生变化,反映了目标相对于雷达距离和多普勒速度的变化。将原始无线信号转换成距离和多普勒速度特征处理流程如下。

步骤 1 信号预处理。雷达接收器接收的原始毫米波信号通常包含许多噪声和干扰,这些干扰会影响雷达系统的性能和精度。为了滤除与手势无关的信号,需要对信号进行带通滤波去除直流分量和高低频噪声。

步骤 2 获取 Frame-Range 特征矩阵。毫米波雷达设备会连续发送一系列的 chirp 信号,经过目标反射后返回雷达接收端。接收端将接收信号与发送信号混频后采样传输至处理器。将采样整形为维度为帧数(F)×每个 chirp 采样点(N)×每帧的 chirp 数(M)×接收天线数(K)的数据块 R 。对每个收到的 chirp 采样点进行快速傅里叶变换 (fast Fourier transform, FFT) 得到频谱图 Range-Profile, 其中每个 FFT-bin 对应一个离散的距离称为 Range-bin, 雷达到目标的距离可通过最高能量对应的频率计算得到,即 $d = \frac{ct_d}{2}$, 其中 c 表示光速, t_d 表示时延。对一帧中所有 chirp 执行 Range-FFT 操作并通过累积一帧中所有 chirp 的 Range-Profile 增强信号能量,并对每一帧进行同样操作以获得距离随时间的变化特征 Frame-Range。

步骤 3 获取 Frame-Doppler 特征矩阵。在 Range-Profile 的基础上,对每个 Range-bin 的所有 chirp 采样点进行快速傅里叶变换,得到对应的频谱图 Doppler-Profile。此时,每个 FFT-bin 对应一个多普勒速度,称为 Doppler-bin。目标的速度可以通过最高能量对应的频率计算,即 $v = \frac{\lambda f_d}{2}$, 其

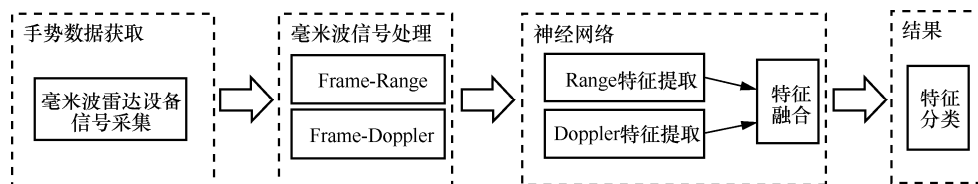


图 1 手语识别系统



中 λ 表示信号的波长, f_d 表示信号的多普勒频移。累积一帧中的所有 Range-bin 的 FFT 的 Doppler-Profile, 并对每一帧进行同样操作以获得速度随时间的变化特征 Frame-Doppler。

步骤 4 数据增强。由于数据采集过程繁杂耗费时间, 因此, 为了增加样本多样性, 本文生成了现有样本在速度、距离、时间维度均有变化的情况。具体来说, 通过对基准样本的 Frame-Doppler 在速度维度和时间维度进行缩放以模拟每次执行手势速度差异, 对 Frame-Range 特征在距离维度上进行上下平移来模拟生成用户在不同距离下执行的手势的数据, 使得训练集中包含了用户在不同位置和以不同速度执行手势的数据, 提升了模型的识别准确度和稳健性。

3 神经网络设计

特征提取和分类网络结构如图 2 所示, 即所设计的神经网络, 深度神经网络的输入为无线信号特征, 在经过特征提取和分类模块后, 得到对无线信号特征进行分类的输出结果。特征提取模块是基于 ResNet (残差网络)^[23] 的, 包括卷积层、归一化层、最大池化层、卷积残差块、恒等残差块、平均池化层、卷积层、归一化层和自适应池

化层。特征分类模块包括全连接层、Dropout^[24] 层和 Softmax 激活层。通过调整超参数, 如卷积核大小、卷积步长等, 选择分类效果最优的模型。网络的优化函数选择 Adam^[25], 用于加速网络训练和提高分类准确性。最终目的是得到对无线信号特征进行分类的准确结果。

3.1 输入数据准备模块

定义输入数据集 D 和第 i 个数据样本:

$$D = \{([X(f)_i, R(f)_i], Y_i)\}_{i=1}^N \quad (1)$$

其中, D 为输入神经网络中数据样本的集合, f 为经过两次 FFT 以及静态物体消除的三维无线信号特征; $|X(f)_i|_i$ 为多普勒速度矩阵沿着距离维度压缩得到的 Frame-Doppler 特征矩阵; $|R(f)_i|_i$ 为无线信号特征沿着 chirp 维度压缩得到的 Frame-Range 特征矩阵; Y_i 为真实样本标签。

3.2 特征提取模块

本文采用残差网络框架作为特征提取的基础框架, 残差网络与普通网络不同的地方就是引入了跳跃连接, 这可以使上一个残差块的信息没有阻碍地流入下一个残差块, 提高了信息流通, 并且避免了网络过深引起的消失梯度问题和退化问题。残差网络由多个浅的网络融合而成, 浅的网

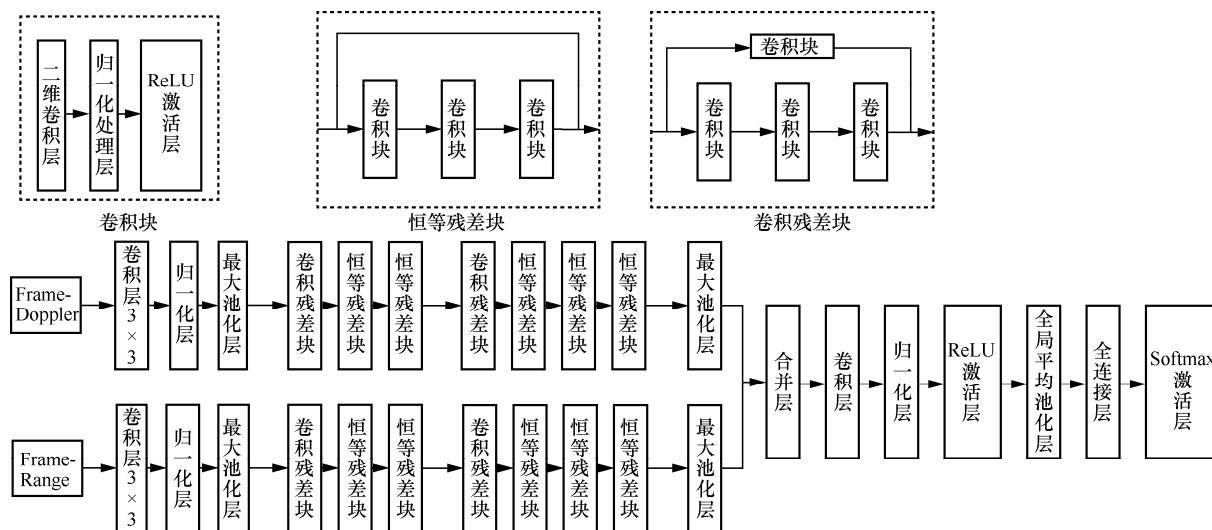


图 2 特征提取和分类网络结构

络在训练时不会出现消失的梯度问题, 所以它能够加速网络的收敛。

首先, 定义基础卷积块, 该块采用卷积层 (Conv2D)、批归一化层和 ReLU 激活函数, 由于本文数据特征尺寸小, 所以选择较小的卷积核 3×3 , 卷积步长为 3。恒等残差块 (identity block) 包含 3 个卷积块的基本路径再短接输入, 采用 ResNet 架构, 设计卷积残差块 (convolutional block) 包含一个 4 个卷积块的基本路径再短接 1 个卷积块, 同样采用 ResNet 架构。通过堆叠多个恒等残差块和卷积残差块提取输入特征的深层表达, 再连接最大池化层 (MaxPooling2D) 以减少参数量, 防止过拟合, 其中最大池化层的窗口也选择最小 2×2 。

然后, 为了进一步提取特征并增强模型的表达能力, 对两个输入特征的高维表达输出联合后添加 1 个卷积块, 为了降低空间维度并提取整体特征表示, 本文添加了平均池化层 (average Pooling2D) 和全局平均池化层 (global average Pooling2D), 其中平均池化层的窗口大小选择 2×2 , 全局平均池化层的窗口大小选择 2×2 。将这些子模块依次连接起来, 形成特征提取模块。

3.3 分类模块

特征分类模块包括全连接层和 Softmax 层。全连接层 (Dense) 可定义神经元数量和激活函数, 全连接层的输入维度应该与特征提取模块的输出维度相匹配, 输出维度为手势的类别数。接着, 为了进一步提高模型的稳健性和防止过拟合, 本文添加了 Dropout 层, 并根据经验和实验将 Dropout 比例设置为 0.5。最后, 添加 Softmax 层获得类别概率分布实现分类。

3.4 模型优化模块

手语识别网络选择 Adam 作为优化函数。首先, 它具有自适应学习率的特性, 能够根据每个参数的梯度自动调整学习率, 有效平衡不同参数的更新速度。其次, Adam 算法在训练初期能够快速收敛到较好的解决方案, 结合了动量的概念,

加速了梯度下降的过程。此外, Adam 算法对稀疏梯度的适应性较好, 对于处理包含大量稀疏梯度的任务具有优势。最后, 相对于其他优化算法, Adam 算法具有较低的内存需求, 更适用于处理大规模数据集和复杂模型。综合考虑这些因素, 选择 Adam 作为优化函数能够有效地优化本文的模型, 提升训练效果和性能。

3.5 特征融合模块

由于某些手势在 Frame-Range 特征上表现不明显, 而在 Frame-Doppler 特征上表现较清晰; 某些手势在 Frame-Doppler 特征上表现不明显, 而在 Frame-Range 特征上表现较清晰——为了实现两特征的互补效果, 将两特征通过 Concatenate 层连接, 之后添加一个卷积块以提取更高级别的特征信息。这有助于模型更好地学习特征之间的关联性, 再添加一个平均池化层可以将特征图降维, 减少模型的参数量, 再添加一个自适应池化层, 将不同大小的特征图转换为固定大小的特征向量, 这有助于保留特征的空间位置信息, 提高模型的泛化能力。

3.6 基于残差网络剔除与环境相关信息模块

在传统的卷积神经网络中, 每一层都是直接对上一层的输出进行卷积操作, 这样会导致高层特征与底层特征之间的相关性很强, 这些相关性往往与环境信息相关, 因此会对模型的决策产生负面影响。而在基于残差网络的特征提取网络中, 每一层不仅对上一层的输出进行卷积操作, 还将上一层的输出与当前层的输入进行残差连接。残差连接可以将底层特征与高层特征分离开, 使得高层特征不再受底层特征的干扰, 从而减少环境信息的影响。具体地, 残差连接可以将当前层的输出 (即高层特征) 与当前层的输入 (包含底层特征) 相加, 得到一个更具表征性的特征向量作为当前层的输出。因此, 经过残差连接后, 底层特征与高层特征被分别保存在输入和输出中, 从而实现了对高层特征和底层特征的有效分离, 使网络更加容易学习到数据本身的内在结构, 而不受环境变化的影响。



4 实验评估

4.1 实验设备及参数设置

为了获取手势活动数据,本文采用德州仪器 TI-IWR1843 传感器芯片,频段为 77~81 GHz,带宽为 4 GHz,斜率为 100 MHz/ μ s,采样率为 2.499×10^6 sample/s,每个 chirp 采样点数为 80,每帧 chirp 的循环数为 255,共采用 3 根发射天线和 4 根接收天线。

4.2 数据收集

为了减少噪声干扰,我们选择在大厅中进行实验,毫米波雷达设备放置在实验受试者的正前方,距离受试者 1.4 m,考虑博物馆观众身高的多样性,设备距离地面 0.9 m。所收集的数据块尺寸为 $130 \times 80 \times 255 \times 12$ 。实验环境布置如图 3 所示。



图3 实验环境布置

4.3 手语识别性能

为了验证本方法的性能,本文进行了手语识别实验。实验对象为 10 名受试者,每人做 38 种手势进行测试。实验使用信号的 Doppler 特征和 Range 特征作为输入特征,对手语进行识别。手语识别性能如图 4 所示,在仅使用 Doppler 特征和 Range 特征进行识别时,平均识别精度分别为 91.48%和 91.56%,这已经表现出较为出色的识别能力。当将 Doppler 特征和 Range 特征进行融合,再次进行手语识别时,平均识别精度达到 94.42%。这表明本文的方法在使用多特征融合时,能够进一步提高手语识别的性能。

为了验证本方法相对于其他方法的性能优势,本文进行了对比实验,使用 10 名受试者的数据进行模型训练和测试,将本文提出的融合特征方法与 mmASL 进行了对比,与其他方法的对比如图 5 所示。在这个实验中,本文所提方法精度可达 94.42%,而 mmASL 的精度只有 85%。这表明本文的融合特征方法在这个实验中比 mmASL 性能更好。另外,当只使用 8 名受试者的数据进行训练,并用未参与训练的两名受试者的数据进行测试时,发现 mmASL 的精度只有 75%,而本文所提融合特征方法的精度是 85%。这表明本文方法比 mmASL 有更好的稳健性,在实际应用中具有更高的可靠性和实用性。

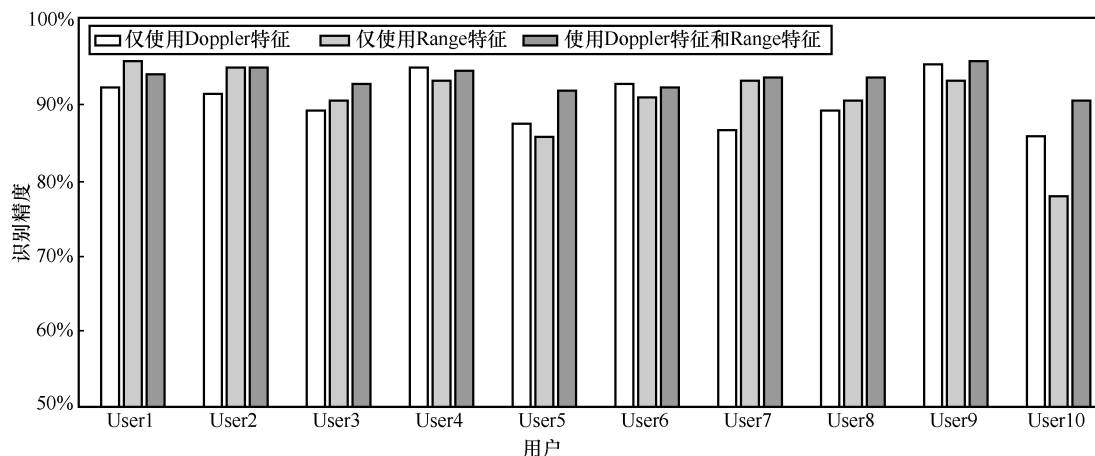


图4 手语识别性能

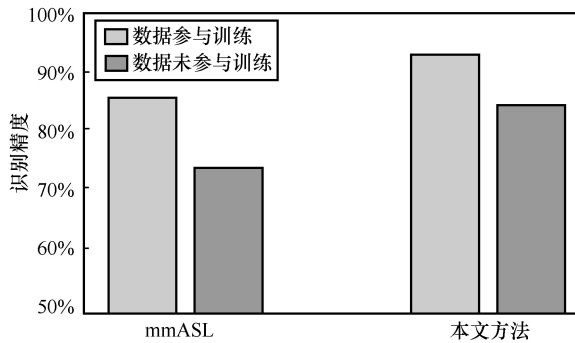


图5 与其他方法的对比

4.4 稳健性验证

为了验证在环境等因素变化时该方法的稳健性，本文在不同环境、距离和角度条件下进行了比较试验。

(1) 跨环境稳健性验证

环境在无线传感系统中起着重要作用，这是因为不同的房间大小、家具放置和设备部署会导致不同的多径效应。为了研究手势识别在不同环境下的表现，本文进行了一个环境稳健性测试，博物馆通常由大厅、走廊和办公室等场景构成，本文选择了3个具有代表性的环境（走廊、大厅和办公室）作为测试地点。这意味着一个环境作为测试集、其他两个环境作为训练集，不同环境下的识别准确率如图6所示，走廊的平均精度较高，具体数值为95.55%，大厅和办公室准确率次之，分别为92.66%和88.66%。在经过数据增强处理之后的结果分别为95.7%、92.3%和91.9%。结果显示，在不同环境下该系统的识别准确率相差不大（5%以内），并且识别准确率都达到了90%以上。这样的结果证明了本文方法对环境的变化是稳健的，具有迁移至博物馆空间中的潜力。

(2) 距离变化稳健性验证

距离（毫米波雷达距离受试者的距离）对毫米波雷达的影响具体体现在反射信号的强度上。一般而言，距离增大会导致接收的信号强度降低，从而影响特征的提取。为了验证系统对距离的稳健性，实验设置了3个距离，分别为：1.4 m、1.8 m、2.2 m。其中，一个作为测试集，另外两个作为训练集，去测试这

3个地方的具体识别率。经过后3个距离的识别准确率分别为88.66%、95.33%、93.55%。经过数据增强，3个距离的识别准确率分别为90.6%、96.2%、95.1%。不同距离下的识别准确率如图7所示，在不同距离下该系统的识别准确率相差不大（7%以内），并且它们的识别准确率都达到90%以上。该结果证明本方法对距离是稳健的。

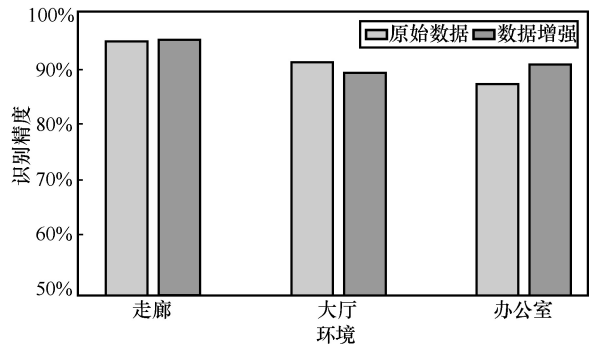


图6 不同环境下的识别准确率

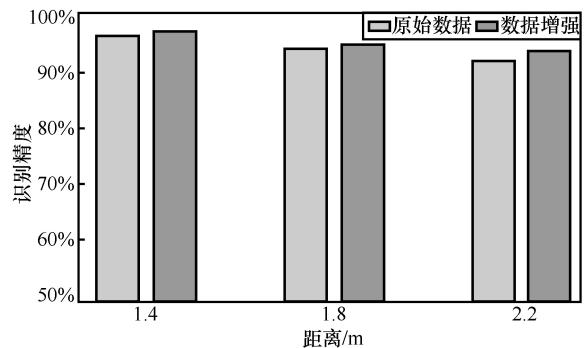


图7 不同距离下的识别准确率

(3) 角度变化稳健性验证

雷达的视角范围有限，同时，面向雷达的不同角度会有不同的强度。因此，实验设置5个角度（毫米波雷达与受试者的水平方位角）进行验证，分别为 -40° 、 -20° 、 0° 、 20° 、 40° ，其他条件与默认条件保持一致。选择一个作为测试集、另外4个作为训练集。5个角度的识别准确率分别为92.2%、96.5%、88.88%、95.56%、90.88%。经过数据增强后，这5个角度的识别准确率分别为93.11%、97.11%、94%、90.88%、90.98%。不同角度下的识别准确率如图8显示， 20° 与 40° 的识别率略低于 -20° 和 -40° 的识别率。一个可能的原



因是,当用户执行手势动作时,用户的右手位于雷达视野(FoV)的角落,导致从人手反射的信号变弱,从而造成识别率略有降低。

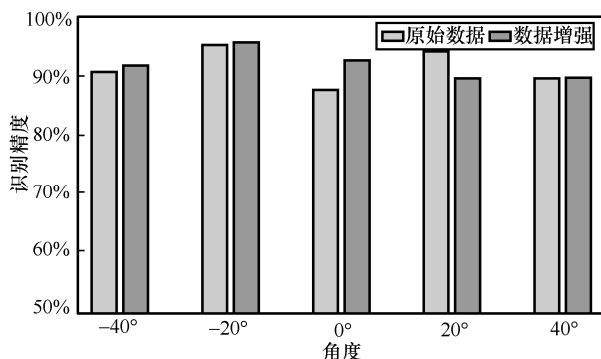


图8 不同角度下的识别准确率

实验结果显示,在不同角度下该系统的识别准确率相差不大(6%以内),并且它们的识别准确率都达到90%以上。该结果证明本文的系统对角度是稳健的。

综上所述,该系统的识别准确率达到90%以上,可以有效地识别手语,并随着环境的变化、距离的变化、角度的变化,识别准确率变化较小,系统对这些影响有较强的稳健性。

5 结束语

本文在智慧博物馆的建设需求下提出了一种基于毫米波雷达的手语识别系统,旨在为听障语障人群提供一种便捷的沟通和导览方式,同时初探基于无线信号的非接触手势识别技术在智慧博物馆空间中的应用潜力。本文提出的系统采用毫米波雷达来收集手语的回波无线信号,并采用数字信号处理和深度学习技术来识别手语。具体而言,特征处理采用原始无线信号经过去噪、去除静态物体信息、经过两次傅里叶变换后提取距离和速度特征作为分类网络的输入特征。使用设计的ResNet分别对两个特征进一步提取深度高维表达联合后输入分类网络进行识别。实验结果表明,该系统可以有效地识别手语,并具有较强的环境稳健性,有潜力为听障语障群体提

供一种新的无接触的无线手语识别方法。未来,本文方法可以在多人监测、可识别手势范围和现场实验方面继续优化,以用于基于手势交互的智慧导览与展示、对观众手势动作的分析研究等,更加匹配博物馆对空间智能感知和观众包容性服务的需求。

参考文献:

- [1] 王春法. 关于智慧博物馆建设的若干思考[J]. 博物馆管理, 2020(3): 4-15.
WANG C F. Some thoughts on the smart museum construction[J]. Museum Management, 2020(3): 4-15.
- [2] DI GIUSEPPANTONIO DI FRANCO P, MATTHEWS J L, MATLOCK T. Framing the past: how virtual experience affects bodily description of artefacts[J]. Journal of Cultural Heritage, 2016(17): 179-187.
- [3] STEIER R, PIERROUX P, KRANGE I. Embodied interpretation: gesture, social interaction, and meaning making in a national art museum[J]. Learning, Culture and Social Interaction, 2015(7): 28-42.
- [4] PUGEAULT N, BOWDEN R. Spelling it out: real-time ASL fingerspelling recognition[C]//Proceedings of 2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops). Piscataway: IEEE Press, 2011: 1114-1119.
- [5] MOHANDES M, ALIYU S, DERICHE M. Arabic sign language recognition using the leap motion controller[C]//Proceedings of 2014 IEEE 23rd International Symposium on Industrial Electronics (ISIE). Piscataway: IEEE Press, 2014: 960-965.
- [6] CAO D, LEU M C, YIN Z Z. American sign language alphabet recognition using Microsoft Kinect[C]//Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE Press, 2015: 44-52.
- [7] YALOWITZ S S, BRONNENKANT K. Timing and tracking: unlocking visitor behavior[J]. Visitor Studies, 2009, 12(1): 47-64.
- [8] LI K H, ZHOU Z Y, LEE C H. Sign transition modeling and a scalable solution to continuous sign language recognition for real-world applications[J]. ACM Transactions on Accessible Computing, 2016, 8(2): 1-23.
- [9] WU J, SUN L, JAFARI R. A wearable system for recognizing American sign language in real-time using IMU and surface EMG sensors[J]. IEEE Journal of Biomedical and Health Informatics, 2016, 20(5): 1281-1290.
- [10] YANG X D, CHEN X, CAO X, et al. Chinese sign language recognition based on an optimized tree-structure framework[J]. IEEE Journal of Biomedical and Health Informatics, 2016, 21(4): 994-1004.

- [11] ABDELNASSER H, YOUSSEF M, HARRAS K A. WiGest: a ubiquitous Wi-Fi-based gesture recognition system[C]//Proceedings of 2015 IEEE Conference on Computer Communications (INFOCOM). Piscataway: IEEE Press, 2015: 1472-1480.
- [12] MA Y S, ZHOU G, WANG S Q, et al. SignFi: sign language recognition using Wi-Fi[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2018, 2(1): 1-21.
- [13] 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. New York: ACM Press, 2019: 313-325.
- [14] ZHANG L, ZHANG Y X, ZHENG X L. WiSign: ubiquitous American sign language recognition using commercial Wi-Fi devices[J]. ACM Transactions on Intelligent Systems and Technology, 2020, 11(3): 1-24.
- [15] WU C S, ZHANG F, WANG B B, et al. mSense: towards mobile material sensing with a single millimeter-wave radio[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(3): 1-20.
- [16] LI H, YANG W, WANG J X, et al. WiFinger: talk to your smart devices with finger-grained gesture[C]//Proceedings of the 2016 ACM International Joint Conference on Pervasive and Ubiquitous Computing. New York: ACM Press, 2016: 250-261.
- [17] YAN B, WANG P, DU L, et al. mmGesture: semi-supervised gesture recognition system using mmWave radar[J]. Expert Systems with Applications, 2023(213): 119042.
- [18] LIU H, ZHOU A, DONG Z, et al. M-gesture: person-independent real-time in-air gesture recognition using commodity millimeter wave radar[J]. IEEE Internet of Things Journal, 2021, 9(5): 3397-3415.
- [19] LI Y D, ZHANG D H, CHEN J B, et al. DI-gesture: domain-independent and real-time gesture recognition with millimeter-wave signals[C]//Proceedings of GLOBECOM 2022 - 2022 IEEE Global Communications Conference. Piscataway: IEEE Press, 2022: 5007-5012.
- [20] SANTHALINGAM P S, HOSAIN A A, ZHANG D, et al. mmASL: environment-independent ASL gesture recognition using 60 GHz millimeter-wave signals[J]. Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies, 2020, 4(1): 1-30.
- [21] SANTHALINGAM P S, DU Y Q, WILKERSON R, et al. Expressive ASL recognition using millimeter-wave wireless signals[C]//Proceedings of 2020 17th Annual IEEE International Conference on Sensing, Communication, and Networking (SECON). Piscataway: IEEE Press, 2020: 1-9.
- [22] 李攀攀, 谢正霞, 乐光学, 等. 基于深度学习的无线通信接收方法研究进展与趋势[J]. 电信科学, 2022, 38(2): 1-17.
LI P P, XIE Z X, YUE G X, et al. Research progress and trends of deep learning based wireless communication receiving method[J]. Telecommunications Science, 2022, 38(2): 1-17.
- [23] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE Press, 2016: 770-778.

- [24] SRIVASTAVA N, HINTON G, KRIZHEVSKY A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. Journal of Machine Learning Research, 2014, 15(1): 1929-1958.
- [25] KINGMA D P, BA J. Adam: a method for stochastic optimization[J]. arXiv preprint, 2014, arXiv:1412.6980.

[作者简介]



赵学荣(1999—), 男, 西北大学信息科学与技术学院硕士生, 主要研究方向为无线感知。



王旋(1993—), 女, 西北大学信息科学与技术学院博士生, 主要研究方向为毫米波人体活动感知以及无源物联网。



刘彤(1998—), 男, 西北大学信息科学与技术学院硕士生, 主要研究方向为无线感知。



郑霞(1979—), 女, 博士, 浙江大学艺术与考古学院副教授, 主要研究方向为博物馆信息化、智慧博物馆。



江翼成(1997—), 男, 浙江大学艺术与考古学院硕士生, 主要研究方向为文化遗产数字化传播、智慧博物馆。