

# 基于 MFCC 特征的 Wi-Fi 信道状态信息人体行为识别方法

蒙倩霞 余江 常俊 浦钰

(云南大学信息学院 云南 昆明 650500)

**摘要** CSI(Channel State Information)可提供被动的人体行为识别方法,根据 CSI 和声音信号传播相似性和共享频谱带宽的特性将 MFCC(Mel-Frequency Cepstral Coefficients)特征提取应用于 CSI 时间序列,并对复杂场景下的视距和非视距的几种日常行为进行识别。该方法对数据去噪、PCA、相位校准处理,从预处理后的信号中提取了 MFCC 统计特征和一个无偏移对数频谱能量,用蚁群和粒子群混合优化 SVM 进行分类识别。实验结果表明,该方法能有效识别复杂场景下的日常行为,在视距情况下,平均识别率达到了 91%。

**关键词** CSI 行为识别 PCA MFCC 支持向量机

**中图分类号** TP3 TN99

**文献标志码** A

**DOI**:10.3969/j.issn.1000-386x.2022.12.019

## HUMAM BEHAVIOR RECOGNITION METHOD BY WI-FI CHANNEL STATE INFORMATION BASED ON MFCC CHARACTERISTICS

Meng Qianxia Yu Jiang Chang Jun Pu Yu

(School of Information Science and Engineering, Yunnan University, Kunming 650500, Yunnan, China)

**Abstract** Channel state information(CSI) can provide passive human behavior recognition methods. According to the characteristics of CSI and sound signal propagation similarity and shared spectrum bandwidth, Mel-frequency cepstral coefficients(MFCC) feature extraction was applied to CSI time series, and several daily behaviors of LOS and NLOS in complex scenes were identified. The data denoising, PCA and phase calibration were processed. The MFCC statistical features and an unbiased logarithmic spectrum energy were extracted from the preprocessed signal, and the SVM based on ant colony and particle swarm optimization was used for classification and recognition. The experimental results show that this method can effectively recognize the daily behavior in complex scenes, and the average recognition rate is 91% in the case of LOS distance.

**Keywords** CSI Behavior recognition PCA MFCC Support vector machine

## 0 引言

基于摄像机的行为识别方法无法监测光线盲区且隐私无法得到保障;基于雷达的部署成本过于昂贵,不适用于普通家庭;基于传感器的方法传感器一般佩戴于腰上,对于特殊人群来说,发生意外的时候无疑会增加危险指数。基于 Wi-Fi 信道状态信息(CSI)的活动识别系统克服了以上方法的缺陷,为活动识别提供了一种新的保密隐私、价格实惠且无须携带设备的方法。Wi-Fi 信号易受到人体行为的影响,不同动作速度的

快慢、维持的时间、身体的摆动幅度都会对 CSI 的变化造成不同的影响。因此可以根据 CSI 的变化去判别相应的行为。

Wi-Fall<sup>[1]</sup>跌倒检测系统分析出不同行为 CSI 幅值变化情况,平均实验精度为 87%,但忽略了相位带来的影响。Emosens 系统<sup>[2]</sup>从 CSI 识别用户的情绪状态,同样使用的是幅值特征。文献[3]提出了一种吸烟检测系统,用于室内环境吸烟行为检测,单个人的吸烟检测正确率为 97.6%,但是多人条件下,性能急剧下降。文献[4]采用 RNN 网络可以有效识别其动作,穿墙识别精度高达 95% 以上。利用机器学习的方法

虽然可以自动提取特征,识别精度也甚好,但是计算的复杂度很高,需要大量的数据训练且提取特征的合理性也难以解释。在最新的研究成果中,文献[5]提出了一种穿墙识别的结构,但仅限理论分析。文献[6]提出的 TS-see 系统同样也实现了穿墙识别。文献[7]提出在多人场景中无设备位置行为特征转换的问题,场景复杂化和多人适应性将是今后研究的重点之一。文献[8]利用 CSI 来识别驾驶员的注意力状态,获得了较好的识别性能。文献[9]利用多通道 CSI 测量数据转化二维图像,并使用深度学习进行训练和分类,这是基于观察而得到的方法。

本文利用 CSI 信号和声音信号传播相似性和两者共享带宽频谱的性质,把 CSI 信号当作语音流来处理,在 CSI 数据中提取了 MFCC 特征向量的统计特征以及提取了对数频谱能量特征,然后用蚁群和粒子群混合优化 SVM (Ant Colony and Particle Swarm Optimization SVM, ACO + POS-SVM)<sup>[10]</sup> 算法进行分类识别。实验结果表明,此方法在会议室场景的视距和非视距情况下都可以获得较好的识别结果。

## 1 基于 CSI 识别基本原理

### 1.1 Wi-Fi 信号传播模型

由于 Wi-Fi 信号以多路径方式传播,环境中的物体会反射无线信号,那么一个发送信号  $x(t)$  就会通过多条路径到达接收器。假设  $x(t)$  通过  $N$  条不同的路径到达接收器,那么接收到的信号  $y(t)$  如式(1)所示。

$$y(t) = e^{-j2\pi\Delta ft} \sum_{k=1}^N a_k(t) e^{-j2\pi\tau_k(t)} x(t - \tau_k(t)) \quad (1)$$

式中: $a_k(t)$  是  $t$  时刻第  $k$  条路径的幅值大小; $e^{-j2\pi\tau_k(t)}$  是第  $k$  条路径上的相移; $\tau_k(t)$  为其传播延迟; $e^{-j2\pi\Delta ft}$  是发送器和接收器之间的载波频率差  $\Delta f$  引起的相移。如图 1 所示。

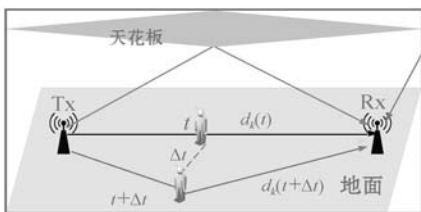


图 1 行为引起的多径变化

当人在  $t$  到  $t + \Delta t$  的时间内在发射机和接收机的视距之间活动时,第  $k$  条路径上的信号由  $d_k(t)$  变化到  $d_k(t + \Delta t)$ ,此时接收端产生传播延迟相移  $\tau_k(t) = d_k(t + \Delta t)/c$ ,则根据  $\lambda = c/f$  可以得到该条路径上的相移为  $e^{-j2\pi d_k(t + \Delta t)/\lambda}$ ,其中: $\lambda$  为波长; $c$  为光速。这说明

路径的变化会导致相应路径上 Wi-Fi 信号的幅度和相位的变化。

### 1.2 CSI 数据模型

无线通信链路信道的特性通常是用信道信息状态来衡量,根据时域卷积性质,可以把式(1)表示为:

$$y(t) = \int_{-\infty}^{\infty} h(\tau, t) x(t - \tau_k(t)) d\tau \quad (2)$$

由式(1)和式(2)可得信道冲击响应(Channel Impulse Response, CIR)<sup>[11]</sup>为:

$$h(\tau, t) = e^{-j2\pi\Delta ft} \sum_{k=1}^N a_k(t) e^{-j2\pi\tau_k(t)} \delta(t - \tau_k(t)) \quad (3)$$

在 IEEE 802.11n 标准中,利用正交频分复用技术可以从物理层解析得到 CSI 数据。根据时频域的转换性质,可得到频域的信道信息响应  $H(f, t)$  为:

$$H(f, t) = \int_{-\infty}^{\infty} h(\tau, t) e^{-j2\pi f\tau} d\tau = e^{-j2\pi\Delta ft} \sum_{k=1}^N a_k(f, t) e^{-j2\pi f\tau_k(t)} \quad (4)$$

式中: $a_k(f, t)$  是  $t$  时刻第  $k$  条路径的衰减和初始相位偏移的复数值表示。

在 MIMO 系统中,设  $NT_X$  为发射端的天线数量,  $NR_X$  为接收端的天线数量。对于每个采样时刻  $t$ ,CFR 值构成  $NT_X \times NR_X \times i$  维的一个矩阵,其中  $i$  为 Wi-Fi 的子载波数。所以,在一段时间内,CSI 由  $NT_X \times NR_X \times i$  维 CFR 值构成。 $H(f, t)$  也可以表示为:

$$H(f, t) = \|H(f, t)\| e^{j\angle H(f, t)} \quad (5)$$

式中: $\|H(f, t)\|$  和  $\angle H(f, t)$  分别表示  $t$  时刻频率为  $f$  状态下的 CSI 的幅度和相位。

### 1.3 MFCC 特征向量

MFCC 向量是基于人耳对声音的频率感知的能力,其对应的梅尔频率与赫兹频率成非线性关系,频率与梅尔频率之间转换公式如式(6)所示<sup>[12]</sup>。

$$m = 2595 \log_{10} \left( 1 + \frac{f}{700} \right) \quad (6)$$

MFCC 向量<sup>[13]</sup>提取过程中首先要对预处理过的信号进行分帧和加窗处理,常用的有汉明窗口如式(7)所示。

$$w(n) = 0.54 - 0.46 \cos(2n\pi/N - 1) \quad (7)$$

式中: $n$  是窗口的长度; $w$  是窗口值。

然后,将各帧信号  $h(n)$  求快速傅里叶变换(FFT)得到各帧的频谱  $H(k)$ :

$$H(k) = FFT[h(n)] \quad (8)$$

将上述的功率谱  $|H(k)|^2$  通过梅尔滤波器组并求出对数能量得到对数梅尔频谱  $S(m)$ :

$$S(m) = \log\left(\sum_{k=1}^{N-1} |H(k)|^2 H_m(k)\right) \quad (9)$$

式中: $N$  是 FFT 的点数; $m$  为滤波器的中心频率。 $H_m(k)$  是三角形重叠的窗口,其描述方程为:

$$H_m(k) = \begin{cases} 0 & k < f(m-1) \\ \frac{k-f(m-1)}{f(m)-f(m-1)} & f(m-1) \leq k \leq f(m) \\ \frac{f(m+1)-k}{f(m+1)-f(m)} & f(m) \leq k \leq f(m+1) \\ 0 & k > f(m+1) \end{cases} \quad (10)$$

在估计了加权频率系数之后,将梅尔频谱  $S(m)$  进行离散余弦变换(DCT),得到 MFCC 系数  $c_g$ :

$$c_g = \sum_{m=1}^{M-1} S(m) \cos\left(\frac{\pi g(m+0.5)}{M}\right) \quad (11)$$

式中: $c_g$  为第  $g$  个 MFCC 序列, $M$  为滤波器的个数。因为人体日常行为引起的 CSI 信号主要集中在低频阶段,因此在本文中提取了前 12 个序列。

## 2 系统设计

本文首先利用巴特沃斯滤波器去除了原始信号中的高频信号,然后再用 PCA 对信号进行处理,结果为一维时间序列。将预处理过的 CSI 数据当作时间流来提取 MFCC 系数特征,为了提升识别的精度,本文把统计特征(MFCC 的均值、平均标准差和平均变化率)作为分类特征量。此外,CSI 数据中的相位信息携带的信息是不可忽视的,为了充分利用相位信息,在特征向量中还增加了无偏移信号的对数能量,这个无偏移信号是经过相位校准之后得到的信号,最后进行数据匹配和行为判别,总体流程如图 2 所示。

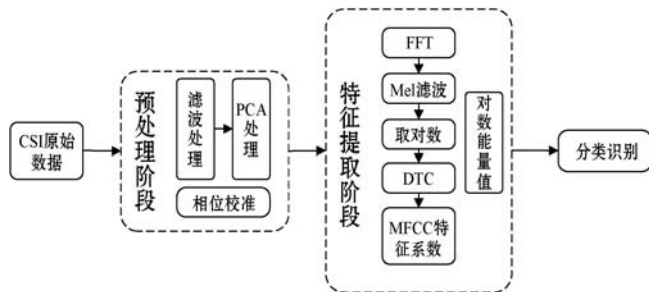


图 2 总体流程

### 2.1 数据预处理

#### 2.1.1 滤波和 PCA 处理

在数据预处理阶段,本文使用巴特沃斯滤波器和 PCA 联合的方式对数据进行预处理,由于多径效应、设备和干扰等多种噪声的存在,原始采集的 CSI 信号毛刺较重,十分不平滑;另外,人的活动属于低频活动,需要把由于环境和设备而带入的高频滤除。经过巴特沃

斯滤波器处理后的数据如图 3 所示。

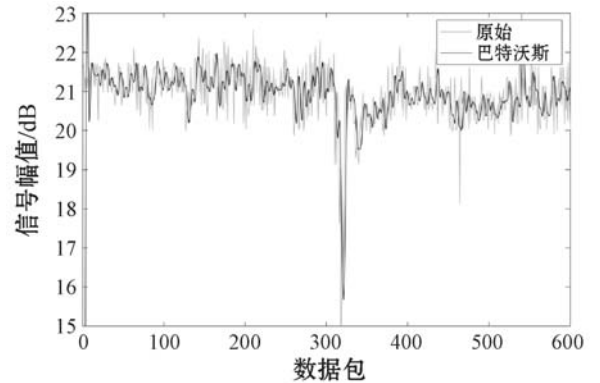


图 3 滤波前后结果对比

在  $t$  到  $t + \Delta t$  的时间内,当信号的传播路径由  $d_k(t)$  变化到  $d_k(t + \Delta t)$  时,假设变化的路径为  $\Delta_k(t)$ ,那么  $\Delta_k(t) = d_k(t + \Delta t) - d_k(t)$ ,即  $d_k(t + \Delta t) = d_k(t) + \Delta_k(t)$ ,其中  $d_k(t)$  是路径的初始长度,设  $\varphi_k = e^{-j2\pi\Delta_k t}$  为子载波的初始相位,那么在  $t + \Delta t$  时刻副载波  $n$  的相位如式(12)所示。

$$\begin{aligned} \cos(2\pi d_k(t + \Delta t)/\lambda_n + \varphi_k) &= \\ \cos(2\pi d_k(t)/\lambda_n + 2\pi\Delta_k(t)/\lambda_n + \varphi_k) &= \\ \cos(2\pi d_k(t)/\lambda_n + \varphi_k) \cos(2\pi\Delta_k(t)/\lambda_n) - & \\ \sin(2\pi d_k(t)/\lambda_n + \varphi_k) \sin(2\pi\Delta_k(t)/\lambda_n) & \quad (12) \end{aligned}$$

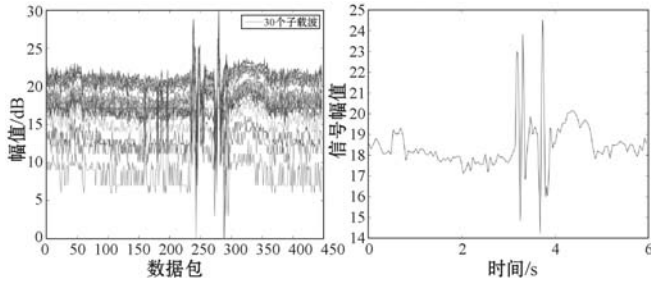
考虑到 Wi-Fi 信道中子载波的波长之间的差异很小,即两个不同子载波的波长存在  $\lambda_1 \approx \lambda_2$  的关系,且  $\Delta_k(t)$  很小,则可以得到以下两式:

$$\cos(2\pi\Delta_k(t)/\lambda_1) \approx \cos(2\pi\Delta_k(t)/\lambda_2) \quad (13)$$

$$\sin(2\pi\Delta_k(t)/\lambda_1) \approx \sin(2\pi\Delta_k(t)/\lambda_2) \quad (14)$$

但是,式(3)中的两个子载波的常数项是不相等的,因为初始路径长度  $d_k(t)$  远大于  $\Delta_k(t)$ ,因此,即使波长略有不同,两个子载波之间的初始相位差不可忽略。因此对于不同的副载波,CSI 是具有不同初始相位的同一组时变波形的线性组合,因此,同一根天线上的 CSI 是相关的。图 4(a) 表示的是走动动作发生时一路天线上 30 个子载波的变化情况,可以直观地看出不同子载波的数据相关性大、冗余度很高。基于 CSI 流的相关性,本文采用了 PCA 对信号进行处理。由上文可知,在采集的 CSI 原始数据中,无论是静止环境还是有行为发生子载波之间都呈现出高度的相关性。但是从采集的角度来说,行为发生的时间只占据了整个采集过程的少部分而静态环境占据绝大部分。在主成分分析时,静态环境的信号与人体行为信号会一起在第一主成分中被捕获。然而,根据式(12),可得知子载波的相位是两个正交分量的线性组合。根据 PCA 处理的正交性,可知在第一主成分中捕获的人体行为信号的所有信息也被捕获在其他主分量中,但是第一

主成分中最多的是静态环境引起的噪声,因此,丢弃第一主成分可以消除静态噪声且人体行为信息又得到保留。本文方法中丢弃了第一主成分,并计算后五个主成分的平均值。经过 PCA 处理以后的走动动作的信号如图 4(b) 所示。



(a) 走动动作 30 个子载波 (b) 走动动作 PCA 处理结果

图 4 PCA 处理结果对比

### 2.1.2 相位校准

CSI 的大量信息存在相位中,相位是不可忽视的一个方面。为了充分利用相位信息,在特征向量中还增加了无偏移信号的对数频谱能量特征,这个无偏移信号是经过相位校准之后得到的信号。子载波在不同的时刻的原始相位无规律可循,也无法使用。由式(3)可知,相位的偏差既有来自于设备自身误差也有来自于人体行为的变化引起的影响。如果将信号传播的路径分为静态和动态的话,在  $t$  到  $t + \Delta t$  的时间内,当信号的传播路径由  $d_k(t)$  变化到  $d_k(t + \Delta t)$  时,式(5)就可以表示为:

$$H(f, t + \Delta t) = e^{-j2\pi\Delta f t} (H_s(f) + \sum_{k=1}^N a_k(f, t) e^{-j2\pi d_k(t + \Delta t)/c}) \quad (15)$$

式中:  $H_s(f)$  表示静态路径的 CSI。本文采用共轭乘法来消除时变随机相位偏移,跌倒动作校准前后的相位如图 5 所示。

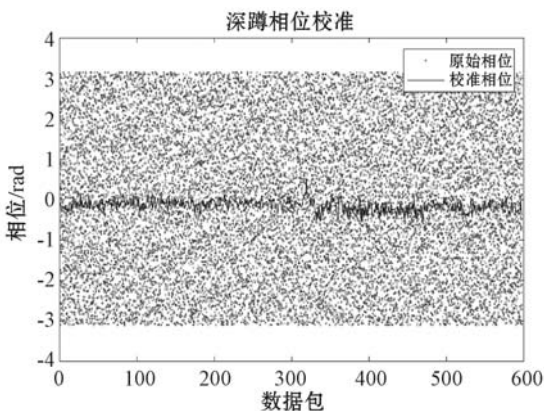


图 5 跌倒动作相位校准

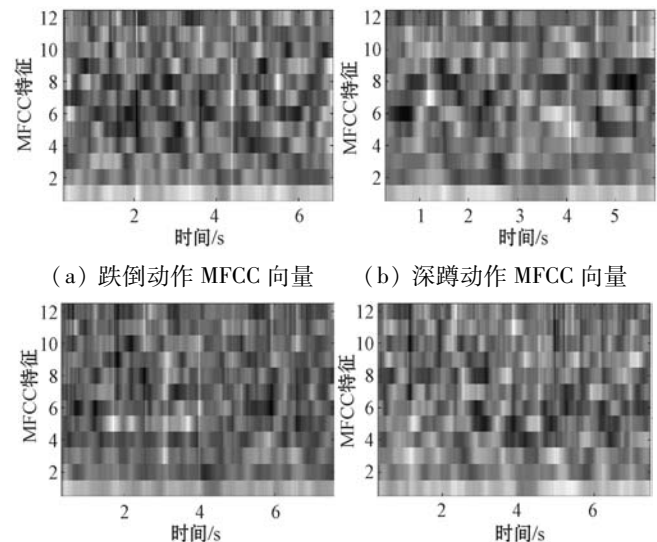
## 2.2 特征提取及行为识别

### 2.2.1 MFCC 特征向量提取

尝试在 CSI 信号中提取 MFCC 特征的原因有以下

四点:(1) 带宽的相关性。在文献[14]对 CARM 数据集的每个日常活动的 CSI 信号平均占用的带宽进行了分析;其中包括了跌倒、深蹲和行走等动作,从统计数据中可以得出所用数据集的所有活动的带宽都在 5 kHz 以内,这和语音信号的带宽和能量分布极为相似。可以说由于活动影响 CSI 信号变化所占用的带宽是人类听力频谱的范围,大约为 2 ~ 20 kHz,这为本文提取 CSI 数据的 MFCC 向量奠定了基础。(2) 信号的相似性。MFCC 向量特征是识别语音信号的一个重要的特征,语音信号能量集中于低频段且变化迅速,从而使用密集的梅尔滤波器组放大这种特性。对于 CSI 信号而言,同样具备这样的性质。例如,当跌倒动作发生时,速度极快,CSI 信号的变化也非常迅速。(3) 传播的相似性,语音和 Wi-Fi 信号分别是声波和电磁波,它们根据波动方程在测量环境中传播。它们遵循来自墙壁和其他障碍物的反射、折射和衍射的相同定律到达了它们的接收器,并由于类似的物理机制而变形。无论是声波和电磁波,都能产生时空共振现象。(4) 在文献[15]中已经证实 OFDM 信号的梅尔频率倒谱既保留了信号的基本特征,也移除了信号的冗余信息,相对原始接收信号的抗噪能力更加鲁棒<sup>[16]</sup>。相关研究表明,基于 MFCC 特征的方法能够有效实现多径信道下 OFDM 信号子载波多种调制方式的识别,且识别性能优于传统方法。

基于声音信号和 CSI 信号的这些共同特征,本文在 CSI 时间序列分类中提取 MFCC 向量特征,不同行为的对数 MFCC 向量如图 6 所示。



(a) 跌倒动作 MFCC 向量 (b) 深蹲动作 MFCC 向量

(c) 跳跃动作 MFCC 向量 (d) 走动动作 MFCC 向量

图 6 MFCC 向量特征

不同行为 CSI 信号的 MFCC 向量互不相同,因此可以将 CSI 信号的 MFCC 向量作为识别不同行为特征参量。为了提高分类的准确性,本文提出使用 MFCC

向量的统计特征作为分类标准,分别是 MFCC 向量的均值、标准差和平均变化率作为分类特征量。根据 1.3 节中对 MFCC 向量的定义, MFCC 向量的均值定义为:

$$E(c) = \sum_{g=1}^G c_g / G \quad (16)$$

式中:  $G$  是 MFCC 向量的个数。

平均标准差定义为:

$$\sigma_g = \left( \sum_{g=1}^G \sqrt{\frac{1}{I} \left( \sum_i c_g^2(i) - \left( \frac{1}{I} |c_g(i)| \right)^2 \right)} \right) / G \quad (17)$$

式中:  $I$  表示帧数。

平均变化率定义为:

$$v_g = \frac{\left( \sum_{i=1}^I \left| c_g(i) - \frac{\sum_{i=1}^I c_g(i)}{I} \right| \right)}{\sum_{i=1}^I c_g(i)} \quad (18)$$

### 2.2.2 对数频谱能量

除了 MFCC 向量是统计特征外,本文还提取了可以忽略相位误差信息的对数频谱能量作为额外的特征。文献[11]研究表明,CSI 的频谱能量  $\|H(f,t)\|^2$  是一系列振幅随动态路径长度变化且频率随动态路径速度变化的余弦信号与直流分量之和,如式(19)所示。

$$\|H(f,t + \Delta t)\|^2 = |H_s(f)|^2 + \sum_{k=1}^N |a_k(f,t + \Delta t)| + \sum_{k=1}^N 2 |H_s(f)| \cdot |a_k(f,t + \Delta t)| \cos\varphi_x \quad (19)$$

经过相位校准后,  $\varphi_x$  可视为一个常数,这说明无偏移信号的  $\|H(f,t)\|^2$  变化只与速度有关,提取  $\|H(f,t)\|^2$  就可以合理地忽略相位误差带来的影响。在我们的实验中,不同行为的速度是有差异的,像走动、坐站动作速度较小,跳跃和跌倒等动作速度较快。可以根据频谱能量将行为分类。但是对于同类速度的行为而言,如果速度相似,那么仅用频谱能量就会出现误分。为减少计算量以及更加放大同类速度行为的差别以便于更好地分类,因此添加了对数频谱能量,定义如式(20)所示。

$$E = \log \|H(f,t)\|^2 \quad (20)$$

式中:  $H(f,t)$  表示经相位校准之后的 CSI。

### 2.2.3 识别方法

本文根据 ACO 和 PSO 优势互补特性,对原始 SVM 算法的惩罚参数和核函数参数进行优化<sup>[10]</sup>,具体流程如图 7 所示。在实验中,ACO 参数设置为:  $\rho = 0.65, Q = 70, c_1 = c_2 = 1.5$ ,种群规模为 30,迭代次数为 45。PSO 参数设置为:  $c_1 = c_2 = 1.5$ ,种群规模为 28,进化次数为 270。

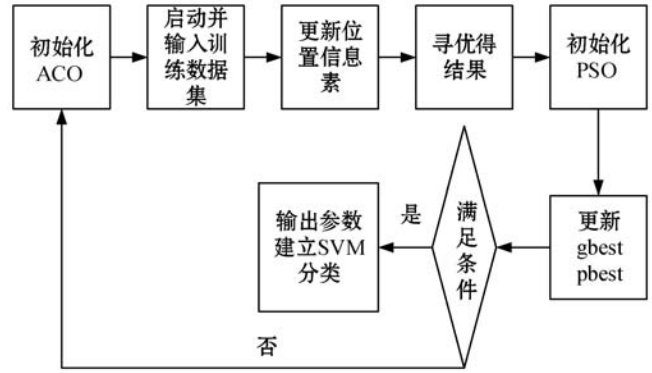


图 7 算法流程

## 3 实验与结果分析

### 3.1 实验部署

实验的收发装置为两台配有 Intel 5300 网卡的主机,其中发射天线网卡上配有 1 根全向天线,接收天线网卡上均配有 3 根全向天线,为了得到稳定且采样率较高的 CSI 数据,采用 Linux 802.11n CSI tool Monitor 模式对网卡固件进行修改,带宽为 40 MHz,中心频率为 5 755 MHz。

实验的场景设置在环境较为复杂空间的会议室,平面图如图 8 所示。把收发装置放于 50 cm 高的桌子的对角,收发装置之间的距离始终保持 4.5 m,本实验中邀请了 3 名志愿者,在收发装置的视距(图 8 中位置 1)和非视距(图 8 中位置 2 和位置 3)位置分别进行坐站、走动、跳跃、深蹲和跌倒五个动作,每名志愿者在每个位置重复每个动作 30 次,每次采集时间为 30 s,第 15 s 的时候开始做动作,第 30 s 时候停止,重复此过程,接着采集下一个,总共采集 1 350 个样本。其中: 1 000 个样本用来训练;350 个样本用来测试。

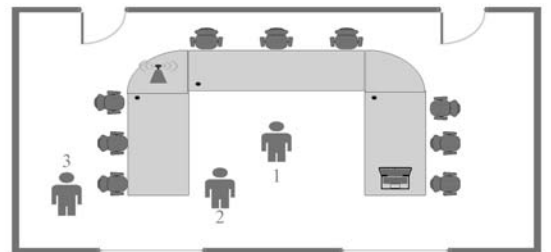


图 8 实验场景

### 3.2 实验结果分析

#### 3.2.1 分类算法的性能分析

表 1 表示的是用 SVM 和 ACO + POS-SVM<sup>[10]</sup> 对志愿者的五种行为(坐站、走动、跳跃、深蹲和跌倒)进行上述特征识别的结果。结果表明,使用 ACO + POS-SVM 比 SVM 平均精度分别提高了 5.1 百分点、4.8 百分点、5.3 百分点、4.5 百分点和 5.1 百分点。整体而

言,使用 ACO + POS-SVM 比 SVM 总平均精度提高了 4.8 百分点,这说明使用 ACO + POS-SVM 算法优于传统的 SVM 算法。用 ACO + POS-SVM 识别效果最好为跌倒行为平均达到了 93.4%,这是因为跌倒行为与其他行为相比,下降趋势更明显、速度更快、能量高度集中,所提取的 MFCC 向量也更容易识别。而跳跃和深蹲两个动作识别率比较低,分别是 85.5% 和 86.3%,由于跳跃和深蹲两个动作都是相似的运动,跳跃动作自下而上,深蹲动作自上而下,而且维持的时间相似以及速度相当,故系统会出现误报情况。

表 1 算法对比(%)

算法	坐站	走动	跳跃	深蹲	跌倒
SVM	83.7	86.8	80.2	81.8	88.3
ACO + POS-SVM	88.8	91.6	85.5	86.3	93.4

### 3.2.2 视距和非视距位置的性能分析

图 9 表示的是基于本文方法对会议室场景不同位置的五种行为的识别结果,其中,在视距条件下(位置 1),坐站、走动、跳跃、深蹲和跌倒平均识别结果分别为 90.2%、93.9%、87.2%、88.4% 和 95.4%,在非视距(位置 2 和位置 3)的行为识别结果会有所下降,但是走动和跌倒的识别结果都在 90% 以上。走动行为在五个行为里持续时间最长,跌倒行为速度最大,也使得这两个行为的识别结果较好。在视距情况下,只有一条直射路径,使得能量高度集中,相比于非视距情况而言,会有更加好的识别效果。在非视距状态下,实验采集了距发射机不同距离的两个位置,位置 2 和位置 3 的平均识别结果分别为 88.3% 和 88%,实验结果表明,非常相似的行为坐站、跳跃和深蹲三个动作的识别结果都在 88% 以上,且在位置 2 和位置 3 的识别结果无明显差距。在实验中,位置 2 和位置 3 距离收发机的距离是有区别的,这也说明本文的识别方法在静态多径没有变化的情况下,如果仅仅由于位置发生了变化,识别效果不会下降太多,且在非视距的位置识别结果具有相似性。同时也说明了本文的识别方法对活动的人体行为识别具有一定的鲁棒性,对于实际情况而言,相比于其他只有视距情况的方法更具有实用性。

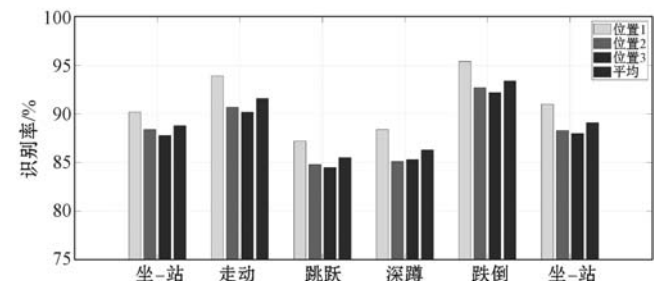


图 9 位置识别结果对比

### 3.2.3 特征对性能的影响

不同行为 CSI 信号的 MFCC 向量互不相同,因此可以将 CSI 信号的 MFCC 向量作为识别不同行为特征参量。为了提高分类的准确性,本文提出了使用 MFCC 向量的统计特征作为分类标准,分别是 MFCC 向量的均值、标准差和平均变化率作为分类特征量。图 10 为单用 MFCC 向量识别和用统计特征识别的结果。结果表明,用统计特征识别准确率分别提高 3.8 百分点、4.4 百分点、4.8 百分点、5.2 百分点和 4.7 百分点。其中跳跃和深蹲动作提高明显,这说明了本文提出的基于 MFCC 向量的统计值比单独用 MFCC 向量识别具有一定的优越性。

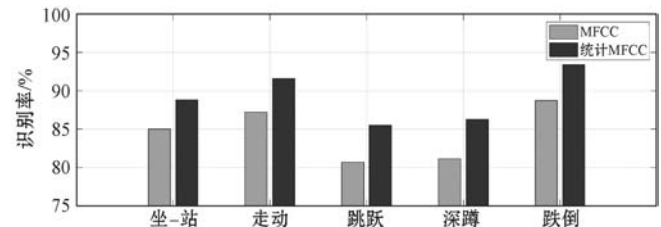


图 10 统计特征对性能的影响

除了 MFCC 向量特征外,本文还额外增加了一个无偏移信号的对数频谱能量作为特征,图 11 表示的是此对数能量对整个识别结果的影响。对于总体数据而言,有对数能量的平均识别结果为 89.1%,反之,识别结果仅为 81.6%。增加了额外对数频谱能量特征,对于比较相似动作的区分是非常有用的。如图 11 所示,跳跃和深蹲动作的识别分别提升了 7.4 百分点和 6.1 百分点,这与 2.2.2 节的理论相符。跌倒动作在视距和非视距情况下平均识别效果可达 93.4%。可见,增加无偏移信号的对数频谱能量特征给整体识别结果带来了提升,使系统更容易区分速度相似的动作。

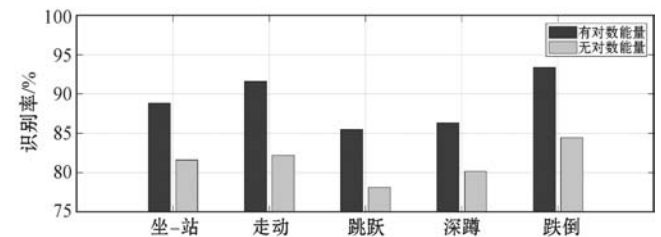


图 11 有无对数能量性能对比

本文还将实验结果与时频特征和小波特征进行了对比,如表 2 所示。结果表明,在预处理相同的前提下,本文方法均优于时频特征和小波特征识别结果。

表 2 方法对比(%)

行为	本文方法	时频特征	小波特征
SVM	82.3	79.5	73.2
ACO + POS-SVM	89.1	83.4	77.6

## 4 结 语

本文参考了日常行为的 CSI 信号与语音信号传播的相似性以及共享带宽的性质,利用梅尔滤波器组能使低频率范围(0 ~ 10 kHz)的能量集中更加明显的特性,提出在 CSI 信号中提取 MFCC 向量统计特征作为识别日常人体行为识别方法。本文首先对数据做了滤波处理,论证了 CSI 信号流之间的相关性并用 PCA 对其处理,相位校准处理;然后在 CSI 中提取了 MFCC 向量统计特征和无相位偏移信号的对数频谱能量作为分类特征;最后用 ACO + POS-SVM 算法对在复杂场景下的视距和非视距情况的数据进行了识别,此方法对日常几个活动的分类达到了较好的效果。但是此实验中,识别的精度还有限,且仅考虑了一个人识别结果,这是比较局限的,未来的工作将会在实时活动识别、多人及更复杂的条件下推进,增加识别系统的广泛性与实用性。

## 参 考 文 献

[1] Wang Y, Wu K, Ni L M, et al. WiFall: Device-free fall detection by wireless networks [J]. *IEEE Transactions on Mobile Computing*, 2017, 16(2): 581 - 594.

[2] Gu Y, Liu T, Li J, et al. EmoSense: Data-driven emotion sensing via Off-the-Shelf WiFi devices [C]//International Conference on Communications, 2018.

[3] Zheng X, Wang J, Shangguan L, et al. Smokey: Ubiquitous smoking detection with commercial WiFi infrastructures [C]//IEEE International Conference Computer and Communications, 2016.

[4] 鲁翔. 基于 WiFi 信道状态信息的室内人员感知研究与实现[D]. 成都:电子科技大学, 2019.

[5] Wang Z, Jiang K, Hou Y, et al. A survey on CSI-Based human behavior recognition in through-the-wall scenario [J]. *IEEE Access*, 2019, 7(8): 772 - 793.

[6] Liu W, Wang S, Lin W. A multi-person behavior feature generation model based on noise reduction using Wi-Fi [J]. *IEEE Transactions on Industrial Electronics*, 2020, 67(6): 5179 - 5186.

[7] Wu X, Chu Z, Yang P, et al. TW-See: Human activity recognition through the wall with commodity Wi-Fi devices [J]. *IEEE Transactions on Vehicular Technology*, 2019, 68(1): 306 - 319.

[8] Akhtar Z U A, Wang H. Wi-Fi-based driver's activity monitoring with efficient computation of radio-image features [J]. *Sensors*, 2020, 20(5): 1381.

[9] Gao Q, Wang J, Ma X, et al. CSI-Based device-free wire-

less localization and activity recognition using radio image features [J]. *IEEE Transactions on Vehicular Technology*, 2017, 66(11): 346 - 356.

- [10] 李爱莲, 郭志斌, 解韶峰, 等. 蚁群和粒子群混合优化 SVM 的钢板表面缺陷分类研究 [J]. *中国测试*, 2020, 46(1): 110 - 116.
- [11] 许杨. 基于信道状态信息的人体行走感知技术研究 [D]. 合肥: 中国科学技术大学, 2019.
- [12] Han W, Chan C, Choy C, et al. An efficient MFCC extraction method in speech recognition [C]//2006 IEEE International Symposium on Circuits and Systems (ISCAS), 2006.
- [13] Ma Y, Zhou G, Wang S, et al. SignFi: Sign language recognition using WiFi [J]. *Proceedings of the ACM on Interactive*, 2018, 2(1): 23 - 28.
- [14] Tegou T, Papadopoulos A, Kalamaras L, et al. Using auditory features for WiFi channel state information activity recognition [J]. *SN Computer Science*, 2020, 1(1): 3.
- [15] 张海川, 雷迎科. 基于 MFCC 的 OFDM 信号子载波调制方式识别方法 [J]. *弹箭与制导学报*, 2017, 36(6): 166 - 170.
- [16] 李兵兵, 杨吉, 刘明骞. DVB-T 系统中 OFDM 信号子载波调制方式的识别 [J]. *华南理工大学学报(自然科学版)*, 2014, 42(3): 21 - 26.

## (上接第 113 页)

- [12] Pan J, Xu J, Ruiz A L, et al. Field-weighted factorization machines for click-through rate prediction in display advertising [C]//2018 World Wide Web Conference, 2018: 1349 - 1357.
- [13] Cheng H T, Koc L, Harmsen J, et al. Wide & deep learning for recommender systems [C]//1st Workshop on Deep Learning for Recommender Systems, 2016: 7 - 10.
- [14] Guo H, Tang R, Ye Y, et al. DeepFM: A factorization-machine based neural network for CTR prediction [C]//26th International Joint Conference on Artificial Intelligence, 2017: 1725 - 1731.
- [15] Lian J, Zhou X, Zhang F, et al. XDeepFM: Combining explicit and implicit feature interactions for recommender systems [C]//24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, 2018: 1754 - 1763.
- [16] Wang R, Fu B, Fu G, et al. Deep & cross network for ad click predictions [C]//ADKDD'17, 2017: 1 - 7.
- [17] Cortes C, Vapnik V. Support-vector networks [J]. *Machine Learning*, 1995, 20(3): 273 - 297.
- [18] Qu Y, Cai H, Ren K, et al. Product-based neural networks for user response prediction [C]//IEEE International Conference on Data Mining, 2016: 1149 - 1154.
- [19] Lobo J M, Jiménez V A, Real R. AUC: A misleading measure of the performance of predictive distribution models [J]. *Global Ecology and Biogeography*, 2008, 17(2): 145 - 151.