

撤稿声明

撤稿文章名: 基于无线传感器网络的室内人体日常动作识别
作者: 罗晓牧, 谭火媛

* 通讯作者: 邮箱: tanhuoyuan@gmail.com

期刊名: 无线通信 (HJWC)
年份: 2017
卷数: 7
期数: 2
页码 (从X页到X页): 53-69
DOI (to PDF): <https://doi.org/10.12677/HJWC.2017.72008>
文章ID: 1730237
文章页面: <https://www.hanspub.org/journal/PaperInformation.aspx?paperID=20317>
撤稿日期: 2017-6-19

撤稿原因 (可多选):

- ☐ 所有作者
☐ 部分作者:
☐ 编辑收到通知来自于
☐ 出版商
☐ 科研机构:
☐ 读者:
☒ 其他:

撤稿生效日期: 2017-6-19

撤稿类型 (可多选):

- ☐ 结果不实
☐ 实验错误
☐ 数据不一致
☐ 分析错误
☐ 内容有失偏颇
☐ 其他:
☐ 结果不可再得
☐ 未揭示可能会影响理解与结论的主要利益冲突
☐ 不符合道德
☐ 欺诈
☐ 编造数据
☐ 虚假出版
☐ 其他:
☐ 抄袭
☐ 自我抄袭
☐ 重复抄袭
☐ 重复发表 *
☐ 侵权
☐ 其他法律相关:
☐ 编辑错误
☐ 操作错误
☐ 无效评审
☐ 决策错误
☐ 其他:
☒ 其他原因:

出版结果 (只可单选)

- ☐ 仍然有效.
☒ 完全无效.

作者行为 失误(只可单选):

- ☐ 诚信问题
☐ 学术不端
☒ 无 (不适用此条, 如编辑错误)

* 重复发表: "出版或试图出版同一篇文章于不同期刊."

历史

作者回应:

☐ 是, 日期: yyyy-mm-dd

☒ 否

信息改正:

☐ 是, 日期: yyyy-mm-dd

☒ 否

说明:

“基于无线传感器网络的室内人体日常动作识别”一文刊登在 2017 年 4 月出版的《无线通信》2017 年第 7 卷第 2 期第 53-69 页上。因统计数据有误等问题, 根据国际出版流程, 编委会现决定撤除此稿件, 保留原出版出处:

罗晓牧, 谭火媛. 基于无线传感器网络的室内人体日常动作识别[J]. 无线通信, 2017, 7(2): 53-69.
<https://doi.org/10.12677/HJWC.2017.72008>

所有作者签名:

罗晓牧 2017.6.11

谭火媛 2017.6.11

Indoor Human Activity Recognition Using Wireless Sensor Networks

Xiaomu Luo*, Huoyuan Tan

School of Medical Information Engineering, Guangzhou University of Chinese Medicine, Guangzhou Guangdong
Email: *woodwood2000@163.com

Received: Apr. 9th, 2017; accepted: Apr. 25th, 2017; published: Apr. 30th, 2017

Abstract

Human locomotion tracking and activity recognition in the indoor environment are fundamental problems for healthy aging. In this paper, we propose a method to deal with these two coherent problems simultaneously by constructing a wireless sensor network (WSN). In the WSN, the Field of View (FOV) of each Pyroelectric Infrared (PIR) sensor is modulated by two degrees of freedom (DOF) segmentation, which provides coarse location information of the human target. Data fusion of the adjacent sensor nodes enlarges the monitored region and improves the human localization accuracy. To incorporate the activity lasting time as a crucial cue for activity recognition, we build a two-layer Random Forest (RF) classifier. The first layer is utilized to label the activity type for each data frame, and the second layer will utilize the lasting time of the same activity as a useful feature for the final activity classification. We conducted experiments in a mock apartment, and the average mean accuracy for 10-fold cross validation of 5 kinds of daily activities is above 93%. The encouraging results show that our method can achieve human tracking and daily activity recognition simultaneously and effectively.

Keywords

Wireless Sensor Network, Pyroelectric Infrared Sensor, Locomotion Tracking, Activity Recognition, Random Forest

基于无线传感器网络的室内人体日常动作识别

罗晓牧*, 谭火媛

广州中医药大学, 医学信息工程学院, 广东 广州
Email: *woodwood2000@163.com

收稿日期: 2017年4月9日; 录用日期: 2017年4月25日; 发布日期: 2017年4月30日

*通讯作者。

文章引用: 罗晓牧, 谭火媛. 基于无线传感器网络的室内人体日常动作识别[J]. 无线通信, 2017, 7(2): 53-69.
<https://doi.org/10.12677/hjwc.2017.72008>

摘要

实现室内人体定位跟踪与动作智能识别在人口老龄化社会具有重要的现实意义。本文提出了一种通过构造无线传感器网络(Wireless sensor network, WSN)同时解决这两个相关问题的方法。在WSN中, 每个热释电红外(Pyroelectric Infrared, PIR)传感器的视场(Field of View, FOV)通过两个自由度(Degrees of freedom, DOF)分割来实现调制, 通过位置信息的编码解码实现了人体目标的粗略定位。通过相邻两个传感器节点的数据融合扩大了监测区域, 同时提高了人体定位的精确度。动作的持续时间是动作识别的一个关键特征, 为此本文构造了一个两层的随机森林(Random Forest, RF)分类器。第一层随机森林用于识别每个数据帧的动作类型, 第二层随机森林利用相同动作的持续时间作为有效的特征判断最终的动作类型。实验在真实的室内环境中进行, 5种日常动作的10折交叉验证平均准确率高于93%。实验结果表明本文提出的方法可以同时有效地实现人体定位跟踪与日常动作识别。

关键词

无线传感器网络, 热释电红外传感器, 定位跟踪, 动作识别, 随机森林

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

1.1. 研究背景

世界人口迅速老龄化, 老年人独居的比例远高于其他年龄段的人[1], 提高老年人的生活质量非常必要并具有重大的社会意义。通过机器学习的方法实现室内的人体定位跟踪与日常动作识别是实现智能家居的基础, 是提高老年人生活质量的有效途径。目前用于日常动作识别的传感器主要有两种: 基于视觉和基于可穿戴的传感器。基于视觉的系统可以观察到完整的人体运动[2] [3], 但是由于其内在的数据关联问题和需要处理大量的数据, 通过视频识别人体日常行为十分具有挑战性[4]。视觉数据通常易受环境因素的影响, 例如变化的光照条件和遮挡。在现实生活中, 人们对传感器的可接受程度也至关重要。生活在摄像头的监控下, 许多人会觉得很不自在, 认为个人隐私受到了侵犯[5]。

相比起摄像头, 人们可能更易于接受可穿戴式传感器。与基于视觉的动作识别系统相比, 基于可穿戴式传感器的系统没有数据关联问题, 并且需要处理的数据量较少。然而人们通常每天都需要换洗衣服, 这很可能使得他们忘记再次戴上传感器[6]。即使电源管理模块经过精心设计, 传感器内的电池也必须定期充电或更换, 这对于用户来说也是不方便的[7]。因此, 当基于视觉的系统不适宜部署, 可穿戴式传感器又不方便使用时, 如何使用简单的传感器实现日常动作识别具有重大的研究意义。

热释电红外(Pyroelectric Infrared, PIR)传感器可以实现泛在感知。人们对 PIR 传感器并不陌生, 因为它们作为安全系统的一部分已经应用在诸多场景, 包括家庭、银行、图书馆等[8]。它们具有很多的优点: 1) 价格便宜并且可以方便地安装在任何的室内环境; 2) 不需要被佩戴或携带, 省去了经常充电的麻烦; 3) 它们是被动的红外传感器, 其性能不受照明变化的影响[9]。然而, 对于单个 PIR 传感器其输出为类正弦波信号, 只能用于检测人体动作是否发生; 所以, 必须重新设计感知模式和分类器算法, 才能充分发挥 PIR 传感器的潜能。

在本文中, 我们提出了一种可以从传感器数据流同时提取位置信息和动作信息的方法。为了监测室内环境, 本文构建了一个无线传感器网络(Wireless Sensor Network, WSN)。在 WSN 中, 传感器节点由 PIR 传感器组成并部署在天花板上。每个 PIR 传感器的视场(Field of View, FOV)经过轴方向和圆形半径这两个自由度(Degrees of Freedom, DOF)的调制可以实现人体目标运动时空信息的编码。相邻传感器节点的数据融合可以提高定位的精度, 同时也可以获取人体运动的速度信息。为了实现人体动作的分类, 本文提出了由两层随机森林(Random Forest, RF)构成的分类器。基于人体目标的位置和移动速度, 第一层随机森林标记每个数据帧的动作类型。为了结合每种动作持续时间的先验知识, 我们使用有限状态机(Finite-State Machine, FSM)来记录相同连续动作帧的持续时间。位置、速度、持续时间这三个特征将作为第二层随机森林的输入, 其输出结果就是最终的动作识别结果。

本文所提出的方法有以下几个优点: 第一, PIR 传感器嵌入在室内环境中而无需随身携带, 可以将传感器的入侵感降低到最低程度; 第二, 动作识别只需要处理低维的数据, 与基于视觉的系统相比计算复杂度显著降低; 第三, 通过非线性的多层分类器进行动作识别, 无需手动预先划分不同动作可能发生的区域。

本文的余下章节安排如下: 第 2 节描述传感模型的设计和实现; 第 3 节给出用于人体定位的数据融合策略; 第 4 节展现两层随机森林的设计, 每层的特征以及数据处理的概述; 第 5 节是本文的实验与讨论。

1.2. 相关工作

近几年来, 研究人员在人体日常动作识别领域取得了巨大的进步。传统的人体日常动作识别是基于视觉信息的。基于视觉识别的典型方法主要有两个步骤, 特征提取和模式识别[4]。然而, 为了避免复杂的图像处理, 一些研究人员应用其它可穿戴或二进制传感器来替代视觉传感器。Wilson 等人[10]提出了同时进行目标跟踪和动作识别的问题。他们使用四种二进制传感器, 包括运动检测器、光断续传感器、压力垫和接触开关; 动态贝叶斯网络用于对在室内的定位跟踪和动作识别进行建模, 并使用 Rao-Blackwellised 粒子滤波器(RBPF)来解决数据关联问题。Zhu 等人[11]结合运动数据和位置信息来推断人体的动作。在他们的研究中, 惯性传感器被附在人体躯干上以提供运动数据, 另一个光学运动捕捉系统提供人体的位置信息。他们分别使用神经网络和隐马尔可夫模型对动作进行粗粒度与细粒度的识别, 进一步使用贝叶斯定理融合动作和位置信息。

由于 PIR 传感器结构简单和对光照差异的鲁棒性, 近年来受到越来越多的关注。在[12]中, Hao 等提出使用侧视的 PIR 传感器节点来对人体目标进行定位。每个传感器节点可以检测移动的热源目标的角位移, 多个传感器节点可以提高检测的准确度。他们使用相同的硬件设备实现了多人跟踪[13], 基于 Expectation-Maximization-Bayesian 的追踪方案和信息融合策略提高了系统的性能, 传感器节点的部署促进了数据-对象-关联的过程。Yang 等[14]在基于信用(Credit-based)的多人定位方法基础上提出了 PIR 传感器跟踪系统。他们使用特殊的光锥来实现传感器节点的花瓣(petal-shape)感知模型, 所有的交点都被视为是测量点。通过聚类分析和具有最高信用(highest credit)的数据关联可以定位人体目标。此方法避免了区域分区和区域分类器。Xiong 等[15]提出了在时域和频域中应用经验模态分解(empirical mode decomposition)和 Hilbert-Huang 变换提取移动人体目标的特征, 提取的特征再使用支持向量机(SVM)进行分类。

然而在上述研究中, PIR 传感器是侧视或放置在地面上的, 这意味着它们在实际部署中容易被家具或其他障碍物遮挡。为了克服这个缺点, Luo 等[16]实现基于安装在天花板的 PIR 传感器节点的人体室内定位。这些 PIR 传感器的 FOV 通过两个 DOF 的空间分割来调制, 这为信息融合提供了灵活的定位模式; 卡尔曼滤波器和平滑器则被用于细化人体运动轨迹。

一些研究人员致力于探索 PIR 传感器用于动作识别的潜力。在[17]中, 伪随机编码的菲涅尔透镜被

用于获取和压缩由 PIR 传感器收集的运动信息, 高维图像序列的识别问题可以转化为低维传感器数据流来处理。Luo 等[18]提出了一种利用视场调制的 PIR 阵列进行异常行为检测的方法。人体活动的时空特征被编码为低维数据流, 并且采用自适应谱聚类(self-tuning spectral clustering)算法来聚类类似的训练样本。One-Class 支持向量机(One-Class SVM)用于描述正常活动和检测异常活动。Guan 等[19]采用 PIR 传感器捕获人体活动引起的红外辐射变化。三个传感模块用于构造一个多视角的红外移动感知系统, 包括一个安装在天花板上的模块和两个在三脚架上彼此面对的模块; 隐马尔可夫模型和支持向量机则用于对六种类型的日常活动进行分类。

上述对 PIR 传感器的研究都分别集中在人体定位或人体行为的识别, 而本文则致力于提供一个框架来同时解决这两个相互关联的问题。

2. 传感器的设计

传感器节点是 WSN 的基础和基本元素。本节将介绍传感器节点设计背后的思想, 包括传感模型的设计和传感器节点的实现。传感器节点设计的目的是捕获人体活动的时空运动特征。

2.1. 感知模型

本文的感知模型源自于参考结构层析成像(Reference Structure Tomography, RST)的范式, 它使用多维调制来编码辐射对象和测量之间的映射[20]。本文感知模型的原理图如图 1 所示。对象空间是指热辐射对象移动的空间, 测量空间指的是放置 PIR 传感器的空间。参考结构指定从对象空间到测量空间的映射[20], 并用于调制每个 PIR 传感器的 FOV。

在使用不透明参考结构的情况下, 可视性函数 $v_j(\mathbf{r})$ 的值是二值型的, 其值取决于在对象空间里的点 $\mathbf{r}$ 是否对第 j 个 PIR 传感器是可见的。

$$v_j(\mathbf{r}) = \begin{cases} 1 & \mathbf{r} \text{ 对于第 } j \text{ 个 PIR 是可见的} \\ 0 & \text{其余情况} \end{cases} \quad (1)$$

假设在测量空间中存在 M 个传感器, 并且它们的 FOV 被复用。因此, 对象空间中的每个点 $\mathbf{r}$ 可以与二进制特征向量 $[v_j(\mathbf{r})] \in \{0,1\}^M$ 相关联, 特征向量指定了对象空间中的点对这 M 个传感器的可见性。在对象空间中, 具有相同特征向量的连续点形成采样胞元。通过这种方式, 三维对象空间 Ω 可以被划分为 L 个离散的非重叠的采样胞元。PIR 传感器的测量值可以写成矩阵的形式:

$$\mathbf{m} = \mathbf{V}\mathbf{s} \quad (2)$$

其中 $\mathbf{m} = [m_j(t)] \in \mathbb{R}^{M \times 1}$ 为测量向量, $m_j(t)$ 是第 j 个 PIR 传感器的测量值; $\mathbf{V} = [v_{ji}] \in \mathbb{R}^{M \times L}$ 是测量矩阵, 并且如果第 i 个采样胞元对第 j 个传感器可见, 则 $v_{ji} = 1$, 否则 $v_{ji} = 0$; $\mathbf{s} = [s_i(t)] \in \mathbb{R}^{L \times 1}$ 是采样胞元的辐射状态。

2.2. 参考结构

传感器节点的设计基于上述的感知模型。一个传感器节点上有 9 个 PIR 传感器, 如图 2 所示。每个 PIR 传感器都配有一个菲涅耳透镜阵列。菲涅耳透镜阵列用于收集其 FOV 内的红外辐射从而增加 PIR 传感器的灵敏度。在进行调制之前, 每个 PIR 传感器的 FOV 为全锥体, 而不透明的掩膜起到参考结构的作用。

在本文的节点设计中, 9 个传感器中的 4 个使用 I 型掩膜, 如图 3(b)所示。每个 PIR 传感器的 FOV 为扇形锥体, 角度为 120° , 仅为全锥的 $1/3$, 如图 3(a)所示。通过这 4 个传感器的复用将传感器节点的监视区域分成 8 个扇形, 每个扇形对应角度 θ 的编码方案列于表 1。

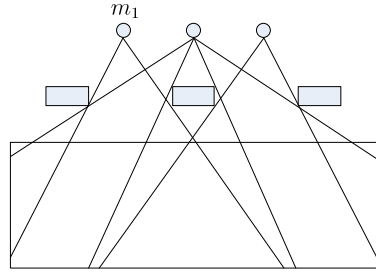


Figure 1. The sensing model: measurement space, reference structure, object space and sampling cells

图 1. 传感器模型: 测量空间, 参考结构, 对象空间和采样胞元

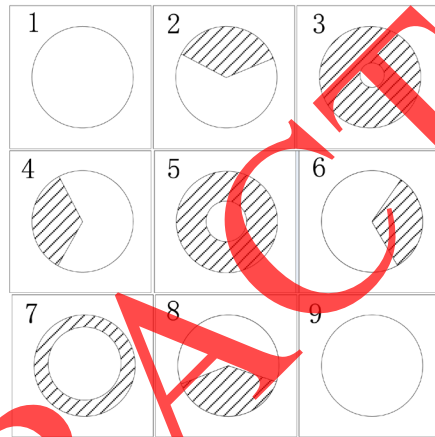


Figure 2. PIR sensor node: There are 9 PIR sensors on one sensor node, 4 of them using type I mask, the rest using type II mask

图 2. PIR 传感器节点: 每个传感器节点有 9 个 PIR 传感器, 其中 4 个使用 I 型掩膜, 其余的使用 II 型掩膜

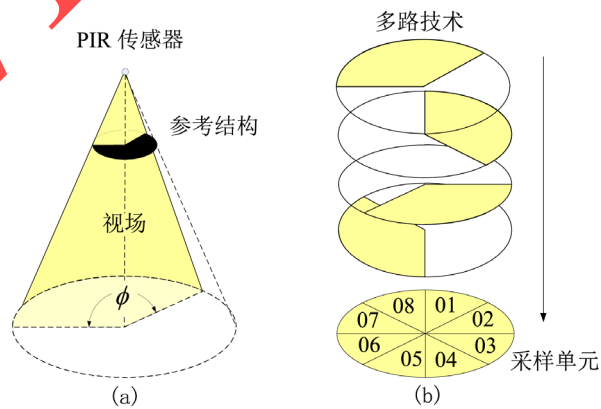


Figure 3. Type I mask: (a) bearing segmentation; (b) multiplexing of 4 PIRs forms 8 sampling cells

图 3. I 型掩膜: (a) 轴方向的分割; (b) 4 个 PIR 传感器形成 8 个采样胞元

Table 1. The coding scheme of angle θ **表 1.** 角度 θ 的编码方案

采样胞元	PIR 6	PIR 2	PIR 4	PIR 8	角度 θ
01	1	0	0	1	67.5°
02	1	0	0	0	22.5°
03	1	1	0	0	337.5°
04	0	1	0	0	292.5°
05	0	1	1	0	247.5°
06	0	0	1	0	202.5°
07	0	0	1	1	157.5°
08	0	0	0	1	112.5°

其余 5 个传感器使用 II 型掩膜掩蔽, 如图 4(a)所示。PIR 传感器的 FOV 仍然是全锥体, 但是其锥角 β 小于原始锥体的锥角, 如图 4(b)所示。通过这 5 个传感器的复用将传感器节点的监视区域分成 5 个环, 这对应于不同的半径 r 。表 2 中列出了半径 r 的编码方案, 其中采用了优先编码策略。

通过使用这两种掩膜, 对象空间被分割为若干个采样胞元, 如图 5(a)所示, 采样胞元的中心 (x, y) 由极坐标 $r \angle \theta$ 表示, 如图 5(b)所示:

$$\begin{cases} x = r \times \sin \theta \\ y = r \times \cos \theta \end{cases} \quad (3)$$

当人体在对象空间中移动时, 将触发若干个 PIR 传感器。根据编码方案, 我们可以解码出人的位置。采样胞元的中心将被认为是人体目标的估计位置; 换言之, 一个传感器节点的定位精度率与采样胞元的投影面积相关。

3. 目标定位

3.1. 信号特征的提取

PIR 传感器的脉冲响应为非线性, PIR 传感器 $m_j(t)$ 的输出为类正弦波, 如图 6 所示。因为人体目标为非刚性可变形的, 所以 $m_j(t)$ 不是平滑的曲线。因此, 不适合使用信号的幅度作为定位的特征。短时能量(STE)已被广泛用于表示类正弦波信号的能量变化[21], 因此本文采用 STE 作为 PIR 信号的特征。第 j 个 PIR 传感器在第 n 个窗口的 STE 被定义为

$$p_j(n) = \sum_{k=0}^{Z_n-1} \|m_j(k) - avSTE_j(n)\| \quad (4)$$

$$\text{with } avSTE_j(n) = \frac{1}{Z_n} \sum_{k=0}^{Z_n-1} m_j(k) \quad (5)$$

其中 $j \in \{1, \dots, M\}$ 是 PIR 传感器的编号, Z_n 为第 n 个窗口中采样点的总数, $avSTE_j(n)$ 为采样点的平均能量, $m_j(k)$ 是第 k 个采样点的信号幅度。在本文的系统实际部署中, $Z_n = 15$ 。

为每个 $p_j(n)$ 设置特定阈值 th_j 以指示 PIR 是否被触发:

$$m_j(n) = \begin{cases} 1 & p_j(n) > th_j \\ 0 & \text{otherwise} \end{cases} \quad (6)$$

Table 2. The coding scheme of radius γ **表 2.** 半径 γ 的编码方案

采样胞元	PIR 1	PIR 5	PIR 9	PIR 3	PIR7	半径 γ (m)
1	1	x	x	x	x	0.0
2	0	1	x	x	x	0.8
3	0	0	1	x	x	1.3
4	0	0	0	1	x	2.3
5	0	0	0	0	1	3.0

“x” 表示 0 或 1。

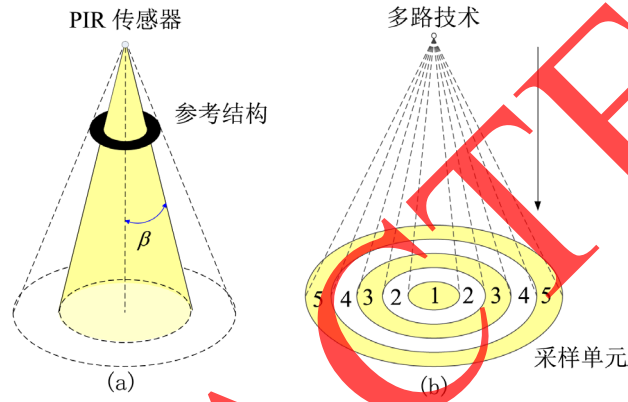


Figure 4. Type II mask: (a) radial segmentation; (b) multiplexing of 5 PIRs forms 5 sampling cells

图 4. II 型掩膜: (a) 径向分割; (b) 5 个传感器的多路复用形成 5 个采样胞元

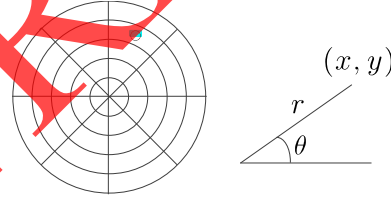


Figure 5. (a) The floor projection of the sampling cells, the segmentation of the object space within one sensor node; (b) the center of the cell, the blue point is denoted by polar coordinates

图 5. (a) 采样胞元在地面上的投影; (b) 采样胞元的中心蓝色点用极坐标表示

对于每个传感器节点的 PIR 传感器阵列的输出 $\mathbf{M}(n) = [m_1(n), \dots, m_5(n)]$, 我们查找表 1 和表 2 确定被触发的采样胞元, 并根据(3)计算人体目标的估计位置。

3.2. 数据融合

在本文的实验部署中, 有 5 个传感器节点安装在天花板上, 形成一个 WSN, 如图 7 所示。每个传感器节点的对象空间相互重叠, 通过数据融合的方法提高定位精度。将 $\mathbf{z}_k$ 定义为第 k 个时间窗口的目标的估计位置 (x_k, y_k) 。定位算法的选择取决于所有传感器节点中同时触发的采样胞元的数量, 如图 8 所示:



Figure 6. The output of PIR sensors in one sensor node for different activities: (a) lying, (b) sitting, (c) standing, (d) walking and (e) transitional activities

图 6. 在做不同动作时一个传感器节点的 PIR 传感器节点输出: (a) 躺, (b) 坐, (c) 站, (d) 走, (e) 过渡动作

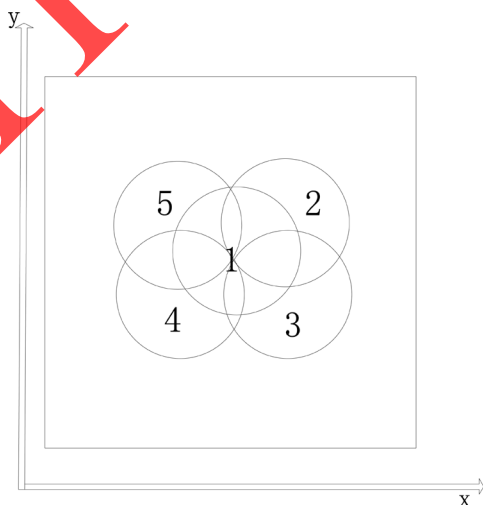


Figure 7. The projection of deployment for the sensor nodes in two-dimensional plan view, where the centers of the circles are the installation position of sensors mounted on the ceiling

图 7. 传感器节点的部署在二维平面上的投影, 其中圆心为传感器在天花板的安装位置

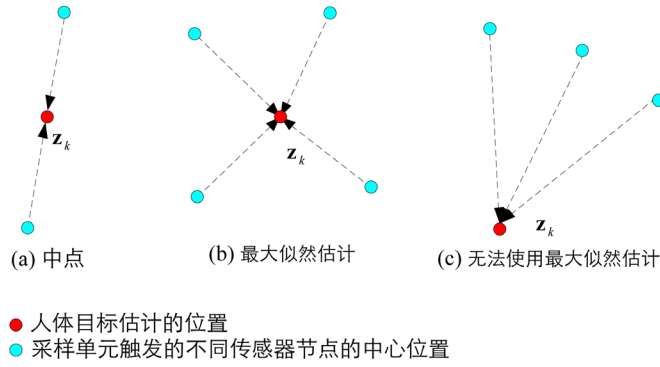


Figure 8. Data fusion of adjacent sensor nodes

图 8. 相邻传感器节点的数据融合

- 触发两个采样胞元：这两个采样胞元中心的中点则视为 z_k ，如图 8(a) 所示。
- 触发两个以上的采样胞元：最大似然估计算法用于计算 z_k ，如图 8(b) 所示。

然而在某种情况下，当人体目标估计位置与三个触发的采样胞元之间的距离是相同时，使用上述的方法计算 z_k 显然是不正确的，如图 8(c) 所示。在这种情况下，我们必须判断估计位置与采样胞元之间的关系：如果估计点 z_k 位于采样胞元中心点所形成的凸区域之外，则这些采样胞元的几何中心将是 z_k 的估计位置。

3.3. 粒子滤波器

在本文中，PIR 传感器的采样频率为 15 赫兹。在使用定位算法之前，数据流将被分割成数据帧，每个数据帧中包含 30 个采样点，时间约为 2 秒。两个连续数据帧之间有 15 个采样点是重叠的，如图 9 所示。在每个数据帧中，窗口的大小为 15 个采样点，定位算法则被应用于这些连续的窗口。因此，在每个数据帧中将得到最多 16 个定位结果。

假设在第 t 个数据帧中有 L 个定位结果，表示为 y_t^l ，其中 $l=1, \dots, L$ 。人体目标的前一个时刻的位置记为 z_{t-1} 。基于粒子滤波器的思想 (Particle Filter, PF) [22]，第 l 个定位结果的权重可以定义为

$$w_t^l = p(y_t^l | z_{t-1}) \quad (7)$$

假设在给定前一个位置 z_{t-1} 的情况下，当前位置 y_t^l 的概率服从二维高斯分布，则(7)可以被改写为

$$w_t^l = N(y_t^l | z_{t-1}, \sigma) \quad (8)$$

其中 σ 是观测模型的标准差。如果 $w_t^l < \gamma$ ， y_t^l 将被丢弃，其中 γ 为验证 y_t^l 是否有效的阈值。因为 w_t^l 的值越小，则其为虚警 y_t^l 产生的概率越高。因此， $L \leq 16$ 。

假设在同一个数据帧内有 L 个有效的位置结果，每个结果的归一化权重为：

$$\tilde{w}_t^l = \frac{w_t^l}{\sum_{l'=1}^L w_t^{l'}} \quad (9)$$

因此，人体目标的当前位置可以表示为

$$z_t = \sum_{l=1}^L y_t^l \tilde{w}_t^l \quad (10)$$

其中 $\tilde{w}_t^l$ 是 y_t^l 的归一化权重。

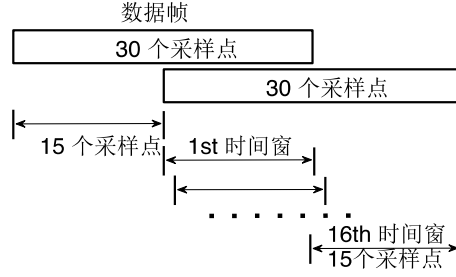


Figure 9. Data frame and window. The size of each data frame is 30 sampling points, and the size of each window is 15, sampling points. The overlap between two successive data frame is 15 sampling points

图 9. 数据帧和窗口。每个数据帧的长度为 30 个采样点，时间窗口的大小为 15 个数据点。两个连续的数据帧之间有 15 个采样点相互重叠

4. 两层的随机森林

4.1. 随机森林

在本文中，随机森林(Random Forest, RF)是我们所构造的多层分类器的基础。RF 可看作 Bagging 算法和决策树算法的结合[23] [24]。在训练阶段，RF 将构建多棵决策树，而每棵决策树的训练样本都是从所有训练样本中随机抽样而来，即采样自举(bootstrap)的算法抽取样本。在训练单棵决策树时，选择所有特征的随机子集，然后从子集中选择一个使信息熵最大的属性用于划分，如图 10 所示。每棵决策树尽可能生长到最深的深度而不进行剪枝。最后将所得的树做平均得到最终结果。RF 是一种无需特征选择、训练时间短、泛化能力强的集成分类器[27]，可用于各种不同的分类任务，是 Bagging 方法最成功的实现之一。决策树的数量和每个节点选择的输入特征的数量是最重要的参数[26]。

4.2. 数据处理流程图

本文所提出的数据处理的流程图，如图 11 所示。从所有传感器节点产生的传感器数据流被分割成数据帧，如图 9 所示。在每个数据帧中，我们根据(10)计算人体目标的位置估计。基于两个连续数据帧的位置 (x_t, y_t) 和 (x_{t-1}, y_{t-1}) ，人体目标在第 t 个数据帧中的移动速度 v_t 可计算为

$$v_t = \sqrt{(x_t - x_{t-1})^2 + (y_t - y_{t-1})^2} / T \quad (11)$$

其中 T 为两个数据帧之间的时间间隔。矢量 $\{x_t, y_t, v_t\}$ 用作第一层 RF 的输入特征；第一层 RF 的输出表示为 L_t^1 。

为了获得相同连续动作的持续时间，我们使用了两个状态的有限状态机(finite state machine, FSM)。FSM 是用于表示有限个状态以及这些状态之间转移和动作等行为的数学模型[34]；其中状态存储关于过去的信息，反映了系统从开始到现在的输入变化；转移表示状态的变更，用使得状态发生改变的条件来描述；动作则是在给定时刻所进行的活动描述。当第一层 RF 的输出改变时，即 L_t^1 的值改变时，表示 FSM 由一个状态跳转至另一个状态，由此可得知 FSM 在同上一个状态所保持的时间。那么连续动作的持续时间 D_t 就可以计算为：

$$\begin{cases} C_t = C_{t-1} + 1, & \text{if } L_t^1 = L_{t-1}^1 \\ C_t = 1 \text{ and } D_{t-1} = C_{t-1}, & \text{if } L_t^1 \neq L_{t-1}^1 \end{cases} \quad (12)$$

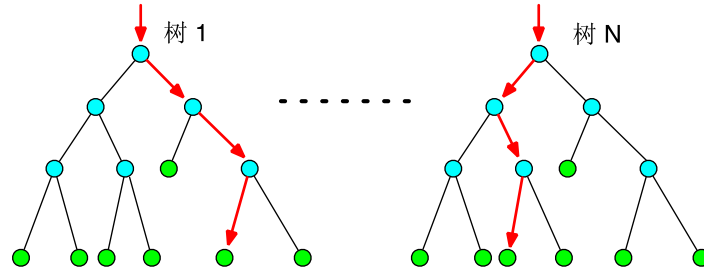


Figure 10. Random Forest. A forest is an ensemble of trees. Each tree consists of split nodes (blue) and leaf nodes (green). The red arrows indicate the different paths that might be taken by different trees for a particular input. The final result is based on the maximum voting among the trees

图 10. 随机森林。一个森林是树的集成。每棵树都包含分割节点(蓝色)和叶子节点(绿色)。每个红色的箭头代表了一个特定输入在不同树中的可能路径。最终的结果是基于所有树的最多数投票

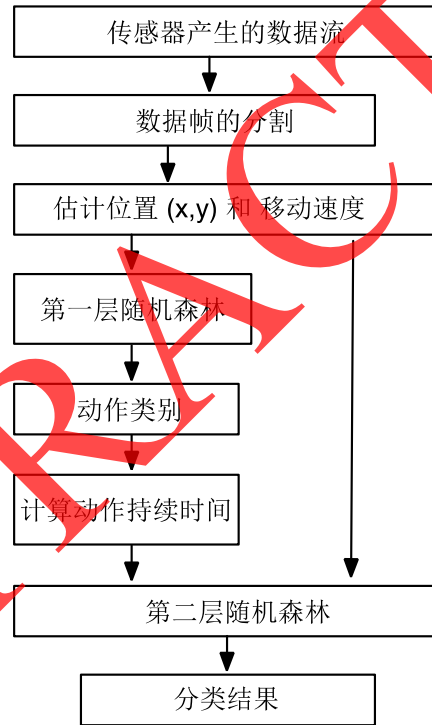


Figure 11. The overview of data processing. The feature vector for the first layer Random Forest is $\{x_t, y_t, v_t\}$. The feature vector for the second layer Random Forest is $\{x_t, y_t, v_t, D_t\}$

图 11. 数据处理流程图。第一层随机森林的特征向量是 $\{x_t, y_t, v_t\}$ 。
第二层随机森林的特征向量是 $\{x_t, y_t, v_t, D_t\}$

其中 C_t 是数据帧 t 的持续时间计数器, D_t 是连续动作的持续时间。例如, 数据帧 $\{m, m+1, \dots, t\}$ 是由第一层 RF 分类出的相同动作, 即 $L_m^1 = L_{m+1}^1 = \dots = L_t^1 \neq L_{t+1}^1$, 则 $D_m = D_{m+1} = \dots = D_t = C_t$ 。

基于第一层 RF 的输出, 我们构造第二层 RF 的特征向量 $\{x_t, y_t, v_t, D_t\}$ 。第二层 RF 的输出是我们系统的最终分类结果。

5. 实验

5.1. 环境设置

本文在一个面积约为 $6\text{ m} \times 6\text{ m}$ 的室内环境中进行了实验。五个传感器节点安装在距离地面 3 米左右的天花板上, 形成一个 WSN, 如图 7 所示。每个 PIR 传感器的采样频率为 15 Hz, 传感器节点生成的数据将通过 2.4 GHz 的 ZigBee 协议发送到汇聚节点[28]。最后, 汇聚节点将数据传输到 PC 主机进行数据处理。这些传感器节点具体的安装位置在表 3 中列出。每个传感器节点覆盖的监视区域大约为地面投影 3m 半径的圆锥。

实验中总共有三个实验人员演示五种日常动作: 行走, 站立, 坐, 躺和过渡动作。过渡动作包括: 坐到站, 站立到坐着, 坐着到躺着, 躺着到坐着等, 共收集了 67 组数据。每组数据集有约持续四分钟的动作。本文使用摄像头记录动作的真实情况, 以评估系统的最终识别结果。

5.2. 实验结果

在本文的实验中, 准确率的计算是基于每个数据帧的识别结果, 如图 9 所示。观察者根据视频标记数据帧的真实类别, 识别结果与真实类别做比较。随机打乱数据集后, 选择 90% 的数据集作为训练数据, 10% 的数据集作为测试数据。本文采用 10 折的交叉验证来评估系统的性能。在每次交叉验证中, 训练集合包含了 14,841 个数据帧, 测试集合有 1487 个数据帧。

一些典型的数据帧如图 13 所示。在每个子列中, 右列是从视频中捕获的快照, 中间列是人在室内空间二维投影上的移动轨迹, 左列分别为从视频中获取的真实类别、RF 的第一层和第二层的分类结果。系统的平均定位误差约为 0.85 m。红色正方形是放椅子的区域, 绿色正方形是放床的位置, 在这两个位置都可能发生坐、躺以及过渡动作。

在左列的每个子图中, 真实类别的标签与动作分类的结果如下: 1. 躺, 2. 坐, 3. 站, 4. 行走, 5. 过渡动作。在图 13(a)中, 人体目标对象站起来, 然后走向床, 动作主要包括坐和行走。在图 13(b)中, 人体目标对象四处走动, 然后坐在椅子上, 动作包括行走和坐。在图 13(c)中, 人体目标对象走到床边, 然后坐在床上, 最终躺在在床上, 动作主要包括行走, 过渡动作和躺。在图 13(d)中, 人体目标对象四处走动, 然后静止站立一段时间, 动作包括行走和站立。在图 13(e)中, 人体目标对象躺在床上一段时间, 然后坐在床上, 这是从躺到坐的过渡动作。

为了比较不同参数设置的效果, 我们计算了 RF 中不同数量决策树 10 折交叉验证的平均准确率和标准差, 如图 12 所示。第二层 RF 明显比第一层 RF 表现要好, 准确率较高, 同时标准差也较低。

根据图 12, 使用 20 棵决策树的随机森林可以在结果的准确率与计算复杂度之间做一个较好的权衡, 本文使用这个参数设置以计算所有活动的混淆矩阵(confusion matrix), 如表 4 和表 5 所示。第一层 RF 和第二层 RF 的平均准确率分别为 82% 和 93%。第一层 RF 和第二层 RF 的标准差分别为 0.05 和 0.01。粗体字的值是对应于特定活动类型的正确分类百分比, 其他数字表示不正确分类的百分比。通过比较这两个表可以看出将每个活动的持续时间作为输入特征添加到第二层 RF 是有效的, 能提高识别的准确率, 同时降低标准差。系统的总体准确率高于 93%, 与基于可穿戴传感器或视频数据的识别方法相当甚至是更高[11] [29] [30] [31]。然而, 由于实验条件的不同, 仅仅比较准确率是不够客观的。

为了比较 RF 与其他分类器的有效性, 我们使用支持向量机(Support Vector Machine, SVM)和朴素贝叶斯(Naive Bayes, NB)分别代替我们的分类器框架中的两层 RF 并进行实验[22] [23]。10 折交叉验证的平均准确率和标准差列于表 6 中。实验表明 SVM 和 NB 的准确率低于 RF。根本原因在于 NB 的决策边界是线性的, 这与在室内空间发生活动的位置是区域性的事实不一致, 如图 13 的中间列所示。在椅子和床

Table 3. Position of sensor nodes
表 3. 传感器的位置

	Node 1	Node 2	Node 3	Node 4	Node 5
位置(x, y)	(0,0)	(2,2)	(2,-2)	(-2,-2)	(-2,2)

Table 4. Mean accuracy of the first layer Random Forest (%)
表 4. 第一层随机森林的平均准确率(%)

真实类别	分类结果					测试准确率
	走	坐	躺	站	过度动作	
走	62.17	37.00	0.74	0.00	0.09	62.17
坐	8.67	90.40	0.80	0.00	0.13	90.40
躺	0.33	2.21	93.80	3.43	0.22	93.80
站	0.00	0.00	0.27	90.25	9.48	90.24
过渡动作	0.06	0.11	0.03	16.73	83.08	83.08

20 棵决策树：平均准确率为 82%，标准差为 0.05。

Table 5. Mean accuracy of the second layer Random Forest (%)
表 5. 第二层随机森林的平均准确率(%)

真实类别	分类结果					测试准确率
	走	坐	躺	站	过度动作	
走	98.75	1.16	0.00	0.00	0.09	98.75
坐	2.01	97.61	0.29	0.00	0.08	97.61
躺	2.21	1.33	93.47	2.77	0.22	93.47
站	0.00	0.00	0.38	87.50	12.13	87.50
过渡动作	0.06	0.14	0.19	9.77	89.84	89.84

20 棵决策树：平均准确率为 93%，标准差为 0.01。

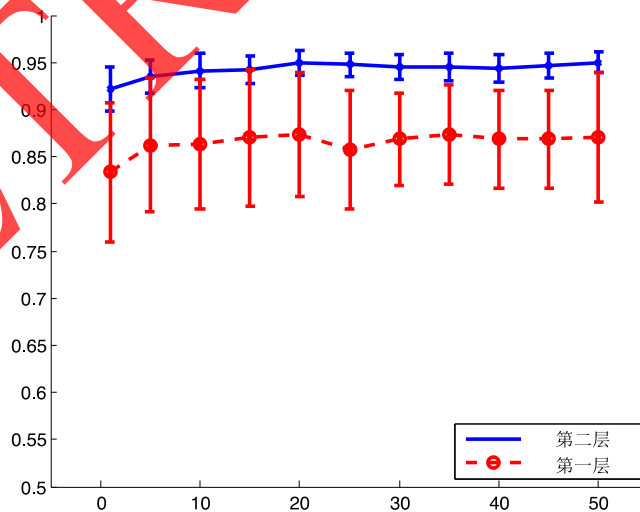


Figure 12. Mean accuracy and standard deviation of 10-fold cross-validation for different number of decision trees in Random Forest. The number of trees includes 1, 10, 15, 20, 25, 30, 35, 40, 45, 50. The blue line is the performance of second layer of RF, and the dash with circle is the performance of the first layer of RF

图 12. 不同棵数的随机森林 10 折交叉验证的平均准确率和标准差。树的棵数分别为 1, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50。蓝色的曲线是第二层随机森林的性能，带有圆的红色虚线是第一层随机森林的性能

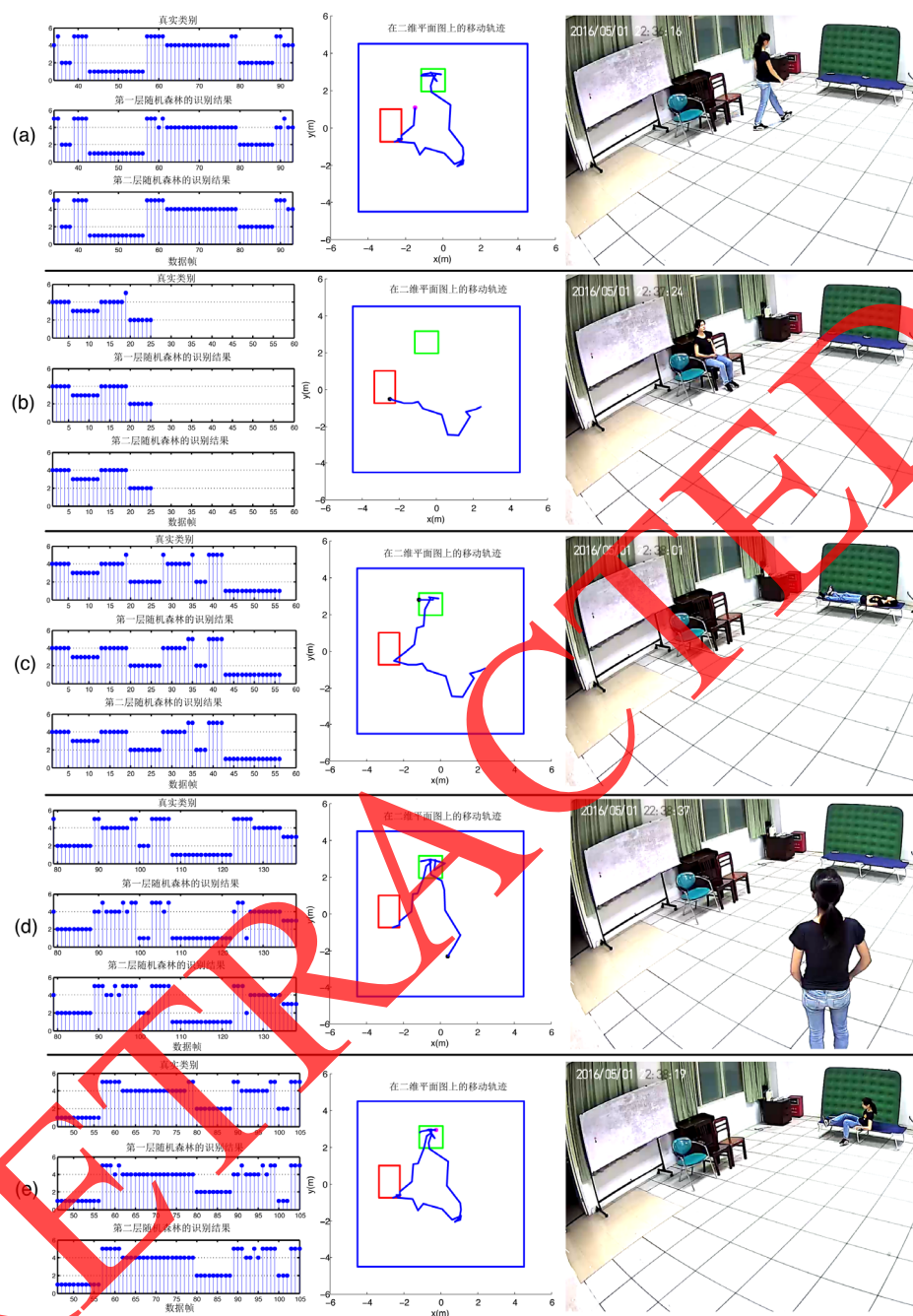


Figure 13. Results captured from video and the server PC. For subfigure (a)-(e), the left column represents the ground truth labels of each second, the results of the first layer random forest, and the results of the second layer random forest, respectively. The middle column is the moving trace of the human target on the map. The red square is the region of chairs, and the green square is the location of the bed. The right column is the snapshot capture from the video, where (a)-(e) represent walking, sitting, lying, standing, and lie-to-sit, respectively. Labels for activity classification result: 1. lying, 2. sitting, 3. standing, 4. walking, 5. transitional activities

图 13. 从视频捕获和服务器识别的结果。对于子图(a)-(e)，左列分别代表真实类别，第一层随机森林的识别结果，第二层随机森林的识别结果。中间列是人体目标的移动轨迹在二维平面上的投影。红色的正方形是放椅子的区域，绿色的正方形是放床的位置。右列是从视频捕获的快照，(a)-(e)分别代表走，坐，躺，站，躺到坐。动作分类结果的标签为：1.躺；2.坐；3.站；4.走；5.过渡动作

Table 6. Comparison of different algorithms**表 6.** 不同算法的比较

	第一层随机森林	第二层随机森林	支持向量机	朴素贝叶斯
平均准确率	0.82	0.93	0.79	0.66
标准差	0.05	0.01	0.03	0.05
训练时间(总秒数/交叉验证)	—	1839.14	1428.41	995.72
测试时间(秒/样本)	—	0.12	5.25e-04	0.06

在每一次交叉验证中, 共有 14841 个数据帧用来做训练, 1487 个数据帧用来做测试。

的区域内, 坐着和躺着发生的概率很高; 在这些区域之外, 行走和站立发生的概率很高。因此, 人体目标的 (x, y) 位置是很好的特征。然而在使用 NB 时, 分别将权重赋给 x 和 y 方向, 在输入特征不是线性可分离的情况下会导致混淆。SVM 的性能优于 NB, 因为由输入特征的内积产生的超平面可以对非线性可分离特征建模。然而, RF 对室内空间的活动区域建模要比 SVM 和 NB 好。RF 由许多决策树组成, 其更易于针对不同活动可能发生的正方形区域进行建模, 因为每棵树的决策阈值可以分别应用于 x 方向和 y 方向。多数表决方案可以提高决策的准确性, 而 RF 的决策边界可以是正方形的, 这与室内空间的布局相一致。此外, RF 还可以组合多样化的特征, 例如位置, 速度和连续活动的持续时间。

本文记录不同算法的训练时间和测试时间, 如表 6 所示。在后台处理器中, 本文使用 Matlab 2013b 进行数据处理。主机的 CPU 是 Intel(R)Core TM i5-6400 2.70GHz 2.70 GHz 和 8.00GB RAM。对于每次交叉验证, 朴素贝叶斯的总训练时间是最短的, 而 RF 是最长的。在测试阶段, SVM 的测试时间最短, RF 最长。然而, 对于 RF, 每个数据帧(2 秒)的测试时间约为 0.12 秒, 这满足实时处理的需求。

5.3. 结论

本文提出了一种利用位置和运动信息之间的协同关系来同时解决人体跟踪和动作识别(STAR)问题的方法[11]。为了充分发挥 PIR 传感器用于智能监护的潜力, 我们通过设计其 FOV 的编码方案来捕获人体目标运动的联合时空信息。本文还设计了对人体行为识别有用的特征, 包括人体位置, 运动速度和相同动作的持续时间。使用两层 RF 框架来对这些特征建模并输出最终的分类结果。通过在真实的室内环境中进行实验, 我们对系统的准确性和稳定性进行评估, 并与其他算法进行比较。本文所提出的方法优点如下: 减少传感器的入侵感, 降低数据通信的负担和数据处理的复杂性, 同时保持动作识别的高准确率。在资源受限的情况下, 如 WSN, 这是至关重要的。在未来的工作中, 研究组将关注多个人体对象和细粒度活动的识别, 进一步研究 PIR 传感器的编码方案, 并且将重新设计参考结构以解决多个目标的数据关联问题。利用更复杂的分层模型, 如多层 CRF [32]和 HMM [33], 对连续活动的顺序约束进行建模。

致 谢

国家自然科学基金资助项目(No.61301294); 广州中医药大学青年英才计划资助项目; 2016 年广东省大学生创新训练计划项目(No. 201610572084)。

参考文献 (References)

- [1] Morgan, L.A., Perez, R., Frankowski, A.C., Nemec, M. and Bennett, C.R. (2016) Mental Illness in Assisted Living: Challenges for Quality of Life and Care. *Journal of Housing for the Elderly*, **30**, 185-198.
<https://doi.org/10.1080/02763893.2016.1162255>
- [2] Turaga, P., Chellappa, R., Subrahmanian, V. and Udrea, O. (2008) Machine Recognition of Human Activities: A Survey. *IEEE Transactions on Circuits and Systems for Video Technology*, **18**, 1473-1488.
<https://doi.org/10.1109/tcsvt.2008.2005594>

- [3] Liu, A.-A., Su, Y.-T., Nie, W.-Z. and Yang, Z.-X. (2015) Jointly Learning Multiple Sequential Dynamics for Human Action Recognition. *PLoS ONE*, **10**, e0130884. <https://doi.org/10.1371/journal.pone.0130884>
- [4] Chaquet, J.M., Carmona, E.J. and Fernandez-Caballero, A. (2013) A Survey of Video Datasets for Human Action and Activity Recognition. *Computer Vision and Image Understanding*, **117**, 633-659. <https://doi.org/10.1016/j.cviu.2013.01.013>
- [5] Ryoo, M.S., Rothrock, B., Fleming, C. and Yang, H.J. (2016) Privacy Preserving Human Activity Recognition from Extreme Low Resolution. arXiv preprint arXiv:1604.03196
- [6] Storm, F.A., Heller, B.W. and Mazza, C. (2015) Step Detection and Activity Recognition Accuracy of Seven Physical Activity Monitors. *PLoS ONE*, **10**, e0118723. <https://doi.org/10.1371/journal.pone.0118723>
- [7] Zhu, C., Sheng, W. and Liu, M. (2015) Wearable Sensor-Based Behavioral Anomaly Detection in Smart Assisted Living Systems. *IEEE Transactions on Automation Science and Engineering*, **12**, 1225-1234. <https://doi.org/10.1109/tase.2015.2474743>
- [8] Chodon, P., Adhikari, D.M., Nepal, G.C., Biswa, R., Gyeltshen, S., et al. (2013) Passive Infrared (PIR) Sensor Based Security System. *International Journal of Electrical, Electronics & Computer Systems*, **14**.
- [9] Lee, M., Guo, R. and Bhalla, A.S. (1998) Pyroelectric Sensors. *Journal of Electroceramics*, **2**, 229-242.
- [10] Wilson, D.H. and Atkeson, C. (2005) Simultaneous Tracking and Activity Recognition (STAR) Using Many Anonymous, Binary Sensors. In: Gellersen, H.W., Want, R. and Schmidt, A., Eds., *Pervasive Computing. Pervasive 2005. Lecture Notes in Computer Science*, Vol. 3468, Springer, Berlin, Heidelberg, 62-79. https://doi.org/10.1007/11428572_5
- [11] Zhu, C. and Sheng, W. (2011) Motion- and Location-Based Online Human Daily Activity Recognition. *Pervasive and Mobile Computing*, **7**, 256-269. <https://doi.org/10.1016/j.pmcj.2010.11.004>
- [12] Hao, Q., Brady, D., Guenther, B., Burchett, J., Shankar, M. and Feller, S. (2006) Human Tracking with Wireless Distributed Pyroelectric Sensors. *IEEE Sensors Journal*, **6**, 1683-1696. <https://doi.org/10.1109/jsen.2006.884562>
- [13] Hao, Q., Hu, F. and Xiao, Y. (2009) Multiple Human Tracking and Identification with Wireless Distributed Pyroelectric Sensor Systems. *IEEE Systems Journal*, **3**, 428-439. <https://doi.org/10.1109/JSYST.2009.2035734>
- [14] Yang, B. and Zhang, M. (2017) Credit-Based Multiple Human Location for Passive Binary Pyroelectric Infrared Sensor Tracking System: Free from Region Partition and Classifier. *IEEE Sensors Journal*, **17**, 37-45.
- [15] Xiong, J., Li, F., Zhao, N. and Jiang, N. (2014) Tracking and Recognition of Multiple Human Targets Moving in a Wireless Pyroelectric Infrared Sensor Network. *Sensors*, **14**, 7209-7228. <https://doi.org/10.3390/s140407209>
- [16] Luo, X., Liu, T., Shen, B., Gao, L., Luo, X., et al. (2016) Human Indoor Localization Based on Ceiling Mounted PIR Sensor Nodes. *13th IEEE Annual Consumer Communications & Networking Conference (CCNC)*, Las Vegas, 9-12 January 2016, 868-874. <https://doi.org/10.1109/CCNC.2016.7444903>
- [17] Liu, T. and Liu, J. (2014) Design and Implementation of a Compressive Infrared Sampling for Motion Acquisition. *EURASIP Journal on Advances in Signal Processing*, **2014**, 20. <https://doi.org/10.1186/1687-6180-2014-20>
- [18] Luo, X., Tan, H., Guan, Q., Liu, T., Zhuo, H.H. and Shen, B. (2016) Abnormal Activity Detection Using Pyroelectric Infrared Sensors. *Sensors*, **16**, 822. <https://doi.org/10.3390/s16060822>
- [19] Guan, Q., Yin, X., Guo, X. and Wang, G. (2016) A Novel Infrared Motion Sensing System for Compressive Classification of Physical Activity. *IEEE Sensors Journal*, **16**, 2251-2259. <https://doi.org/10.1109/JSEN.2016.2514606>
- [20] Brady, D., Pitsianis, N. and Sun, X. (2004) Reference Structure Tomography. *Journal of the Optical Society of America A, Optics, Image Science, and Vision*, **21**, 1140-1147. <https://doi.org/10.1364/JOSAA.21.001140>
- [21] Lu, L., Zhang, H. and Jiang, H. (2002) Content Analysis for Audio Classification and Segmentation. *IEEE Transactions on Speech and Audio Processing*, **10**, 504-516. <https://doi.org/10.1109/TSA.2002.804546>
- [22] Barber, D. (2012) Bayesian Reasoning and Machine Learning. Cambridge University Press, Cambridge.
- [23] Fernandez-Delgado, M., Cernadas, E., Barro, S. and Amorim, D. (2014) Do We Need Hundreds of Classifiers to Solve Real World Classification Problems. *Journal of Machine Learning Research*, **15**, 3133-3181.
- [24] Ooi, S.Y., Tan, S.C. and Cheah, W.P. (2016) Temporal Sampling Forest (TS-F): An Ensemble Temporal Learner. *Soft Computing*, 1-14.
- [25] Breiman, L. (2001) Random Forests. *Machine Learning*, **45**, 5-32. <https://doi.org/10.1023/A:1010933404324>
- [26] Feng, Z., Mo, L. and Li, M. (2015) A Random Forest-Based Ensemble Method for Activity Recognition. *37th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, Milan, 25-29 August 2015, 5074-5077. <https://doi.org/10.1109/EMBC.2015.7319532>
- [27] Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M. and Moore, R. (2013) Real-Time Human Pose Recognition in Parts from Single Depth Images. *Communications of the ACM*, **56**, 116-124.

- <https://doi.org/10.1145/2398356.2398381>
- [28] Alliance, Z. Zigbee Specification. ZigBee Alliance. <http://www.zigbee.org/>
- [29] Brdiczka, O., Reignier, P. and Crowley, J.L. (2007) Detecting Individual Activities from Video in a Smart Home. In: Apolloni, B., Howlett, R.J. and Jain, L., Eds., *Knowledge-Based Intelligent Information and Engineering Systems. KES 2007. Lecture Notes in Computer Science*, Vol. 4692, Springer, Berlin, Heidelberg, 363-370. https://doi.org/10.1007/978-3-540-74819-9_45
- [30] Jalal, A., Kamal, S. and Kim, D. (2014) A Depth Video Sensor-Based Life Logging Human Activity Recognition System for Elderly Care in Smart Indoor Environments. *Sensors*, **14**, 11735-11759. <https://doi.org/10.3390/s140711735>
- [31] Saeb, S., Kording, K. and Mohr, D.C. (2015) Making Activity Recognition Robust against Deceptive Behavior. *PLoS ONE*, **10**, e0144795. <https://doi.org/10.1371/journal.pone.0144795>
- [32] Liao, L., Fox, D. and Kautz, H. (2007) Hierarchical Conditional Random Fields for GPS-Based Activity Recognition. In: Thrun, S., Brooks, R. and Durrant-Whyte, H., Eds., *Robotics Research*, Springer, Berlin, Heidelberg, 487-506. https://doi.org/10.1007/978-3-540-48113-3_41
- [33] Raman, N. and Maybank, S. (2016) Activity Recognition Using a Supervised Non-Parametric Hierarchical HMM. *Neurocomputing*, **199**, 163-177. <https://doi.org/10.1016/j.neucom.2016.03.024>
- [34] 有限状态机. 维基百科[EB/OL]. <https://zh.wikipedia.org/wiki/有限状态机>

期刊投稿者将享受如下服务:

1. 投稿前咨询服务 (QQ、微信、邮箱皆可)
2. 为您匹配最合适的期刊
3. 24 小时以内解答您的所有疑问
4. 友好的在线投稿界面
5. 专业的同行评审
6. 知网检索
7. 全网络覆盖式推广您的研究

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: hjwc@hanspub.org