

密级:_____



中国科学院大学
University of Chinese Academy of Sciences

博士学位论文

自适应行为识别中的迁移学习方法研究

作者姓名: _____ 赵中堂

指导教师: _____ 陈益强 研究员

_____ 中国科学院计算技术研究所

学位类别: _____ 工学博士

学科专业: _____ 计算机应用技术

研究所: _____ 中国科学院计算技术研究所

2013 年 11 月

**Transfer Learning Methods
for Adaptive Activity Recognition**

**By
ZHAO Zhongtang**

**A Dissertation Submitted to
The University of Chinese Academy of Sciences
In partial fulfillment of the requirement
For the degree of
Computer Application Technology**

**Institute of Computing Technology
Nov., 2013**

声 明

我声明本论文是我本人在导师指导下进行的研究工作及取得的研究成果。尽我所知，除了文中特别加以标注和致谢的地方外，本论文中不包含其他人已经发表或撰写过的研究成果。与我一同工作的同志对本研究所做的任何贡献均已在论文中作了明确的说明并表示了谢意。

作者签名：

日期：

论文版权使用授权书

本人授权中国科学院计算技术研究所可以保留并向国家有关部门或机构送交本论文的复印件和电子文档，允许本论文被查阅和借阅，可以将本论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存、汇编本论文。

（保密论文在解密后适用本授权书。）

作者签名：

导师签名：

日期

摘 要

行为识别是普适计算领域中重要的研究内容，具有广阔的应用前景，如“空巢”老人生活远程监护、肥胖病人热量消耗估计、心脑血管等慢性病人的疾病监测、老年人的跌倒检测与报警等。现有的行为识别方法一般分为离线训练和在线识别两个阶段，在离线阶段基于标定的行为数据训练行为识别模型，在在线阶段应用离线阶段训练的模型对未标定的行为数据进行预测。然而，现有的行为识别方法存在一些技术难点，比如离线阶段基于某个加速度传感器建立的模型不能在在线阶段很好地区分来自其他不同型号参数加速度传感器的数据，我们称之为非特定传感器行为识别问题；离线阶段基于某些特定位置的行为数据建立的模型不能在在线阶段很好地区分新位置的行为数据，我们称之为非特定位置行为识别问题；离线阶段基于一人或者多人的行为数据建立的模型不能在在线阶段很好地区分新用户的行为数据，我们称之为非特定人行为识别问题；离线阶段基于某些特定类别行为数据构建的模型在在线阶段不能区分新类别行为数据，我们称之为非特定类别行为识别问题。本文以迁移学习思路为指导，针对不同问题分别提出各自的解决方法，具体包括以下几个方面：

(1) 提出了基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法。针对由于加速度传感器参数不同而导致的源域（离线阶段）和目标域（在线阶段）数据分布不同的问题，该方法通过检测静止状态，计算每种加速度传感器静止时的合成加速度，应用各自的静止合成加速度去规范化各自的加速度数据，达到统一量纲的目的，极大地减少相同行为在时域特征的不同。然后，采用在每秒时间窗口内进行重采样的方法，使得各种传感器的采样频率相同，极大地减小相同行为在频域特征的差别。实验结果表明该方法能够将不同参数加速度传感器的输出值映射到一致的数据空间，从而能够实现模型在内置有异质加速度传感器的便携式移动终端间的迁移，达到构建非特定传感器行为识别模型的目的。

(2) 提出了基于特征迁移的非特定位置行为识别方法。针对不同位置加速度传感器感知到的数据分布不同的问题，该方法首先利用主成分分析技术，寻找不同位置数据的主成分，生成各个位置数据的共有特征，并基于极速学习机算法离线训练一个行为识别模型。然后，在线识别来自不同位置的数据，给每个样本一个预测标号和对应的置信度，并选择高置信度的样本在线增量更新已有模型，当高置信度样本中含有新位置的样本时，更新后的模型就逐渐拥有了对新位置样本的区分能力。实验结果表明本文提出的方法能够在保持模型对已有位置行为数据识别精度的同时提高模型对新位置行为数据的识别能力，达到构建非特定位置行为识别模型的目的。

(3) 提出了基于参数迁移的非特定人行为识别方法。针对使用模型的用户和训练

模型的人员行为数据分布不同的问题，我们首先假设每个人的行为识别模型结构（本文指决策树模型）是相同的，不同之处在于模型的参数的值（本文指决策树非叶子结点的值），由此我们提出在源域和目标域之间迁移决策树及其参数，并针对新用户的数据，对参数的值进行调整，使模型成为一个个性化的行为识别模型。实验结果表明该方法能够在在线阶段自动地利用用户的数据对模型进行更新，达到针对目标用户构建个性化模型的目的。

（4）提出了基于结构迁移的非特定类别行为识别方法。针对目标域中新类行为数据的出现而导致的整体数据分布发生变化的问题，我们基于源域数据训练具有疑似异常行为识别能力的极速学习机，用其对目标域数据进行分类，若某些样本的预测置信度处于异常区间之内，则认为其为疑似异常行为。若当前用户积累了一定数量疑似异常行为数据，系统将其上传到“云端”，通过无监督聚类算法可将它们划分成不同的类别，并邀请相关专家对每个类别提供标号，在此基础上采用类增量极速学习的方法更新模型，使得模型越来越能描述目标域的数据。实验结果表明由于该方法能够保持从源域数据学习得到的结构知识，因此能够在保持模型对已有类别行为识别精度的基础上，逐渐加入对新类别行为的识别能力，能够达到构建非特定类别行为识别模型的目的。

（5）面向实际应用的需求，实现了基于智能移动终端的社区老人健康感知公共服务原型系统。在日常生活中可以识别老人的行为，检测老人的运动量，为促进其合理运动提供帮助；在老人跌倒时，可以及时地发出警报，极大程度地缩短从老人跌倒到得到救助之间的时间，实现对老人健康的监护。

关键字：普适计算，行为识别，迁移学习，增量学习，极速学习

ABSTRACT

Activity recognition is one of the most important research topics in ubiquitous computing, which has been applied in many fields, such as monitoring the life of empty-nest elderly, caring patience who is suffering cardio-cerebrovascular diseases, evaluating the energy expenditure, detecting the fall of elderly, and so on. The existing methods of activity recognition generally include two phases, i.e. offline training phase and online prediction phase. In the offline phase, an activity recognition model is trained based on the labeled activity data. In the online phase, the model trained in the offline phase is used to predict the labels of the unlabeled activity data. However, there are some technical challenges to the existing methods of activity recognition, for examples, the classifier trained on the samples from a specific accelerometer cannot effectively classify the samples from another accelerometer that has different parameters, the classifier trained on the samples collected from some specific locations of the user cannot effectively distinguish the samples from different locations unseen in the training phase, the classifier trained on the samples from some specific people cannot effectively recognize the samples from new users, the classifier trained on some specific types of activities cannot distinguish the samples of new activity and the existing incremental learning methods are time costing or space costing. To address these problems, some methods are proposed under the transfer learning framework as follows.

(1) Still state detection and time-frequency domain transformation based data preprocessing method is proposed to address the problems caused by different parameters of different accelerometers. Through detecting the still state, each accelerometer's synthesized acceleration of still state can be calculated. Then, each synthesized acceleration can be scaled to the same dimension of quantity by dividing the acceleration of still state. In the following step, the synthesized samples in a one second time window are interpolated to the same window size, which can vastly decrease the feature distributions of different accelerometers. The experimental results show that the data from sensors of different parameters can be mapped into the uniform data space. Based on this space, a non-specific activity recognition model is built, which can run on different mobile terminals without modification.

(2) Extreme learning machine based displacement free activity recognition model is presented. The data distribution of accelerometer is highly affected by varying device locations and orientations, which will degrade the performance of recognition model. To solve this problem, we propose a fast, robust and device displacement free activity recognition

model in this paper, which combines principal component analysis(PCA) and extreme learning machine (ELM) method to realize adaptive placement activity recognition. Evaluated on real-world activity data, the experiment results show that this method can increase the ability of recognizing data from new locations, and maintain the ability of recognizing data from known locations as well.

(3)A parameter transferred cross-people activity recognition model is proposed. In activity recognition, the model learned from a specific person often cannot yield accurate results used on a different person. To solve the cross-people activity recognition problem, we suppose that the activity recognition models of each user should have the same structure, but have different parameters. Therefore, the model in source domain(offline phase) can be transferred into target domain(online phase). With the new user's samples, the parameter of transferred model can be adapted to fit the user. Experimental results show that the proposed activity recognition model can be adapted to the user in the online phase, and based on this method a cross-people activity recognition model can be built.

(4)A structure transferred non-specific-class activity recognition model is presented. To address the different distribution caused by the new activities emerging in target domain, a model trained in offline phase is used to predict the samples in target domain in online phase. In the predicted results, some samples' confidence may be among the abnormal interval and can be recognized as suspected abnormal activities. The abundant unlabeled suspected abnormal activity samples of current user can be uploaded to the "cloud" to run a clustering program, then the specialists in the domain of activity recognition can be employed to label each cluster. A new class incremental extreme learning machine is triggered to update the old activity recognition model and the model may fit the data in target domain increasingly. It can be seen from the experiments that the method can not only learn the ability of recognizing new type activities, but also maintain the ability of distinguishing the known type activities.

(5)A mobile phone based health care prototype system for the elderly is designed and implemented. In the daily life, the system can recognize activities of the elderly, calculate amount of their exercise and encourage them to perform moderate exercise. When the elderly is falling down unfortunately, the system can send an alert to their guardians immediately, which can shorten the interval between falling and rescue.

Key Words: Ubiquitous Computing, Activity Recognition, Transfer Learning, Incremental Learning, Extreme Learning

目 录

摘 要.....	I
ABSTRACT	III
目 录.....	V
第一章 绪论	1
1.1 研究背景	1
1.2 研究现状	3
1.2.1 个人日常行为识别及其流程.....	3
1.2.2 行为识别中的传感器设备	5
1.2.3 行为识别方面的研究现状.....	6
1.2.4 迁移学习方面的研究现状.....	8
1.3 行为识别中的四个实际问题	10
1.4 本文的贡献	11
1.5 论文组织	13
第二章 基于加速度传感器的行为数据采集和特征提取	15
2.1 数据采集要求	15
2.1.1 现有的公共数据集.....	15
2.1.2 数据采集设备的多样性.....	17
2.1.3 数据采集要求.....	20
2.2 数据采集过程	21
2.3 数据分割和标定	22
2.4 特征提取	23
2.4.1 时域特征.....	23
2.4.2 频域特征.....	24
2.5 样本集统计信息	26
2.6 小结	30
第三章 面向异质加速度传感器数据的预处理方法.....	31
3.1 问题的提出	31
3.2 相关研究工作	33
3.3 面向异质加速度传感器数据的预处理方法	34
3.3.1 合成加速度的规范化方法.....	35

3.3.2 基于时间窗的采样频率一致化方法	35
3.4 实验及结果分析	36
3.4.1 数据准备	36
3.4.2 量纲对数据分布的影响	37
3.4.3 采样频率对数据分布的影响	39
3.4.4 方法的效果	42
3.5 小结	44
第四章 基于特征迁移的非特定位置行为识别方法	45
4.1 问题的提出	45
4.2 相关研究工作	46
4.3 基于特征迁移的非特定位置行为识别模型	48
4.3.1 合成加速度数据	49
4.3.2 基于 PCA 的特征降维处理	49
4.3.3 极速学习机分类器	50
4.3.4 样本的自标定	53
4.3.5 模型的在线增量更新	53
4.4 实验及结果分析	54
4.4.1 分类器性能比较	55
4.4.2 未进行 PCA 变换时的交叉位置行为识别实验	56
4.4.3 进行 PCA 变换后的交叉位置行为识别实验	57
4.4.4 自适应调整后模型对新旧位置数据的识别能力实验	58
4.5 小结	60
第五章 基于参数迁移的非特定人行为识别方法	61
5.1 问题的提出	61
5.2 相关研究工作	62
5.3 一种基于迁移学习的嵌入决策树算法	63
5.3.1 算法框架	63
5.3.2 二叉决策树的构建及迁移	64
5.3.3 新用户高置信度样本的筛选过程	65
5.3.4 二叉决策树模型的更新过程	66
5.4 实验及结果分析	68
5.4.1 数据准备	68
5.4.2 决策树模型对个人数据的区分能力	68
5.4.3 不同用户两两之间模型的交叉验证实验	69

5.4.4 算法性能.....	70
5.4.5 决策树模型和训练样本数量的关系.....	70
5.4.6 个性化模型和新用户样本的关系.....	71
5.5 小结	71
第六章 基于结构迁移的非特定类别行为识别方法.....	73
6.1 问题的提出	73
6.2 相关研究工作	73
6.3 基于结构迁移的非特定类别行为识别方法	74
6.3.1 算法框架.....	74
6.3.2 具有疑似异常行为识别能力的极速学习机.....	75
6.3.3 新类别样本的发现和标定.....	77
6.3.4 极速学习机模型结构的迁移和更新.....	77
6.4 实验及结果分析	80
6.4.1 数据准备.....	80
6.4.2 模型选择.....	81
6.4.3 识别异常行为的实验.....	82
6.4.4 分类器性能 VS 新样本数量.....	82
6.4.5 行为识别模型稳定性实验.....	83
6.4.6 批量学习与增量学习的性能比较实验.....	84
6.5 小结	85
第七章 基于智能移动终端的社区老人健康感知公共服务系统.....	87
7.1 背景	87
7.2 系统架构	88
7.3 系统的基本功能	89
7.3.1 行为识别.....	89
7.3.2 热量消耗估计.....	90
7.3.3 跌倒报警.....	90
7.4 系统功能的设计与实现	91
7.4.1 行为识别功能的实现.....	91
7.4.2 行为统计分析功能的实现.....	93
7.4.3 热量消耗估计功能的实现.....	95
7.4.4 跌倒检测功能的实现.....	97
7.5 小结	98
第八章 结束语.....	99

8.1 本文工作总结	99
8.2 未来工作展望	100
参考文献	103
致 谢	111
作者简介及在学期间发表的学术论文与研究成果	113

第一章 绪论

1.1 研究背景

个人日常行为识别具有重要的研究意义。在“空巢”老人生活远程监护、肥胖病人热量消耗估计、心脑血管等慢性病人的疾病监测、老年人的跌倒检测与报警等领域有着广阔的应用前景，具体而言：

(1) 个人日常行为识别在老年人远程监护领域有广阔的应用前景。第六次全国人口普查结果显示，全国 60 岁及其以上人口已占 13.26%。根据国际公认标准，60 岁以上人口比例超过 10%，或者 65 岁以上人口超过 7%，就算进入老龄化社会[岳俊峰,2004]。按此标准，我国已经步入了老龄化社会。随着老龄化社会的到来，养老问题日益成为一个社会问题。在各种养老方式中，“居家养老”是人们的首选。但是随着社会人口结构的调整，“空巢”老人的比例日益增大，这些老年人日常起居无人看护，一旦发生意外，后果不堪设想。因此，自动识别老年人日常行为，发现异常行为能及时报警，是实现“居家养老”的重要保障，这对于像我国这样的人口大国意义尤其重大。

(2) 个人日常行为识别在肥胖病人热量消耗估计领域有重要的市场需求。据统计，2007 年美国人中大约有 30%罹患肥胖症。肥胖症导致的高血压、冠心病、糖尿病、痛风、胆石症、脂肪肝等疾病，大约每年消耗了 1000 亿美元。而当年中国人中超重者的比例大约为 28.9%。并非仅有中美有此现象，肥胖已成为困扰世界的难题。大量研究证明，运动是一个能够保持健康的好方法，只要长期和有规律地进行运动，就能使人们保持健康。减肥也是这样，适量运动有助于燃脂，减重，强身健体。然而，目前没有便携的工具能准确地检测用户每天的日常行为，进而计算用户运动量和热量消耗，也就无法科学地衡量用户摄入的能量和消耗的能量之间的关系，从而无法给予用户科学的指导。因此自动识别个人日常行为，寻找行为产生的运动量与热量消耗之间的运动关系，是当前的研究热点，有着重要的市场需求。

(3) 个人日常行为识别在心脑血管等慢性病人的疾病监测方面能够发挥越来越重要的作用。据统计，全球每年因冠心病、脑梗塞等心脑血管疾病死亡的人数约为 1750 万，占全球总死亡的 29%，而这些死亡患者有 80%处于发展中国家。据卫生部统计，我国中老年人高血脂患者有 2 亿人、高血粘患者有 1.5 亿人、高血压患者有 1.5 亿人、冠心病患者有 1 亿人、脑中风患者有 2000 万人、高血糖和糖尿病患者有 9000 万人，且患病人数和发病率呈现逐年增加的趋势。心脑血管相关疾病现已成为我国需要重点关注和亟待解决的公共卫生领域的重大问题，做好该病的预防性工作控制其发生发展、进而提高人们生活质量的的关键。心脑血管疾病具有发病率高、致残率高、死亡率高、复发率高、并发症多等“四高一多”的特点，但目前病人一般是在症状已经较为明显甚至病情已

然严重时才入院进行诊治。而流行病学数据显示,有效改善入院前病情监护水平比住院强化治疗更有意义,同时,治疗越早康复效果越好。因此,必须缩短初始症状出现到确定治疗方法的时间周期,逐步从院内监护扩展到院外长期实时监护,及早诊断,及时救治。基于此,实时准确地监测心脑血管等慢性病人的个人日常行为,发现异常后能够及时地报警,能够有效地保护他们的健康。

(4) 个人日常行为识别能够检测老人的跌倒行为并及时报警[Shi,2012]。老年人由于身体衰弱,平衡能力较差,跌倒已成为威胁他们健康的重要因素。除了平时加强锻炼,防患于未然外,在跌倒事实已经形成后,及时地通知其监护人或者医疗救助机构,也能极大地减轻身体受到的危害。据统计,中国每年有不少于 2000 万老年人不幸跌倒,大约有不少于 2500 万人次的跌倒,由此造成的住院治疗费用高达 50 亿元以上,社会代价大约为 160 到 800 亿元。跌倒已经是导致老年人死亡的首位因素。意外伤害中大约有一半就是老年人跌倒,而跌倒导致的死亡率已经高居第 4 位,排在其前面的且处于第一位的是心脑血管疾病,处于第二位的是肿瘤疾病,处于第三位的是呼吸道疾病。老人意外伤害多在家中发生,占 57.75%,在公共场所发生的占 21.13%,在街道马路发生的占 18.31%,在其他场所发生的占 2.82%。故实时监测老人的个人日常行为,准确地检测出跌倒行为并及时报警,能够极大地缩短从跌倒发生到获得救助之间的时间,为挽救老人生命争取更多的时间。

基于便携设备的非干扰行为识别正成为普适可行的方案。随着微电子机械系统(Micro Electrical Mechanical Systems, MEMS)的发展,微型加速度传感器得以研发并被嵌入到小型移动设备中,这大大促进了基于便携式设备的行为识别系统的研究和部署。已有少数设备能够识别个人日常行为,比如文献[Consolvo,2006]中提到的计步器(Step Counter)以用户走过的步数量化用户的行为,非常适合测量散步和跑步的运动量。一个名为 BeWell 的基于手机的健康监护应用程序[URL1]能够从睡眠、运动、社交等三个方面来评估用户的行为,并对用户当天的每种行为给予一个位于[0,100]之间的分值,用户可根据分值调整自己的行为。Zhang[Zhang,2006]利用内置在手机中的加速度计来获取三维数据,采用 One-class SVM、KFD(Kernel Fisher Discrimination)和 K-NN 方法来检测身体方向的变化。Wolf[Wolf,2007]通过在身体上放置一个三轴的加速度计传感器,来检测身体的方向,进而可以判断当前是否为跌倒等异常行为。文献[Lymberopoulos,2009]提出一种利用传感网络识别独居老人生活模式的方法,利用布置在房间内的传感网络,识别用户睡觉、做饭、看电视、上卫生间的频率和持续时间,进而能识别出其生活规律是否被打破,是否发生了异常。博能公司的产品通过心率来量化用户消耗的卡路里数量[Nachman,2010],非常适合有氧运动项目。文章[Larson,2012]利用移动终端内置的声音传感器来检测用户的肺活量,文章[Lu,2012]还可感知用户承受的心理压力情况。本文设计的基于移动终端的便携式行为识别系统[Zhao,2011],通过分析三轴加速度传感器的数据,能够准确识别多种日常行为,比如静止、散步、跑步、上下楼梯等,并能够统计每

种行为的累计时间；还可以识别出用户的跌倒行为，在用户尤其是老年人跌倒无助时，及时地通过移动通信系统向其监护人和医疗卫生机构发送求救信号，为老年人的安全提供保障。

1.2 研究现状

1.2.1 个人日常行为识别及其流程

近年来，移动电话、掌上电脑等便携式移动设备得到了普遍应用，成为人体感知和计算的新载体[Want,2009]，便携式计算的概念也应运而生。对于便携式计算，目前还没有比较规范、明确与完备的定义。其与穿戴式计算¹有共通之处，都强调“以人为中心”的计算，其又与穿戴式计算有所区别，更强调简单、易用、对人无干扰性。因此便携式计算的概念与 Mark Weiser 提出的普适计算²的概念最为契合，是普适计算领域最为重要的分支之一。

Mann 在文章[Mann,1998]和 Starner 等人在文章[Starner,1998]中首次指出，通过解析人体携带的传感器所获得的数据，能够得到一些复杂的与个人相关的上下文信息，从而构造出新奇的、自适应的、个性化的应用系统。便携式系统能够从用户的角度以第一人称的方式感知用户的行为，比如姿势或者举止，从而引发了基于行为或者交互的应用系统的发生和发展[Davies,2008; Lukowicz,2010a]。这些系统不仅能基于用户所处的情景主动地提供有用的信息或必要的帮助，还支持用户通过自然的手势和身体动作进行交互。

与便携式计算相关的几个应用领域包括但不限于：工业辅助[Stanford,2002]，基于手势输入的人机交互[Kallio,2006]，用于个性化健康监护的行为监控[Pentland,2004]，用于体育训练辅助的运动分析[Chi,2004]等。

现存便携式系统能够识别出非常多的行为或者手势，但是基于不同传感器的系统识别的行为又有所不同。行为按照其复杂程度不同又可分为低级行为和高级行为。低级行为包括运动方式（比如散步、跑步、静止、上楼梯、下楼梯等）、姿势（例如坐姿、站姿等）、手势（例如伸手拿一个物体、开门等）。高级行为是指由一系列动作组成的复杂行为（例如泡茶由拿杯子、放茶叶、冲热水等一系列动作构成）。下面是两个简单介绍日常行为识别系统的例子：

¹对于便携式计算，目前尚无较规范、明确和完备的定义。追本溯源，其应是由加拿大多伦多大学的 Mann 博士于 1996 年首次提出，其本意是强调计算范式方面的转变[Mann,1996]。在 Mann 教授看来，计算机不应是与用户割裂开来的冰冷的工具，而应是用户身体的自然延伸，如影随形随时随地为用户提供更多的感知、反馈和计算能力，以弥补和增强用户自身的能力。我们认为，“穿戴式”计算，就字面看来，它关注的绝非是向用户体内移植各种传感器或者芯片，而是强调，“衣着”将会把各种技术编织并融入自身的纤维，使其从人们的视线中消失，为人们提供无缝的非侵入式的服务。

²普适计算最早起源于 1988 年 Xerox PARC 实验室的一系列研究计划。在该计划中美国施乐(Xerox)公司 PARC 研究中心的 Mark Weiser 首先提出了普适计算的概念。1991 年 Mark Weiser 在《Scientific American》上发表文章“The Computer for the 21st Century”，正式提出了普适计算(ubiquitous computing)。Mark Weiser 指出：“The most profound technologies are those that disappear. They weave themselves into the fabric of everyday life until they are indistinguishable from it.”

1、识别七种动作模式的行为识别系统[Lester,2006]。该系统仅采用一个加速度传感器采集数据，通过建立模型，可以识别的行为包括静坐、站立、散步、上楼梯、下楼梯、乘电梯上楼、乘电梯下楼。

2、识别手势的行为识别系统[Kallio,2006]。该系统也是仅采用一个加速度传感器采集数据，可识别画□（正方形）、×号、○型、∞、曲线等五种手势。

行为识别问题通常被抽象为机器学习问题[Ward,2006; Bao,2004]。用户随身携带设备，在实验室或者真实世界环境中做出各种预先定义好的动作。设备感知用户的动作，记录由动作触发的各种传感器数据。同时通过各种方法记录各种动作对应的语义信息，这种赋予传感器数据语义信息的过程称为标定过程，该过程通常由相关人员现场手工记录完成，也可通过视频或者音频的手段在事后通过回放的方式完成。数据采集和标定工作完成后，基于标定好的数据提取特征、训练行为识别模型并进行性能验证。训练行为识别模型就是寻找传感器感知到的数据和用户的行为之间的映射关系，达到从传感器数据预测用户行为的目的。

如图 1.1 所示，传统的基于传感器的行为识别系统的运作过程分为两个阶段，离线训练模型阶段（Offline Phase）和在线识别阶段（Online Phase）。

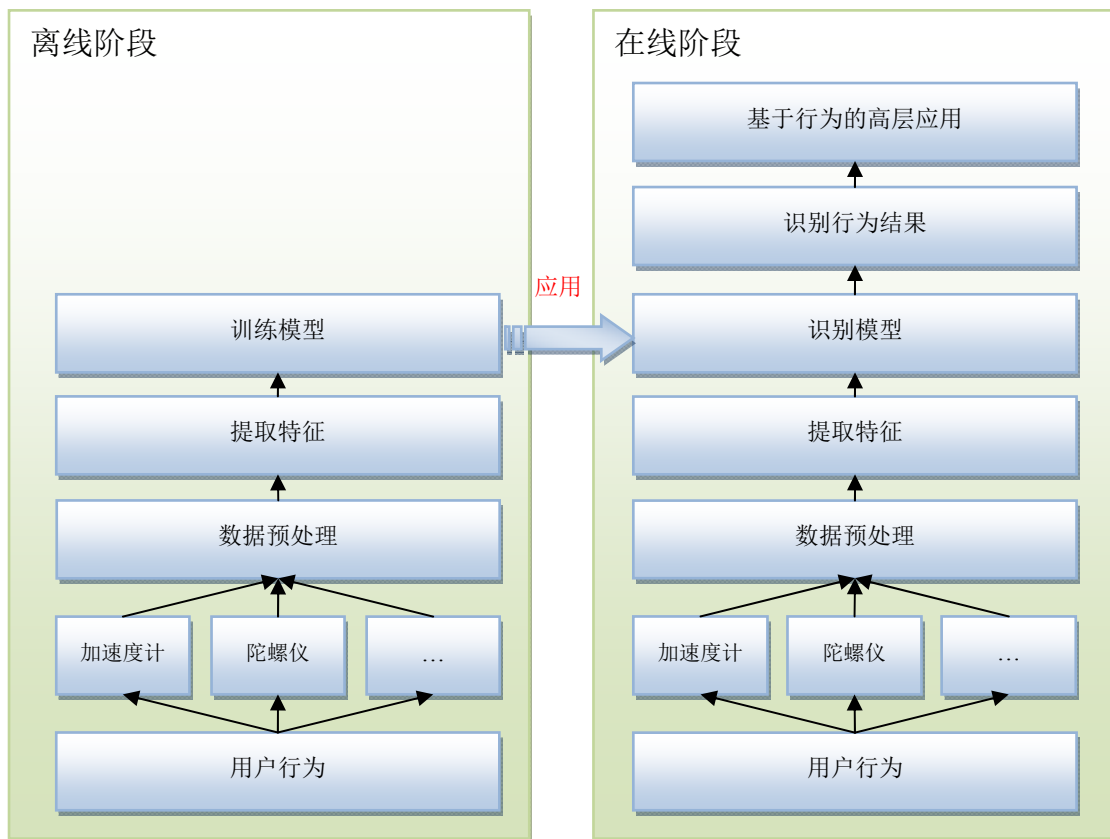


图 1.1 传统行为识别系统的运作过程

离线训练阶段的步骤如下：

- 1、传感器数据采集：获取传感器的原始数据，存储在文件 S 中；
- 2、数据预处理：进行原始传感器数据的预处理，典型的转换方式有校准、去噪、传感器级别数据融合(sensor-level data fusion)；
- 3、数据流的分割：将数据流分割成段，存储在独立的单元 W 中，使得每段尽可能的包含一个行为（动作、手势等）。在数据流中，段可以用他们的起止时间表示。最常用的分割方法是基于滑动窗口的方法；
- 4、特征提取：在分割后的每段上进行特征提取，目的是降低样本的维数和寻找更具代表性的特征，最终形成特征向量 X ；
- 5、建立分类模型：将特征向量映射为预先定义好的类别 $X \rightarrow \langle c, p \rangle$ 。通常情况下，会给出该向量隶属于某类 c 的概率 p ，这些概率会用于决策融合。

在线识别阶段的步骤如下：

- 1、传感器数据采集：获取传感器的原始数据，存储在向量 S 中；
- 2、数据预处理：进行原始传感器数据的预处理，典型的转换方式有校准、去噪、传感器级别数据融合；
- 3、数据流的分割：将数据流分割成段，存储在独立的单元 W 中，使得每段尽可能的包含一个行为（动作、手势等）。在数据流中，段可以用他们的起止时间表示。最常用的分割方法是基于滑动窗口的方法；
- 4、特征提取：在识别出的每段上进行特征提取，目的是降低样本的维数和寻找更具代表性的特征，最终形成特征向量 X ；
- 5、行为分类：用离线阶段训练的分类器对获得的样本分类，识别出对应的行为类别；
- 6、高层应用：基于识别出的动作，构建高层应用程序，例如老人生活远程监护、肥胖病人热量消耗估计、心脑血管病人的运动量估计、老年人跌倒检测等。

1.2.2 行为识别中的传感器设备

欲识别用户的行为，须先感知用户的行为；欲感知用户的行为，须有感知用户行为的工具。MEMS 技术的发展为感知现实世界提供了各种各样的传感器，且这方面的技术还在不断地发展、进步。对于便携式计算系统来说，其能否大行其道，首先需要考虑的是用户的体验，必须让用户感觉到舒适，而不是感觉不便。因此，传感器应该体积小巧、携带方便、不易发现。选择传感器的时候，还必须考虑到各方面性能之间的均衡，比如便携性、计算能力、功耗情况、通信要求、要获取的信息内容等。例如，摄像机在便携式计算中很少被使用，因为视频分析需要耗费大量的计算能力，反而是那些需要轻量级计算的传感设备更受欢迎。

常用的传感器是加速度传感器和惯性测量单元(Inertial Measurement Units, IMUs)。

加速度传感器的规格较小、功率较低。IMUs 由多种传感器构成,包括加速度传感器、磁力计、陀螺仪等,因此其不仅可以感知携带者的加速度情况,还能感知携带者相对于某个参考点的方向。通常会把 IMUs 配置在用户身体的不同部位,用以重建用户的身体运动模型。再者,由于不同的行为会产生不同的响声(例如使用咖啡机的声音和刷牙的声音)[Ward,2006],所以麦克风有时也是进行行为识别的一种有效手段。文献[Roggen,2010]中列出了一些常用的传感设备。

衣物是无侵扰配置传感设备的最好载体。例如,可以将 IMUs 集成到工人的夹克中。有些研究机构正在尝试研制可感知的纺织纤维,一旦成功,即可实现真正的无侵扰的服装感知系统[Tognetti,2006]。

为提高行为识别系统的性能,流行的做法是使用多个多模传感设备[Stiefmeier,2008],也可充分利用周围环境的传感设备来弥补自身的不足。近来,欧盟发起资助了一个名为 OPPORTUNITY 的项目[Roggen,2009],提出了机会行为识别的概念(Opportunistic Activity Recognition)。该项目尝试构建一个无需专门布置传感器、仅利用环境中偶然可得传感设备进行行为识别的系统,这将进一步解决便携性问题。同时该项目也对机器学习算法提出了新的需求,它要求必须能在不同的传感器领域共享行为识别模型。

1.2.3 行为识别方面的研究现状

(1) 基于视频监控的行为识别

基于二维视频的动作行为识别方面已有大量工作存在[徐光佑,2009]。Liu 等人[Liu,2004]使用了与人体朝向呈 90° 的固定视角,基于距离投影用高斯分布表示每一个动作,用以分析坐、站、躺等几个日常生活中最基本的动作。Park 等人[Park,2005]用高斯混合模型来表示人体的五个主要部分(头、脸、躯干、胳膊和下身)的颜色分布,并用 Kalman 滤波器和椭圆拟合实时地更新参数,最后训练动态贝叶斯网络来检测人的动作和姿态,并从中总结出表达式来表示动作行为。Chung 等用竖直方向和水平方向上的相关的一对投影表示用户当前的体态[Chung,2008],Robertson 等采用基于光流的描述子描述动作[Robertson,2005],然后逐个与样本集中的样本进行匹配以识别动作的类型。Turaga 等也是用提取光流的方法为每一帧中的动作提取特征[Turaga,2007]。Wang 等人在识别办公室中异常行为的研究中[Wang,2007]先分割出人体区域,然后对相应区域采用 R 变换[Kim,2008],提取出动作的形状,R 变换具有旋转不变性和尺寸不变性,可以应对用户距离摄像机的远近不同而产生的尺寸变化。以上所述工作都是将摄像头的视角固定后提取二维运动特征来表示动作,这时虽然有利于观测用户的动作,但也限制了用户的活动范围,从而难以适应复杂多变的实际应用环境。现实中,观测对象的活动范围较广,不同时刻所处的位置变化大,不可避免地导致视角的多变性;同时用户的生活环境不会是空旷的,家具等物品也会遮挡人体。为解决这些问题,就需要建立支持分布式计

算的视觉系统,通过融合多摄像机的信息来克服由于活动范围广、视角多变以及遮挡等因素带来的问题。

Campbell 等人基于立体视觉理论构建了一个识别三维手势的系统[Campbell,1996]。Jin 等人建立了基于三维模型的动作识别系统[Jin,2007]。然而,三维模型通常具有参数较多、训练较复杂、计算量较大的缺点。

总之,基于视频的行为识别无需用户携带任何设备,因此对用户没有侵入性,应用前景好;但视频设备一般安装在固定位置,只对固定视角的对象具有很好的识别性,再加上光线、遮挡等因素,会造成识别困难,而且并非任意地方都有视频设备,会造成大量的监控“死角”。

(2) 基于多传感器穿戴式设备的行为识别

在穿戴式行为识别中,可穿戴在用户身体上的传感器包括声音传感器、加速度传感器、温度传感器等。考虑到加速度传感器具有能检测所附着物体运动速度变化的特点,故在行为识别中得到了最为广泛得运用。文献[Bao,2004]在人体不同部位配置 5 个两轴加速度传感器,采用决策表、基于实例的学习、决策树、朴素贝叶斯算法等分类算法,识别 20 个不同动作,平均识别准确率为 77.31%。在文献[Mantyla,2001; Ravi,2005; Ward,2011]中,他们利用用户穿戴在身上的多个加速度传感器识别站立、刷牙、行走等日常行为;在文献[Lukowicz,2004b]中,作者利用穿戴式的声音传感器和加速度传感器来监控木材加工厂中工作人员的加工流程;在文献[Ohgi,2000; Bachlin,2009]中,作者构建了自由泳辅助训练系统,他们让运动员在手腕部位佩戴内置有加速度传感器的设备,基于采集到的数据分析运动员在训练过程中的动作特点。在文献[Albinali,2010; Lester,2009; Nachman,2010]中,作者利用穿戴在用户身上的多个传感器识别人体动作,从而估计用户的能量消耗。此外基于穿戴式行为识别技术, Bourke 和 King 在检测老年人跌倒 [Bourke,2008;King,2010]、Aziz 和 McIlwraith 在帮助人们进行体育锻炼 [Aziz,2009;McIlwraith,2009]、Bachlin 在指导运动员训练[Bachlin,2009]等方面都有较深入的研究。

总之,穿戴式设备的出现可以弥补视频监控具有“死角”的不足,因此可作为视频监控功能的扩充。随着新型生理监测传感器的出现,可以提供视频监控不能提供的功能。然而,目前的多传感器穿戴式传感设备由于价格昂贵穿戴位置较多,而导致用户体验不佳,不方便实用,仅仅应用于特殊的领域,还未在人民大众中普及。

(3) 基于单个移动终端的行为识别

CardioTrainer 是一个基于移动设备的应用[URL2],能够成为用户的健身伙伴,监督其在跑步、骑车、远足、滑雪等任意户外运动方面的能量消耗情况。睡眠周期时钟应用[URL3]基于生物学知识,能够分析用户的睡眠模式,当其处于浅睡眠状态时将其唤醒。Nokia 的计步器应用[URL4]能够计算用户步行时走过的步数和路程。Nokia 公司的 SportsTracker[URL5]基于 GPS 信号能够记录用户的运动时间、路程及消耗的热量。然而,

这些移动应用程序要求用户在进行某一项运动前必须手动地设置运动类型。在文献[Lester,2009]中,提出了一种基于单个加速度传感器的系统,该系统能够自动识别用户的运动行为,并计算消耗的热量。Chen[Chen,2010]提出了用手机的加速度传感器检测用户的6种动作从而检测用户的运动状态变化点,以此决定是否开启和关闭WiFi扫描设备,达到节能的目的。文献[Ravi,2005]使用一个三轴加速度传感器识别8个不同动作,对比了基于生成模型的方法、基于区分模型的方法和基于元模型的方法,平均识别准确率为92.45%。文献[Yang,2009]利用手机内置的三轴加速度传感器能够识别静止、走路、骑车等6种行为,并自动地生成运动日志,他们采用决策树算法,平均识别准确率为90.6%。Altun等人[Altun,2010]基于人们的行为数据,对各种分类算法做了详细的比较研究,包括贝叶斯决策,决策树,最小二乘法,KNN(K Nearest Neighbor, K近邻)算法,DTW(Dynamic Time Wrap, 动态时间弯折)、SVM(Support Vector Machine, 支持向量机)和人工神经网络等。Jamie等人在文献[Jamie,2011]中提出了用于衡量行为识别性能的准则,为衡量行为识别算法的性能奠定了很好的基础。Sun等人[Sun,2010]使用内置有三轴加速度传感器的手机,将其分别放置在腰部周围的多个口袋,采集了多个人的行为数据,他们先对加速度进行合成操作,然后提取多个特征,采用SVM建立行为识别模型,取得了很好的识别效果。此外国内的清华大学、浙江大学、南京大学、中国科学技术大学、上海交通大学、北京邮电大学、华南理工大学、大连理工大学、西北工业大学、重庆大学、电子科技大学、中科院软件所、中科院自动化所等单位利用行为识别技术在关爱人们身体健康[苗强,2009;徐经纬,2012]、检测老年人跌倒[石欣,2010]、手势识别[Chen,2011;薛洋,2011;PanGang,2010;Wu,2009]、步态识别[Meng,2011]、基于生理计算的人机交互[王宏安,2011]等领域获得了广泛应用。我们关注老年人健康看护问题,在行为识别[Zhao,2011;Chen,2012]、跌倒检测[Zhao,2010]等方面做了部分研究工作,有了一定的工作积累。

总之,目前移动终端设备,尤其是智能手机,嵌入了大量的传感器设备,能够检测用户的多项健康监控数据。且手机大量普及,已成为多数人日常生活不可或缺的一部分,因此基于手机的健康监控应用已成为当前研究的一个热点。

然而,在实际研究过程中,我们发现基于移动终端的行为识别还面临如何构建非特定传感器、非特定位置、非特定人和非特定类别等行为识别模型的问题,这些问题由于不满足传统机器学习对数据的要求而需要寻求迁移学习方法的支持。

1.2.4 迁移学习方面的研究现状

迁移学习的概念在1995年被提出,并在学术界获得广泛关注。早期对迁移学习的定义过于宽泛,曾出现过许多专用术语,如Life-Long Learning, Learning to Learn, Multi-task Learning, Meta Learning, Knowledge Transfer等。

2005年,美国国防部高级研究机构DARPA给予迁移学习一个新的定义:迁移学

习系统能够对先前任务中学到的知识和技巧进行总结，并将它应用到一个新的任务中去。传统机器学习模型的训练和测试都在同一个域 (Domain) 中进行，而在迁移学习中，进行模型构建的领域被称为源域，模型被应用的领域被称为目标域。同年，NIPS 给出了一个比较具有代表性的定义：Transfer learning emphasizes the transfer of knowledge across domains, tasks, and distributions that are similar but not the same。中文解释为：迁移学习强调在相似但存在差别的不同领域、任务和分布之间进行知识迁移[Pan,2010]。

目前，在几个顶级会议中已经出现了关于迁移学习的文章，最著名的，比如 Data Mining 领域的 ACM KDD, IEEE ICDM, PKDD 等等；Machine Learning 领域的 ICML、NIPS、ECML、AAAI、IJCAI 等等；Machine Learning 和 Data Mining 方面的应用领域 ACM SIGIR、WWW、ACL 等等。

香港科技大学研究机器学习的国际著名专家杨强教授和他的学生合作的论文“A Survey on Transfer Learning” [Pan,2010]中，对迁移学习进行了系统的总结。他们按照数据域和目标域数据标定情况的不同把迁移学习分为三类：归纳迁移学习 (Inductive Transfer Learning)、直推式迁移学习 (Transductive Transfer Learning) 和无监督迁移学习 (Unsupervised Transfer Learning)。根据迁移的内容不同，又可把迁移学习分为四类：实例迁移、特征迁移、参数迁移和相关知识迁移。

下面我们形式化地描述一下迁移学习领域中涉及的概念。

定义 1 (域, Domain)。一个域 (用 D 表示) 包含两方面的内容：特征空间 \mathcal{X} 和边缘概率分布 $P(X)$ ，其中 $X = \{x_1, x_2, \dots, x_n\} \in \mathcal{X}$ ，表示一个具体的样本。

定义 2 (任务, Task)。给定一个特定的域 $D = \{\mathcal{X}, P(X)\}$ ，一个任务包含两方面的内容：一个标号空间 \mathcal{Y} 和一个目标预测函数 $f(\cdot)$ ，任务用 $T = \{\mathcal{Y}, f(\cdot)\}$ 表示。目标预测函数可以从标定好的训练集中学习得到，训练集中包含多个形如 (x_i, y_i) 的数据，其中 $x_i \in \mathcal{X}, y_i \in \mathcal{Y}$ 。

定义 3 (源域, 目标域)。我们用符号 $D_S = \{(x_S^i, y_S^i) | x_S^i \in \mathcal{X}_S, y_S^i \in \mathcal{Y}_S\}_{i=1}^{N_S}$ 表示源域，用符号 $D_T = \{(x_T^i, y_T^i) | x_T^i \in \mathcal{X}_T, y_T^i \in \mathcal{Y}_T\}_{i=1}^{N_T}$ 表示目标域。

定义 4 (迁移学习, Transfer Learning)。给定源域 D_S 及其学习任务 T_S ，目标域 D_T 及其学习任务 T_T ，迁移学习的目的是在 $D_S \neq D_T$ 或者 $T_S \neq T_T$ 的条件下运用源域 D_S 及其任务 T_S 帮助提高目标域目标预测函数 $f_T(\cdot)$ 的学习。

定义 5 (归纳迁移学习, Inductive Transfer Learning)。给定源域 D_S 及其学习任务 T_S ，目标域 D_T 及其学习任务 T_T ，归纳迁移学习的目的是在 $T_S \neq T_T$ 的条件下运用源域 D_S 及其任务 T_S 帮助提高目标域目标预测函数 $f_T(\cdot)$ 的学习。

定义 6 (直推式迁移学习, Transductive Transfer Learning)。给定源域 D_S 及其学习任务 T_S ，目标域 D_T 及其学习任务 T_T ，直推式迁移学习的目的是在 $D_S \neq D_T$ 且 $T_S = T_T$ 的条件下运用源域 D_S 及其任务 T_S 帮助提高目标域目标预测函数 $f_T(\cdot)$ 的学习。

定义 7 (无监督迁移学习, Unsupervised Transfer Learning)。给定源域 D_S 及其学

习任务 T_S ，目标域 D_T 及其学习任务 T_T ，无监督迁移学习的目的是在 $T_S \neq T_T$ 且源域和目标域均没有标定数据的条件下运用源域 D_S 及其任务 T_S 帮助提高目标域目标预测函数 $f_T(\cdot)$ 的学习。

目前的文献中，只有在归纳迁移学习和直推式迁移学习中存在实例迁移的情况，在无监督的迁移学习中由于源域和目标域均不存在已标定样本，故不存在实例迁移的情况。在归纳迁移学习中，通过度量源域或者目标域的未标定样本与目标域标定样本之间的相似度，为未标定样本分配类标号以及权重，相似度越大对训练目标模型越有利，相应的权重就应被加大，否则权重就应被削弱，然后利用新标定的样本辅助目标预测函数的学习。在直推式迁移学习中，根据目标域和源域实例分布的不同，计算源域每个标定样本隶属于目标域的权重，然后根据加权的机器学习方法将源域中的标定样本迁移到目标域中，从而能够构建目标域的目标预测函数。

在归纳迁移学习、直推式迁移学习和无监督迁移学习中都有特征迁移的成功案例。基于特征的迁移学习方法是把源域和目标域的数据从原始的高维特征空间映射到低维特征空间，使得在低维空间中源域数据与目标域数据具有相同的数据分布，这样就可以在低维空间利用已标定的源域样本训练分类器，对目标域测试数据进行预测。

参数迁移学习方法假设在源域和目标域的模型之间共享一些参数或者先验知识，这样只要在其中某一个域能够计算出该参数或者先验知识，其他域的模型就可以相应求解。目前只有在归纳迁移学习中有相关的文献出现。

和其他三种迁移学习方法不同，相关知识迁移针对的是具有相关关系的源域和目标域，这两个域的数据不是独立同分布，它们之间存在多种关系，比如网络数据、社交网络数据同属此类。解决此类问题通常采用统计相关学习技术（Statistical Relational Learning Techniques）。

1.3 行为识别中的四个实际问题

现有行为识别技术尚存实用问题。在实际的开发应用过程中，我们发现由于获取到的行为数据会因为不同的传感器、不同的位置、不同的人 and 不同的行为而不同，从而影响行为识别模型的识别精度，由此我们发现基于便携式移动设备的行为识别存在四个实际问题。

第一，离线阶段基于某个传感器建立的模型在在线阶段不能很好地区分来自其他不同参数传感器的数据。与行为识别相关的传感器的参数主要有量纲和采样频率，它们影响着样本的时域和频域特征，导致基于某个特定传感器行为数据训练的识别模型不能直接应用于其他传感器数据。那么，能否找到一个非特定传感器的行为识别模型，不需要用户的参与就能自适应地识别来自不同参数传感器的数据？我们称之为**非特定传感器行为识别问题**。

第二，离线阶段基于某些特定位置的行为数据建立的模型不能很好地区分新位置的

行为。用户佩戴移动设备的位置变化不定，有时拿在手中，有时放在口袋里，有时放在包中等等，不一而足。采集的训练数据只能涵盖有限的位置，而不能涵盖所有。这样，模型就不能很好地区分新位置的行为数据。那么，能否找到一个非特定位置的行为识别模型，不需要用户的参与就能自适应地自我调整逐渐具有识别所有位置数据的能力？我们称之为**非特定位置行为识别问题**。

第三，离线阶段基于一人或者多人的行为数据建立的行为识别模型在在线阶段不能很好地区分新用户的行为。在基于机器学习的行为识别领域，通常会得出如下结论：当用户的行为数据在训练集中出现时，模型对该用户行为就具有较高的区分能力；反之，当用户的行为数据不在训练集中出现时，模型就不能很好地区分该用户的行为。在现实世界中，不可能采集到足够多具有不同身体条件的人们行为数据，于是基于该不完备训练集合得到的模型对新用户的行为数据的区分能力就有可能不会太高，那么，能否找到一个非特定人的行为模型，当用户使用该模型时，其能够自适应地满足用户的需求？我们称之为**非特定人行为识别问题**。

第四，离线阶段基于特定类别的行为数据建立的行为识别模型在在线阶段不能识别未在训练数据集中出现的新行为。对用户的行为建模时，往往是对用户当前的行为进行建模，能够反映其当前行为模式。随着时间的推移，用户的行为特点会发生变化，必须用新的行为数据更新模型，模型才能与时俱进。更重要的是，在新时期，用户有可能更改了生活习惯，比如对新的运动形式发生了兴趣，并经常从事这种运动。由于该种行为没有出现在已有模型的训练数据中，模型就不具有识别新类行为的能力。那么，能否找到一个快速的非特定类别的行为模型，当用户使用模型时，能够随着新类行为数据的加入而自适应地进行增量学习，使之具有识别新类行为的能力？我们称之为**非特定类别行为识别问题**。

1.4 本文的贡献

本文采用迁移学习的思想对行为识别中存在的问题进行研究。传统的机器学习方法假设所要解决的问题需要满足两个条件：1) 用于训练模型的已标定训练样本和用于测试模型的未标定测试样本必须满足独立同分布的条件；2) 必须有大量的足够多的训练样本才能得到好的模型。在研究基于移动终端的行为识别问题过程中，我们发现这两个条件很难得到有效的保证：1) 在行为识别模型的应用过程中，由于条件的多变性，很难保证测试样本和训练样本是独立同分布的；2) 一般情况下，在应用行为识别模型的过程中，测试样本是没有任何标签的。根据行为识别流程，我们自下而上地分析了系统应用过程中存在的问题，分别从由传感器参数不同、位置不同、用户不同和样本类别标签不同而导致的数据分布不同的角度出发，提出了相应的解决方案。首先针对传感器参数不同导致的识别准确率下降问题，我们提出了基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法，能够将各个参数不同的传感器的数据映射到相同的数据

空间。第二，针对由于使用过程中移动终端放置位置不同而导致的数据分布不同问题，我们提出了基于特征迁移的非特定位置行为识别方法，能够较好地提高系统的识别准确率。第三，针对使用模型的用户和训练模型的人员的行为数据分布不同的问题，我们提出了基于参数迁移的非特定人自适应行为识别方法，该方法使得模型能够随着用户使用时间的增加，不断地自我调整，越来越能反映用户自身行为的特点。第四，对于如何有效添加对新类行为数据的区分能力，我们提出了基于增量学习的行为识别方法，用少量标定数据训练分类模型，要求用户以较少代价参与标定分类模型不确定的数据，并用新标定的数据在线更新分类模型。

(1) 提出了基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法

针对由于加速度传感器参数不同而导致的源域（某个加速度传感器的数据）和目标域（另一个参数不同的加速度传感器的数据）数据分布不同的问题，该方法通过检测传感器的静止状态，计算每种加速度传感器静止时的合成加速度，用各自的静止合成加速度去规范化各自的加速度数据，达到统一量纲的目的，极大地减少相同行为在时域特征的差别。然后，采用在每秒时间窗口内进行重采样的方法，使得各种传感器的采样频率相同，极大地减小相同行为在频域特征的差别。实验结果表明该方法能够将异质加速度传感器的输出值映射到统一的数据空间，从而能够实现模型在内置有异质加速度传感器的便携式移动终端间的迁移，达到构建非特定传感器行为识别模型的目的。

(2) 提出了基于特征迁移的非特定位置行为识别方法

针对不同位置加速度传感器感知到的数据分布不同的问题，该方法首先利用主成分分析技术，寻找不同位置数据的主成分，生成各个位置数据的共有特征，并离线训练一个行为识别模型。然后，在线识别来自不同位置的数据，给每个样本一个预测标号及相应的置信度，并选择高置信度的样本在线增量更新已有模型，使之既能保持对已有位置的识别精度，又能提高对新位置数据的识别能力。实验结果表明本文提出的方法能够在保持模型对已有位置数据识别精度的同时，又能提高模型对新位置数据的识别能力，达到构建非特定位置行为识别模型的目的。

(3) 提出了基于参数迁移的非特定人行为识别方法

为解决使用模型的用户和训练模型的人员行为数据分布不同的问题，我们提出了基于参数迁移的非特定人行为识别方法。我们假设每个人的行为识别模型结构（本文指决策树模型）是相同的，不同之处在于模型的参数的值（本文指决策树非叶子结点的值），因此我们在源域和目标域之间迁移决策树及其参数，针对新用户的数据，对参数的值进行调整，使模型成为一个个性化的行为识别模型。实验结果表明该方法能够在在线阶段利用使用者的数据自动地对模型进行更新，达到针对目标用户构建个性化模型的目的。

(4) 提出了基于结构迁移的非特定类别行为识别方法

针对目标域中新类行为数据的出现而导致的整体数据分布发生变化的问题，我们基于源域数据训练具有疑似异常行为识别能力的极速学习机，用其对目标域数据进行分

类，若某些样本的预测置信度处于异常区间之内，则认为其为疑似异常行为。若当前用户积累了一定数量疑似异常行为数据，系统将其上传到“云端”，通过无监督聚类算法可将它们划分成不同的类别，并邀请相关专家对每个类别提供标号，在此基础上采用类增量极速学习的方法更新模型，使得模型越来越能描述目标域的数据。实验结果表明由于该方法能够保持从源域数据学习得到的结构知识，因此能够在保持模型对已有类别行为识别精度的基础上，逐渐加入对新类别行为的识别能力，达到构建非特定类别行为识别模型的目的。

1.5 论文组织

本文由八章组成，各章节组织如下：

第一章介绍了论文的背景，提出了本文所研究的问题及问题的研究现状。

在第二章，介绍了基于加速度传感器的行为数据的采集和特征提取，得到的样本集是后续各章节实验的基础。

第三章提出了基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法，能够将不同参数加速度传感器的数据映射到统一的数据空间。

在第四章中，针对基于某些特定位置样本的行为识别模型不能很好地识别新位置样本的问题，提出了基于特征迁移的位置无关识别方法。

第五章，针对基于某些用户样本的行为识别模型不能很好地识别新用户样本的问题，提出了基于参数迁移的非特定人行为识别方法。

在第六章，为解决目标域中新类数据的出现而导致的数据分布不同问题，我们提出了基于结构迁移的非特定类别行为识别方法。

第七章，基于本文的部分研究成果，我们构建了一个基于智能移动终端的社区老人健康感知公共服务原型系统。

第八章，对本文做出的工作进行回顾和总结，并且指出了一些未来可能的研究方向。

第二章 基于加速度传感器的行为数据采集和特征提取

机器学习研究的前提条件之一是需要有大量的训练数据，因此基于机器学习方法研究行为识别模型，需要事先获取各种行为数据并正确标定。本章首先调研了行为识别领域已有的公共数据集，分析其数据特点，然后介绍我们针对本文所研究问题而构建的行为数据集，分别介绍数据采集方案、数据采集过程、数据分割与标定和行为数据特征提取等，最后以表格形式展示了行为数据集的统计信息。

2.1 数据采集要求

2.1.1 现有的公共数据集

在行为识别领域，根据研究问题的深度和广度不同，麻省理工学院、瑞士联邦工学院等研究机构分别构建了自己的行为数据集，并积极地与业界分享。

如图 2.1 所示，麻省理工学院在文献 [Bao,2004] 中提到他们设计了一个基于多个加速度传感器的行为识别系统，并发布了相关的数据集³。系统中共有 5 个独立的加速度传感器，分别固定配置在测试者的右侧手腕、右侧腰部、右侧脚踝、左侧手臂和左侧大腿五个部位，识别目标为 20 种日常运动模式。虽然融合多个加速度传感器的系统能够在一定程度上提高行为识别的准确率，但是在身体的多个位置配置传感器的要求，会对佩戴者造成侵扰，使他们感到不便。此外，随着加速度传感器数量的增加，系统的构建成本、数据的传输和计算成本也会增加。因此，目前来看，多加速度传感器系统的实用性较差，大多处于实验室的研究阶段。



图 2.1 麻省理工学院行为识别系统中传感器放置位置

³ 麻省理工学院加速度传感器数据集: <http://web.media.mit.edu/~nawyn/downloads/LingAcceleration.html>



图 2.2 瑞士联邦工学院穿戴式计算实验室的数据集

如图 2.2 所示，瑞士联邦工学院穿戴式计算实验室（Wearable Computing Lab）也发布了一系列基于加速度传感器的行为数据集⁴。其中，Skoda mini checkpoint 数据集是在汽车维护场景下采集得到的，共采集了一个测试者的数据，该测试者在左右胳膊上各绑定 10 个加速度传感器，分别采集了 10 种动作行为。BodyAttack fitness dataset 数据集包含了一个用户的运动数据，该用户在左腿上绑定 10 个加速度传感器，做出 6 种动作，这些动作大多和腿部运动相关。HCI gestures dataset 数据集共包含 5 个动作的数据，分别是画上三角、下三角、正方形、圆形、无穷大符号，这些动作由一个测试者完成，该测试者在右臂上绑定了 10 个加速度传感器。

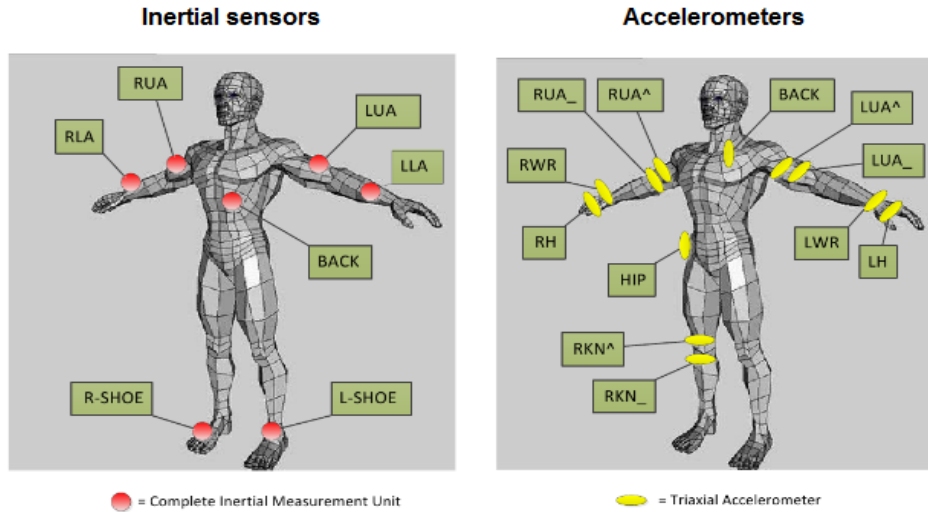


图 2.3 Opportunity 行为识别公共数据集对应的数据采集设备及其穿戴位置

瑞士联邦工学院穿戴式计算实验室在机会行为识别[Roggen,2010; Lukowicz,2010a]方面（Opportunity 项目）也做出了探索。Opportunity 项目获得了欧盟的资助，在项目的推进过程中，他们取得了大量研究成果。他们基于采集到的数据集的子集发起了一个世界范围的行为识别竞赛（Activity Recognition Challenge），让参与者在数据集⁵上进行研究，并比较各自的研究成果。

⁴ ETH Wearable Computing Lab 行为识别数据集: <http://www.wearable.ethz.ch/resources/Dataset>

⁵ Opportunity Activity Recognition Challenge 数据集: <http://www.opportunity-project.eu/challenge>

如图 2.3 所示,竞赛数据集中涉及到的传感器有:一套含有 5 个惯性传感单元(Inertial Sensors)的夹克,两个固定在左右鞋子上的传感单元,12 个固定在四肢和腰部的加速度传感器。

数据集中共包含 4 个人的数据,每个人完成 5 组不同的动作,第一组数据是在一个锯木厂按照一定的顺序做一组动作,其他四组数据是测试者日常的行为。

文献[Ravi,2005]使用一个放置在臀部的三轴加速度传感器采集了 2 个人 8 种不同动作(站立、散步、跑步、上楼梯、下楼梯、起立、用吸尘器打扫卫生、刷牙)。Yang 在文献[Yang,2009]中提到他们基于 Nokia 手机的系统,采集了 4 人的站立、静坐、散步、跑步、驾车、骑车等 6 种行为,且他们对用户如何携带设备没做硬性规定。Sun 等人[Sun,2010]使用内置有三轴加速度传感器的手机,将其分别放置在腰部周围的多个口袋,采集了 7 个人的 7 种行为数据(静止、散步、跑步、骑自行车、驾车、上楼梯和下楼梯)。然而,他们都没有公开自己的数据集。

由于我们研究基于移动终端的行为识别系统,关注如何构建非特定加速度传感器、非特定位置、非特定人和非特定类别的行为识别模型。因此采集的数据应该涵盖不同的加速度传感器、不同的位置、不同的人 and 不同的行为类别等方面的特性。而上述数据集,有的是基于固定在身体上的传感器采集,其使用场景与使用移动终端的场景有显著差别,不能满足我们的需要;有的虽然是基于移动终端构建,但是没有公开发布,无法获得使用权。基于此,我们有必要构建适合我们研究问题的数据集。

2.1.2 数据采集设备的多样性

目前,智能移动终端尤其是智能手机已成为大家日常生活中重要的一部分,手机与人之间有着如影随形的关系,最能反映人们的日常行为,因此我们研究基于智能手机的行为识别系统。

2.1.2.1 智能手机及其操作系统的多样性

据艾媒咨询⁶统计,2012 年第一季度中国的智能手机市场中不同品牌的占有率状况如图 2.4 所示。

⁶ 艾媒咨询: <http://www.iimedia.cn/29575.html>



图 2.4 中国智能手机市场品牌占有率统计

从图 2.4 中可以看出，诺基亚、三星、苹果的市场占有率之和为 59.3%，处于市场的主导地位。但三家却采用不同的手机操作系统，据艾媒咨询统计，2012 年第一季度中国智能手机市场操作系统占有率状况如图 2.5 所示。

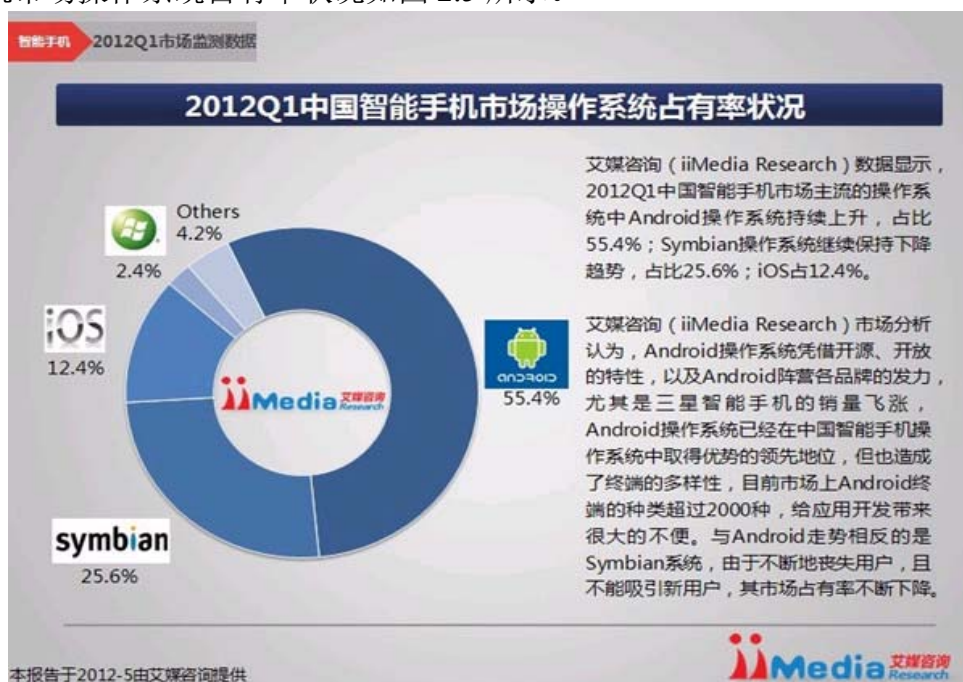


图 2.5 中国智能手机市场操作系统占有率状况

从图 2.5 可以看出，中国智能手机市场主流的操作系统中，Android 操作系统的市场占有率持续上升，占 55.4%；Symbian 系统延续下降趋势，占比 25.6%；iOS 系统占 12.4%，合计占有率为 93.4%。

综合图 2.4 和图 2.5 可以看出，三星和其他智能手机销量飞涨，带动 Android 操作

系统在中国智能手机操作系统中取得优势的领先地位。诺基亚智能手机虽然销量下滑，但依然有一定的市场份额，使得 Symbian 操作系统有 25.6% 的市场率。苹果智能手机带来的非凡体验也为其操作系统 iOS 带来 12.4% 的市场占有率。总而言之，当前虽然手机品牌纷繁芜杂，但对应的操作系统相对集中在 Android、Symbian 和 iOS。因此研究这些操作系统下加速度传感器软件接口的特性，进而分析对应手机平台加速度传感器的量纲、采样频率等参数，具有一定的代表性。

2.1.2.2 获取加速度数据软件接口的多样性

根据各种手机操作系统及其相应的程序开发接口，我们可以很容易地获取加速度传感器数据。我们将各个手机的六个面分别放置在一个水平面上，用软件记录下静止时的加速度数据，在记录过程中，对每个数据均标上对应的时间戳。根据记录的数据，我们可以计算加速度传感器每个轴正负方向的数值、计算零点漂移值、统计一段时间内加速度传感器的采样频率。表 2.1 列出了 N95、HTC、iphone、itouch 的相应参数，是从各种手机采集到的数据中得到的统计信息，在表中对比了不同手机操作系统下加速度传感器的参数。不同加速度传感器静止时读数的差异是它们具有不同量纲的表现，而采样频率的差异则是由硬件和操作系统软件共同导致的。数据量纲和采样频率的差异不同程度地影响了后续步骤提取到的行为特征。

表 2.1 不同手机平台下获得的加速度数据的相关参数

设备名	操作系统	加速度采集 程序语言	静止读数	零点 漂移	采样频率
N95	Symbian	C++	57 或者 -57 左右	≈ 0	32Hz 左右，波动不大
HTC	Android	Java	9.8 或者 -9.8 左右	≈ 0	[40Hz, 65Hz]， 波动较大
iphone	iOS	Objective-C	1.0 或者 -1.0 左右	≈ 0	可调，设定到 100Hz 时，实 际采样频率为 91Hz 左右，波 动不大
itouch	iOS	Objective-C	1.0 或者 -1.0 左右	≈ 0	可调，设定到 32 时，实际采 样频率为 32 左 右，波动不大

为了涵盖大部分的手机类型，使得我们研究的成果更具有普适性，经过前面的调研，

我们发现 Android、Symbian 和 iOS 操作系统手机的市场占有率超过了 90%。因此我们有必要研究能够横跨这三个平台的通用识别模型。对于上述三种操作系统的手机，我们分别选择 HTC、N95 和 iPhone 手机作为研究对象。

三种操作系统平台上的开发语言各不相同，在 HTC、N95 和 iPhone 上，我们分别用 Java 语言、C++语言和 Objective-C 语言开发了加速度数据采集程序，程序采集界面如图 2.6 所示。



图 2.6 三大主流智能手机及其加速度数据采集程序

为了方便后续过程的处理，三种加速度数据采集程序采用相同的数据存储格式。在采集过程中，每采集到加速度传感器三个轴的数据，就将其存储到文本文件中。每行数据代表一个记录，格式如表 2.2 所示。

表 2.2 原始加速度数据存储格式

日期	时间	x 轴数据	y 轴数据	z 轴数据
2012-06-07	10:46:01	-30	56	96

2.1.3 数据采集要求

针对不同用户具有不同的行为特点，我们采集了课题组 12 位人员的数据。

为了研究相同行为在不同位置的数据表现，我们必须采集手机位于不同位置时的数据。根据相关调查研究，裤袋、腰部和手中是用户放置手机最频繁的地方。因此我们采

集了左裤袋、右裤袋、左手、右手、左腰部、右腰部等 6 个位置的数据。

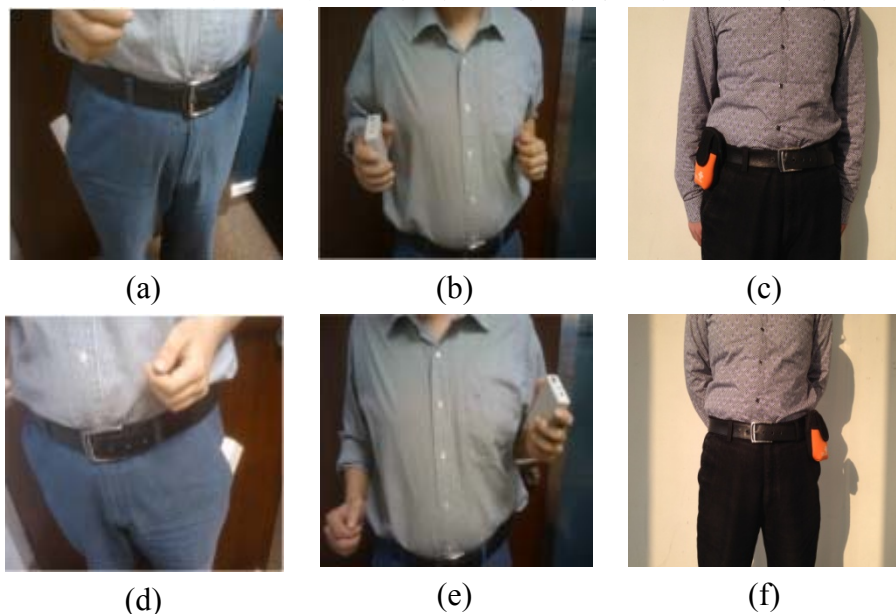


图 2.7 采集加速度数据时手机放置的位置

日常生活中，人们常见的与运动相关的行为有静止、散步、跑步、上楼梯和下楼梯。因此，我们主要采集这五种行为的数据。

2.2 数据采集过程

在采集过程中，一人携带设备做各种动作，另一人详细记录动作的开始时间、结束时间、动作类别、传感器类别、传感器位置。

采集路线为，在中国科学院计算技术研究所一楼静止 1 分钟，走楼梯到 8 楼，再静止一分钟，走楼梯到 1 楼，静止 1 分钟，散步 5 分钟到中关村体育场，静止 2 分钟，绕体育场跑步 5 分钟，最后静止 2 分钟。

每个人把每个手机放在每一个位置，重复上述路线一次。根据采集人员的体力，分多天采集，每天采集不同的人。因此采集时间累计达到了 30 天。

数据采集过程的记录格式如下表所示。

表 2.3 数据采集过程中的记录格式

姓名	日期	手机	位置	时间段	行为
小王	2012-10-07	N95	右手中	8: 30-9: 00	散步

2.3 数据分割和标定

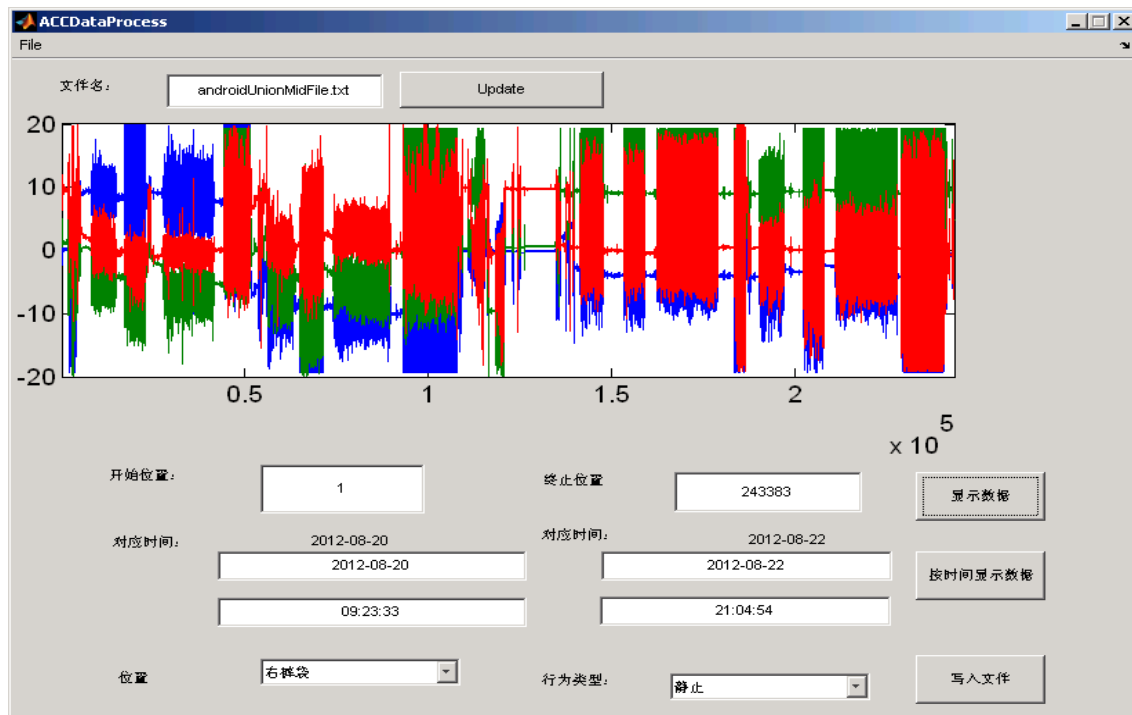


图 2.8 数据分割和标注软件

为了便于分割和标定数据，我们设计了如图 2.8 所示的软件。将每个原始数据文件通过点击 Update 按钮导入系统，三轴加速度数据就被显示在界面上。如果单靠观察，不能区分开所有的行为，因此必须利用数据自身的信息和采集过程中记录的信息。

数据采集过程中记录的可用于分割行为数据的信息包括行为名称、行为的起始和结束时间，可以根据它们确定某种行为的开始和结束时间，这种方法的缺点在于：用于记录的时钟和采集数据的手机的时钟虽然经过调校，但不是完全同步的，它们之间会有若干秒的差异，如果按照记录的时钟信息去自动分割行为数据，会为其引入噪音数据，最终对识别模型的性能带来负面影响。

在数据采集过程中，我们刻意安排每个采集人员在某种行为的开始前和开始后静止 1 至 2 分钟，就是要能够清楚地在图中看出行为之间的分割情况。然而，即使是这样，由于每秒钟能够采集 30 至 100 个数据，在图中寻找行为的精确到秒的起止时间也是需要不停地调整时间窗口，非常耗时。

在研究过程中，我们发现了静止行为的规律，反映在数据上就是：如果一秒钟内合成加速度数据的方差小于均值的 5%，则可认为该一秒钟对应的数据为静止状态。基于这条启发式规则，我们可以自动地寻找到所有的静止状态数据，并用连续的静止状态数据将所有行为数据分割开来。这样对于每一段行为数据，我们去查看记录，寻找对应的行为名称，完成数据的分割和标定。

2.4 特征提取

常用的特征提取技术是滑动窗口法。文献中大多采用窗长 2 秒，步长 1 秒的滑动窗口技术提取特征。这样做的前提条件为传感器的采样频率必须是固定的，如果采样频率不固定，会导致每个时间窗口内的数据量不同，会影响到提取到的特征的分布情况。当采样频率由于硬件或者软件的原因确有不同时，我们可以采取插值或者抽样的技术来解决这个问题。假设传感器的采样频率为 f Hz，则一个窗口中恰好包括 $2 \times f$ 个数据。后文还会讲到，在窗口中提取频域特征时，通常会采用快速傅里叶变换(Fast Fourier Transform, FFT)的方法，而该方法要求数据的个数为 2 的 n 次幂，因此，我们通常选取的窗口大小为：

$$window\ size = 2^{\lceil \log_2(2 \times f) \rceil}$$

其中函数 $\lceil x \rceil$ 返回不小于 x 的最小整数。

当窗长确定后，就可在每个窗口内提取特征，常见的加速度数据特征可分为两大类：时域(Time Domain)特征和频域(Frequency Domain)特征。

2.4.1 时域特征

为了便于描述各个特征，我们假设每个窗口内的 n 个三轴加速度数据用 W 表示：

$$W = \{(a_x^1, a_y^1, a_z^1), (a_x^2, a_y^2, a_z^2), \dots, (a_x^n, a_y^n, a_z^n)\}$$

其中 (a_x^n, a_y^n, a_z^n) 表示第 n 个加速度数据分别在 x 、 y 、 z 轴上的读数。

(1) 合成加速度

a_x^n, a_y^n, a_z^n 分别表示某时刻加速度数据在三个轴上的读数，这三个值是有方向的。如果加速度传感器的位置固定，其方向信息就有很大的价值，比如可以用来衡量其附着物的姿态。在某些场景下，比如识别动作的剧烈程度，则没必要考虑方向信息的影响，只需考虑合成加速度值即可：

$$a_{syn}^n = \sqrt{(a_x^n)^2 + (a_y^n)^2 + (a_z^n)^2}$$

所以一个窗口内所有数据的合成加速度可表示为：

$$W_{syn} = \{a_{syn}^1, a_{syn}^2, \dots, a_{syn}^n\}$$

(2) 信号向量幅值 (Signal Vector Magnitude, SVM) [Karantonis,2006]

即窗口内合成加速度的均值，可用来表示窗口内的平均幅度。

$$SVM = \frac{1}{n} \sum_{i=1}^n a_{syn}^i$$

(3) 均值 (Mean)

窗口内 x 轴加速度均值为：

$$a_x^{mean} = \frac{1}{n} \sum_{i=1}^n a_x^i$$

它代表 x 轴加速度数据在一个时间窗口内的平均特征。

同理，可以计算 y 轴和 z 轴的平均加速度，也可用该方法计算合成加速度的均值。

(4) 标准差 (Standard Variance)

标准差反映数值相对于平均值 (Mean) 的离散程度。

窗口内 x 轴加速度标准差为：

$$a_x^{\text{std}} = \sqrt{\frac{1}{n-1} \sum_{i=1}^n (a_x^i - a_x^{\text{mean}})^2}$$

同理可计算 y 轴和 z 轴的加速度标准差。

(5) 众数 (Mode)

是一个窗口内出现次数最多的数值，有时众数在一组数中有多个。

(6) 最大值 (Maximum)

(7) 最小值 (Minimum)

(8) 范围 (Range)

指的是最大值与最小值的差。可用于区分运动剧烈程度不同的行为，譬如“跑步”的范围比“散步”大。

(9) 相关系数 (Correlation Coefficient)

$$\rho_{x,y} = \frac{\text{cov}(x,y)}{\sigma_x \sigma_y}$$

$\text{cov}(x,y)$ 表示向量 x、y 的协方差， σ_x 、 σ_y 分别表示向量 x、y 的标准差， $\rho_{x,y}$ 衡量加速度传感器两个轴之间的相关性，值为正说明正相关，为负说明负相关，为 0 表示不相关。这个特征对于区分仅在一个方向的运动特别有效。

(10) 过均值点个数 (Mean-Crossings)

过均值点指的是那些和均值相同的点，过均值点个数就是相应的个数。

(11) 信号幅值面积 (Signal Magnitude Area, SMA) [Mathie,2003]

$$\text{SMA} = \frac{1}{t} \left(\int_0^t |x(t)| dt + \int_0^t |y(t)| dt + \int_0^t |z(t)| dt \right)$$

SMA 指的是三个轴的幅值与坐标轴围成部分的面积之和。该特征一般被用于区分静止状态和运动状态。

2.4.2 频域特征

频域分析技术通常被用于发现传感器信号中的重复性信息，这种重复性通常对应着行为中的周期性，比如散步和跑步都是周期性运动。最常使用的信号变换技术是傅里叶变换，它允许人们在频域表示时域信号的一些特性，比如直流分量 (Direct Current, DC)、主要频率成分等。在频谱表示方法中，信号的周期由谱峰位置表示。比如，一个周期为 0.5 秒左右的信号，将会在傅里叶变换系数对应的 2Hz 处具有较高的峰值。这种频谱分

析技术主要由快速傅里叶变换（Fast Fourier Transform, FFT）算法计算。

直流分量 DC，是傅里叶变换系数的第一个分量，是信号的均值，一般比其他系数要大得多。

关键系数的谱分析（Spectral Analysis of Key Coefficients），有些研究人员利用一些谱系数的和作为区分特定行为的特征。比如，从 0.5Hz 到 3Hz 的系数被用来区分跑步和散步行为[Santos,2009; Welbourne,2005]。

在文献[Mathie,2003]中，作者设计了一个从信号谱中检测平均步速的算法，该算法在 0.7 至 3Hz 的范围中寻找一个频率峰值。该峰值和一个基准噪音值对比，如果信噪比大于一个预先设定的阈值，则认为此峰值处对应的频率就是人的步频。还有一些作者采用最大峰值处对应的频率来判断用户是在散步还是在跑步，如果是跑步，还能判断跑步的步频[Kawaharaq,2005; Kawaharaq,2007; Kawaharaq,2009]。

功率谱密度（Power Spectral Density, PSD）。PSD[Wang,2009]被定义为时间序列数据自相关数据的傅里叶变换，用以描述数据在频域的能量分布。PSD 的幅度统计特征定义如下：

$$\text{幅度 (Amplitude)} : \mu_{amp} = \frac{1}{N} \sum_{i=1}^N C(i)$$

$$\text{标准差 (Standard deviation)} : \sigma_{amp} = \sqrt{\frac{1}{N} \sum_{i=1}^N (C(i) - \mu_{amp})^2}$$

$$\text{偏度 (Skewness)} : \gamma_{amp} = \frac{1}{N} \sum_{i=1}^N \left(\frac{C(i) - \mu_{amp}}{\sigma_{amp}} \right)^3$$

$$\text{峰度 (Kurtosis)} : \beta_{amp} = \frac{1}{N} \sum_{i=1}^N \left(\frac{C(i) - \mu_{amp}}{\sigma_{amp}} \right)^4 - 3$$

其中， $C(i)$ 是第 i 个频率窗口的 PSD 幅度， N 是频率窗口数。

同理可定义 PSD 的形状统计特征如下：

$$\text{幅度 (Amplitude)} : \mu_{shape} = \frac{1}{S} \sum_{i=1}^N i C(i)$$

$$\text{标准差 (Standard deviation)} : \sigma_{shape} = \sqrt{\frac{1}{S} \sum_{i=1}^N (i - \mu_{shape})^2 C(i)}$$

$$\text{偏度 (Skewness)} : \gamma_{shape} = \frac{1}{S} \sum_{i=1}^N \left(\frac{i - \mu_{shape}}{\sigma_{shape}} \right)^3 C(i)$$

$$\text{峰度 (Kurtosis)} : \beta_{shape} = \frac{1}{S} \sum_{i=1}^N \left(\frac{i - \mu_{shape}}{\sigma_{shape}} \right)^4 C(i) - 3$$

其中, $S = \sum_{i=1}^N C(i)$ 。

此外, 谱能量和信息熵在其他文献中也有被用到, 虽然在本文中没有采用, 我们在此也一并介绍。

谱能量 (Spectral Energy), 是谱系数平方的均值。文章[Nham,2008]用其识别交通模式, 比如散步、骑自行车、跑步、驾驶汽车等。

信息熵 (Information Entropy), 可由傅里叶变换系数 (不包括直流分量) 计算得到 [Ho,2004]。熵用于区别那些具有相似能量值的不同行为。比如 Bao 等人[Bao,2004]使用频域熵来区分慢跑和骑车, 这两种行为具有相似的能量级别。

2.5 样本集统计信息

基于滑动窗口技术, 我们从原始数据中分割出以 2 秒为单位的窗口, 每一个窗口中的数据先经过合成加速度运算, 消去朝向的影响, 然后基于合成加速度向量, 提取多个特征, 以组成一个样本。提取的特征为: (1) 时域的最大值、最小值、均值、标准差、众数、过零点个数、范围、信号幅值面积等 8 个特征; (2) FFT 变换后的 64 个系数; (3) 功率谱密度 PSD 的 8 个特征。总计 80 个特征。

样本个数的统计信息如表 2.4、表 2.5、表 2.6 所示。其中 L1~L6 分别代表右裤袋、右腰部、右手、左手、左腰部、左裤袋。Obj_01~Obj_12 分别代表 12 个数据采集人员, A1~A5 分别代表静止、散步、跑步、上楼梯、下楼梯。

表 2.4 中共包含 84983 个样本, 表 2.5 中共包含 83446 个样本, 表 2.6 中共包含 86580 个样本。

表 2.4 基于 Android 平台手机采集的样本信息

	L1					L2					L3					L4					L5					L6				
	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5
Obj_01	102	330	303	112	104	146	323	224	124	103	143	274	143	135	109	66	304	279	129	124	429	296	306	135	128	295	302	302	26	128
Obj_02	336	300	613	140	142	644	306	312	138	154	365	327	326	137	141	333	322	329	153	147	495	520	596	270	256	310	314	314	142	145
Obj_03	254	304	303	139	133	294	289	306	141	277	386	312	302	135	124	290	280	107	140	131	429	296	306	135	128	363	302	302	141	128
Obj_04	121	304	303	139	133	343	305	306	130	148	417	304	304	155	141	367	333	316	148	123	497	292	301	129	138	519	607	605	265	251
Obj_05	294	295	303	131	123	247	295	306	131	122	331	308	303	161	117	246	299	303	138	119	410	304	286	157	140	306	303	308	159	125
Obj_06	337	300	307	145	133	318	308	288	143	126	288	308	304	143	124	324	304	308	136	130	316	303	307	139	132	383	302	303	143	131
Obj_07	254	304	303	139	133	294	289	306	141	277	386	312	302	135	124	290	280	107	140	131	429	296	306	135	128	363	302	302	141	128
Obj_08	348	316	297	116	138	315	303	279	117	137	282	307	306	139	132	233	322	316	141	120	343	317	306	126	139	383	302	288	125	131
Obj_09	345	305	302	147	144	318	313	311	147	143	247	299	310	121	145	252	314	301	141	167	315	297	308	147	149	301	302	303	145	145
Obj_10	271	319	306	145	121	394	310	288	143	110	312	305	315	123	112	293	313	307	122	108	467	308	309	144	128	445	325	306	68	130
Obj_11	426	419	172	83	149	389	437	181	84	68	295	180	183	88	74	273	185	159	83	72	236	221	163	161	139	366	417	175	102	166
Obj_12	330	244	183	80	64	304	171	180	76	64	280	248	194	79	72	147	213	187	90	72	443	191	183	64	68	382	190	182	82	67

表 2.5 基于 IOS 平台手机采集的样本信息

	L1					L2					L3					L4					L5					L6				
	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5
Obj_01	44	324	234	104	103	367	304	52	119	124	119	292	307	130	107	92	351	245	136	105	447	302	303	141	125	192	334	271	113	109
Obj_02	527	305	313	138	153	403	327	322	135	141	355	321	329	148	146	500	549	323	135	140	368	301	314	139	145	348	300	312	138	141
Obj_03	328	291	305	143	138	300	304	52	135	124	410	280	304	137	133	366	297	304	136	126	395	302	303	141	125	313	304	301	140	130
Obj_04	311	304	302	130	148	501	301	303	151	141	412	330	314	144	125	358	292	310	130	135	351	300	304	129	127	354	306	304	132	125
Obj_05	319	296	73	131	123	389	305	305	149	116	330	299	303	136	119	329	303	290	160	142	375	291	307	158	125	316	299	302	144	132
Obj_06	320	306	289	143	125	334	309	304	142	123	302	304	310	133	131	274	302	305	140	131	355	302	303	144	132	335	300	308	276	130
Obj_07	328	291	305	143	138	300	304	52	135	124	410	280	304	137	133	366	297	304	136	126	395	302	303	141	125	313	304	301	140	130
Obj_08	359	302	281	112	136	376	307	303	139	134	282	321	302	139	118	261	321	309	126	140	338	301	299	128	132	347	315	296	116	140
Obj_09	357	314	309	145	142	330	298	307	118	144	233	303	301	142	155	191	305	311	149	151	338	302	303	142	145	340	301	300	147	143
Obj_10	220	291	591	143	138	352	301	312	122	109	300	312	310	121	109	324	305	312	142	128	376	486	313	140	132	313	320	303	146	120
Obj_11	465	419	167	80	144	390	437	176	83	69	281	180	182	86	76	298	184	154	80	71	251	180	155	163	138	381	457	172	101	156
Obj_12	301	244	182	80	64	298	171	183	77	63	262	249	192	80	73	127	213	187	86	72	442	190	182	66	68	366	191	183	77	66

表 2.6 基于 Symbian 平台手机采集的样本信息

	L1					L2					L3					L4					L5					L6				
	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5	A1	A2	A3	A4	A5
Obj_01	197	264	298	78	96	68	299	284	125	108	58	334	204	123	94	58	324	224	123	94	89	309	248	117	89	69	309	204	123	88
Obj_02	415	328	327	138	139	375	323	329	148	149	485	563	327	137	141	302	303	313	142	146	353	301	316	137	141	515	308	313	136	153
Obj_03	461	309	304	138	125	431	281	305	138	133	369	298	305	137	129	296	303	305	142	128	307	304	305	142	136	346	292	307	141	136
Obj_04	517	300	313	152	144	415	331	315	149	125	411	292	312	131	136	252	301	307	130	130	339	307	307	142	129	284	303	303	128	147
Obj_05	495	307	305	154	118	237	301	305	132	119	318	303	287	166	139	264	292	311	167	134	317	300	303	144	132	399	298	379	136	124
Obj_06	268	310	302	143	121	343	308	313	135	130	276	302	310	142	132	279	304	308	144	132	335	301	309	148	132	275	308	288	155	128
Obj_07	461	309	304	138	125	431	281	305	138	133	369	298	305	137	129	296	303	305	142	128	307	304	305	142	136	346	292	307	141	136
Obj_08	337	308	303	142	133	312	319	302	141	119	813	596	610	274	258	253	307	290	128	132	339	315	298	116	139	357	305	286	116	136
Obj_09	373	297	305	120	143	279	302	305	146	156	187	304	310	149	151	245	302	302	146	146	306	303	302	149	141	328	312	311	147	142
Obj_10	367	304	313	122	111	352	312	308	120	111	261	319	312	142	129	304	490	315	142	130	278	358	305	145	121	349	311	290	146	109
Obj_11	372	417	170	102	153	336	441	180	83	68	313	184	150	81	69	240	180	182	81	74	266	220	153	164	139	454	425	167	79	142
Obj_12	993	559	512	229	203	431	281	305	138	133	120	211	201	92	77	176	247	193	74	75	307	304	305	142	136	280	247	181	80	68

2.6 小结

基于机器学习算法构建行为识别模型，需要有大量的已精确标定样本。本章首先介绍了目前国际上已经公开的几个行为数据集，分析了其不能满足我们研究需要的原因。然后，基于我们面临的挑战问题对数据集的要求，精心地采集了较大规模的行为数据，为后续章节的研究奠定了良好的基础。

第三章 面向异质加速度传感器数据的预处理方法

本章首先例证了对异质加速度传感器数据进行预处理的必要性，回顾了已有的加速度传感器数据处理方法，提出了基于静止状态检测和时频域变换的数据预处理方法，并通过实验验证了方法的有效性。

3.1 问题的提出

构建基于智能终端(尤其是手机)的行为识别模型必须考虑其宿主设备数据的特性。为了直观地了解不同设备数据的特性，我们将基于 Android 系统、Symbian 系统和 iOS 系统的三个手机(对应的手机型号分别是 HTC、N95、iphone)绑定在一起，放置在数据采集者的右裤袋中，分别在其站立不动和慢跑时采集了相应的加速度数据。通过手工分割，分别将静止和慢跑时的加速度数据存储在不同的文件中，并通过调用相应 MATLAB 命令，将静止数据和慢跑数据的合成加速度分别描绘在如图 3.1 和图 3.2 所示的图片上。

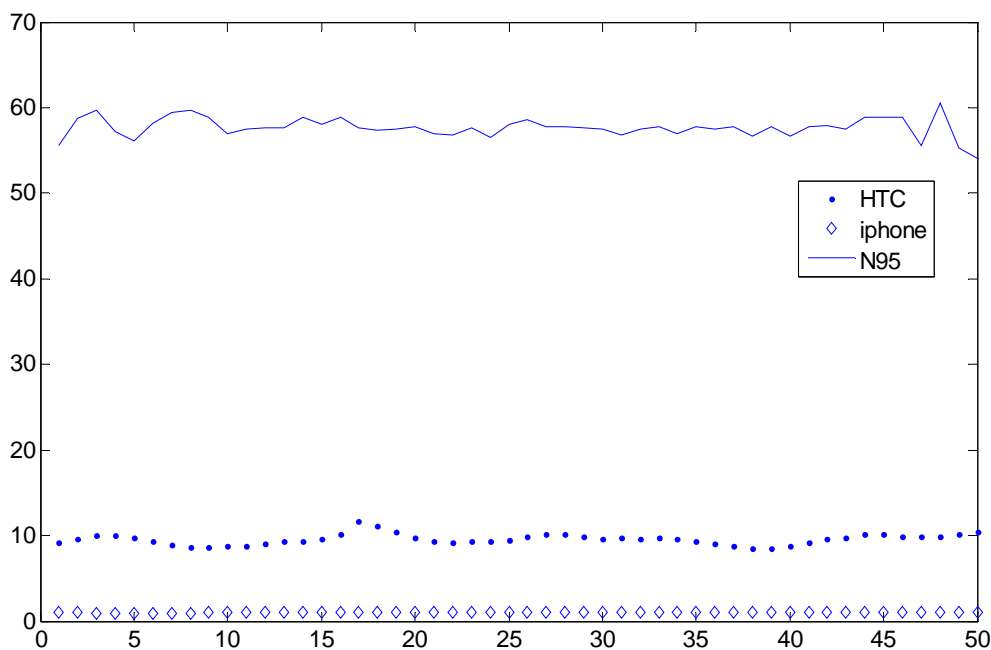


图 3.1 三种手机操作系统平台下静止时的合成加速度

从图 3.1 可以看出，静止时 iPhone 手机的合成加速度大约为 1 左右，由静止时重力加速度的大小为 $1g$ 可知 iPhone 手机读取到加速度的单位是 g ；HTC 手机的合成加速度为 10 左右，由静止时重力加速度的大小为 9.8 米/秒^2 可知 HTC 手机读取到的加速度的单位是 米/秒^2 ；N95 手机的合成加速度大约为 57，其单位未知。这些读数的不同可以

理解为不同手机平台下的 API 接口度量加速度的量纲不同。三种平台的手机通过各自的加速度数据采集程序读取到的静止时数据为什么存在这么明显的差距，我们想探究其原因，然而通过查阅相关的资料，并没有官方明确的声明。虽然如此，三种不同手机平台下采集的加速度传感器数据的量纲不同，是一个客观存在的现实。

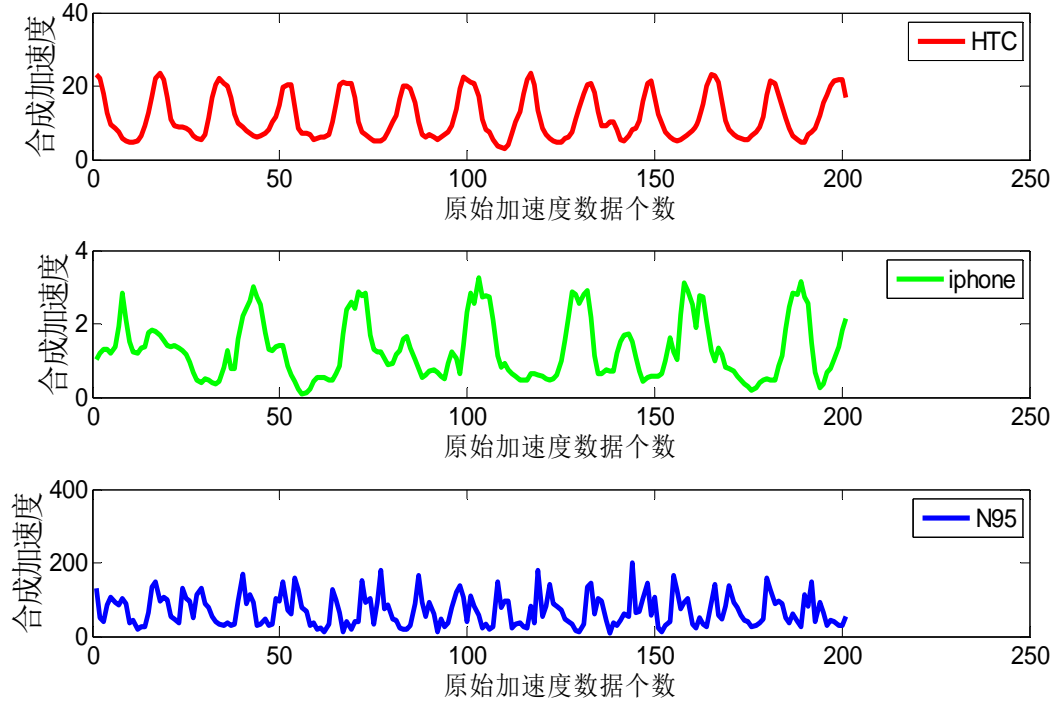


图 3.2 三种手机操作系统平台下跑步时的合成加速度

图 3.2 中所描绘的曲线表现的是 Android 平台手机 (HTC)、iOS 平台手机 (iphone)、Nokia 平台手机 (N95) 在随用户慢跑时分别采集到的 200 个数据的合成加速度值。跑步是周期性的运动，表现在加速度数据的波形图上，呈现出波峰波谷交替的形态。从图 3.2 可以清楚地看出，三种平台手机在 200 个数据内感知到的波峰波谷数目是不同的，也就是所包含的用户慢跑经过的步数是不同的。由于三个手机是绑定在一起的，因此无论是哪个手机，每两个波峰之间的时间差是相同的，于是可以推算三个手机采集 200 个数据所耗费的时间是不同的，从而可以进一步推知三个手机的加速度采样频率是不同的。

综上所述，不同手机平台的静止时加速度（量纲）和采样频率是不同的。

在本文，具有不同参数的加速度传感器我们统称为异质加速度传感器。异质加速度传感器的量纲和采样频率的不同，决定了不同加速度传感器采集到的数据分布的不同，数据分布的不同必然导致在一个传感器数据上训练的行为识别模型不能有效地识别其他参数不同传感器的数据，这样就有悖于建立一个普适于设备的行为识别模型的初衷，是我们必须直面的问题。

3.2 相关研究工作

在处理加速度传感器数据方面，有一些文献做出了探索。

Yang 在文章[Yang,2010]中指出，他们通过观察发现,当 N95 手机静止放置在平面上时,手机内置的加速度传感器与重力垂直的另两个方向的读数并非为 0,而是有一定的数值,他们称之为跳动错误（Jittering Error），而这两个数值要远远小于与重力方向平行的加速度传感器轴上的读数。为了进一步减少跳动错误对最终行为识别模型识别精度的影响，他们首先采用窗口大小为 5 的均值去噪方法减少数据中的噪音，然后将去噪后的数据缩小 5 倍进一步减少噪音的值。实验表明，该方法能够有效地对数据进行平滑，减少加速度数据的噪音，提高模型的识别精度。该方法的不足之处在于，对结果缩小 5 倍，是一个经验值，需要反复实验得到，对于其他类型手机，应该缩小多少倍，没有一个统一的方法。

Yang 在文章[Yang,2010]中提到对实验用 N95 手机内置加速度传感器产生的数据进行校准的方法。他们首先寻找一个与重力方向垂直的平面，然后将手机的六个平面分别放置在该平面上，并记录下加速度传感器每个轴的读数。假设三个轴正方向的读数分别为 (a_1, a_2, a_3) ，负方向的读数分别为 (a'_1, a'_2, a'_3) ，则据此可以计算加速度传感器每个轴的偏置和标度因子,我们以计算 x 轴的偏置和标度因子为例：

$$offset_x = \frac{a_1 + a'_1}{2}$$

$$scale_x = \frac{a_1 - a'_1}{2}$$

该方案的不足之处在于可操作性不强，首先是不能够保证平面是和重力方向绝对垂直的，其次是手机未必是标准的六面体形状，这两个条件若得不到满足，则必定会引入新的误差，而存在的误差会进一步导致结果的不准确。为了减少引入的误差，通常这个校准过程要重复几次，取多次实验结果的平均值作为最终的偏置和标度因子。

在文章[Lukowicz,2004a]中，作者提到，在基于加速度传感器的穿戴式应用中，一个潜在的实际问题就是加速度传感器校准问题。每一个加速度传感器都有自己特定的灵敏度（ s ）和偏移量（ o ），通过公式 $a=(x-o)/s$ 可以把加速度 x 转化为以 g 为量纲的值。通常情况下，不同加速度传感器的 o 、 s 值是不同的，而且还会受到温度漂移（Thermal Drift）的影响。他们认为加速度由两部分构成：静态加速度和动态加速度，静态加速度是由重力加速度引起的，动态加速度是由外部其他作用力引起的。他们进一步指出，这两种加速度通常难以分解开来。他们还指出，在静止状态下，三轴加速度传感器的读数 x ， y ， z 和各个轴上的 o_x ， o_y ， o_z ， s_x ， s_y ， s_z 的值满足下面的关系：

$$\left(\frac{x-o_x}{s_x}\right)^2 + \left(\frac{y-o_y}{s_y}\right)^2 + \left(\frac{z-o_z}{s_z}\right)^2 = 1$$

上式具有 6 个不同的参数，因此，如果能测量出 6 个不同位置静止时的加速度数据，

就能够采取诸如牛顿迭代（Newton Iteration）的方法求出这 6 个参数。这种方法的难点在于：1）如何在用户不参与交互的情况下，分割出 6 个不同位置的静态加速度；2）基于牛顿迭代或者梯度下降的方法在求解过程中容易陷入局部最小（Local Minimum）的境地。对于第一个问题，他们以 0.5 或者 1 秒为窗口，在此窗口内，如果三个轴的数据标准差（Standard Deviation）都不高于均值的 0.5%，则认为是处于静止状态。对于在某些情况下算法容易陷入局部最小的情况，他们期望将来能有更先进的数值计算方法来提高算法的性能。

在文献[Junker,2004]中，作者用分别固定配置在臀部（hip）和右膝关节（right knee）处的两个 3 轴加速度传感器，提取 4 个轴（垂直的和向前的）上的特征，分类散步、上楼梯、下楼梯三类动作，频率在 10Hz 到 100Hz 之间分别取值，分辨率（resolution）在 1bit 到 16bits 之间分别取值，在频率和分辨率的各种组合下测试模型的识别精度，实验结果表明，当频率为 20、分辨率为 2bits 以上时，并不能显著提高模型的分类精度，即 20Hz 频率对于区分散步、上楼梯、下楼梯已经足够了。

我们在探索面向非特定加速度传感器的行为识别模型时发现，不同量纲和采样频率的原始数据如果不考虑采样频率间的差别，仅按照现存的滑动窗口的方法提取特征，则在加速度传感器 A 数据上建立的模型，对加速度传感器 B 的数据的识别率较低，达不到实用的要求。具体实验结果见 3.4.3 节。

3.3 面向异质加速度传感器数据的预处理方法

根据研究的问题及其解决方法的需要，我们提出了基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法，方法流程如图 3.3 所示。



图 3.3 基于静止状态检测和时频域变换的数据预处理方法

3.3.1 合成加速度的规范化方法

设备在静止的条件下，其合成加速度应为重力加速度 $1g$ ，根据这一准则，可以将不同量纲的传感器数据统一起来。

假设静止时 x 、 y 、 z 轴的加速度数值构成向量 (a_x, a_y, a_z) ，各个轴的偏置值 $offset$ 为 (o_x, o_y, o_z) ，各个轴的标度因子 $scale\ factor$ 分别为 (s_x, s_y, s_z) ，则它们满足公式：

$$\left(\frac{a_x - o_x}{s_x}\right)^2 + \left(\frac{a_y - o_y}{s_y}\right)^2 + \left(\frac{a_z - o_z}{s_z}\right)^2 = 1$$

在实际实验中我们发现 o_x, o_y, o_z 的值远远小于 a_x, a_y, a_z 的值，或者说它们的存在不足以影响后续数据的精度，即可认为它们可以忽略不计（相信随着生产工艺的提高，这个偏置会越来越小），则公式可近似为：

$$\left(\frac{a_x}{s_x}\right)^2 + \left(\frac{a_y}{s_y}\right)^2 + \left(\frac{a_z}{s_z}\right)^2 = 1$$

在通常情况下， x 、 y 、 z 轴的 $scale\ factor$ 近似相同，我们用 s 表示，则公式可演变为：

$$\left(\frac{a_x}{s}\right)^2 + \left(\frac{a_y}{s}\right)^2 + \left(\frac{a_z}{s}\right)^2 = 1$$

即

$$\frac{\sqrt{(a_x)^2 + (a_y)^2 + (a_z)^2}}{s} = 1$$

进一步

$$s = \sqrt{(a_x)^2 + (a_y)^2 + (a_z)^2}$$

即， s 为静止时的加速度的合成值。

于是，对于任意状态下的合成加速度数据 (a'_x, a'_y, a'_z) ，其经过规范化的合成加速度为：

$$\frac{\sqrt{(a'_x)^2 + (a'_y)^2 + (a'_z)^2}}{s}$$

即，可以采用除以静止时合成加速度的方法将加速度数据规范化到静止时量纲为 1 的情境，使得不同 $offset$ 和 $scale\ factor$ 的传感器数据有了可比性。

3.3.2 基于时间窗的采样频率一致化方法

1 重采样前时间窗口的选定

散步、跑步、上楼梯、下楼梯等动作具有周期性，表现在根据加速度传感器采集到的数据所描绘的曲线上，是该曲线呈现波峰波谷交替出现的形态。当我们考虑对这些行为数据进行操作或者处理的时候，最直观的想法是以一个行为周期为单位进行处理。文献中大多采用 2 秒钟的时间窗口[Bao,2004; Ravi,2005]，即大家普遍认为在 2 秒钟的时间

窗口中采集到的数据足以涵盖一个动作周期。

在加速度传感器采样频率恒定的情况下，每秒钟采集到的数据个数是相同的，因此可以直接采集数量为采样频率 2 倍的数据，作为 2 秒钟时间窗口内的数据。然而目前我们面临的问题是采样频率是变化的，这样就不能简单地以固定数量作为时间窗口。

在数据采集过程中，每一个数据对应有一个时间戳，带有时间戳的数据格式如下： $(year, month, day, hour, minute, second, a_x, a_y, a_z)$ ，我们将具有相同时间戳的数据视为同一秒，取出其中的数据 $\{(a_x^i, a_y^i, a_z^i) | 1 \leq i \leq m\}$ ，其中 m 是该 1 秒窗口中的数据个数。按照这样划分时间窗口的方式，则不同的时间窗口内的数据个数是变化的。

2 重采样后数据窗口长度的选定

对于基于加速度传感器数据的应用，从原始数据提取特征是必不可少的阶段，如第一章所述，常用的提取特征的方法是滑动窗口法。而滑动窗口的大小是恒定不变的，并且为了应用快速傅里叶变换（Fast Fourier Transform, FFT），通常取窗口的大小为 2 的整数幂。根据在研究过程中积累的经验，我们将窗口的大小设定为 64，即加速度传感器的采样频率为 32。

3 一秒钟时间窗口内的重采样操作

通过上面的讨论，我们根据经验要将具有不同采样频率的加速度传感器采集的数据重采样到频率为 32Hz，即：如果采样频率低于 32Hz 则进行插值操作，如果采样频率高于 32Hz 则进行下采样操作。我们采取线性插值的方法将数据插值到频率为 32Hz，线性插值的方法不仅能够有效的补齐数据，很好地保持原有的波形，而且时间复杂度低，计算量小。

3.4 实验及结果分析

我们采用决策树算法建立行为识别模型，测试我们的方法是否能够有效地将不同加速度传感器的数据规范化到一致的数据空间。本部分的安排如下：1) 量纲对数据分布影响实验；2) 采样频率对数据分布影响实验；3) 我们的方法对数据分布影响实验。

实验的硬件平台是一台普通的 PC 机，CPU 主频为 2.6GHz，软件平台是 MATLAB2009a。决策树算法的实现采用的是 MATLAB2009a 自带的分类与回归树（Classification and Regression Tree）[Breiman,1984]，它根据信息熵理论构建成一棵二叉决策树，在面对诸如存在缺失值、变量数多等问题时表现得非常稳健，且具有训练速度快的优点。该软件包还提供有显示二叉树结构的程序，能够以直观的图示呈现生成的决策树模型，且二叉树结构简单，简单明了，易于分析其中存在的问题，非常适合我们的需要。

3.4.1 数据准备

由于我们仅考虑加速度传感器的不同参数对行为数据分布的影响，因此我们尽量减

少其他因素对实验的影响，采集的是同一个人、同一个位置、不同设备的数据。数据采集过程中所采用的设备有 iPhone、iTouch、N95、HTC。采集数据时，数据采集者将四种设备同时放置在自己的右裤袋中。采集的行为包括静止、散步、跑步、上楼梯、下楼梯等，为获得较充分的样本，要求每种行为的持续时间不低于 20 分钟。每种行为之间要求有接近一分钟的静止行为，以方便后续步骤对数据的分割，采集人在采集过程中记录下每种行为的开始和结束时间。

对采集到的数据用本文开发的分割程序进行手工分割，每种设备的数据分别存储到各自的文件中。

我们采取滑动窗口的技术提取特征。窗口的大小为 2 秒，滑动步长为 1 秒。这种采用滑动窗口方法提取加速度传感器数据特征的技术在文献中得到了广泛应用[Bao,2004; Ravi,2005]。提取的特征为：（1）时域的最大值、最小值、均值、标准差、众数、过零点个数、范围、信号幅值面积等 8 个特征；（2）FFT 变换后的 64 个系数；（3）功率谱密度 PSD 的 8 个特征。总计 80 个特征。

3.4.2 量纲对数据分布的影响

从表 2.1 可以看出，满足条件的设备是 N95 和 iTouch，它们的采样频率都是 32Hz 左右，且波动不大；N95 静止时读数的绝对值为 57 左右，iTouch 静止时读数的绝对值为 1 左右，有较大的差别。分别采取窗口大小为 64 个数据（2 秒）、滑动步长为 32 个数据（1 秒）的滑动窗口技术提取特征。

当未对合成加速度进行规范化处理时，用 N95 和 iTouch 各自样本的一半数据作为训练集建立决策树模型，另一半作为测试集测试模型的精度，得到的结果如表 3.1 所示。

表 3.1 合成加速度未规范化时 N95 和 iTouch 模型的分类精度

设备名	训练精度	测试精度
N95	0.9846	0.9566
iTouch	0.9941	0.9768

从表 3.1 可以看出，即使是未对合成加速度进行规范化处理，在各个设备自身数据集上建立的模型均能取得良好的分类性能，说明无论加速度传感器静止时读数是多少，采样频率是多少，基于其采集到的静止、散步、跑步、上楼梯、下楼梯等五种行为数据都具有很好的可分性。

下面在上述数据集上进行模型的交叉验证实验，即用 N95 数据建立的模型去分类 iTouch 的数据和用 iTouch 数据建立的模型去分类 N95 的数据，实验结果如表 3.2 所示。

表 3.2 合成加速度未规范化时 N95 和 itouch 模型的交叉验证

训练集	测试集	测试精度
N95 数据	itouch 数据	0.2311
itouch 数据	N95 数据	0.2293

从表 3.2 可以看出, 未对合成加速度进行规范化处理时, 用 N95 数据建立的模型分类 itouch 的数据的识别率只有 23%左右, 按照机器学习理论, 这是由于作为测试数据的 itouch 数据集和作为训练集的 N95 数据集的分布不同造成的, 同理用 itouch 数据建立的模型分类 N95 的数据也存在相同的结果。由此实验可以验证“加速度传感器的量纲(静止时的读数不同反映量纲的不同)不同会影响到数据的分布情况, 导致异质加速度传感器数据分布的不同, 进而会影响后续行为识别模型的迁移性”。

表 3.3 是用 N95 数据建立的模型区分 itouch 数据时的混淆矩阵 (Confusion Matrix), 表 3.4 是用 itouch 数据建立的模型区分 N95 数据时的混淆矩阵 (Confusion Matrix)。

表 3.3 N95 的模型区分 itouch 数据的 Confusion Matrix

	静止	散步	跑步	上楼梯	下楼梯
静止	550	0	0	0	0
散步	550	0	0	0	0
跑步	550	0	0	0	0
上楼梯	304	0	0	0	0
下楼梯	426	0	0	0	0

表 3.4 itouch 的模型区分 N95 数据的 Confusion Matrix

	静止	散步	跑步	上楼梯	下楼梯
静止	0	0	444	0	0
散步	0	0	696	0	0
跑步	0	0	630	0	0
上楼梯	0	0	414	0	0
下楼梯	0	0	564	0	0

从表 3.3 和表 3.4 可以看出, N95 的模型把 itouch 的数据全部区分为静止, itouch 的模型把 N95 的数据全部区分为跑步, 这是因为二者的数据量纲不同, N95 的静止时合成加速度比 itouch 跑步时的合成加速度还要大。

计算 N95 和 itouch 静止时的平均合成加速度, 分别为 55.5886 和 1.0048。在提取特征之前, 用它们分别对 N95 和 itouch 的合成加速度数据进行规范化处理。在此基础上用滑动窗口技术提取特征, 形成不同设备的样本集, 在各自的样本集上随机选取其中的一半数据作为训练集用来训练模型, 选取剩余的一半作为测试集用来测试模型, 实验结果如表 3.5 所示。

表 3.5 合成加速度规范化后 N95 和 itouch 模型的分类精度

设备名	训练精度	测试精度
N95	0.9942	0.9600
itouch	0.9966	0.9697

如表 3.5 所示, 规范化后基于 N95 和 itouch 的训练集训练的模型在各自的测试数据集上均取得了良好的分类精度。与表 3.2 对比分析, 可以发现, 合成加速度规范化前后, 各个传感器上的行为识别模型对各自数据的识别精度是一致的, 即虽然规范化操作将各自的数据映射到了新的数据空间, 但并未改变各类行为数据的可区分性。

在规范化后的数据集上进行模型的交叉验证实验, 用在 N95 的训练集上训练的模型去分类 itouch 的测试数据和用在 itouch 的训练的模型去分类 N95 的测试数据, 实验结果如表 3.6 所示。

表 3.6 合成加速度规范化后 N95 和 itouch 模型的交叉验证

训练集	测试集	测试精度
N95 数据	itouch 数据	0.8924
itouch 数据	N95 数据	0.8599

从表 3.2 和表 3.6 可以看出, N95 和 itouch 的加速度传感器数据经过规范化处理后, 模型之间的交叉验证性能得以大大提高, 从 22% 左右提升到了 85% 以上, 即模型的迁移能力得以大幅度地提升。这反映了表 3.6 中列出的测试集的分布与训练集的分布在经过规范化处理后已趋于一致, 也验证了规范化方法的有效性。

综上所述, 当加速度传感器的采样频率相似而静止时读数不同时, 会导致异质加速度传感器数据分布的不同, 通过对合成加速度的规范化处理, 能够使得异质加速度传感器的数据分布趋于一致。

3.4.3 采样频率对数据分布的影响

从表 2.1 可以看出, 满足条件的设备是 iphone 和 itouch, iphone 和 itouch 的采样频率通过程序设置可以调整, 我们设置其为 100Hz 时, 实际采样频率为 91 左右, itouch 的采样频率为 32Hz 左右, 相差较大; 它们静止时的读数绝对值都为 1 左右, 几乎没有差别。按照时间窗口提取特征, 即对每 2 秒的窗口内数据提取特征, 而不论不同窗口内数据个数是否相同。因为 iphone、itouch 的采样频率基本上稳定, 因此各自相邻时间窗内的数据量仅有细微波动。

当未对合成加速度进行采样一致化处理时, 用各自样本的一半作为训练集建立决策树模型, 另一半作为测试集测试模型, 得到的结果如表 3.7 所示。

表 3.7 采样频率未一致化时 iphone 和 itouch 模型的分类精度

设备名	训练精度	测试精度
iphone	0.9961	0.9814
itouch	0.9950	0.9782

从表 3.7 可以看出,即使是未对合成加速度进行采样一致化处理,即使是相邻窗口数据个数稍有波动,在各个设备自身数据集上建立的模型也均能取得良好的分类性能。

在上述数据集上进行交叉验证实验,分别用基于 iphone 训练数据建立的模型去分类基于 itouch 的测试数据和用基于 itouch 的训练数据建立的模型去分类基于 iphone 的测试数据,实验结果如表 3.8 所示。

表 3.8 采样频率未一致化时 iphone 和 itouch 模型的交叉验证

训练集	测试集	测试精度
iphone 数据	itouch 数据	0.7630
itouch 数据	iphone 数据	0.7220

从表 3.8 可以看出,虽然模型的交叉验证精度均可达到 70%以上,但与识别自身数据的 90%以上的精度相比,还有一定差距。按照机器学习的观点,这是由于表 3.8 中测试集的数据分布与训练集的数据分布存在差异造成的。

我们在表 3.9 和表 3.10 中分别列出 iphone 模型区分 itouch 数据的 Confusion Matrix,和 itouch 模型区分 iphone 数据的 Confusion Matrix,以观察是什么行为之间的误分类导致整体识别能力的不高,即那些行为之间的数据分布相似。

表 3.9 iphone 模型区分 itouch 数据的 Confusion Matrix

	静止	散步	跑步	上楼梯	下楼梯
静止	550	0	0	0	0
散步	0	2	0	35	513
跑步	0	0	548	0	0
上楼梯	0	0	0	304	0
下楼梯	0	0	0	14	412

表 3.10 itouch 模型区分 iphone 数据的 Confusion Matrix

	静止	散步	跑步	上楼梯	下楼梯
静止	550	0	0	0	0
散步	0	12	4	0	577
跑步	0	1	592	0	0
上楼梯	0	0	0	249	82
下楼梯	0	32	0	0	480

从表 3.9 和表 3.10 可以看出,两种模型都几乎把所有散步数据错分为下楼梯数据。

考虑到决策树是靠非终端结点上属性的阈值来分类的性质,有必要找到用以区分散步和下楼梯两种行为的属性。我们用分类和回归树软件包自带的程序将基于 iphone 训练集训练的 二叉决策树描绘出来,其形态如图 3.4 所示。

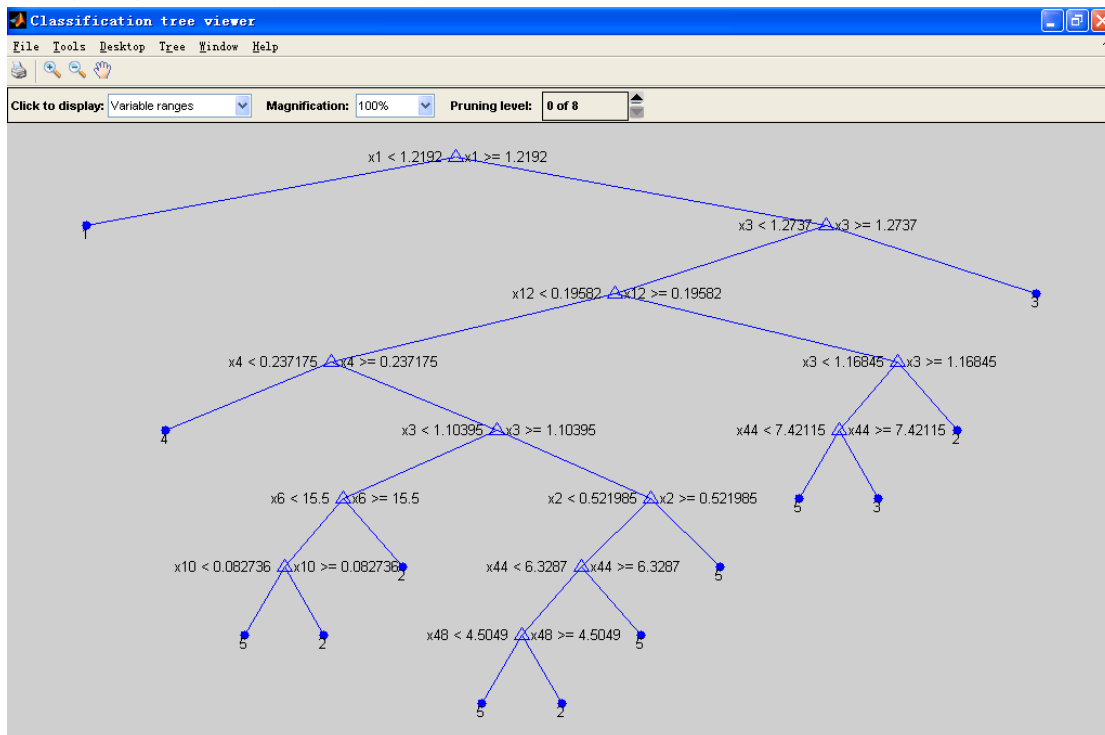


图 3.4 iphone 行为识别模型的形态

图 3.4 中,叶子结点 1、2、3、4、5 分别代表静止、散步、跑步、上楼梯、下楼梯等五种行为。出现在分类散步(2)和下楼梯(5)两种行为路径上的特征有 x_1 (最大值)、 x_3 (均值)、 x_4 (标准差)、 x_6 (过 0 点次数)、 x_{10} (傅里叶变换系数)、 x_{12} (傅里叶变换系数)、 x_{44} (傅里叶变换系数)、 x_{48} (傅里叶变换系数),其中 x_{10} 、 x_{12} 、 x_{44} 、 x_{48} 跟窗口中的数据个数相关,即与加速度传感器的采样频率相关。

因此本实验验证了“采样频率会导致异质加速度传感器的数据分布不一致,进而在后续步骤影响到行为识别模型的迁移性”。

我们对 iphone 和 itouch 的数据进行了重采样操作,将他们 1 秒钟窗口内的数据都重采样到 32Hz,然后重新进行了交叉验证实验,实验结果如表 3.11 所示。

表 3.11 采样频率一致化后 iphone 和 itouch 模型的交叉验证

训练集	测试集	测试精度
iphone 数据	itouch 数据	0.8759
itouch 数据	iphone 数据	0.9641

从表 3.11 可以看出,数据经过重采样操作,模型交叉验证的识别精度均得到了较大幅度提高。这种结果说明了使不同加速度传感器的采样频率一致能够进一步减少同类行为数据在频域特征上的差距,从而使数据的分布更一致。

3.4.4 方法的效果

从 3.4.2 节和 3.4.3 节的实验可以看出，异质加速度传感器数据分布同时受到加速度传感器量纲和采样频率的制约，应同时对二者进行考虑。在本部分，我们验证本章提出的方法对异质加速度传感器数据进行预处理后的效果。

(1) 消除朝向和角度的影响

当手机处于口袋中时，并不能保证手机的朝向和角度保持不变，而朝向和角度的改变会导致加速度传感器读数的多变性，进而影响数据分布的复杂性，进一步影响到后续步骤建立的行为识别模型的复杂性。为了消除朝向和角度的影响，我们首先对每个时刻的加速度进行合成。当某个时刻加速度传感器的读数为 (a_x, a_y, a_z) 时，合成加速度为：

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2}$$

(2) 减少量纲的影响

从表 2.1 可以看出，各种设备采集到的加速度的量纲是不同的。为了消除量纲的影响，我们以各种设备静止时的合成加速度为基准，对所有行为的合成加速度做规范化操作。比如若规范化前静止时的合成加速度为 a_{still} ，任意一个合成加速度为 a ，则规范化后静止时的合成加速度为 $\frac{a_{still}}{a_{still}} = 1$ ，任意一个合成加速度为 $\frac{a}{a_{still}}$ 。

为寻找静止时的加速度，我们观察一秒钟内的数据，当每个轴上数据的方差均不超过均值的 0.5% 时，认为设备处于静止状态，然后计算该一秒钟内合成加速度的均值作为静止时的合成加速度的估计值。

各个设备静止时的加速度如表 3.12 所示。

表 3.12 各种设备静止时的合成加速度值

设备名	静止时合成加速度估计值
N95	55.5886
itouch	1.0048
iphone	1.0067
HTC	9.96

(3) 减少采样频率不同的影响

我们从加速度数据中提取的特征都是时域或者频域的统计量，涉及到原始数据的数量，因此，为了保持每个样本提取特征过程的一致性，必须使得每个窗口中的数据量相同，即使得数据的采样频率恒定。既然无法控制手机操作系统提供的接口的采样频率，我们采用对获取到的数据进行下采样或者插值的方法使得采样频率达到一致。

在获取加速度数据时，我们同时对每一个数据标定一个时间戳，根据时间戳寻找每一个一秒钟窗口的数据，对这些数据进行下采样或者插值处理，使得它们都规整到同一

采样频率。在我们的实验中，采样频率统一为 32Hz。

经过量纲规范化和采样频率一致化后，在已有的各个传感器的样本集上进行实验。每个样本集平均分为两部分，一部分作为训练集，另一部分作为测试集。上述过程重复 10 次，以平均精度作为模型的测试精度。

表 3.13 各个传感器上模型的测试精度

设备名	测试精度
N95	0.9810
itouch	0.9812
iphone	0.9859
HTC	0.9904

从表 3.13 可以看出，经过变换后，在每个设备采集到的数据集上训练的模型对自身数据仍具有很好的区分能力，说明在变换后所形成的新的数据空间上，这五类行为仍具有较强的可分性。

(4) 对异质加速度传感器数据分布影响实验

本部分验证在采用我们的方法后异质加速度传感器数据分布是否趋于一致，由于无法直接从数据中直接计算出分布函数，不能直接比较，我们通过模型间的交叉验证，即验证在一个传感器上训练的模型，去分类其他传感器的样本会取得什么样的效果，来反推测试集是否和训练集的分布一致。实验结果如表 3.14 所示。

表 3.14 模型的交叉验证

训练设备名	测试设备名	测试精度
N95	itouch	0.9712
N95	iphone	0.9385
N95	HTC	0.8318
itouch	N95	0.9538
itouch	iphone	0.9641
itouch	HTC	0.9157
iphone	N95	0.9096
iphone	itouch	0.8759
iphone	HTC	0.9073
HTC	N95	0.9506
HTC	itouch	0.9309
HTC	iphone	0.9786
平均		0.9273

从表 3.14 可以看出, 经过变换, 对各个传感器的合成加速度读数和采样频率进行了调整, 交叉验证的结果都达到了 83% 以上, 平均达到 92.73%, 从机器学习的角度可以得出测试集和训练集数据分布接近一致的结论, 从而验证了“我们的方法能使得异质加速度传感器数据映射到一个具有一致分布的数据空间”。

3.5 小结

本章调研了目前中国智能手机市场的品牌份额及手机操作系统的份额情况, 了解到手机市场品牌繁多但操作系统相对集中。基于智能终端(尤其是手机)的行为识别模型必须考虑其宿主设备之间的差异性。我们通过大量实验多方面验证了加速度传感器读数的量纲和采样频率对模型的迁移性有着较大的影响。我们提出了一种基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法, 能将不同参数传感器的输出值映射到分布一致的数据空间, 基于此数据空间就能构建一个普适于设备的行为识别模型。

第四章 基于特征迁移的非特定位置行为识别方法

4.1 问题的提出

传统的穿戴式行为识别一般都将加速度传感器固定在人体的特定位置，甚至传感器各个轴的朝向都固定不变，但基于移动设备的行为识别则不应有这种限制，因为不同的用户会有不同的携带设备的方式（比如手中、口袋中、包中等），甚至同一个用户在不同的时间也会有不同的携带方式。然而，这样的要求会增加行为识别的复杂性。

1 设备朝向不同对数据的影响

部署在同一位置的多个设备，如果其朝向不同，在用户做某一个动作时，加速度传感器各个轴上的加速度分量也会不同。如图 4.1 所示。

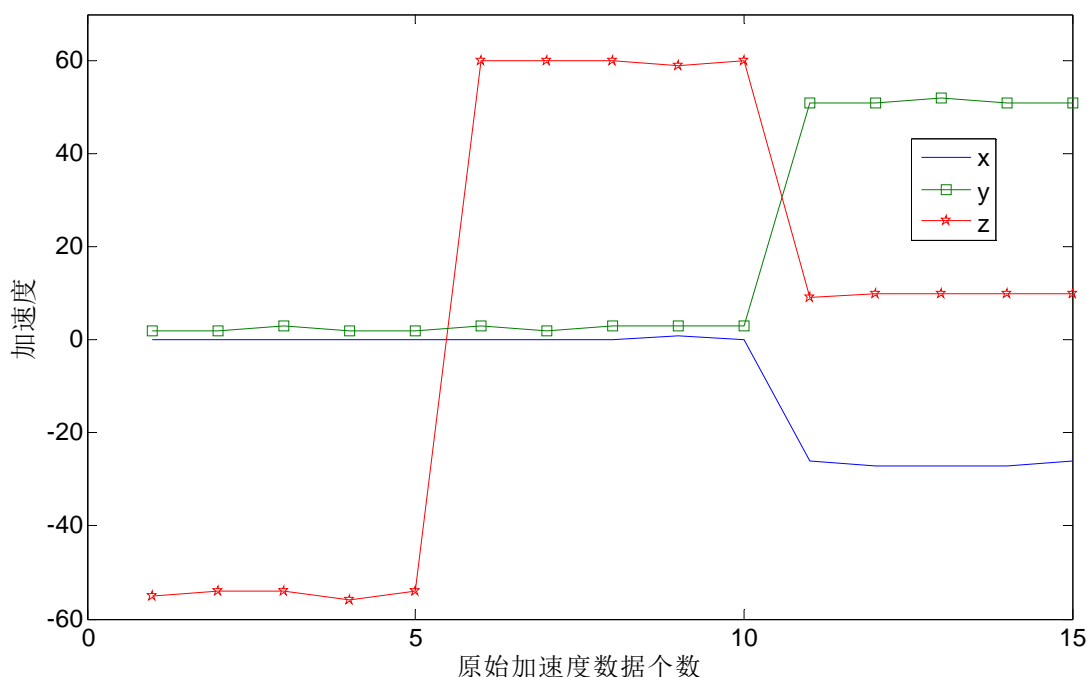


图 4.1 不同朝向 N95 手机静止时加速度传感器数据特性

图 4.1 表示的是 N95 手机静止时的加速度传感器数据，区间[1,5]内的数据是在手机屏幕朝向与重力方向相反时采集到的，区间[6,10]内的数据是在手机屏幕朝向和重力方向相同时采集到的，可以看到，由于手机所在平面相对平整，x、y 轴的读数都几乎为 0；在区间[11,15]的数据是手机倾斜放置时采集的，此时重力在各个轴上都有分量产生。

2 设备携带位置不同对数据的影响

当多个设备被同时部署在人体不同的位置，在用户做同样行为动作时，它们会被施以不同的力，从而有不同的加速度；因此不同位置的加速度也是不相同的。如图 4.2 所

示。

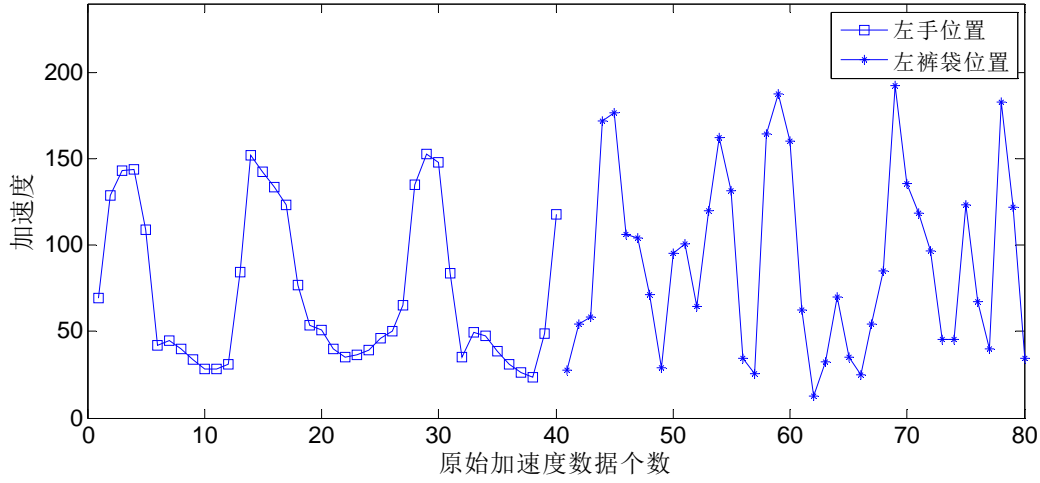


图 4.2 不同位置 N95 手机跑步时加速度传感器数据特性

图 4.2 绘出了不同位置 N95 手机感知到的跑步时的加速度传感器数据特性，区间 [1,40] 对应的是左手位置，区间 [41,80] 对应的是左裤袋位置，它们在幅值上具有明显的不同，因为在不同位置感知到的剧烈程度是不同的。

为了处理设备朝向和位置不同导致的问题，本文提出一种基于特征迁移的非特定位置行为识别模型。首先，在离线阶段，采集不同位置的多种行为的数据并进行标定，为了消除设备朝向的影响，基于合成加速度值来提取特征。其次，在提取了大量特征之后，采用主成分分析（Principal Component Analysis, PCA）技术消除噪音特征，合成鲁棒的低维特征。在此，基于 PCA 后的特征基于极速学习机算法（Extreme Learning Machine, ELM）训练行为识别模型。最后，在在线阶段采用 ELM 模型识别出新位置高置信度的样本，并利用 ELM 极速学习的特点在线增量更新旧的识别模型，得到非特定位置的新的识别模型。

本章后续部分内容安排如下：在 4.2 节中，描述相关的工作。4.3 节介绍非特定位置行为识别模型的概貌和相关技术知识，给出模型的自适应调整和在线增量学习过程。实验结果将在 4.4 节中展示。最后，在 4.5 节中对本章进行了总结。

4.2 相关研究工作

1 在解决设备朝向影响方面的相关研究工作

关于设备朝向的问题，近几年有相关研究工作如下：Reddy 等人在识别交通模式 [Reddy,2008] 时建议采用从合成加速度序列中提取特征，以消除加速度的方向信息对模型识别能力的影响。Mizell 在文章 [Mizell,2003] 指出，从一段合理时间区间内的加速度平均值可得到重力相关分量的一个合理估计，从而可采用加速度分解的方式将动态加速度投影到垂直和水平方向上，从而将加速度传感器坐标系映射到重力坐标系，然后从垂

直和水平方向分别提取特征。Yang 在文章[Yang,2009]中采用文章[Mizell,2003]中的方法提取特征,识别用户的行为(包括坐、站、散步、跑步、开车、骑自行车),他们实验结果表明基于加速度分解的方法要略好于基于加速度合成的方法。但是 Wang 等人[Wang,2010]指出在识别诸如骑自行车、乘公交车、驾车、静止、乘地铁、散步等 6 种特定交通模式时,加速度合成的方法要好于加速度分解的方法;他们还指出在基于加速度分解的方法中,重力加速度分量的错误估计会降低模型的识别精度。

Kunze 等人在文章[Kunze,2009]中提出了一种基于 PCA 技术判断用户水平运动方向的方法。他们认为当判断出加速度传感器各个轴的方差接近于 0,且合成加速度接近 9.81 米/秒^2 时,可以判断此时加速度传感器主要受垂直方向重力的影响,由此可以找到重力平面与加速度传感器坐标系之间的夹角关系。根据水平面与重力平面垂直的特性,可以计算出水平平面与加速度传感器坐标系之间的夹角关系。将一段时间内的加速度数据投影到水平平面,用 PCA 技术获取加速度变化最大的方向就是用户前进的方向。结合磁力计,可以找到用户前进的方向(相对于地球坐标系)。该方法从另一方面验证了 PCA 可以用来处理加速度传感器数据。

2 在解决设备携带位置不同导致影响方面的相关研究工作

关于设备位置的问题,相关的研究工作有:Kunze[Kunze,2005]提出了一种位置识别的方法,当用户行走的时候,如果在手腕、头部、裤子口袋、上衣口袋配置有加速度传感器,则这些传感器的读数所反映的行走模式是不同的,根据这些不同的模式可以反推出传感器的位置信息。然而,对于基于移动设备的行为识别来说,这种方法是不可行的,因为它需要事先定义好固定的位置。Kunze[Kunze,2008]还指出,由旋转导致的加速度是位置敏感的,加速度计和陀螺仪的配合使用有助于减少这种敏感性。时岳等[时岳,2013]提出了一种基于旋转模式识别设备随身佩戴位置的方法。该方法使用加速计与陀螺仪两种传感器进行数据融合,计算出旋转半径、角速度幅度和重力加速度并提取特征。在分类时,使用随机森林作为分类器,并与使用支持向量机的方案进行了对比。为了检验其效果,在包含 3 种佩戴位置和 13 种用户活动种类的公开数据集上进行了实验。实验结果显示,该方法能够达到平均 95.39% 的交叉验证准确率。

为了同时解决设备位置和朝向不同的问题,Sun[Sun,2010]做了一些尝试。为解决设备朝向的问题,他们在合成加速度序列上提取朝向不敏感的特征。为解决设备位置的问题,他们采用了扩大训练数据集的方法,尽可能涵盖常用位置的样本,为此他们采集了所有口袋位置(具体是上衣左右下摆处的两个口袋,裤子左右大腿前侧的两个口袋和左右臀部的两个口袋)的 7 种行为(静止、散步、跑步、上楼梯、下楼梯、骑自行车、自驾车)样本,并构造了支持向量机(Support Vector Machine, SVM)用以分类。然而,他们仅仅假设用户将手机放置在腰部周围的 6 个口袋里,而在现实中用户可能将手机放在更多的位置,例如手中、皮带上的手机袋中等。我们认为,既然在训练阶段无法覆盖所有可能的位置,不如从动态模型的角度出发,根据使用过程中采集到的数据,不断地

对模型进行更新，使模型逐渐适合用户所有经常放置手机的位置。

3 在模型自适应调整方面的研究工作

有些文献采用了识别模型自适应调整的方法，比如 Lai[Lai,2010]采用身体姿态分析流程图来判断哪种行为正在发生。为了使得识别算法能够符合用户的个人习惯，即具有个性化，他们应用减法聚类算法（Subtractive Clustering Method, SCM）计算每个姿势习惯性的倾斜角度的中心位置，以此设定姿态判定的阈值。Zhao[Zhao,2011]等人提出了一种基于迁移学习的方法，该方法集成了决策树和 K-Means 算法来进行模型的自适应调整。然而，上述论文都只是考虑到不同用户之间的自适应调整，没有考虑到不同位置之间如何自适应调整的问题。

4.3 基于特征迁移的非特定位置行为识别模型

我们提出了一种基于特征迁移的非特定位置行为识别模型，能够在用户使用的过程中，预测用户当前的行为，并给出该行为的置信度。然后根据用户的高置信度样本自动地更新模型，尤其是当用户的样本中包含来自新位置的样本时，模型就会随着自动更新的过程逐渐具备了对新位置样本的识别能力。该模型由两步构成，如图 4.3 所示，图中虚线部分表示第一步，实线部分表示第二步。

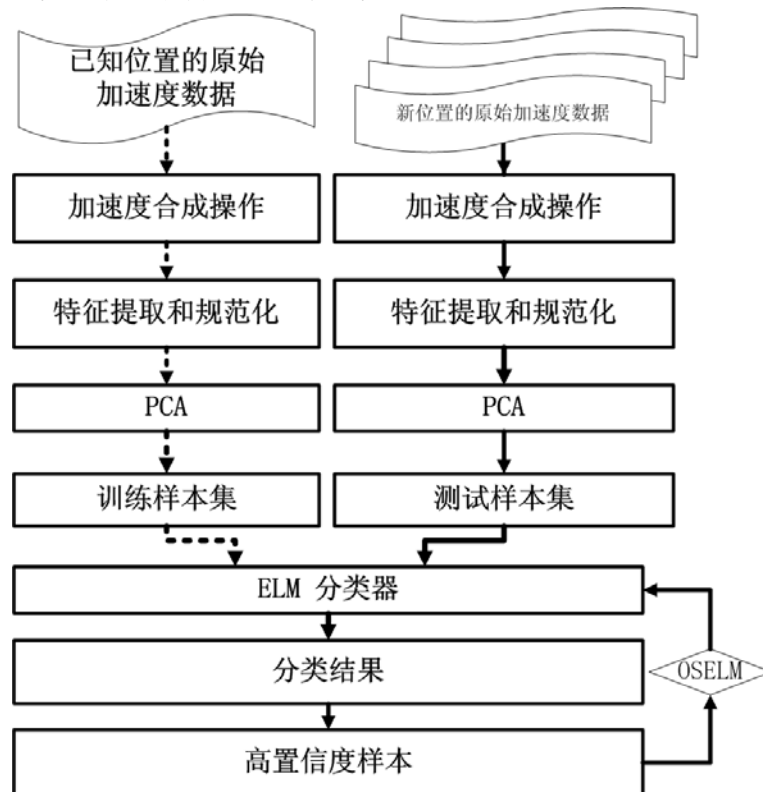


图 4.3 基于 ELM 的非特定位置行为识别模型

第一步：离线训练 ELM 模型。首先对采集并标定后的特定位置原始三轴加速度数据进行合成运算，以消除朝向的影响，在此基础上提取时域和频域的特征，然后应用

PCA 进行降维以保存有用和鲁棒的特征。结合降维后的特征向量和对应的标号信息，训练 ELM 分类器，作为行为识别的模型。

第二步：在线行为识别和自适应更新模型。当离线阶段训练的 ELM 模型被应用到诸如智能手机的移动终端时，应用系统能够按照离线阶段的方法获取用户的未标定样本。应用 ELM 模型对未标定样本进行分类，能够得到初步的分类结果。根据分类结果，能够估算每个样本被正确分类的置信度。然后选取那些置信度超过某个阈值的样本，采取增量学习的方法对模型进行更新。由于用以更新的样本可能来自新的位置，这样原来基于来自特定位置数据的行为识别模型就会自动调整以适应新位置的样本，成为非特定位置的行为识别模型。

4.3.1 合成加速度数据

加速度传感器能够检测出电流的变换，并将其转换成模拟输出电压，该电压值正比于加速度的大小。对于三轴加速度传感器而言，其输出电压可在相互垂直的三个方向分解，形成三个轴上的读数，分别为 a_x ， a_y ， a_z 。数据采集程序能够获取到的加速度值就是上述三个。从这三个值，我们可以根据公式（4.1）反推出合成加速度的值：

$$a = \sqrt{a_x^2 + a_y^2 + a_z^2} \quad (4.1)$$

从公式（4.1）可以看出， a 是合成加速度的大小，是一个不小于 0 的值，根据物理学知识，其是加速度传感器三个轴上加速度的合成操作结果。本章和其他文献[Mizell,2003; Sun,2010]中的处理方法一样，仅关注 a 的值的大小，而不关注其方向的信息。所以基于合成加速度大小的行为识别模型是与智能终端的朝向无关的。

4.3.2 基于 PCA 的特征降维处理

基于合成加速度序列，采取滑动窗口的方法提取特征，在每个滑动窗口中，共包含 64 个数据，提取的特征为：（1）时域的最大值、最小值、均值、标准差、众数、过零点个数、范围、信号幅值面积等 8 个特征；（2）FFT 变换后的 64 个系数；（3）功率谱密度 PSD 的 8 个特征。总计 80 个特征。

由第 4.1 节的示例可知，不同位置的数据具有不同的特点，则对于某个位置各种行为的样本具有较好区分能力的特征对于其他位置未必有效，即某些特征是位置相关的。要构建非特定位置的行为识别模型，就要尽可能地去掉那些冗余的特征，去掉那些位置相关的特征，仅保留那些对所有位置的样本都具有最大区分能力的特征。主成分分析技术 PCA[Wang,2009]具有这样的特性。

下面简要介绍一下 PCA 变换技术。

用 X 表示规范化的样本特征矩阵，其维数为 $M \times N$ ，其中 M 是样本数， N 是每个样本的特征数。PCA 变换可表示为 $Y_{M \times K} = X_{M \times N} \cdot B_{N \times K}$ ，其中 $K < N$ 。 $B = \{b_1, b_2, \dots, b_K\}$ 是一系列线性无关和正交的向量的集合，可以如下计算得到[Wang,2009]：

Step 1: 计算原始特征矩阵的协方差矩阵, 如下:

$$S = \frac{1}{M} (X - \mu)^T (X - \mu)$$

其中 μ 是所有样本的平均值。

Step 2: 计算协方差矩阵 S 的特征值和特征向量。

Step 3: 根据特征值的大小对特征向量排序, 选取其中最大的 K 个特征值 $\lambda_1, \lambda_2, \dots, \lambda_k$ 对应的特征向量构造变换矩阵 B 。 K 值的选择可以由下面的公式决定:

$$\frac{\sum_{i=1}^K \lambda_i}{\sum_{i=1}^M \lambda_i} \geq 1 - \eta$$

其中 η 表示能量损失值。

4.3.3 极速学习机分类器

在提取特征和经过 PCA 变换后, 我们可以选取特定的学习方法完成特征到行为类别之间的映射, 比较流行的学习方法有基于实例[Mitchell,2003]的学习方法(如最近邻等)、基于统计的学习方法(如支持向量机等)[Cristianini,2004]等。最近邻方法由于所有的计算都发生在在线分类的过程中, 当实例库中的实例数量较大时, 分类新实例的开销就会比较大。支持向量机在训练时需要通过遍历的方式选择最优的参数以得到最优的效果, 时间开销比较大, 而对于我们的问题, 需要在便携式设备上在线预测用户行为样本的类标号, 判断该样本属于所预测类别的置信度, 然后用高置信度的行为样本去在线更新已有模型, 想要在便携式设备这样的资源受限设备上完成这些运算, 要求学习方法需要具有高效性。

极速学习机 ELM 是最近出现的一种机器学习方法, 它具有计算速度快、泛化能力强的特点[Huang,2011; Huang,2006; Huang,2004; Huang,2004a], 吸引了大量研究人员的注意, 近年来在增量 ELM、并行 ELM [He,2013]、半监督 ELM 等方面得到了发展。虽然 ELM 既可以处理回归问题又可以处理分类问题, 我们在此仅关注其分类的能力。

如图 4.4 所示, ELM 是一个具有单隐结点层的前向神经网络, 输入层 (Input Layer) 包含所有输入神经元, 每个输入神经元对应样本的一个特征; 隐结点层 (Hidden Layer) 包含所有隐层神经元; 输出层 (Output Layer) 包含所有输出神经元, 每个输出神经元对应多分类问题的一个类别。对于含有 N 个样本的训练集合 $\mathbf{X} = \{(x_i, t_i) | x_i \in R^n, t_i \in R^m, i = 1, 2, \dots, N\}$, 其中样本 $x_i = \langle f_1^1, f_1^2, \dots, f_1^n \rangle$ 是一个行向量, 其长度 n 对应着特征的个数; t_i 是一个行向量, 其长度 m 对应着类别的个数, 当 x_i 的类标号为 m' 时, 向量 t_i 的第 m' 位为 1, 其他各位全部为 0。

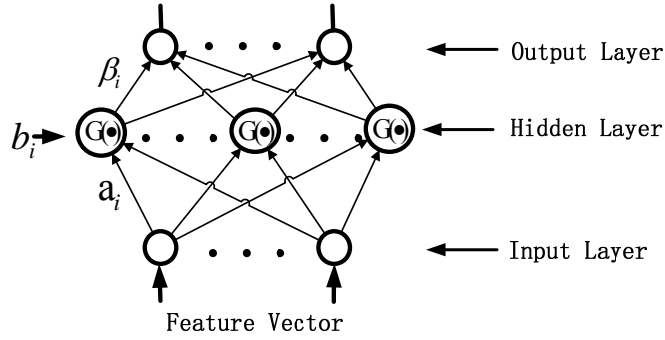


图 4.4 ELM 算法的网络结构

用单隐层前向神经网络对该问题建模，选定输入结点个数为 n ，隐层结点个数为 \tilde{N} ，输出结点个数为 m ，激励函数为 $G(a, b, x)$ 。对每一个样本 x_i ，有：

$$[G(a_1, b_1, x_i), \dots, G(a_{\tilde{N}}, b_{\tilde{N}}, x_i)] \begin{bmatrix} \beta_{11} & \beta_{21} & \dots & \beta_{m1} \\ \beta_{12} & \beta_{22} & \dots & \beta_{m2} \\ \vdots & \vdots & \dots & \vdots \\ \beta_{1\tilde{N}} & \beta_{2\tilde{N}} & \dots & \beta_{m\tilde{N}} \end{bmatrix} = [o_{i1}, o_{i2}, \dots, o_{im}] \quad (4.2)$$

其中：

$a_i = [a_{i1}, a_{i2}, \dots, a_{in}]$ 表示第 i 个隐结点到所有输入结点之间权值组成的向量；

b_i 表示第 i 个隐结点的阈值；

a_i 和 b_i 的值随机生成，其值位于 $[-1, 1]$ 之间。

$$\beta_i = \begin{bmatrix} \beta_{i1} \\ \beta_{i2} \\ \vdots \\ \beta_{i\tilde{N}} \end{bmatrix}$$

表示第 i 个输出结点到每一个隐结点之间的权值组成的向量；

$o_i = [o_{i1}, o_{i2}, \dots, o_{im}]$ 表示 m 个输出结点的值组成的向量。

则对于训练集合 $\mathbf{X} = \{(x_i, t_i) | x_i \in \mathbb{R}^n, t_i \in \mathbb{R}^m, i = 1, 2, \dots, N\}$ 中所有 N 个样本，每个样本均可以写成式 (4.2) 的形式。可以把所有 N 个等式写成如式 (4.3) 的矩阵形式：

$$\begin{bmatrix} G(a_1, b_1, x_1) & \dots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ G(a_1, b_1, x_2) & \dots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_2) \\ \vdots & \vdots & \vdots \\ G(a_1, b_1, x_N) & \dots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_N) \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \dots & \beta_{m1} \\ \beta_{12} & \beta_{22} & \dots & \beta_{m2} \\ \vdots & \vdots & \dots & \vdots \\ \beta_{1\tilde{N}} & \beta_{2\tilde{N}} & \dots & \beta_{m\tilde{N}} \end{bmatrix} = \begin{bmatrix} o_{11} & o_{12} & \dots & o_{1m} \\ o_{21} & o_{22} & \dots & o_{2m} \\ \vdots & \vdots & \dots & \vdots \\ o_{N1} & o_{N2} & \dots & o_{Nm} \end{bmatrix} \quad (4.3)$$

对于训练集合 \mathbf{X} ，其对应的单隐层前向网络的优化目标是使得预测误差趋近于 0，即：

$$\sum_{i=1}^N \|o_i - t_i\| = 0$$

即存在 β_i 、 a_i 和 b_i ，使得：

$$\begin{bmatrix} G(a_1, b_1, x_1) & \dots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ G(a_1, b_1, x_2) & \dots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_2) \\ \vdots & \vdots & \vdots \\ G(a_1, b_1, x_N) & \dots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_N) \end{bmatrix} \begin{bmatrix} \beta_{11} & \beta_{21} & \dots & \beta_{m1} \\ \beta_{12} & \beta_{22} & \dots & \beta_{m2} \\ \vdots & \vdots & \dots & \vdots \\ \beta_{1\tilde{N}} & \beta_{2\tilde{N}} & \dots & \beta_{m\tilde{N}} \end{bmatrix} = \begin{bmatrix} t_{11} & t_{12} & \dots & t_{1m} \\ t_{21} & t_{22} & \dots & t_{2m} \\ \vdots & \vdots & \dots & \vdots \\ t_{N1} & t_{N2} & \dots & t_{Nm} \end{bmatrix}$$

上式可以采用矩阵的形式，写得更加紧凑一些，即：

$$H\beta = T$$

其中：

$$H = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_N) \end{bmatrix}_{N \times \tilde{N}}$$

$$\beta = [\beta_1 \quad \cdots \quad \beta_m]_{\tilde{N} \times m} = \begin{bmatrix} \beta_{1,1} & \cdots & \beta_{m,1} \\ \vdots & \ddots & \vdots \\ \beta_{1,\tilde{N}} & \cdots & \beta_{m,\tilde{N}} \end{bmatrix}_{\tilde{N} \times m}$$

$$T = \begin{bmatrix} t_{11} & t_{12} & \cdots & t_{1m} \\ t_{21} & t_{22} & \cdots & t_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ t_{N1} & t_{N2} & \cdots & t_{Nm} \end{bmatrix}_{N \times m}$$

按照传统的方法，训练一个单隐层前向网络，我们需要找到特定的 β 、 a_i 和 b_i ($i=1,2,3,\dots,\tilde{N}$)，使得：

$$\|H(\hat{a}_1, \dots, \hat{a}_{\tilde{N}}, \hat{b}_1, \dots, \hat{b}_{\tilde{N}})\hat{\beta} - T\| = \min_{\beta, a_i, b_i} \|H(a_1, \dots, a_{\tilde{N}}, b_1, \dots, b_{\tilde{N}})\beta - T\|$$

Huang 在文献[Huang,2006]中指出， a_i 和 b_i 的值无需调整，只需在算法开始时随机生成即可，那么问题就转换为：

$$\|H(a_1, \dots, a_{\tilde{N}}, b_1, \dots, b_{\tilde{N}})\hat{\beta} - T\| = \min_{\beta} \|H(a_1, \dots, a_{\tilde{N}}, b_1, \dots, b_{\tilde{N}})\beta - T\|$$

则该问题的最小范数二乘解为：

$$\hat{\beta} = H^+ T$$

其中 H^+ 是矩阵 H 的 Moor-Penrose 伪逆。

当用 ELM 分类器分类新样本 x 时，ELM 输出结点的值可以计算如下：

$$TY_{1 \times m} = [G(a_1, b_1, x), \dots, G(a_{\tilde{N}}, b_{\tilde{N}}, x)]_{1 \times \tilde{N}} \cdot \beta_{\tilde{N} \times m}$$

其中， m 是输出结点的个数，也是分类问题中类的个数。相应的， TY 就是一个含有 m 个值的行向量。

那么 ELM 如何根据输出向量 TY 决定样本 x 的类标号呢？根据 ELM 的训练过程可知， TY 的哪个分量距离 1 最近，那么该分量相应的下标就是样本 x 的类标号，即：

$$j = \operatorname{argmin}_{1 \leq j \leq m} |1 - TY_j|$$

即 j 就是样本 x 的类标号。

综上所述，在本章中，训练阶段的输入为经过 PCA 变换后的特征及其对应的类标号，每个样本对应的类标号用一个 m (m 是模型能够预测的类别的个数) 维的向量表示，当样本的类标号为 m' 时，标号向量的第 m' 位为 1，其他各位全部为 0，训练的结果是 ELM 预测模型。ELM 预测模型的输入是一个新样本经过 PCA 变换的特征，输出是对该样本预测的结果，为一个 m 维的向量 TY 。根据 TY 可以计算出样本的类标号。

4.3.4 样本的自标定

进一步,还可以从向量 TY 中计算出样本 x 隶属于某个行为 j 的置信度, $confidence$ 。 TY 的每个分量与 1 都有一定的距离, 该距离的大小可看做样本隶属于对应类别的一个度量, 距离越小则越接近该类, 直观地我们也可以用该距离的倒数作为样本隶属于对应类别的度量, 倒数值越大则越接近该类。因此 $confidence$ 的值可以通过下面的步骤得到:

- 1) 计算 TY 每个分量与 1 的距离。

$$D = |TY - 1|$$

- 2) 计算 D 中每个分量的倒数, 记做:

$$D_Inverse = \frac{1}{D}$$

- 3) 计算 $D_Inverse$ 中最大值在 $D_Inverse$ 中所占的比重, 以此作为样本 x 的 $confidence$ 。

$$confidence = \frac{\argmax D_Inverse}{\sum D_Inverse}$$

根据 $confidence$ 可以进一步衡量 x , 如果 $confidence$ 的值小于某一个阈值, η , 则 x 可被认为是和模型中任何一个类别都不相似的新类样本或者是含噪音严重的数据。相反的, 如果 $confidence$ 大于阈值 η , 则 x 作为 j 类的新样本保留下来, 用以对模型的增量更新。

4.3.5 模型的在线增量更新

离线阶段训练的 ELM 模型能最大地反映训练数据隐含的知识, 对训练数据具有很好的区分能力。对我们的问题而言, 当便携式设备放置在与训练数据来自位置不同的地方时, 其产生的行为数据的分布不能保证与训练数据的分布完全一致, 一般会存在一定的偏差。为了使 ELM 模型能更准确地预测新位置的样本, 需要用新位置的高置信度样本去更新已有模型, 该更新过程应该是无需用户参与的, 应该是自动地进行的, 所以更新过程应该是一个在线更新过程。而便携式设备是资源受限的, 不可能采用大量的数据去重新训练模型, 因此增量式学习是有必要的。因此我们需要的是一个能够自动选择高置信度样本且能够在线增量学习的 ELM 模型。

下面我们简要介绍一下在线增量极速学习机 (Online Sequential ELM, OSELM)。

前面描述的 ELM 算法以整个训练数据集作为输入, 一次性得到训练模型。为了得到在线增量 ELM 算法, Huang 等人给出了 OSELM 算法, 描述如下:

第一步: 设初始数据集为: $\mathbf{x}_0 = \{(x_i, t_i)\}_{i=1}^{N_0}$, 随机选择的输入权重为 a_i , 偏置值为 b_i , 选取的激励函数为 $G(a, b, x)$, 隐结点的个数为 \tilde{N} 。那么按照 ELM 算法, 隐结点层的输出矩阵 H_0 可计算为:

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_{N_0}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0}) \end{bmatrix}_{N_0 \times \tilde{N}}$$

相应的，输出权值为： $\beta_0 = K_0^{-1} H_0^T T_0$ ，其中 $K_0 = H_0^T H_0$ ； $T_0 = [t_1, t_2, \dots, t_{N_0}]^T$ 。

第二步，设新来的下一块数据为： $\mathbf{x}_1 = \{(x_i, t_i)\}_{i=N_0+1}^{N_0+N_1}$ ，那么采用与第一步相同的 a_i, b_i, G, \tilde{N} ，同样可以计算出 H_1 ：

$$H_1 = \begin{bmatrix} G(a_1, b_1, x_{N_0+1}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0+1}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_{N_0+N_1}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0+N_1}) \end{bmatrix}_{N_1 \times \tilde{N}}$$

进而可以计算 K_1 ：

$$\begin{aligned} K_1 &= \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \\ &= \begin{bmatrix} H_0^T & H_1^T \end{bmatrix} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \\ &= K_0 + H_1^T H_1 \end{aligned}$$

这样，更新后的输出权值可计算如下：

$$\begin{aligned} \beta_1 &= K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \\ &= K_1^{-1} (K_1 \beta_0 - H_1^T H_1 \beta_0 + H_1^T T_1) \\ &= \beta_0 + K_1^{-1} H_1^T (T_1 - H_1 \beta_0) \end{aligned}$$

可以看出， β_1 是关于 β_0, K_1, H_1 和 T_1 的函数，而跟先前的数据集 \mathbf{x}_0 无关，这意味着，先前数据使用完毕即可弃之不用，不必继续保留，可以节省存储空间。

综上所述，本章首先基于已标定的经过 PCA 变换的特定位置的行为样本离线训练 ELM 模型 ELM_A，在在线预测阶段，用 ELM_A 对新样本 x 进行预测，得到 x 的预测类标号 j 及 x 隶属于类 j 的置信度 confidence，当 confidence 的值大于某一个阈值 η （本文中指 50%）时，将其放入高置信度样本集 HConf，随后基于 OSELM 算法用 HConf 集合对 ELM_A 进行在线增量更新，当 HConf 集合中包含新位置的样本时，特定位置行为识别模型 ELM_A 就具有了识别新位置样本的能力，我们称其经过自适应地调整成为了非特定位置的行为识别模型 ELM_B。

4.4 实验及结果分析

本部分，基于表 4.1 所示的数据集我们验证了本章提出的基于特征迁移的非特定位置行为识别方法。共验证了四方面的内容：1）比较了 ELM 和 NN、SVM 算法的性能，突出了 ELM 算法具有训练和预测速度快的优点；2）分别比较了 ELM、NN、SVM 在未对数据进行 PCA 变换时的交叉验证实验结果，验证了不同位置行为数据分布具有较大差异的特点；3）分别比较了 ELM、NN、SVM 在对数据进行 PCA 变换后的交叉验证

实验结果,验证了 PCA 变换对发现不同位置公共特征的作用;4)比较了自适应调整后模型对新旧位置数据的识别能力,验证了我们方法的有效性。实验的硬件平台是一台普通的 PC 机, CPU 主频为 2.6GHz, 软件平台是 MATLAB2009a。ELM 源代码下载自 Huang 教授的个人主页⁷。

本部分实验基于第二章采集的部分数据。由于我们仅考虑同一个传感器在不同位置时的情况,因此选取的数据都来自 iphone 手机。数据集中包含 12 位采集人员的数据,每人将手机分别放在左手、右手、左腰部、右腰部、左裤袋、右裤袋等 6 个位置(在实验中我们把左手、右手合称为手部,左裤袋、右裤袋合称为裤袋,左腰部、右腰部合称为腰部),在每个位置都要完成静止、散步、跑步、上楼梯、下楼梯等 5 种行为,每种行为的总的持续时间不低于 5 分钟。表 2.5 中详细列出了最终每个人、每个位置、每种动作的样本数,此处给出一个关于 iphone 数据集的总的统计结果,如表 4.1 所示。

表 4.1 基于 iphone 的行为数据统计信息

	右裤袋	右腰部	右手	左手	左腰部	左裤袋
静止	3879	4340	3696	3486	4431	3918
散步	3687	3668	3471	3719	3559	3731
跑步	3351	2671	3458	3354	3389	3353
上楼梯	1492	1505	1533	1556	1632	1670
下楼梯	1552	1412	1425	1467	1519	1522
总计	13961	13596	13583	13582	14530	14194

4.4.1 分类器性能比较

为了测试 ELM 算法对于行为识别问题的性能,本部分比较了它和另外两个流行的分类器的效果,这两个分类器分别是支持向量机(Support Vector Machine, SVM)和最近邻算法(Nearest Neighbor, NN)。基于表 4.1 中描述的样本,选取右裤袋位置样本,分别建立不同的分类器,并采用了 10 折交叉验证技术对它们的性能进行测试。ELM 的隐结点数被设置为 100,激励函数采用的是 sigmoid。为得到最优的 SVM 模型,在选取 SVM 的参数 c (c 是惩罚系数,代表算法对误差的容忍程度,值越高表明越不能容忍出现误差)和 g (径向基函数作为 kernel 时的参数 γ ,隐含地决定了数据映射到新的特征空间后的分布)时,我们选择了遍历的方法,最终选择的 c 值是 32, g 值是 0.5。NN 算法不需要参数的设定,当有新样本到来时,用其和样本库中的所有样本进行比较,选取距离最近的样本的类标号作为新样本的类标号。在表 4.2 中列出了各个分类器的训练时间(以秒为单位)、测试时间(以秒为单位)和识别精度。

⁷ Source codes and some references of ELM can be found at www.ntu.edu.sg/home/egbhuang

表 4.2 分类器性能比较

	ELM	SVM	NN
训练时间（秒）	21.26	1850.77	\
测试时间（秒）	0.84	50.99	506.39
识别精度	88.33%	86.01%	90.38%

如表 4.2 所示，ELM、SVM 和 NN 的识别精度相差不大。然而，ELM 的训练时间和测试时间要远小于 SVM。对于 13961 个样本，采用 10 折交叉验证的方法，ELM 总的训练时间为 21.26 秒，而 SVM 为 1850.77 秒，这也说明了 ELM 具有训练速度快的特点。ELM 总的测试时间为 0.84 秒，也明显低于 SVM 和 NN 的测试时间（NN 是一种消极的学习算法，没有训练阶段，也就没有对应的训练时间）。总之，在识别精度相似的情况下，ELM 具有相对快得多的训练和测试速度，这对于行为识别这样需要实时判断的问题有重要的意义。

4.4.2 未进行 PCA 变换时的交叉位置行为识别实验

为了全面获取一个位置的识别模型对其他位置数据的识别结果，我们进行了交叉位置模型识别实验。对于每一个位置，训练一个分类器，并对所有位置进行测试。表 4.3 列出了 ELM 行为识别模型在交叉验证情形下的混淆矩阵。表的第一列表示训练位置，第一行表示测试位置。从表中我们可以看出，某个特定位置的识别模型对于来自该位置的样本具有较高的分类效果，但对于其他位置的样本识别效果较差。性能下降最大的情况是手部的模型对来自裤袋的样本数据的分类，降幅大约为 27%；性能下降最小的情况是裤袋的模型应用到来自腰部的样本数据，降幅大约为 5%。

表 4.3 ELM 不同位置模型交叉验证实验结果

	裤袋	腰部	手部
裤袋	87.48%	82.61%	66.52%
腰部	76.53%	88.93%	65.49%
手部	62.45%	65.31%	89.84%

同理，SVM 模型进行交叉验证的混淆矩阵列在表 4.4 中。与表 4.3 对比来看，对于对角线上的实验结果，即用特定位置的模型分类来自该位置的样本，SVM 和 ELM 取得近似的分类精度，但对于来自其他位置的样本，SVM 却比 ELM 逊色得多，这在一定程度上反映了 SVM 容易发生过拟合，而 ELM 的泛化能力较强。

表 4.4 SVM 不同位置模型交叉验证实验结果

	裤袋	上衣口袋	手部
裤袋	90.84%	76.89%	61.14%
腰部	55.77%	92.84%	56.28%
手部	50.43%	54.45%	94.40%

表 4.5 是 NN 分类器的交叉验证效果，由于 NN 是一个基于实例的消极（lazy）学习方法，当训练集和测试集相同时，其测试精度达到了 95% 以上。对比表 4.3，可以看出对于交叉位置的测试，NN 的表现也不如 ELM。

表 4.5 NN 不同位置模型交叉验证实验结果

	裤袋	上衣口袋	手部
裤袋	96.12%	70.32%	56.84%
腰部	66.09%	95.98%	57.96%
手部	55.29%	63.15%	96.87%

4.4.3 进行 PCA 变换后的交叉位置行为识别实验

正如前文第 4.4.1 节所述，已经提取出 80 个特征。在这些特征里，一些是对分类结果有用的，而一些可能是不利于分类的噪音。为了消除噪音以得到鲁棒性好的特征，引入了 PCA 进行特征降维。

图 4.5 中列出了 PCA 变换后的主成份个数和保存的能量之间的关系。可以看出，当能量损失率设置在 0.05 时，30 个主成份就可以满足要求。因此，在以下的实验中，选取的主成份都是 30 个。

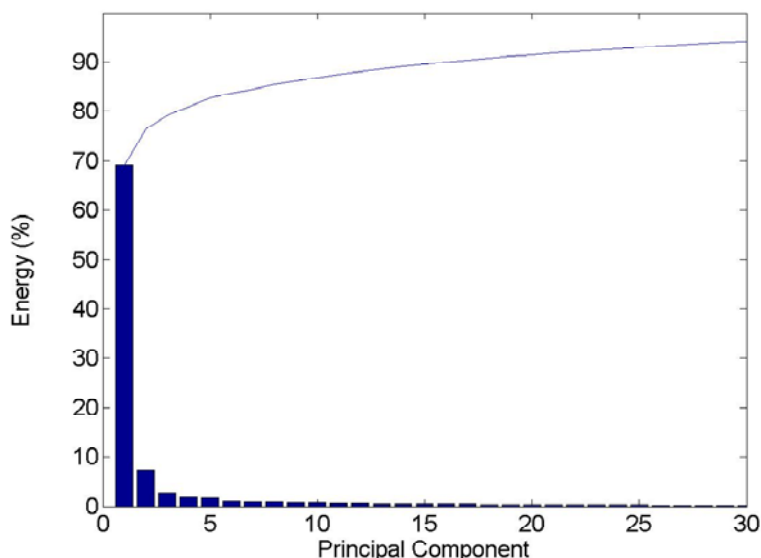


图 4.5 PCA 主成分与能量之间的关系

为了直观地观察 PCA 降维后对模型性能的提升情况，我们基于降维后的数据重现了第 4.4.2 节的实验。对比 PCA 降维前后 ELM 分类器的性能，如表 4.3 和 4.6 所示，降维后交叉位置测试的精度提高了 1% 到 2%。这种现象表明，PCA 能消除特征空间中的噪音数据，得到鲁棒性好的特征。

表 4.6 PCA 降维后 ELM 不同位置模型交叉验证实验结果

	裤袋	腰部	手中
裤袋	87.09%	82.91%	68.77%
腰部	79.06%	88.85%	66.67%
手中	63.66%	66.72%	89.73%

表 4.7 PCA 降维后 SVM 不同位置模型交叉验证实验结果

	裤袋	上衣口袋	手部
裤袋	86.35%	77.26%	61.94%
上衣口袋	56.98%	87.30%	56.48%
手部	50.89%	55.53%	89.56%

对比表 4.4 和表 4.7, 可发现应用 PCA 降维后, 基于 SVM 的分类器模型对于交叉位置的识别精度也略有提高。尽管如此, 对比表 4.6 和 4.7, 仍然可以看出, ELM 无论在已知位置还是在未知位置的识别精度都要好于 SVM。当用 NN 作为分类器时, 从表 4.5 和 4.8 也可以发现相同的情况。从上述实验可以看出, PCA 降维能够去除噪音提高特征的鲁棒性, ELM 较之于其他两种分类器, 表现最好, 具有最好的泛化能力。

表 4.8 PCA 降维后 NN 不同位置模型交叉验证实验结果

	裤袋	上衣口袋	手部
裤袋	90.22%	78.90%	67.72%
上衣口袋	75.30%	89.00%	67.30%
手部	58.52%	64.29%	91.17%

从第 4.4.1 节到第 4.4.3 节的实验可以看出, ELM 具有更快的训练速度, 更强的泛化能力。因此, 在后续的实验中, 我们将以 ELM 作为我们的分类器。

4.4.4 自适应调整后模型对新旧位置数据的识别能力实验

本节的实验旨在测试 ELM 模型对新位置的自适应调整能力。用 A、B、C 代表手部、腰部、裤袋等三个位置, 对应的数据集分别用 $Data_A$, $Data_B$ 和 $Data_C$ 表示。每一个数据集被随机地平分为两部分, 分别为 $Data_{A1}$ 和 $Data_{A2}$, $Data_{B1}$ 和 $Data_{B2}$, $Data_{C1}$ 和 $Data_{C2}$ 。 $Data_{C1}$ 中的数据标号在模型自适应调整过程中不可见, $Data_{C2}$ 中的数据及其标号用于测试模型的识别能力。

不失一般性, 我们首先假设 A、B 是已知位置, C 是未知位置。用 $Train_{AB}$ 表示 $Data_{A1} + Data_{B1}$ 。 $Train_{AB}$ 用于训练一个初始的 ELM 分类器, $Data_{C1}$ 用于调整初始模型以适应新位置的样本。 $Test_{AB} = Data_{A2} + Data_{B2}$, 用于测试新旧模型对已知位置样本的分类能力, $Data_{C2}$ 用于测试它们对新位置样本的分类能力。当用初始模型对新样本 $Data_{C1}$ 分类时, 如果样本的分类置信度 η 大于 0.5, 则将样本添加到一个新的集合 $HConf_{C1}$ 中。

随后, 使用 $HConf_{C1}$ 作为增量样本集合, 应用 OSELM 算法增量更新已有 ELM 模型得到新的 ELM 模型。

初始模型和新模型对于已知位置样本的区分能力如表 4.9 所示, 可以看出, 新旧模型对已知位置样本具有基本相同的区分能力, 自适应调整后, 模型识别能力略有下降, 是因为位置数的增加导致了问题更加复杂, 模型识别能力降低也在情理之中。初始模型和新模型对于新位置的区分能力如表 4.10 所示, 经过自适应调整, 新模型的识别精度大约提高了 19%, 这是因为原模型中没有新位置的样本信息, 识别能力自然不高, 自适应调整后模型中具有了新位置样本的知识, 故对新位置样本的识别能力有大幅度提高。

表 4.9 新旧模型对于已知位置 A&B 的识别精度

	自适应调整前	自适应调整后
模型	Train _{AB} 训练的 ELM	ELM 模型+Data _{C1} +OSELM
测试数据	Test _{AB}	Test _{AB}
识别精度	86.71%	84.73%

表 4.10 新旧模型对于新位置 C 的识别精度

	自适应调整前	自适应调整后
模型	Train _{AB} 训练的 ELM	ELM 模型+Data _{C1} +OSELM
测试数据	Data _{C2}	Data _{C2}
识别精度	67.58%	86.76%

当 B、C 作为已知位置, 而 A 作为新位置时, 相应的实验结果如表 4.11 和 4.12 所示。自适应调整之后, 模型的精度大约提高了 5%。

表 4.11 新旧模型对于已知位置 B&C 的识别精度

	自适应调整前	自适应调整后
模型	Train _{BC} 训练的 ELM	ELM 模型+Data _{A1} +OSELM
测试数据	Test _{BC}	Test _{BC}
识别精度	85.62%	84.72%

表 4.12 新旧模型对于新位置 A 的识别精度

	自适应调整前	自适应调整后
模型	Train _{BC} 训练的 ELM	ELM 模型+Data _{A1} +OSELM
测试数据	Data _{A2}	Data _{A2}
识别精度	80.31%	85.63%

接下来是 A、C 作为已知位置而 B 作为新位置的实验结果, 如表 4.13 和 4.14 所示, 自适应调整后比自适应调整前的识别精度大约提高了 7%。

表 4.13 新旧模型对于已知位置 A&C 的识别精度

	自适应调整前	自适应调整后
模型	Train _{AC} 训练的 ELM	ELM 模型+Data _{B1} +OSELM
测试数据	Test _{AC}	Test _{AC}
识别精度	87.04%	86.49%

表 4.14 新旧模型对于新位置 B 的识别精度

	自适应调整前	自适应调整后
模型	Train _{AC} 训练的 ELM	ELM 模型+Data _{B1} +OSELM
测试数据	Data _{B2}	Data _{B2}
识别精度	75.62%	82.81%

从表 4.10、4.12 和 4.14 可以看出,当已有模型对未知位置样本的识别精度达到 67% 以上时,用新位置的样本对模型进行进行自适应调整后,对新位置样本的识别精度都达到了 82% 以上。从表 4.6 可以看出,当用某一个位置的数据训练的模型识别该位置的数据时,识别精度在 88% 左右,而自适应调整后对新位置数据的识别精度已经接近了该精度。表 4.10、4.12 和 4.14 的对比结果也表明了,在已有模型的识别精度达到一定程度时,其原始识别精度越低,经过自适应调整后提高的空间就越大,从侧面也反映了我们方法的有效性。

4.5 小结

本章基于内置有加速度传感器的智能手机研究用户的行为识别。现存一些相似工作假定内置有加速度传感器的设备位置固定不变,本文打破了该限制,力图识别用户非特定位置的动作行为。

我们提出了一种基于特征迁移的非特定位置行为识别方法来处理智能手机放置位置和朝向多变的问题。在此模型中,具有快速学习能力的 ELM 算法被用于离线训练特定几个位置的模型,通过在线增量学习的方法逐渐调整其适应新的位置。实验结果表明,即使没有新位置任何的先验知识,模型的自适应调整也能够显著提高模型的识别精度。

第五章 基于参数迁移的非特定人行为识别方法

5.1 问题的提出

在诸如移动健康看护之类的应用程序中，基于移动终端的行为识别有着巨大的发展潜力。在欧美国家，为了帮助人们保持健康并提高身体素质，医生往往给他的病人开出运动处方，比如散步、跑步、爬楼梯等。而运动处方最重要的衡量指标就是运动时间。为了使人们能够知道自己是否严格按照医生的运动处方锻炼，他必须能够容易地获知各种运动的累计时间。这对于人们提高自己的运动质量尤为重要。而且，大量的民众拥有健康的体魄也能够减少整个政府的医疗预算。基于此，我们设计了一款基于移动终端的行为报告系统，其能够识别一系列日常行为动作，并统计每种动作的累计时间。我们通过嵌入在移动终端的三轴加速度传感器用以分类诸如静止、散步、跑步和上下楼梯之类的动作行为，同时记录下各行为准确的起止时间。

但是，我们在研究中发现，从某一个或者某一些特定人的数据训练得到的分类模型应用于另外一个人的行为数据时，并非总能产生很好的分类效果。探究其原因，在于每个人都有自己的个性化行为，即通常所说的“习惯”，如图 5.1 所示。这样对于相同参数的加速度传感器，放置到不同用户的同一个位置（比如口袋），对于相同的概念行为（比如散步），感知到的加速度数据的分布是不同的。这种分布的不同使得建立普适于用户的行为识别模型具有一定的挑战性。我们称这种问题为非特定人行为识别问题。

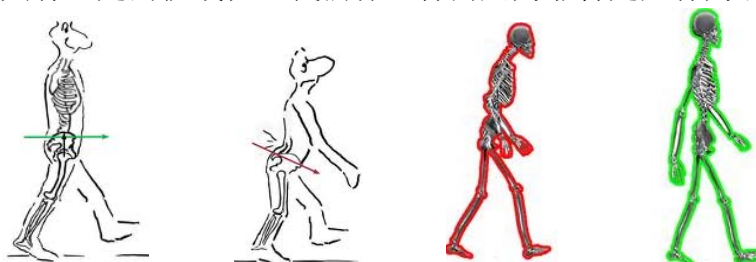


图 5.1 用户的个性化行为

为了解决非特定人行为识别问题，本章提出了一种基于迁移学习的嵌入决策树 TransEMDT(**T**ransfer learning **E**Mbedded **D**ecision **T**ree)算法，该算法集成了决策树分类器和 K-Means 聚类算法用以形成个性化的行为识别模型。经过在真实世界数据集上的测试，实验结果表明，TransEMDT 算法能够取得良好的效果。

本文提出的方法新颖，简单快捷，能够嵌入到计算、存储、能量受限的移动终端中使用。同时我们也测试了系统的有效性，探索如何在保持系统识别精度的情况下尽可能地减少训练集的大小和迭代的次数。

5.2 相关研究工作

医疗保健市场上已有少量设备能够估算运动量的多少，比如文献[Consolvo,2006]中提到的计步器（Step Counter）以用户走过的“步数”量化用户的行为，非常适合测量散步和跑步的运动量。再比如博能公司[Nachman,2010]的产品通过心率来量化用户消耗的卡路里数量，非常适合有氧运动项目。但是，这些方法只是提供给用户一些最终的统计数字，而不能给出每种运动的细节。并且，当使用该产品时，用户必须佩戴专门的心率带，对于他们来说，显得有些不便。再者，这些设备都没有考虑到用户的个性化行为问题。

现存的一些文献对基于加速度传感器的行为识别给予了关注。在文献[Bao,2004; Ravi,2005; Ward,2011]中，他们使用多个设备来区分不同的行为，获得了较高的识别精度。Chen 在文章[Chen,2010]中使用一个智能手机来区分 6 种行为，目的是寻求状态的变化点，用以决定打开或者关闭 WiFi 的时机，实现低功耗定位的目的。上述两种方法之所以能够获得较高的识别准确度，是因为它们的训练数据和测试数据来自同一批数据，这些数据遵从相同的概率分布模型。然而这些文献均没有考虑到非特定人行为识别问题。

也有相关文献关注到了非特定人行为识别问题。比如，在文献[Albinali,2010]中，作者使用多个传感器识别行为类别。当使用特定对象的数据作为训练数据和测试数据时，能够得到 49% 的分类精度。然而当用多个人的数据进行训练，而用另外一个人的数据测试时，只能得到 26% 的测试精度。然而对如何解决这个问题，他们并没有提出相应的方法。

其他领域的研究人员在研究与以上类似的数据域相同而标号空间不同的识别问题时，采用了迁移学习技术。比如，Zheng 等人[Zheng,2009]通过 Web 搜索技术，在源域和目标域的行为之间找到一个相似度函数，基于此函数他们把源域的标定数据引入到目标域中，完成了知识的迁移。就我们的问题而言，源域和目标域具有相同的标号空间，并且我们更关注于寻找能运行于移动终端之上的具有更高计算效率和更低功耗的的算法。

在文献[Stikic,2008]中，作者采用协作学习方法将未标定数据引入到训练集合中，能够得到较好的模型。然而他们的算法要同时使用多个分类器。这些分类器建立在不同的特征集合上，采用投票的方法决定某个未标定样本是否出现在训练集合中。因此这种方法比较复杂，功耗较高，不太适用于基于智能终端的应用程序。

Lee[Lee,2007]提出了迁移决策树算法（Transfer in Decision Trees, TDT）。TDT 假设源域的属性恰巧是目标域的一个子集。它首先在源域学习一个分类模型，必要时将其迁移到目标域。Torrey 等人在文献[Torrey,2005]中提出一种增强学习的方法，能够把源域的知识迁移到目标域中。他们需要一个真正的教师在迁移过程中进行指导。而我们的方法是无需任何手工输入的。

对于移动健康看护这样的应用，若要为广大人民群众接受，其应该具有下面的优点：

1) 其应该是便携的，不会对用户造成干扰，不会引起用户的不便，比如可以安装到用户的移动终端（尤其是手机）上面。

2) 其对新用户的行为应该具有较高的识别精度，或者具有自适应更新的能力，能随着用户的使用而逐渐地迁移为适合新用户的个性化行为识别模型。

为达到此目的，我们提出了一种基于迁移学习的嵌入决策树方法，能够基于移动终端构建应用程序，并解决非特定人行为识别问题。

5.3 一种基于迁移学习的嵌入决策树算法

为了获取非特定人的个性化行为识别模型，我们提出了一种基于迁移学习的嵌入决策树算法—TransEMDT(**T**ransfer learning **E**Mbedded **D**ecision **T**ree)。

5.3.1 算法框架

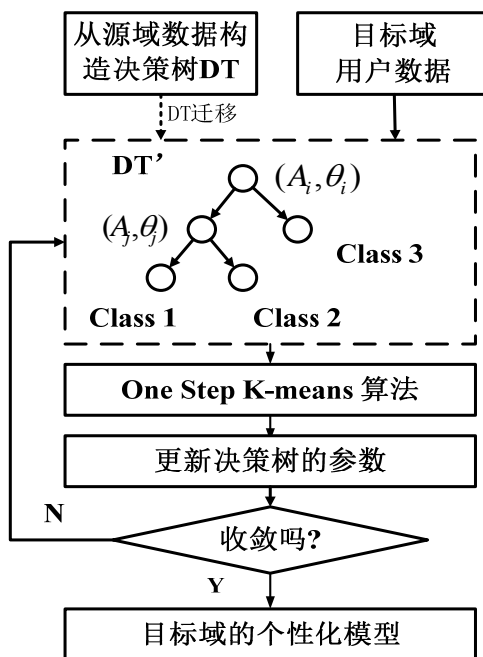


图 5.2 TransEMDT 算法框架

如图 5.2 所示, TransEMDT 模型由三层构成: DT 层(二叉决策树层, Decision Tree Layer), One-Step K-Means 聚类算法层和二叉决策树更新层。在构建目标域用户的个性化模型过程中,正是通过调整二叉决策树的非终端叶子结点的阈值实现的。该算法描述如下:第一步,从源域用户的样本训练得到决策树模型 DT, DT 是一棵二叉决策树,每个非终端结点对应一个分割属性 A 及其值 θ_A 。第二步,用 DT 区分目标域用户的样本,并用分类结果作为目标域用户数据的初始类中心,以便在第三步执行一个 One-Step K-Means 算法。在我们的应用中,DT 的每一个叶子结点对应一个类。第三步,执行 One-Step K-Means 算法,将上述目标域用户所有的样本在 DT 的叶子结点间重新分配。

第四步，用每个叶子结点上置信度最高的样本更新 DT，我们假设距离每个类中心最近的 K' 个样本的置信度最高（距离类中心的距离与置信度成反比）。第五步，重复执行第 2 步到第 4 步，直至算法收敛。这样，我们就得到了目标域用户的个性化模型。

5.3.2 二叉决策树的构建及迁移

1 二叉决策树的构建过程

在离线阶段，从训练数据集构建二叉决策树 DT，其构建过程如下：

第一步：计算所有属性的信息增益，选择信息增益最大的属性作为根结点；

第二步：由根结点属性及其取值，构建左右分支；

第三步：采用递归的方法，对各分支的子集重复第一步和第二步，直到子集中数据具有相同的类标号或者数据个数少于某个阈值。

2 属性选择的度量规则

如何选择属性及其对应的阈值进行分裂，是构建二叉决策树的关键所在。本文属性选择的度量规则涉及到熵、信息增益两个概念，我们在下面给予简单介绍。

对于已标定样本集合 S ，按照信息论的思想，对 S 进行分类所需要的期望信息为：

$$Entropy(S) = \sum_{i=1}^m -p_i \log_2 p_i$$

其中 p_i 是 S 中属于类别 i 的样本的比例。 $Entropy(S)$ 又称为集合 S 的熵。

由于要构建二叉决策树，对于集合 S ，当选定分类属性 A 及其阈值 T_i 将集合分成两部分后，要将两部分都准确分类还需要的期望信息为：

$$Entropy_A(S) = \frac{|S_{v < T_i}|}{|S|} Entropy(S_{v < T_i}) + \frac{|S_{v \geq T_i}|}{|S|} Entropy(S_{v \geq T_i})$$

如果把熵理解为当前分类的混乱程度，那么信息增益就容易理解了，信息增益可以理解为在用属性 A 及其阈值 T_i 对 S 分类后混乱减少的程度，即：

$$\begin{aligned} Gain(S, A, T_i) &= Entropy(S) - Entropy_A(S) \\ &= Entropy(S) - \frac{|S_{v < T_i}|}{|S|} Entropy(S_{v < T_i}) - \frac{|S_{v \geq T_i}|}{|S|} Entropy(S_{v \geq T_i}) \end{aligned}$$

使得属性 A 信息增益最大的 T_i ，就是属性 A 对应的阈值：

$$[A, T, Gain(S, A, T)] = \operatorname{argmax}_{T_i} Gain(S, A, T_i) \quad (5.1)$$

那么对于当前结点面临的属性集合为 B 时，对 B 中的每个属性都要按照公式 (5.1) 计算其增益，最终选择具有最大增益的属性及其阈值作为当前结点的分裂属性和阈值。

如此自上而下进行递归处理，就可以得到一棵二叉决策树。

综上所述，DT 是一棵二叉决策树，每一个非终端结点都有自己的决策条件 (A_i, θ_i) ， A_i 表示决策属性， θ_i 表示决策阈值，当 A_i 的值小于阈值时，样本将会取道左子树，否则取道右子树。

3 二叉决策树的迁移

基于训练集合构建的二叉决策树行为识别模型，如果不加修改的应用到新用户的移动终端且用于分类用户的行为数据，由于用户行为差异性，必然导致识别精度不高。我们提出的基于参数迁移的非特定人行为识别方法，在保持二叉决策树结构不变的情况下，可以用新用户的高置信度样本在线更新决策树非终端结点的阈值，使之越来越适应新用户的行为数据。

5.3.3 新用户高置信度样本的筛选过程

一方面，离线阶段训练的基于二叉决策树的行为识别模型对新用户样本的分类能力要高于随机猜想的精确度。另一方面，K-Means 聚类算法能够将样本聚成若干个类，但是其聚类质量、耗费时间和初始的 K 个类中心点有关，一旦选择不当，算法就会收敛得很慢。在我们的解决方案中，结合了决策树和 K-Means 算法的优点，以二叉决策树的分类结果作为 K-Means 算法的初始聚类条件，实现了较快速的聚类。在 K-Means 聚类过程中，通常是采用样本与聚类中心之间的距离衡量样本是否属于某一个类，样本距离的计算方法在较大程度上影响着聚类的结果。本文结合二叉决策树的特点提出了一种基于各类最小区分属性的样本间距离计算方法，在此基础上实现更为准确的聚类算法。

1 基于最小区分属性集的距离计算方法

从图 5.2 可以看出，未标定样本经过二叉决策树模型 DT 分类后将被分配到相应的叶子结点，即被冠以相应的类标号。同时，从二叉决策树的构建过程可知，出现在从叶子结点到根结点上的属性就是能最大区分该类和其他类的最小属性集，用这些属性衡量同一类的样本会更准确。于是我们假设每类数据都具有自己的属性集合，当计算某个样本和该类的类中心间的距离时，只需相应的属性值参与计算，而不需要全部的属性值都参与计算。

对于样本 $x=\{x_1, x_2, \dots\}$ ， A_j 是 x 的第 j 个属性， x_j 是相应的值。对于第 i 个叶子结点，我们为其设一个属性向量 ω_i ，表示对应的属性是否属于第 i 个叶子结点的最小属性集，若属于则用 1 表示，否则用 0 表示，其第 j 个元素由下述公式表示：

$$\omega_{ij} = \begin{cases} 1, & A_j \in P_i \\ 0, & A_j \notin P_i \end{cases} \quad (5.2)$$

其中， P_i 是第 j 个叶子结点到根结点的路径。

我们考察每个叶子结点上的样本，计算其类中心。那么样本 x 到第 i 个类中心 μ_i 的距离可如下计算：

$$D(x, \mu_i, \omega_i) = \frac{|\omega_i \cdot (x - \mu_i)|^2}{\sum_{j=1}^{|\omega_i|} \omega_{ij}} \quad (5.3)$$

其中， $\omega_i \cdot x = (\omega_{i1}x_1, \omega_{i2}x_2, \dots)$ 。为了消除不同属性个数的影响，在式 (5.3) 中用属性个数对距离进行了规范化。

那么，样本 x 的类标号的下标就是 j :

$$j = \operatorname{argmin}_{j \in [1, m]} D(x, \mu_j, \omega_j)$$

其中 m 表示叶子结点的个数， μ_j 表示第 j 个叶子结点所有样本的均值。

Label_j 就是决策树第 j 个叶子结点上所有样本的标号。正如下一节将要描述的那样，每当在叶子结点间重新分配样本时，就需要重新计算样本到各个叶子结点上样本均值的距离。

2 基于 One-Step K-Means 算法的高置信度样本选择方法

下面介绍一下基于 One-Step K-Means 算法的高置信度样本选择方法：

第一步，用二叉决策树模型区分目标用户的未标定数据， $D_{\text{tar}} = \{x_{\text{tar}}^i\}_{i=1}^{N_2}$ ，将其划分成 m 个类， $V_{\text{tar}} = \{(\text{Label}_j, \omega_j, \mu_j, V_j)\}_{j=1}^m$ ，其中 m 是叶子结点的个数， Label_j 第 j 个叶子结点中所有样本的标号， ω_j 可由公式 (5.2) 计算得到， μ_j 是第 j 个叶子结点上样本的中心点， V_j 用来存储第 j 个叶子结点上的所有样本。

第二步，为 K-Means 算法准备初始类中心。在一个外部循环内部，我们希望该算法仅被执行一次，所有的样本将会寻找离它最近的叶子结点，正如在 K-Means 算法中描述的那样。我们首先为每个叶子结点 j 计算其初始类中心：

$$\mu_j = \frac{\sum_{i=1}^{|V_j|} x_i^j}{|V_j|}$$

其中 x_i^j 是类 j 的第 i 个样本。然后，基于每个样本到类中心的距离，我们把其重新分配到叶子结点。样本新标号的计算标准如下：

$$j = \operatorname{argmin}_{j \in [1, m]} D(x, \mu_j, \omega_j)$$

第三步，样本 x 被重新分配到第 j 个叶子结点并存储在 V_j 中。这样就完成了 One-Step K-Means 算法。

经过以上 One-Step K-Means 算法过程，我们可以得到标定的数据集。可以想象，经过 One-Step K-Means 算法，有可能出现样本被错误分类的情况，我们从每个叶子结点选取距离类中心最近的 K' 个样本，作为高置信度样本，用其更新决策树模型。

5.3.4 二叉决策树模型的更新过程

当新用户的样本被重新分配到新的叶子结点且高置信度样本被选取后，我们可以用一个自下而上的过程更新所有非终端结点的阈值。对于 DT 的一个非终端结点，其分类属性为 A_i ，我们收集其左分支上所有的样本对应属性 A_i 的值并存储在集合 LSamples 中，收集其右分支上所有样本对应属性 A_i 的值并存储在集合 RSamples 中。既然属性 A_i 的值 θ_i 必须要把这两个集合分开，因此该值必须位于 LSamples 的最大值和 RSamples 的最小值之间，可以如下计算：

$$\theta_i = \frac{\text{Max}(\text{LSample}) + \text{Min}(\text{RSample})}{2}$$

当所有的非终端结点均被更新完毕，就可以得到一个更新后的模型 DT'。该过程一直重复直到算法收敛或者达到设定的循环次数。

作为总结，我们在算法 1 中对 TransEMDT 进行了详细描述。第一步，从标定的样本生成 DT 模型。从第 2 步到第 4 步，将 DT 模型迁移到新用户，并准备 K-Means 算法的初始条件。第 5 步，K-Means 算法被执行一次。第 6 步，调整 DT 模型。第 7 步，第 2 步到第 6 步重复执行直至生成个性化模型。

算法 1: TransEMDT 算法

Input: 源域 $D_{src} = \{(x_{src}^{(i)}, y_{src}^{(i)})\}_{i=1}^{N_1}$ ，其中 $y_{src}^{(i)}$ 是样本 $x_{src}^{(i)}$ 的类标号。源域中共有 N_1 个样本且所有样本来自一个或者少数几个人。目标域 $D_{tar} = \{x_{tar}^{(i)}\}_{i=1}^{N_2}$ 中所有样本来自于与源域中不同的人。Thd 是一个阈值，决定着算法迭代的次数。 K' 是每次迭代中选取的置信度最高的样本的个数。Times 是迭代次数的上界，是决定迭代次数的另一个条件。

Output: 新用户的个性化模型

Begin:

1. 从标定的样本集合学习得到 DT 模型
2. 对 DT 的每一个叶子结点，找到与其相关的属性，形成向量 ω_i ，表示每个属性对于该类的重要程度。该向量被用来计算叶子结点上所有样本的中心。

$t=0$;

Do

3. 对目标域 D_{tar} 中的每个样本分类，得到结果: $V_{tar} = \{(Label_j, \omega_j, \mu_j, V_j)\}_{j=1}^m$ ，其中 $V_j = \{V_j^{(i)}\}_{i=1}^{|V_j|}$ 。
4. 设置 One-Step K-Means 算法的初始条件，即第 j 个叶子结点数据的均值:

$$\mu_j = \frac{\sum_{i=1}^{|V_j|} x_i^j}{|V_j|}$$

5. 对每个样本 $x_{tar}^{(i)}$ ，通过计算距离 $D(x_{tar}^{(i)}, \mu_j, \omega_j)$ ，寻找距离其最近的叶子结点，并据此改变其类标号。
6. 通过一个自下而上的过程更新每个非终端结点:

对每个叶子结点，选择其中 K' 个离其中心最近的叶子结点，用以更新非终端结点的阈值。

For DT 中的每个非终端结点对应的属性 A_i

令 LSamples 表示其左分支上的所有样本对应于属性的 A_i 值，RSamples 表示其右分支上所有样本对应于属性的 A_i 值，按如下更新其阈值:

$$\theta_i = \frac{\text{Max}(\text{LSample}) + \text{Min}(\text{RSample})}{2}$$

End For

$t=t+1$;

While $\sum_{j=1}^m \sum_{i=1}^{|V_j|} D(V_j^{(i)}, \mu_j, \omega_j) > \text{Thd}$ or $t < \text{Times}$

7. 输出个性化模型

End Begin

5.4 实验及结果分析

本部分，基于表 2.6 所示的数据集我们验证了本章提出的 TransEMDT 方法，共验证五方面的内容：1) 验证决策树模型对个人行为数据的区分能力；2) 验证不同用户行为数据分布之间的差异性；3) 对比应用 TransEMDT 算法前后模型的识别能力，验证方法的有效性；4) 观察决策树模型识别能力与训练样本数量之间的关系；5) 观察个性化模型与新用户样本数量之间的关系。实验 4 和实验 5 的目的是为了寻找最优的训练模型的样本的数量，不盲目地扩大训练数据集的大小，达到减少移动终端能量消耗的目的。实验的硬件平台是一台普通的 PC 机，CPU 主频为 2.6GHz，软件平台是 MATLAB2009a。

5.4.1 数据准备

本实验中，我们的数据采集设备是 Nokia N95 智能手机，利用其内置的加速度传感器采集数据，最终得到一个行为数据库。数据库中共有 12 个参与者的 5 种行为数据，手机放置的位置是右裤袋。采用滑动窗口的方法提取特征数据。虽然 N95 内置加速度传感器的采样频率最高可以达到 100Hz 或者 400Hz，但由于考虑到能耗等原因，Nokia 提供的 API 函数只能以 32Hz 的频率采集数据。我们选择的窗口长度是 2 秒，重叠窗口是 1 秒，这样每个窗口中可以包含完整的动作。在早期的文献如[Bao,2004]中已经展示了这种滑动窗口方法提取特征的效果。在每个窗口中，最终提取的特征有：(1) 时域的最大值、最小值、均值、标准差、众数、过零点个数、范围、信号幅值面积等 8 个特征；(2) FFT 变换后的 64 个系数；(3) 功率谱密度 PSD 的 8 个特征。总计 80 个特征。

每种行为的样本数统计在表 2.6 中，此处不再重复列出。

5.4.2 决策树模型对个人数据的区分能力

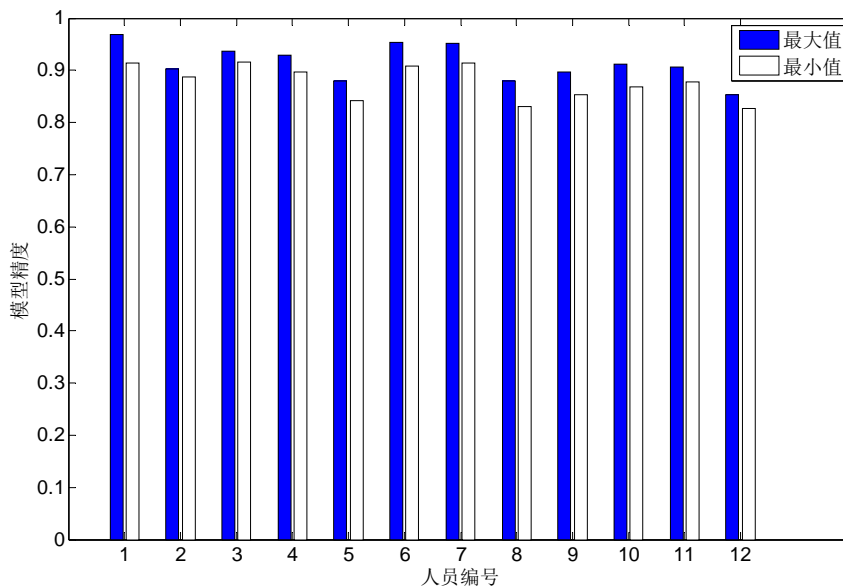


图 5.3 决策树模型对个人数据的区分能力

对于 12 个人在右裤袋位置的数据，我们验证决策树模型对数据的区分能力。对于每个人的数据，随机地平分为两部分，一部分用以训练模型，另一部分用以测试模型，该过程重复 10 次。实验结果如图 5.3 所示。

从图 5.3 可以看出，对于每个人的数据，10 次实验结果的最好和最差识别精度相近；对于不同的数据采集人员，每个人的行为识别模型精度都能达到 80% 以上，说明对于每个人的静止、散步、跑步、上楼梯、下楼梯等 5 类行为，都是具有可分性的。

5.4.3 不同用户两两之间模型的交叉验证实验

对于 12 个人在右裤袋位置的数据，本实验验证一个人数据的决策树模型对其他人数据的区分能力。对于每两个人的数据，其中一人的数据用以训练模型，另一人的数据用以测试模型。实验结果如图 5.4 所示。

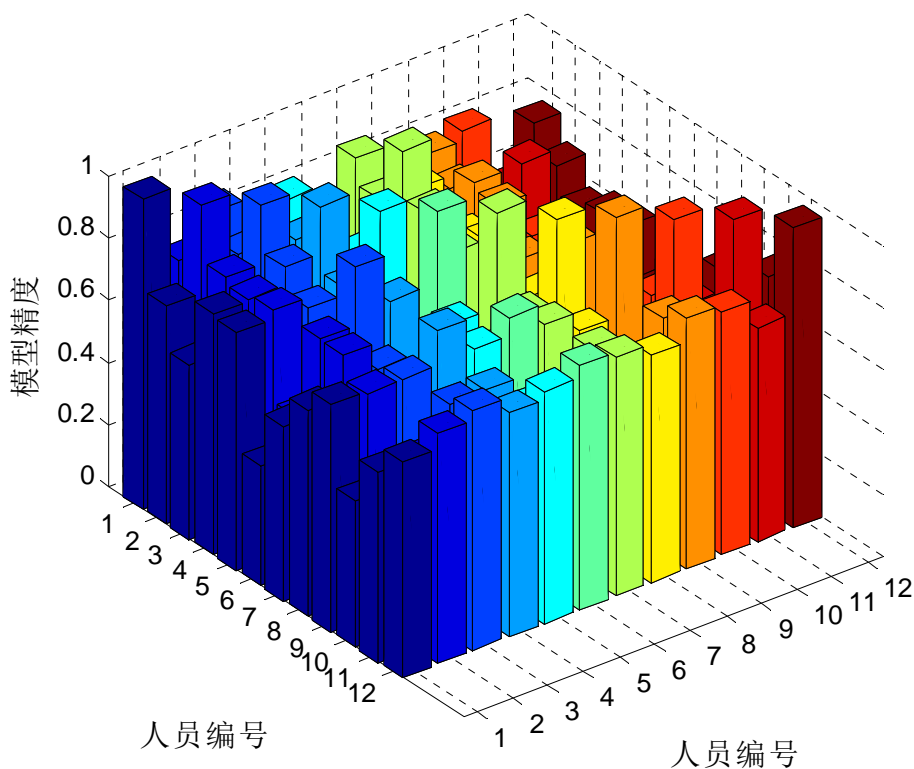


图 5.4 不同人决策树模型的交叉验证

如图 5.4 所示，用每个人的数据进行自验证的时候，都能得到很高的分类精度，平均达到了 95% 以上。然而做交叉验证的时候，分类精度却发生了很大的波动，少数几个在 40% 左右，整体上在 60%~70% 之间，换句话说，用某个人精度很高的识别模型去识别其他人的行为，并不总能得到很好的效果，这验证了人的行为具有个性化的特点。

5.4.4 算法性能

为了详细的测试 TransEMDT 算法的性能，我们在整个数据集上进行了实验，实验结果见表 5.1。首先，我们用用户 A 的样本训练决策树模型 DT。然后，用模型分类用户 B 的样本，对分类结果用 One-Step K-Means 算法聚类，在聚类结果中每类选择 K' 个高置信度样本来更新模型 DT。该步重复执行 5 次。最后，用户 B 的剩余数据用来测试更新后的模型。

表 5.1 TransEMDT 算法性能

用户对	应用算法前	应用算法后
132	63.30%±4.35%	83.26%±6.27%

从表 5.1 可以看出，对于从 12 个用户中选择 2 个用户，共有 132 种排列方式。应用 TransEMDT 算法后，行为识别算法的平均性能大约提高了 20%。

5.4.5 决策树模型和训练样本数量的关系

在实验进行过程中，第一步，我们用用户 A 的 K 个样本训练得到模型。正如前面算法描述中所言，我们称之为 DT 模型。第二步，我们用用户 A 的剩余样本测试 DT 模型，获知其分类精度。第三步，我们用用户 B 的所有样本测试模型 DT，获知另一个分类精度。用户 A 和用户 B 随机地从所有 12 个用户中选择。上述过程重复 10 次，得到其平均精度并列在图 5.5 中。为了观察 K 值对 DT 精度的影响程度，我们令其值从 15 变化到 900。

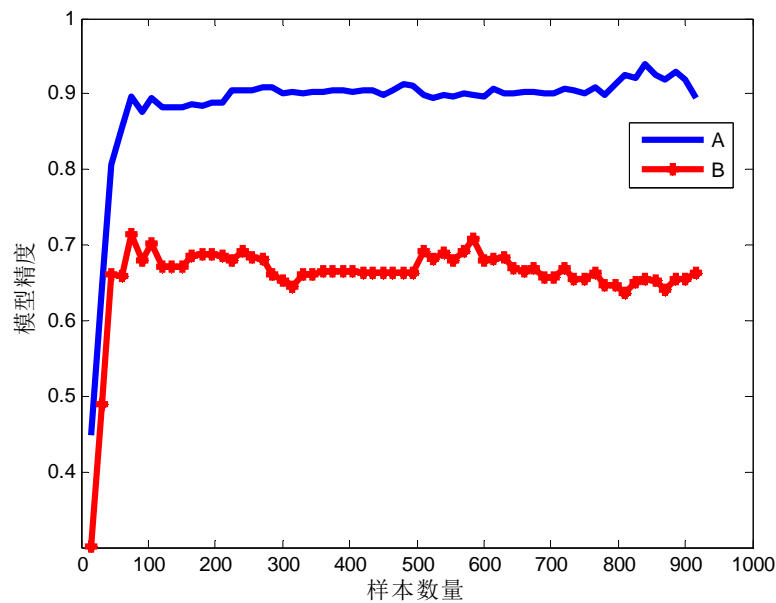


图 5.5 决策树模型性能 VS 样本数量

如图 5.5 所示，当 K 从 15 变化到 100 时，DT 模型的精度发生了显著的提升变化，

当 K 超过 100 时, DT 模型精度几乎保持不变。也就是说, 用 100 个训练数据, 即每类 20 个, 就可以得到一个精度较高的分类模型, 并且该模型对另一个用户具有中等的分类精度 (大约为 63%)。

5.4.6 个性化模型和新用户样本的关系

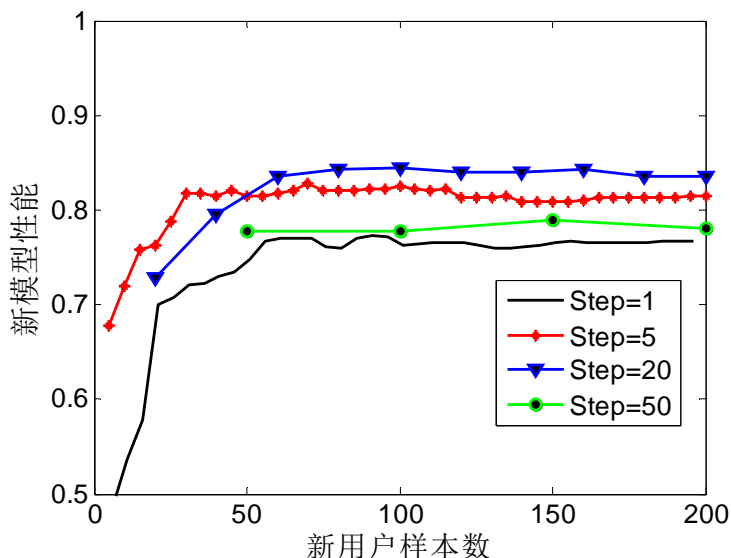


图 5.6 个性化模型精度 VS 新用户样本数量

本节, 我们研究 DT 模型的迁移能力及 TransEMDT 算法的效用。考虑到智能终端的功耗问题, 我们的算法希望能在模型的精度和计算量之间找到折中的方法, 目标是能够找到既能保持较高的分类精度又能减少能量消耗的合适的未标定样本数量和迭代次数。在实验方案中, 我们使用了变量“Step”, 表示每次迭代中使用多少未标定样本来更新 DT 模型。不同的“Step”值所对应的模型精度从图 5.6 中可以直观地看出来。可以看到:

- 1) 无论“step”值是大是小, 算法 TransEMDT 总能有效收敛。
- 2) 当“step”的值取 20 时, 能够得到一个优化的模型。这是因为, 当“Step”较小时, 每次引入到模型中的新用户的有效信息太少; 当“Step”较大时, 又容易把噪音数据引入到模型中。

5.5 小结

为解决基于移动终端的非特定人行为识别问题, 本章提出了一种基于迁移学习的嵌入决策树算法 TransEMDT, 实验表明:

- 1) 我们提出的 TransEMDT 模型能够迁移用户的行为知识, 从而提高了非特定人行为识别的精度。
- 2) 训练数据的样本数目能够影响到分类器的性能。当样本数量大于 100 (即每类样本的数量大于 20 时), 分类器的性能达到最优。因此不必引入过多的训练样本, 这对于

构建具有低功耗要求的应用程序是非常重要的。

3) 在迁移 DT 模型的过程中, 如果每次添加 20 个未标定样本, 经过 5 次迭代就可以得到一个优化的个性化模型。该发现表明, 无需迭代多次就可以构建一个个性化的模型。减少迭代次数就意味着减少能量的消耗, 这对于资源受限的设备尤为重要。

尽管我们已经研究了基于移动终端的非特定人行为识别问题, 在该领域仍然有许多开放的问题值得进一步研究:

1) 将该算法应用于非二叉决策树: 这个问题较之于二叉分类树更具有挑战性, 因为此时在计算非终端结点的阈值时再使用左右分支最大最小值的均值就未必正确。

2) 考虑更大的数据集和更复杂的场景: 后续工作中, 我们将采集 50 个用户的数据。这 50 个用户还将具有不同的身体条件, 比如性别、年龄、身高等等。我们将采集他们在现实生活场景中更为真实的数据。

第六章 基于结构迁移的非特定类别行为识别方法

6.1 问题的提出

在实际应用背景下,某些行为识别模型需要对已知类别行为的识别能力进行增量学习,比如第四章的非特定位置行为识别和第五章的非特定人行为识别。在另外一些场景下,有时需要改变模型识别行为的种类,增加其对新行为的识别能力。比如人在成长的过程中,幼儿阶段大多是走路,童年时代学会跑步、上楼、下楼等,青少年时代又逐渐学会打篮球、踢足球等,中年时代忙于工作以各种交通工具为主,老年时代身体衰弱而从事一些类似于门球等的休闲运动等。

新行为的增加必然导致用户待识别行为数据整体分布的改变,使得该分布与训练数据的分布存在差异,如此一来,已有模型不仅不能识别出新行为,还会将新行为误分为已有行为,导致误警率的增加。

本章的重点在于提出了基于结构迁移的类增量极速学习机算法。该算法将在源域训练的 ELM 模型迁移到目标域,在应用过程中随着目标域中新行为标定数据的出现而不断以类增量的方式调整模型,在调整过程中,从源域数据学习得到的 ELM 模型结构得以保持,类增量学习导致的变化体现在模型输出结点的增加和输出权值矩阵的扩展,而不会影响到输入结点、隐结点及其之间的结构,能够极大地保留源域的相关知识。通过类增量方式得到的模型不仅能够识别新类别的行为,也能够保持对已有类别行为的识别能力。

本章后续部分内容安排如下:在 6.2 节中,描述增量学习相关的工作。6.3 节介绍基于结构迁移的非特定类别行为识别方法,给出模型结构的自适应调整 and 在线增量学习过程。实验结果将在 6.4 节中展示。最后,在 6.5 节对本章进行了总结。

6.2 相关研究工作

传统的机器学习方法属于批量学习,即训练模型时需要全部数据的参与,这种特性使其不能满足现实中的某些情况,比如时间序列数据。对于时间序列数据,数据是随着时间变化的(比如人们的行为特征会随着年龄的增长而变化),这种性质决定了任何一个阶段训练的模型如果不加修正都不能比较准确地预测未来的数据,必须根据新时期新数据不断地对模型进行增量学习,才能反映当前数据的特点。

1 基于径向基函数网络 and 自适应共振理论的增量学习相关研究工作

增量学习算法得到了广泛研究,近年来在前馈神经网络中也涌现了许多研究成果:增量径向基函数网络[Fritzke,1994](Radial Basis Function networks, RBF networks)是一种新型的 RBF 网络,它能够根据错分类别的累计错误信息动态地插入新的径向基函

数中心，并根据 Neural-Gas 算法选择新中心的半径。GAP-RBF[Huang,2004b]算法也是一种具有 RBF 结点的前向神经网络，能够自动从网络中删除那些不重要的结点。虽然它也力求简化增量学习算法以提高学习速度，但由于其需要事先知道输入数据的分布情况或者采样范围，因此对于许多应用来说其学习过程依然很慢。ARTMAP 是一种基于自适应共振理论的增量学习网络，模糊 ARTMAP[Carpenter,1991] (FAM) 是它的一个变种。FAM 吸纳了模糊集理论，用以决定 ARTMAP 网络的动态变化。概率 FAM[Lim,1995] 混合了 FAM 和概率神经网络 (Probabilistic Neural Networks, PNNs) 以达到在线学习和预测的目的，它通过引入 PNN 提高了 FAM 的分类精度。上述方法的主要缺点在于学习速度慢，易受噪音数据的影响；容易导致过拟合，泛化能力不强。文献[Wang,2012]基于 PNN 和可调节的模糊聚类算法 (Adjustable Fuzzy Clustering algorithm, AFC) 实现了一种增量学习方法，该方法可以轻易地从新的训练样本中获取知识以提高模型的识别能力；还可以自由地增加对新类别的识别能力，删除对不感兴趣类别的识别能力；可以无需保存以往使用的训练数据以节省存储空间。然而，观察其实验结果，对于大约 1000 个样本，采用 5 折交叉验证的方法，其增量学习的过程需要 264.3 秒，对于智能终端之类的资源受限设备，其计算时间是无法忍受的。

2 基于极速学习机算法的增量学习相关研究工作

新加坡南洋理工大学的 Huang 教授等人提出了极速学习机算法 (Extreme Learning Machine, ELM)，ELM 属于单隐层前向神经网络中的一种，与传统的单隐层神经网络的求解算法不同，它不需要通过反复的迭代过程调整神经元的输入权重和偏置值，而是通过求解线性系统的 Moore-Penrose 伪逆和范数最小二乘解的方法计算输出权重。因此，ELM 具有计算速度快、鲁棒性强、泛化能力强的特点，但是其只能批量学习。在此基础上，文献[Liang,2006]提出了在线增量极速学习机算法 OSELM (Online Sequential ELM)，该方法能够以增量的方式处理数据，只有那些新出现的单个或者成块数据需要参与学习，用新出现样本学习到的输出矩阵 $\Delta\beta$ 去增量更新已有的输出矩阵 β ，这样得到的新的输出矩阵 β' 不仅保持了已有数据的知识，又从新数据中学习到了新的知识，并且已经参与过学习的数据无需再次参与训练，可以节省存储空间。然而，OSELM 只能处理模型中已有类别的样本，不能处理新出现类别的样本。

我们基于 OSELM 算法，并对其进行改进，提出了能同时学习已有类别和新类别知识的增量学习算法。

6.3 基于结构迁移的非特定类别行为识别方法

6.3.1 算法框架

为了使识别模型既能保持对已有类别行为的识别能力，又能逐步增加对新类别行为的识别能力，我们提出了一种基于结构迁移的非特定类别行为识别框架，如图 6.1 所示。

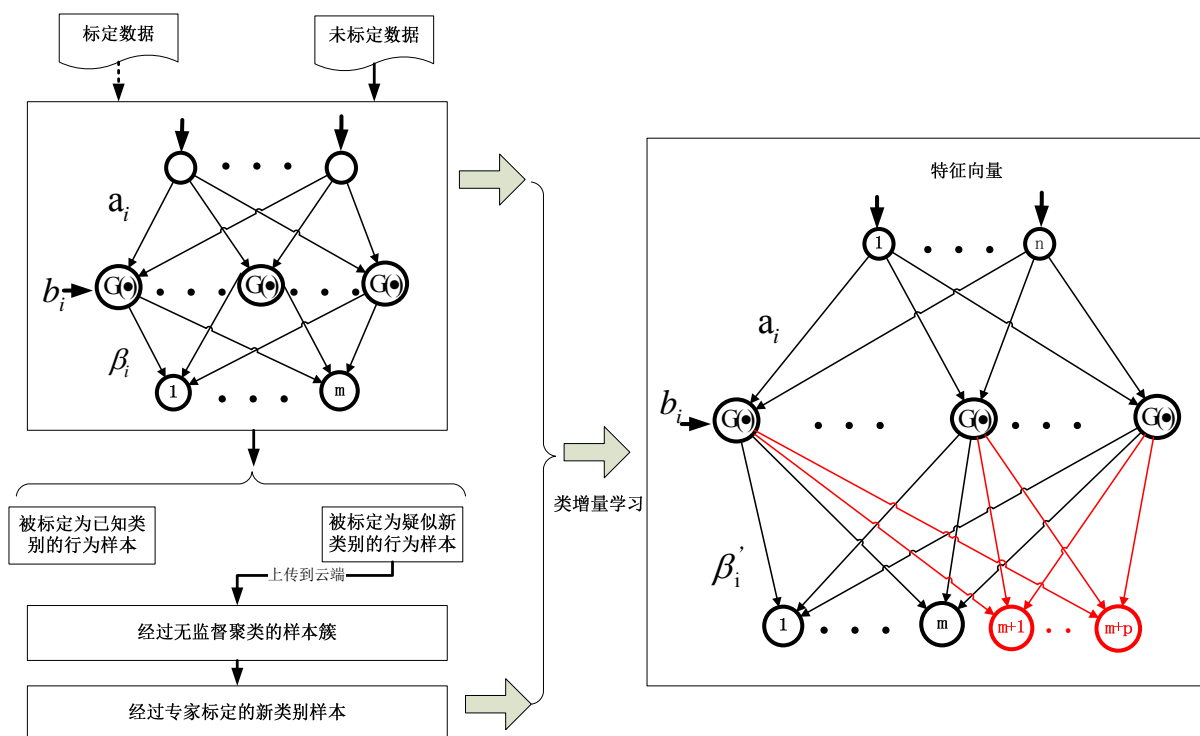


图 6.1 基于结构迁移的非特定类别行为识别框架

初始时，在离线阶段，如图中上部虚线部分所示，我们可以利用预研阶段采集的部分行为数据进行标定，或者收集已经公开发布的公共数据集，建立基于极速学习机（Extreme Learning Machine, ELM）的具有疑似异常行为识别能力的行为识别模型，此时的模型可以仅识别少数几种行为。我们将其部署到智能终端上，由用户携带自然地做出各种行为。如图中实线部分所示，按照先从上到下的顺序来看，行为识别模型首先对采集到的用户行为数据进行初步判断，对于置信度处于正常区间内的样本，系统直接对其标定；对于置信度处于异常区间（正常区间之外）的样本，系统标定其为疑似新行为样本。当疑似样本积累到一定数量，系统将其上传到“云端”，在“云端”发起无监督的聚类过程，将疑似行为样本分成不同的簇。通过专家在“云端”的标注，发现新的行为类别。然后采用基于结构迁移的类极速增量学习方法更新现有识别模型。

6.3.2 具有疑似异常行为识别能力的极速学习机

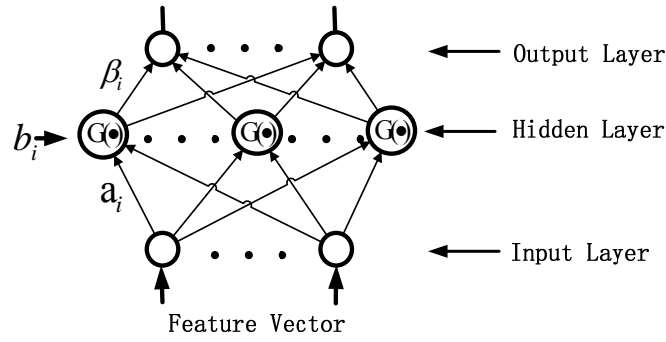


图 6.2 ELM 算法的网络结构

图 6.2 展示了 ELM 算法的网络结构，它是一个具有单隐结点层的神经网络，和其他机器学习算法相比，它不仅具有较好的识别精度，还具有较少的运算时间。ELM 无需训练调整神经元的输入权重和偏置值，它通过求线性系统 Moore-Penrose 伪逆和范数最小二乘解的方法计算输出权重[Huang,2011; Huang,2004; Huang,2006; Huang,2004a]。在预测过程中，通过寻找输出值最大的结点来决定最终的输出。尽管 ELM 算法能同时处理分类问题和回归问题，对我们来讲，仅考虑 ELM 的分类功能。

单隐层 ELM 的训练过程如算法 1 所示。

算法 1: ELM 算法 (Extreme Learning Machine)

Input: 设训练集为 $\mathbf{X} = \{(x_i, t_i) | x_i \in \mathbb{R}^n, t_i \in \mathbb{R}^m, i = 1, 2, \dots, N\}$ ，激励函数为 $G(a, b, x)$ ，隐结点数为 \tilde{N} 。

Output: 输出权重 β

Begin

1. 随机生成输入权重 a_i 和偏置值 b_i , $i = 1, 2, \dots, \tilde{N}$ 。
2. 计算隐结点层的输出矩阵 H :

$$H = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_N) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_N) \end{bmatrix}_{N \times \tilde{N}}$$

3. 计算输出权重矩阵 $\beta = H^+ T = (H^T H)^{-1} H^T T$ ，其中 H^+ 是矩阵 H 的 Moore-Penrose 伪逆， $T = [t_1, t_2, \dots, t_N]^T$

End

当用 ELM 分类器区分新样本 x 时，ELM 输出结点的值可以计算如下：

$$TY_{1 \times m} = [G(a_1, b_1, x), \dots, G(a_{\tilde{N}}, b_{\tilde{N}}, x)]_{1 \times \tilde{N}} \cdot \beta_{\tilde{N} \times m}$$

其中， m 是输出结点的个数，也是分类问题中类的个数。相应的， TY 就是一个含有 m 个值的行向量。

然后，ELM 分类器通过比较向量 TY 中哪个分量距离 1 最近，并将其下标 j 作为样本 x 的类标号。

$$j = \underset{1 \leq j \leq m}{\operatorname{argmin}} |1 - TY_j|$$

进一步，还可以从向量 TY 中计算出样本 x 的置信度， $confidence$ ，用以衡量样本 x 隶属于类 j 的程度。 $confidence$ 的值可以通过下面的步骤得到：

- 1) 计算 TY 每个分量与 1 的距离。

$$D = |TY - 1|$$

- 2) 计算 D 中每个分量的倒数，记做：

$$D_Inverse = \frac{1}{D}$$

- 3) 计算 $D_Inverse$ 中最大值在 $D_Inverse$ 中所占的比重，以此作为样本 x 的 $confidence$ 。

$$confidence = \frac{\argmax D_Inverse}{\sum D_Inverse}$$

每个离线采集并标定的样本都有一个对应的置信度 $confidence$ ，对于同一类样本的置信度集合，我们可以假设他们服从正态分布，从而可以计算它们的均值 μ 和标准差 σ ，按照正态分布的 3σ 原则： $P(\mu - \sigma < X \leq \mu + \sigma) = 68.3\%$ ， $P(\mu - 2\sigma < X \leq \mu + 2\sigma) = 95.4\%$ ， $P(\mu - 3\sigma < X \leq \mu + 3\sigma) = 99.7\%$ ，我们称区间 $[\mu - 3\sigma, \mu + 3\sigma]$ 为正常区间，即当置信度处于该区间内时，则认为 ELM 模型对行为样本的预测结果是可信的；我们称区间 $(-\infty, \mu - 3\sigma)$ 或者 $(\mu + 3\sigma, +\infty)$ 为异常区间，即当置信度处于该区间内时，则认为 ELM 模型对行为样本的预测结果是不准确的，其为一个疑似的新行为样本。该疑似样本或者为一个新行为样本，或者为已有行为的样本（只是由于用户自身的原因，比如疾病、衰老等，使得行为与先前行为不同）。

综上所述，引入了正常区间和异常区间后，使得 ELM 模型具有了识别疑似异常行为的能力。

6.3.3 新类别样本的发现和标定

当疑似样本积累到一定数量，系统将其上传到“云端”，在“云端”发起无监督的聚类过程，将疑似行为样本分成不同的簇。通过专家在“云端”的标注，发现新的行为类别。然后采用基于结构迁移的类极速增量学习方法更新现有识别模型。

6.3.4 极速学习机模型结构的迁移和更新

当有了当前用户的新类别行为数据，就可以对原有模型进行更新。正如在 6.2 节中所述的原因，我们需要一个类增量学习算法在线增量更新行为识别模型。本文我们提出了一个基于结构迁移的类增量极速学习机算法，基于在源域训练的 ELM 模型，我们将其结构迁移到目标域，用含有新类别行为样本的增量数据在原有结构上进行扩展和更新，得到满足目标域行为识别的新的 ELM 结构。

我们首先介绍现存的在线增量极速学习机算法（Online Sequential ELM, OSELM），分析其不能满足我们需求的原因，在此基础上，我们对算法进行了改进，提出了基于结构迁移的类增量极速学习机算法。

6.3.4.1 在线增量极速学习机算法

前面 6.3.2 节描述的 ELM 算法以整个训练数据集作为输入，一次性得到训练模型。为了得到在线增量 ELM 算法，Huang 给出了在线增量极速学习机算法，描述如下：

第一步：设初始数据集为： $\mathbf{x}_0 = \{(x_i, t_i)\}_{i=1}^{N_0}$ ，随机选择的输入权重为 a_i ，偏置值为 b_i ，选取的激励函数为 $G(a, b, x)$ ，隐结点的个数为 \tilde{N} 。那么按照 ELM 算法，隐结点层的输出矩阵 H_0 可计算为：

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_1) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_1) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_{N_0}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0}) \end{bmatrix}_{N_0 \times \tilde{N}}$$

相应的，输出权值为：

$$\beta_0 = K_0^{-1} H_0^T T_0$$

其中： $K_0 = H_0^T H_0$ ， $T_0 = [t_1, t_2, \dots, t_{N_0}]^T$ 。

第二步：设新来的下一块数据为： $\mathbf{x}_1 = \{(x_i, t_i)\}_{i=N_0+1}^{N_0+N_1}$ ，此时的数据 \mathbf{x}_1 中不含有与前面数据 \mathbf{x}_0 类标号不同的数据，那么采用同第一步相同的 a_i ， b_i ， G ， \tilde{N} ，同样可以计算出 H_1 ：

$$H_1 = \begin{bmatrix} G(a_1, b_1, x_{N_0+1}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0+1}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_{N_0+N_1}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_{N_0+N_1}) \end{bmatrix}_{N_1 \times \tilde{N}}$$

进而可以计算 K_1 ：

$$\begin{aligned} K_1 &= \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \\ &= \begin{bmatrix} H_0^T & H_1^T \end{bmatrix} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix} \\ &= K_0 + H_1^T H_1 \end{aligned}$$

这样，更新后的输出权值可计算如下：

$$\begin{aligned} \beta_1 &= K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}^T \begin{bmatrix} T_0 \\ T_1 \end{bmatrix} \\ &= K_1^{-1} (K_1 \beta_0 - H_1^T H_1 \beta_0 + H_1^T T_1) \\ &= \beta_0 + K_1^{-1} H_1^T (T_1 - H_1 \beta_0) \end{aligned} \quad (6.1)$$

从公式 (6.1) 可以看出， β_1 是 β_0 、 K_1 、 H_1 和 T_1 的函数，而跟先前的数据集 \mathbf{x}_0 无关，这意味着，当前数据使用完毕即可弃之不用，而不必继续保留，可以节省存储空间。可以看出，当没有新类数据出现时，可以用一个样本增量更新模型，也可以用一批数据增量更新模型。然而，当有新类别数据出现时，OSELM 却不能处理。

6.3.4.2 基于结构迁移的类增量极速学习机算法

对于我们的新问题，我们不仅希望模型能对已有类别行为的识别能力进行增量学习，也希望能在模型中添加对新类别的识别能力。然而每增加一类数据，不仅需要保持

原有的结构（输入结点、隐结点及其之间的联系），也需要更改 ELM 网络的结构，即需要增加输出结点的个数。基于实际的需要，我们提出了**基于结构迁移的类增量极速学习算法**，将已有 ELM 模型结构迁移到含有新类别数据的域，并用新类别的数据对已有 ELM 结构进行更新，使之能够处理新出现类别的标定样本。如图 6.3 所示。

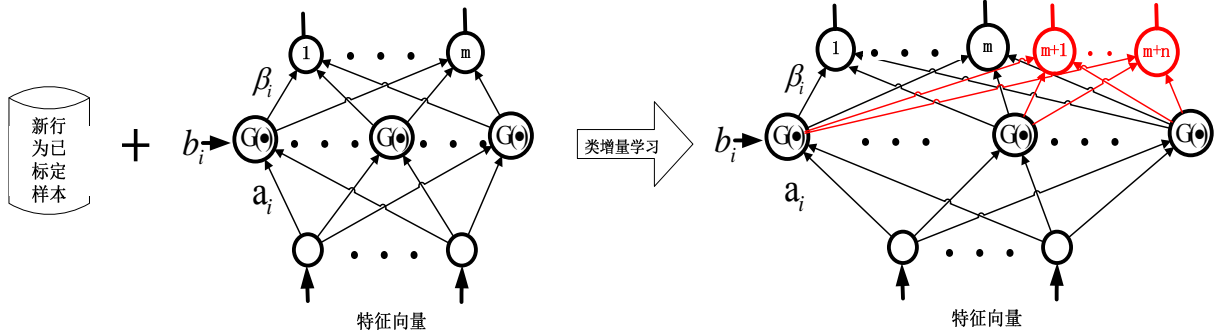


图 6.3 基于结构迁移的类增量 ELM 算法

下面详细介绍一下算法的理论推导过程。

假设我们已有部分已标定的数据为 $S = \{(x_s^{(i)}, y_s^{(i)}) | i = 1, 2, \dots, N_0\}$ ，某个用户提交了一部分新类的数据为 $D = \{(x_d^{(i)}, y_d^{(i)}) | i = 1, 2, \dots, N_1\}$ 。为简化问题的描述，也不失一般性，我们假设 D 中数据的类标号只有一个，且仅比 S 的最大的类标号多一。

根据 ELM 算法，用 S 训练分类模型，可以得到：

$$\beta_0 = K_0^{-1} H_0^T T_0,$$

其中，

$$H_0 = \begin{bmatrix} G(a_1, b_1, x_s^{(1)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_s^{(1)}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_s^{(N_0)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_s^{(N_0)}) \end{bmatrix}_{N_0 \times \tilde{N}}$$

$$T_0 = \begin{bmatrix} t_1^T \\ \vdots \\ t_{N_0}^T \end{bmatrix}_{N_0 \times m}$$

$$K_0 = H_0^T H_0$$

那么对于 D ，我们可以计算得到 H_1 ，

$$H_1 = \begin{bmatrix} G(a_1, b_1, x_d^{(1)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_d^{(1)}) \\ \vdots & \ddots & \vdots \\ G(a_1, b_1, x_d^{(N_1)}) & \cdots & G(a_{\tilde{N}}, b_{\tilde{N}}, x_d^{(N_1)}) \end{bmatrix}_{N_1 \times \tilde{N}}$$

对于 D 的每个标号，用多维向量表示，其应该比 T_0 中的类标号多一列。我们用 T_1 表示：

$$T_1 = \begin{bmatrix} 0 & \cdots & 01 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 01 \end{bmatrix}_{N_1 \times (m+1)}$$

如果将 S 和标定后的 D 同时当做训练集训练 ELM 模型，得到的 β_1 应该如下式：

$$\beta_1 = K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}_0^T \begin{bmatrix} T_0 \cdot M \\ T_1 \end{bmatrix}$$

其中,

$$K_1 = \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}_0^T \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}$$

$$M = \begin{bmatrix} 1 & \cdots & 00 \\ \vdots & \ddots & \vdots 0 \\ 0 & \cdots & 10 \end{bmatrix}_{m \times (m+1)}$$

M 是一个转换矩阵, 作用是在矩阵 T_0 的右面添加一列 0, 使得 T_0 和 T_1 的列数相同。

β_1 的推导如下:

$$\begin{aligned} \beta_1 &= K_1^{-1} \begin{bmatrix} H_0 \\ H_1 \end{bmatrix}_0^T \begin{bmatrix} T_0 \cdot M \\ T_1 \end{bmatrix} \\ &= K_1^{-1} \left(\begin{bmatrix} H_0^T & H_1^T \end{bmatrix} \begin{bmatrix} T_0 \cdot M \\ T_1 \end{bmatrix} \right) \\ &= K_1^{-1} (H_0^T T_0 M + H_1^T T_1) \\ &= K_1^{-1} (K_0 K_0^{-1} H_0^T T_0 M + H_1^T T_1) \\ &= K_1^{-1} (K_0 \beta_0 M + H_1^T T_1) \\ &= K_1^{-1} [(K_1 - H_1^T T_1) \beta_0 M + H_1^T T_1] \\ &= K_1^{-1} (K_1 \beta_0 M - H_1^T T_1 \beta_0 M + H_1^T T_1) \\ &= \beta_0 M + K_1^{-1} H_1^T (T_1 - H_1 \beta_0 M) \end{aligned}$$

从上式可以看出, β_1 和数据集 S 已经不再直接相关, 因此可以通过增量学习的方法改进模型, 使其能够适应新出现的类别。

在实际应用过程中, 已有类别数据和新类别数据可以同时出现在增量数据中, 算法可以根据增量数据中的类标号确定是否有新类别的出现, 若有则按照类增量极速学习算法进行 ELM 结构的扩展, 否则就按照 OSELM 算法进行模型的在线增量学习。

6.4 实验及结果分析

本部分, 基于上述数据集我们验证了本章提出的类增量极速学习算法。共验证了五方面的内容: 模型选择, 识别异常行为的实验, 类增量极速学习性能分析, 类增量极速学习稳定性分析, 增量学习与批量学习的对比分析等。实验的硬件平台是一台普通的 PC 机, CPU 主频为 2.6GHz, 软件平台是 MATLAB2009a。ELM 源代码下载自算法提出者 Huang 教授的主页⁸。

6.4.1 数据准备

为了验证类增量极速学习算法的性能, 我们采用第二章描述的基于 N95 的数据集进

⁸ Source codes and some references of ELM can be found at www.ntu.edu.sg/home/egbhuang

行实验。在该数据集中, 有 12 个人参与了数据采集, 每个人分别采集了静止、散步、跑步、上楼梯、下楼梯等 5 种动作, 在采集数据过程中, 手机放置在采集者的右裤袋中。采用滑动窗口的方法提取特征数据。虽然 N95 内置加速度传感器的采样频率最高可以达到 100Hz 或者 400Hz, 但由于考虑到能耗等原因, Nokia 加速度传感器的内置 API 函数只能以 32Hz 的频率采集数据。我们选择的窗口长度是 2 秒, 重叠窗口长度是 1 秒, 这样每个窗口中可以包含完整的动作。在早期的文献如[Bao,2004]中已经展示了这种滑动窗口方法提取特征的效果。在每个窗口中, 最终提取的特征有: (1) 时域的最大值、最小值、均值、标准差、众数、过零点个数、范围、信号幅值面积等 8 个特征; (2) FFT 变换后的 64 个系数; (3) 功率谱密度 PSD 的 8 个特征。总计 80 个特征。最终得到的样本信息如表 6.1 所示。

表 6.1 行为样本信息

行为类别	类标号	样本数量
静止	1	5256
散步	2	4012
跑步	3	3756
上楼梯	4	1656
下楼梯	5	1611

6.4.2 模型选择

模型选择是指选择最优的网络结构和选择机器学习算法中最优的参数, 它一般是模型相关的。对于本章的算法而言, 只有隐层结点的个数是需要优化选择的。

为了选择最优的隐层结点个数, 采取了训练和验证的方法。数据集被随机均分为两部分, 一部分作为训练数据, 另一部分作为测试数据。隐层结点的个数在 1 到 300 之间遍历, 相应的实验结果如图 6.4 所示。

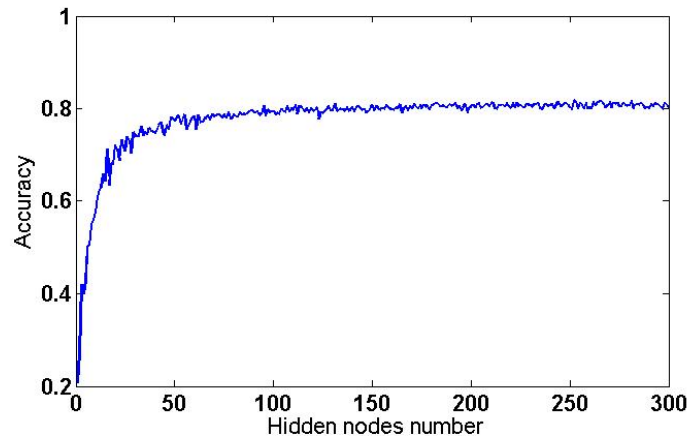


图 6.4 ELM 模型选择

如图 6.4 所示，当隐层结点的个数大于 100 时，模型就能达到最好的性能且比较稳定。因此，在后续实验中，我们选择 100 个隐层结点作为最优的模型结构。

6.4.3 识别异常行为的实验

选取数据集中的一种行为作为异常行为，其他四种行为作为已知行为，用已知行为数据训练具有疑似异常行为识别能力的 ELM，用异常行为数据进行测试。该过程重复 5 次，每次用一种不同的行为作为异常行为，剩余行为作为已知行为，测试结果取平均值，如下表 6.2 所示。

表 6.2 模型识别异常行为的试验结果

训练数据	测试数据	平均测试精度
随机四种行为数据	剩余一种行为数据	98.02%

从表 6.2 可以看出，模型对异常行为的识别能力达到了 98% 以上，验证了该方法对不在训练集中出现的行为具有较高的识别能力。

6.4.4 分类器性能 VS 新样本数量

在本实验中，我们关注的是分类器对新类的分类能力，关注它是如何随着新类样本的增多而不断地增强的。在增量学习过程中，每次用多少样本学习也是我们关注的问题。

实验过程中，我们进行了 12 轮实验，在每轮中，选取一个人静止、散步、跑步、上楼梯这四类动作训练一个模型，然后逐渐加入下楼梯的数据进行增量学习，记录了不同增量数据时分类模型的分类精度。最终，用 12 轮数据的平均值来衡量系统的性能。

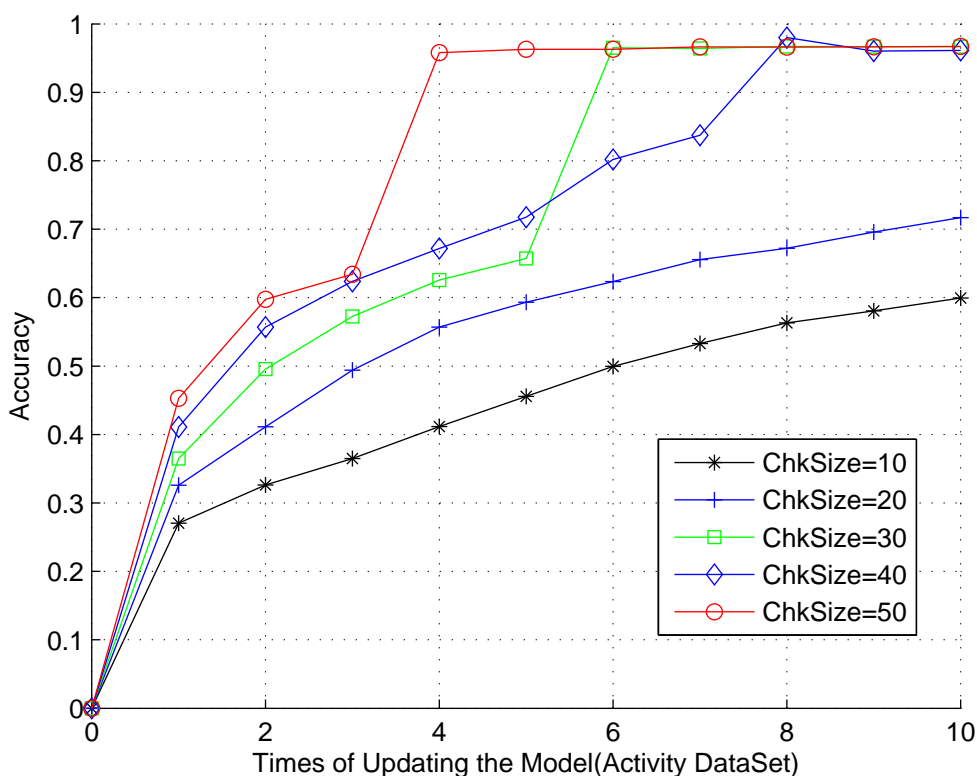


图 6.5 分类器性能与增量数据之间的关系

从图 6.5 中可以看出, 无论增量块的大小如何, 随着增量次数的增加, 模型的识别精度都有逐步提高的趋势。只是在增量块不大于 20 的时候, 识别精度提高的速度较慢; 在增量块大于等于 50 时, 识别能力提高的速度就较快, 大约经过 4 到 5 次的增量学习, 就能达到较高的识别精度。

6.4.5 行为识别模型稳定性实验

在 6.4.4 节中, 我们用固定的 4 类数据建立模型, 用固定的一类数据来更新模型, 以此验证了模型精度对增量数据集大小的依赖性。本节, 我们验证算法的稳定性, 即以任意一类行为数据作为新类, 其他类作为已知类, 通过增量学习得到模型是否稳定。实验结果如图 6.6 所示。

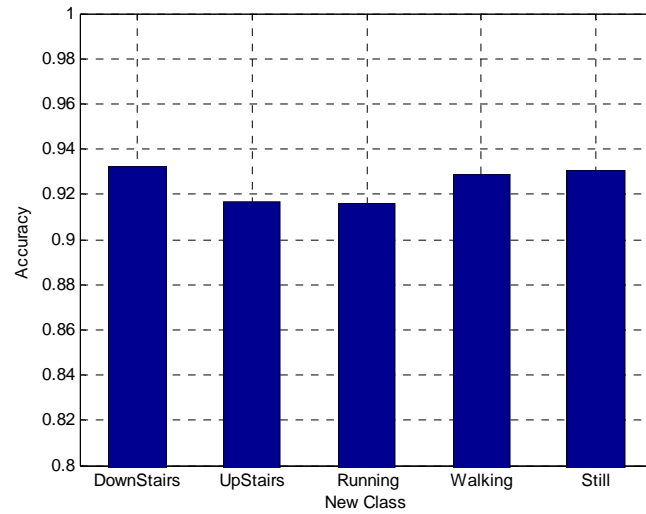


图 6.6 行为识别模型稳定性实验

实验结果表明,无论哪一类作为新类对原有模型更新,更新后的模型对新类数据的识别率均能达到 90%以上,说明这种方法是一个普适可行的方法,不是一个仅对某类敏感的方法。也说明我们的行为识别模型是可以通过这种方法得到增量更新的,能够逐渐学习到对新类的识别能力。

6.4.6 批量学习与增量学习的性能比较实验

本实验验证增量学习和批量学习的学习效果。

对于每个人的数据,初始时任意选取两个位置训练模型,测试其训练精度作为二类行为的批量学习精度,对于增量学习,我们也以此精度作为二类问题的精度。

然后加入第三类数据,对于批量学习,是将其和前两类数据同时作为训练集训练分类模型,得到三类问题的训练精度;对于增量学习,是在二分类模型的基础上用增量数据基于本章的类增量算法更新模型,得到一个训练精度。

同理,分别加入第四类、第五类数据,得到相应的批量学习精度和增量学习精度。

通俗的讲,对于 A、B、C、D、E 五种行为,取其不同的排列方式,对于每一种排列,先取前两种行为建立模型,然后逐渐加入第三类、第四类、第五类。该过程对于每个人的每种排列重复一次,最后取模型性能的均值。图 6.7 列出了实验结果。

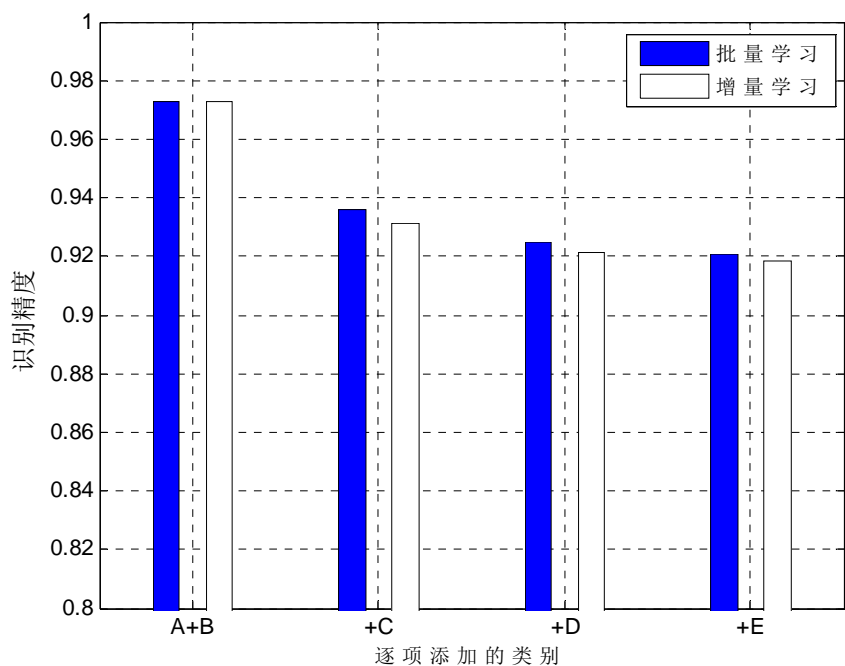


图 6.7 批量学习 VS 增量学习

从图 6.7 可以看出，对于批量学习和增量学习，随着新类别的增加，总体识别精度都呈下降的趋势，这是因为随着类别的增加，识别任务越来越困难。对于我们的 5 类行为，从二分类模型到三分类模型，性能下降得较快，随后下降的速度有所放缓。

对于增量学习，与批量学习相比，其性能略低，这是因为增量学习每次涉及到的数据数量有限，对总体规律的反映能力略差。

6.5 小结

本章针对行为识别过程中需要对新类别数据进行增量学习的问题，提出了一种基于结构迁移的非特定类别行为识别方法，重点研究了一种能够对新类样本进行增量学习的算法。该算法基于增量学习，能够处理模型自动更新的问题。通过在行为数据集上详细的实验，验证了算法的有效性。

如何自动地识别新类别行为数据是一个值得研究的课题，在将来的研究工作中，我们将研究基于“众包”理论的行为样本采集和标定方法，期望发现一种能够准确、快速获取、标定各种行为的方法，为构建用户的个性化行为识别模型提供更好的服务。

第七章 基于智能移动终端的社区老人健康感知公共服务系统

目前,我国的老龄化、空巢化、失能等问题日益突出,形式相当严峻。然而,我国医疗基础设施还非常不完善,医疗资源相对匮乏,面向老年人健康的服务十分缺乏。依靠先进的科学技术大力发展面向老年人的医疗健康服务,提高老年人日常监护水平和独立生活能力,已经成为十分迫切的社会需求。

我们设计了一个基于智能移动终端的社区老人健康感知公共服务系统,在日常生活可以识别出老人的行为,估计他们的能量消耗情况,在老人跌倒时,还可以及时地发出警报,极大程度地减轻他们受到的伤害。

7.1 背景

人口老龄化当前已成为一个世界性的问题,指一个地区或国家 60 岁及以上或 65 岁及以上人口数量占该地区或国家人口总数的百分比。根据联合国的传统标准,60 岁以上人口占总人口 10%以上,或 65 岁以上人口占总人口 7%以上的社会被称为老龄化社会或老年型国家。

根据 2010 年全国第六次人口普查的数据公报,我国 60 岁及以上人口已达 1.78 亿人,占总人口的 13.26%,比 2000 年时人口普查的数据上升了 2.93%,其中超过 65 岁的人口大约为 1.19 亿,占总人口的 8.87%,比 2000 年人口普查上升了 1.91 个百分点,现在已经远远越过 7%的“红线”。数据表明:我国自 2000 年起跨入老龄化社会,近十年来由于中国的生育水平较低,老龄化进程有逐步加快的趋势。

保持身体健康是老年人独立、幸福生活的前提条件之一,而锻炼是达到此目的的有效途径。运动处方⁹是指导大众有计划、有目的和科学地锻炼的一种方法。而如何衡量用户真实的运动量是否符合运动处方的规定,是一个至关重要的问题。运动量太小,起不到锻炼的作用;运动过量,容易造成运动损伤,过犹不及。已有少量设备能够估算运动量的多少,比如计步器(Step Counter)以用户走过的步数量化用户的行为,非常适合测量散步和跑步的运动量。再比如博能公司的产品通过心率来量化用户消耗的卡路里数量,非常适合有氧运动项目。但是,这些方法只是给用户一些统计的数字,而不能给出每种运动的细节。并且,在使用博能公司的产品时,用户必须佩戴专门的心率带,对于某些用户来说,显得有些不便。

老年人由于身体衰弱,平衡能力较差,跌倒已成为威胁他们健康的重要因素。除了

⁹ 运动处方的概念最早是美国生理学家卡波维奇在 20 世纪 50 年代提出的。20 世纪 60 年代以来,随着康复医学的发展及对冠心病等的康复训练的开展,运动处方开始受到重视。1969 年世界卫生组织开始使用运动处方术语,从而在国际上得到认可。运动处方的完整概念是:康复医师或体疗师,对从事体育锻炼者或病人,根据医学检查资料(包括运动实验和体力测验),按其健康、体力以及心血管功能状况,用处方的形式规定运动种类、运动强度、运动时间及运动频率,提出运动中的注意事项。

平时加强锻炼，防患于未然外，在跌倒事实已经造成后，及时地通知其监护人或者医疗机构，也能极大地减轻身体受到的危害。据统计，中国每年有不少于 2000 万老年人不幸跌倒，大约有不少于 2500 万人次的跌倒，由此造成的住院治疗费用高达 50 亿元以上，社会代价大约为 160 到 800 亿元。跌倒已经是导致老年人死亡的首位因素。意外伤害中大约有一半就是老年人跌倒，而跌倒导致的死亡率已经高居第 4 位，排在其前面的且处于第一位的是心脑血管疾病，处于第二位的是肿瘤疾病，处于第三位的是呼吸道疾病。老人意外伤害多在家中发生，占 57.75%，在公共场所发生的占 21.13%，在街道马路发生的占 18.31%，在其他场所发生的占 2.82%。

我们选择智能移动终端作为老年人行为识别系统的平台，是因为其具有下述优点：

1) 不可替代性。移动终端尤其是手机已成为日常生活中重要的一部分，人们出门可能会忘记带钥匙，很少会忘记带手机。2) 便携性。手机可以拿在手里，放在口袋内，放在包中，挂在胸前，佩戴在腰间，人们不会感觉到不便。3) 智能性。现有智能手机内置有多种传感器，如摄像头、麦克风、加速度传感器、陀螺仪等等，随着科技的发展，内置的传感器将会越来越多。智能手机厂商还会提供操作这些传感器的 API 接口，供人们开发应用程序。4) 经济性。选择智能手机作为行为识别系统的载体，对于用户来说，不必支出额外的花费去购买新的设备。对于开发人员来说，不必购买或者研发新型的传感器设备，可减少研发费用。

7.2 系统架构

我们开发了一个社区老人健康感知公共服务系统，整个系统涉及三部分：移动终端 APP、后台服务器以及用户，用户主要包括老人和医生。系统提供的功能模块主要有：行为识别、能量消耗估计和跌倒报警。

1、移动终端

移动终端是整个系统的核心，所有的交互（包括医疗传感器与终端的交互、人与终端的交互、服务器与终端的交互）都需要通过移动终端。在移动终端上安装有一套我们开发的用于老人健康监控的软件，实现了系统的所有功能，包括行为识别、能量消耗估计、跌倒报警的人机交互接口。移动终端实时获取监控到的数据，既可以实时地将其以合适的方式在移动终端界面上显示给老人，又可以将其通过无线传输的方式保存在服务器端供系统和医生进一步分析。

2、后台服务器

后台服务器是整个系统的数据存储中心，用户在使用该系统的过程中，移动终端采集到的所有的数据都会被上传到后台服务器并存储在数据库中。后台服务器还提供了对数据进行进一步统计的功能，当用户需要某个时间段健康数据的统计信息时，可以向服务器发送请求，服务器会将结果发送到移动终端并显示给用户。同时后台服务器还提供了医生查询数据的接口，医生可以通过后台服务器匿名地获取用户的数据，并进行分析，

然后给出健康建议，再通过后台服务器反馈给终端用户。

3、用户

用户包括老人和医生。

老人是系统的主要用户，也是系统提供服务的对象。老人在日常生活中只需随身携带移动终端即可，它会自动采集老人的运动数据并计算相应的能量消耗，相应的数据也会存储到服务器端，并在老人查询的时候以图表的方式予以显示。

医生也是系统的用户之一。医生可以根据老人的请求分析老人的健康数据，并给出健康建议，系统会将健康建议反馈给老人。老人通过系统获取想要的健康监护信息，医生则通过系统获取提供健康建议所得的报酬。

7.3 系统的基本功能

系统的基本功能主要有：行为识别、能量消耗估计和跌倒报警。我们在本节分别给予简单介绍。

7.3.1 行为识别



图 7.1 行为识别和行为统计分析功能展示

如图 7.1 所示，系统提供了行为识别的功能，该功能可以实时识别老人的运动状态，主要包括静止、走路和跑步这三种状态。

有些医学指标习惯采用运动的持续和累计时间衡量用户的运动量。如运动处方。在行为识别的基础上，系统对每次识别的相应行为的时间进行累计，实时统计每种行为的持续时间，并将其记录下来，然后上传到后台服务器中。

也有些医学指标习惯采用用户走过或者跑过的“步数”衡量用户的运动量。比如中国的谚语“饭后百步走，活到九十九”；再比如 2007 年 9 月，我国卫生部向全国人民群

众发出倡议，提出“每日一万步；吃动两平衡；健康一辈子”的口号。系统利用“人在走路和跑步的过程中，加速度曲线会呈周期性波动”的规律，通过检测加速度曲线上波峰的个数，利用阈值的方式来进行步数的统计。

7.3.2 热量消耗估计

系统提供了热量消耗估计功能，该功能根据行为统计数据，利用运动医学上的热量消耗公式，估算出老人运动所消耗的热量，包括走路所消耗的热量以及跑步所消耗的热量。当用户查询热量消耗的统计数据时，系统会将一周或者一个月的热量数据进行统计并返回给用户。同时用户还可以根据统计的热量消耗数据，分析自己的运动与饮食是否平衡。对于老年人来说，其运动种类较少，大多是静止、散步、跑步三种运动，因此我们主要计算该三种运动的热量消耗情况。如图 7.2 所示。



图 7.2 能量消耗估算功能展示

7.3.3 跌倒报警

跌倒报警的功能是利用对加速度数据进行实时建模和分析而实现的，它利用多阈值协同工作的方式，即跌倒过程中失重和超重两个状态和相应状态的时间差来判断跌倒是否发生。

一个典型的跌倒过程如图 7.3 所示。加速度会有一个失重（左边的波谷所示）和一个超重（右边的波峰所示）的过程，并且超重过程和失重过程在一个很小的时间范围内发生。所以，我们设置三个阈值即可实现，一个超重阈值，一个失重阈值和一个时间阈值。如果在一秒内发生的失重和超重均超过了阈值，我们认为是发

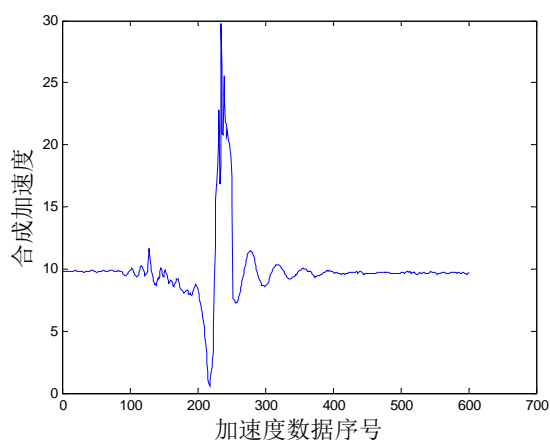


图 7.3 一个典型的跌倒过程的加速度曲线

生了跌倒。同时，为了减少误报率，防止手机因小规模振动导致的误报，我们还设置了一个时间下限阈值，即超重和失重之间的时间间隔需满足合理的条件。

在跌倒检测的基础上，系统提供了跌倒报警的功能，系统可以实时地识别用户是否跌倒，在发现用户的疑似跌倒行为时，系统会发出报警声，提示并询问用户是否跌倒，用户如果确认或者不响应，系统会发送一条短信给用户设置的一个紧急联系人，告诉紧急联系人用户在哪里跌倒了。如果是系统误报，用户可以将报警取消。如图 7.4 所示。



图 7.4 跌倒报警功能展示

7.4 系统功能的设计与实现

7.4.1 行为识别功能的实现

行为识别功能的实现是利用移动终端内置的加速度传感器所采集的加速度数据进行实时的判别。要实现行为识别的功能，我们首先利用智能终端采集老人的行为数据，然后对行为数据进行训练，过程如下：

1) 老人行为数据的采集

首先采集老人的行为数据，主要包括静止、走路和跑步这三种运动状态的数据。采集的数据是智能终端获取的加速度计的原始值，即三个方向上的值。这些值不是真实的加速度值，只是系统通过 API 函数读取到的加速度的值。

2) 将采集的加速度数据转换为真实的加速度值

接着需要将采样的加速度计的值转换为真实的加速度值，将手机三个方向静止放置，就可以获得三个方向加速度值，计算每个方向的系数和偏移。所有的加速度通过乘以这个系数并加上偏移就可以得到真实的加速度值。

3) 选择一个适当的时间窗口

我们需要分析一段时间内的运动状态，就需要对一段时间内的加速度值作为一个样本点来处理，所以需要选取一个适当的时间窗口，我们选择 1 秒钟为一个时间窗口，即 32 个加速度值。

4) 对每个时间窗口内的数据进行均值滤波

滤波的目的是消除采样的加速度数据中的噪声，以防止一些比较大偏差的数据影响训练的结果。

5) 对滤波后的数据计算合成加速度大小

由于采样的加速度值是三个方向的加速度，根据力学原理，将三个方向的加速度值进行平方和，即可得到总的合成加速度。

6) 提取合成加速度的特征

对窗口内的加速度值提取特征，我们主要提取了如下四个特征：

- 时间窗口内加速度的均值
- 时间窗口内加速度的标准差
- 时间窗口内加速度的过均值点个数
- 时间窗口内加速度的频域的峰值（需对数据进行傅立叶变换）

7) 对提取出来的特征用决策树算法进行训练得出模型参数

最后用决策树算法对特征进行训练，得出决策树模型的参数，具体的训练方法这里不再赘述。

训练的模型即可用于在线行为识别，识别过程如下：

1) 老人行为数据的采集

首先实时采集老人的行为数据，主要包括静止、走路和跑步这三种运动状态的数据。采集的数据是加速度计的原始值。

2) 将采集的加速度数据转换为真实的加速度值

接着需要将采样的加速度计的值转换为真实的加速度值，将手机三个方向静止放置，就可以获得三个方向加速度值，计算每个方向的系数和偏移。所有的加速度通过乘以这个系数并加上偏移就可以得到真实的加速度值。

3) 选择一个适当的时间窗口

我们需要分析一段时间内的运动状态，就需要对一段时间内的加速度值作为一个样本点来处理，所以需要选取一个适当的时间窗口，我们选择 1 秒钟为一个时间窗口，即 32 个加速度值。

4) 对每个时间窗口内的数据进行均值滤波

接着需要对时间窗口内的数据进行均值滤波，滤波的目的是消除采样的加速度数据中的噪声，以防止一些比较大偏差的数据影响训练的结果。

5) 对滤波后的数据计算线性加速度大小

由于采样的加速度值是三个方向的加速度，根据力学原理，将三个方向的加速度值进行平方和，即可得到总的加速度。

6) 提取加速度值的特征

对窗口内的加速度值提取特征，我们主要提取了如下四个特征：

- 时间窗口内加速度的均值

- 时间窗口内加速度的标准差
- 时间窗口内加速度的平均交叉率
- 时间窗口内加速度的频域的峰值（需对数据进行傅立叶变换）

7) 利用决策树模型参数对特征进行判别，得出识别出来的运动

利用决策树模型对每一个特征进行判别，得出一个识别结果，在这里我们只能识别三种运动状态：静止、走路和跑步。在我们的实现中，考虑到计算复杂性和模型识别精度之间的折衷关系，只用到了标准差和频谱峰值作为决策树的参数，通过这两个参数就可以较好实现这三种运动状态的识别。

整个行为识别的流程图如图 7.5 所示。

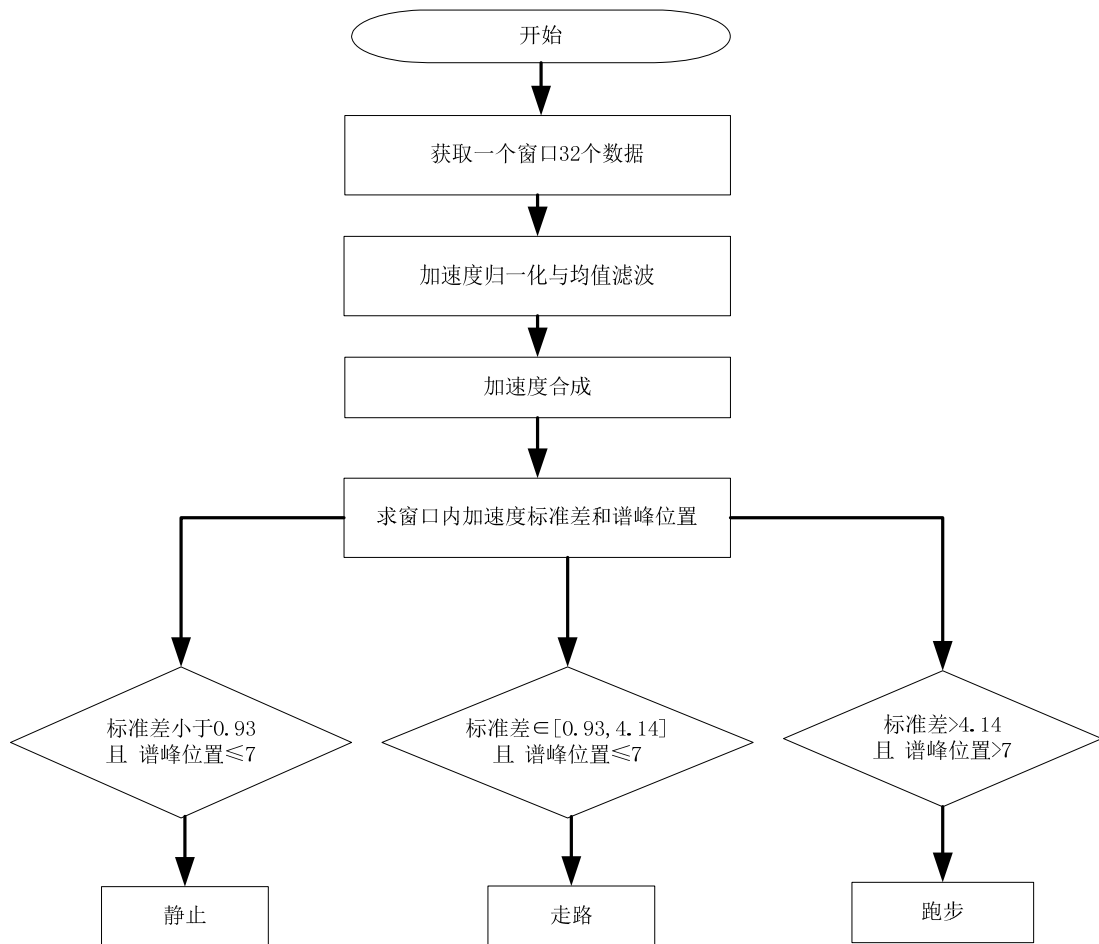


图 7.5 行为识别算法流程图

7.4.2 行为统计分析功能的实现

行为统计分析功能可以统计一周或者一个月内用户各种运动状态的时间，以及运动的步数。对于时间的统计，是在行为识别的基础上，对每次识别的相应行为的时间进行累计。利用阈值的方式来进行步数的统计，人在走路和跑步的过程中，加速度曲线会呈周期性波动，每走一步都会有一个明显的峰值。波峰的个数就是走过的步数。计步的过

程如下：

1) 实时采集行为数据

实时采集老人的行为数据，即加速度数据，能够通过读取加速度传感器三个轴上的数据得到。

2) 将加速度数据转换为真实的加速度值

将三个方向的加速度计采集的值转换为三个方向的加速度的真实值。并计算三个方向加速度值的平方和，获取总的加速度。

3) 选取 32 个数据的时间窗口

以 32 个数据为一个窗口，对数据进行处理，因为我们所采用智能终端的采样率为 32。

4) 对数据进行中值滤波

对数据进行中值滤波的目的是滤除数据中的噪声，以保证数据不会出现大的偏差。

5) 计算窗口内数据的方差

计算窗口内数据的方差，方差就是我们数波峰的样本点。每个时间窗口内的数据计算出来的方差会形成一个方差序列。

6) 对方差序列进行分析，数波峰

对方差序列进行分析，当一个波峰超过阈值的时候，就认为用户走了一步，那么当前行为的步数值就加一。

计步过程的流程图 7.6 如下：

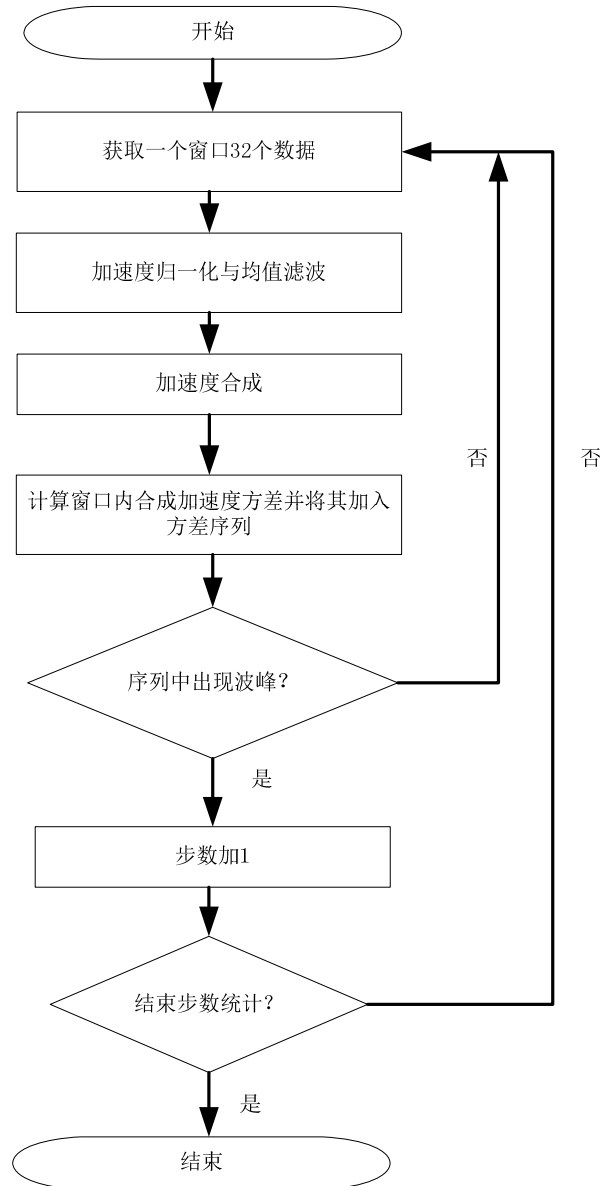


图 7.6 计步算法流程图

7.4.3 热量消耗估计功能的实现

热量消耗估计，是利用行为统计分析的结果，并利用医学上的热量消耗估算公式，估算热量消耗值，对于老年人来说，其运动种类较少，大多是静止、散步、跑步三种运动，因此我们主要计算该三种运动的热量消耗情况。

人在运动时消耗的热量与其所消耗的氧气量成正比，大约为氧气量的 5.01 倍，即：

$$Calories = 5.01 * VO_2 \dots\dots\dots(7.1)$$

其中 VO_2 表示用户消耗的氧气量，单位是升。Calories 表示消耗的热量。

从式 7.1 可以看出， VO_2 值是估算热量消耗的关键。根据权威机构研究，人体运动时消耗的氧气主要由三部分构成，一部分是人体基本消耗，即处于静止状态呼吸平稳时的消耗，用 R 表示；一部分是由于水平运动而产生的消耗，用 H 表示；另一部分是由

于竖直方向运动而产生的消耗，用 V 表示，因此 VO_2 可以写成如下形式：

$$VO_2 = R + H + V \dots\dots\dots(7.2)$$

那么热量的估算公式就是：

$$Calories = 5.01 * (R + H + V) \dots\dots\dots(7.3)$$

其中 R 的计算公式如下：

$$R = 3.5ml * Weight(kg) * Time(minutes) / 1000 \dots\dots\dots(7.4)$$

$Weight$ 表示体重，单位是千克 (kg)， $Time$ 表示持续时间，单位是分钟 (minute)。

对于 H 和 V 的计算，走路和跑步时不同，走路时的计算方法如下：

$$H = 0.1ml * speed(m/min) * Weight(kg) * Time(minutes) / 1000 \dots\dots\dots(7.5)$$

$$V = 1.8ml * speed(m/min) * Weight(kg) * Time(minutes) / 1000 * grade \dots\dots\dots(7.6)$$

其中的 $grade$ 如下图所示。

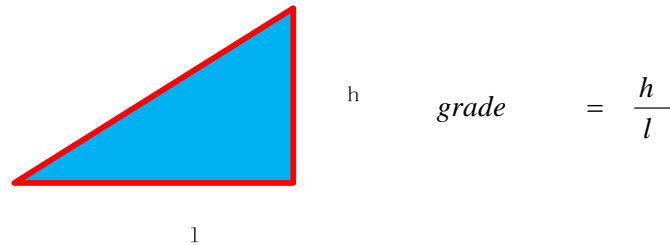


图 7.7 grade 的计算

公式中的速度是通过身高、步数和时间估算出来。由于我们的运动主要是在平地上完成，不包括上下楼梯等运动，所以不需要估算公式中的 V 值。

跑步时的计算方法如下：

$$H = 0.2ml * speed(m/min) * Weight(kg) * Time(minutes) / 1000 \dots\dots\dots(7.7)$$

$$V = 0.9ml * speed(m/min) * Weight(kg) * Time(minutes) / 1000 * grade \dots\dots\dots(7.8)$$

除了前面的系数和走路时不一样，其他都是一样的。

速度由如下公式计算：

$$Speed = K * height * avg_steps_per_second \dots\dots\dots(7.9)$$

即速度由系数 K 乘以身高乘以平均每秒钟的步数而得到。其中 K 对男性来说是 0.415，对女性来说是 0.413， $Height$ 是身高，单位是厘米。每秒种的步数 $avg_steps_per_second$ 通过运动的步数和时间可以很容易算出来。

系统中利用行为统计功能得出的运动的步数和时间信息，再加上用户提供的性别体重身高就可以计算所消耗的热量的大致估计值。

7.4.4 跌倒检测功能的实现

跌倒检测的功能是利用对加速度数据进行实时建模和分析而实现的，它利用多阈值协同工作的方式，即跌倒过程中失重和超重两个状态和相应状态的时间差来判断跌倒是否发生。

在跌倒过程中，加速度会有一个超重和一个失重的过程，并且超重过程和失重过程在一个很小的时间范围内发生。所以，我们设置三个阈值即可实现，一个超重阈值，一个失重阈值和一个时间阈值。在系统实现中，我们选取的超重阈值是 30，失重阈值是 5，时间阈值用数据点表示，由于我们的系统基于的智能终端的采样频率为 32，因此我们设置的数据点的阈值也是 32，即在一秒内发生的失重和超重，我们认为是发生了跌倒。同时，为了减少误报率，防止手机因小规模振动导致的误报，我们还设置了一个时间下限阈值，即数据点个数为 8，数据点个数必须同时保证大于 8，小于 32，我们认为发生了一次跌倒。识别跌倒的过程如下：

- 1) 获取一个采样点。
- 2) 根据采样点计算加速度值 a 。
- 3) 如果长度为 32 的队列 Q 不满，则将 a 加入队尾并继续获取样本点；如果队列已满，则删除队头，并将 a 插入队尾。
- 4) 如果当前加速度值（即队尾元素）大于 30，再判断距离其 8 到 32 个元素距离的元素的加速度值是否小于 5，若有则判断有跌倒发生，否则没有跌倒发生，继续获取样本点。

跌倒检测算法流程如图 7.8 所示。

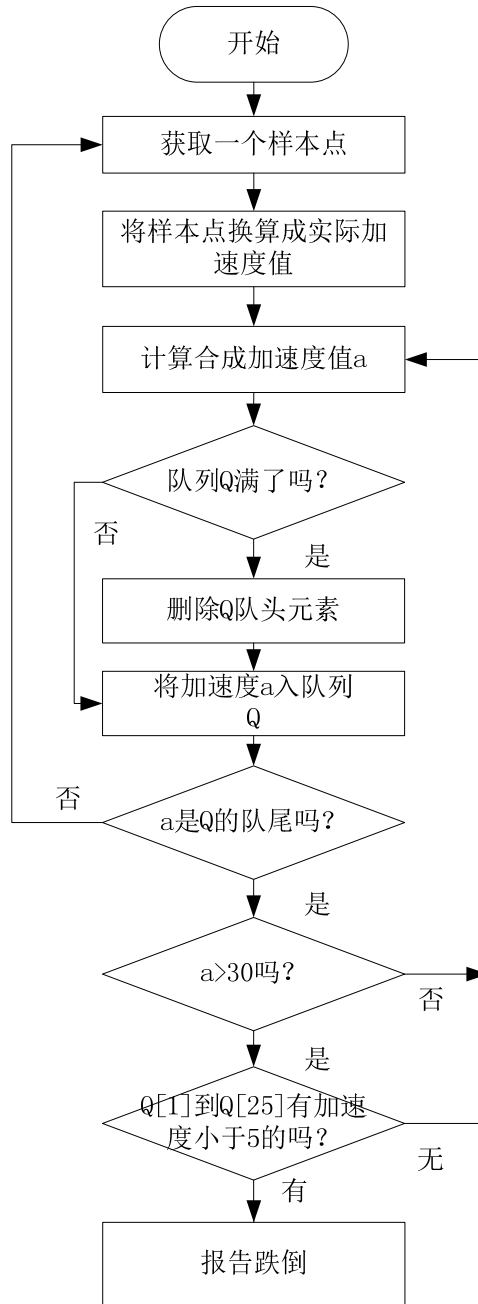


图 7.8 跌倒检测算法流程图

7.5 小结

针对目前我国的老龄化、空巢化、失能等问题，基于前面各章的部分研究成果我们设计了一个基于智能移动终端的社区老人健康感知公共服务系统，在日常生活中可以检测老人的运动量，为促进其合理运动提供帮助；在老人跌倒时，可以及时地发出警报，极大程度地减轻他们受到的伤害。

第八章 结束语

本章首先对本文在非特定传感器行为识别、非特定位置行为识别、非特定人行为识别以及非特定类别行为识别方面的研究工作和取得的结果进行简要总结，然后对下一步工作进行展望。

8.1 本文工作总结

基于移动终端的自动行为识别在“空巢”老人生活远程监护、肥胖病人热量消耗估计、心脑血管等慢病病人的健康恢复、老年人的跌倒检测与报警等领域具有广阔的应用前景。在研发过程中，我们发现基于移动终端的行为识别存在实用性问题，即由于在线测试阶段数据（目标域数据）和离线训练阶段数据（源域数据）来自不同的传感器、不同的位置、不同的人、不同的行为类别，使得测试数据和训练数据的分布不同，从而导致离线训练的模型在实际应用中不能取得较好的结果。本文基于迁移学习的思想，对各个问题分别提出了解决方法。本文的研究结果体现在下面几个部分：

1. 面向异质加速度传感器数据的预处理方法

为解决非特定传感器行为识别问题，即由于加速度传感器参数不同导致的源域和目标域数据分布不同的问题，我们提出了基于静止状态检测和时频域变换的异质加速度传感器数据预处理方法。通过静止状态检测方法，计算每种加速度传感器静止时的合成加速度，用各自的静止合成加速度去规范化各自的加速度数据，达到统一量纲的目的，极大地减少同类行为样本在时域特征的区别。然后，采用在每秒时间窗口内进行插值的方法，使得各种传感器的采样频率相同，极大地减小同类行为样本在频域特征的差别。实验结果表明该方法能够将不同参数加速度传感器的输出值映射到统一的数据空间，从而能够实现模型在内置有异质加速度传感器的便携式终端间的迁移，达到构建非特定传感器行为识别模型的目的。

2. 基于特征迁移的非特定位置行为识别方法

为解决非特定位置行为识别问题，即不同位置加速度传感器感知到的数据分布不同的问题，我们提出了基于特征迁移的非特定位置行为识别方法。首先利用主成分分析技术，寻找能反映不同位置数据特征的主成分，生成各个位置数据的共有特征，并在离线训练阶段训练得到一个行为识别模型。然后，在线识别来自不同位置的数据，给每个样本一个预测标号及相应的置信度，并选择高置信度的样本在线增量更新已有模型，当高置信度样本中含有新位置的样本时，更新后的模型就具有了对新位置样本较好的区分能力。实验结果表明本文提出的方法能够在保持模型对已有位置数据识别精度的同时，又能提高模型对新位置数据的识别能力，达到构建非特定位置行为识别模型的目的。

3. 基于参数迁移的非特定人行为识别方法

为解决非特定人的行为识别问题，即使用模型的用户和训练模型的人员行为数据分布不同的问题，我们提出了基于参数迁移的非特定人行为识别方法。我们认为每个人的行为识别模型结构（本文指决策树模型）是相同的，不同之处在于模型的参数的值（本文指决策树非叶子结点的值），因此我们在源域和目标域之间迁移决策树及其参数，针对新用户的数据，对参数的值进行调整，使模型成为一个个性化的行为识别模型。实验结果表明该方法能够在在线阶段自动地利用使用者的数据对模型进行更新，达到针对目标用户构建个性化模型的目的。

4. 基于结构迁移的非特定类别行为识别方法

为解决非特定类别行为识别问题，即目标域中新类数据的出现而导致的数据分布不同问题，我们提出了基于结构迁移的非特定类别行为识别方法。在已有模型的情况下，对目标域数据进行分类，若某些样本的预测置信度处于异常区间时，模型可认为其是疑似新行为样本。当累计了一定数量的疑似新行为样本，系统就将其上传到云端，并发起无监督聚类算法，并请专家对聚类结果指定类标号，在此基础上，采取基于结构迁移的类增量极速学习方法更新模型，使得模型越来越能描述目标域的数据。实验结果表明该方法能够在保持模型对已有类别行为识别精度的基础上，逐渐加入对新类别行为的识别能力，满足资源受限型便携式设备的要求。

8.2 未来工作展望

基于行为识别的老年人健康监护系统通常用于长期监控，会源源不断地产生数据，最终形成“大数据”（Big Data）。从大数据中对老年人行为的识别越精确，越能准确了解和掌握老年人的身心状况，越能为他们的健康保驾护航。因此，在今后的研究过程中，我们将在下述方面展开研究：

1) 真实场景下老年人日常行为数据采集

中国已进入老龄化社会，关注老年人的身心健康成为子女和社会关注的重点。老年人的行为最能反映他们的健康状况，通过监控和分析老年人的行为，从而对他们的健康状况作出判断是我们的最终研究目标。本文由于各种条件限制仅关注了少数几种常见的行为，如静止、散步、跑步、上楼梯、下楼梯等，现实中老年人从事的行为是多种多样的，比如家务劳动、交通模式、娱乐活动、人际交往等，这些行为也都值得关注。在将来的研究工作中，我们将选择某个社区中的老人为研究对象，有计划地采集他们自然状态下的行为数据，期望建立起具有较多行为种类、较大规模的行为数据库。

2) 基于“众包”理论的可扩展行为识别系统的构建

人一生中的行为是不断变化的，这种变化不仅反映到行为类别上，对于同一种行为，在不同的人生阶段也会表现出不同的特征。因此构建一个能感知用户行为的变化，并自适应调整用户行为识别模型的系统是非常有必要的。然而，自适应调整模型，需要有最

能反映用户当前行为的标定数据，而这恰恰又是问题的难点所在，因为标定数据的代价较大，除非必要，用户不愿意去标定自己的行为数据。在今后的研究工作中，我们期望基于“众包”理论，研究能吸引用户主动参与标定数据的激励机制，建立能自动标定用户行为数据的理论和方法，当发现能反映用户当前行为特点的数据时，采用本文提出的非特定类别行为识别方法自适应地更新用户的个性化模型。

3) 特征极速增量学习方法研究

移动终端技术的发展日新月异，其中一个重要表现在于终端内置的传感器层出不穷，譬如加速度传感器、陀螺仪、电子罗盘、GPS、GSM、蓝牙、WiFi、距离传感器等等。每一个传感器的数据都具有自己的特点，能反映人们行为的某一方面，譬如加速度传感器善于衡量惯性运动，陀螺仪善于衡量旋转运动，GPS 善于检测用户的位置等等。因此在已有模型中融合越来越多的新传感器信息，能够有效地提高和扩展模型的识别能力。此外，对于同一种传感器数据，随着研究的深入，也会发现新的特征，能够帮助提高模型识别的精度。这两种情况，都是拓展了行为样本的特征空间，然而，特征空间的转变必定导致 ELM 模型结构的改变，如何根据新特征动态的调整已有模型是我们未来研究的方向之一。

4) 类别大小不均衡数据的增量学习方法研究

类别大小不均衡数据对分类器算法造成的影响，在 SVM 算法中研究得较多。研究表明，如果一个类中的样本数少，那么相比较于样本数大的类，其被错分的几率就比较大。如此一来，在行为识别的场景中，某些应用的识别精度就会得不到保证。比如跌倒检测应用，跌倒行为比正常行为要少得多，同时也重要得多，如果将跌倒行为错分为正常行为将可能导致跌倒老人不能得到及时救治，造成不可估量的损失。因此，需要研究综合考虑类别大小不均衡的增量学习算法

参考文献

- [Albinali,2010] F. Albinali, S. S. Intille, W. Haskell and M. Rosenberger. Using wearable activity type detection to improve physical activity energy expenditure estimation[C]. In Proceedings of the 12th ACM international conference on Ubiquitous computing, 2010.
- [Altun,2010] K. Altun, B. Barshan, and O. Tuncel. Comparative study on classifying human activity with miniature inertial and magnetic sensors[J]. Pattern Recognition, vol. 43, pp. 3605–3620, 2010.
- [Aziz,2009] O. Aziz, B. Lo, J. Pansiot, L. Atallah, G. Z. Yang, and A. Darzi. From computers to ubiquitous computing by 2010: Health care[J]. Phil. Trans. Roy. Soc. A: Math., Phys. Eng. Sci., vol. 366, pp. 3805–3811, May 2009.
- [Bachlin,2009] M. Bachlin, K. Forster and G. Troster. SwimMaster: a wearable assistant for swimmer[C]. In Proceedings of the 11th international conference on Ubiquitous computing (UbiComp '09). ACM, New York, NY, USA, 215-224.
- [Bao,2004] L. Bao and S. S. Intille. Activity recognition from user-annotated acceleration data[C]. In Proc. 2nd Int. Conf. Pervasive Computing, 2004, pp. 1–17.
- [Blanke,2010] U. Blanke and B. Schiele. Remember and transfer what you have learned-Recognizing composite activities based on activity spotting[C]. In Proc. 14th IEEE Int. Symp. Wearable Computers, 2010.
- [Bourke,2008] A. K. Bourke and G. M. Lyons. A threshold-based fall detection algorithm using a bi-axial gyroscope sensor[J]. Med. Eng. Phys., vol. 30, pp. 84–90, 2008.
- [Breiman,1984] L. Breiman, J. Friedman, R. Olshen, and C. Stone. Classification and Regression Trees[J]. Boca Raton, FL: CRC Press, 1984.
- [Campbell,1996] L.W. Campbell, D.A. Becker and A. Azarbayejani. Invariant features for 3D gesture recognition[C]. In Proceedings of International Conference on Automatic Face and Gesture Recognition, Killington, Vermont, USA, 1996: 157-162.
- [Carpenter,1991] G. A. Carpenter, S. Grossberg and D. Rosen. Fuzzy ART: Fast stable learning and categorization of analog patterns by an adaptive resonance system[J]. Neural Netw., vol. 4, pp. 759–771, 1991.
- [Chen,2010] Y. Q. Chen, J. Qi, Z. Sun and Q. Ning. Mining user goals for indoor location based services with low energy and high qos[J]. Computational Intelligence, 26(3):318–336, 2010.
- [Chen,2011] Y. Q. Chen, M. J. Liu, J. F. Liu, Z. Q. Shen, W. Pan. Slideshow: Gesture-aware PPT presentation[C]. ICME 2011: 1-4.
- [Chi,2004] E. H. Chi, J. Song and G. Corbin. Killer app of wearable computing: Wireless force sensing body

- protectors for martial arts[C]. In Proc. 17th Annual ACM Conf. User Interface Software and Technology, 2004, pp. 277–285.
- [Chung,2008] P. C. Chung and C. D. Liu. A daily behavior enabled hidden Markov model for human behavior understanding[J]. Pattern Recognition, 2008, 41 (5) : 1572-1580.
- [Consolvo,2006] S. Consolvo, K. Everitt, I. Smith and J.A. Landay. Design requirements for technologies that encourage physical activity[C]. In Proceedings of the CHI, pages 457–466, Montreal, Canada, April 2006.
- [Cristianini,2004] N. Cristianini 等著, 李国正等译。支持向量机导论。北京: 电子工业出版社, 2004。
- [Davies,2008] N. Davies, D. P. Siewiorek and R. Sukthankar. Activity-based computing[J]. IEEE Pervasive Comput., vol. 7, no. 2, pp. 20–21, 2008.
- [Digalakis,1995] V. Digalakis, D. Rtischev and L. Neumeyer. Speaker adaptation using constrained estimation of gaussian mixtures[J]. Speech and Audio Processing, IEEE Transactions on, vol. 3, pp. 357-366, sep 1995.
- [Fritzke,1994] B. Fritzke. Fast learning with incremental RBF networks[J]. Neural Process. Lett., vol. 1, no. 1, pp. 2–5, 1994.
- [He,2013] Q. He, T. Shang, F. Zhuang, Z. Shi: Parallel extreme learning machine for regression based on MapReduce. Neurocomputing 102: 52-58 (2013).
- [Ho,2004] J. Ho. Interruptions: using activity transitions to trigger proactive messages[D]. Massachusetts Institute of Technology, 2004.
- [Huang,2004] G.B. Huang, Q.Y. Zhu and C.K. Siew. Extreme learning machine[C]. in Technical Report ICIS/03/2004, Jan. 2004.
- [Huang,2004a] G.B. Huang, Q.Y. Zhu and C.K. Siew. Extreme learning machine: A new learning scheme of feedforward neural networks[J]. IJCNN, vol. 2, pp. 985-990, 2004.
- [Huang,2004b] G.B. Huang, P. Saratchandran and N. Sundararajan. An efficient sequential learning algorithm for growing and pruning RBF (GAP-RBF) networks[J]. IEEE Trans. Syst., Man, Cybern., B, Cybern., vol. 34, no. 6, pp. 2284–2292, Nov. 2004.
- [Huang,2006] G.B. Huang, Q.Y. Zhu and C.K. Siew. Extreme learning machine: Theory and applications[J]. Neurocomputing, vol. 70, pp. 489-501, 2006.
- [Huang,2011] G.B. Huang, D.H. Wang and Y. Lan. Extreme learning machines: a survey[J]. International Journal of Machine Learning and Cybernetics, vol. 2, no. 2, pp. 107-122, 2011.
- [Jamie,2011] A. W. Jamie, L. Paul, and W. G. Hans. Performance metrics for activity recognition[J]. ACM Trans. Intell. Syst. Technol., 2, January 2011.
- [Jin,2007] N. Jin and F. Mokhtarian. Image-based shape model for view-invariant human motion recognition[C]. In Proceedings of IEEE Conference on Advanced Video and Signal Based Surveillance, London, UK, 2007: 336-341.

- [Junker,2004] H. Junker, P. Lukowicz and G. Troster. Sampling frequency, signal resolution and the accuracy of wearable context recognition systems[C], In Proceeding of Eighth International Symposium on Wearable Computers(ISWC2004), vol. 1, pp. 176- 177, 31 Oct.-3 Nov. 2004
- [Karantonis,2006] D. Karantonis, M. Narayanan, M. Mathie, N. Lovell and B. Celler. Implementation of a real-time human movement classifier using a triaxial accelerometer for ambulatory monitoring[J]. IEEE Trans Inform Technol Biomed 10(1):156–167,2006.
- [Kallio,2006] S. Kallio, J. Kela, P. Korpipaa and J. Mantyjarvi. User independent gesture interaction for small handheld devices[J]. Int. J. Pattern Recogn. Artif. Intell., vol. 20, no. 4, pp. 505–524, 2006.
- [Kawaharaq,2005] H.S.Y. Kawahara, H.M. Hisashi and T. Aoyama. A context-aware collaborative filtering algorithm for real world oriented content delivery service[C]. In: Proceedings of ubiPCMM,2005.
- [Kawaharaq,2007] Y. Kawaharaq, H. Kurasawa and H. Morikawa. Recognizing user context using mobile handsets with acceleration sensors[C]. In: IEEE international conference on portable information devices (PORTABLE'07), pp 1–5.
- [Kawaharaq,2009] Y. Kawaharaq, N. Ryu and T. Asami. Monitoring daily energy expenditure using a 3-axis accelerometer with a low-power microprocessor[J]. Int J Hum Comput Interact 1(5):145–154,2009.
- [Kim,2008] K. Kim and G.G. Medioni. Distributed visual processing for a home visual sensor network[C]. In Proceedings of IEEE Workshop on Applications of ComputerVision , CopperMountain, Colorado, USA, 2008:1-6.
- [King,2010] R. C. King, L. Atallah, C. Wang, F. Miskelly, and G. Z. Yang. Elderly risk assessment of falls with BSN[C]. In Proc.of 2010 Int. Conf. Body Sensor Networks, 2010, pp. 30–35.
- [Kunze,2005] K. Kunze, P. Lukowicz, H. Junker and G. Troster. Where am i: Recognizing on-body positions of wearable sensors[C]. Lecture Notes in Computer Science, vol. 3479, pp. 257-268. Springer Berlin / Heidelberg(2005).
- [Kunze,2008] K. Kunze and P. Lukowicz. Dealing with sensor displacement in motion-based onbody activity recognition systems[C]. In: Proceedings of the 10th international conference on Ubiquitous computing, UbiComp '08, pp. 20-29.ACM, New York, NY, USA (2008).
- [Kunze,2008] K. Kunze, P. Lukowicz, K. Partridge, B. Begole: Which Way Am I Facing: Inferring Horizontal Device Orientation from an Accelerometer Signal. ISWC 2009: 149-150.
- [Lai,2010] C.F. Lai, Y.M. Huang, J.H. Park and H.C. Chao. Adaptive body posture analysis for elderly-falling detection with multisensors[J]. Intelligent Systems, IEEE 25(2), 20 -30(2010).
- [Larson,2012]E. C. Larson et al.. SpiroSmart: Using a Microphone to Measure Lung Function on a Mobile Phone[C]. UbiComp 2012.
- [Lee,2007] J. W. Lee and C. Giraud. Transfer learning in decision trees[C]. In Proceedings of the International Joint Conference on Neural Networks, 2007.
- [Leggetter,1995] C. Leggetter and P. Woodland. Maximum likelihood linear regression for speaker

- adaptation of continuous density hidden markov models[J]. *Computer Speech and Language*, vol. 9, no. 2, pp. 171-185, 1995.
- [Lester,2006] J. Lester, T. Choudhury and G. Borriello. A practical approach to recognizing physical activities[C]. In *Proc.of Pervasive Computing*, 2006, pp. 1–16.
- [Lester,2009] J. Lester, C. Hartung, L. Pina, R. Libby, G. Borriello and G. Duncan. Validated caloric expenditure estimation using a single body-worn sensor[C]. In *Proceedings of the 11th international conference on Ubiquitous computing (UbiComp '09)*.
- [Liang,2006] N.Y. Liang, G.B. Huang, P. Saratchandran and N. Sundararajan. A fast and accurate online sequential learning algorithm for feedforward networks[J]. *Neural Networks, IEEE Transactions on*, vol. 17, pp. 1411-1423, nov. 2006.
- [Lim,1995] C. P. Lim and R. F. Harrison. Probabilistic fuzzy artmap: An autonomous neural network architecture for Bayesian probability estimation[C]. In *Proc. of 4th Int. Conf. Artif. Neural Netw.*, vol. 409, pp. 148–153, Jun.1995.
- [Liu,2004] C.D. Liu, P.C. Chung and Y.N. Chung. Human home behavior interpretation from video streams[C]. In: *Proceedings of the 2004 IEEE International Conference on Networking, Sensing & Control* [C], Taipei, Taiwan, China, 2004: 192-197.
- [Lu,2012]H. Lu et al. StressSense: Detecting Stress in Unconstrained Acoustic Environments using Smartphones[C]. *UbiComp 2012*.
- [Lukowicz,2004a] P. Lukowicz, H. Junker and G. Troster. Automatic calibration of body worn acceleration sensors[J]. *PerCom2004*.
- [Lukowicz,2004b] P. Lukowicz and J.A Ward. Recognizing workshop activity using body worn microphones and accelerometers[C]. In *2nd International Conference on Pervasive Computing*, Vienna, Austria, Springer, 2004, pp.18-22.
- [Lukowicz,2010a] P. Lukowicz et al. Recording a complex, multi modal activity data set for context recognition[C]. In *Proceeding of 1st Workshop on Context-Systems Design, Evaluation and Optimisation at ARCS*, 2010, 2010.
- [Lukowicz,2010b] P. Lukowicz, O. Amft, D. Roggen and J. Cheng. On-body sensing: From gesture-based input to activity-driven interaction[J]. *IEEE Comput.*, vol. 43, no. 10, pp. 92–96, 2010.
- [Lymberopoulos,2009]D. Lymberopoulos, A. Bamis, A. Savvides. A Methodology for Extracting Temporal Properties from Sensor Network Data Streams[C]. *MobiSys'09*. P193-206.
- [Mann,1996] S. Mann. Smart clothing: The shift to wearable computing[J]. *Commun. ACM*, vol. 39, no.8, pp. 23–34, 1996.
- [Mann,1998] S. Mann. Humanistic computing: Wearcom as a new framework and application for intelligent signal processing[C]. In *Proc. IEEE*, vol. 86,no. 11, pp. 2123–2151, 1998.
- [Mantyla,2001] J. Mantyla, J. Himberg and T. Seppanen. Recognizing human motion with multiple

- acceleration sensors[C], In IEEE International Conference on Systems, Man and Cybernetics, Tucson, USA, IEEE Press, 2001, pp.747-752.
- [Mathie,2003] M. Mathie. Monitoring and interpreting human movement patterns using a triaxial accelerometer[D]. Ph.D. thesis, University of New South Wales, 2003.
- [McIlwraith,2009] D. McIlwraith and G. Z. Yang. Body sensor networks for sport, wellbeing and health[J]. In Sensor Networks, G. Ferari, Ed. Berlin, Germany: Springer-Verlag, 2009, pp. 349–381.
- [Meng,2011] X. Meng, S. Sun, L. Ji, J. Wu, and Wai-Choong Wong. Estimation of Center of Mass Displacement based on Gait Analysis[C]. 8th International Conference on Body Sensor Networks (BSN 2011), 23-25 May 2011, Dallas, Texas, USA.
- [Mitchell,2003] T. M. Mitchell(卡内基梅隆大学)著, 曾华军等译。机器学习。北京: 机械工业出版社, 2003.
- [Mizell,2003] D. Mizell. Using gravity to estimate accelerometer orientation[C]. In: Proceedings Of the 7th IEEE International Symposium on Wearable Computers(ISWC) 2003, USA(2003)
- [Nachman,2010] L. Nachman, A. Baxi, S. Bhattacharya, V. Darera, P. Deshpande, N. Kodlapura, V. Mageshkumar, S. Rath, J. Shahabdeen and R. Acharya. Jog falls: A pervasive healthcare platform for diabetes management[C]. In Proceedings of the Pervasive, pages 94–111, 2010.
- [Nham,2008] B. Nham, K. Siangliulue and S. Yeung. Predicting mode of transport from iphone accelerometer data[M]. CS 229: Machine Learning Final Projects, Stanford University, 2008.
- [Ohgi,2000] Y. Ohgi, M. Yasumura, H. Chikawa and C. Miyaji. Analysis of stroke technique using acceleration sensor IC in freestyle swimming[J], The Engineering of SPORT, Blackwell Publishing, 2000, pp.503-511.
- [PanGang,2010] G. Pan, J. Wu, D. Zhang, Z. Wu, Y. Yang, S. Li: GeeAir: a universal multimodal remote control device for home appliances. Personal and Ubiquitous Computing 14(8): 723-735 (2010)
- [Pan,2010] S. J. Pan and Q. Yang. A survey on transfer learning[J]. IEEE Transactions on Knowledge and Data Engineering, 22(10):1345–1359, 2010.
- [Park,2005] S. Park and M. Trivedi. Driver activity analysis for intelligent vehicles: issues and development framework[C]. In: Proceedings of IEEE Intelligent Vehicles Symposium, LasVegas, Nevada, USA, 2005: 644-649.
- [Pentland,2004] A. Pentland. Healthwear: Medical technology becomes wearable[J]. Computer, vol. 37, no. 5, pp. 42–49, 2004.
- [Platt,1991] J. Platt. A resource-allocating network for function interpolation[J]. Neural Comput., vol. 3, pp. 213–225, 1991.
- [Ravi,2005] N. Ravi, N. Dandekar, P. Mysore and M. Littman. Activity recognition from accelerometer data[C]. In proceedings of the 17th Annual Conference on Innovative Applications of Artificial Intelligence, Pittsburgh, USA, AAAI Press, 2005, pp.141-1546.

- [Reddy,2008] S. Reddy, J. Burke, D. Estrin, M. Hansen and M. Srivastava. Determining transportation mode on mobile phones[C]. In IEEE International Symposium on Wearable Computers (ISWC 2008). Pittsburgh, Pennsylvania, September 2008, pp. 25-28 (2008)
- [Robertson,2005] N. Robertson and I. Reid. Behavior understanding in video: a combined method. In: Proceedings of IEEE International Conference on Computer Vision[C], Beijing, China, 2005: 808-815.
- [Roggen,2009] D. Roggen, K. Forster, A. Calatroni, T. Holleczeck, Y. Fang, G. Troster, P. Lukowicz, G. Pirkel, D. Bannach, K. Kunze, A. Ferscha, C. Holzmann, A. Riener, R. Chavarriaga, J. del and R. Millan. Opportunity: Towards opportunistic activity and context recognition systems[C]. In Proc. Of 3rd IEEE WoWMoM Workshop Autonomic and Opportunistic Communications, 2009.
- [Roggen,2010] D. Roggen, A. Calatroni, M. Rossi, T. Holleczeck, K. Forster, G. Troster, P. Lukowicz, D. Bannach, G. Pirkel, A. Ferscha, J. Doppler, C. Holzmann, M. Kurz, G. Holl, R. Chavarriaga, M. Creatura and R. Millan. Collecting complex activity data sets in highly rich networked sensor environments[C]. In Proc. 7th Int. Conf. Networked Sensing Systems, New York: IEEE Press, 2010, pp. 233–240.
- [Santos,2009] A.C. Santos, L. Tarrataca, J.M.P. Cardoso, D.R. Ferreira, P.C. Diniz and P. Chainho. Context inference for mobile applications in the UPCASE project[C]. In Proceedings of the 2nd international conference on mobile wireless middleware, operating systems, and applications (MOBILWARE 2009). Springer, no. 7 in LNICST, pp 352–365.
- [Shi,2012] Y. Shi, Y. Shi, X. Wang: Fall Detection on Mobile Phones Using Features from a Five-Phase Model. UIC/ATC 2012: 951-956
- [Stanford,2002] V. Stanford. Wearable computing goes live in industry[J]. IEEE Pervasive Comput. Mag., vol. 1, no. 4, pp. 14–19, 2002.
- [Starner,1998] T. Starner, J. Weaver and A. Pentland. Real-time American sign language recognition using desk and wearable computer based video[J]. IEEE Trans. Pattern Anal. Machine Intell., vol. 20, no. 12, pp. 1371–1375, 1998.
- [Stiefmeier,2008] T. Stiefmeier, D. Roggen, G. Ogris, P. Lukowicz and G. Troster. Wearable activity tracking in car manufacturing[J]. IEEE Pervasive Comput. Mag., vol. 7, no. 2, pp. 42–50, 2008.
- [Stikic,2008] M. Stikic, K.V. Laerhoven and B. Schiele. Exploring semi-supervised and active learning for activity recognition[C]. In Proceedings of the IEEE Int. Symp. on Wearable Computers, 2008.
- [Sun,2010] L. Sun, D. Zhang, B. Li, B. Guo and S. Li. Activity recognition on an accelerometer embedded mobile phone with varying positions and orientations[C]. In Proceedings of the 7th international conference on Ubiquitous intelligence and computing, UIC'10, pp. 548-562. Springer-Verlag, Berlin, Heidelberg (2010)
- [Tabbone,2006] S. Tabbone, L. Wendling and J.P.A Salmon. new shape descriptor defined on the Radon transform[J]. Computer Vision and Image Understanding, 2006, 102 (122) : 42-51.
- [Tognetti,2006] A. Tognetti, N. Carbonaro, G. Zupone and D. Rossi. Characterization of a novel data glove

- based on textile integrated sensors[C]. In Proc. Of 28th IEEE Int. Conf. Engineering in Medicine and Biology Society, 2006, pp. 2510–2513.
- [Torrey,2005] L. Torrey, T. Walker, J. Shavlik and R. Maclin. Using advice to transfer knowledge acquired in one reinforcement learning task to another[C]. In Proceedings of the Sixteenth European Conference on Machine Learning, pages 412–424, 2005.
- [Turaga,2007] P. K. Turaga, A. Veeraraghavan and R. Chellappa. From videos to verbs: mining videos for activities using a cascade of dynamical systems[C]. In Proceedings of Conference on Computer Vision and Pattern Recognition, Minneapolis, Minnesota, USA, 2007: 1-8.
- [URL1]BeWell. Website: <https://play.google.com/store/apps/details?id=org.bewellapp>
- [URL2] Cardio trainer. Website <http://www.worksmartlabs.com/cardiotrainer/about.php>
- [URL3] Sleep cycle alarm clock. Website: <http://www.lexwarelabs.com/sleepcycle/>
- [URL4] Nokia step counter. Website <http://betalabs.nokia.com/betas/view/nokia-step-counter>
- [URL5] Nokia sports tracker. Website <https://sportstracker.nokia.com/nts/main/index.do>
- [Wang,2007] Y. Wang, K. Huang and T.N. Tan. Abnormal activity recognition in office based on R transform[C]. In Proceedings of IEEE Conference on Image Processing, San Antonio, TX, USA, 2007: I-341-344.
- [Wang,2009] X.L. Wang. High accuracy distributed target detection and classification in sensor networks based on mobile agent framework[D]. Phd thesis, University of Tennessee,2009.
- [Wang,2010] S. Wang, C. Chen and J. Ma. Accelerometer based transportation mode recognition on mobile phones[C]. In Proceedings of 2010 Asia-Pacific Conference on Wearable Computing Systems (APWCS) , pp. 44 -46 (2010).
- [Wang,2012] Z. Wang, M. Jiang, Y. Hu and H. Li. An Incremental Learning Method Based on Probabilistic Neural Networks and Adjustable Fuzzy Clustering for Human Activity Recognition by Using Wearable Sensors[J]. Information Technology in Biomedicine, IEEE Transactions on , vol.16, no.4, pp.691-699, July 2012.
- [Want,2009] R. Want. When cell phones become computers[J]. IEEE Pervasive Comput., vol. 8, no. 2, pp. 2–5, 2009.
- [Ward,2006] J. Ward, P. Lukowicz, G. Troster and T. Starner. Activity recognition of assembly tasks using body-worn microphones and accelerometers[J]. IEEE Trans. Pattern Anal. Machine Intell., vol. 28, no. 10, pp. 1553–1567, 2006
- [Ward,2011] J.A. Ward, P. Lukowicz and H.W. Gellersen. Performance metrics for activity recognition[J]. ACM Trans. Intell. Syst. Technol., 2, January 2011.
- [Welbourne,2005] E. Welbourne, J. Lester, A. LaMarca and G. Borriello. Mobile context inference using low-cost sensors[C]. In Proceedings of the first international workshop on Location- and Context-Awareness (LoCA 2005), Springer, LNCS, vol 3479, pp 254–263.

- [Wolf,2007]K.H. Wolf, A. Lohse, M. Marschollek, R. Haux. Develeopment of a Fall Detector and Classifier based on a Triaxial Accelerometer Demo Board[C]. In Proceedings of UbiComp 2007, 2007, pp. 210-213.
- [Yang,2009] J. Yang. Toward physical activity diary: motion recognition using simple acceleration features with mobile phones[C]. In Proceedings of the 1st International Workshop on Interactive Multimedia for Consumer Electronics(IMCE '09), Beijing, China, pp. 1-10 (2009)
- [Yang,2010] J. Yang, H. Lu, Z. Liu and P. Boda. Physical Activity Recognition with Mobile Phones: Challenges, Methods and Applications[M]. invited book chapter for the Springer book entitled "Multimedia Interaction and Intelligent User Interfaces: Principles, Methods and Applications", 2010.
- [Zhang,2006] T. Zhang, J. Wang, P. Liu and J. Hou. Fall Detection by Embedding an Accelerometer in Cellphone and Using KFD Algorithm[J]. International Journal of computer Science and Network Security, Vol.6(10), 2006.
- [Zhao,2010] Z. Zhao, Y. Chen and J. Liu. Fall Detecting and Alarming Based on Mobile Phone[C]. In Proceedings of 2010 Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, (2010),494-497.
- [Zhao,2011] Z. Zhao, Y. Chen, J. Liu, Z. Shen and M. Liu. Cross-People Mobile-Phone Based Activity Recognition[C]. the International Joint Conference on Artificial Intelligence(IJCAI2011), Barcelona, Spain, July 16-22, 2011.
- [Zheng,2009] V. W. Zheng, D. H. Hu and Q. Yang. Cross-domain activity recognition[C]. In Proceedings of Ubiquitous computing, pages 61-70. ACM, 2009.
- [苗强,2009]苗强,周兴社,於志文,倪红波.一种非觉察式的睡眠行为识别技术研究[A]. NCMT2009、CHCI2009、PCC2009 论文集[C]. 2009。
- [倪训博,2010]倪训博,赵德斌,姜峰,程丹松。Viterbi 和 DTW 算法的关系分析——在非特定人手语识别中的应用[J]。计算机研究与发展。2010(02)。
- [石欣,2010]石欣,熊庆宇,雷璐宁。基于压力传感器的跌倒检测系统研究[J]。仪器仪表学报。2010(03)。
- [时岳,2013]时岳,喻纯,史元春。基于旋转模式的移动设备佩戴位置识别方法[J]。软件学报。2013(08)。
- [王宏安,2011]王宏安,田丰,戴国忠。基于生理计算的人机交互[M]。中国计算机学会通讯。2011(8)。
- [徐光佑,2009]徐光佑,曹媛媛。动作识别与行为理解综述[J]。中国图象图形学报。2009, 14(2): 189-195。
- [徐经纬,2012]徐经纬。基于无线传感器的人体行为识别研究[D]。南京大学。2012。
- [薛洋,2011]薛洋。基于单个加速度传感器的人体运动模式识别[D]。华南理工大学。2011。
- [岳俊峰,2004]岳俊峰,胡望社。面临人口老龄化的居住区设计[J]。重庆工业高等专科学校学报。2004(06)。

致 谢

自二零零九年九月踏入我无限向往的中国科学院计算技术研究所以来，匆匆已是四载有余。日子一天天过去，我能感觉到计算所人的精神和他们所研究的先进技术在一点一点向我脑海中迁移，感觉自己如雨后春笋般获得了快速成长。往事恍如昨日，一切尽在心头，师生的感情，同学的情谊，那些人，那些事，陪伴了我的成长，丰富了我的人生。几年来，感受到了大家对我无私的帮助和无微不至的关怀，在此请允许我发自内心地向你们表达我的谢意。

感谢我的恩师陈益强研究员，是您带我进入了心目中的中国计算机的圣堂，是您引导我找到了科研的方向，是您给我提供了宽松的科研环境，是您在科研的道路上送了我一程又一程。“桃李不言，下自成蹊”，您的言传身教已为学生树立了榜样，我也将秉持这种精神投入到将来的教学科研中去。您渊博的专业知识，敏捷又严密的思维，对科研大局准确的把握以及对研究方向的深刻理解，都让学生我钦佩不已，都激励着学生我奋起直追。感谢您几年来对我的帮助，期待在今后的教学科研道路上能继续接受您的指导。

在刘军发博士的具体带领下，我接触了基于 WiFi 信号的室内定位研究，参与了 Brain Fiber 分类的竞赛，发表了跌倒检测的 Demo Paper，人生第一次真正地用 Machine Learning 解决实际的问题。军发博士的敦厚宽容，多才多艺，都值得我好好学习。对军发博士给予我生活上的帮助表示衷心的感谢，您帮我度过了曾经的困难时期。感谢王双全博士的精心指导。王博丰富的经历，对人生的领悟，启迪了我曾经的迷茫，有拨云见日的感觉。王博对科研的孜孜不倦，精益求精，帮助我在具体的研究过程中不断提高。王博略带川韵的普通话，总能将研究的问题阐述得清清楚楚，博得阵阵掌声，是我学习的好榜样。感谢纪雯博士，您真诚的鼓励不止一次地温暖了我。感谢工程师张绘国、颜庆聪和邢云冰师兄，从与你们的交谈中总能得到巨大的收获。

感谢计算所普适计算中心感知计算研究课题组这个优秀的团队，为我提供了良好、宽松的学习氛围。感谢实验室同学：陈振宇、高兴宇、何文静、张博宁、于汉超、谷洋、黄美玉、忽丽莎、刘明杰、唐晓庆、胡琨、杨华、唐焱、周敏、陈涛、蒋鑫龙、尹苓琳、陈超、范玉广、张静、肖强、黄陈、刘振中、刘定俊、贾立平、李啸海和其他同学的帮助，在一起学习生活的日子是人生不可多得的美好回忆。感谢雷俊老师、赵雪艳老师在工作中给予的关照和帮助。

感谢香港科技大学的杨强老师，新加坡南洋理工大学的申志奇教授和黄广斌教授，他们作为机器学习领域的知名专家教授，给予了我耐心的指导，从他们那里受益良多。感谢新加坡南洋理工大学的于涵博士，帮我为论文加工润色。

感谢我的室友诸峰，我的同学赵国光、张博、许营坤、熊海泉等，能与你们相遇相

识，共同战斗，共度美好时光，是我一生中最美的回忆。

最后要特别感谢我的爱人马倩女士，你承担了家庭的全部负担，在我几乎没有收入的情况下，无怨无悔地给予我精神和物质的支持，使我没有后顾之忧，得以顺利完成学业。

赵中堂

2013 年 11 月

作者简介及在学期间发表的学术论文与研究成果

作者简介:

姓名: 赵中堂 性别: 男 出生日期: 1978.9.26 籍贯: 河南浚县

2009.9 – 2013.12	中科院计算技术研究所	计算机应用技术专业	博研
2004.9 – 2009.8	郑州航院	计算机科学与技术系	教师
2001.9 – 2004.7	郑州大学	计算机软件与理论专业	硕研
1997.9 – 2001.7	郑州大学	计算机科学与技术	本科

攻读博士学位期间发表的论文

会议论文

- [1] Zhao, Z., Chen Y., Liu, J., Shen, Z., Liu, M. “Cross-People Mobile-Phone Based Activity Recognition” the International Joint Conference on Artificial Intelligence(IJCAI2011), Barcelona, Spain, July 16–22, 2011. (人工智能领域顶级会议)
- [2] Zhao, Z., Chen Y., Liu, J. “Fall Detecting and Alarming Based on Mobile Phone”. 2010 Symposia and Workshops on Ubiquitous, Autonomic and Trusted Computing, (2010), Xi'an, China, October 26-29, 2010, Pages 494–497. (EI 检索)
- [3] Zhao, Z., Chen Y., Wang S., Chen Z. “FallAlarm: Smart Phone Based Fall Detecting and Positioning System”. Procedia CS 10: 617–624 (2012). (ISTP 检索)

期刊论文

- [1] Chen Y., Zhao, Z., Wang S., Chen Z. “Extreme learning machine based device displacement free activity recognition model”. Soft Comput. 16(9): 1617–1625 (2012). (SCI 检索)
- [2] Liu, J., Chen Y., Liu, M., Zhao Z. “SELM: Semi-supervised ELM with application in sparse calibrated location estimation”, Neurocomputing, Volume 74, Issue 16, September 2011, Pages 2566–2572. (SCI 检索)
- [3] Zhao, Z., Chen Z., Wang S., Chen Y. “A Class Incremental Extreme Learning Machine for Activity Recognition”. (首先在国际会议 ELM2013 上交流, 然后被大会推荐到 SCI 检索期刊 Cognitive Computation, 已接收)

攻读博士学位期间参加的科研项目

- [1] 2011/01-2013/12 北京市自然科学基金项目“基于多模传感的老人摔倒检测模型研究”。
负责摔倒模型的构建。
- [2] 2012/01-2015/12 国家自然科学基金面上项目“高动态室内无线环境中渐进式自适应定位方法研究”。
负责免标定协同算法研究。
- [3] 2011/01-2013/12 国家自然科学基金青年基金项目“面向海量数据处理的多核学习算法及其并行优化方法研究”。
负责优化方法研究。

攻读博士学位期间的获奖情况

- [1] 2011 年中国科学院研究生院“三好学生”