

融合多模信息感知的低功耗行为识别*

齐娟^{1,2}, 陈益强¹, 刘军发¹⁺, 孙卓^{1,2}

¹(中国科学院 计算技术研究所 普适计算中心,北京 100190)

²(中国科学院 研究生院,北京 100049)

Power-Efficient Activity Recognition Based on Multi-Modal Information Sensing and Fusion

QI Juan^{1,2}, CHEN Yi-Qiang¹, LIU Jun-Fa¹⁺, SUN Zhuo^{1,2}

¹(Institute of Computing Technology, The Chinese Academy of Sciences, Beijing 100190, China)

²(Graduate University, The Chinese Academy of Sciences, Beijing 100049, China)

+ Corresponding author: E-mail: liujunfa@ict.ac.cn

Qi J, Chen YQ, Liu JF, Sun Z. Power-Efficient activity recognition based on multi-modal information sensing and fusion. Journal of Software, 2010,21(Suppl.):39-50. <http://www.jos.org.cn/1000-9825/10005.htm>

Abstract: Activity recognition has many potential applications in pervasive computing field, such as medical care, intelligent home/office, business service, etc. Recently, sensor based activity recognition has attracted much attention, due to its advantages of ubiquity and less intrusion. This paper studies sensor based activity recognition. In particular, it recognizes high-level user activities from multi-mode sensor data using machine learning methods. This paper proposes an activity recognition framework, which makes use of low-level wireless network signal and acceleration data to answer questions like “where is the user”, “what is the user doing” and “what is the user going to do”. Three main fusion algorithms are included in this framework. Considering the trend of pervasive computing, this paper uses the mobile phones and their embedded sensors to collect activity information. Effectiveness of the proposed algorithms is confirmed on real world collected dataset.

Key words: activity recognition, sensor, machine learning, information fusion

摘要: 行为识别在普适计算领域有着极大的应用前景,可广泛应用于医疗监护、智能家居/办公、商业服务等方面.其中基于传感器的行为识别因其分布范围广、不具侵扰性等优点,已成为目前的研究热点之一.采用机器学习理论和方法,提出了一种基于多模传感信息感知和融合的行为识别层次框架.该框架综合加速度和无线网络信号两种传感器信息、利用多种基于融合的认识方法,能同时解决“用户在哪里”、“用户在做什么”、“用户将要去做什么”等行为相关问题.采用智能手机作为实验平台,利用其内置的多种传感器收集用户的行为信息,更符合普适计算的发展趋势.最后通过实际采集的数据和大量的实验说明了各种方法的有效性.

关键词: 行为感知;传感器;机器学习;信息融合

* Supported by the National Natural Science Foundation of China under Grant No.90820303 (国家自然科学基金); the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z305 (国家高技术研究发展计划(863)); the Co-building Program of Beijing Municipal Education Commission of China (北京市教育委员会共建项目)

Received 2009-07-15; Accepted 2010-07-09

普适计算技术力图把以计算机为中心的计算转变为以人为中心的计算,使得计算机能够在无形中为用户提供智能的、透明的服务^[1].为了提供适时适地的服务,计算机必须能够了解周围的环境以及与用户的交互状态等,即具有上下文感知计算的能力.因此上下文感知计算是普适计算环境中不可或缺的技术,只有通过它才可以为用户提供及时的切合需要的服务^[2].而行为感知技术是感知计算的一个重要分支,重点对用户行为进行感知和识别,在智能和个性化服务方面发挥着重要作用,如医疗监护、智能家居/办公、商业服务、社交领域等.在医疗监护领域,可以用于老人、小孩、残疾人士的监护以及病情的趋势判断等^[3].在智能家居/办公领域,行为感知的应用包括根据用户的日常行为习惯自动地调整环境,进行事件提醒等^[4,5].

随着电子与通信技术的发展,传感器开始覆盖人类生活的各个方面,例如无线接入点、射频发生器、加速度计.这些传感器能够在无形中记录人们的行为轨迹,基于这些传感器的行为识别具有分布范围广、不具侵扰性等优点,已成为目前研究的热点^[6-12].

基于传感器的行为识别多采用传感器技术和数据挖掘/机器学习算法相结合的方法,从底层的传感器信息推断高层的用户行为.但人类的行为各式各样,一种合理的表示方式是从底层细节行为过渡到高层泛化行为的层次结构.位于层次结构最底层的是短暂的运动状态(静止、行走、跑动等),这些运动状态包含了较多的行为细节信息;运动状态的不同排列组合又构成诸如冲咖啡、交谈等稍长的行为;这些行为进一步构成高层的泛化行为(在家、工作、购物等).单一类型的传感器,只能提供用户行为某方面的信息,导致其识别只局限于某种类型的行为.单纯地依赖某一种传感器不能全面有效地识别用户行为,如果能有效融合各种传感器信息则能够充分利用各种传感器的长处,从而可以识别各个层次的用户行为并达到更好的识别效果.

电量对于各种移动设备和传感器而言是稀缺资源,为了提供无处不在的服务,行为识别还需要尽量降低电量的消耗,以延长可服务时间.低功耗是行为识别需要考虑的衡量指标之一,而以往的工作只关注于识别度的提高,忽略了低功耗的考虑.

因此,本文主要对基于多模传感信息感知和融合的低功耗行为识别展开研究.本文第 1 节介绍行为识别的研究现状,并进一步分析依旧面临的问题.第 2 节提出基于多模信息感知与融合的行为识别解决框架和几种融合算法.第 3 节基于实际数据对各种方法的有效性进行验证.第 4 节对整篇论文进行总结,并给出将来的研究方向.

1 研究现状分析

行为识别的早期工作主要集中在视频分析方面,通过在环境中布置摄像头并使用计算机视觉和图像处理技术对人物进行跟踪和行为识别.然而基于计算机视觉的行为识别具有易受光照等因素影响、数据量大、涉及隐私等缺点,不能被用户普遍接受.基于传感器的行为识别成为目前研究的热点,具有应用范围广、不具侵扰性等优点.基于传感器的行为识别方法目前主要分为 3 类:基于加速度的方法、基于位置信息的方法和基于物品使用(object usage)的方法.

基于加速度的行为识别主要通过通过对加速度计信号的分析来推测用户行为.三维加速度计可以测量 x, y, z 三个方向上加速度值,可进一步用于进行运动探测或者人体姿态感知.这类工作中最为全面详细的为文献^[1],实验者在身上 5 个不同的部位绑上加速度计,并从事不同行为,同时记录加速度值.然后使用分类器算法对用户行为识别,实验结果表明决策树方法表现最好,准确率大约为 84%.在用户身上多个部位绑上加速度计会给用户造成诸多不便,不符合实际应用的要求.不同部位的加速度在有着不同的效果,其中大腿和腰部在行为识别上的效果最好^[13].因此很多研究工作开始考虑减少加速度计的数目,文献^[8]使用一个放置在腰间的加速度计来识别站立、走路、跑动、上下楼等 8 类行为.他们将行为识别问题看作是一个分类问题,实验分析了各种分类器算法以及组合分类器方法的性能,指出利用 Plurality Voting 组合分类器的方法表现最优,能够达到 90% 以上的准确率.与之类似,目前已有的这类工作需要用户在用户身体的不同部位绑定特定的装置,并不符合普适计算的应用需求.并且基于加速度的方法对于识别细节行为状态如站立、行走等比较有效,但识别上层的行为还未成熟应用.

基于位置信息的行为识别利用行为与其所在的位置的联系,通过定位技术来识别一些泛化的行为(如“在

家”、“在公司”、“上班”)。例如,基于 GPS 数据判断和预测用户的交通状态^[12]、识别出对用户有重要意义的地点和发生在这些地点的行为^[14-16]。基于室内定位技术,也出现了一些行为识别的研究工作^[10,11],他们利用 WiFi 无线网络推断高层的用户目标,能够达到 70%~90%的目标识别率。然而基于位置的行为识别往往对位置和行为之间做强行的约束,如在办公室对应工作,在会议室对应开会。然而位置和行为未必是唯一对应关系,例如在办公室也可以有会客、打电话等多种行为。

基于物品使用的行为识别基于一种假设:行为可以通过用户所操作的一系列物品来表示,如烹饪的相关行为往往涉及拿起或者放下某些餐具。这类识别方法多在物品贴上 RFID 标签,用户则佩戴着一个 RFID 读取器,这样用户所接触的物品就可以通过 RFID 读取器记录下来。然而基于物品使用的行为识别方法对于那些并不需要接触到物品的很多行为(如散步)是无效的。另外,这种方法只能在极小的范围内发挥作用,因为目前在行为场所涉及的所有物品上安置 RFID 成本较高,部署起来也相对麻烦。

多模信息融合指综合不同类型的信息来达到比单一信息更好的性能。该技术在计算机视觉领域得到了广泛的应用,多项融合音视频数据的方法被提出。在定位领域,全球定位系统(GPS)和惯性导航系统(INS)的有效结合是信息融合最为直观的例子。在行为识别领域,多模信息融合的技术也越来越受到重视。文献[14]提出了一种结合 RFID 和加速度计进行行为识别的算法,即优先使用 RFID,当其失效时再使用加速度计。但这只是一种简单的相加融合方法,并没有充分利用传感器信息之间的相互关系。

综上所述,目前基于传感器的行为识别研究还有以下问题:

- 忽略了多种传感器信息的综合利用。单类型的传感器只能提供用户行为某方面的信息,因此只对一种传感信息进行分析不能全面地识别行为。即使有部分研究工作开始考虑信息融合的方法,也只是采取了简单的相加思想,没有考虑各种信息之间的相关性。
- 不恰当的设备选择对用户造成不便。目前的研究多采用在用户身上绑置特定装置的方法来采集数据,对用户的生活带来不必要的麻烦。采用手机、PDA 等用户经常携带的设备,并直接利用其内置的各种传感器将是理想选择,也是普适计算的发展趋势。
- 忽略了低功耗的需求。以往的工作往往只关注于行为识别的精度,然而电量对于手持设备而言是稀缺的资源,低功耗能够提供更长的服务时间,这也是服务质量重要的指标之一。

2 解决方案

目前对行为识别的研究取得了一些进展和研究成果,但依然存在很多问题需要研究解决。这一方面是由于人类行为本身的复杂性,种类繁多很难定义;另一方面是因为底层传感器信号到高层行为信息之间有较大的间隔,使得识别具有挑战性。通过对相关工作的分析,本文提出了基于多模信息感知与融合的行为识别解决方案。该方案能有效识别各个层次的用户行为,同时解决以下几个问题:用户在哪里、用户在做什么、用户将要去做什么。

为此本文提出了如图 1 所示的层次模型,该模型共包括 4 层,每层间通过一个推断模型连接。这 4 层分别是:

- 加速度。手机中内置的加速度计能够测量三维加速度,进一步用于识别不同的运动状态。在这一层我们提出了第 1 种融合算法 MotRec:基于位置约束的用户运动状态感知方法。该方法基于加速度识别运动状态,并通过分析位置信息和运动状态信息之间的相关性进一步优化,从而提高识别精度。
- 无线网络信号。从各个无线接入点接收到信号,无线网络信号可以用于位置感知。在这一层我们提出了第 2 种融合算法 PowerLoc:基于关键点检测的低功耗用户位置感知方法。通过结合加速度信息和无线网络信号,减少不必要的定位次数,从而能够在较大范围内实现用户位置感知且降低设备的电量消耗。
- 目标,即用户将要去做什么。在这一层我们提出了第 3 种融合算法 GoalRec:基于多模信息融合的行为预测方法。GoalRec 通过历史记录建立行为预测模型,并将其用于实时行为识别,综合利用加速度和无线网络信号两种传感器信息推断用户的目标行为。
- 活动序列。活动被定义为用户所在的位置及其运动状态(如在走廊走动,在办公桌前静坐等)。通过引入活

动序列作为中间层来降低推断算法的复杂性.

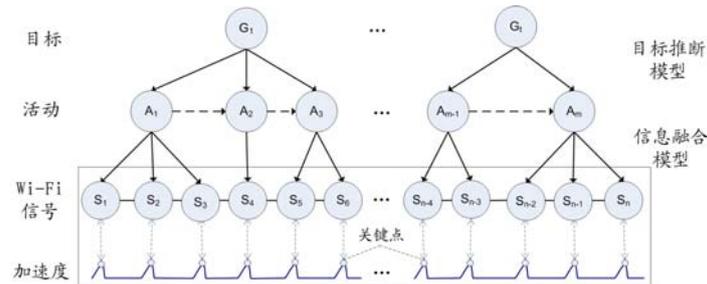


图 1 行为识别结构模型

上面给出了行为识别结构模型的整体框架,接下来本文对各种融合方法进行详细介绍.

2.1 基于位置约束的用户运动状态感知方法

该方法通过手机内置加速度计采集到的加速度数据来识别静止、行走、跑动、上/下楼梯、上/下电梯 7 种运动状态.依赖手机内置的加速度计来采集数据,虽然方便用户使用,但给行为识别算法带来了困难.一方面不能确定设备在用户身上的哪个部位,另一方面不能确定设备的朝向,因此无法确认内置加速度计的朝向位姿.以上两个方面被称为手机和用户之间存在相对运动,它使得加速度计的坐标系并不是固定的,会导致行为识别准确率的下降.为解决上述问题,本文提出了基于位置约束的用户运动状态感知方法 MotRec.算法流程包括加速度数据采集、相对运动消除、特征提取、分类器模型以及添加位置约束的优化,最终识别出用户的运动状态.

针对诺基亚手机平台,我们专门开发了数据采集软件,能实时地同步采集加速度和无线网络两种传感器信号,并将其存储到文本文件中供后续分析之用.加速度采集部分利用此软件完成.图 2 所示为部分运动状态下采集到的加速度序列.

相对运动消除模块是为了克服手机位置、朝向不固定带来的识别率下降问题,主要利用内置加速度计的特点.目前手机内置的加速度计多为 MEMS(micro electromechanical system,微电子机械系统)加速度计,它测量到的加速度值包括两个部分:由重力引起的静态加速度和由用户运动引起的动态加速度,其中只有动态加速度对于识别用户的运动状态是有效的.

在这个问题中存在如图 3 所示两个坐标系:加速度计坐标系 (x,y,z) 和期望坐标系 (v,s,f) .加速度计坐标系与手机放置的方式密切相关,因为手机可能被用户放置为任意朝向,所以加速度计坐标系是不固定.运动状态识别算法希望得到的是以用户为基准的运动坐标系 (v,s,f) ,其中 v 代表垂直方向、 f 代表水平方向(即用户前进方向)、 s 代表侧方向.由于用户一般很少做横向运动,侧方向 s 通常可以不用考虑,重点集中于垂直方向 v 和用户前进方向 f .

为了求出动态加速度在 v 和 f 两个方向上的分量,MotRec 采用了如下加速度分离算法:

输入:长度为 N 的加速度窗口序列 $A = a_1 a_2 \dots a_N$, 其中 $a_i = (a_{ix}, a_{iy}, a_{iz})$.

输出:动态加速度的垂直分量 $P = p_1 p_2 \dots p_N$ 和水平分量 $H = h_1 h_2 \dots h_N$.

步骤:

1. 求出静态加速度.对窗口内的 N 个加速度数据求平均,得到静态分量 $d_i = (a_{ix} - v_x, a_{iy} - v_y, a_{iz} - v_z)$, 其中, v_x, v_y, v_z 分别是 3 个坐标轴上加速度的平均.

2. 分离出动态加速度.加速度是静态加速度和动态加速度的叠加,所以动态加速度的第 i 个元素 $d_i = (a_{ix} - v_x, a_{iy} - v_y, a_{iz} - v_z)$. d_i 由用户的运动产生,而不是由重力因素产生.

3. 分解动态加速度.使用向量点积得到 d_i 在垂直方向 v 的投影 $p_i = \frac{d_i \cdot v}{v \cdot v}$, 即 p_i 为动态加速度 d_i 的垂直分量.因为加速度是垂直分量和水平分量的和,所以水平分量为 $h_i = d_i - p_i$.

4. 如此得到垂直分量序列 $p = p_1 p_2 \dots p_N$ 和水平分量序列 $h = h_1 h_2 \dots h_N$.

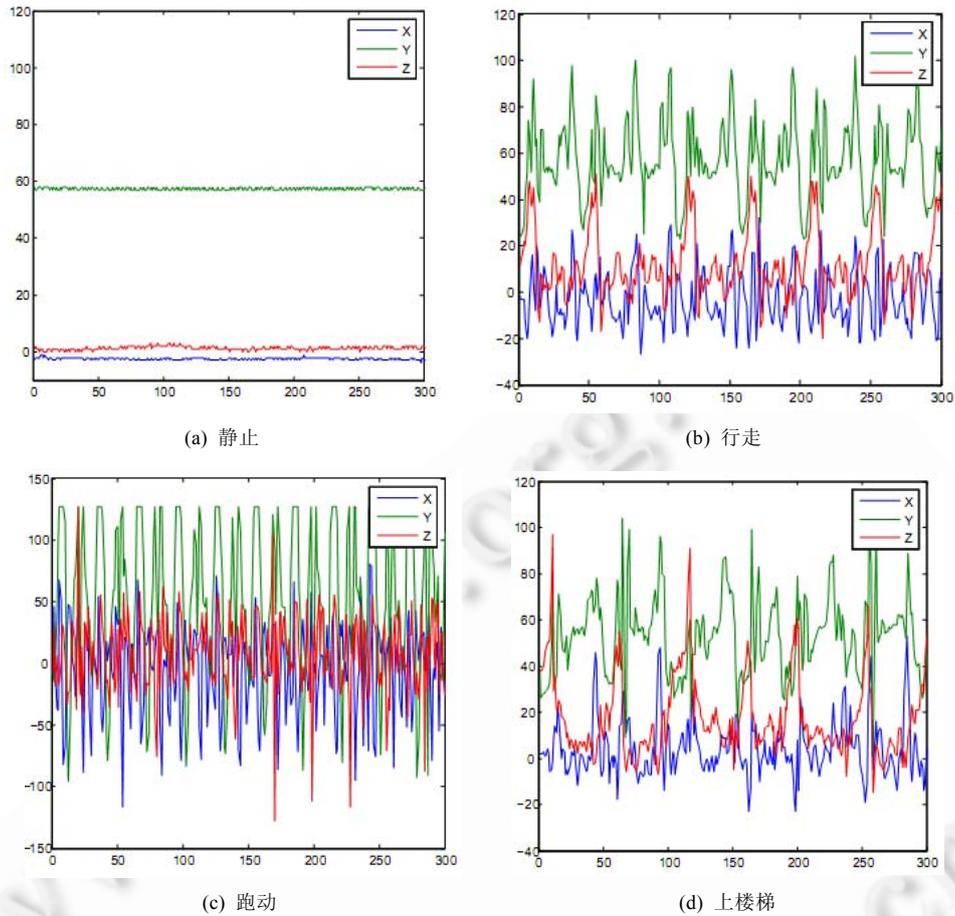


图 2 不同运动状态对应的加速度序列模式



图 3 运动状态识别中的两个坐标系

提取出动态加速度的垂直分量 p 和水平分量 h 之后, MotRec 在这两个方向上分别进行特征提取. 基于对加速度序列模式的分析, 本文主要提取了 {均值、标准方差、相关系数} 3 种主要特征. 下面以垂直分量 p 为例, 介绍如何提取各个特征. 均值为整个窗口中加速度的平均值. 对于长度为 N 的加速度窗口序列, 均值为整个窗口中加速度的平均值 $\bar{p}_x = \frac{1}{N} \sum_i^N p_{ix}$; 标准方差为 $\sigma_x = \sqrt{\frac{1}{N} \sum_i^N (p_{ix} - \bar{p}_x)^2}$, 它可以有效区分具有不同浮动范围的运动状态;

相关系数是每对数据轴之间协方差和标准方差的比率: $corr(x, y) = \frac{cov(p_x, p_y)}{\sigma_x \sigma_y}$, 对于识别涉及多维加速度数据

变化的运动状态非常有效.

用户处于不同运动状态时,用户身上的加速度计所得到的加速度值也往往呈现出不同的模式.运动状态识别算法希望根据不同的加速度模式识别出各个运动状态,这实际上是一个分类问题,每一运动状态对应一个类,数据则是在一段时间内采集到的加速度序列.本文采用机器学习的方法来识别运动状态,主要包括两个阶段:(1) 离线训练阶段.使用多种分类器算法对训练样本进行学习,捕捉运动状态和加速度模式之间的对应关系,从而建立起运动状态分类器模型.(2) 在线识别阶段.实时采集的加速度序列作为分类模型的输入,输出即为运动状态序列.我们主要采用了 K 最近邻(K -NN)、决策树(C4.5)、支持向量机(SVM)和朴素贝叶斯(NB)这 4 种分类器算法.

优化模块主要利用位置信息对运动状态识别的结果进行优化从而进一步提高识别率.位置和运动状态之间确实存在着很强的关联性,如走廊里不可能出现上下楼梯的状态,办公室内不可能出现上下电梯等,位置信息可以用来进一步优化运动状态的识别结果.

2.2 基于位置约束的用户运动状态感知方法

基于无线网络信号强度的定位技术根据移动终端接收到的信号强度进行位置估计,其正式定义如下:用户处于布有无线局域网的环境中,环境中共有 n 个无线网络接入点,用户携带的移动终端可以不断地从这些接入点接收到信号.在某个时刻 t_i ,移动终端从 n 个接入点接收到的信号强度形成一个 n 维的向量 $s_i=(s_{i1},s_{i2},\dots,s_{in})$,其中 s_{ij} 是来自第 j 个接入点的信号强度.位置感知的目标是根据 s_i 估计出移动终端此刻所在位置的坐标 $l_i=(x_i,y_i)$.

室内无线信号的传播由于反射、折射等原因产生多路径效应,不再完全符合信号衰减模型,为此各种基于机器学习的定位方法被先后提出,并取得了不错的定位精度^[10,11,16-18].虽然目前已存在广泛的研究工作存在,但用于行为识别的定位技术具有其特殊性,依然是一个值得研究的问题.首先,用户的活动不可能局限于某个狭小的范围内,往往会涉及较大的应用环境,但已有的基于机器学习的定位技术由于需要采集训练数据限制了其应用范围,收集带有位置信息的训练样本是一件需要花费大量人力和时间的昂贵工作.其次,用于行为识别的定位技术必须尽量做到低功耗以延长服务时间,已有的定位技术很少考虑这个问题.

针对上面讲到的大范围应用以及电量消耗问题,本文提出了基于关键点检测的低功耗用户位置感知方法.用户在做某项事情的过程中需要经过很多位置,对于推测最终的目标行为而言,并不是每一个位置都发挥着同样的作用.在一条位置轨迹上存在一些有着重要作用的关键点,只需要在这些关键点上定位准确,就能达到较好的目标识别率.为此本文提出了基于关键点的定位技术,只在关键点启动无线网络扫描进行定位,这可能会导致粗糙的定位精度,但由于考虑了各个位置的重要性,对最终行为识别的精度不会造成太大的影响.首先由于只需要在关键点采集训练数据,这种定位技术极大减少了训练样本的采集量,降低了定位系统在大环境下部署的困难.其次因为只需要在关键点才扫描无线网络,它能极大减少扫描的次数,节省手机电量.

人们很容易理解关键点的概念也比较容易判断哪些点是关键点,通常会将门口、电梯口等位置看作关键点.但定位算法是运行在手机上的,手机如何自动知道某个位置是不是关键点?是否该启动无线网络扫描进行定位呢?经过观察发现,用户在经过关键点时一般都会发生运动状态的改变,如停下来开门、等电梯等,因此用户的运动状态信息对于是否定位有着直接的影响.基于此本文提出了基于运动状态的关键点探测方法,该方法规定只有当手机探测到运动状态发生变化时,才启动无线网络扫描进行定位.第 2.1 节已经证明运动状态能够通过设备内置的加速度计进行识别,因此基于运动状态的关键点探测方法通过加速度判断用户的运动状态,运动状态发生变化时认为探测到了关键点,启动无线网络扫描进行定位.

2.3 基于多模信息融合的行为预测方法

目标识别(行为预测)模型融合多种传感信息来预测用户将来的行为,首先给出目标识别问题的形式化定义.设 X_i 为用户行为过程中采集到的包含 T 个样本的传感信号序列 $X_{i1}X_{i2}\dots X_{iT}$,其中每一个样本 $X_{ij},j\in[1,2,\dots,T]$ 代表 t_j 时刻采集到的多维信号向量.每一个信号序列 X_i 都对应一个目标行为 $G_i\in G$,其中 G_i 表示用户将要去做的行为,而 G 表示在一个环境中所有目标行为的集合.

行为预测算法包括以下两个阶段:

- 离线训练阶段.一系列采集好的训练数据 X ,包含若干条信号序列 X_i 和其对应的目标行为 G_i .我们的期望是通过这些数据训练出一个行为预测模型 M ,使用模型 M 可以得到条件概率 $P(X_{ij}/G_i)$.
- 在线识别阶段.给出一个新的目标行为未知的信号序列 Y ,推断出其最可能对应的目标 \hat{G} , 即 $\hat{G} = \arg \max_{G_k} P(G_k | Y, M)$.

本文提出的行为识别模型中 X 包含两种信息:加速度值和无线网络信号.由于从复杂的底层传感信号序列直接推断出上层的目标具有较大的难度,本文引入活动序列作为中间状态层,将目标识别问题划分为两个步骤.第 1 步由底层传感器信号序列推断出所对应的活动序列,第 2 步由活动序列推断出其最可能对应的目标行为.

我们定义活动为用户所在的位置及其运动状态,如在走廊走动,在办公桌前静坐等.这样不同的位置与不同的运动状态进行组合就得到了不同的活动.在不同的应用环境中用户可能从事的活动是不同的,需要根据实际环境通过人工统计建立规则库,我们称其为活动库,它包含当前环境下所有可能的活动.利用第 2.1 节描述的使用加速度推断出用户运动状态的方法 MotRec,以第 2.2 节使用无线网络信号推断用户所在位置的算法 PowerLoc,就能得到用户的位置和运动状态信息.进一步结合这两种信息以及建立起的活动库便可得到活动序列 $A=A_1A_2\dots A_m$,其流程如图 4 所示.



图 4 活动序列推断

给定一个目前已识别出的活动序列 $A = A_1A_2\dots A_m$, 目标识别的任务在于推断出最为可能的目标 \hat{G} . 本文采用了 N -gram 方法^[15], N -gram 是一个概率模型,在自然语言处理中被广泛用于预测一个句子中下一个最可能出现的词.将 N -gram 引入到目标识别问题中,目标函数变为

$$\tilde{G} = \arg \max P(G | A_1, A_2, \dots, A_m) = \arg \max P(G | A_{1:m}).$$

应用贝叶斯规则后,上面的公式变为

$$\tilde{G} = \arg \max \frac{P(A_{1:m} | G)P(G)}{P(A_{1:m})} = \arg \max P(A_{1:m} | G)P(G).$$

为了计算简单我们取 $N=2$,这样变成一个 Bigram 模型.根据链式规则(chain rule),计算 $P(A_{1:m} | G)$, 则上面公式变为

$$\tilde{G} = \arg \max P(A_{1:m} | G)P(G) = \arg \max P(G)P(A_1 | G)P(A_2 | A_1, G)\dots P(A_m | A_{m-1}, G).$$

3 实验结果

实验环境为某大学的整个校园,整个校园被 IEEE802.11b 无线网络覆盖.用户可以在整个校园内自由活动,主要的行为集中在 4 个区域(实验室、会议室、教室、餐厅)以及连通它们的走廊、电梯等.本文行为识别研究的实验设备采用了 Nokia N95 智能手机, N95 内置有加速计 LIS302DL 和无线网卡,满足我们对传感器的要求.数据采集通过我们开发的数据采集软件 SensorFusion 完成.

3.1 基于位置约束的用户运动状态感知方法效果

在实验采用的校园环境中,我们采集了 4 个用户在两周内的 86 条加速度序列,每一条序列都进行了标定.用户将 Nokia N95 手机随身携带于口袋里,并记录其各种运动状态的起始与终止时间.

针对上面提到的所有分类器算法,主要采用了 4 种不同实验设置来测试基于位置约束的用户运动状态感知方法的性能:

- 设置 1:单个用户多天的数据,混合并进行交叉验证.

- 设置 2: 多个用户多天的数据, 混合并进行交叉验证.
- 设置 3: 单个用户某天的数据作训练数据, 相同用户其他天的数据做测试数据.
- 设置 4: 单个用户某天的数据作训练数据, 不同用户且不同天的数据做测试数据.

关于分类器模型的建立, 采用了 Weka 工具包, 验证部分则采用 10 倍交叉验证的方法. 表 1 显示了各种分类器算法在不同实验设置下的准确率. 可以看到 C4.5 在各种设置情况下都表现最优, NB, SVM, KNN 的表现依次降低. 对同一用户而言即使在不同的时间, 准确度也可以达到 90% 以上. 但对于实验设置 3 和设置 4 即不同用户的数据而言, 准确度有所下降. 这说明基于加速度的运动状态识别是用户依赖的, 基于用户 A 采集的训练数据得到的分类器, 作用于用户 B 时不再十分有效.

表 1 运动状态识别准确率

分类器	准确率(%)			
	设置 1	设置 2	设置 3	设置 4
KNN	94.83	81.74	92.46	62.14
C4.5	95.92	85.20	93.46	67.60
SVM	95.15	79.62	92.16	63.00
NB	95.86	84.96	91.69	64.00

表 1 中的结果是对所有运动状态识别率做平均之后的结果, 实际上各种运动状态的识别率是有所不同的. 为了分析各个运动状态效果上的差异, 我们采用了混合矩阵的表示方式. 混合矩阵给出的是实际运动状态和被分类器识别成的运动状态的对别效果, 行属性表示实际的运动状态, 列属性表示被识别成的运动状态. 表 2 是实验设置 2 情况下的混合矩阵, 以第 1 行数据为例, 总共 293 段标定为静止的加速度序列, 其中有 281 段被正确识别为静止状态, 3 段被识别成行走状态, 1 段被识别成下楼梯状态, 8 段被识别为电梯状态. 由此我们可以看出, 静止状态的识别率为 $281/293=95.9\%$, 而且静止状态容易被误识别为电梯状态. 通过整个表格的观察, 还可以看出容易被误判的状态还包括跑动与上下楼梯、行走与上下楼梯等.

表 2 优化前运动状态识别混合矩阵

实际状态	被识别为的运动状态					
	静止	行走	跑动	上楼梯	下楼梯	电梯
静止	281	3	0	0	1	8
行走	0	49	0	3	0	6
跑动	0	1	7	2	0	0
上楼梯	0	1	0	4	0	1
下楼梯	0	2	1	0	9	1
电梯	5	1	0	1	0	13

在第 2.1 节曾提出使用位置信息做进一步优化的方法, 利用位置和运动状态之间的相关性消除易混运动状态之间的误判. 针对采用的实验环境, 我们统计定义了实验环境中每个位置不可能出现的运动状态(如办公室内不可能出现上下楼梯的运动状态), 并利用这些约束对运动状态识别的结果进行了进一步优化. 关于如何得到位置信息, 下一节中将进行阐述, 在此先给出已知位置信息情况下对运动状态识别率进行优化后的效果. 表 3 所示混合矩阵是对表 2 优化后的效果, 两者是针对同一数据集, 在同一实验设置下进行的实验结果. 可以看到经过优化, 易混运动状态之间的误判得到了明显的消除. 以第 1 行为例, 优化前有 8 段静止状态的序列被识别为电梯状态, 优化后这种误判不再存在.

表 3 优化后运动状态识别混合矩阵

实际状态	被识别为的运动状态					
	静止	行走	跑动	上楼梯	下楼梯	电梯
静止	290	3	0	0	0	0
行走	0	55	0	3	0	6
跑动	0	1	7	2	0	0
上楼梯	0	0	0	6	0	0
下楼梯	0	2	1	0	10	0
电梯	2	1	0	0	0	17

3.2 基于关键点检测的低功耗用户位置感知方法效果

基于关键点的定位技术,该技术能够减少不必要的无线网络扫描和定位次数,同时期望不会对最终行为识别精确度造成太大影响.该技术在手机上的具体实现则是通过用户运动状态探测关键点并自适应地调整定位模块,只在关键位置进行定位.在此部分,我们在实际数据集上对算法进行了验证.

图 5 给出的是关键点(圆圈标识)的探测效果图.可以看到我们的运动状态识别算法能够基本有效地识别各个运动状态,意味着能够有效捕捉到关键点.同时,因为固定窗口分割算法的影响,有些关键点的探测有稍微的延迟,不过都在可接受范围之内.

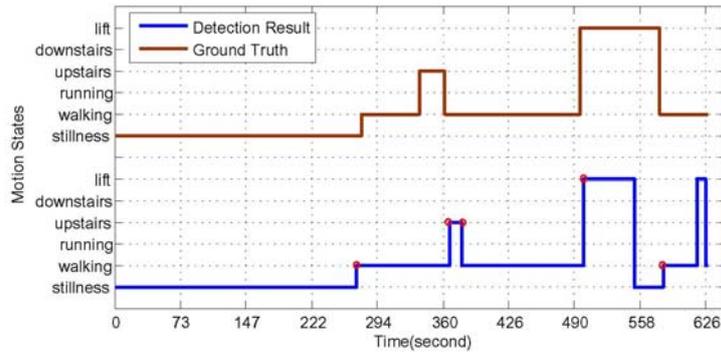


图 5 关键点探测算法效果

上面给出的是一个直观的例子,同时我们根据人工标定的 86 条加速度序列(标注出每一种运动状态的起始和终止点),对关键点探测算法的效果进行了统计.主要使用了两种类似搜索引擎里采用的查全率(recall)和查准率(precision)两种衡量指标,统计结果表明关键点探测算法的查全率和查准率分别为 92.4%和 89.1%.其中查全率定义为: $Recall = \text{探测到的正确关键点数目} / \text{序列中实际含有的关键点数目}$;查准率定义为: $Precision = \text{探测到的正确关键点数目} / \text{探测出的所有关键点数目}$.

对于基于关键点的定位技术是否能够节省电量,我们在这一部分进行了验证.在实际采集的 86 条序列中,采用了不同的运动状态数目设置来扫描无线网络,并对其耗电量进行了统计.为适应手机计算能力,此处采用了基于 K-NN 的运动状态识别算法.图 6 显示电量节省随着运动状态数目变化的情况,基数是每 10 秒扫描一次的固定周期设置.运动状态数目为 2 时即只考虑静止和运动两种状态,然后逐步加入跑动、上楼梯、下楼梯和电梯等运动状态.可以看到随着区分的运动状态数目不断增加,节省的电量呈不断减少趋势.运动状态数目为 2 时可节省电量近 31%,数目为 6 时可节省电量 24%,都达到了不错的电量节省效果.

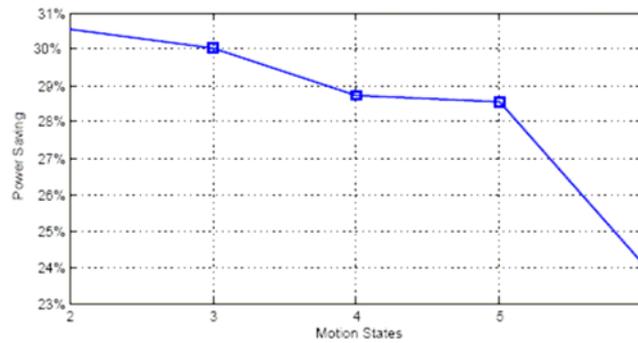


图 6 基于关键点的定位技术的省电量效果

为了衡量定位算法的性能,我们采用误差距离(error distance)作为衡量指标,所谓误差距离为定位算法给出

的估计位置和实际位置之间的距离.上文描述的 3 种定位算法(K-NN,贝叶斯,隐马尔科夫模型)的定位准确度如图 7 所示.误差距离被设为较大的值(例如 10m)是有情可原的,因为我们的应用环境是整个校园,而不是狭小的某个房间的实验环境.可以看到隐马尔可夫因为考虑了序列信息在 3 种算法中表现最好,贝叶斯其次,K-NN 表现最差.

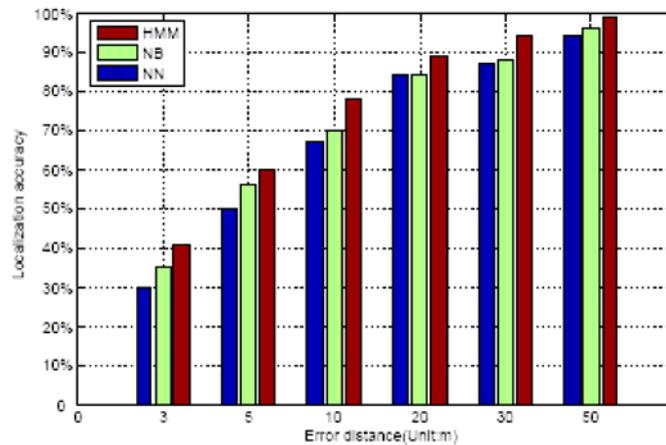


图 7 基于关键点的定位技术的定位精度

3.3 基于多模信息融合的行为预测方法效果

用户在校园里活动时,其携带的手机能够通过加速度计实时采集加速度序列信息,同时能够通过无线网卡收集各个无线接入点的信号强度.我们共采集了一位学生的 742 条日常活动轨迹,每一条轨迹都标定了其对应的目标行为.在所采集的数据集中共有 16 个目标行为,部分目标行为列在表 4 中.为了将传感器信息转化为活动序列,需要根据环境人工建立活动库.针对校园环境我们共定义了 142 个活动,部分活动列在表 5 中.

表 4 目标行为列表

标号	名称
G1	从办公到会议室开会
G2	从图书馆到餐厅就餐
G3	从图书馆到教室上课
G4	从办公室到健身房
G5	从办公室到图书馆

表 5 活动列表

标号	名称
A1	坐在办公桌前
A2	在走廊 1 走动
A3	坐在办公室沙发上
A4	在走廊 2 走动
A5	乘坐 12 号电梯

为了测试行为预测算法,我们在所采集的 742 条行为轨迹上进行了 4 倍交叉验证.由于行为预测算法的底层为基于关键点的定位技术,所以目标识别的效果跟关键点探测的结果密切相关.对同一条传感器信号序列而言,能够识别的运动状态越多,探测到的关键点也会越多,已知的位置信息也会越多,最终目标识别率会相对较高.反之,能够识别的运动状态越少,探测到的关键点也会越少,已知的位置信息也会越少,最终目标识别率也会相对较低.图 8(a)显示的是目标识别率随运动状态数目变化的情况.一方面,可以看到随着运动状态数的增加,目标识别率确实呈上升趋势.如运动状态数为 1 时识别率为 65%,运动状态数为 7 时识别率为 78%.另一方面,目标识别的基础是第 3 节提出的基于关键点的定位,该定位技术出于大范围应用和耗电量节省方面的考虑,只在比

较重要的位置进行定位.但从效果图上可以看到虽然基于粗略定位,GoalRec 依然达到了可接受的目标识别率.

作为服务质量(QoS)的重要衡量指标,目标识别率当然越高越好,但它并不是唯一的衡量指标,例如可用服务时间也是必须考虑的标准之一.为了延长服务时间,需要控制设备的电量消耗.我们曾研究了电量消耗和运动状态数目之间的关系,如图 8(b)所示,随着运动状态数目的增加电量消耗不断上升,与目标识别率呈相反趋势.目标识别率和电量消耗作为两个相互矛盾的 QoS 衡量指标,可以根据不同的应用需求在两者之间寻找平衡点.

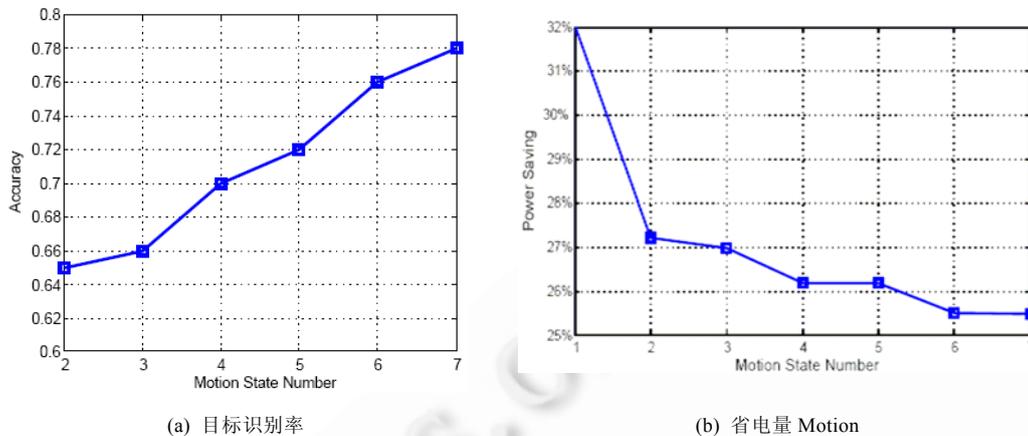


图 8 目标识别率和省电量之间的平衡

4 结束语

本文通过全面分析已有行为识别研究的局限,提出了基于多种传感信息融合的行为识别框架.该框架能有效识别各种行为并提高了识别的精确度,同时具有低功耗的优点,更符合普适计算的发展趋势.目前的方法只是融合了加速度和无线网络两种信息,将来可以通过融合更多的传感器信息进一步提高行为识别的准确度,如通过陀螺仪添加角速度信息、通过温度传感器添加周围环境信息等.

References:

- [1] Xu GY, Shi YC, Xie WK. Pervasive/Ubiquitous Computing. Chinese Journal of Computers, 2003,26(9):1042-1044 (in Chinese with English abstract).
- [2] Li R, Li RF. A survey of context-aware computing and its system infrastructure. Journal of Computer Research and Development, 2007,44(2):269-276 (in Chinese with English abstract).
- [3] Bricon-Souf N, Newman CR. Context awareness in health care: A review. Int'l Journal of Medical Informatics, 2007,76(1):2-12.
- [4] Helal S, Mann W, El-Zabadani H, King J, Kaddoura Y, Jansen E. The gator tech smart house: A programmable pervasive space. IEEE Computer, 2005,38(3):50-60.
- [5] Harris C, Cahill V. Exploiting user behaviour for context-aware power management. In: Proc. of the Int'l Conf. on Wireless and Mobile Computing, Networking, and Communications. 2005. 122-130.
- [6] Chai XY, Yang Q. Multiple-Goal recognition from low-level signals. In: Proc. of the AAAI 2005. 2005. 3-8.
- [7] Tapia EM, Intille SS, Larson K. Activity recognition in the home setting using simple and ubiquitous sensors. In: Proc. of the PERVASIVE 2004. LNCS 3001, 2004. 158-175.
- [8] Ravi N, Dandekar N, Mysore P, Littman ML. Activity recognition from accelerometer data. In: Proc. of the 17th Innovative Applications of Artificial Intelligence Conf. Menlo Park: AAAI Press, 2005. 11-18.
- [9] Patterson DJ, Fox D, Kautz H, Philipose M. Fine-Grained activity recognition by aggregating abstract object usage. In: Proc. of the IEEE Int'l Symp. on Wearable Computers. 2005. 44-51.
- [10] Yin J, Chai XY, Yang Q. High-Level goal recognition in a wireless lan. In: Proc. of the AAAI 2004. 2004. 578-584.
- [11] Yin J, Shen D, Yang Q, Li ZN. Activity recognition through goal-based segmentation. In: Proc. of the AAAI 2005. 2005. 28-34.

- [12] Patterson DJ, Liao L, Fox D, Kautz H. Inferring high-level behavior from low-level sensors. In: Proc. of the 5th Int'l Conf. on Ubiquitous Computing (Ubicomp). 2003. 73-89.
- [13] Bao L, Sintille SS. Activity recognition from user annotated acceleration data. In: Proc. of the 2nd Int'l Conf. on Pervasive Computing. 2004. 1-17.
- [14] Stikic M, Huỳnh T, van Laerhoven K, Schiele B. ADL recognition based on the combination of RFID and accelerometer sensing. In: Proc. of the Int'l Conf. on Pervasive Computing Technologies for Healthcare 2008. 2008. 258-263.
- [15] Chen SF, Goodman J. An empirical study of smoothing techniques for language modeling. In: Proc. of the 34th Annual Meeting on Association for Computational Linguistics. Morristown: Association for Computational Linguistics, 1996. 310-318.
- [16] Chen YQ, Yang Q, Yin J, Chai XY. Power-Efficient access-point selection for indoor location estimation. IEEE Trans. on Knowledge and Data Engineering, 2006,18(7):877-888.
- [17] Dong M, Yang Z, Zhang J, Wang N. Signal strength based WLAN. Journal of Computer Applications, 2004,24(12):49-52 (in Chinese with English abstract).
- [18] Chen YG, Li XH. Signal strength based indoor geolocation. Acta Electronica Sinica, 2004,32(9):1456-1458 (in Chinese with English abstract).

附中文参考文献:

- [1] 徐光祐,史元春,谢伟凯. 普适计算. 计算机学报, 2003,26(9):1042-1044.
- [2] 李蕊,李仁发. 上下文感知计算及系统框架综述. 计算机研究与发展, 2007,44(2):269-276.
- [17] 董梅,杨曾,张健,王能. 基于信号强度的无线局域网定位技术. 计算机应用, 2004,24(12):49-52.
- [18] 陈永光,李修和. 基于信号强度的室内定位技术. 电子学报, 2004,32(9):1456-1458.



齐娟(1985-),女,河北衡水人,硕士生,主要研究领域为人机交互,普适计算.



刘军发(1973-),男,博士,助理研究员,主要研究领域为人机交互,普适计算.



陈益强(1973-),男,博士,研究员,博士生导师,主要研究领域为人机交互,普适计算.



孙卓(1985-),男,硕士生,主要研究领域为人机交互,普适计算.