

基于多模融合的半监督场景识别方法

沈 鸿^{1,2,3} 刘军发^{1,2,3} 陈益强^{1,2,3} 蒋鑫龙^{1,2,3} 黄正宇^{2,3}

(中国科学院大学 北京 100190)¹ (中国科学院计算技术研究所泛在计算系统研究中心 北京 100190)²
(北京市移动计算与新型终端重点实验室 北京 100190)³

摘 要 场景识别是普适计算中的一项重要研究内容,旨在通过识别智能手机用户所在位置的场景,为用户提供精准的个性化服务并提升服务的质量。在实际环境中,精确的场景识别存在两个问题:(1)基于单模传感器数据或无线信号数据的分类效果不佳、普适性不足;(2)场景识别的精度需要依赖大量标定数据,导致成本较高。针对这些问题,提出一种基于多模融合的半监督场景识别方法,该方法充分利用 Wi-Fi、蓝牙和传感器的多模特征来提高识别精度。相比基于单模数据的识别,融合特征将静态场景的分类精度提升了 10%,并且本文通过构建半监督的学习方法解决了动态场景中数据采集成本高的问题,在将标定数据量减少一半的基础上将识别精度提高至 90% 以上。实验数据表明,在利用 Wi-Fi、蓝牙、传感器的互补优势的基础上,引入半监督的学习方法能够提升场景识别的精确度且降低在某些场景下采集数据的成本,从而有效地提升了场景识别的精度和普适性。

关键词 场景识别, Wi-Fi, 蓝牙, 多模融合, 半监督学习

中图分类号 TP311 文献标识码 A DOI 10.11896/jsjcx.191200500C

Semi-supervised Scene Recognition Method Based on Multi-mode Fusion

SHEN Hong^{1,2,3} LIU Jun-fa^{1,2,3} CHEN Yi-qiang^{1,2,3} JIANG Xin-long^{1,2,3} HUANG Zheng-yu^{2,3}

(University of Chinese Academy of Sciences, Beijing 100190, China)¹

(Research Center for Ubiquitous Computing Systems, Institute of Computing Technology, Chinese Academy of Sciences, Beijing 100190, China)²

(Beijing Key Laboratory of Mobile Computing and Pervasive Device, Chinese Academy of Sciences, Beijing 100190, China)³

Abstract Scene recognition is an important part of pervasive computing. It aims to provide users with accurate personalized service and improve service quality by identifying the location of the smartphone users. In the actual environment, there are two problems in accurate scene recognition. Firstly, based on single mode sensor data or wireless signal data, classification effect is not good enough, and its generalization is not enough. Secondly, scene recognition accuracy depends on a large number of labeled data, resulting in high cost. In view of these problems, a semi-supervised scene recognition method based on multi-mode fusion was proposed. The method makes full use of the complementary information of Wi-Fi, Bluetooth and sensors to improve the accuracy of recognition. Compared with the recognition based on single mode data, fusion feature can increase the static scene classification accuracy by 10%. In this paper, a semi-supervised learning method was constructed to solve the problem of high data acquisition cost in dynamic scene, and the classification accuracy is over 90% by reducing half of the labeled data. The results show that introducing semi-supervised learning method based on the complementary advantages of Wi-Fi, Bluetooth and sensors information can reduce data collecting cost and improve scene recognition accuracy to some extent, and thus highly increase its recognition accuracy and universality.

Keywords Scene recognition, Wi-Fi, Bluetooth, Multi-mode fusion, Semi-supervised learning

1 前言

近年来,随着 Wi-Fi、蓝牙等无线网络技术的发展以及移

动终端的小型化和普及化,场景识别逐渐成为研究热点。场景识别旨在感知移动设备所处的场景,进而为用户提供自适应的内容和个性化的动态服务^[1]。例如,当感知到用户处于

到稿日期:2019-01-11 返修日期:2019-04-15 本文受国家自然科学基金面上项目(61572004,61472399,61572471)资助。

沈 鸿(1995—),男,硕士生,主要研究方向为普适计算、无线定位,E-mail:shenhong@ict.ac.cn;刘军发(1973—),男,博士,副研究员,CCF 高级会员,主要研究方向为无线定位、机器学习,E-mail:liujunfa@ict.ac.cn(通信作者);陈益强(1973—),男,博士,研究员,博士生导师,主要研究方向为普适计算、人机交互;蒋鑫龙(1989—),男,博士,助理研究员,主要研究方向为无线定位、机器学习;黄正宇(1991—),男,硕士,主要研究方向为无线定位、机器学习。

商场中时,可以为用户提供智能化的广告和精准的商品推荐。

现有的场景识别方法主要分为如下几类:1)基于环境中无线信号(如低功耗蓝牙、Wi-Fi、GPS等)^[2-3]的场景识别方法;2)基于用户动作信息(如运动传感器数据)的场景识别方法;3)基于场景特点信息(如场景中的温度信息)的场景识别方法。

在基于环境中无线信号的场景识别领域,Chen等^[4]利用环境中动态的蓝牙信号,通过提取手机所处环境中蓝牙的数量、静态蓝牙的比例等特征识别用户所在场景。这种方法具有一定的识别率,但是由于无线信号在某些场景中的覆盖范围不足,因此很难被广泛应用。

在基于用户动作信息的场景识别领域,叶舒雁等^[5]选取运动传感器和温度传感器等数据作为场景感知的依据,进而分类出用户所处的场景。阮锦佳等^[6]首先挖掘出不同传感器在不同环境及不同行为下的特征,然后设计基于多支持向量机的融合模型对室内和室外场景进行识别,最终获得了较好的识别率。这种方法虽然可以完成场景分类识别的任务,但是手机传感器采集的数据会因手机姿态不固定而存在较大的随意度。

在基于场景特点信息的场景识别领域,主要是利用基于场景信息的指纹算法^[7]进行识别。指纹算法主要分为离线训练阶段和在线预测阶段。离线训练阶段的任务是采集场景的指纹数据,例如 Wi-Fi 信号、蓝牙信号、温度、湿度等,然后将一组指纹信息及其对应的场景类别形成(特征,场景)映射并存入指纹数据库。在线预测阶段的任务是采用匹配算法,将终端设备实时收集到的指纹数据与指纹数据库中的指纹信息进行匹配,进而确定终端的实际位置。Krumm^[8]利用设备所处环境中的温度来区分室内和室外,区分度达到了 80% 以上。毕树超等^[9]提出了一种基于 Wi-Fi 指纹识别用户所在场景的方法,在一定程度上能够识别用户所在的室内场景。但这种方法需要采集大量的标定数据作为指纹,极大地增加了成本,并且分类效果高度依赖于离线训练阶段的数据与在线预测阶段的数据是否满足同一分布。由于现实环境具有较大的波动性,Wi-Fi、蓝牙、温度、湿度等信号都会随着时间的变化而发生明显的变化,导致离线阶段训练的模型难以保持时效性。

针对现有工作的不足,本文主要进行以下几方面的研究。1)动态场景和静态场景的二分类问题。如果在某种场景下,用户的位置在短时间内容易发生较大变化,本文将某定义为动态场景,如地铁、公交车等;如果在某种场景下,用户的位置在短时间内的变化范围有限,本文将其定义为静态场景,如办公室、电影院等。2)基于监督学习的静态场景细分类识别以及基于半监督学习的动态场景细分类识别。对于动、静态场景的细分类,需要对动、静态场景进行细类场景的数据采集,即细粒度场景的数据采集。在日常生中,数据采集者处于静态场景的时间远远多于处于动态场景的时间,因此细粒度静态场景的数据采集成本相对较低且数据量较大,所以本文采用监督学习对静态场景进行细分类识别;而细粒度动态场景数据的采集成本较高且数据量较少,所以引入半监督学习方法对动态场景进行细分类识别,以大大减少采集动态场景细

粒度标定数据的成本。算法框架如图 1 所示,其中, x_1, x_2, \dots, x_n 表示特征值的输入。

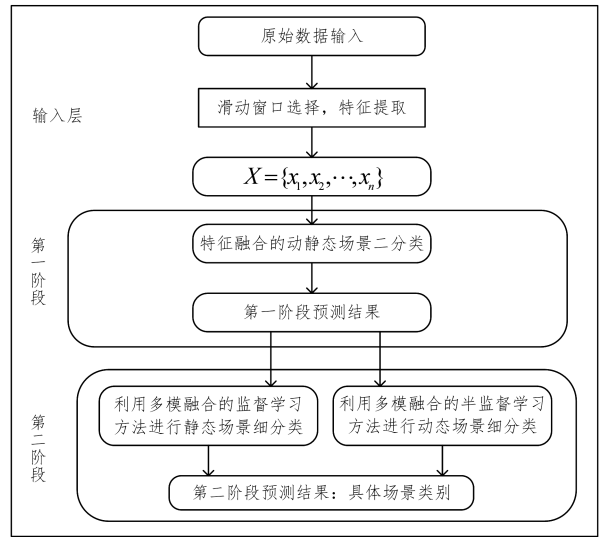


图 1 整体框架图

Fig. 1 Overall framework diagram

本文在已有研究成果的基础上提出了一种基于多模融合的半监督场景识别方法。相比于已有的成果,本文的主要贡献是:1)充分融合了 Wi-Fi、蓝牙、传感器 3 类数据的互补优势,相比于单模数据,分类效果有了明显提升;2)针对动态场景采集标定数据成本较高的问题,引入了半监督的学习方法,较大地减少了标定工作量,同时使得模型的识别率较单独利用标定数据的识别率有了明显提升。

2 数据处理与特征提取

在复杂的环境中,Wi-Fi 和蓝牙信号受信源不足、多径传播、其他电子设备信号以及建筑物的干扰,导致手机接收到的信号强度值存在随机性和不稳定性^[10];同样,手机的电量、网络延迟等因素,也不可避免地使得传感器采集的数据存在异常或者缺失的情况^[11]。因此,采集的原始数据无法直接放入模型中进行训练以及分类,需要进行预处理和特征提取。

2.1 异常数据的处理

本文采取了高斯滤波的方法将异常数据过滤。假定同一环境对无线信号源的影响近似为高斯分布^[12],如对于 Wi-Fi、蓝牙信号,其信号强度服从高斯分布;而对于传感器数据,其采集到的原始信号数据(如温度传感器采集到的温度数据)服从高斯分布。假定某个信号源序列为 $X = \{x_1, x_2, \dots, x_n\}$,其均值 μ 、标准差 σ 分别表示为:

$$\mu = \frac{1}{n}x, \sigma = \sqrt{\frac{1}{n-1}(x-\mu)^2}$$

概率密度函数表示为:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad (1)$$

对异常数据进行过滤的步骤如下:

1)统计在相同场景下来自同一信源的信号序列,计算高斯分布参数 μ 和 σ ;

2)根据概率密度公式计算 $f(x)$, 确定概率范围,过滤不在范围内的信号。

2.2 缺失数据的处理

通过对采集自现实环境中的数据进行观察,将缺失数据分为两种。(1)在某些场景下搜索不到某种网络信标,如在某些动态场景下,由于蓝牙信源不足,手机一直扫描不到蓝牙信号。为了提高算法的普适性,本文在对动态场景进行细分类识别时,不会从蓝牙信号中提取特征,详见第 6.1 节。(2)信号随机缺失。本文针对这种情况,使用在滑动窗口内提取特征的方法来消除缺失值带来的影响。滑动窗口的长度是一个很重要的参数,过长可能会影响在实际运用过程中的实时性,过短则有可能掩饰 Wi-Fi、蓝牙或者传感器在反映环境信息方面的特性。本文将滑动窗口的大小设置为 60 s。滑动窗口示意图如图 2 所示,其中 T_1, T_2, T_3 为滑动窗口的大小,都等于 60 s。

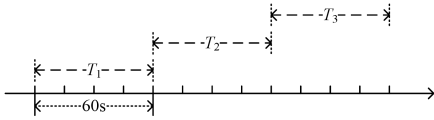


图 2 滑动窗口示意图

Fig. 2 Schematic diagram of sliding window

本文将滑动窗口的大小设置为 L , 滑动窗口的步长设置为 L , 当对滑动窗口进行更新时,在当前的滑动窗口内提取特征并对其进行分类识别。因滑动窗口较大,所以在在一个时间窗口中手机会采集到多条数据,本文从这多条数据中提取特征(如信号数据的平均值、标准差等)的方法可以有效地解决单条数据特征随机缺失的问题,从而消除缺失值的影响。

2.3 特征提取

本文选取与分类相关度较高的特征作为模型的输入,共提取 44 维信号特征,其中 Wi-Fi 信号特征 4 维,蓝牙信号特征 4 维,传感器信号特征 36 维。

对于 Wi-Fi 信号,表 1 给出一些基本符号及其定义。

表 1 基本符号及其定义

Table 1 Basic symbol and its definition

符号	定义
Mac_i	第 i 时间窗口采集到的 Wi-Fi mac 序列: $\{mac_i^1, mac_i^2, \dots, mac_i^n\}$
d_K	Wi-Fi mac 序列 K 中 mac 的个数
Rss_j^i	第 i 个时间窗口中第 j 个 Wi-Fi mac 所对应的信号强度

说明:在经过了数据预处理后,一个时间窗口内不会出现没有扫描到 Wi-Fi mac 的数据,即无论在训练数据还是测试数据中, $d_{Mac_i} \neq 0$ 。

根据表 1 的符号定义,本文针对 Wi-Fi 信号提取了如下 4 个特征。

(1) Wi-Fi mac 的个数 d_{Mac_i}

Wi-Fi mac 的个数反映了终端设备所处环境中 Wi-Fi 设备的密度。

在不同环境中,建筑物等的阻挡影响了 Wi-Fi 信号的传播,从而使得不同环境下的 Wi-Fi 信号值有着明显的差异,对此本文提取了在一个时间窗口内的 Wi-Fi 信号强度的均值和标准差。

(2) Wi-Fi 信号强度的均值

$$Rss_i^? = \frac{1}{N} \sum_{j=1}^N Rss_j^i \quad (2)$$

(3) Wi-Fi 信号强度的标准差

$$Rss_i^? = \sqrt{\frac{1}{N-1} \sum_{j=1}^N (Rss_j^i - Rss_i^?)^2} \quad (3)$$

式(2)、式(3)中的 $N = d_{Mac_i}$ 。

(4) Wi-Fi mac 的变化率

由于人在不同环境中的运动情况不同,通过判断 Wi-Fi mac 的变化率可以知道用户是处于动态环境还是静态环境。变化率 CR 的定义如下:

$$CR = 1 - \frac{d_A}{d_B} \quad (4)$$

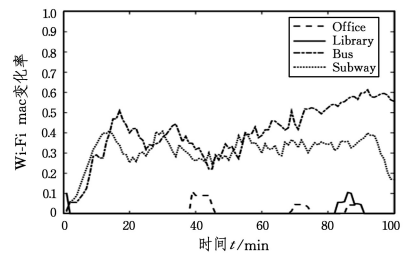
其中, $A = \{x | x \in Mac_i \cap x \in Mac_{i-1}\}$, $B = \{x | x \in Mac_i \cup x \in Mac_{i-1}\}$ 。

对于蓝牙信号,提取其特征的方法与提取 Wi-Fi 特征的方法相同,本文不再详细介绍。对于传感器,本文通过相关性分析,选取加速度(三轴)、方向(三轴)、温度、湿度、光照强度共 5 个传感器,共 9 维传感数据,并提取每维传感数据的平均值(mean)、极差(range)、标准差(standard deviation)和中值(median) 4 个特征量,对每个窗口得到 36 维传感特征值。

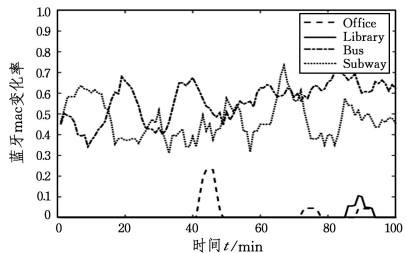
3 动静态场景的二分类

经过信号预处理和特征提取后,首先融合 Wi-Fi、蓝牙、传感器特征对动静态场景做二分类预测。以往的研究主要是利用单模的运动传感器或者无线信号来区分动静态场景,这种方法虽然能完成区分的任务,但是运动传感器的数据与手机的放置位置和姿态有着很强的相关性,并且无线信号亦存在变化性、波动性较大的问题。因此,本文充分利用 Wi-Fi、蓝牙、传感器特征的互补优势,进而对动、静态场景做二类区分,这种方法提升了系统的普适性和分类精度。

为了验证加入 Wi-Fi 和蓝牙特征能提升区分动静态场景的分类准确率,本文分析了在办公室、图书馆两种静态场景及公交车、地铁两种动态场景下 Wi-Fi、蓝牙信号特征的差别。经实验分析, Wi-Fi 和蓝牙 mac 的变化率对于区分动、静态场景有明显的效果,实验结果如图 3 所示。



(a) Wi-Fi 的 mac 变化率



(b) 蓝牙的 mac 变化率

图 3 不同场景下的 Wi-Fi mac 变化率和蓝牙 mac 变化率

Fig. 3 Change rates of Wi-Fi mac and BLE mac in different scenes

从图中可以看出,在公交车和地铁这类动态场景中,Wi-Fi 和蓝牙 mac 的波动非常大;而在办公室和图书馆这类静态场景中,Wi-Fi 和蓝牙的 mac 几乎保持不变。由此,本文选择 Wi-Fi 和蓝牙 mac 变化率特征作为动态和静态场景分类模型的重要特征。

为了验证融合特征下动静态场景二分类的有效性,本文取 Wi-Fi mac 的变化率、蓝牙 mac 的变化率和 36 维传感器特征共 38 维特征数据,观察几种不同的特征组合(传感器、传感器+Wi-Fi、传感器+蓝牙、传感器+Wi-Fi+蓝牙)在几种经典分类算法即 J48 决策树,IBK,Naive Bayes,LibSVM 下的实验效果,结果如图 4 所示。从图中可以得知,在不同算法下,进行特征融合的效果都要优于单传感器的分类识别效果,且将 Wi-Fi、蓝牙、传感器 3 类特征进行融合时的识别率最高。融合特征下几种常见分类算法相比单模特征的识别率平均提升了 10% 左右,相比 Naive Bayes 算法提升了 25%。融合特征充分利用了 Wi-Fi、蓝牙、传感器所包含的信息,因此分类识别效果最好,从而证明了本文所提出的融合特征方法的有效性。

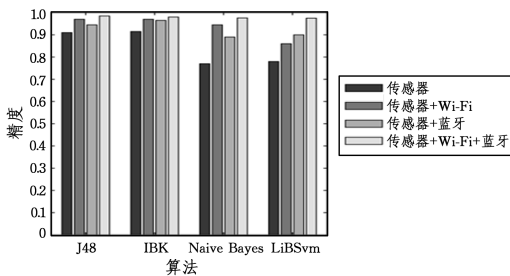


图 4 不同特征组合下动静态场景分类算法结果的对比

Fig. 4 Comparison of classification algorithm results for Dynamic and Static Scene in different feature combinations

4 静态场景的细分类

在对场景进行动静态区分后,需要对动静态场景做进一步的细分,本节主要融合 Wi-Fi、蓝牙、传感器特征对静态场景进行细分。

本文描述的场景为公交、地铁、公园、居民区这 4 种常见的动态场景,以及办公室、图书馆、电影院、医院这 4 种常见的静态场景。由于采集静态场景数据比较方便,数据量较大,本文采用传统的监督学习方法对其进行细类区分;由于对动态场景采集细粒度标定数据的成本较高,因此本文引入了半监督的思想对其进行细类区分,这部分内容将在下面的章节中详细描述。

在区分静态场景时,Wi-Fi、蓝牙、传感器数据的不同特征对不同场景的区分度不同。为了筛选出对静态场景感知分类影响较大的信号特征,本文利用 Weka 对静态场景下的特征值进行排序。其中,Wi-Fi 信号强度均值、蓝牙设备的变化率、光照强度均值是与静态场景分类识别最强相关的 3 个特征,从图 5 中可以清晰地看出 4 种静态场景数据在 3 个特征下的分布。

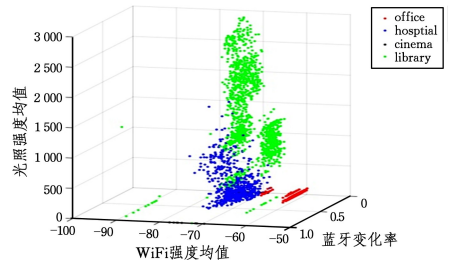


图 5 3 个特征在不同场景下的分布(电子版为彩色)

Fig. 5 Distribution of three features in different scenes

为了验证融合特征的有效性,选取排序后的前 25 维特征(Wi-Fi mac 变化率、Wi-Fi 信号强度均值、Wi-Fi mac 个数、蓝牙 mac 变化率、蓝牙信号强度均值、蓝牙 mac 个数、加速度传感器三轴的平均值和中值、方向传感器三轴的平均值和标准差、温度平均值、温度标准差、光照强度平均值、光照强度中值、光照强度标准差、湿度平均值和温度中值)。针对这 25 维特征,利用 4 种不同分类算法进行建模,并采用十折交叉算法来验证分析,然后从 25 维特征中取出属于传感器类的 19 维特征进行相同实验,算法运行结果如表 2 和表 3 所列。

表 2 基于传感器特征下静态场景的分类结果

Table 2 Classification results of static scenes based on sensor

Algorithm	features	
	Accuracy/%	Modeling Time/s
J48	86.01	1.616
IBK	85.97	5.579
Naive Bayes	68.66	1.266
LibSvm	73.92	15.63

表 3 基于融合特征下静态场景的分类结果

Table 3 Classification results of static scenes based on fusion

Algorithm	features		
	Accuracy/%	Modeling Time/s	Accuracy change/%
J48	94.83	2.275	8.82
IBK	93.18	7.846	7.21
Naive Bayes	80.27	1.586	11.61
LibSvm	82.92	23.89	9

可见,基于融合特征的方法能够明显提高静态场景分类的效果,精度平均提升了在 10% 左右,4 种分类算法对静态场景的分类识别率都达到了 80% 以上,其中 J48 算法的精确度甚至接近于 95%,且建模时间也在可接受范围内,由此证明了本文提出的融合特征静态下场景识别方法的有效性。

5 动态场景的细分类

为了解决单模数据分类精度较低及动态场景下采集细粒度标定数据成本过高的问题,本文在融合特征的基础上先后引入极速学习机(Extreme Learning Machine,ELM)模型以及半监督极速学习机(Semi-Supervised ELM,SS-ELM)。

5.1 极速学习机

ELM^[13]是由黄广斌教授于 2004 年提出的一种单隐层的前馈神经网络。相对于传统神经网络算法而言,ELM 具有学习速度快并且泛化性能好的特点,其网络结构如图 6 所示。

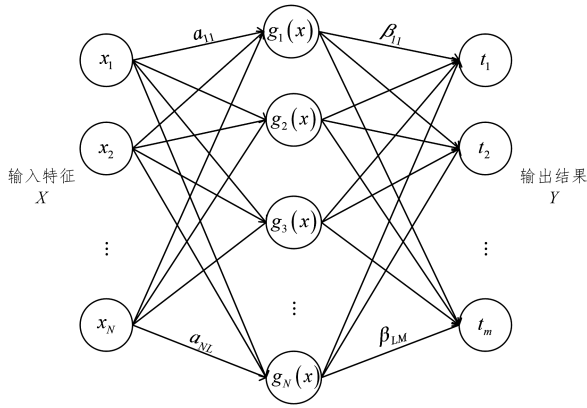


图6 ELM结构图

Fig. 6 Network structure of ELM

对于一个输入向量 $\mathbf{x} \in \mathbf{R}^n$ 和具有 L 个隐层节点的 ELM 输出表达为:

$$f_L(\mathbf{x}) = \sum_{i=1}^L \beta_i G(\mathbf{a}_i, b_i, \mathbf{x}), \beta_i \in \mathbf{R}^m, \mathbf{a}_i \in \mathbf{R}^n, b_i \in \mathbf{R} \quad (5)$$

其中, $\beta_i = [\beta_{i1}, \beta_{i2}, \dots, \beta_{im}]^T$ 表示第 i 个隐层节点与输出节点的连接权重, $G(\mathbf{a}_i, b_i, \mathbf{x})$ 表示第 i 个隐层节点的输出, $\mathbf{a}_i = [a_{i1}, a_{i2}, \dots, a_{in}]^T$ 代表第 i 个输入节点与隐层节点的连接权重, \mathbf{a}_i 和 b_i 是第 i 个隐层节点的偏置。对于每一个隐藏节点, 引入激励函数 $g(x)$, 则第 i 个隐藏节点的输出为:

$$G(\mathbf{a}_i, b_i, \mathbf{x}) = g(\mathbf{a}_i \cdot \mathbf{x} + b_i) \quad (6)$$

对于一个有 N 个训练样本的数据集 $\{(x_j, t_j) \mid j = 1, 2, \dots, N\}$, $\mathbf{x}_j = [x_{j1}, x_{j2}, \dots, x_{jn}]^T$, 其中 $x_i \in \mathbf{R}^n$ 表示输入信号的特征值, $t \in \mathbf{R}^m$ 表示分类标签。假定矩阵的各分量可以如下表示:

$$\mathbf{H} = \begin{bmatrix} G(\mathbf{a}_1, b_1, x_1) & \cdots & G(\mathbf{a}_L, b_L, x_1) \\ \vdots & & \vdots \\ G(\mathbf{a}_1, b_1, x_N) & \cdots & G(\mathbf{a}_L, b_L, x_N) \end{bmatrix}_{N \times L}$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_L^T \end{bmatrix}_{L \times m}, \mathbf{T} = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m} \quad (7)$$

根据文献[13]可知, \mathbf{a}_i 和 b_i 只需在训练初始阶段随机给定, 而不需要在训练过程中通过训练得到。因此, 上式中, \mathbf{H} 和 \mathbf{T} 是已知参数, 只需要求解参数 β 。对于 β 的求解, 可以将上述模型看成是一个线性模型, 从而利用最小均方误差进行求解。

目标函数为:

$$\arg \min_{\beta} (\|\mathbf{H}\beta - \mathbf{T}\|) \quad (8)$$

求解得:

$$\beta = \mathbf{H}^\dagger \mathbf{T} = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{T} \quad (9)$$

其中, \mathbf{H}^\dagger 是 \mathbf{H} 的广义逆矩阵[14-15]。

综合可得 $\mathbf{H}\beta = \mathbf{T}$ 。上述模型将会以近乎零误差得到参数 β 。

5.2 半监督极速学习机(SS-ELM学习机)

SS-ELM 是 ELM 的一种延伸, 主要思想是将标注样本和未标注样本结合起来进行训练以提升学习的性能; 主要出发点是因为人工采集标注样本数据的成本过高, 而未标注样本

的获取非常容易, 因此通过利用少量的标注样本和大量的未标注样本来达到良好的识别精度。

根据结构风险最小化理论[16], 为了达到较好的泛化能力, 学习方法需要平衡学习函数 f 与经验风险。本文引入拉普拉斯图算子[17]做流形约束, 目的是通过构建拉普拉斯图算子来使用无标注样本数据, 进而利用流形正则化方法来实现半监督的能力[18]。本文使用图的平滑度函数 $S(f)$ 来描述模型的复杂度[19], 定义为:

$$S(f) = \sum_{i \sim j} W_{ij} (f_i - f_j)^2 = \mathbf{f}^T \mathbf{L}_a \mathbf{f} \quad (10)$$

其中, $i \sim j$ 表示顶点 i 是顶点 j 的邻居之一, f_i 和 f_j 分别是顶点 i 和顶点 j 上观察到的函数值, L_a 是图的拉普拉斯算子。根据文献[20]:

$$\mathbf{L}_a = \mathbf{D} - \mathbf{W}, D_{ii} = \sum_{j=1}^{l+n} W_{ij}, W_{ij} = e^{-\|x_i - x_j\|^2 / 2\delta^2} \quad (11)$$

考虑到经验风险和控制模型的复杂度, 将损失函数转换为:

$$\arg \min_f \frac{1}{2} \{ \|\beta\|^2 + \|\mathbf{f} - \mathbf{T}\|^2 + \lambda \mathbf{f}^T \mathbf{L}_a \mathbf{f} \} \quad (12)$$

式中的后半部分为流形约束项。为了简化计算, 由 $\mathbf{H}\beta = \mathbf{f}$ 可得:

$$\arg \min_{\beta} l(\beta) = \frac{1}{2} \{ \|\beta\|^2 + \|\mathbf{J}\mathbf{H}\beta - \mathbf{T}\|^2 + \lambda (\mathbf{H}\beta)^T \mathbf{L}_a \mathbf{H}\beta \} \quad (13)$$

利用式(13)对 β 进行求导, 可得:

$$\beta = (\mathbf{I} + \mathbf{H}^T \mathbf{J}\mathbf{H} + \lambda \mathbf{H}^T \mathbf{L}_a \mathbf{H})^{-1} \mathbf{H}^T \mathbf{J}\mathbf{T} \quad (14)$$

其中, λ 为 0 时, 未标定数据将会被忽略, 式(14)会退化成式(9)。

5.3 基于融合特征的半监督方法

根据 2.3 节, 通过特征提取得到融合特征。假定 Wi-Fi、蓝牙、传感器得到的融合特征表示为:

$$\mathbf{x}_i = \{r_w, r_b, r_s\} \quad (15)$$

经筛选后, Wi-Fi 信源、蓝牙信源和传感器信源的个数分别为 $w=3, b=2$ 和 $s=20$, 从而向量 \mathbf{x}_i 可以表示为:

$$\mathbf{x}_i = [r_{w_1}, r_{w_2}, r_{w_3}, r_{b_1}, r_{b_2}, r_{s_1}, \dots, r_{s_{20}}]^T \quad (16)$$

由式(16)可知, 当特征维度为 n 时, $\mathbf{x}_i \in \mathbf{R}^n$, 对于 l 条标定训练样本数据集 $\{(x_i, t_i) \mid x_i \in \mathbf{R}^{w+b+s}, t_i \in \mathbf{R}^m, i=1, 2, \dots, l\}$, u 条非标定样本数据集 $\{(x_i) \mid x_i \in \mathbf{R}^{w+b+s}, i=1, 2, \dots, u\}$, 整个方法的估计过程如下:

- 1) 随机给 \mathbf{a}_i 和 b_i 赋值, 设置隐层节点数为 L ;
- 2) 计算矩阵 \mathbf{H} (\mathbf{H} 的维度为 $(l+u) \times L$);
- 3) 计算图的拉普拉斯算子 L_a , 且 L_a, D, W 的维度均为 $(l+u) \times (l+u)$;
- 4) 计算 β ;
- 5) 利用训练得到的模型 $\mathbf{f} = \mathbf{H}\beta$ 进行预测。

6 实验与验证分析

本文采集数据的终端是魅族 Pro5, 操作系统为 Android 5.1.6, 利用自主开发的 Wi-Fi、蓝牙、传感器同步采集器对室外 4 个场景(公园、地铁、公交车、居民区)采集数据, 总共采集有效数据 4100 条, 其中公园 1098 条、地铁 976 条、公交 1036、

居民区 990 条,训练数据与测试数据的数量比为 1:1。

6.1 融合特征的有效性证明

在某些动态场景下(如公园等空旷场地),由于蓝牙信源不足导致手机在此场景下扫描不到蓝牙信号,所以动态场景预测部分主要是融合 Wi-Fi、传感器的特征,利用半监督的方法对动态场景进行细分类识别。图 7 为 Wi-Fi、传感器、Wi-Fi+传感器 3 种特征组合在不同的动态场景下利用 SS-ELM 进行模型训练和分类识别的结果。

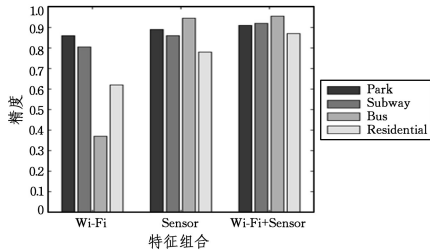


图 7 SS-ELM 算法下不同特征组合的结果对比

Fig. 7 Comparison of results of different feature combinations under SS-ELM algorithm

从图 7 中可以看出:只使用 Wi-Fi 特征进行场景的分类识别时,难以达到较好的识别效果,这是因为在动态场景下 Wi-Fi 的覆盖范围有限,且不同动态场景下的区分度不明显,而在不同场景下,由于不同终端的运动差异性导致传感器信号的差异性较大,因此单利用传感器特征的分类效果优于单使用 Wi-Fi 特征的分类效果;另外,将 Wi-Fi 特征和传感器特征融合后的分类效果最佳,每种场景都有较明显的提升,这是因为融合特征充分利用了 Wi-Fi 和传感器特征的互相关特征,从而使得分类效果有明显的提升,结果证明了融合特征策略的有效性。

6.2 标定数据量对分类识别效果的影响

为了验证标定数据量对 SS-ELM 分类精度的影响,保持测试样本的个数不变,将训练样本分为标注样本 $Train_l$ 和未标注样本 $Train_u$,未标注样本不包含场景标签。在 $Train_l$ 为 200,400,600 时,将 u/l 分别设置为 0,0.5,1,1.5,2 和 2.5,其中 u 为未标注样本的数量, l 为标注样本的数量,对模型进行训练,然后测试分类识别的效果,结果如图 8 所示。

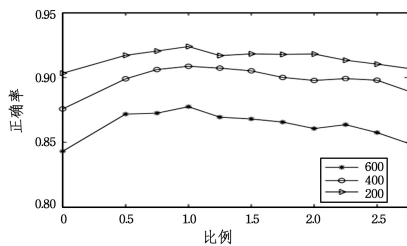


图 8 不同数据比例下分类结果的比较

Fig. 8 Comparison of classification result under different data proportions

分析图 8 可知,当标定数据与未标定数据的比例固定时,随着标定数据的增加,分类识别率逐步上升,表明标定数据量越多,分类模型的效果越好。另外,当标定数据一定时,分类识别率随着未标定数据的增加先上升后下降,这是因为当未

标定数据量的比例较低时,未标定数据和标定数据的相互结合提升了模型的效果;当未标定数据量的比例过高时,训练模型的效果由未标定数据控制,训练数据会偏离标定数据所处的分布,从而使得模型的分类效果下降。从图 8 中可知,当标定数据与未标定数据的数量比为 1:1 时,模型的分类效果最佳,当标定数据量为 400 时,识别效果达到了 90% 以上,从而证明了文中所提半监督方法能够充分利用未标定的数据来提升模型的效果。

6.3 半监督与其他分类算法的比较

为了验证不同算法的分类识别效果,本文选取了贝叶斯网络(Bayesian Network, BN)、 k 近邻中的 IBK、决策树算法中的 J48、ELM 来与本文引入的 SS-ELM 算法进行比较。保持测试数据不变,将训练数据中的标定数据量取值为 400, $u/l=1:1$ (其中 u 表示标定数据量, l 表示未标定数据量),由于 Naive Bayes, IBK, J48, ELM 不能用于半监督学习中,因此只能用标定数据进行训练。利用上述几种算法对模型进行训练以及类别预测,并对比不同算法下的训练时间、预测时间、分类识别率,最终结果如表 4 所列。

表 4 动态场景的分类结果

Algorithm	Accuracy/%	Train Time/ms	Test Time/ms
SS-ELM	90.8	53.7	59.9
ELM	87.5	12.3	46.8
IBK	87.0	35	72
Naive Bayes	85.2	32	64.5
J48	85.6	27	65

分析表 4 可知,本文引入的 SS-ELM 算法和 ELM 算法在分类效果上优于其他算法,且 SS-ELM 算法优于 ELM 算法,因为 SS-ELM 是在 ELM 的基础上充分利用了未标定数据对模型进行训练,既保留了 ELM 的优势,又体现了未标定数据的作用。在训练耗时方面,SS-ELM 由于引入了未标定数据进行训练,因此训练时长有所增加,但在将模型训练好后,其测试阶段的时间仅次于 ELM。综上所述,SS-ELM 在将标定数据减少一半的基础上使识别精度达到了 90% 以上,从而证明了引入的半监督分类方法能够充分利用未标定数据提升模型的识别率。

表 5 为融合 Wi-Fi 和传感器特征后,SS-ELM 分类算法结果的混淆矩阵。从混淆矩阵中可以看出,SS-ELM 算法对公园、地铁、公交等动态场景的识别率达到了 90% 以上,对静态场景居民区的识别率也达到了 82.7%,以上结果证明了本文引入的 SS-ELM 算法对动态场景分类的有效性。

表 5 SS-ELM 算法的混淆矩阵

Model	Accuracy/%			
	Park	Subway	Bus	Residential
Park	91.4	0.5	2.4	4.9
Subway	1.0	93.3	3.3	2.5
Bus	0.9	1.9	95	2.1
Residential	2.7	2.3	12.2	82.7

结束语 本文针对利用单模数据通过传统机器学习方法

识别用户终端所在场景时存在普适性不足、识别精度较低的问题,提出了一种基于多模融合的半监督场景识别方法。该方法首先对采集的 Wi-Fi、蓝牙、传感器信号数据进行预处理;然后提取融合特征,利用传统机器学习方法对动态场景进行大类区分并对静态场景进行细类区分,同时引入半监督方法充分利用动态场景中未标定的数据训练模型,进而实现动态场景的细类区分。实验表明,通过提取融合特征能够大大地提升区分度,引入半监督学习方法大大地减少了动态场景下采集细粒度标定数据的成本,也提升了动态场景的区分度。

参 考 文 献

- [1] KAPITSAKI G M,PREZERAKOS G N,TSELIKAS N D, et al. Context-aware service engineering: A survey[J]. *Journal of Systems & Software*,2009,82(8):1285-1297.
- [2] CHEN Z Y. Mining individual behavior pattern based on significant locations and spatial trajectories[C]// *IEEE International Conference on Pervasive Computing and Communications Workshops*. Lugano:IEEE,2012:540-541.
- [3] CHEN Z,ZHOU J,CHEN Y,et al. Combing multiple linear regression and manifold regularization for indoor positioning from unique radio signal[C]// *Pervasive Computing*. Taipei:IEEE,2009:611-614.
- [4] CHEN Y Q,CHEN Z Y,LIU J F,et al. Surrounding context and episode awareness using dynamic Bluetooth data[C]// *ACM Conference on Ubiquitous Computing*. For Worth:ACM,2012:629-630.
- [5] YE S Y,ZHANG W Z,QI T L,et al. A Sensor and User Behavior Data Analysis Based Method of Mobile Learning Situation Perception[J]. *Journal of Computer Research and Development*,2016,52(12):2721-2728. (in Chinese)
叶舒雁,张未展,齐天亮,等. 一种基于传感器与用户行为数据分析的移动学习场景感知分类方法[J]. *计算机研究与发展*,2016,52(12):2721-2728.
- [6] RUAN J J,LUO D,LUO H Y. Indoor and outdoor scene recognition algorithm based on support vector machine multi-classifier [J]. *Journal of Computer Application*,2015,35(11):3135-3138. (in Chinese)
阮锦佳,罗丹,罗海勇. 基于支持向量机多分类器的室内外场景感知算法[J]. *计算机应用*,2015,35(11):3135-3138.
- [7] YOUSSEF M A,AGRAWALA A,SHANKAR A U. WLAN Location Determination via Clustering and Probability Distributions[C]// *IEEE International Conference on Pervasive Computing and Communications*. IEEE,2003:143-150.
- [8] KRUMM J,HARIHARAN R. TempIO: Inside/Outside Classification with Temperature[C]// *Second International Workshop on Man-Machine Symbiotic Systems*. Kyoto,2004.
- [9] BI S C,YUAN K,CHEN J L,et al. A method of user's scene recognition based on WIFI portrait; CN 201610373922. 6[P]. 2016. (in Chinese)
毕树超,袁凯,陈津来,等. 一种基于 WIFI 画像识别用户所在场景的方法;CN 201610373922. 6[P]. 2016
- [10] LIU Z W,XU F Y,WANG Z X. A new multi-user indoor location algorithm based on difference distance-loss model with parameter fitting [J]. *Chinese Journal of Radio Science*,2008,23(6):1090-1094. (in Chinese)
刘召伟,徐凤燕,王宗欣. 基于参数拟合的室内多用户定位算法 [J]. *电波科学学报*,2008,23(6):1090-1094
- [11] MILETTE G,STROUD A. *Professional Android Sensor Programming*[M]. Wrox Press Ltd,2012.
- [12] ZHU X M,LI S R,LI Z B. WKNN Indoor Positioning Method Based on Distribution Overlap of Rssi[J]. *Journal of Xuzhou Institute of Technology(Natural Sciences Edition)*,2017,32(3):48-52. (in Chinese)
朱雪梅,李石荣,李泽彬. 基于 RSSI 分布重叠的 WKNN 室内定位方法[J]. *徐州工程学院学报(自然科学版)*,2017,32(3):48-52
- [13] HUANG G B,ZHU Q Y,SIEW C K. Extreme learning machine: a new learning scheme of feedforward neural networks [C]// *IEEE International Joint Conference on Neural Networks*. Budapest:IEEE,2005:985-990.
- [14] BANERJEE K S. Generalized Inverse of Matrices and Its Applications[J]. *Technometrics*,1971,15(1):197-197.
- [15] CVETKOVIĆ D M,DOOB M,SACHS H. Spectra of graphs: theory and application[J]. *Math Hungar*,1995,6(2):191-195.
- [16] CHERKASSKY V. The Nature Of Statistical Learning Theory [J]. *IEEE Transactions on Neural Networks*,1997,8(6):1564-1564.
- [17] JERIBI A. *Spectral Graph Theory*[M]// *Spectral graph theory*. Springer,1997:212.
- [18] XU R,JIANG F,YAO H X. Overview of manifold learning [J]. *Journal of Intelligent Systems*,2006,1(1):44-51. (in Chinese)
徐蓉,姜峰,姚鸿勋. 流形学习概述[J]. *智能系统学报*,2006,1(1):44-51.
- [19] BELKIN M,MATVEEVA I,NIYOGI P. Regularization and Semi-supervised Learning on Large Graphs[C]// *International Conference on Computational Learning Theory*. Berlin:Springer,2004.
- [20] BELKIN M,NIYOGI P,SINDHWANI V. Manifold Regularization: A Geometric Framework for Learning from Labeled and Unlabeled Examples[J]. *Journal of Machine Learning Research*,2006,7(1):2399-2434.