

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

叶舒雁 张未展 齐天亮 李 静 郑庆华

(西安交通大学计算机科学与技术系 西安 710049)

(陕西省天地网技术重点实验室(西安交通大学) 西安 710049)

(zhangwzh@mail.xjtu.edu.cn)

A Sensor and User Behavior Data Analysis Based Method of Mobile Learning Situation Perception

Ye Shuyan, Zhang Weizhan, Qi Tianliang, Li Jing, and Zheng Qinghua

(Department of Computer Science and Technology, Xi'an Jiaotong University, Xi'an 710049)

(Shaanxi Province Key Laboratory of Satellite and Terrestrial Network Technology (Xi'an Jiaotong University), Xi'an 710049)

Abstract As the popularity of the smart phones and mobile technologies, more and more people begin to use smartphones to learn and get new knowledge. Mobile learning has played a critical role in the field of education for a few years. The effectiveness of mobile learning reflects in the ability of perceiving different learning contexts and then provides matched learning resource. Context awareness has become a research hotspot, but the most important is learning situation perception. We can provide proper learning resources according to the specific learning situation. Because of the mobility and complexity of mobile learning, it's difficult to perceive learning situation. The thesis proposes a method to perceive learning situations by combining sensor data and learning operation data and conducts some experiments. It chooses and calculates some sensor data eigenvalues and learning operation index eigenvalues as the inputs of the classification algorithms, the learning situations that students provide as training set data. The result shows that combining sensor data and learning operation data to perceive learning situations can improve the accuracy of the learning situation perception, which proves the feasibility and effectiveness of learning situation perception based on sensor data and learning operations.

Key words mobile learning; mobile sensor; learning operation; learning context perception; context classification

摘 要 随着智能手机和移动互联网的普及,使用智能移动终端进行学习的用户也逐渐增多,移动学习在数字教育领域占据着越来越重要的地位.移动学习的有效性体现在情境感知的能力,即能够感知不同学习情境并提供相应合理的学习内容.因而,移动学习中的情境感知技术已经成为一个研究热点.学习场景的感知是移动学习情境感知的重点,但是由于移动学习的动态性和复杂性,准确的场景感知具有一

收稿日期:2016-08-19;修回日期:2016-10-28

基金项目:国家重点研发计划项目(2016YFB1000903);国家自然科学基金项目(61472317,61428206,61472315,61532015,61532004);教育部创新团队发展计划资助项目(IRT13035);陕西省科技统筹创新工程重点实验室项目(2013SZS05-Z01)

This work was supported by the National Key Research and Development Program of China (2016YFB1000903), the National Natural Science Foundation of China (61472317, 61428206, 61472315, 61532015, 61532004), the Innovative Research Team in University of Ministry of Education of China (IRT13035), and the Co-ordinator Innovation Project for the Key Lab of Shaanxi Province (2013SZS05-Z01).

定的难度. 基于实际的移动学习环境, 提出了一种根据传感器与学习操作行为对学习场景进行感知分类的方法, 处理并分析了由移动学习客户端采集到的传感器数据和学习操作行为日志数据, 对比了以传感器数据特征值与学习操作行为特征值共同作为输入特征值的多种场景感知分类算法. 结果表明: 对比仅使用传感器数据作为分类算法输入特征值的结果, 结合学习操作行为日志和传感器数据一起作为学习场景分类感知的依据, 可以显著提高移动学习场景的感知分类效果.

关键词 移动学习; 移动传感器; 学习操作; 学习场景感知; 场景分类

中图法分类号 TP391

随着移动终端设备的普及, 近几年移动学习应用也呈现出爆发性的增长趋势, 对比传统学习方式, 移动学习有着非常显著的特点: 多变的学习场景. 现有场景感知的研究主要是从原始环境、服务设计开发、可用性研究等方面开展, 比如将地理位置信息与带宽预测结合, 从而优化视频自适应服务技术^[1], 考虑场景因素对于 QoE 的影响, 进而提高视频质量^[2], 利用移动用户位置信息, 提高移动网络服务的可靠性和准确性^[3]等. 通过考虑终端位置、用户相关数据等场景相关信息感知场景, 进而提供自适应的内容、个性化的动态服务^[4], 可以显著地提高用户的积极性和服务的有效性. 通过获取用户手机的各类传感器数据信息并采用一定的数据挖掘算法来学习感知用户情境信息, 进一步与移动学习应用结合成为目前教育学习领域研究的热点^[5]. 在此基础上, 基于具体学习情境向学习者推送适合的学习资源, 以实现个性化学习是移动学习发展的趋势^[6].

不同学习场景下学生的学习操作行为也有一定的差异^[7], 因此本文在不影响用户体验的情况下在后台采集移动学习用户手机传感器数据并记录学生学习操作行为, 在传感器数据感知场景的基础上, 进一步考虑综合传感器数据和学习操作日志共同来感知学生的学习情境以获取更好的场景感知效果.

1 传感器数据与用户行为数据的采集和预处理

1.1 传感器数据与学习操作行为采集

不同于一般的移动学习应用, 本文中使用的自主研发的客户端, 在学生在学习过程中会在后台采集学生课程视频学习过程中的一些学习操作数据, 比如开始学习课程的时刻、暂停操作的时刻和暂停时长、拖动操作的时刻和拖动的时长、结束学习的时刻和学习时长. 通过这 4 类操作记录数据可以完整地还原出某个学生的学习过程. 由于研究课题的需要, 客户端在学生课程视频的过程中还会采集手机上

8 类传感器数据, 分别是加速度传感器、重力传感器、陀螺仪传感器、旋转矢量传感器、温度传感器、光线传感器、湿度传感器和声音传感器. 本文后续的场景感知分类首先是基于从服务器上下载的原始传感器数据, 然后在此基础上以学习操作行为日志分析出的一些学习操作为指标, 比如学习时长、暂停以及拖动的次数和时长这些行为指标连同传感器数据一起作为场景感知分类的依据, 从而达到优化场景感知分类的效果.

数据采集完成后, 就可以针对数据进行具体分析了. 然而由于手机使用过程中的移动性、电量消耗、网络传输等因素, 传感器数据的获取并不稳定, 当传感器的能量较低时采集到的数据通常是不可用的脏数据、噪声数据或者直接数据丢失^[8], 所以原始传感器数据必须经过特定的处理才能进行分析. 本文对传感器数据的处理主要包括脏数据处理以及对缺失值的插补. 本文利用不同传感器属性之间的相关性和传感器数据自身的时间特性来针对缺失数据进行插补. 除此之外, 传感器数据也需要与从服务器中提取的移动学习操作数据进行匹配.

1.2 传感器脏数据处理

尽管目前移动设备中传感器已经有了很大的改进, 但是由于手机电量和用户移动的影响, 传感器数据中不可避免地存在噪声^[9], 数据还是要依赖各种形式的平滑和均衡, 也就是通常说的低通滤波, 它可以滤除高频率的噪声, 而“放行”低频率和改变缓慢的变化^[8]. 而高通滤波可以去除对静态和缓慢变化的背景数据的强调, 它强调的是更高频率和瞬变分量. 因此本文首先使用加权平均法作为低通滤波算法来实现对数据的平滑作用, 然后从传感器数据中减去低通滤波后的结果, 从而得到高通滤波的最终数据.

另外, 对于变化剧烈的数据, 计算简单移动平均线对单个数据点的峰值能起到更好的平滑作用, 可以更好地分析数据. 简单移动平均线 (simple moving

average, SMA)也被称为滚动平均值或移动平均值,它能够很方便地找到数据序列中最近的 k 个数据的算术平均值,其中整数 k 代表平均“窗口”的大小.只有 k 个值全部被采集到后,该方法才能奏效.对前 $k-1$ 个值,平均值可以取任意值(例如 0),或者将 SMA 的计算延迟到采集到 k 个值之后进行.平滑参数和窗口大小的选择需要权衡,在定义 1 个合适的窗口大小时,最好能够平滑在给定的时间段内采集到的所有数据值,而不仅仅考虑最新的,本文中选取最近的 60 个值作为窗口大小.

1.3 传感器数据缺失值处理

由于移动手机的特殊性,传感器数据的缺失是无法避免的.对于缺失的传感器数据,本文考虑到传感器数据的时间序列特性和其他一些相关特性,因此插值在准确性上明显不够^[10],会对后续的数据建模带来偏差,影响场景感知的正常进行^[11].传感器数据的缺失模式通常也是随机缺失模式,数据缺失时间之间不存在线性关系,某属性值这一刻缺失,下一时刻可能就可以正常采集.在对已有的传感器数据进行分析 and 挖掘后,本文根据不同传感器属性值之间的相关性和传感器数据自身的时间特性来针对缺失的传感器数据进行插补,取得了良好的数据插值效果.

由于温度、湿度和光照强度这 3 个传感器属性数据之间存在极大的相关性,因此可以通过分析传感器属性之间的相关性并加以利用来恢复某些传感器数据缺失值.通过一元回归模型和多元回归模型的使用,本文成功对一些场景相关传感器的缺失数据,比如温度、湿度、光强进行了预测和插补^[12].

利用不同类型传感器数据之间的属性相关性可以预测部分数据,然而原始数据中也存在多种传感器数据在某一时刻数据都缺失的情况,现有的缺失数据插补方法没有考虑到传感器数据的时间信息.由于手机传感器采集到的环境数据如温度、光强、湿度等都是些连续变化的物理量,即相邻时刻的数据值相差不大,所以传感器数据作为时间序列数据,数据之间通常具有一定的时间相关性^[13],利用这种时间相关性来预测丢失的传感器数据,准确率应该相对较高.因此本文根据传感器数据的时间序列特性,通过对相邻时刻数据构建线性插值函数对某些缺失值进行估计和预测.

缺失值插值完成后,本文需要对数据完整性进行检查,检查传感器数据的采样率和序列长度以及

传感器类型的完整性.确保分析的每个传感器序列文件都是基于采样率 20 Hz、时长为 10 min,且包含所有需要的传感器类型的数据.

1.4 学习操作行为日志预处理

由于移动学习客户端是在判断网络状况良好的情况下才进行学习操作行为数据的发送,所以可能导致学习操作行为发生时间和操作数据被上传到服务器的时间并不一致,学习操作在服务器数据库中的顺序并不一定对应操作真实发生的顺序.在此本文需要对操作日志进行一定的处理,根据操作行为发生的时间对操作进行排序,整理出正确的操作行为发生顺序.学生 1 次学习过程中产生的学习操作行为日志中的开始操作和结束操作的时刻对应于传感器采集开始和结束的时刻,这样使学习操作日志数据和传感器日志数据文件可以对应起来,从而可以确定某学习场景下的学习操作行为数据.

2 基于传感器与学习行为的场景感知

2.1 传感器数据特征值选取与计算

经过第 1 节除噪和缺失值插补后的传感器数据并不能直接作为场景感知分类算法的输入,为了降低建模时间,提高分类效果,本文基于原始传感器数据处理选取出某些对分类结果有较大影响的特征值作为算法建模的依据^[14].移动终端采集的传感器数据是时间上连续的序列,本文选取的是时长超过 10 min 的数据,可以采用滑动窗口的方法进行特征提取.

设 $\mathbf{A}_{3a}$ 为 3 轴加速度的信号强度矢量,可以为

$$\mathbf{A}_{3a} = \sqrt{\mathbf{A}_x^2 + \mathbf{A}_y^2 + \mathbf{A}_z^2},$$

其中, $\mathbf{A}_x, \mathbf{A}_y, \mathbf{A}_z$ 分别为 x, y, z 轴方向的加速度.同理, 3 轴陀螺仪的旋转矢量也使用了相同的计算方法.

在情境感知研究中常用的特征值有:均值、方差、相关系数等;本文选取了 6 种特征值:平均值(mean)、标准偏差(standard deviation)、中值(median)、偏度(skewness)、峰态(kurtosis)和 4 分位点范围(inter-quartile range),这 6 个统计学特征被证明是利用传感器数据进行分类的比较有效的选择^[15].

2.2 基于传感器数据的场景感知方法

由于数据量非常大,静态场景和动态场景的比例差距也很大,考虑到静态场景和动态场景的数据很容易区分,本文选择对传感器数据采用 2 次分类策

略. 首先确定场景的基本类型: 动态或者静态; 然后再进行第 2 次分类, 分别在这 2 类场景中判断出具体的场景类型.

经实验分析, 静态场景和动态场景仅根据加速度传感器数据的变化即可判断. 如图 1 所示, 在图书

馆和办公室场景中手机 3 轴加速度传感器的数据几乎保持恒定不变, 而在公交或地铁中手机 3 轴加速度传感器的数据变化剧烈, 波动很大. 因此通过计算加速度传感器数据的标准差, 很容易区分出静态场景和动态场景.

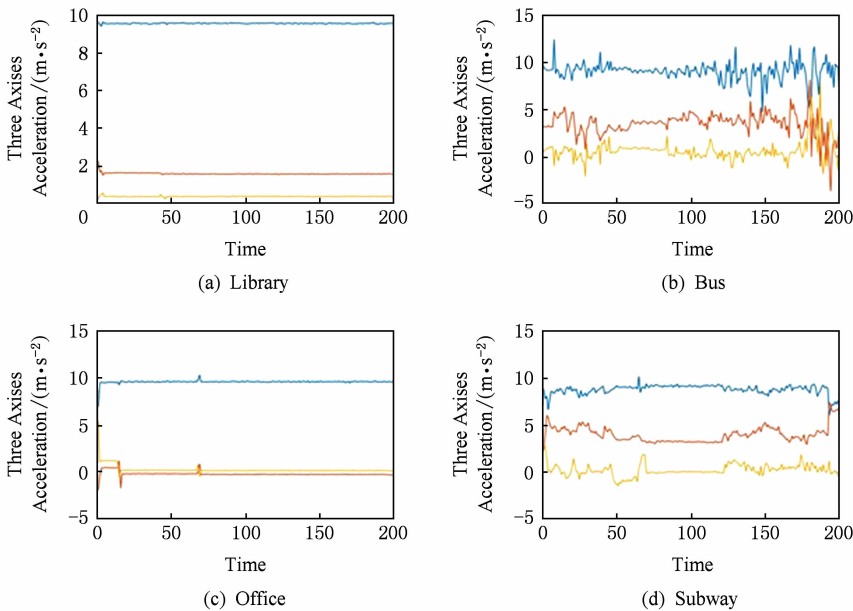


Fig. 1 The triaxial acceleration data of different scenarios.

图 1 不同场景下的 3 轴加速度数据

本文分别随机抽取了 200 个静态和 200 个动态场景的传感器数据文件, 并计算出每个场景对应的 3 轴加速度传感器数据的标准差特征值, 得到如图 2 所示两大类场景的分类结果的散点图. 很明显地看到, 静态场景的 3 轴加速度标准差集中在坐标原点附近, 动态场景的 3 轴加速度标准差离原点较远, 比较分散. 通过对已有场景的加速度传感器数据的统计分析, 得出当 3 轴加速度 x, y, z 方向上传感器数据标准差有任意 1 个超过 2.0 的即为动态场景, 否则为静态场景. 分类的准确率达到 96% 以上.

针对已区分的静态和动态场景, 本文仍需要再在这 2 个大类里面进行细化分类, 目标是区分出具体的场景类型. 由于本文采集了 8 种传感器类型的数据, 根据 2.1 节选取的 6 个特征量共计算得到 48 个特征值.

本文选择了随机森林(random forest, RF)、J48 决策树、随机委员会(random committee, RC)、贝叶斯网络(Bayesian network, BN)和 k 最近邻算法中的 IBk 共 5 种分类算法来对传感器数据进行建模.

由于不同传感器对于不同场景分类的影响程度不同, 因此首先要分别筛选出对动静态场景影响较大的传感器属性值. 使用 Weka 对 4 种静态场景下的传感器数据特征值进行排序, 对静态场景影响较大的排名前 10 的属性分别是: 光强度中值、音量 4 分位点范围、温度中值、音量中值、光强度偏度、光强度标准差、温度峰态、光强度 4 分位点范围、光强度平均值、光强度峰态. 由此可知, 环境传感器数据对静态场景的感知分类起着关键作用. 使用 5 种不同分类算法对所有 48 个传感器特征值数据建模并采用十折交叉验证进行验证, 算法运行结果如表 1 所示:

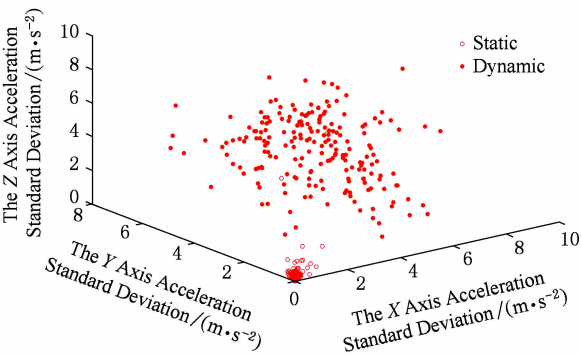


Fig. 2 Dynamic and static scene classification.

图 2 动静态场景分类

Table 1 Classification Results of Static Learning Situations

表 1 静态场景分类算法结果

Classification Algorithm	Accuracy/%	Modeling Time/s
RF	83.6	7.38
J48	81.7	7.04
RC	85.8	9.12
BN	65.1	0.36
IBk	83.5	0.00

可见,使用 RC 算法建模的性能最好,准确度最高,建模时间也在可以接受的合理范围.表 2 为使用传感器特征值集合时 RC 算法分类结果的混淆矩阵.

Table 2 Confusion Matrix of RC

表 2 RC 算法分类结果混淆矩阵

Practical	Model			
	Home	Office	Library	Outdoor
Home	5483	116	130	136
Office	237	5726	221	109
Library	166	273	2095	307
Outdoor	258	217	199	1008

由于所有 6 种动态场景中,不同场景所占比例区别较大,公交场景的比例 31.55% 远远高于飞机场景的比例 4.34%,因此不能直接采用分类算法^[16].这是一个不平衡类的分类问题,本文采用 smote 抽样方法,过分抽样复制数据量少的样本,主要是火车、飞机和其他场景的数据,使各样本的数据量保持大致平衡,然后再对所有数据样本采用不同分类算法建模.同样使用 Weka 对所有动态场景下的传感器数据特征值进行排序,得出的对动态场景分类影响较大的排名前 10 的属性分别是:加速度中值、加速度标准差、加速度 4 分位点范围、旋转矢量中值、加速度峰态、音量偏度、陀螺仪峰态、重力中值、旋转矢量偏度、重力 4 分位点范围,说明对动态场景分类结果影响较大的是加速度传感器数据、陀螺仪数据以及旋转矢量传感器数据的特征值,这些都是动作行为传感器.使用不同分类算法对所有传感器数据特征值建模并采用十折交叉方法进行验证,表 3 为算法运行结果.

根据表 3 得出:IBk 算法对动态场景的分类感知效果最好.本文选取所有的传感器数据特征值建立全特征值集合,表 4 为使用全特征值集合时 IBk 分类算法分类结果的混淆矩阵.

Table 3 Classification Results of Dynamic Learning Situations

表 3 动态场景分类算法结果

Classification Algorithm	Accuracy/%	Modeling Time/s
RF	82.1	0.49
J48	81.8	0.47
RC	82.9	0.56
BN	78.5	0.00
IBk	84.3	0.00

Table 4 Confusion Matrix of IBk

表 4 IBk 算法分类结果混淆矩阵

Practical	Model					
	Bus	Car	Subway	Train	Plane	Others
Bus	147	13	6	0	0	4
Car	10	142	6	8	2	0
Subway	9	11	138	0	0	6
Train	8	0	7	137	4	5
Plane	0	0	15	10	139	4
Others	5	8	6	9	0	131

3 融合传感器与行为数据的场景感知优化

不同场景下学生的学习操作行为通常也呈现出明显的差异化.对 2015-08-12—2016-03-10 期间产生的 77875 条 Android 移动学习客户端的学习操作行为记录进行统计分析.图 3 是根据学习操作行为日志计算出的不同场景下的一些学习操作行为指标,包括学生在不同场景下的平均暂停和拖动的时长、平均暂停和拖动的次数以及平均学习时长;图 3 的横坐标表示不同的学习场景:1 代表办公室,2 代表家里,3 代表图书馆,4 代表室外,5 代表汽车,6 代表公交,7 代表飞机,8 代表地铁,9 代表火车,10 代表其他学习场景.

由图 3 可以看出不同场景下的学习行为呈现出很大的差异,例如学生在图书馆和办公室等安静的学习场所的单次学习时长通常比交通工具上的学习时长相对长些.在动态场景比如火车上的暂停时长相对长些.所以本文进一步考虑综合学习操作行为日志计算出的一些学习行为指标,将学习行为指标和传感器数据一起作为场景感知分类的特征值以达到更好的移动学习场景感知分类效果.

本文基于学习操作行为日志计算出特征值:某场景下的学习时长、暂停次数、暂停时长以及拖动次

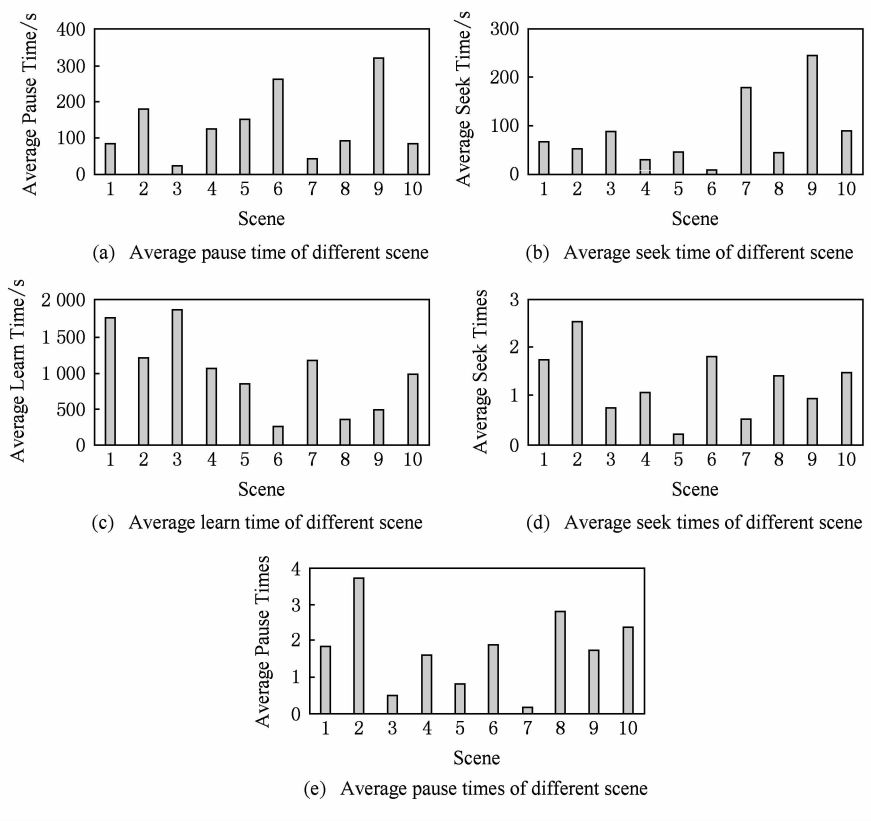


Fig. 3 Learning operation in different contexts.

图 3 不同场景下的学习行为

数、拖动时长共 5 个学习操作行为指标作为结合传感器数据特征值与学习操作行为日志的移动学习场景感知分类的特征值。

本次同样采用 2 次分类:第 1 次分类同第 2 节中介绍的一样,首先根据 3 轴加速度传感器数据的标准差来区分出静态场景和动态场景;第 2 次分类中算法的输入不只有传感器数据特征值,本节将计算出来的 5 个学习操作行为指标连同传感器数据特征值一起作为分类算法的输入,同样将场景类型数据作为训练集数据。

首先进行特征选择,将所有静态场景的 5 个学习操作行为指标连同 48 个传感器数据特征值导入 Weka,得到排名靠前的 10 个特征值,对比只有传感器特征值的情况,发现学习过程中的暂停次数、暂停时长、总的学习时长对于静态场景的分类也有影响,而拖动次数和拖动时长对静态场景分类结果影响不大。

采用 5 种不同分类算法对所有 48 个传感器数据特征值和 5 个学习行为指标进行建模,并采用十折交叉方法进行验证,算法运行结果如表 5 所示,结果表明 RC 算法对静态场景的感知分类效果最好。将算法结果与仅使用传感器数据特征值作为分类依

据的结果对比,每个算法的准确率都有不同程度的提升,RC 算法的准确率可以达到 88.5%,说明结合学习操作行为日志之后的场景感知分类效果更好。

Table 5 Classification Results of Static Learning Situations with Learning Operation Data

表 5 加入学习行为数据后的静态场景分类算法结果			
Classification Algorithm	Accuracy/%	Modeling Time/s	Accuracy Change/%
RF	86.5	8.51	2.9 ↑
J48	85.9	7.28	4.2 ↑
RC	88.5	10.50	2.7 ↑
BN	75.6	0.50	10.5 ↑
IBk	85.6	0.00	2.1 ↑

“↑” means the improvement of accuracy.

综合所有 48 个传感器数据与 5 个学习操作行为指标建立全特征值集合,表 6 为采用全特征值集合时 RC 分类算法分类结果的混淆矩阵。对比表 3 中数据,每个静态场景被正确分类的个数都有一定程度的增加,进一步验证了结合学习操作行为日志后可以提高学习场景感知的准确度。

Table 6 Confusion Matrix of RC with Learning Operation Data

表 6 加入学习行为数据后的 RC 算法分类结果混淆矩阵

Practial	Model			
	Home	Office	Library	Outdoor
Home	5 589	103	108	65
Office	214	5 862	125	92
Library	137	242	2 186	276
Outdoor	228	183	145	1 126

同基于传感器数据的场景感知分类方法一样，首先采用 smote 抽样方法使各样本的数据量保持大致平衡；然后进行特征选择，将所有动态场景的 5 个学习操作行为指标连同 48 个传感器数据特征值导入 Weka，得到排名靠前的 10 个特征值：加速度中值、加速度标准差、学习时长、加速度 4 分位点范围、暂停时长、陀螺仪偏度、拖动次数、拖动时长、加速度峰态、音量偏度。特征选择结果表明学习时长、暂停时长、拖动次数和拖动时长等学习操作行为指标对动态场景的分类结果影响较大。

采用不同分类算法对所有传感器数据特征值和学习操作行为指标数据进行建模，并采用十折交叉验证，算法运行结果如表 7 所示，可以看出 IBk 算法的准确率最高，分类性能最好，对动态场景的感知分类效果最好。

Table 7 Classification Results of Dynamic Learning Situations with Learning Operation Data

表 7 加入学习日志后的动态场景分类算法结果

Classification Algorithm	Accuracy/%	Modeling Time/s	Accuracy Change/%	Modeling Time Change/s
RF	85.5	0.65	3.4 ↑	0.16 ↑
J48	85.3	0.52	3.5 ↑	0.05 ↑
RC	86.2	0.98	3.3 ↑	0.42 ↑
BN	80.6	0.00	2.1 ↑	
IBk	88.7	0.00	4.4 ↑	

“↑” in the fourth column means the improvement of accuracy, in the fifth column it means the modeling time increased.

将算法结果与仅使用传感器数据特征值作为分类依据的结果对比，IBk 算法进行动态场景分类感知的效果最好，选取的学习操作行为指标可以明显提高动态场景感知分类的准确率，优化场景感知分类效果。

表 8 为采用 IBk 分类算法对动态场景进行感知分类结果的混淆矩阵。对比表 5 中的数据，每个动态场景被正确分类的个数都有一定程度的增加，进一

步验证了结合学习操作行为日志后可以提高学习场景感知的准确度。

Table 8 Confusion Matrix of IBk with Learning Operation Data

表 8 加入学习行为数据后的 IBk 算法分类结果混淆矩阵

Practial	Model					
	Bus	Car	Subway	Train	Plane	Others
Bus	154	9	4	0	0	3
Car	8	150	4	5	1	0
Subway	9	8	146	0	0	1
Train	5	0	7	144	0	5
Plane	0	0	12	10	146	0
Others	3	6	4	7	0	139

4 结束语

本文研究了基于移动学习系统的学习场景感知过程，通过采集学生使用 Android 移动学习客户端进行课程视频学习过程中的传感器数据和学习操作行为采集，并对原始数据进行一定的预处理和分析；然后选取合适的特征值来作为场景感知分类的依据，进而挖掘出用户所在的学习场景，为后续根据不同场景提供不同的学习内容以达到最优的学习效果提供一定的借鉴意义。

参 考 文 献

[1] Hao Jia, Zimmermann R, Ma H. GTube: Geo-predictive video streaming over HTTP in mobile environments [C] // Proc of the 5th ACM Multimedia Systems Conf. New York: ACM, 2014: 259-270

[2] Awobuluyi O, Nightingale J, Wang Qi, et al. Video quality in 5G networks: Context-aware QoE management in the SDN control plane [C] //Proc of the 3rd Int Workshop on Multimedia Communication & Networking. Piscataway, NJ: IEEE, 2015: 1657-1662

[3] Liu Shudong, Meng Xiangwu. Approach to network services recommendation based on mobile user's location [J]. Journal of Software, 2014, 25(11): 2556-2574 (in Chinese)
(刘树栋, 孟祥武. 一种基于移动用户位置的网络服务推荐方法[J]. 软件学报, 2014, 25(11): 2556-2574)

[4] 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

[5] Gellersen H, Beigl M, Schmidt A. Sensor-based context-awareness for situated computing [C] // Proc of Workshop on Software Engineering for Wearable and Pervasive Computing (ICSE'00). Limerick, Ireland: ICSE, 2000: 77-83

[6] Wang Yuankai. Context awareness and adaptation in mobile learning [C] // Proc of the 2nd IEEE Int Workshop on Wireless and Mobile Technologies in Education. Los Alamitos, CA: IEEE Computer Society, 2004: 154-158

[7] Krause A, Smailagic A, Siewiorek D P. Context-aware mobile computing: Learning context-dependent personal preferences from a wearable sensor array [J]. IEEE Trans on Mobile Computing, 2006, 5(2): 113-127

[8] Milette G, Stroud A. Professional Android Sensor Programming [M]. Indianapolis, Indiana: Wiley, 2012: 127-181

[9] Hernandez M, Stolfo S. Real-world data is dirty: Data cleansing and the merge/purge problem [J]. Data Mining and Knowledge Discovery, 1998, 2(1): 9-37

[10] Li Yuanyuan, Parker L. A spatial-temporal imputation technique for classification with missing data in a wireless sensor network [C] // Proc of IEEE/RSJ Int Conf on Intelligent Robots and Systems. Piscataway, NJ: IEEE, 2008: 3272-3279

[11] Kong Linghe, Xia Mingyuan, Liu Xiaoyang, et al. Data loss and reconstruction in sensor networks [C] //Proc of the IEEE INFOCOM'13. Piscataway, NJ: IEEE, 2013: 1654-1662

[12] Xu Ke, Lei Jianjun. Estimating algorithm for missing values based on attribute correlation in wireless sensor network [J]. Journal of Computer Applications, 2015, 35(12): 3341-3343 (in Chinese)
(许可, 雷建军. 基于属性相关性的无线传感网络缺失值估计方法[J]. 计算机应用, 2015, 35(12): 3341-3343)

[13] Pan Liqiang, Li Jianzhong, Luo Jizhou. A temporal and spatial correlation based missing values imputation algorithm in wireless sensor networks [J]. Chinese Journal of Computers, 2010, 33(1): 1-11 (in Chinese)
(潘立强, 李建中, 骆吉州. 传感器网络中一种基于时-空相关性的缺失值估计算法[J]. 计算机学报, 2010, 33(1): 1-11)

[14] Hong Jongyi, Suh E H, Kim S J. Context-aware systems: A literature review and classification [J]. Expert Systems with Applications, 2009, 36(4): 8509-8522

[15] Zhang Mi, Sawchuk A. A feature selection-based framework for human activity recognition using wearable multimodal

sensors [C] //Proc of the 6th Int Conf on Body Area Networks. Brussels: ICST, 2011: 92-98

[16] Wang Shuo, Yao Xin. Diversity analysis on imbalanced data sets by using ensemble models [C] //Proc of the 2009 IEEE Symp on Computational Intelligence and Data Mining. Piscataway, NJ: IEEE, 2009: 324-331



Ye Shuyan, born in 1991. Master candidate. Her main research interest is multimedia systems for e-learning.



Zhang Weizhan, born in 1977. Associate professor and PhD supervisor. His main research interests include multimedia systems for e-learning, peer-to-peer computing, analysis and application of big data and wireless networks.



Qi Tianliang, born in 1993. Master candidate. His main research interests include personalization service in e-learning, analysis and application of big data.



Li Jing, born in 1993. Master. Her main research interest is multimedia systems for e-learning.



Zheng Qinghua, born in 1969. Professor, PhD supervisor, and vice-president of Xi'an Jiaotong University. His main research interests include multimedia distance education, computer network security, intelligent e-learning theory and algorithm.