

一种增强现实分场景推送情景感知服务的方法*

林一^{1,2}, 刘越^{1,2}, 王涌天^{1,2}, 贺长宇^{1,2}

¹(北京理工大学 光电学院, 北京 100081)

²(北京市混合现实与新型显示工程技术研究中心(北京理工大学), 北京 100081)

通讯作者: 刘越, E-mail: liuyue@bit.edu.cn



摘要: 针对目前增强现实浏览器中由于标签使用不当阻碍用户认知操作的问题,提出了一种分场景推送情景感知服务的方法.该方法根据用户寻找兴趣点并检索内容的认知流程划分场景,采用情景感知四层服务框架分别创建了3个场景的功能模块,通过关联3个场景模块,构建了完整的移动增强现实浏览器系统.实验结果表明,与同类浏览器相比,基于上述方法设计的增强现实浏览器系统在分类准确度上平均提高了13%,用户对系统预测效果的平均满意度提高了26%.

关键词: 移动增强现实浏览器;情景感知服务;用户体验;认知流程
中图法分类号: TP391

中文引用格式: 林一,刘越,王涌天,贺长宇.一种增强现实分场景推送情景感知服务的方法.软件学报,2016,27(8):2115–2134. <http://www.jos.org.cn/1000-9825/4846.htm>

英文引用格式: Lin Y, Liu Y, Wang YT, He CY. Method of pushing context-aware service of augmented reality by scene classification. Ruan Jian Xue Bao/Journal of Software, 2016, 27(8): 2115–2134 (in Chinese). <http://www.jos.org.cn/1000-9825/4846.htm>

Method of Pushing Context-Aware Service of Augmented Reality by Scene Classification

LIN Yi^{1,2}, LIU Yue^{1,2}, WANG Yong-Tian^{1,2}, HE Chang-Yu^{1,2}

¹(School of Optics and Electronics, Beijing Institute of Technology, Beijing 100081, China)

²(Beijing Engineering Research Center of Mixed Reality and Advanced Display (Beijing Institute of Technology), Beijing 100081, China)

Abstract: This dissertation presents a method of pushing context-aware service based on scene classification to overcome the improper use of labels in augmented reality browsers caused by the interference with user's cognitive operations. User's scene is separated by using the proposed method on the basis of user's cognitive processes when searching and retrieving points of interest. The function module of three scenes is built by using the framework of four-layer context-aware services respectively, then the overall construction of the augmented reality system is completed by associating three scenario modules. Experimental results of comparison with the existing popular browsers show that the classification accuracy of the system constructed by using the proposed method is improved by 13% on average, and the user's average satisfaction with the prediction results of the system is increased by about 26%.

Key words: mobile augmented reality browser; context-aware service; user experience; cognitive processes

随着计算技术、通信技术、网络技术和半导体技术的发展,在经历了主机计算和桌面计算的阶段后,计算模式已步入普适计算的时代.普适计算首先由美国学者 Weiser 提出^[1],IBM 亦于 1999 年概括了普适计算的概念,即,计算资源普遍存在于环境之中.普适计算中存在的大量数据,只有在经过数据处理后才能被用户使用.推荐系统^[2]能够建立用户与检索项之间的二元关系,通过分析用户操作习惯,可从海量数据中寻找并推送用户感兴趣

* 基金项目: 国家高技术研究发展计划(863)(2013AA013904); 国家自然科学基金(61370134)

Foundation item: National High-Tech R&D Program of China (863) (2013AA013904); National Natural Science Foundation of China (61370134)

收稿时间: 2014-12-19; 修改时间: 2015-03-09; 采用时间: 2015-03-26

趣的内容,缓解了普适计算信息过载的状况。

情景描述的是一个设备或用户所处的态势和环境^[3],普适计算随时随地的特点指的就是情景。Dey^[4]认为,用桌面计算机和移动设备在普适计算环境中使用情景的应用都可以被称为情景感知。Byun^[5]定义了情景感知系统,即,如果一个系统能抽取、解释并运用场景信息适配于当前情景所适用的功能,它就是情景感知系统。引入情景感知计算的推荐系统被称作情景感知推荐系统(context-aware recommender system,简称 CARS),它不但满足了个性化的需求,而且也提高了检索的精度,体现了普适计算随时随地为用户提供定制化信息的核心思想。

移动增强现实浏览器作为典型的 CARS,收集传感器从周边环境捕获的情景数据,通过情景推理检索出用户当前需要的信息,以增强现实的交互形式推送并展示服务信息。早期运用情景感知计算的移动导览系统功能单一,1996年,美国佐治亚理工的 Gregory 等人^[6]提出的移动情景感知游览向导 Cyberguide 能够依据用户所在位置推送简单的情景服务。该系统未将各类传感设备集成到移动终端内部,给使用带来不便。Raento 等人^[7]以用户为中心开发的 ContextPhone 移动软件平台能够帮助开发者轻松创建应用程序,将现有技术集成到日常生活中。2010年,Coppola 等人^[8]重新定义了情景感知浏览器(context-aware browser,简称 CAB)的组成,作者提出的 MoBe 系统加入搜索引擎、网页内容和自动下载模块,能够根据环境的变化对网页内容进行更细化的检索;同时,通过检索和滤波处理减小了传输带宽,降低了隐私和安全风险。Van Woensel 等人^[9]提出了名为 SCOUT 的移动应用框架,在框架结构中,系统依据检测层获取的实体信息在定位层中采用亲近/偏远的策略,然后在环境层中创建实体模型与关系模型,根据环境模型的变化获取信息;最后将查询结果传递给应用层。然而,该研究仅改进了中间件,缺乏对连接中间件的增强现实可视化的研究。Lee 等人^[10]提出了一种部署情景感知框架的新方法。这个框架由“情景感知与自适应层”、“接口层”和“基于增强现实的可视化与交互层”结构组成。由新方法构建的浏览器在车辆服务、工程协作与家庭服务等 3 类情景环境中能提供虚实融合的增强现实可视化交互操作。不足之处在于:增强现实可视化与交互是建立在有标志点的基础上,在无标志点的户外复杂环境下,该系统将难以维持正常的工作。由于户外环境下的增强现实交互需要结合基于位置服务(location based service,简称 LBS)与基于图像的方法进行定位与识别,这项研究成果所提出的框架结构还存在局限,只有在重新调整结构之后,才能保证对户外增强现实服务的支持。

随着移动计算的发展,使用移动终端随时随地联网互动成为普遍现象,融入增强现实技术的移动浏览器以其便捷的服务模式迅速得到广泛认可。然而 2011 年,奥地利格拉茨大学的 Grubert 等人^[11]的调查显示:增强现实服务在吸引用户关注的同时,却出现了使用频率低和使用持续时间短的现象。这种矛盾状况反映在现实中,表现为增强现实浏览器安装率高而使用率低。Olsson 等人^[12]在线调查了用户对移动增强现实应用的想法,结果显示:用户对增强现实应用的新颖性表现出极大地好奇,然而使用上的不稳定是最主要的薄弱点。Dunser 等人^[13]在关于增强现实领域评估技术的论文中指出,当前非常缺乏对人机交互的研究和增强现实实证研究。由软硬件多方面因素造成的地理信息标签的误定位带来的感知混淆、由交换数据在移动网络中传输不流畅造成的显示中断与操作迟滞、由不合理的界面设计给用户带来的认知负担等,这些都会影响用户对增强现实浏览器的交互体验。因此,如何运用认知科学的理论改进增强现实浏览器的系统设计理念、调动用户使用积极性,是本文所要研究的问题。

综上所述,本文旨在改善移动增强现实浏览器的用户感受,从源头分析用户需求,发现用户的认知规律,提出一种分场景推送情景感知服务的方法。该方法根据用户寻找兴趣点并检索内容的认知流程划分场景,采用情景感知四层服务框架创建了 3 个场景的功能模块,关联 3 个子模块完成对移动增强现实浏览器系统的整体构建。使用新方法构建的移动增强现实浏览器系统一方面提升了移动增强现实浏览器分类准确度,另一方面改善了用户在检索兴趣点信息过程中的体验。

1 增强现实浏览器的情景感知服务框架

情景感知服务通过传感器和无线通信等技术手段使便携的移动计算设备实时感知当前所处的环境,在对场景信息做出相应判断之后,系统主动向使用者推送服务,将增强现实技术应用于智能手机平台后,可以运用移

动增强现实浏览器为用户提供情景感知服务.增强现实浏览器的情景感知服务框架如图 1 所示,该框架由情景感知信息供应层、情景感知信息采集层、情景感知服务层和增强现实服务层这 4 个层次组成,其主要功能见表 1.

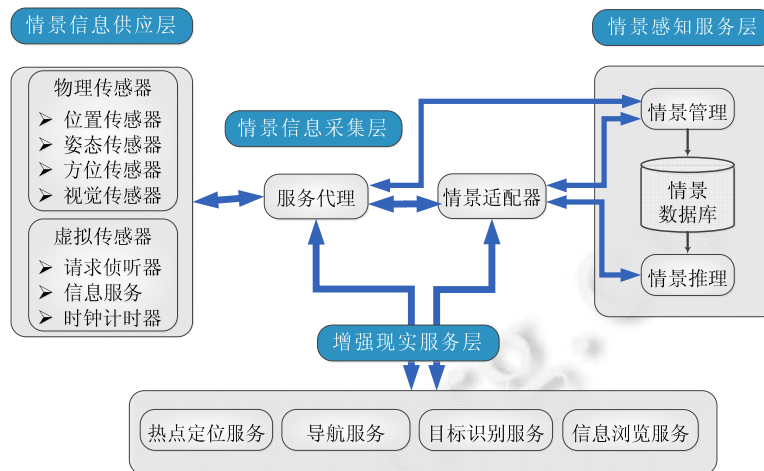


Fig.1 Context-Aware services framework of augmented reality browser

图 1 增强现实浏览器的情景感知服务框架

Table 1 Context-Aware service layer and its main functions

表 1 情景感知服务层及其主要功能

感知服务层	主要功能
情景信息供应层	提供情景的物理与虚拟传感数据信息
情景信息采集层	采集与过滤传感数据,传递请求和命令
情景感知服务层	管理场景,依据数据库推理并聚合场景信息
增强现实服务层	面向情景提供增强现实服务

下文对情景感知服务框架的具体组成部分及其工作原理进行详细叙述.

1.1 情景信息供应层

嵌入了计算、信息设备和传感器装置的工作或生活空间被称为智能空间,它具有通用和便捷的互动接口,为用户能够便利地获得计算机系统的服务提供支持^[14].在系统对智能空间中的情景信息做出正确推断后,情景感知服务模块才能主动地向用户推送个性化服务.场景信息由众多复杂的原始传感器数据归纳而来,在情景感知服务框架中的所有层次,都是建立在基础的情景感知信息供应层之上.

根据属性的不同,移动增强现实浏览器情景感知信息供应层的传感信息可分为硬件与软件两类数据.因此,情景感知信息供应层的数据包含物理设备捕获的传感数据和软件模块生成的虚拟传感数据两部分.

(1) 硬件捕获的传感数据

如表 2 所示,智能手机的物理传感器包括位置传感器、方向传感器和光学传感器等.其中,方向传感器使用手机内嵌的 GPS 模组进行地理定位,通过 WGS-84 地心坐标系(world geodetic system—1984 coordinate system)表征手机的位置传感数据^[15]获取手机的位置传感数据.方向传感器包括三轴陀螺仪、三轴加速度计和三轴磁传感器等,其中,三轴陀螺仪采集手机的旋转角加速度,通过积分可以得到手机的姿态角;三轴加速度计感知重力加速度在体坐标系 3 个轴上的分量,然后,通过余弦算法计算出手机的俯仰角和横滚角;三轴磁传感器测量地球磁场在体坐标系 3 个轴的分量,基于俯仰角和横滚角得到手机当前的航向角^[16].光学传感器通过光学器件采集到达感光元件的光信息,生成并发送图像数据^[17].智能手机可以通过自带的视觉传感器,即摄像头捕获图像或视频,并将传感数据用于增强现实服务的图像处理与跟踪.

Table 2 Physical sensors, corresponding devices and sensing data of smart phone**表 2** 智能手机的物理传感器、对应设备及传感数据

传感器	物理设备	传感数据
位置传感器	GPS 模组	经纬度数值和海拔信息
方向传感器	陀螺仪、加速度计和地磁场传感器	方向角、俯仰角和横滚角
光学传感器	手机摄像头	图像或视频数据

本文对世界坐标系、手机体坐标系和手机屏幕坐标系定义如下:

- 如图 2(a)所示,世界参考坐标系中 X 轴指向地球磁北方向、 Z 轴垂直指向地心方向, X 轴和 Z 轴向量叉积可得出 Y 轴方向;
- 手机体坐标系中 X 轴水平垂直向上, Y 轴水平向左, Z 轴垂直于屏幕指向屏幕背面之外,如图 2(b)所示;
- 手机屏幕坐标系主要用于手机屏幕上定位并绘制增强信息,如图 2(c)所示,将屏幕左上方作为坐标系的原点,水平向右为 X 轴正方向,垂直向下为 Y 轴正方向.

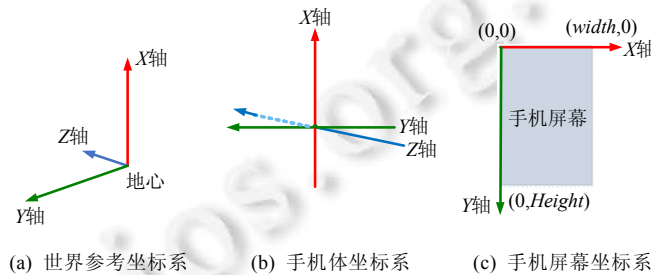
**Fig.2** Coordinate systems of mobile sensors

图 2 手机传感器的坐标系统

手机的旋转可以用欧拉角和四元数两种方式表示.使用欧拉角表示物体(如手机)的旋转时通过以下 3 个步骤实现:首先,绕手机体坐标系的 Z 轴旋转 ψ ;然后,绕手机体坐标系的 Y 轴旋转 θ ;最后,绕手机体坐标系的 X 轴旋转 γ .坐标系的转换可以通过旋转矩阵 R 表示,具体计算公式如式(1)所示.

$$R = \begin{bmatrix} \cos\psi & 0 & -\sin\psi \\ 0 & 1 & 0 \\ \sin\psi & 0 & \cos\psi \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\gamma & \sin\gamma \\ 0 & -\sin\gamma & \cos\gamma \end{bmatrix} \quad (1)$$

ACC_x , ACC_y 和 ACC_z 是重力加速度在加速度计 3 个轴上的分量,通过 X 轴的重力加速度分量,可以由公式(2)得出俯仰角 θ .

$$\theta = \arcsin(ACC_x) \quad (2)$$

根据 Y 轴的重力加速度分量 ACC_y 和俯仰角 θ ,由公式(3)可以计算出横滚角 γ .

$$\gamma = \arcsin\left(\frac{ACC_y}{g \cos(\theta)}\right) \quad (3)$$

其中, g 表示重力加速度常量.之后,利用磁传感器 3 个轴上的分量 MAG_x , MAG_y 和 MAG_z 可以计算出手机的航向角 ψ ,具体计算公式如(4)所示.

$$\psi = \arctan\frac{M}{N} \quad (4)$$

其中,

$$M = MAG_x \cos(\theta) + MAG_y \sin(\theta) + MAG_z \cos(\theta) \sin(\gamma) \quad (5)$$

$$N = MAG_y \cos(\gamma) + MAG_z \sin(\gamma) \quad (6)$$

使用欧拉角表示物体旋转时,会不可避免地出现奇点问题,即,公式(3)中,当 $\cos(\theta)=0$ 时,计算结果会产生奇

异值.为避免出现该问题,可以使用如公式(7)所示的四元数法表示物体旋转.

$$R = \begin{bmatrix} q_1^2 - q_2^2 - q_3^2 + q_4^2 & 2(q_1q_2 - q_3q_4) & 2(q_1q_3 + q_2q_4) \\ 2(q_1q_2 + q_3q_4) & -q_1^2 + q_2^2 - q_3^2 + q_4^2 & 2(q_2q_3 - q_1q_4) \\ 2(q_1q_3 - q_2q_4) & 2(q_2q_3 + q_1q_4) & -q_1^2 - q_2^2 + q_3^2 + q_4^2 \end{bmatrix} \quad (7)$$

其中, q_1 、 q_2 和 q_3 表示四元数中的虚数部分, q_4 表示四元数中的实数部分.四元数状态方程表示为

$$X = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ q_4 \end{bmatrix} \quad (8)$$

可以使用三轴陀螺仪的测量值 $GYRO_x$ 、 $GYRO_y$ 和 $GYRO_z$ 对四元数值进行更新,通过公式(9)中的增益矩阵 A ,可以得到手机当前的姿态信息.

$$A = \begin{bmatrix} 1 & dGYRO_z & -dGYRO_y & -dGYRO_x \\ -dGYRO_z & 1 & dGYRO_x & -dGYRO_y \\ dGYRO_y & dGYRO_x & 1 & -dGYRO_z \\ dGYRO_x & dGYRO_y & dGYRO_z & 1 \end{bmatrix} \quad (9)$$

其中, d 表示惯性传感器每帧的刷新时间.

由于使用陀螺仪角速度积分得到的姿态角会受到随时间增加的漂移误差影响,需要使用卡尔曼滤波器^[18]对漂移误差进行抑制.综上所述,通过卡尔曼滤波器融合陀螺仪、加速度计和磁传感器测量数据得到精度较高的手机姿态角.

(2) 软件生成的传感数据

虚拟传感器是依照一定的机制进行数据侦听和事件处理的软件模块,为情景感知提供模拟的传感数据,通过从已知的网络服务和信息服务中抽取情景感知信息,用于对高层的场景信息的推理与融合.

1.2 情景信息采集层的服务代理

情景感知信息采集层收集情景感知信息供应层提供的所有传感数据,并将所收集的数据传递给情景感知服务器,同时接收情景感知服务器发来的控制请求并传递给传感器层.情景感知信息的采集分为主动模式和被动模式两种基本形式,其中:主动模式主动采集传感器信息并上传给情景感知中间件,例如,按照系统时钟向服务器定时发送关键帧图片进行图像识别处理;而被动模式的采集以控制命令的方式向传感信息供应层索取情景感知信息,例如,当服务层的应用程序发送读取当前位置地理所在信息的请求时,系统采集手机的 GPS 经纬度数据.

情景感知信息采集层的主要工作由服务代理模块完成,如图 3 所示.

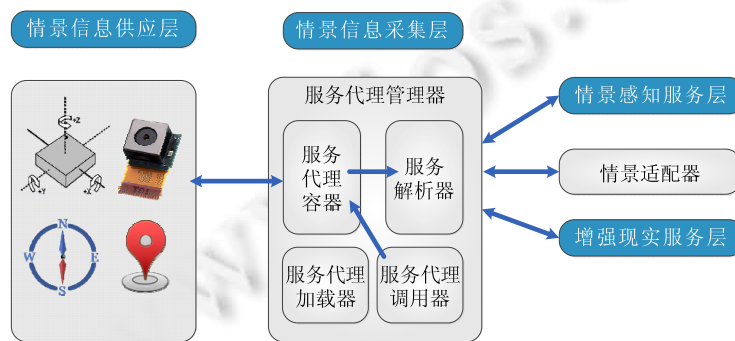


Fig.3 Service agent structure

图 3 服务代理结构

服务代理将外部传感器和设备连接到服务框架内,向情景感知适配器、情景感知服务层和增强现实服务层传输传感器信息;同时,按照情景感知服务层、情景感知适配器和增强现实服务层发送的控制命令操纵设备并处理应用.服务代理加载器将各种服务代理加载到服务代理容器中,之后,通过传感器解析器过滤或聚合服务代理获取的信息.这些被解析的信息以事件的形式被传输给服务器,最后由服务代理调用器发出服务器提出的服务代理请求^[19].

1.3 情景感知服务层

情景感知服务层处理接收到的情景感知信息,基于事件的数据通过情景感知规则的推理推断出需要执行的服务,并调用服务数据库的情景感知服务数据.图4显示出了情景感知服务层结构,其中,情景感知管理器由事件管理器和任务管理器组成.一旦场景状态改变,真实环境中的传感数据将发生变化,采集层的服务代理将收集到的所有情景感知信息传递给事件管理器,并随之生成一个事件.每一个事件带有时间、地点、人物等参数.情景感知管理器基于事件的数据推断出隐性知识,并推送到任务管理器执行.任务管理器依照事件-状态-行动(event-condition-action,简称ECA)的任务规则处理每一个任务,其中,事件-状态用来表示环境和用户/程序的需求的变化,行动表示当事件和状态满足条件时采取的任务^[20].

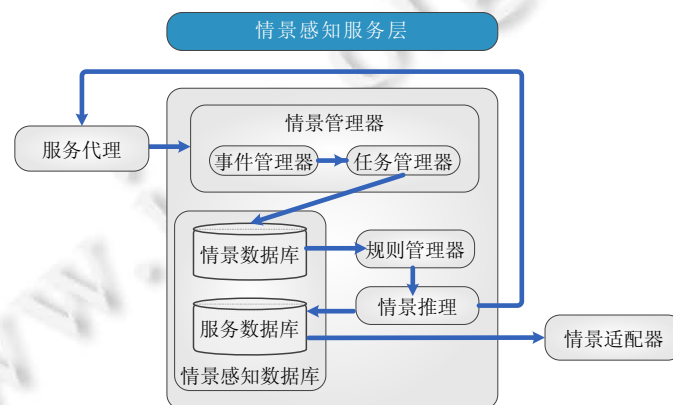


Fig.4 Context-Aware service layer structure

图4 情景感知服务层结构

当环境发生变化促使新的情景感知数据被存储到数据库中时,只要任务管理器发起推理任务,就会立刻触发情景感知推理操作.任务管理器发起任务时,调用规则管理器将抽象的情景感知数据推送到推理器中进行查找和匹配运算的情景感知推理.根据情境模型的不同,主要有基于规则和基于本体的两种推理方法.其中,

- 基于规则的推理方法的实现有两种方式:一种是将逻辑规则写入程序代码中,通过程序自动实现推理;另一种是通过逻辑规则和推理引擎共同实现推理^[21];
- 基于本体的推理方法通过本体描述语言将顾客的情境、情境之间的关系和推理规则定义成本体、本体的属性和约束关联关系,通过本体的推理实现对情境的推理^[22].

1.4 情景适配器

软件自适应是指为了保证持续、高质量地提供服务,软件在运行时检测环境变化和自身状态,据此对自身行为进行主动调整的活动^[23].增强现实浏览器应当根据用户当前所处的情景感知提供高层次个性化的服务.然而,该服务一方面极其依赖手机性能,另一方面随场景实际状况的不同不断发生改变,因此,采用自适应服务的情景感知适配器具有重要意义.情景感知适配器的结构如图5所示,从中可以看出,情景感知适配器与情景感知服务器进行通信获取查询和推理所需要的情景感知信息,并进而提供符合自适应情景感知的服务.情景感知适配器包括两组传输通道模块和一个服务内容生成模块,其中,服务内容生成模块处理的结果被情景感知适配器

传递给服务代理和增强现实服务层.

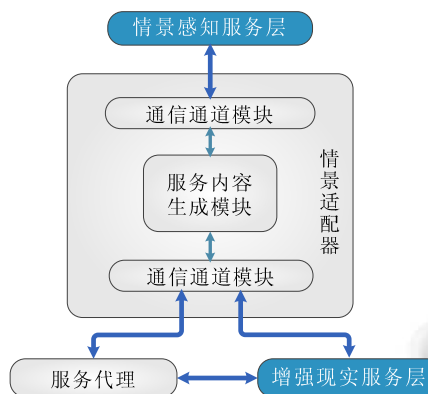


Fig.5 Structure diagram of context-aware adaptor

图 5 情景适配器结构图

1.5 增强现实服务层

增强现实服务层为用户提供具有虚实融合特性的个性化的普适服务,目前,增强现实技术使用基于地理位置信息的方法进行目标定位,使用基于图像识别的方法进行兴趣点识别.基于图像识别的方法又分为有标志点和无标志点识别两类.本文研究的移动增强现实浏览器侧重于户外环境下的应用,其增强现实服务依托于 LBS 模式,使用无标志点识别的方法识别兴趣点,因此,增强现实服务层的服务可划分为兴趣点定位、导航、目标识别和信息浏览这 4 种.为便于用户对兴趣点的逐步认知与理解,图 6 显示出了增强现实服务层中情景感知服务的推送流程.

- 首先,根据设备所在地理位置定位用户周边的兴趣点;
- 然后,为用户提供到达被选中兴趣点的路径导航服务;
- 用户到达兴趣点区域后,通过目标识别服务获取被扫视建筑物的信息;
- 最后,得到建筑物详细信息的浏览服务.



Fig.6 Context-Aware service push process in augmented reality service layer

图 6 情景感知服务在增强现实服务层中的推送流程

增强现实空间是虚拟元素与真实环境相互融合的环境,增强现实服务层从情景感知适配器得到匹配各自服务的情景感知服务内容,这些虚拟元素包括二维图片文字标签、三维模型和交互控制组件(按钮、对话框等),同时也直接从情景感知供应层获取能够表现增强现实空间中真实环境所需的传感信息,如表达视觉环境的视频背景层需要手机摄像头的图片序列帧数据.

2 多场景情景感知的移动增强现实浏览器的构建

2.1 系统总体架构

本文提出的虚拟现实与增强现实混合式移动导览系统(mobile augmented reality browser,简称 MARB)对目前业内常用的定位、识别后再跟踪的技术方案^[24,25]进行了改进.首先,使用 GPS 和传感器对指定范围内的所有兴趣点进行粗略定位,当用户选中目标后自动生成导航路径,在用户到达兴趣点所在地后进行观察点周围的复

杂场景的识别以及像素级的精确跟踪.以下的系统框架主要针对场景识别和精确跟踪注册进行论述.

本文构建的 MARB 系统框架如图 7 所示,浏览器系统运行后首先开启视觉传感器捕获具有环境特征的关键帧图像并传输至场景识别服务器;之后,已预先完成样本训练的服务器对图像进行快速识别,将以字符串编号形式表示的场景识别结果和场景特征信息传输给图像渲染单元;情景感知推理器根据场景编号的变化判断当前情景感知是否已被改变.在使用场景特征信息和摄像头后续传递的序列帧计算出姿态数据后,跟踪注册模块将旋转、平移信息传递给渲染模块实现虚拟标签的渲染和实时跟踪显示.当用户点击虚拟标签时,渲染模块依照场景编号从数据服务器下载该场景对应的图片、文本或模型信息,然后通过图像渲染单元分类将这些场景信息绘制展示.

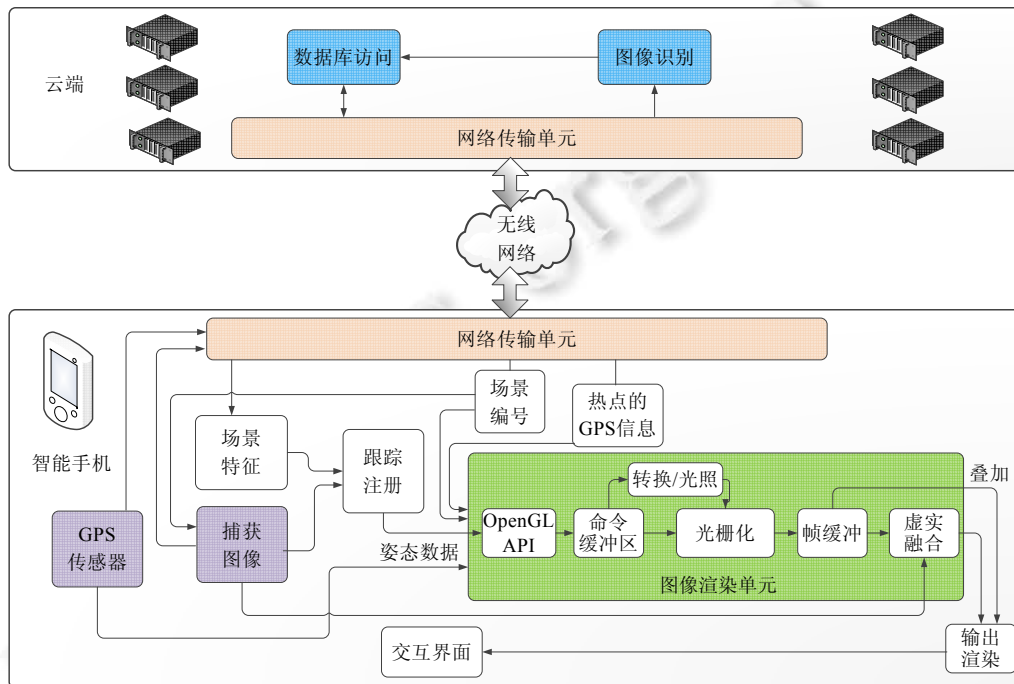


Fig.7 Overall framework diagram of the MARB system

图 7 MARB 系统总体框架图

2.2 信息表示与场景划分

2.2.1 增强现实内容的信息表示

增强现实通过向嵌入物理世界的三维信息空间提供直观和自然的实时交互来实现普适计算.为了使用户更好地体验系统推送的情景感知服务,移动增强现实浏览器模块的构成不仅应当依照情景感知的服务框架进行组织,还需要考虑对情景感知信息的表达方式进行合理的规划.注释(annotation)作为增强现实技术中提供物理环境相关信息的最佳表现形式,已成为当前增强现实系统中增进对真实世界理解的一个主流组件.对服务信息的表示包含数量、表现、布局和视图组合这 4 个检验指标^[26].

(1) 数量指标

如果真实世界背景和虚拟信息的复杂度在增强现实中不能达成平衡,合成的场景会阻碍用户的认知功能.在一个界面内同时显示大量信息,会导致用户难以集中注意力,从而增加了认知学习的精神负担.

(2) 表现指标

标签的表现可分为二维和三维两种,其中,二维标签包含文字、图片或图片组、视频和图片文字混合的信息提示形式,三维标签包含三维形态的平面标签和立体模型两种呈现形态.

(3) 布局指标

一方面,注释与目标物距离的最大值将确保信息的可读性;另一方面,坐标系的选择会影响到系统对于用户视域内物体可视部分的识别和标签的放置^[27]。目前常用的移动增强现实浏览器系统有两类坐标系:一种是基于用户头部的(head-stabilized)坐标系,即以用户视点为基准放置标签;另一种是基于世界的(world-stabilized)坐标系,即以真实世界的位置为基准放置标签^[28]。

(4) 视图组合指标

采用缩放与平移、总图与细节和焦点与上下文这 3 种可视化技术^[29]。在单一显示屏幕上组合多个视图,这不但可以充分利用显示空间,而且提高了用户搜寻信息的效率。

表 3 给出了主流移动增强现实浏览器的注释对比结果。通过对目前国内外主流的 6 款移动增强现实浏览器的注释进行调查与研究,发现了浏览器使用注释的规律,即,标签由 2~3 个控制标签和多个兴趣点标签组成,其中:控制标签包含菜单、雷达图和被选中兴趣点的概要说明;兴趣点标签分为图文混合的二维标签和基于标志点的三维模型,兴趣点标签的数量一般不超过 15 个。为了方便用户对空间方位的理解,浏览器标签的布局全部采用基于用户视点的坐标系放置标签。缩放与平移没有作为浏览器的视图组合方式出现,总图与细节因其直观和易于表现的特点被所有浏览器采用,部分浏览器在地图导航服务中还采用了焦点与上下文的组合方式。

Table 3 Comparison of annotation in popular augmented reality browser

表 3 主流增强现实浏览器的注释对比

	Juniao	Layar	Wikitude	City Lens	触景	城市一点即通
数量	2+POI 20>POI>4	3+POI 7>POI>2	3+POI 5>POI>2	3+POI 17>POI>5	2+POI 5>POI>1	2+POI 12>POI>1
表现	图文混合 三维模型	图文混合 三维模型	图文混合	图文混合	图文混合 视频	图文混合
布局	基于用 户视点	基于用 户视点	基于用 户视点	基于用 户视点	基于用 户视点	基于用 户视点
视图组合	总图与细节	总图与细节 焦点与上下文	总图与细节	总图与细节	总图与细节	总图与细节 焦点与上下文

从表 3 的分析结果可以看出,当前主流的移动增强现实浏览器存在以下 3 点问题。

- 为缓解界面交互元素的复杂度,目前大部分浏览器筛选兴趣点标签显示数量,如按兴趣点类型(娱乐、餐厅等)或距离范围(公里数)过滤标签个数。采用动画组织形式简化标签数量,如,“城市一点即通”通过合并重叠标签的方法避免兴趣点的遮挡。但当兴趣点标签多达 20 个的时候,仍然会给用户的认知体验带来较大的不便。
- 在标签的表现上,图文混合的模式能够较好地为用户传达兴趣点的简要信息。有的增强现实浏览器使用不同的颜色块作为兴趣点标签,这将导致交互界面的混乱。虽然 City Lens 的标签用颜色的不同区分兴趣点的种类,然而过于繁杂的色相会增加用户对目标的认知成本,降低人眼检索目标的时效,因此,采用色相一致且有别于背景色彩的标签配色方案才是较为可取的方法。
- 表 3 中的移动增强现实浏览器对兴趣点的定位都采用 GPS 定位的方式辨认周边兴趣点,对兴趣点的跟踪则通过对手机姿态数据的反推计算出兴趣点与观察视点之间的距离和方位。基于机器视觉的图像识别常被运用于室内环境中,在有标物体上叠加三维模型。虽然目前户外环境下,基于机器视觉的无标识别的方法还存在许多暂未解决的困难,如光照变化、遮挡以及传输效率等,但它在识别精度上高于当前常用的 GPS 定位的方法。

2.2.2 MARB 场景的划分

针对上述问题,本文在搭建移动增强现实浏览器系统时提出了场景的概念。此处所述的场景可以理解为对情景感知的划分,用户为达成某个目标必须执行多个步骤的任务,而任务在执行过程中所处的环境(情景感知)会发生变化。在一个环境中,把用户为实现某个目标需要执行的多组任务定义为一个场景。

由于用户计划完成一个目标时将经历在不同环境下执行多组任务的过程,因此,系统可以依据对当前环境

的感知与推理向用户提供服务来协助用户完成不同的任务.使用场景对任务进行分组不但便于系统向用户推送感知服务,同时有利于用户在交互环境中加深对操作对象的认知,运用多场景的概念能有效解决当前移动增强现实浏览器存在的问题.一方面,就信息表示的数量指标而言,由于单一场景中展示的所有标签被归类并分散到多个场景中,降低了交互界面的复杂度,并从操作逻辑上简化了用户主动区分兴趣点类型的认知过程;另一方面,单场景多任务分解为多场景多任务降低了系统情景感知推理规则的复杂度,使得情景感知推荐系统能够针对特定的场景推送定制的服务.

增强现实浏览器采用兴趣点方式为用户提供获取信息的入口.兴趣点的种类繁多,包含房产、娱乐、餐饮、购物和交通等类型.每种类型的兴趣点又囊括了海量的信息,例如,房产类型的兴趣点包含了楼盘信息、小区环境、户型结构等子类数据.使用场景划分的方法将众多的兴趣点子类信息加以整理并归类到不同场景中,用户在交互过程中逐步地获取信息并最终达成需求目标.为满足用户了解兴趣点详细信息的需求,首先将操作任务进行有序的划分,然后确定不同的场景,从而针对具体场景制定相应的情景感知服务.本文以查询房产信息为例,以用户执行任务的行为逻辑为次序,将众多任务划分为兴趣点定位与选择、寻路导航与目标识别这 3 个场景.其中,兴趣点定位与选择场景仅推送兴趣点类型选择和兴趣点概况展示服务,由用户自行比较和筛选后选择中意的兴趣点;寻路导航场景提供从用户当前所在位置到达被选中兴趣点的地图导航服务;目标识别与浏览场景为用户使用摄像头扫视的建筑物提供目标识别服务,将被识别建筑物的详细信息分类并展示.

2.2.3 MARB 的情景感知服务与模块构成

(1) 兴趣点定位与选择

兴趣点定位与选择是用户启动 MARB 后最先进入的场景,用户在该场景内的目标是查找并选定兴趣点,因此,用户首先执行的任务是了解自身所处的位置;然后,以所在位置为搜索圆心扫描指定半径范围内的所有兴趣点;最后,在查看兴趣点信息概况后做出选择决定.根据用户在此场景中需要执行的任务,MARB 必须提供用户定位、兴趣点类型选择、兴趣点定位与概况展示这 3 项情景感知服务.

MARB 系统的兴趣点定位与选择场景的模块构成如图 8 所示,手机客户端包括定位与跟踪模块、图像捕获模块、情景感知适配器和渲染模块,情景感知服务层的位置服务数据库与情景感知推理器分布于云端.

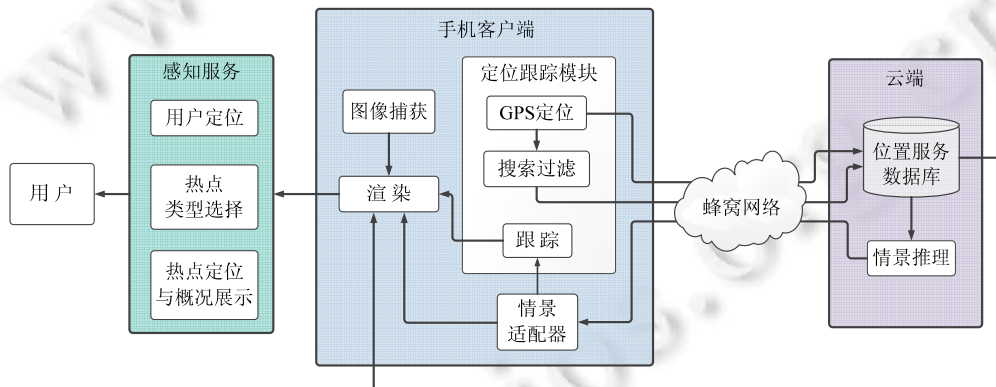


Fig.8 Module structure and context-aware service in the scenario of POI's positioning and selection

图 8 兴趣点定位与选择场景的模块构成及情景感知服务

为使 MARB 分析传感数据并更好地推算出用户关注的信息,在设计情景感知服务模块的构成之前,应预先通过操作次序的设定引导用户将需求以分类数据的形式提供给情景感知服务层,在兴趣点定位与选择场景中,使用菜单过滤方式实现了系统对用户于当前环境下服务诉求的收集,通过定位与选择数据的抽取,能够有效地缩小 MARB 在情景感知服务推理阶段所需参照的判断条件的范围.运行 MARB 后,手机摄像头与 GPS 开启,摄像头捕获的传感信息以序列帧图像的形式用于增强现实场景的背景层,GPS 向云端的位置服务数据库索取用

户当前位置的经纬度数据,渲染模块使用位置数据,以标签形式表现用户与兴趣点的位置关系.以用户所在位置为中心,MARB 调用搜索过滤模块,以范围(公里数)和类型(兴趣点类型)为参数进行筛选,并将搜索结果发送给云端的位置服务数据库.在完成情景感知推理后,云端把数据库查询结果以可扩展标记语言(extensible markup language,简称 XML)表单形式传递给情景感知适配器,之后,情景感知适配器将 XML 表单解析为兴趣点经纬度和兴趣点概述.兴趣点的概述数据被情景感知适配器发送到渲染模块用于标签的绘制,兴趣点的经纬度数据传输给跟踪模块后,跟踪模块先推算出标签在手机屏幕内绘制的位置,然后参照方向传感器捕获的数据计算出标签的实时跟踪数据,最后将位置数据与跟踪数据发送给渲染模块.

(2) 寻路导航

用户在首个场景中选中兴趣点标签并发出导航请求后,MARB 进入寻路导航场景,寻路导航场景通过在平面地图上绘制当前位置到达指定地点的最短路径的方法为用户提供查看导航路径的服务.为方便用户了解搭乘不同交通工具抵达目标位置的相关信息,该场景不但提供多种导航方式(如步行、公交、自驾等)的选择服务,还为每一种抵达方式绘制路径,用户能够查看路径上经过的节点信息(如公交站、地铁站、十字路口等).

在寻路导航场景中,客户端的模块构成包括情景感知适配器、地图调用、导航方式选择、路径生成和渲染,云端的模块构成包括位置服务数据库和情景感知推理(如图 9 所示).由于寻路导航的目的是为了指引用户到达兴趣点所在地,因此与大部分常见的专用导航程序提供的服务比较相似,MARB 对用户在该场景中服务需求的判定较为清晰,主要包括抵达方式选择、导航路线标注与路线节点提示.其中,抵达方式的选择便于系统计算出最优的行进路线,而路径节点的查看以补充信息的方式为用户对抵达细节的需求提供便利.兴趣点标签被选中后,用户执行导航命令切换场景,进入寻路导航场景后,情景感知适配器将起点和终点的经纬度数据以及用户选定的导航方式发送给云端服务器.服务器调用位置服务数据库进行情景感知推理,推理结果包括地图信息(包含起止点区域的地图)和最短路径上每个节点的经纬度信息.接收到推理结果后,情景感知适配器先调用并绘制出地图,而后按节点的先后连接次序生成导航路径并在地图上绘制.用户查看路径上每一个节点时,情景感知适配器从云端的位置服务数据库获取被选点的描述信息,然后直接调用渲染模块将文字在地图上绘制显示.

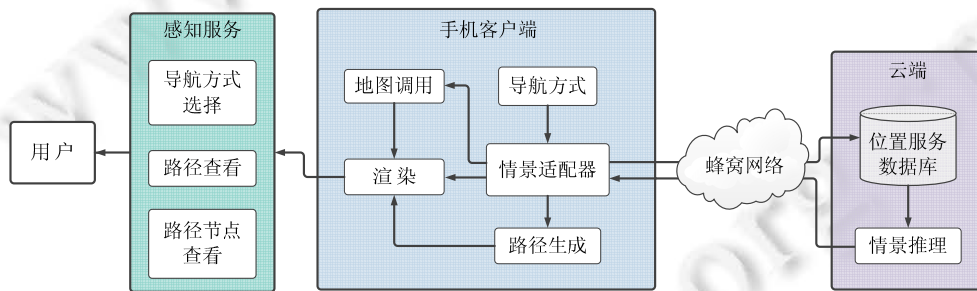


Fig.9 Module structure and context-aware service in the scenario of pathfinding navigation

图 9 寻路导航场景的模块构成及情景感知服务

(3) 目标识别与浏览

根据普遍的认知习惯,用户到达兴趣点楼盘的场景之后通常先识别楼房,然后查看楼房详细信息,并通过模型互动加深了解.在本文提出的 MARB 系统中,通过二维图片或文字标签的方式展示房屋的详细信息(楼房结构、户型面积、销售价格、物业信息等).为了带来更加直观丰富的浏览体验,在交互过程中,用户可以触控旋转或缩放楼房的三维模型,与建筑物模型的互动是对增强现实浏览器所展示详细信息的进一步补充.因此在目标识别与浏览场景中,MARB 需要提供目标识别、详细信息展示和模型交互这 3 项情景感知服务.

如图 10 所示:在目标识别与浏览场景中,手机客户端包括图像捕获、侦听器、情景感知推理、情景感知适配器和渲染模块,云端包括识别服务数据库与资源服务数据库.用户抵达兴趣点所在位置后,为增进对感兴趣建筑的了解,通常希望先认识视觉接触到的所有目标,而后详细查看目标各类信息,最后做出选择判断.为了满

是用户在做出最终选择前获得判断所需的先决条件, MARB 应当给予准确的识别服务和丰富的细节展示, 因此, 模块设计必须遵循目标识别与浏览场景中用户的认知方式, 依照交互过程中用户需求的先后顺序推送对应的服务, 即, 根据先识别目标后展示信息的次序分别推送情景感知服务. 当用户位于兴趣点楼盘的建筑物区域时, MARB 开启手机摄像头对目标建筑物进行精确识别. 首先, 通过图像捕获模块采集一帧关键帧图片, 在对图片进行降采样处理后, 转换为字节数组发送给识别服务器; 识别服务器调用识别服务数据库中的建筑物训练样本同关键帧图片进行对比, 找到匹配对象后, 把识别结果以 ID 编号的方式传输给客户端情景感知推理单元; 在完成情景感知推理后, 情景感知适配器将建筑物识别结果通过渲染模块显示, 从而最终完成目标识别服务. 如果情景感知推理结果判断没有查询到对应建筑物信息, 则情景感知适配器请求侦听器触发命令控制图像捕获模块, 继续采集关键帧进入重新识别目标的流程.

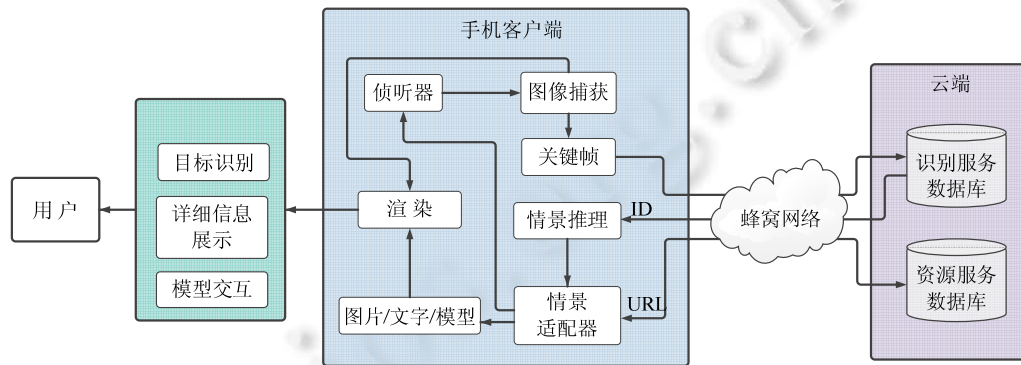


Fig.10 Module structure and context-aware service in the scenario of POI's recognition and browsing

图 10 兴趣点识别与浏览场景的模块构成及情景感知服务

3 效用评价实验

为验证本文所提出的移动增强现实浏览器原型系统的性能并评测其可用性, 首先, 根据本节提出的分场景方法构建原型系统框架; 然后, 运用可视化技术将系统模块的功能以交互界面的形式展现给用户; 最后, 将本文构建的原型系统与其他增强现实浏览器系统作为实验对象同时参与效用评估实验.

按照上文所述情景感知服务模块的组成结构, 在 Eclipse 和 Unity 集成开发环境中编制了移动增强现实浏览器的原型系统 MARB; 然后, 使用信息可视化的方法设计出 MARB 系统模块交互界面的数据视图; 最后, 采用交互设计的方法组织离散的数据视图生成完整的 MARB 用户交互界面(如图 11~图 13 所示).



Fig.11 Interface of context-aware service in the scenario of POI's positioning and selection

图 11 兴趣点定位与选择场景的情景感知服务界面



Fig.12 Interface of context-aware service in the scenario of pathfinding navigation

图 12 寻路导航场景的情景感知服务界面



Fig.13 Interface of context-aware service in the scenario of POI's recognition and browsing

图 13 兴趣点识别与浏览场景的情景感知服务界面

3.1.1 实验对象与方法

(一) 实验样本与评测对象

选定 28 人作为测试人员参加移动增强现实浏览器效用评测实验,其中,男女性别受试者比例各半.测试者职业为学生与上班族,年龄范围分布在 20 岁到 32 岁之间.考虑到先验知识在实验中对用户体验的影响因素,实验开始前,受试者填写调查问卷的个人信息调查部分.问卷统计结果表明:样本中 20 人曾有增强现实类浏览器使用经历,剩余 6 人没有任何类似软件的使用经历.

本次实验的评测对象选择目前已正式投入商用并且在测试区域能够提供兴趣点服务的增强现实类浏览器应用,包括 Junaio 浏览器、城市一点即通和本文提出的 MARB 系统.

(二) 效用评估指标

当前,对情景感知推荐系统的性能评估主要以推荐精确度为客观评价指标,包括 MAE, Precision, Recall, F1, P@N, MAP, AUC, ROC, RSME, nDCG 和 MRR 等^[30],其中,MAE 和 RSME 属于预测精确度指标,其他属于分类精确度指标.此外,由于需要获得用户使用浏览器的体验感受,为得到用户满意度的评价指标,还需要结合调查问卷获取用户对推荐效果的主观评价数据,衡量用户对推荐满意度、有效性等用户感受性的指标.

分类准确度指标可以用来评价推荐系统能够正确预测用户喜欢或者不喜欢某个商品的能力,衡量分类准确度主要指标的方法有 Recall, Precision 和 F1 这 3 种.其中,Recall 反映系统检索到的相关文件在系统所有相关的文件总数中所占的比率,衡量的是系统的查全率,又称为召回率或灵敏度;Precision 是系统检索到的相关文件与系统所有检索到的文件总数的比率,表示每个单独推荐项被检索的精度,衡量的是系统的查准率,又称精度、准确率;F1 又被称为平衡 F 分数(balanced F score),它是对召回率和准确率的一种加权平均,其数值范围在 0~1 之间,更高的数值表示更好的推荐效果^[31].

测试集中的每一个会话都有一组推荐结果,用 Rec 表示,则推荐结果相应地可以表示为

$$Rec := \{I(i), R(i)\}_{i=1}^N \quad (10)$$

其中, N 是浏览器推荐的数量,未被推荐的实例被称为隐藏项,用 $Hidden$ 表示; $I(i)$ 表示 i 个用户的索引值, $R(i)$ 为其对应的推荐结果.基于测试集中所有会话的推荐集和隐藏项, $Recall$, $Precision$ 和 $F1$ 表示为

$$Recall = \frac{|Hidden \cap Rec|}{|Hidden|} \quad (11)$$

$$Precision = \frac{|Hidden \cap Rec|}{|Rec|} \quad (12)$$

$$F1 = \frac{2 \times Recall \times Precision}{Recall + Precision} \quad (13)$$

图 14 是对以上 3 种衡量分类准确度评价方法的说明,图中的 A, B, C 和 D 需要测试者人工标注.

由于阈值设定对推荐结果存在影响,为评测推荐系统的性能,需要设定阈值对推荐结果的序列进行分类,以查看推荐系统检索到相关已推荐项的概率,其中,大于阈值的实例分为正类,小于阈值的结果分为负类.如果实例为正类且系统推荐成正类,则称它为真正类(true positive,简称 TP);如果实例为负类且系统推荐为正类,则称为假正类(false positive,简称 FP).对应于图 14,可以构造一个 2×2 的混淆矩阵(confusion matrix)来表示测试集中各个实例的分布状况.因此在图 15 中,根据增强现实浏览器推荐的和未被推荐的项目可以归纳出 4 种判定,分别是真正类、假负类(false negative,简称 FN)、假正类和真负类(true negative,简称 TN),图 15 中,1 代表正类,0 代表负类.

	相关的	无关的	
已推荐项	A	B	$P(\text{查准}) = A/A+B$
未推荐项	C	D	
	$R(\text{查全}) = A/A+C$		$F1 = 2RP/R+F$

Fig. 14 Recall, Precision and F1

图 14 Recall, Precision 与 F1

	相关的	无关的		
	1	0	总计	
已推荐项	1	TP	FP	推荐的正类 TP+FP
未推荐项	0	FN	TN	推荐的负类 FN+TN
总计	相关的正类 TP+FN	相关的负类 FP+TN	TP+FP+TN+F N	

Fig. 15 Four cases of dichotomous questions

图 15 二分问题的 4 类情况

根据图 15 的 4 类情况,可以使用以下 3 个通用的性能评估标准来衡量浏览器的性能.

$$Recall = \frac{TP}{TP + FN} \quad (14)$$

$$Precision = \frac{TP}{TP + FP} \quad (15)$$

$$FPR = \frac{TN}{FP + TN} \quad (16)$$

通过公式(2)~公式(21)计算得到的 $Recall$ 又被称为真正类率(true positive rate,简称 TPR),公式中, TP 表示推荐系统已推荐的与当前情境相关的项, FN 表示未被推荐的相关项.因此, $Recall$ 计算的是浏览器所推荐出的正实例占有所有正实例的比例. $Precision$ 计算的是已推荐的相关项在所有被推荐项总数中的比率,能够反映系统推荐的准确率.真负类率(true negative rate,简称 TNR)也被称为特异度(specificity),是把实际为假值的推荐为假值的概率.假正类率(false positive rate,简称 FPR)是错误推荐的负例的数目与所有负例数目的比值,也称误报率. FPR 值可以用 $1-TNR$ 表示.

(三) 实验方法

在实验中,安排测试者使用 Junaio 浏览器、城市一点即通和本文提出的 MARB 系统这 3 款增强现实浏览

器在指定位置搜索兴趣点.选定目标后,查询兴趣点的相关信息.28位受测者分4轮,每轮在不同的指定地点执行上述任务操作.每位测试者选中兴趣点后只查询该目标的一种信息(售价、地址、联系方式等),在4轮测试过程中,为每位测试者配置一名观察员,观察员在整个测试过程中记录用户的操作流程和用户实时反馈的信息.每完成一轮测试后,观察者让用户及时填写调查问卷的第2部分,即情景感知服务使用调查部分,记录结果用于主观分类精度指标的结论分析.实验完成后,统计在4轮测试中系统对每位测试者在每步操作流程中反馈的推荐项,通过整理调查问卷收集的数据,比对浏览器设计预期推荐的项,可以分析得出客观分类精度指标的实验结论.

3.1.2 结果分析

(一) 客观分类精度指标分析

根据计算出的误报率和准确率,可以绘制出受试者工作特征曲线(receiver operating characteristic curve,简称 ROC).ROC 曲线是根据不同的二分类方式(分界值或决定阈),以 TPR(准确率)为纵坐标、FPR(误报率)为横坐标绘制的曲线.由于坐标系中 ROC 曲线的横坐标为虚报概率、纵坐标是击中概率,受试者在某种应激条件下,由于采取不同的决断准则将得出不同结果,根据收集到的结果,评测人员能够绘制出 ROC 曲线^[32].给定一个二分类模型和它的阈值,就能从所有实例(正例/负例)的相关项和推荐项中计算出一个坐标点($X=FPR, Y=TPR$).

观察由连续的坐标点连接而成的 ROC 曲线,可以用定量的数据对受试对象做出客观的分析,从而得出定性的结果.

表4为 MARB 测试对象的实验数据,首先把前10%的数值作为阈值,即,把前10%的实例设定为正类;接着,根据浏览器向用户推荐的兴趣点数量(TP)计算出无关的推荐项(FP),使用 Recall 计算该实验对象分类准确度指标的灵敏度比率;然后,采用相同的方法计算相关的未推荐项(FN)和无关的未推荐项(TN),通过调用 FPR 计算实验对象的误报率;最后,按照10%的间隔增加阈值,随着判断点阈值不断的改变,表中显示出了该被测浏览器的灵敏度指标和误报率指标的变动状况.将反映增强现实浏览器情景感知分类精度指标的连续点对相连,可以绘制出如图16所示的 MARB 的 ROC 曲线.

Table 4 Experimental data tables of MARB

表4 MARB 的实验数据统计表

判断点阈值	实验项	相关的已推荐项的数目	无关的已推荐项的数目	误报率(%)	准确率(%)
10	630	479	151	3.04	36.23
20	630	284	346	9.99	57.72
30	630	223	407	18.18	75.58
40	630	90	540	29.04	81.37
50	630	62	568	40.46	86.08
60	630	58	572	51.96	90.47
70	630	56	574	63.50	94.70
80	630	25	605	75.67	96.60
90	630	23	607	87.87	98.34
100	625	22	603	100	100

按照同样的方法可得出另外两个测试对象的实验数据,它们的统计列表见表5和表6.

Table 5 Experimental data tables of junaio

表5 Junaio 的实验数据统计表

判断点阈值	实验项	相关的已推荐项的数目	无关的已推荐项的数目	误报率(%)	准确率(%)
10	630	271	359	7.22	20.50
20	630	304	326	13.77	43.49
30	630	170	460	23.02	56.35
40	630	90	540	33.88	63.16
50	630	107	523	44.40	71.26
60	630	103	527	55.00	79.05
70	630	58	572	66.50	83.43
80	630	96	534	77.24	90.70
90	630	50	580	88.90	94.48
100	625	73	552	100	100

Table 6 Experimental data tables of cityoneclick

表 6 城市一点即通的实验数据统计表

判断点阈值	实验项	相关的已推荐项的数目	无关的已推荐项的数目	误报率(%)	准确率(%)
10	630	331	229	6.01	25.04
20	630	176	454	15.14	38.35
30	630	128	502	25.24	48.03
40	630	182	445	34.18	62.03
50	630	153	477	43.78	73.60
60	630	134	496	53.75	83.74
70	630	67	563	65.07	88.80
80	630	56	574	76.61	93.04
90	630	55	575	88.18	97.20
100	625	37	588	100	100

图 16 显示出了由 3 个测试对象的统计数值绘制的对应 ROC 曲线,图中由原点(0,0)到右上角(1,1)的对角线被称为机会线(chance line/guessing line),它把二维的 ROC 空间对半划分为左上与右下两个区域.机会线以上的点表示一个较好的分类结果,而在机会线以下的点反映了较差的分类结果^[33].由图 15 可以看出:3 款增强现实浏览器的 ROC 曲线均位于机会线上方,表明它们都有较好的分类效果.客观评价测试对象分类性能优劣的一个重要方法是曲线下的面积(area under the curve,简称 AUC),表示分类器正确判断正类样本的值高于负类样本的几率.AUC 的取值范围在 0~1 之间,AUC 取值越大的分类器正确率越高^[33].

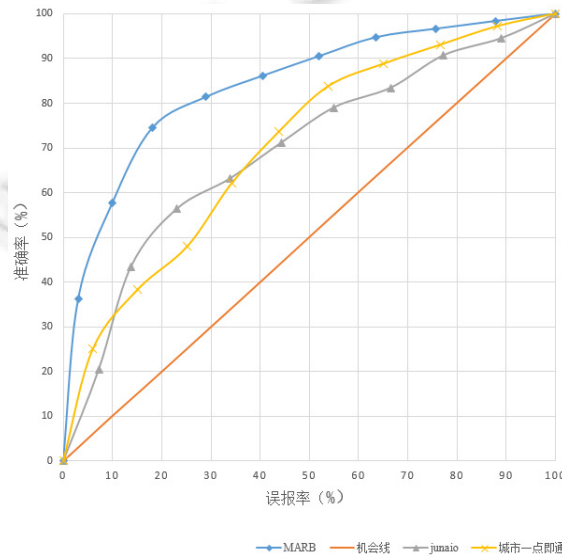


Fig.16 Distribution of ROC curve in the evaluation experiments

图 16 评测实验的 ROC 曲线分布

使用 AUC 判断分类器预测性能的标准如下:

- 当 $AUC=1$ 时,表明分类器能够在任何场景中做出完美的预测.由于实验过程中产生的误差是难以避免的,因此现实中不存在这种情况;
- 当 $AUC \leq 0.5$ 时,分类器预测结果等同于或劣于随机猜测,说明分类器没有发挥应有的作用;
- 只有当 $0.5 < AUC < 1$ 时,分类器预测结果优于随机猜测,数值越接近 1,预测效果越好.

通过计算每条曲线的 AUC 数值,能够比较出各浏览器分类性能的好坏.将浏览器所推荐出的兴趣点的数量表示为分值,所有的兴趣点表示为实例,用 M 代表正类实例的个数、 N 代表负类实例的个数,通过下式可以计算出在所有的 $M \times N$ 个正负样本对中有多少组的正实例的分值大于负实例的分值:

$$AUC = \frac{\sum_{ins_i \in \text{positiveclass}} rank_{ins_i} - \frac{M \times (M + 1)}{2}}{M \times N} \tag{17}$$

上式首先按照从大到小的顺序将分值排序,将分值最大的实例的名次设为 n ,分值次大的实例的名次设为 $n-1$,以此类推,直至名次为 1;然后,将正类的实例相加,再减去正类实例的分值为最小的那部分,得到的就是所有实例中有多少对正类实例的分值大于负类实例的分值;最后,除以 $M \times N$ 就能计算出 AUC 的值。

计算结果显示:3 条 ROC 曲线之中,MARB 的 AUC 面积最大.与同类浏览器相比,MARB 在分类准确度上提高了约 13%,表明 MARB 向用户推荐的信息分类准确度效果最佳.这得益于在 MARB 模块的构成中采用了场景划分的方法对不同场景中显示标签的数目和分类预先进行规划,系统推送标签的先后次序比较符合用户的使用逻辑,因此,用户在使用 MARB 检索兴趣点时获取的标签更为合理.搜寻兴趣点操作方式的不同简化了后续查看兴趣点详细信息的操作步骤,从而改善了用户体验。

(二) 主观分类精度指标分析

每一次实验完成后,受试者需要额外完成一份简短的调查报告.报告内只有两项问题需要测试者填写:一个问题是受试者对浏览器推荐效果的整体印象,另一个问题是在使用过程中对浏览器时间延误的感受.为保证调查问卷没有偏向性,对两项主观感受程度的评分使用从“非常有用”到“完全没有用”的李克特五点量表(Likert scale)的形式。

实验结果见表 7,测试者的投票分布状况展示了用户使用 3 个受测浏览器的主观感受,其中,“预测效果”与“时间延迟”是最能直观地评测情景感知推荐可用性的两个参数.从均值与标准差的数值可见,MARB 是所有测试对象中可用性最好的浏览器.Junaio 与城市一点即通在预测效果上基本持平,虽然约占 50%的测试者对这两个浏览器的总体推荐效果有正面的评价,然而两者同 MARB 相比均存在一定的差距,MARB 对预测效果的平均满意度比其他两个测试对象高出约 32%.采用场景划分构建模块的方法使得 MARB 能够更好地向用户提供推荐标签,在提高推荐效率的同时,也降低了系统的反应时间.MARB 和城市一点即通在时间延迟项的用户感受相差不多,然而 Junaio 在这一项的可用性上却表现为中等偏下水平.均值为负数,说明过半的用户在使用 Junaio 时明显感受到推荐滞后的情况.其原因在于,Junaio 在兴趣点服务供应方面的不足.由上述对预测效果和操作延迟两个指标投票数据的分析可知:MARB 分类准确度的提升,使得增强现实浏览器对标签的推荐与显示符合用户交互操作时的思维逻辑,分场景推送感知服务对标签的使用更为合理.该方法不但优化了操作的步骤,而且能有效促进了用户对兴趣点的认知。

Table 7 System usability evaluation table
表 7 系统的可用性评测表

		完全没用				非常有用		
		-2	-1	0	1	2	Average	St.Dev.
MARB	预测效果	4.5%	5.1%	36%	24.4%	29.9%	0.70	1.09
	时间延迟	13.1%	17.9%	20.7%	31.2%	17.1%	0.21	1.29
Junaio	预测效果	9%	14%	20.9%	31.7%	24.4%	0.49	1.25
	时间延迟	23.3%	21.8%	25%	20.6%	9.3%	-0.29	1.28
城市一点即通	预测效果	6.6%	14.4%	33.1%	25.3%	20.6%	0.39	1.15
	时间延迟	9.6%	21.2%	30.9%	24.8%	13.5%	0.11	1.17

4 总结与展望

目前已有的移动增强现实类浏览器在标签的显示与布局上存在诸多问题,给用户在操作上带来了认知的混淆与交互的不便.本文提出的 MARB 以用户对兴趣点检索并浏览兴趣点的操作逻辑为次序,将众多的操作行为划分为兴趣点定位与选择、寻路导航与目标识别这 3 个场景,而后,分别在 3 个场景中有针对性地采用情景感知服务框架逐一构建提供情景感知服务的模块架构,提升每个模块在场景中推送增强现实服务的效率.为提高增强现实服务器训练图片的效率、最大化服务器组共有的内存计算资源,本文提出的 MARB 采用基于分布

式聚类的词汇树识别算法进行兴趣点识别,运用基于 BRISK 特征的跟踪注册算法实现了虚拟标签对兴趣点的实时跟踪。

在情景感知系统的效用评测中,针对 MARB,Juniao 和城市一点即通这 3 款移动增强现实浏览器开展了评测实验,以验证推荐系统的分类精度效用。按照 10%间隔设置判断点的阈值,将 FPR 和 TPR 的点对映射在二维的 ROC 空间内绘制曲线。通过对比 3 条曲线的 AUC 数值,得出受测浏览器分类精度的客观分析结果,即,AUC 面积值最大的 MARB 优于其他测试对象。受益于分场景的情景感知服务四层框架对结构的合理规划,MARB 推荐信息的分类精确度得到提高,简化了用户的操作步骤,改善了用户体验。在对情景感知推荐的可用性评测实验中,通过收集问卷调查的数据,检验了浏览器分类精度指标的用户主观感受。李克特五点量表的统计结论表明,MARB 在预测效果和时间延迟上均优于 Juniao 和城市一点即通。说明面向不同场景有针对性地构筑情景感知推荐服务模块,能够优化兴趣点标签的组织方式和推送效率,可以为用户带来流畅的使用体验。

虽然通过使用本文提出的新方法可以解决移动增强现实浏览器系统在分类准确度和时间延迟方面给用户体验带来的问题,但是设计者在界面设计阶段的人为错误产生的设计偏差同样会导致一系列人因问题,从而阻碍用户对兴趣点的认知,影响交互体验,这是本研究课题下一步将要研究的方向。此外,由于本文所提出的方法当前仅针对科研项目要求的指标进行了小范围的实验性测试,未来将逐步地向公众开放测试,收集用户在实际使用过程中的各类数据,以验证该研究的有效性。

致谢 在此,我们向对本文的研究工作给予支持和建议的同行,尤其是北京理工大学光电学院光电成像技术与系统教育部重点实验室科研项目组的老师和同学表示感谢。

References:

- [1] Weiser M. The computer for the 21st century. *Scientific American*, 1991,265(3):94-104. [doi: 10.1038/scientificamerican0991-94]
- [2] Adomavicius G, Tuzhilin A. Towards the next generation of recommender systems: A survey of the state-of-the-art and possible extensions. *Trans. on Knowledge and Data Engineering*, 2005,17(6):734-749. [doi: 10.1109/TKDE.2005.99]
- [3] Schmidt A, Beigl M, Gellersen HW. There is more to context than location. *Computers & Graphics*, 1999,23(6):893-901. [doi: 10.1016/S0097-8493(99)00120-X]
- [4] Dey AK. Providing architectural support for building context-aware applications [Ph.D. Thesis]. Atlanta: Georgia Institute of Technology, 2000.
- [5] Byun HE, Cheverst K. Utilizing context history to provide dynamic adaptations. *Applied Artificial Intelligence*, 2004,18(6): 533-548. [doi: 10.1080/08839510490462894]
- [6] Abowd GD, Atkeson CG, Hong J, Long S, Kooper R, Pinkerton M. Cyberguide: A mobile context-aware tour guide. *Wireless Networks*, 1997,3(5):421-433. [doi: 10.1023/A:1019194325861]
- [7] Raento M, Oulasvirta A, Petit R, Toivonen H. ContextPhone: A prototyping platform for context-aware mobile applications. *Pervasive Computing*, 2005,4(2):51-59. [doi: 10.1109/MPRV.2005.29]
- [8] Coppola P, Mea VD, Gaspero LD, Menegon D, Mischis D, Mizzaro S, Nazzi E, Scagnetto I. The context-aware browser. *Intelligent Systems*, 2010,25(1):38-47. [doi: 10.1109/MIS.2010.26]
- [9] Woensel VW, Casteleyn S, Troyer OD. A framework for decentralized, context-aware mobile applications using semantic Web technology. *Lecture Notes in Computer Science*, 2009,5872:88-97. [doi: 10.1007/978-3-642-05290-3_18]
- [10] Lee JY, Seo DW, Rhee G. Visualization and interaction of pervasive services using context-aware augmented reality. *Expert Systems with Applications*, 2008,35(4):1873-1882. [doi: 10.1016/j.eswa.2007.08.092]
- [11] Grubert J, Langlotz T, Grasset R. Augmented reality browser survey. Graz: Institute for Computer Graphics and Vision, Graz University of Technology, 2011. 1-11.
- [12] Olsson T, Salo M. Online user survey on current mobile augmented reality applications. In: *Proc. of the 10th IEEE Int'l Symp. on Mixed and Augmented Reality (ISMAR 2011)*. Basel: IEEE, 2011. 75-84. [doi: 10.1109/ISMAR.2011.6092372]
- [13] Dünser A, Grasset R, Billinghurst M. A survey of evaluation techniques used in augmented reality studies. Technical Report, TR-2008-02, Christchurch: Human Interface Technology Laboratory New Zealand, 2008. 767-777. [doi: 10.1145/1508044.1508049]

- [14] Wang X, Dong JS, Chin C, Hettiarachchi SR. Semantic space: An infrastructure for smart spaces. *Pervasive Computing*, 2004,3(3): 32–39. [doi: 10.1109/MPRV.2004.1321026]
- [15] Board J. Future global satellite systems for Intelsat. *Journal of the British Interplanetary Society*, 1985,38(9):407–411.
- [16] He C, Hu X, Liu Y, Wang Y, Cheng H. A novel drift compensating method for orientation measurement system in VR applications. In: *Proc. of the Int'l Instrumentation and Measurement Technology Conf. (I2MTC 2012)*. IEEE, 2012. 2482–2487. [doi: 10.1109/I2MTC.2012.6229539]
- [17] Szeliski R. *Computer Vision: Algorithms and Applications*. New York: Springer-Verlag, 2010. 210–212.
- [18] Welch G, Bishop G. An introduction to the Kalman filter. In: *Proc. of the Siggraph Course*. Los Angeles: ACM Press, 2001. 7–14.
- [19] Kim H, Cho YJ, Oh SR. CAMUS: A middleware supporting context-aware services for network-based robots. In: *Proc. of the IEEE Workshop on Advanced Robotics and its Social Impacts*. Nagoya: IEEE, 2005. 237–242. [doi: 10.1109/ARSO.2005.1511658]
- [20] Lee JY, Rhee GW, Kim H, Lee KW, Suh YH, Kim K. Convergence of Context-Awareness and Augmented Reality for Ubiquitous Services and Immersive Interactions. Berlin: Springer-Verlag, 2006. 466–474. [doi: 10.1007/11751632_51]
- [21] Mo T, Li WP, Wu ZH, Chu WJ. Framework of context-aware based service system. *Chinese Journal of Computers*, 2010,33(11): 2084–2092 (in Chinese with English abstract). [doi: 10.3724/SP.J.1016.2010.02084]
- [22] Wang XH, Zhang DQ, Gu T, Pung HK. Ontology based context modeling and reasoning using OWL. In: *Proc. of the 2nd IEEE Annual Conf. on Pervasive Computing and Communications Workshops*. Orlando: IEEE, 2004. 18–22. [doi: 10.1109/PERCOMW.2004.1276898]
- [23] Ding B, Wang HM, Shi DX. Constructing software with self-adaptability. *Ruan Jian Xue Bao/Journal of Software*, 2013,24(9): 1981–2000 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4432.htm> [doi: 10.3724/SP.J.1001.2013.04432]
- [24] Engelke T, Becker M, Wuest H, Keil J, Kuijper A. MobileAR Browser—A generic architecture for rapid AR-multi-level development. *Expert Systems with Applications*, 2013,40(7):2704–2714. [doi: 10.1016/j.eswa.2012.11.003]
- [25] Macintyre B, Hill A, Rouzati H, Gandy M, Davidson B. The argon ar web browser and standards-based ar application environment. In: *Proc. of the 10th IEEE Int'l Symp. on Mixed and Augmented Reality (ISMAR 2011)*. Basel: IEEE, 2011. 65–74. [doi: 10.1109/ISMAR.2011.6092371]
- [26] Li N, Duh HBL. Cognitive Issues in Mobile Augmented Reality: An Embodied Perspective. New York: Springer-Verlag, 2013. 109–135. [doi: 10.1007/978-1-4614-4205-9_5]
- [27] Mogilev D, Kiyokawa K, Billingham M, Pair J. AR Pad: An interface for face-to-face AR collaboration. In: *Proc. of the Conf. on Human Factors in Computing Systems*. ACM Press, 2002. 654–655. [doi: 10.1145/506443.506530]
- [28] Billingham M, Bowskill J, Dyer N, Morphett J. An evaluation of wearable information spaces. In: *Proc. of the IEEE 1998 Virtual Reality Annual Int'l Symp.* Atlanta: IEEE, 1998. 20–27. [doi: 10.1109/VRAIS.1998.658418]
- [29] Robert S. *Information Visualization-Design for Interaction*. Pearson Higher Isia Education, 2007. 166–174.
- [30] Wang LC, Meng XW, Zhang YJ. Context-Aware recommender systems. *Ruan Jian Xue Bao/Journal of Software*, 2012,23(1):1–20 (in Chinese with English abstract). <http://www.jos.org.cn/1000-9825/4100.htm> [doi: 10.3724/SP.J.1001.2012.04100]
- [31] Domingues MA, Jorge AM, Soares C. Using contextual information as virtual items on top-*n* recommender systems. In: *Proc. of the Workshop on Context-Aware Recommender Systems (CARS 2009) in Conjunction with the 3rd ACM Conf. on Recommender Systems (RecSys 2009)*. New York: ACM Press, 2009. 124–129.
- [32] Hanley JA, Mcneil BJ. The meaning and use of the area under a receiver operating characteristic (ROC) curve. *Radiology*, 1982, 143(1):29–36. [doi: 10.1148/radiology.143.1.7063747]
- [33] Zou KH, O'malley AL, Mauri L. Receiver-Operating characteristic analysis for evaluating diagnostic tests and predictive models. *Circulation*, 2007,115(5):654–657. [doi: 10.1161/CIRCULATIONAHA.105.594929]

附中文参考文献:

- [21] 莫同,李伟平,吴中海,褚伟杰.一种情境感知服务系统框架. *计算机学报*, 2010,33(11):2084–2092. [doi: 10.3724/SP.J.1016.2010.02084]
- [23] 丁博,王怀民,史殿习.构造具备自适应能力的软件. *软件学报*, 2013,24(9):1981–2000. <http://www.jos.org.cn/1000-9825/4432.htm> [doi: 10.3724/SP.J.1001.2013.04432]

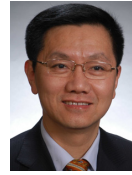
- [30] 王立才,孟祥武,张玉洁.上下文感知推荐系统.软件学报,2012,23(1):1-20. <http://www.jos.org.cn/1000-9825/4100.htm> [doi: 10.3724/SP.J.1001.2012.04100]



林一(1982-),男,福建福州人,博士,讲师,主要研究领域为情景感知计算,虚拟现实与增强现实,人机交互.



刘越(1968-),男,博士,教授,博士生导师,CCF 专业会员,主要研究领域为人机交互,虚拟现实与增强现实,摄像机位置和姿态的准确跟踪,三维立体显示,摄像机标定.



王涌天(1957-),男,博士,教授,博士生导师,CCF 高级会员,主要研究领域为虚拟现实,增强现实.



贺长宇(1987-),男,博士生,主要研究领域为姿态跟踪,数据融合,虚拟现实与增强现实.

www.jos.org.cn

www.jos.org.cn