

面向未来的智能家居前沿进展

付心仪¹, 张鹤¹, 薛程^{1,2}, 孙童心^{1,2}

1. 清华大学未来实验室, 北京 100084

2. 清华大学美术学院, 北京 100084

摘要 随着人工智能、物联网、大数据等科技的迅猛发展, 智能家居正式进入了3.0时代, 深度依赖用户大数据进行深度学习和数据挖掘, 迫切需要面向智能家居的智能交互、智能感知和复杂决策的新技术新方法作为支撑。综合国内外前沿相关研究, 提炼出未来智能家居所需要的3大要素: “感知”“思考”和“执行”。并以此为脉络, 综述了国内外智能家居前沿平台的研发进展, 评议了智能家居环境中的数据收集和特征提取研究、智能家居场景下的智能感知研究、面向智能家居的多模态人机交互与创新场景研究的相关成果, 介绍了清华大学未来实验室在智能家居领域的相关研究工作。

关键词 智能家居; 情感计算; 智能感知; 多模态交互

随着人工智能、物联网、大数据、5G等科技的迅猛发展, 信息社会正在逐渐向智能社会转型过渡, 新冠肺炎疫情的暴发更是推动了世界“宅经济”的发展, 居家办公、居家隔离等措施, 极大地增加了人们在家庭环境下的活动时间, 这也使得智能家居的发展步入了快车道。根据 Statista 统计数据和预测数据显示, 2020 年全球智能家居市场价值达到约 789 亿美元, 2021 年全球智能家居市场价值预计为 1020 亿美元, 而到 2025 年全球智能家居市场价值预计为 1823 亿美元, 其中涉及智能家居控制和连接、设备、安全问题、家庭娱乐、舒适和照明以及能源管理^[1]。

“十四五”规划建议, 坚持把发展经济着力点放在实体经济上, 坚定不移建设制造强国、质量强国、网络强国、数字中国。在时代、产业和政策三重刺激下, 智能强国发展迫在眉睫, 智能家居作为民生息息相关的重要领域, 具有广泛发展空间与创新前景。

智能家居是与人类活动高度相关的科研领域, 是学科高度交叉的综合研究方向, 涉及人机交互、人工智能、计算机视觉、物联网、用户体验、心理学等学科和领域的综合交叉融合, 其相关实验平台建设和理论与技术创新, 对于深度理解人类行为的模式和情感交流方式具有重要作用, 符合下一代人工智能“以人为本”的核心诉求。

收稿日期: 2021-11-30; 修回日期: 2022-07-05

基金项目: 国家自然科学基金项目(6217070900)

作者简介: 付心仪, 助理研究员, 研究方向为自然人机交互、情感计算、文化遗产数字化等, 电子信箱: fuxy@tsinghua.edu.cn

引用格式: 付心仪, 张鹤, 薛程, 等. 面向未来的智能家居前沿进展[J]. 科技导报, 2023, 41(8): 36-52; doi: 10.3981/j.issn.1000-7857.2023.08.004

起始于20世纪90年代互联网时代的智能家居1.0,依赖于有线计算机和互联网,实现对家居安防、家电、灯光的自动化控制,并不具备智能交互、感知和执行能力。随着2014年物联网时代的到来,智能家居步入2.0时代,Wifi、4G、蓝牙、云计算等技术逐步成熟,催生了智能家居的感知能力,智能设备在自动化的基础上,可以感知温度、烟雾、光线等环境变化,并可以完成部分主动决策,但此阶段智能家居依然处于App+设备的点对点连接模式。2020年,随着5G技术的正式发牌,智慧物联网时代到来,人工智能+物联网已经成为了各大产业巨头的主要赛道,智能家居也正式进入了3.0时代,智能家居即将构建云端互联网平台,实现真正的智能交互、智能感知和复杂决策。

在智能家居3.0时代,平台和数据是基础,算法是核心,交互是关键。相关智能产品将广泛依赖于用户大数据进行深度学习和数据挖掘,精准刻画用户画像,进行智能化、个性化的服务提供。基于情感计算、语音识别、计算机视觉、知识图谱、自然语

言处理等方法的多模态人机交互,实现“人机物”的深度融合,是打造智能家居3.0的关键路径。

本文综合国内外前沿相关研究,总结和提炼出未来智能家居所需要的3大要素,即:“感知”“思考”和“执行”,三者关系如图1所示。“感知”即对智能家居环境中的各类人、机、物和环境等数据进行采集、处理和封装,作为用户画像和机器学习的基础;“思考”即通过大数据和人工智能技术,对家居环境内人的行为、情感、发生的事件等进行识别和判断,并产生如何与用户交互的决策;“执行”即智能家居体系内的智能家电等智能硬件,根据决策结果与用户产生自然人机交互,优化用户体验。以这3个要素为脉络展开综述,首先对国内外前沿智能家居平台现状进行总结,随后按照“感知-思考-执行”的逻辑,依次对智能家居环境中的数据采集研究、智能家居场景下的智能感知与复杂决策研究和面向智能家居的多模态人机交互与用户体验研究分别进行综述,最后对未来实验室在智能家居领域的相关工作进行阐述。



图1 感知-思考-执行逻辑思维导图

1 国内外前沿智能家居平台现状

智能家居是物联网系统的一个重要应用场景,其作用于人居环境下,利用多种信息传感器采集人居环境内的各种数据,并使用互联网协议相互自主通信,建立人、设备、空间和场景的广泛连接,进而

实现组成家庭智慧家居空间的目的。

智能家居平台(简称“平台”)依托系统集成、人工智能、射频识别、传感器网络、云计算、边缘计算和联邦(协作)学习等技术和算法,使用集成硬件在应用层面进行完备开发,形成可以全面支撑人居环境中对于场景联动、人因个性、隐私安全、健康安

全、决策指导、社交行为和协作互联需求的自动化、个性化生态平台。平台作为智能家居生长环境、应用环境、分析环境的基础性框架和根基,涉及诸多领域的研究,包括但不限于计算机科学、数据科学、电子工程、艺术设计、人机交互、社会科学等。

从学术界的视角来看,近年来随着物联网的兴起,硬件技术与算力的提升和算法模型的逐步优化,与智能家居物联网架构、可用性、安全性、隐私性、便捷性、美学和技术应用相关的话题被学术界广泛讨论。平台依托网关将设备和数据中心、计算中心连接,并在网络边缘设备上部署分布式学习算法降低延迟和提高预测准确率。Firouzi等^[2]提出利用模糊逻辑控制器和网络边缘控制器对智能家居物联网中的应用提供低级智能的模糊逻辑控制器结构,考虑能源消耗、价格费率、电池等级和需求能源堵的因素,在网络上分散模糊逻辑推理以提高物联网控制器在边缘设备上分析数据的不确定性和噪音的鲁棒性。Mpeis等^[3]介绍了一种物联网本地化开源众包架构 Amyplace 4.0 IoT 应用技术(如红外、蓝牙、视觉、声音分析、无线射频识别、惯用测量单元、超宽带、传感器网络、无线局域网等)组成混合系统并存储在高密度数据库中。平台后端包含服务器、数据存储、网络组成的轻量级“模型-视图-控制器”框架,支持各种数据的存储接口,并包含身份验证等安全措施。结合使用 HTML5、CSS3 和 AngularJS 等构建的 Web 应用程序,允许式设计交互模式和可视化数据平台。

从工业界的角度来看,自进入工业 4.0^[4]和物联网时代起,智能家居市场经历了“野蛮生长”,企业也从生产单一功能(终端)产品向构建完整产业生态体系逐渐过度,由单一智能化产品向整体场景联动转变。大量智能家居设备厂商已经意识到,仅满足业务需求的单一设备难以在物联网时代向用户提供全面、便捷和舒适的服务,通过将智能设备连入整体平台是使其服务更连续化、智能化的一个重要手段。根据统计数据显示,2019 年全球范围内有 620 多家涉足智能家居业务的物联网平台企业^[5],其中不仅包含从电子制造业发展而来的公

司,如美的、海尔等,也包含以欧瑞博为代表的物联网智能设备厂商,苹果、亚马逊等众多大型互联网公司也先后加入物联网智能家居的赛道。

完备的平台是企业智能家居市场中有效提高占有率、留存用户的重要方式之一,其中平台对家居环境下设备的接入情况尤为重要。为此,产业界利用各自的行业优势分成了不同的战略方向。首先是如美的、海尔等传统家电制造企业,依托于长期积累的设备制造工艺和较高的市场占有率,以设备终端为基点铺设平台。另一方面,大型互联网/科技公司依托自身的流量数据、前沿技术和研发能力重点发展平台端,并通过合作、开源等方式将多种设备接入平台协议,形成以智能音箱等为控制中心、本地和云数据计算中心的智能家居生态系统,如百度公司对话式人工智能系统 DuerOS 和阿里巴巴公司的阿里智能等。除此之外,一些同时具有一定制造经验和互联网科技背景的公司融合了两者的优势,正在使用自身产品(或其他接入产品)全面铺设智能家居生态,并使用原生移动设备、语音设备对平台进行不断优化,达到更好的可操作性、连接性和用户体验,如小米公司旗下的米家生态和苹果家庭套件。

综上所述,当前学术界和工业界平台的相关研究和建设仍存在良莠不齐的情况,学术界缺乏综合整体的平台研究成果,工业界仍在以增加接入平台的设备数量和种类为重点发展方向。未来的平台应逐步整合各方的区域性优势,进一步降低生产、制造、装配成本,提升平台稳定性、安全性、智能性和兼容性,从面向单一家庭到连接多家庭、社区和连接其他物联网产物的以家庭住宅为中心的智能家居物联网平台。

2 智能家居环境中的数据研究

数据是未来智能家居构建的基础,结合人工智能和大数据方法,可为用户提供更加个性化体验。在智能家居环境中,获取的数据主要包含人体姿态数据、面部数据、语音数据、定位数据、环境数据等。

2.1 人体姿态数据

人体姿态包含2D姿态和3D姿态^[6-7],通常通过RGB相机或深度相机采集视频数据,再经由算法生成得到人体姿态特征。RGB相机相对成本更低、技术方案成熟、可操作性更强,但对后续处理算法要求更高。当前,随着计算机视觉技术的蓬勃发展,多视角RGB相机已经逐渐成为主流的解决方案。根据不同的获取人体姿态的目的,RGB相机有不同的配置方式。例如,对于人体2D姿态数据的获得,只需配置单视角相机即可,对于人体动作和姿态的精确识别检测和3D坐标的获取,则需要以环绕方式配置多视角相机^[8-10]。

人体姿态数据在智能家居环境中应用范围和潜力很大。基于姿态数据的人体识别可以进行活动分析,对个体行为进行追踪并生成个体活动日志^[11]。通过对人体特定关节点进行追踪,得到速度、加速度等信息,可以得到人体行为模式,进行健康监测和指导^[12]。针对老年用户,基于计算机视觉的跌倒监测也是目前居家环境中主流的检测方式^[13]。也有一些团队进行基于姿态的情感识别,来帮助交互终端更好的理解用户的情感需求^[14]。目前来说,基于视觉的姿态识别技术无法对用户的隐私进行保护,一些研究正使用基于超宽带和Wifi等信号来进行人体姿态的识别,但识别准确率较低,并且对用户环境要求较高^[15]。

2.2 面部数据

面部数据主要应用在人脸识别^[16]和面部表情识别^[17-20]中,智能家居环境中的面部数据的采集,与人体姿态数据采集较为相似,主要使用RGB相机获取二维图像数据,也有少部分产品使用了深度相机,直接对面部关键点进行分析处理。

人脸识别是一种用于身份验证的计算机视觉技术,技术上分为基于图片的人脸识别和基于视频的人脸识别,两者的识别过程是类似的。首先,从图片或者视频中检测到人脸,随后进行人脸的特征提取,包括基于边缘、线条和曲线的通用性方法、基于特征模板的方法、结构匹配方法等,最后对特征进行检测进行验证^[16]。

面部识别在智能家居的应用相对比较局限,当

前较为成熟的产品主要是基于人脸识别的智能门锁^[21],但由于安全性原因,相关产品也往往采用多模态交叉的方式进行验证。部分产品应用使用了基于面部表情的情感计算,辅助剖析用户的情感状态^[22]。清华大学未来实验室正在研究基于面部的家居应用产品,例如基于面部特征的美妆镜研究,通过增强现实技术,辅助用户化妆。总体而言,面部数据在智能家居领域的潜力还有待挖掘。

2.3 语音数据

家居环境下语音数据的采集主要依赖于麦克风类设备,以嵌入到其他应用产品的形式呈现,如智能音箱、智能电视、智能平板等。语音数据的处理主要分为语音识别和语音情感分析2部分^[23-25],语音识别即对获取的语音数据进行语义理解的过程,语音情感分析则是在语音识别的基础上对语音的语气、语调、语义中所表达的情感进行识别。

语音识别目前在智能家居中是主流的交互方式,通过解析用户的语音指令,控制设备为用户提供相应的服务。比如智能音箱作为智能家居的交互门户,可以辅助用户控制全屋的设备。当前的交互方式主要基于对语音内容的识别,而基于语音情感的分析应用还相对较少,这也是智能家居领域未来基于语音交互研究的一个热点方向。

2.4 定位数据

定位数据即表示物体之间空间关系的一种信息。定位系统通过感知空间位置属性,利用相关算法计算物理实体的物理位置或逻辑位置信息。随着通信技术和电子制造工艺的不断发展,室内定位技术包含了众多的解决方案,识别精度也在不断的提升,主要有以下几种的解决方案:(1)视觉定位,通过使用可移动相机的图像数据进行定位,确定固定相机中的图像待测目标的位置,需要选择三维建筑模型、图像、投影目标、部署目标等作为参考点^[26];(2)红外线定位,基于有源信标、红外成像的不同,使用红外线进行定位。有源信标即通过使用若干台红外接收机,而物体将附有身份信息的信号进行发射^[27];(3)超声波定位,通过使用反射式测距法,采用多变定位来对物体进行定位,根据接受波和发射波的时间差进而推算出距离差异,达到定

位的目的^[28]; (4) Wifi 信号定位, 通过利用无线局域网, 不需要额外部署设备, 主流的方式是使用基于接收信号的强度指示的指纹进行定位, 但环境的复杂程度对于定位结果有较大影响^[29]; (5) 超宽带定位, 包含超宽带接收器、参考标签、其他标签, 通过发送给超窄脉冲来传输数据, 可以获得较大的带宽, 并且发射功率较低^[30]; (6) 蓝牙和紫蜂定位, 通过在室内环境部署静态参考点, 利用质心法、临近探测法、多变定位等技术进行定位^[31]; (7) 电容地板定位, 通过在地面铺设电容地板来进行人体的实时追踪, 但难以对于不同人群进行身份识别^[32]。

室内定位技术弥补了传统定位技术的不足, 能够为用户在家居环境提供更优的体验。主要具有如下几个方面的应用潜力: 可以根据人体不同身体部位进行分别定位, 实现人体的姿态识别^[33]; 基于用户的位置信息的时间序列, 可以得到室内行为的活动分布图, 进而进行活动分析和整体家居布局优化; 对于特定物品进行定位, 可以进行数字孪生和虚实结合的系统构建。定位技术可以在保护隐私的同时, 获取到用户的多维度数据, 在人机交互领域目前受到了广泛的关注。

2.5 环境数据

随着越来越多的电子设备接入到住户环境中, 识别环境信息中的风险因素也变得十分重要, 传感器的发展也使得监测环境数据变得可能。目前, 智能家居环境监测系统主要包含温度传感器、湿度传感器、气体传感器(可检测二氧化碳、可燃气体、易挥发性气体等)、智能水表、电表、天然气表等。

环境传感器获取的数据维度较为单一, 主流的数据处理流程为: 预先设置报警阈值, 当指标超过阈值时触发报警信号, 并与其他设备进行联动。近年来, 也有一些研究利用机器学习技术对传感器的数据进行特征挖掘, 比如基于用电设备的电量追踪, 来评估家居火灾风险^[34]。针对独自居住的老年人, 通过记录其水表读数的变化来分析老年人的生命安全状况^[35]。

随着人工智能技术的发展, 大规模的环境数据将不仅用来监测和报警, 也与其他模态数据进行耦合, 对用户的身体健康状态、行为偏好等进行评估。

2.6 现存问题和未来展望

随着电子技术、软件算法和通信网络的蓬勃发展, 智能家居呈现出了百花齐放的态势, 市场上出现了各种满足用户不同需求的产品。然而, 由于行业内的生态环境不统一, 不同公司间存在技术沟壑和信息封锁, 相关技术平台和产品难以形成进行统一的标准化集成和互相通信。数据的采集和应用面临着数据来源广泛、数据类型差异大、技术协议错综复杂的现状, 这对于用户来说存在学习成本高、交互体验差等问题, 而对于产业界和社会来讲, 则会造成硬件资源浪费、数据资源浪费等现象, 构建能够兼容众多产品的平台、协议乃至健康的生态圈是非常严峻且急迫的需求。

3 智能家居场景下的智能感知研究

智能家居场景下的智能感知研究是研究智能家居的基础性、必要性工作, 是完善智能家居场景下应用功能、拓展应用、协作和交互多样化的重要手段, 是以数据为驱动结合技术建立以人为本的智能家居环境的前提条件。相关研究为智能家居提供了感知计算能力、数据处理能力, 并从感知中获得需求、解决需求。

3.1 家庭级的计算与数据中心

搭建面向家庭用户的计算中心和数据中心是智能家居物联网中的重要一环, 意在将数据量和计算范围控制在以家庭为单位的人居场景下, 形成数据采集、处理和计算的流式处理, 与更高等级的数据中心协作, 形成独立的数据采集、存储、计算能力。随着智能家居设备的存储、计算能力的增长, 利用设备资源设计合适于家庭尺寸的计算中心、数据中心是必要的, 这可以加强本地资源的利用率并且有效保护数据隐私。

智能家居数据中心收集场景、设备和人因等多模态数据, 并存入半/非结构化的本地或云端数据库。目前, 常见的解决方案包括使用由网络附加存储^[36]、云端处理、虚拟/物理交换机所组成的直接存在于家庭单位的数据及计算中心。在已有研究中多采用软件定义网络的方法, 基于 OpenFlow 网络

通讯协议将数据平面和控制平面分离,并通过开放接口来对网络进行控制^[37],将人工智能(AI)代理和深度学习算法模型引入传统软件定义网络架构中^[38],从而应对动态网络化境、流量控制和多任务处理的挑战。

虚拟机架构配置是多源异构云数据中心效率的重要相关项,在不同情况下其配置方案的最优选择是不同的,这取决于可用硬件功能和发送请求的虚拟机特性,如CPU核心数、内存量、存储大小等。其中,数据中心物理机的性能起关键作用^[39]。Omer等^[40]展示了一种针对于CPU处理性能和内存,考虑请求优先级并对功率和流量进行感知,以达到最小化功耗、资源浪费和网络带宽使用量的启发式算法解决方案,在与首次拟合递减^[41]、节能和服务质量感知虚拟机配置^[42]和优先级感知虚拟机分配^[43]比较后的结果显示,减少了最多29%的总网络消耗、18%的功耗和68%的资源浪费。

通常,用户驻地设备雾计算层^[44]的处理能力高于物联网层,配合应用联邦学习^[45]、边缘计算^[46]和分布式计算^[47]可以解决服务延迟、非内网带宽限制、数据隐私等问题^[48]。Yosuf等^[49]提出了新的物联网雾云架构,其展示了在家庭范围内使用内部交换机和服务器,在接入点、路由器、网关等位置部署雾单元进行计算的方案。

信息安全直接威胁到整个物联网系统,人居环境下产生的高维数据更可能涉及到公民的个人隐私,由智能家居物联网组成的家庭级别的数据中心和计算中心的安全和隐私问题是不容忽视的^[50]。随着由智能家居用户驻地设备、传感网、计算中心产生的大量密集型数据和未来更多的数据增长空间,一些数据保护方案逐渐过时并难以满足智能家居物联网环境下的隐私敏感数据的需求^[51]。密钥管理^[52-54]、异常检测^[55-57]、隐私协议、政策执行^[58]是目前应对数据安全威胁常用方法。如Samanta等^[59]构建了针对云环境(包括私有云)基于支持向量机的优化树加密策略,利用自然启发式算法支持向量机通过计算位置点与其最近邻点之间的距离计算未知量,通过在高维空间中投影来预测最佳超平面,并线性地分离所有特征向量以生成加密密钥。针

对智能家居物联网数据中心, Lu等^[60]提出了一种雾计算增强的轻量级保护数据聚合的方案,该方案采用Paillier同态加密算法、中国余数定理、单向哈希链技术将混合物联网设备的数据聚合为一个,并且在早期边缘加入虚假数据,以增强物联网数据中心的安全性。

3.2 情感感知

情感感知是智能家居学习、了解人类真实生活状态、协同协作、更好的改善人居环境的必要手段,其感知主体为智能家居环境下的人类真实情感。未来智能家居环境下将包含多种类设备和密集式传感器,这给予了情感感知应用的基础条件。情感感知的策略和融合的数据模态类型除了音频、视频^[61]、文本^[62]等,还在智能家居环境下增加了步态数据、生理信号和社交媒体数据等。通过前文提到的智能家居环境中的数据研究方式方法,在人居环境下可以采集多种模态的连续人类行为数据,将数据放置在已部署学习算法的边缘云/设备上进行标注,再对多模态数据进行提取、融合^[63]和情感感知。

情感感知分为单模态情感感知和多模态情感感知,常用的感知方法包括特征提取、降维分析、机器学习等^[64]。当前相关研究已经验证,多模态情感感知与单一模态的情感感知相比,其受到外界因素的影响更小、识别准确率更高^[65]。Jia等^[66]提出了一种双流异构图递归神经网络对多模态生理信号进行融合情感感知的方法,用图变换器网络、图卷积神经网络和时域依赖的门控循环单元组成流模型,进而将融合的时空域特征用于情感感知。在真实家居环境中,多模态数据会涉及不同模态的时间序列数据,其异步性增加了多模态数据融合的难度。Lv等^[67]基于跨模态转化器提出了渐进式模态强化方法,通过模态差异强化其特征,生成强公共信息以增强异步数据情况下对人类情感的预测。在智能家居环境下已有一些基于情感感知的应用,如个性化家庭音频系统,实时语音情感分析^[68],智能光环境^[69]、智能医疗保健系统^[70]等。

3.3 事件感知

应用于智能家居环境下的事件判断分为独立事件判断与连续性事件判断,事件判断结果和进一

步指令通过智能家居物联网的用户驻地设备、传感网和计算中心协作处理并广播。首先,从感知层面对智能家居环境下的空间、设备、人类进行感知,主要获取传感网连接的用户驻地设备数据、环境传感器数据、人类行为和生理数据,再通过计算中心快速进行事件判断。通过对智能家居系统权限的增强,使智能家居中的设备彼此相关,进而通过分析上下文信息达到事件判断、预测和联动的目的。

独立事件判断指通常通过单一或多个的用户驻地设备、传感器或人类行为所进行的针对单一事件判断,最常见的应用为设备用量检测和人居环境下的异常检测。通过在设备中内置的多种传感器来判断设备的运行情况并单独执行不同的自动化任务,事件对象包括各种状态信息及位置信息、集线器信息和事件源设备信息^[71]。Sikder等^[72]提出了Aegis+,一个由数据采集模块、上下文生成模块和异常检测模块组成的安全架构。更多的案例如能源用量检测^[73]、门窗检测、天然气检测、烟雾检测^[74]、设备启停、健康检测^[75-76]、跌倒检测^[77]、入侵检测等^[78-79]。

连续事件判断是基于多重环境传感器、控制器模块、数据管道和服务层的多重协作实现的,对异构智能环境中生成的连续、广泛、互相传递的高级信息活动通过分布式算法、深度学习、对抗神经网络、联邦学习、边缘计算等方式识别、判断、预测连续的人居活动^[80]。如Lesani等^[81]提出的基于行为模式识别智能家居使用人群的传感器事件包识别方法。这种连续事件判断是在条件随机场对用户不确定行为所进行的判断。Lesani等^[81]使用概率图形模型处理现实世界中的不确定性问题并使用贝叶斯网络和隐马尔可夫模型对所提取特征进行建模,引入了事件包和事件序列,通过环境传感器提取居民行为模式,包括震动信息^[82-83]、步态信息^[84]、语音信息^[85]、视频图像信息^[86]、WiFi信号^[87]、超声波信号^[88]、无线射频识别^[89]及生理信号等^[90],并考虑事件频率,进而对日常活动进行判断和进一步协作。

3.4 行为感知

智能家居环境下的行为感知,所预期感知的是环境下人类的真实非表演行为,其本质是由情感或

事件所触发的在智能家居环境下的人类行为。一方面,传统集成嵌入式传感器的可穿戴设备^[91](如智能手表、手环、传感器集成人工制品、便携式生理设备等)和智能手机^[92]等可以在日常生活中获取大量的有效人类行为数据。另一方面,得益于传感器网络,应用视觉^[93-94]、嗅觉、音频、惯性单元、环境^[95]、信道状态信息^[96]等多模态数据在智能家居环境下结合深度学习算法^[97]进行全天候行为感知和识别。行为感知在智能家居研究中发挥着重要作用,可应用在识别日常生活活动^[98],例如协助做饭^[99]或帮助痴呆症患者洗手^[100]等场景中。在现实生活中,通过传感器数据识别和理解人类行为,结合“以人为本的设计”,给人类社会带来了巨大的效益^[101]。

大多数用于行为感知的技术要求居民佩戴带有传感器的设备,例如智能手机或智能手表等惯性传感器,或者通过在智能家居中安装摄像头来识别用户行为。Voit等^[102]提出一种更隐蔽且隐私友好的方法,利用嗅觉传感器来监测活动期间产生的气味,例如烹饪、洗澡活动或用餐。与工作环境等相对较开放式的室内交互环境相比,智能家居大多是封闭空间,有利于气味数据的收据和测量。清华大学未来实验室嗅觉计算组^[103]进行了基于气味采集的嗅觉输入界面的探索,提出了根据气味的采集、识别、存储和回放等一系列设计准则,为将嗅觉传感器应用于智能家居系统中人类日常生活活动的识别提供了可能性。

3.5 现存问题和未来展望

如何通过全面调用传感器网络来处理复杂决策问题是目前智能家居领域的难点。智能家居物联网中的复杂决策通常以人居环境下的所产生的密集型数据为驱动,结合信息融合方法,经由数据计算中心处理,对复杂和高维事件进行决策并直接下放指令至物联网环境下的设备,以个性化方式辅助协作或模拟用户行为,如数据融合的贝叶斯决策网络^[104]。目前,利用传感器网络和边缘计算等方法已经可以在智能家居环境下完成一些线性决策,如能源管理^[105]、对象管理^[106]和一些简单的非线性决策,如设备启停、温度调节等^[107]。而拥有更好用户体验的复杂决策涉及诸多特征信息和问题,如用户

适配问题、物联网系统平台通信、风险分析等。随着决策复杂度的提升,所需的决策网络和特征数量将会大大提升。Wieland 和 Zeiner^[108]在物联网系统中提出使用结构化、标准化的层次分析法^[109],通过独立标准进行加权的方式获得目标决策。智能家居物联网中的复杂决策方案在学术界和工业界仍在不断探索中,其中难点和痛点主要来源于数据集(由于采集效率、隐私化尚未有连续、完整、非表演性的、广泛的大规模数据集用于模型训练)、人因影响(尚未有完全准确的可以识别人类意图的方案)、权重判别(尚未有全面可靠的对于事件、特征的权重分配方案)、决策政策(尚未有支持机器决策和对其决策负责的政策)。清华大学未来实验室智能家居课题组针对以上难点和痛点,在以人为本的理念上、以数据为驱动的原则上,通过建立家庭级别的计算和数据中心,应用行为感知、情感感知和事件感知进行研究,旨在探索出全面可靠的解决方案。

4 面向智能家居的多模态人机交互与创新场景研究

4.1 多模态人机交互

智能家居经历互联网时代、物联网时代和智慧物联网时代,其交互模式逐渐由传统交互模式发展为基于 App+智能设备的点对点交互模式,目前正向多种模态的交互模式方向发展。

当前,智能家居系统主要通过智能手机等显示终端上的图形用户界面进行交互与控制,使用户能够以信息可视化的方式定义传感器时间和执行器动作之间的关系,例如,一个基于任务条件触发的网络服务平台^[110]和一种基于流的可视化编程开发工具^[111]。现代语音用户界面的兴起,使得控制智能设备的方式正在从以控制器为中心的模式转变为以人为中心的模式,用户在智能环境中的交互操作更加自然。然而,用于智能环境中的自然语言理解存在着种种挑战,自然语言理解模型需要理解各种话语结构,包括语法错误、口语陈述。自然语言的多样性体现在以多种语法形式表达或间接表达同

一目标,如何识别相关的用户意图^[112]限制了面向复杂智能环境的语音界面开发,大多数智能语音产品不同程度上约束了用户的指令发出的语法结构^[113]。Noura 等^[114]通过理解用户在智能家居环境直接和间接的目标,提出了一种面向目标的基于自然语言的语音界面原型,不限制用户智能使用设备命令。

4.2 创新应用场景

传统的智能家居场景专注于提高产品及系统的功能性及可用性,随着物联网技术的发展,人机交互领域的相关研究也在探索如何将智能家居场景的体验塑造为更加令人愉快^[115]、有意义^[116]和个性化^[117-118],使其更符合现代的生活方式。Jensen 等^[119]借用 Nelson 与 Stolterman 的“心之所需”(desiderata)概念延伸出理想智能家居场景的 3 种典型画像,分别是助手、优化者和享乐主义者。这些洞察反映在家庭实践及家居场景的可持续发展中。情感计算在智能家居领域被广泛讨论,通过监测用户的实时情绪而提供积极的反馈,Dang 等^[120]构建了情绪感知智能镜,通过灯光反馈促进调节用户自我反思、改善丰富日常生活和提升幸福感。智能设备在家庭中无处不在,为亲子互动和育儿场景提供了潜力。Beneteau 等^[121]观察到父母会通过智能音箱来实现育儿目标,通过对语音界面的使用以促进教育和培养沟通技巧。

许多研究者从多居住者等复杂家庭结构的角度来对智能家居的应用场景进行考量。随着智能设备技术的稳定发展,这些设备已经集成到用户家庭的长期基础设施中^[122],智能环境所包含的技术手段及交互方式将如何与复杂且不断变化的人际关系和使用情境相互作用,是设计智能家居创新应用场景时亟需考虑的重要维度。由于家庭成员的动态变化,每个家庭都会经历不同的社交情况,极大地影响了智能设备的共享场景及使用模式,不应为智能家居中的用户赋予家中给所有智能设备相同的角色^[123],例如,儿童可能被授予访问电视权限的子集,但是儿童房间里的照明系统的管理、智能家居系统需要根据特定的家庭关系和智能设备的使用情况,在家庭成员之间适当分配和共享权限^[124]。不当的权限管理可能导致潜在的经济损失、家庭成

员隐私的侵犯以及由于隐私敏感内容的传播而损害家庭内部的价值体系。Park 和 Lim^[125]通过为期一周的参与式用户研究,阐释了人工智能音箱作为家庭共享技术的潜力,基于家庭成员对于其的共同拥有权,参与者期望其在不同场景下实现家庭社交和个人隐私的无缝衔接与支持。Geeng 与 Roesner^[126]指出智能家居的使用反应了家庭中的现有关系动态和权力结构(如父母和孩子),在智能家居场景的设计中应考虑到不同的关系类型,包括非传统家庭单位中可能存在的各种关系,比如室友、房东和租户、家庭雇员等。目前市场中的智能设备仍然在适应和探索细粒度权限的概念,缺少对以家庭为单元的智能家居共享场景的深入挖掘,以最大限度地帮助减少共同居住者之间使用智能家居的差异性。

4.3 无障碍智能家居的发展

智能设备得益于与技术的发展与推动,智能灯具、智能音箱以及智能家电在家庭中的应用越来越普及,智能设备的多样性和关联性也可以通过一些解决方案进行整合,以形成更加统一且自然的交互和用户体验。然而,尽管语音是人类最自然的交流方式之一,但这种单模态的解决方案并不适用于所有人,例如有听力障碍的人群,也不适用于所有场景,如远距离交互和噪音场景下。

智能家居系统的交互设计方法也需要考虑到居住者的能力和需求的潜在多样性,并从包容性的视角提升智能家居系统的可及性和提供无障碍的解决方案,老年人、残障人士及儿童或孕妇等特殊用户群体可能会需要特殊的交互模式。例如, Bempong 等^[127]为听力受损人士提供了一个集成系统,可以访问和控制智能家居; Pawale 和 Chidrawar 的工作^[128]针对残障人士创建了基于视觉显示的交互界面。

然而,现有的智能家居交互的无障碍解决方案大多局限于一台设备集中控制系统或者面向单一障碍人群。为解决此问题, Almeida 等^[129]采用模块化和解耦的多模态交互架构提出了对未来多模态无障碍智能家居的愿景,并通过以用户为中心的设计方法绘制以家庭为单位的用户建模,以提高对目

标用户的需求和环境感知,选择更具适应性的交互模态。

多模态交互同样可以用以解决用户多样性^[130]和上下文异构等问题,这些解决方案允许多样化用户在复杂场景中与智能家居环境进行交互^[131]。多模态交互也被用于开发适合老年人的无障碍智能家居环境中^[132]。随着全球人口老龄化的加剧,越来越多的人可以受益于自动化物联网系统来帮助他们在家居环境中进行独立的日常活动。大多数智能家居等复杂系统是通过搭载在智能手机或其他触摸屏终端上的应用程序控制的,操作过程对于老年用户群体来说过于复杂。Rus 等^[133]设计了一款老年用户友好的智能家居控制立方装置,通过结合视觉触觉反馈的实体交互界面和手势操作实现远程控制。Dermikan 和 Olgunt^[134]对于老年人及有身体和视觉障碍的成年人共 161 位参与者进行了调查,以确定家庭环境中的哪些调整与这些用户最相关。现有的无障碍智能家居系统旨在通过成为辅助生活设备和监控系统来增强老年人及残障人士在自己家中的独立性,以用户友好的形式帮助他们生活^[135]。Kon 等^[136]基于 6 个日常生活活动的主要类别(安全、健康、体育活动、个人卫生护理、社会参与和娱乐休闲)对技术进行分析后,发现多数研究缺少对后 3 个主题的关注,大部分都集中在对于安全健康有威胁的基础需求的迭代。

随着智能家居技术的发展,未来的研究者与设计开发人员可以更全面地探究无障碍智能家居的真实需求和应用场景的细分,使得老年人及残障人士将可以有机会定制智能家居,以满足他们的个性化需求,提供最佳的无障碍智能家居辅助环境。

4.4 现存问题和未来展望

在智能家居环境下,有关多模态人机交互问题、场景研究及无障碍智能家居的相关研究与人类活动具有更直接的关系,以人为本的智能计算和设计是相关领域的内在核心,如何平衡用户体验、设计方案和伦理问题是相关领域一直存在的挑战。在多模态人机交互问题中,基于单模态的交互方式已经使用户感到厌倦^[137]。与此同时,基于单模态的交互方式在某些特定情景下并无法提供最优的用

户体验^[138],如何利用多模态通道提高人机交互能力和用户体验是未来需要关注的重点。

随着越来越多新兴技术的发展和新一代智能设备的推广,人类在智能家居环境中的各项需求不断增多,所衍生的新型场景愈发复杂。目前,多数研究仅针对于单一场景,尚未有对于智能家居环境中场景综合的、全面的整理,这可能使得不同场景之间的上下文关系被忽略。虽然对于场景相关问题的研究一直持续至今,但因其具有一定迭代更新的属性,且人类生活无时无刻均处于场景空间内,结合技术和人类社会需求的创新场景研究将成为未来讨论的重点。

随着以人为本的设计理念和对于弱势群体的进一步关注,适老化问题和无障碍设计问题被广泛关注,其中,可用性問題一直为相关研究的难点。同时,随着现代社会的发展和社会观念的转变,对于无障碍智能家居的个性化需求也急剧增加。目前的研究多关注于上文所提及的安全、健康主题内,对于无障碍智能家居相关的广泛社会福祉问题,在未来是具有巨大潜力的研究方向。

5 未来实验室在智能家居领域的工作

清华大学未来实验室智能家居课题组正在致力于探究面向智能家居的用户行为和情感理解以及多模态人机交互理论与方法。目前,课题组正在建设未来智能家居的原型空间实验平台(图2),该平台包括全套家居智能硬件,12种以上的全屋传感器,1 PB的存储服务器和高速计算服务器中心,这是国内外首个领先的面向科研用途的综合性智能家居实验平台,支持多人同时入住、多模态的数据采集和多种维度的人机交互体验。同时,课题组正在开展深度的用户体验研究,依托基础实验平台,邀请被试组独立入住实验平台,在不受外界因素干扰的前提下采集用户多源异构大数据,并对采集的数据进行脱敏、清洗、标注、特征提取、融合等处理,以形成国内外首个PB量级的人类自然行为与情感海量数据库。在平台和数据的基础上,课题

组正在开展基于智能感知的多模态人机交互理论与方法研究,融合听觉、视觉、触觉等模态,构建多元信息和多模态交互的耦合模型,实现自然交互和优质体验的智能家居典型示范应用。

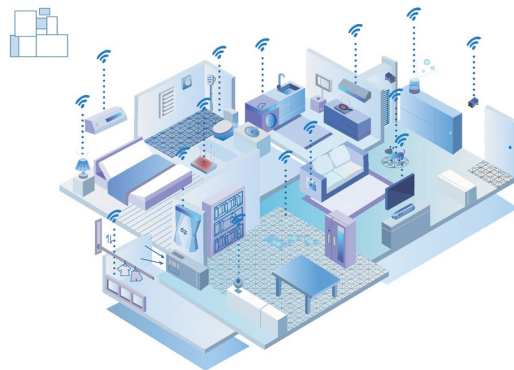


图2 清华大学未来实验室未来智能家居原型空间实验平台设计与实景图

6 结论

从国内外前沿智能家居研究平台发展现状、智能家居环境中的数据研究、智能家居场景下的智能感知研究、面向智能家居的多模态人机交互与创新场景研究4个方面,对面向未来的智能家居前沿研究展开了综述,研究所讨论的内容包含在由“感知”“思考”和“执行”3个维度所组成的智能家居研究框架内,由浅及深、由整体到具体、由技术端到人因角度讨论了覆盖智能家居研究中的主要痛点、难点、高相关性及高关注度问题。

通过对国内外前沿研究的总结,梳理并提出了智能家居相关研究的潜在问题和未来发展方向展望:(1) 未来智能家居平台应进一步整合资源,形成具有高集成化、兼容性、稳定性的物联网中心平台;(2) 应从硬件和数据层面打破现有壁垒,减

少数据、硬件资源成本,进一步实现互联互通;(3) 应建立非表演训练数据集,全面结合传感器网络进行以数据为驱动的多模态感知及复杂决策研究;(4) 应由单一模态转变为多模态的人机交互方式研究;(5) 应发展基于需求的创新场景研究和无障碍设计研究。

参考文献(References)

- [1] Lasqueti-Reyes J. Smart home report 2021[R]. Hamburg: Statista, 2021.
- [2] Firouzi R, Rahmani R, Kanter T. An autonomic iot gateway for smart home using fuzzy logic reasoner[J]. *Procedia Computer Science*, 2020, 177, 102-111.
- [3] Mpeis P. The Anyplace 4.0 IoT localization architecture [C]//2020 21st IEEE International Conference on Mobile Data Management (MDM). Versailles, France: IEEE, 2021: 218-225.
- [4] Crisostomi E. Analytics for the sharing economy: Mathematics, engineering and business perspectives[M]. Cham: Springer, 2020: 309-333.
- [5] Lueth K. IoT Platform companies landscape 2019/2020: 620 IoT platforms globally[EB/OL]. (2019-12-23)[2021-10-18]. <https://iot-analytics.com/iot-platform-companies-landscape-2020/>.
- [6] Toshev A, Szegedy C. DeepPose: Human pose estimation via deep neural networks[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Columbus: IEEE, 2014: 1653-1660.
- [7] 邓益依, 罗键欣, 金凤林. 基于深度学习的人体姿态估计方法综述[J]. *计算机工程与应用*, 2019, 55(19): 22-42.
- [8] 付心仪, 蔡天阳, 薛程, 等. 基于 BGRU-FUS-NN 神经网络的姿态情感计算方法研究[J]. *计算机辅助设计与图形学学报*, 2020, 32(7): 1070-1079.
- [9] Dong J, Jiang W, Huang Q, et al. Fast and robust multi-person 3d pose estimation from multiple views[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach: IEEE, 2019: 7792-7801.
- [10] Kadkhodamohammadi A, Gangi A, de Mathelin M, et al. A multi-view RGB-D approach for human pose estimation in operating rooms[C]//2017 IEEE winter conference on applications of computer vision (WACV). Santa: IEEE, 2017: 363-372.
- [11] Zhou Z, Chen X, Chung Y C, et al. Activity analysis, summarization, and visualization for indoor human activity monitoring[J]. *IEEE transactions on circuits and systems for video technology*, 2008, 18(11): 1489-1498.
- [12] Chen S, Yang R R. Pose Trainer: correcting exercise posture using pose estimation[J]. *CoRR*, 2020, abs/2006.11718.
- [13] Huang Z, Liu Y, Fang Y, Horn B K. Video-based fall detection for seniors with human pose estimation[C]//2018 4th International Conference on Universal Village (UV). Boston: IEEE, 2018: 1-4.
- [14] Fu X, Xue C, Yin Q, et al. Gesture based fear recognition using nonperformance dataset from VR horror games [C]//2021 9th International Conference on Affective Computing and Intelligent Interaction (ACII). Nara: IEEE, 2021: 1-8.
- [15] Sun L, Fu S, Wang F. Decision tree SVM model with Fisher feature selection for speech emotion recognition [J]. *EURASIP Journal on Audio, Speech, and Music Processing*, 2019, 2019(1): 1-14.
- [16] Zhao W, Chellappa R, Phillips P J, et al. Face recognition: A literature survey[J]. *ACM computing surveys (CSUR)*, 2003, 35(4): 399-458.
- [17] Wold S, Esbensen K, Geladi P. Principal component analysis[J]. *Chemometrics and Intelligent Laboratory Systems*, 1987, 2(1-3): 37-52.
- [18] Delac K, Grgic M. Face recognition, I-tech education and publishing[J]. Vienna, Austria, 2007.
- [19] Ojala T, Pietikainen M, Maenpaa T. Multiresolution gray scale and rotation invariant texture classification with local binary patterns[J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2002, 24(7): 971-987.
- [20] Agarwal S, Mukherjee D P. Facial expression recognition through adaptive learning of local motion descriptor [J]. *Multimedia Tools and Applications*, 2017, 76(1): 1073-1099.
- [21] 陶娟. 智能家居门锁技术综述[J]. *现代制造技术与装备*, 2020(3): 162-164.
- [22] Tan L, Zhang K, Wang K, et al. Group emotion recognition with individual facial emotion CNNs and global image based CNNs[C]//Proceedings of the 19th ACM International Conference on Multimodal Interaction. Glasgow: Association for Computing Machinery, 2017: 549-552.
- [23] Özseven T. Investigation of the effect of spectrogram im-

- ages and different texture analysis methods on speech emotion recognition[J]. *Applied acoustics*, 2018, 142: 70–77.
- [24] Zhang S, Zhang S, Huang T, et al. Speech emotion recognition using deep convolutional neural network and discriminant temporal pyramid matching[J]. *IEEE transactions on multimedia*, 2018, 20(6): 1576–1590.
- [25] Jin Q, Li C, Chen S, et al. Speech emotion recognition with acoustic and lexical features[C]//*Proceedings of 2015 IEEE International Conference on Acoustics, Speech and Signal Processing*. South Brisbane: IEEE, 2015: 4749–4753.
- [26] Mautz R, Tilch S. Survey of optical indoor positioning systems[C]//*2011 international conference on indoor positioning and indoor navigation*. Guimaraes: IEEE, 2011: 1–7.
- [27] Want R, Hopper A, Falcao V, et al. The active badge location system[J]. *ACM Transactions on Information Systems (TOIS)*, 1992, 10(1): 91–102.
- [28] 吴雨航, 吴才聪, 陈秀万, 等. 几种室内无线定位技术简介[N]. *中国测绘报*, 2008–02–01.
- [29] 赵永翔. 基于无线局域网的室内定位系统研究[D]. 武汉: 武汉大学, 2010.
- [30] 阮陵, 张翎, 许越, 等. 室内定位: 分类, 方法与应用综述[J]. *地理信息世界*, 2015(2): 8–14.
- [31] 裴凌, 刘东辉, 钱久超. 室内定位技术与应用综述[J]. *导航定位与授时*, 2017, 4(3): 1–10.
- [32] Valtonen M, Maentausta J, Vanhala J. Tiletrack: Capacitive human tracking using floor tiles[C]//*2009 IEEE international conference on pervasive computing and communications*. Galveston: IEEE, 2009: 1–10.
- [33] Xiao L, Pan H. Human Activity Recognition System Based on WiFi Signal[J]. *Journal of Beijing University of Posts and Telecommunications*, 2018, 41(3): 119–124.
- [34] 陆永菊, 吴桂初, 吴自然. 基于灰色神经网络的剩余电流预测研究[J]. *软件导刊*, 2020, 19(10): 61–65.
- [35] Nie X, Fan T, Wang B, et al. Big data analytics and IoT in operation safety management in under water management[J]. *Computer Communications*, 2020, 154: 188–196.
- [36] Ou Z, Song M, Hwang Z H. Is cloud storage ready? Performance comparison of representative IP-based storage systems[J]. *Journal of Systems and Software*, 2018, 138: 206–221.
- [37] Shirmarz A, Ghaffari A. Performance issues and solutions in SDN-based data center: A survey[J]. *The Journal of Supercomputing*, 2020, 76(10): 7545–7593.
- [38] Mao B, Tang F, Fadlullah Z M, et al. An intelligent route computation approach based on real-time deep learning strategy for software defined communication systems[J]. *IEEE Transactions on Emerging Topics in Computing*, 2019, 9(3): 1554–1565.
- [39] Fu Q, Sun E, Meng K, et al. Deep Q-learning for routing schemes in SDN-based data center networks[J]. *IEEE Access*, 2020, 8: 103491–103499.
- [40] Omer S, Azizi S, Shojafar M, et al. A priority, power and traffic-aware virtual machine placement of IoT applications in cloud data centers[J]. *Journal of Systems Architecture*, 2021, 115: 101996.
- [41] Keller G, Tighe M, Lutfiyya H. An analysis of first fit heuristics for the virtual machine relocation problem[C]//*2012 8th International Conference on Network and Service Management(CNSM) and 2012 Workshop on Systems Virtualization Management(SVM)*. Las Vegas: IEEE, 2012: 406–413.
- [42] Wang S H, Huang P P W, Wen C H P, et al. EQVMP: Energy-efficient and QoS-aware virtual machine placement for software defined datacenter networks[C]//*The International Conference on Information Networking 2014 (ICOIN2014)*. Phuket: IEEE, 2014: 220–225.
- [43] Son J, Buyya R. Priority-aware VM allocation and network bandwidth provisioning in software-defined networking (SDN)-enabled clouds[J]. *IEEE Transactions on Sustainable Computing*, 2018, 4(1): 17–28.
- [44] Rani R, Kumar N, Khurana M, et al. Storage as a service in fog computing: A systematic review[J]. *Journal of Systems Architecture*, 2021, 116: 102033.
- [45] Li T, Sahu A K, Talwalkar A, et al. Federated learning: Challenges, methods, and future directions[J]. *IEEE Signal Processing Magazine*, 2020, 37(3): 50–60.
- [46] Shi W, Cao J, Zhang Q, et al. Edge computing: Vision and challenges[J]. *IEEE internet of things journal*, 2016, 3(5): 637–646.
- [47] Li B, Yu J. Research and application on the smart home based on component technologies and Internet of Things [J]. *Procedia Engineering*, 2021, 15: 2087–2092.
- [48] Chang C, Srirama S N, Buyya R. Indie fog: An efficient fog-computing infrastructure for the internet of things [J]. *Computer*, 2017, 50(9): 92–98.
- [49] Yusuf B A, Musa M, Elgorashi T, et al. Energy efficient distributed processing for IoT[J]. *IEEE Access*, 2020, 8: 161080–161108.

- [50] Devarakonda S, Halgamuge M N, Mohammad A. Research anthology on privatizing and securing data[M]. USA: IGI Global, 2021: 1672–1694.
- [51] Kouicem D E, Bouabdallah A, Lakhlef H. Internet of things security: A top-down survey[J]. *Computer Networks*, 2018, 141: 199–221.
- [52] Sicari S, Rizzardi A, Miorandi D, et al. Internet of things: Security in the keys[C]//*Proceedings of the 12th ACM Symposium on QoS and Security for Wireless and Mobile Networks*. Malta: Association for Computing Machinery, 2016: 129–133.
- [53] Chien H. Group-oriented range-bound key agreement for internet of things scenarios[J]. *IEEE Internet of Things Journal*, 2018, 5(3): 1890–1903.
- [54] Roman R, Alcaraz C, Lopez J, et al. Key management systems for sensor networks in the context of the Internet of Things[J]. *Computers & Electrical Engineering*, 2011, 37(2): 147–159.
- [55] Raza A, Hardy L, Roehrer E, et al. GPSPiChain-Blockchain and AI based self-contained anomaly detection family security system in smart home[J]. *Journal of Systems Science and Systems Engineering*, 2021, 30: 433–449.
- [56] Vigoya L, Fernandez D, Carneiro V, et al. Annotated dataset for anomaly detection in a data center with IoT sensors[J]. *Sensors*, 2020, 20(13): 3745.
- [57] Yamauchi M, Ohsita Y, Murata M, et al. Anomaly detection in smart home operation from user behaviors and home conditions[J]. *IEEE Transactions on Consumer Electronics*, 2020, 66(2): 183–192.
- [58] Sicari S, Rizzardi A, Miorandi D, et al. Security policy enforcement for networked smart objects[J]. *Computer Networks*, 2016, 108: 133–147.
- [59] Samanta D, Alahmadi A H, Karthikeyan M P, et al. Cipher block chaining support vector machine for secured decentralized cloud enabled intelligent IoT architecture [J]. *IEEE Access*, 2021, 9: 98013–98025.
- [60] Lu R, Heung K, Lashkari A H, Ghorbani A A. A lightweight privacy-preserving data aggregation scheme for fog computing-enhanced IoT[J]. *IEEE Access*, 2017, 5: 3302–3312.
- [61] Guanghui C, Xiaoping Z. Multi-modal emotion recognition by fusing correlation features of speech-visual[J]. *IEEE Signal Processing Letters*, 2021, 28: 533–537.
- [62] Singh P, Srivastava R, Rana K P S, et al. A multimodal hierarchical approach to speech emotion recognition from audio and text[J]. *Knowledge-Based Systems*, 2021, 229: 107316.
- [63] Jiang Y, Li W, Hossain M S, et al. A snapshot research and implementation of multimodal information fusion for data-driven emotion recognition[J]. *Information Fusion*, 2020, 53: 209–221.
- [64] Zhang J, Yin Z, Chen P, et al. Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review[J]. *Information Fusion*, 2020, 59: 103–126.
- [65] Imani M, Montazer G A. A survey of emotion recognition methods with emphasis on E-learning environments[J]. *Journal of Network and Computer Applications*, 2019, 147: 102423.
- [66] Jia Z, Lin Y, Wang J, et al. HetEmotionNet: Two-stream heterogeneous graph recurrent neural network for multi-modal emotion recognition[C]//*Proceedings of the 29th ACM International Conference on Multimedia*. Virtual: Association for Computing Machinery, 2021: 1047–1056.
- [67] Lv F, Chen X, Huang Y, et al. Progressive Modality Reinforcement for Human Multimodal Emotion Recognition From Unaligned Multimodal Sequences[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Nashville: IEEE 2021: 2554–2562.
- [68] Kang D, Seo S. Personalized smart home audio system with automatic music selection based on emotion[J]. *Multimedia Tools and Applications*, 2019, 78(3): 3267–3276.
- [69] Soheilian M, Fischl G, Aries M. Smart lighting application for energy saving and user well-being in the residential environment[J]. *Sustainability*, 2021, 13(11): 6198.
- [70] Mano L Y, Faiçal B S, Nakamura L H, et al. Exploiting IoT technologies for enhancing Health Smart Homes through patient identification and emotion recognition[J]. *Computer Communications*, 2016, 89–90: 178–190.
- [71] Fernandes E, Jung J, Prakash A. Security analysis of emerging smart home applications[C]//*2016 IEEE symposium on security and privacy (SP)*. San Jose: IEEE, 2016: 636–654.
- [72] Sikder A K, Babun L, Uluagac A S. Aegis+ a context-aware platform-independent security framework for smart home systems[J]. *Digital Threats: Research and Practice*, 2021, 2(1): 1–33.
- [73] Bhatt D, Hariharasudan A, Lis M, et al. Forecasting of energy demands for smart home applications[J]. *Ener-*

- gies, 2021, 14(4): 1045.
- [74] Salhi L, Silverston T, Yamazaki T, et al. Early detection system for gas leakage and fire in smart home using machine learning[C]//2019 IEEE International Conference on Consumer Electronics (ICCE). Las Vegas: IEEE, 2019: 1–6.
- [75] Wang J, Spicher N, Warnecke J M, et al. Unobtrusive health monitoring in private spaces: The smart home[J]. *Sensors*, 2021, 21(3): 864.
- [76] Abreu J, Oliveira R, Garcia-Crespo A, et al. TV interaction as a non-invasive sensor for monitoring elderly well-being at home[J]. *Sensors*, 2021, 21(20): 6897.
- [77] Tafeeque M, Koita S, Spicher N, et al. Multi-camera, multi-person, and real-time fall detection using long short term memory[C]//Medical Imaging 2021: Imaging Informatics for Healthcare, Research, and Applications. California: International Society for Optics and Photonics, 2021: 1160109.
- [78] Nadafa R A, Hatturea S M, Bonala V M, et al. Home security against human intrusion using Raspberry Pi[J]. *Procedia Computer Science*, 2020, 167: 1811–1820.
- [79] Malche T, Maheshwary P. Internet of Things (IoT) for building smart home system[C]//2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC). Palladam: IEEE, 2017: 65–70.
- [80] Demongivert C, Bouchard K, Gaboury S, et al. A distributable event-oriented architecture for activity recognition in smart homes[J]. *Journal of Reliable Intelligent Environments*, 2021(7): 215–231.
- [81] Lesani F S, Fotouhi Ghazvini F, Amirkhani H. Smart home resident identification based on behavioral patterns using ambient sensors[J]. *Personal and Ubiquitous Computing*, 2021, 25(1): 151–162.
- [82] Mukhopadhyay B, Anchal S, Kar S. Person identification using structural vibrations via footfalls for smart home applications[J]. *IEEE Internet of Things Journal*, 2021, 8(17): 13384–13396.
- [83] Hu Z, Zhang Y, Pan S. Vibration-based indoor occupant gait monitoring with robot vacuum cleaners[C]//International Conference on Internet-of-Things Design and Implementation(IoTDI' 21). Charlottesville: Association for Computing Machinery, 2021: 247–248.
- [84] Chapron K, Bouchard K, Gaboury S. Real-time gait speed evaluation at home in a multi residents context[J]. *Multimedia Tools and Applications*, 2021, 80: 12931–12949.
- [85] Nguyen M S, Vo T L. Resident identification in smart home by voice biometrics[C]//International Conference on Future Data and Security Engineering. Ho Chi Minh City: Springer, 2018: 433–448.
- [86] Fang S, Islam T, Munir S, et al. EyeFi: Fast human identification through vision and WiFi-based trajectory matching[C]//2020 16th International Conference on Distributed Computing in Sensor Systems (DCOSS). Marina del Rey: IEEE, 2020: 59–68.
- [87] Mo H, Kim S. A deep learning-based human identification system with wi-fi csi data augmentation[J]. *IEEE Access*, 2021, 9: 91913–91920.
- [88] Mokhtari G, Zhang Q, Nourbakhsh G, et al. BLUE-SOUND: A new resident identification sensor—Using ultrasound array and BLE technology for smart home platform[J]. *IEEE Sensors Journal*, 2017, 17(5): 1503–1512.
- [89] Kim K, Li S, Heydariaan M, et al. Feasibility of LoRa for smart home indoor localization[J]. *Applied Sciences*, 2021, 11(1): 415.
- [90] Guinea A S, Boytsov A, Mouline L, et al. Continuous identification in smart environments using wrist-worn inertial sensors[C]//Proceedings of the 15th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services. New York: Association for Computing Machinery, 2018: 87–96.
- [91] Nweke H F, Teh Y W, Al-Garadi M A, et al. Deep learning algorithms for human activity recognition using mobile and wearable sensor networks: State of the art and research challenges[J]. *Expert Systems with Applications*, 2018, 105: 233–261.
- [92] Hassan M M, Uddin M Z, Mohamed A, et al. A robust human activity recognition system using smartphone sensors and deep learning[J]. *Future Generation Computer Systems*, 2018, 81: 307–313.
- [93] Kim K, Jalal A, Mahmood M. Vision-based human activity recognition system using depth silhouettes: A smart home system for monitoring the residents[J]. *Journal of Electrical Engineering & Technology*, 2019, 14(6): 2567–2573.
- [94] Dang L M, Min K, Wang H, et al. Sensor-based and vision-based human activity recognition: A comprehensive survey[J]. *Pattern Recognition*, 2020, 108: 107561.
- [95] Ranieri C M, MacLeod S, Dragone M, et al. Activity recognition for ambient assisted living with videos, inertial units and ambient sensors[J]. *Sensors*, 2021, 21(3): 768.
- [96] Wang D, Yang J, Cui W, et al. Multimodal CSI-based

- human activity recognition using GANs[J]. *IEEE Internet of Things Journal*(Early Access). 2021, 8(24): 17345–17355.
- [97] Yadav S K, Tiwari K, Pandey H M, et al. A review of multimodal human activity recognition with special emphasis on classification, applications, challenges and future directions[J]. *Knowledge-Based Systems*, 2021, 223: 106970.
- [98] Zhang M, Sawchuk A A. USC-HAD: A daily activity dataset for ubiquitous activity recognition using wearable sensors[C]//*Proceedings of the 2012 ACM conference on ubiquitous computing*. Pittsburgh Pennsylvania: Association for Computing Machinery, 2012: 1036–1043.
- [99] Yordanova K, Whitehouse S, Paiement A, et al. What's cooking and why? Behaviour recognition during unscripted cooking tasks for health monitoring[C]//*2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*. Kona: IEEE, 2017: 18–21.
- [100] Mihailidis A, Boger J N, Craig T, et al. The COACH prompting system to assist older adults with dementia through handwashing: An efficacy study[J]. *BMC geriatrics*, 2008, 8(1): 1–18.
- [101] Kim E, Helal S, Cook D. Human activity recognition and pattern discovery[J]. *IEEE pervasive computing*, 2009, 9(1): 48–53.
- [102] Dang C T, Seiderer A, André E. Theodor: A step towards smart home applications with electronic noses [C]//*Proceedings of the 5th international Workshop on Sensor-based Activity Recognition and Interaction*. Berlin: Association for Computing Machinery, 2018: 1–7.
- [103] 路奇, 吴昊, 梁婉, 等. 基于气味采集的嗅觉输入界面探索[J]. *计算机辅助设计与图形学学报*, 2020, 32(7): 1018–1025.
- [104] Constantinou A C, Fenton N. Things to know about Bayesian networks: Decisions under uncertainty, part 2 [J]. *Significance*, 2018, 15(2): 19–23.
- [105] Yao L, Damiran Z, Lim W H. Energy management optimization scheme for smart home considering different types of appliances[C]//*2017 IEEE international conference on environment and electrical engineering and 2017 IEEE industrial and commercial power systems Europe (EEEIC/I&CPS Europe)*. Milan: IEEE, 2017: 1–6.
- [106] Bergeron F, Bouchard K, Gaboury S, et al. Tracking objects within a smart home[J]. *Expert Systems With Applications*, 2018, 113: 428–442.
- [107] Aheleroff S, Xu X, Lu Y, et al. IoT-enabled smart appliances under industry 4.0: A case study[J]. *Advanced engineering informatics*, 2020, 43: 101043.
- [108] Wieland G, Zeiner H. A survey on criteria for smart home systems with integration into the analytic hierarchy process[C]//*International Conference on Decision Support System Technology*. Loughborough: Springer, 2021: 55–66.
- [109] Dos Santos P H, Neves S M, Sant'Anna D O, et al. The analytic hierarchy process supporting decision making for sustainable development: An overview of applications[J]. *Journal of cleaner production*, 2019, 212: 119–138.
- [110] IFTTT[EB/OL]. [2021–10–27]. <https://ifttt.com>.
- [111] OpenJS Foundation. Node-RED[EB/OL]. [2021–10–27]. <https://nodered.org>.
- [112] Li T J J, Labutov I, Myers B A, et al. *Studies in conversational UX design*[M]. Cham: Springer, 2018: 119–137.
- [113] Catania V, Delfa G C L, Monteleone S, et al. GOOSE: Goal oriented orchestration for smart environments[J]. *International Journal of Ad Hoc and Ubiquitous Computing*, 2019, 32(3): 159–170.
- [114] Noura M, Heil S, Gaedke M. Natural language goal understanding for smart home environments[C]//*Proceedings of the 10th International Conference on the Internet of Things*. Malmö Sweden: Association for Computing Machinery, 2020: 1–8.
- [115] Jordan P W. Human factors for pleasure in product use [J]. *Applied ergonomics*, 1998, 29(1): 25–33.
- [116] Hassenzahl M. Experiences before things: A primer for the (yet) unconvinced[C]//*CHI'13 Extended Abstracts on Human Factors in Computing Systems*. Paris: Association for Computing Machinery, 2013: 2059–2068.
- [117] Holtzblatt K. What makes things cool? Intentional design for innovation[J]. *Interactions*, 2011, 18(6): 40–47.
- [118] Raptis D, Bruun A, Kjeldskov J, et al. Converging coolness and investigating its relation to user experience[J]. *Behaviour & Information Technology*, 2017, 36(4): 333–350.
- [119] Jensen R H, Strengers Y, Kjeldskov J, et al. Designing the desirable smart home: A study of household experiences and energy consumption impacts[C]//*Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems*. New York: ACM Digital Library,

- 2018: 1–14.
- [120] Dang C T, Aslan I, Lingenfelter F, et al. Towards som-aesthetic smarthome designs: Exploring potentials and limitations of an affective mirror[C]//Proceedings of the 9th International Conference on the Internet of Things. Bilbao: Association for Computing Machinery, 2019: 1–8.
- [121] Beneteau E, Boone A, Wu Y, et al. Parenting with Alexa: exploring the introduction of smart speakers on family dynamics[C]//Proceedings of the 2020 CHI conference on human factors in computing systems. Honolulu: Association for Computing Machinery, 2020: 1–13.
- [122] Bowles N. Thermostats, locks and lights: Digital tools of domestic abuse[N]. *The New York Times*, 2018: 23.
- [123] He W, Golla M, Padhi R. Rethinking access control and authentication for the home internet of things (IoT) [C]//27th {USENIX} Security Symposium ({USENIX} Security 18). Baltimore: {USENIX} Association, 2018: 255–272.
- [124] Kanchi S, Karlapalem K. A multi perspective access control in a smart home[C]//Proceedings of the Eleventh ACM Conference on Data and Application Security and Privacy. Virtual: Association for Computing Machinery, 2021: 321–323.
- [125] Park S, Lim Y K. Investigating user expectations on the roles of family-shared AI speakers[C]//Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. Honolulu: Association for Computing Machinery, 2020: 1–13.
- [126] Geeng C, Roesner F. Who's in control? Interactions in multi-user smart homes[C]//Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems. Glasgow Scotland: Association for Computing Machinery, 2019: 1–13.
- [127] Bempong J, Stainslow J, Behm G. Accessible smart home system for the deaf and hard-of-hearing[EB/OL]. [2021–10–28]. https://www.rit.edu/ntid/nyseta/sites/rit.edu/ntid.nyseta/files/docs/fullpapers_PDFs/StanislawJoe-FullPaper.
- [128] Satish R P, Chidrawar S K. Real time implementation of accessible display design to control home area network[Z]. 2015.
- [129] Almeida N, Silva S, Teixeira A. Multimodal interaction for accessible smart homes[C]//Proceedings of the 8th International Conference on Software Development and Technologies for Enhancing Accessibility and Fighting Info-exclusion. Thessaloniki: Association for Computing Machinery, 2018: 63–70.
- [130] Dahl D A. Multimodal interaction with W3C standards [M]. Cham: Springer International Publishing, 2016.
- [131] Almeida N, Teixeira A, Silva S, et al. The am4i architecture and framework for multimodal interaction and its application to smart environments[J]. *Sensors*, 2019 19(11): 2587.
- [132] Spinsante S, Cippitelli E, De Santis A, et al. Multimodal interaction in a elderly-friendly smart home: A case study[C]//International Conference on Mobile Networks and Management. Würzburg: Springer, 2014: 373–386.
- [133] Rus S, Helfmann S, Kirchbuchner F, et al. Designing smart home controls for elderly[C]//Proceedings of the 13th ACM International Conference on Pervasive Technologies Related to Assistive Environments. Corfu: Association for Computing Machinery, 2020: 1–10.
- [134] Demirkan H, Olguntürk N. A priority-based "design for all" approach to guide home designers for independent living[J]. *Architectural Science Review*, 2014, 57 (2): 90–104.
- [135] de Oliveira G A A, de Bettio R W, Freire A P. Accessibility of the smart home for users with visual disabilities: an evaluation of open source mobile applications for home automation[C]//Proceedings of the 15th Brazilian symposium on human factors in computing systems. São Paulo: Association for Computing Machinery, 2016: 1–10.
- [136] Kon B, Lam A, Chan J. Evolution of smart homes for the elderly[C]//Proceedings of the 26th International Conference on World Wide Web Companion. Perth: International World Wide Web Conferences Steering Committee, 2017: 1095–1101.
- [137] Ahire S, Rohs M. Tired of wake words? Moving towards seamless conversations with intelligent personal assistants[C]//Proceedings of the 2nd Conference on Conversational User Interfaces. Bilbao, Spain: Association for Computing Machinery, 2020: 1–3.
- [138] Pradhan A, Lazar A, Findlater L. Use of intelligent voice assistants by older adults with low technology use [J]. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 2020, 27(4): 1–27.

A review of the frontier research on future smart home

FU Xinyi¹, ZHANG He¹, XUE Cheng^{1,2}, SUN Tongxin^{1,2}

1. The Future Laboratory, Tsinghua University, Beijing 100084, China

2. Academy of Arts & Design, Tsinghua University, Beijing 10084, China

Abstract With the rapid development of artificial intelligence, the Internet of Things and big data, etc., the smart home enters the Era 3.0, which profoundly relies on user-based big data mining and deep learning. At present, smart interaction, intelligent perception, and complex decision-making in the smart home urgently need new supporting technologies and methods. This paper integrated a large amount of frontier research worldwide and summarized and refined the three elements required for the future smart home: "perception" "thinking" and "execution". This paper began with an overview of the frontier platforms for smart homes in China and abroad, followed by an elaboration and review of the research on data collection, feature extraction, perception in smart home scenarios, and multimodal human-computer interaction and innovation scenarios for smart homes, respectively. The authors presented a series of related research work of Future Lab in the smart home field for researchers in related fields.

Keywords smart home; affective computing; intelligent perception; multimodal interaction ●



(责任编辑 徐丽娇)