

面向新型智能设备的交互技术

伍楷舜 邹永攀
深圳大学

关键词：交互技术 声波感知 振动感知

新型智能设备

近年来，随着传感器、微机电系统 (Micro-Electro-Mechanical System, MEMS)、电池等技术的发展，新型智能设备如雨后春笋般不断涌现。这些设备外观各异、功能多样、硬件组成各具特色，被广泛地应用于智慧医疗、智慧家居和智能交通等各类生活场景中，不仅给人们的生活带来了便捷，更拓宽了人们对于未来生活的想象空间。

这些智能设备（如图1所示）既可以以手表、腕带、眼镜、衣服、鞋子等形态被穿戴在用户身上，与其融为一体，对其运动、生理等参数进行监测；也可以以音箱、台灯、冰箱等形态部署于家居场景中，与物理环境有机融合，实现设备互联、环境监测、智能控制等功能；还可以以各种车载设备的形式配

备于车辆上。这些功能的实现都依赖于用户与设备之间的交互，包括用户认证、功能控制和文本输入三方面内容。用户认证是合法使用智能设备的第一步，它是设备对用户合法性的甄别过程，能够有效地对设备中的数据进行隐私保护，对于记录用户生理数据的智能设备尤为重要。功能控制是指用户操控智能设备的基本功能，比如调节音量和屏幕亮度、切换画面等。文本输入技术则允许用户向智能设备传达文本信息与指令等。

新型智能设备可以以多样的形态为用户所使用，其交互需求也存在多样性。由于智能穿戴设备的屏幕尺寸较小，且配备传感器的种类和数量相对较少，基于触屏的文本输入方式对于此类设备来说存在效率低和体验差等缺点。对于无屏设备，传统的触屏交互更是无能为力。因此，传统的基于触摸



图1 新型智能设备示例

屏技术的交互方式已经难以作为统一的交互接口满足用户与设备的所有交互需求。

交互问题制约着新型智能设备在现实生活中的应用,使得它们无法真正地发挥出潜能。探索适合于新型智能设备的交互技术既给相关领域的研究带来了新的机遇,也是促使此类设备得到更普遍应用进而推动相关产业进一步发展的必由之路。然而,由于新型智能设备存在硬件能力、尺寸、传感器配置等多方面的限制,设计合适的交互技术充满难度与挑战。

面向移动设备的交互技术

在面向移动设备的交互技术这一研究领域,研究者们已经开展了诸多工作,提出了一系列提升用户与设备之间交互效率与体验的技术方法。尽管这些交互技术并不是专门为新型智能设备而设计的,却给相关领域的研究工作带来了启发。

基于内置传感器的交互技术

语音识别是一项目前被广泛应用在智能设备中进行文本输入和功能控制的交互技术。该技术利用智能设备内置的麦克风录制用户的语音,借助语音识别技术让设备“听懂”用户所说的话,从而实现文本或指令的输入。近年来,随着与深度学习技术的结合,语音识别的最新性能已然逼近人类水平,其识别精度高达97%。然而,从实用的角度来看,该技术存在着在公共场合隐私性差、噪声环境中性能下降明显、易受用户生理/心理状态和说话口音等因素影响等缺点,大大限制了其在现实生活中的普及与应用。

智能设备内置的其他类型传感器,比如惯性测量单元也被研究者们用来感知与识别用户肢体动作。该技术虽然无需用户佩戴额外的硬件设备,但由于传感器对动作的感知粒度较为粗糙,它要求用户佩戴着智能设备做出较大幅度的动作以抵抗肢体无关动作的干扰。对用户而言,此种交互过程不自然、体验不良,并不具备很强的实用性。

基于无线电信号的交互技术

研究者们还提出使用无线信号收发装置,如RFID^[1]和60 GHz设备^[2],利用信号在物理空间中传播的多径效应来设计人机交互系统。然而,由于此类硬件设备难以与智能设备有机结合,因而并不适合于移动设备的交互系统设计。

基于Wi-Fi信号的人机交互技术近年来得到了学术界的广泛重视。其基本假设是Wi-Fi信号收发装置作为通信基础设施广泛分布在人类的活动空间,而智能设备也具备Wi-Fi信号收发功能且无须额外的硬件支持。基于Wi-Fi信号的交互技术是利用信号在物理空间中传播的多径效应,建立用户肢体运动与接收信号变化之间的关联,实现用户与设备之间的交互。研究者们深入到Wi-Fi的物理层,利用其信道状态信息(Channel State Information, CSI)来实现基于Wi-Fi的手势识别。CSI刻画了Wi-Fi信号传输信道的状态,反映出因手势动作造成的信道状态的变化。相比于其他物理量,如接收信号强度指示(Received Signal Strength Indication, RSSI),CSI对于信道状态的表征更加精细和丰富,对于动作的识别能力也更强。因为CSI具有上述优点,研究者在基于Wi-Fi CSI的动作识别上做了大量的研究工作,在ACM MobiCom、IEEE INFOCOM等国际顶级会议和期刊上发表了一系列高水平的论文^[3]。尽管基于Wi-Fi CSI的人机交互思路新颖独特,但是在实际应用中却存在根本性瓶颈需要突破。例如,Wi-Fi信号在整个空间中经多径传播,除了目标的运动会改变传播路径引起CSI改变外,环境中无关人员或物体的运动甚至环境布局的改变也会对CSI产生影响,干扰系统对于目标动作的识别^[4]。

在用户认证领域,传统智能设备的用户认证技术主要包括指纹识别、面部识别和虹膜识别等。尽管这些技术都已经很成熟,准确率和可靠性等性能指标均达到很高水准,但是这些技术对于传感器配置、CPU运算能力有较高的要求,并不完全适用于新型智能设备。

面向新型智能设备的交互技术

近年来,针对新型智能设备的交互难题,研究者们提出了一些新颖的交互技术。其中,基于声波感知和振动感知的交互系统设计颇受关注,成为该领域的一个研究热点。

基于声波感知的交互技术

声波感知是利用智能设备普遍具有的声学传感器(麦克风、扬声器),根据应用需求,以特定方式对声波信号进行调制后经由扬声器发射出去,并利用麦克风接收回波信号,通过对回波信号的分析实现对用户肢体动作的识别或追踪。在此过程中,声波信号在空间中经多径传播被用户肢体(手部)反射回来,使得回波信号中带有肢体运动的信息(如图2所示),再利用物理建模或机器学习的方法,建立回波信号变化与肢体运动之间的关系,最终实现对肢体动作追踪或识别的目的。

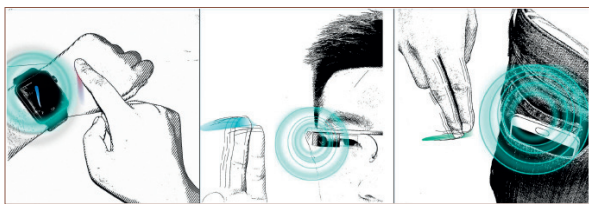


图2 基于声波感知的交互技术示例

以物理建模的方式建立回波信号变化与肢体运动之间的关系,主要依赖于准确提取出与用户肢体反射的回波信号对应的相位变化,并基于此相位变化得到其位移改变,进而实现对肢体运动毫米级别的追踪和定位,该技术方案主要被应用于文本输入、游戏操控等人机交互场景。在这一研究领域,南京大学、德州奥斯汀大学等高校团队在ACM MobiCom、ACM MobiSys等计算机网络和移动计算领域著名学术会议发表了多篇论文,做了许多有益的尝试^[5,6]。

建立回波信号变化与肢体动作间关系的另一条路径是利用机器学习算法,从回波信号中提取特征,经过特征工程筛选,再运用合适的学习模型,获得动作识别的结果。该技术方案主要被应用于用户认

证或功能控制等交互场景^[7]。这种技术路线的不足是需要模型训练时收集大量的训练数据,且很难保证模型在不同场景下的适用性等。

相比于以往的移动设备交互技术,基于声波感知的交互具有以下鲜明的特点。(1)普适性好:智能设备中普遍内置麦克风和扬声器,无须额外的硬件设备;(2)识别粒度精细:高频声波信号的分辨率高,能够识别毫米级别的细微运动;(3)鲁棒性高:借助合适的信号调制方式,系统能够抵抗较强的外在干扰。因此,此项技术适用于设计新型智能设备的交互系统。

基于振动感知的交互技术

基于振动感知的交互技术(如图3所示),主要是利用手指敲击物体(比如手背、桌面等)产生的振动信号经由固体传播被振动传感器所拾取,从而进行交互系统的设计。该技术目前的主要应用场景包括文本输入与用户认证,其优点是相关硬件成本低、尺寸小、易于与新型智能设备相结合。

基于振动的文本输入技术的基本原理是:用户在平面上不同位置处敲击时产生的振动信号与敲击位置紧密相关,通过区分振动信号的差异就可以对敲击位置进行准确定位或判别。将平面上的不同位置与基本文本单元相关联就可以进一步实现文本输入的功能。此种方式相当于将智能设备的交互界面进行了拓展,使得交互过程更加便捷高效。

基于振动信号的用户认证则是利用个体敲击平面时产生信号的特异性实现对用户的识别。该技术方案潜在假设在于:一方面,不同个体的敲击动作存在差异,敲击产生的振动信号可以唯一地区分不同的用户;另一方面,当以用户的手背作为敲击平面时,由于个体肢体构造的差异,敲击点到振动传感器之间的传播将呈现出个体独特性,进一步增强振动信号的独特性,强化了用户认证的性能。在具体的系统实现中,将利用低成本、便携式振动传感器来获取信号,并借助机器学习算法对敲击位置或者用户进行识别。在此研究方向上,深圳大学伍楷舜教授团队、美国罗

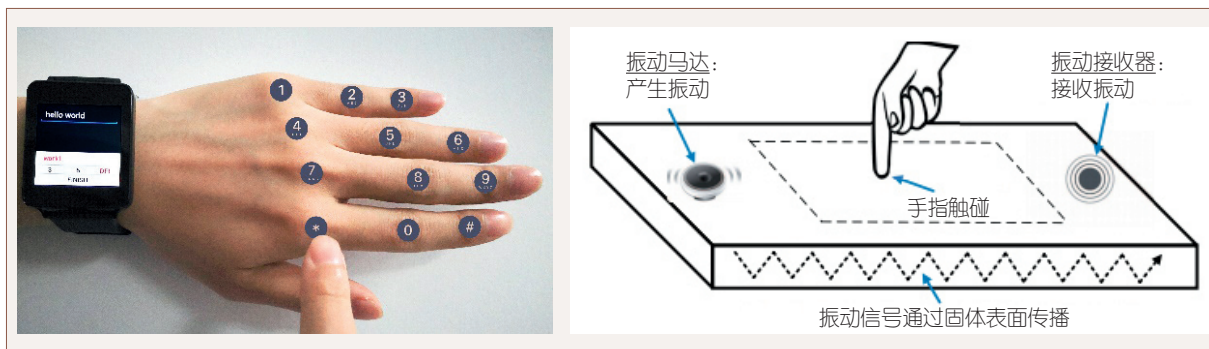


图3 基于振动信号的文本输入与用户认证技术

格斯大学的 Yingying Chen 教授团队做了一些很有意义的探索，相关成果发表于 ACM MobiCom 和 ACM CCS 等国际知名会议上^[8,9]。

挑战与发展

对于新型智能设备的交互而言，基于声波或振动信号的交互技术具有较大的应用潜力，在设计满足多种交互需求的交互系统方面也更具有优势。然而，当前的研究并没有完全解决新型智能设备的交互难题。为了推动相关技术的应用落地，仍然需要解决多方面的挑战。

首先是能量消耗问题。由于新型智能设备电池容量普遍有限，续航时间较短，而基于声波感知的交互方式需要由设备发出经过调制的声波，能量消耗较大，对设备的续航有较大的影响。如何最大程度地优化算法、减少能量消耗，同时保证算法的准确性和可靠性，具有非常大的挑战。

其次是鲁棒性问题。尽管相比于 Wi-Fi 感知，基于声波和振动信号的感知方式不容易受到周边无关人员、物体或者所处环境改变的影响，但是却容易受到用户自身运动的干扰。目前此类系统的设计与评估大都是在用户处于静止状态的场景，没有解决移动状态下进行交互的问题。

最后是用户体验的问题。交互系统的设计归根结底是以用户体验为核心。当前基于声波或振动感知的交互设计都较少地讨论用户体验问题，在今后的研究工作中值得关注和探索。



伍楷舜

CCF 专业会员。深圳大学特聘教授。主要研究方向为物联网、移动计算等。
wu@szu.edu.cn



邹永攀

CCF 专业会员。深圳大学助理教授。主要研究方向为可穿戴与移动计算、人机交互等。
yongpan@szu.edu.cn

参考文献

- [1] Yang L, Chen Y, Li X Y, et al. Tagoram: Real-time tracking of mobile RFID tags to high precision using COTS devices[C]// *International Conference on Mobile Computing & Networking*. ACM, 2014.
- [2] Wei T, Zhang X. mTrack: High-precision passive tracking using millimeter wave radios[C]// *Proceedings of ACM Mobicom*, Paris, France, 2015.
- [3] Wang H, Zhang D, Ma J, et al. Human respiration detection with commodity WiFi devices: Do user location and body orientation matter?[C]// *ACM International Joint Conference on Pervasive & Ubiquitous Computing*. ACM, 2016.
- [4] Zou Y, Liu W, Wu K, et al. Wi-Fi radar: Recognizing human behavior with commodity Wi-Fi[J]. *IEEE Communications Magazine*, 2017, 16(2):381-393.
- [5] Wang W, Liu A X, Sun K. Device-free gesture tracking using acoustic signals[C]// *International Conference on Mobile Computing & Networking*. ACM, 2016.

更多参考文献：<http://dl.ccf.org.cn/cccf/list>