

自然人机交互

关键词：自然人机交互

特邀编辑：史元春
清华大学

背景

人机交互是人与计算机之间为完成某项任务所进行的信息交换过程。计算机形态和使用情境(context)日益复杂多样,交互技术已经成为终端和应用创新的核心竞争力,自然交互是发展趋势。我们希望,人们与手持设备、家居设备、穿戴设备、机器人、无人车,在很多场景中以更自然的模态(比如用语音,用语义丰富的手势,甚至是我们日常的行为)发生互动,人们能获得可理解性与感受效果俱佳的信息反馈。所谓的自然,是在信息呈现和交互表达上,最大程度地符合人对现实世界已有的认知,信道充分,并能降低甚至无须学习成本,而在表达上,还体现在人不需要很精准的表达,可以是某种模糊的表达和传达的方式,而机器端能够给我们正确的理解和精准的服务。

传统的图形用户界面的交互理论主要面向桌面计算环境,交互信道有限,交互接口和交互情境固

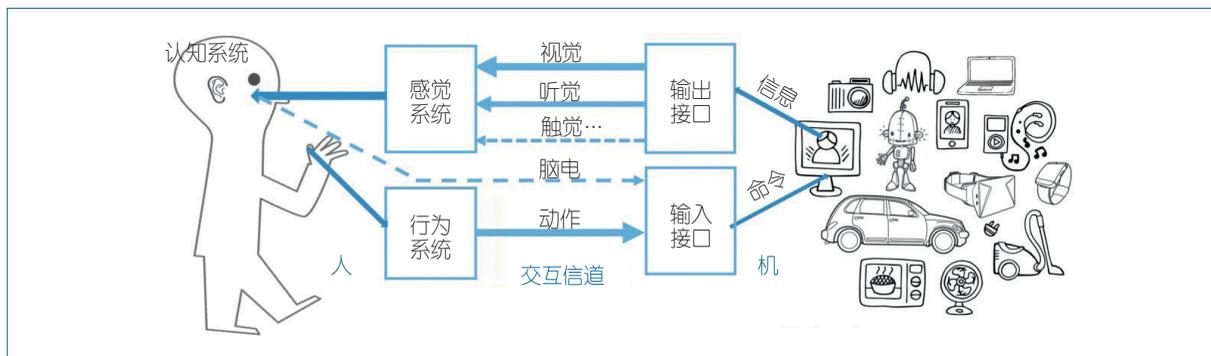
定,刻画范围窄,难以支持普适计算环境中通道综合多样化、环境云端融合化的自然交互。

已有人机交互理论研究

人机交互基础理论主要包括交互评价体系、交互认知机理、可计算的交互模型、交互情境感知等。

1. 交互评价体系方面,认知心理学家唐纳德·诺曼(Donald Norman)提出的以用户为中心的人机界面评价原则很有影响力,但它是针对以鼠标键盘输入、以屏幕输出的传统交互界面,而自然交互过程更加紧密和频繁,事件类型和关系复杂,现有方法不适用于多模态交互自然性的评价。

2. 交互认知方面,传统的用户信息处理模型¹围绕工作记忆、长时记忆、感知和运动系统来解释人处理信息的过程,但其描述的交互通道范围有限(感知只有视听觉的基本处理,运动只有使用工具



人机信息交互示意图

的小范围准确操控动作), 量化粗略。

3. 可计算交互模型方面, 费兹定律 (Fitts's law) 是应用最广泛的人体运动控制模型, 解释了指点空间目标时速度和精度之间的关系。然而该模型无法对连续交互运动过程的细节进行描述。

4. 交互情境感知方面, 研究者开始融合各种简单情境信息, 以推理人所处的活动、人的意图等具有更深层语义的情境。研究者也发现, 符合情境和用户心理模型的交互通道和界面将改进用户体验, 但心理模型 (mental model) 具有内隐性, 目前还没有测定用户心理模型的方法。

近年来, 交互理论最新进展主要包括: 扩展费兹定律解释自然交互通道上的精确控制能力, 比如手指、压力控制; 对自然交互动作中多点观测的行为的信道容量进行测量; 通过生理参数信号对交互认知的状态进行评价, 替代主观评价方法等。这些研究都一致地指向了更加深层次的问题, 即在交互系统中, 用户通过自然的交互通道进行输入输出时, 自然性如何定义, 自然交互是否有优化目标, 如何进行优化等。

人机交互自然性的核心问题

人机交互自然性的核心科学问题体现在人的基本交互信息处理能力、人机交互信道和交互情境三个递进层次上。

交互自然性的生理心理机理及计算模型

在行为上, 交互自然性体现为可用行为与生理结构的适合性, 包括多通道输入行为的操作复杂度和疲劳度机理, 探索与用户生理结构匹配的交互可达范围; 在感知上体现为信息呈现与知觉过程的协调性, 包括多模态信息到达率、复杂程度与用户信息处理速度和加工能力之间的协调性, 通过心理、生理与行为指标揭示信息过载对交互自然体验的影

响规律; 在认知上体现为接口界面与注意记忆决策的相容性, 包括用户加工多模态界面信息的注意选择和分配机制, 解决多模态海量信息与用户有限注意资源之间的矛盾。

自然交互模态的信道特征与交互语义

自然交互是人与机通过多通道输入多模态输出来完成信息交换。自然的信息交换过程需要以人的信息加工能力为约束, 研究自然交互模态的信道特征与交互语义, 以及输入输出接口融合机理。输入通道基于人体运动控制行为构建, 包括全身肢体动作、手势、精细肌肉控制、压力输入、眼动、语音等多种输入通道; 输出模态则包含视、听、触在内的人体主要感知来源。然而不同通道与不同模态在适合表达的数据类型以及信息传递速率上有差异, 需要探索多通道多模态人机信息交换的本质, 建立交互语义自然性和高效性的统一表示。

基于情境的交互路径优化原理

交互自然性与交互情境以及用户对于交互系统状态的认知有着重要关系, 属于高级认知层次的自然性问题。交互情境反映了交互的客观环境, 不仅涵盖了客观物理的环境约束, 也包括交互环境中的接口资源和交互设备。心理模型刻画了个体主观因素对交互的影响, 信道属性则表达了不同交互途径的特性与语义。为了达到整个交互过程的自然性, 核心是在不同情境下交互接口与界面都要满足用户心理、交互能力等方面的自然性约束, 即取得情境 - 心理 - 信道相容匹配。情境 - 心理 - 信道相容匹配策略与方法将为交互路径的生成与优化提供关键支持与依据。因此, 需要着重研究情境、心理、信道三者的相容匹配策略与方法, 包括情境心理关联的用户动态模型, 交互信道与用户心理的匹配策略与方法, 交互信道与交互情境的匹配策略与方法, 以及情境 - 心理 - 信道迭代式相容计算与优化策略。

¹ Human Information Processing Model, 出现于 20 世纪 80 年代初。

专题文章介绍

人机交互是一个典型的学科交叉的研究领域,从计算的角度看,是面向人的性能的计算接口和过程上的优化,人的性能的可计算性发觉和交互接口与过程的优化,需要多个相关学科的研究者密切合作,在研究问题和研究方法上融合互补,是取得突破性进展的必要模式。本期专题邀请了国内在自然人机交互领域与本人有较多合作研究经历的相关专家撰文,针对上述研究论题,从人的感知、认知和行为等方面展开讨论,并结合实际应用探讨自然交互理论和方法的最新进展、发展趋势和挑战。

在行为建模方面,人的动作(手势、头动、眼动)是自然输入的主要方式。对动作的表示和分析,以及基于动作的意图理解,都是自然人机交互的重要内容。清华大学计算机系副研究员喻纯等人撰写的《自然文本输入中的动作建模》,以自然文本输入这一具有挑战性的任务为应用背景,介绍复杂手部动作的建模技术以及输入意图的贝叶斯推理方法,展示了概率统计方法在自然交互动作建模和理解上的能力,具有很强的参考价值。

在感知方面,本期特别针对视频、虚拟现实(VR)和触觉邀请了三篇文章。中科院软件所研究员王宏安、马琴霞撰写的《面向可视媒体的自然交互》,以大规模视频数据的理解为研究目标。由于视频内容具有结构复杂、信息量大的特点,无法像文本那样被自动检索和有效利用,对于视频内容更深层次的交互分析技术目前尚未成熟。文章从草图交互、可视隐喻、可视分析等角度介绍交互式草图摘要、面向视频内容的螺旋摘要和视频地图等可视媒体的交互与分析方法。

清华大学工业工程系教授饶培伦撰写了《VR交互的人因研究》一文。虚拟现实作为一种重要的新兴计算机形态,随着Oculus、HTC Vive、PSVR、Google Glass、HoloLens等商用头戴式显示设备的发行,虚拟现实、增强现实、混合现实等技术逐渐开始实用。该文主要讨论了VR在输入和输出技术、手段上的发展,并关注VR应用过程中需要注意和

预防的一些问题。

北京航空航天大学教授王党校撰写了《触觉人机交互:从桌面式到穿戴式》一文。人与自然界交互时,触觉时时刻刻都在发挥着作用,软硬、摩擦、冷暖等信息必须通过触摸来感知。人类更复杂的情感交流也必须通过触摸实现。与此形成巨大反差的是,在人与机器交互时,触觉交互仍然处在萌芽期,计算机通过显示器、头盔、立体声耳机等设备可以给人呈现逼真的视觉体验和听觉体验,但在交互过程中用户能够获得的触觉体验极为贫乏。该文对触觉技术的分析有助于我们较全面深入地理解这个新的、正在发展的交互通道。

在认知方面,心理模型是人机交互中的基本概念,是用户理解交互系统和做出交互决策的依据。而交互系统“自然”与否的关键,在于用户是否能够更快地掌握其正确的使用方法(学习时间短),使用时心理负荷是否更低,以及是否产生了更愉快的情感体验。中科院心理所副研究员张警吁等人撰写的《自然交互的认知机理与心理模型》,从人机交互过程的一般心理规律出发,并以空管系统为例,阐述心理模型这一关键概念。

清华大学心理系副教授张丹等人撰写的《自然交互的生理计算》,阐述了通过测量人体生理信号以实时计算用户心理状态的生理计算(physiological computing)方法。生理信号是交互行为的物质基础,是心理状态的最直接体现。与心理状态的主观报告及行为评测方法相比,生理计算理论上可以得到更加客观、准确的评测结果。若在交互行为过程中全面测量来自中枢和外周神经系统的生理信号,有望提供交互心理状态更全面、丰富的信息。 ■



史元春

CCF 常务理事、会士, CCCF 前副主编。
清华大学计算机系“长江学者”特聘教授,
人机交互与媒体集成研究所所长, 清华
大学全球创新学院 GIX 院长。主要研究
方向为人机交互、普适计算等。
shiyu@tsinghua.edu.cn