

机器人自然交互 理论与方法

JIQIREN ZIRAN JIAOHU LILUN YU FANGFA

杜广龙 张平 著



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

国家自然科学基金项目(61403145)

中国博士后科学基金面上项目(2014M550436)

人机共融技术系列丛书

机器人自然交互 理论与方法

杜广龙 张平 著



华南理工大学出版社
SOUTH CHINA UNIVERSITY OF TECHNOLOGY PRESS

· 广州 ·

内 容 简 介

本书是作者在自然人机交互领域多年研究成果的总结。该书围绕面向机器人的智能人机交互技术展开,述及人机交互概论、人机交互发展历程、受限人机交互技术分类;重点介绍了非受限人机交互技术,包括:基于多传感器的智能感知技术、自然人机交互模型、智能辅助控制技术、基于多智能体的人机协作模型、非受限力反馈模型和多机器人自然交互与协作模型。同时对人机交互技术未来发展进行了展望。通过理论与实践的结合,力图为从事人机交互领域研究的人员提供参考,并起到抛砖引玉的作用。

本书可作为高等院校智能人机交互和人工智能等专业的本科生、研究生的教学用书,也可作为从事智能人机交互的工程技术和研究人员了解现代智能人机交互技术动向及应用新技术的参考书。

图书在版编目(CIP)数据

机器人自然交互理论与方法/杜广龙,张平著. —广州:华南理工大学出版社, 2017. 3

(人机共融技术系列丛书)

ISBN 978 - 7 - 5623 - 5146 - 7

I. ①机… II. ①杜… ②张… III. ①人工智能 - 研究 IV. ①TP18

中国版本图书馆 CIP 数据核字(2016)290920 号

机器人自然交互理论与方法

杜广龙 张 平 著

出 版 人: 卢家明

出版发行: 华南理工大学出版社

(广州五山华南理工大学 17 号楼, 邮编 510640)

<http://www.scutpress.com.cn> E-mail: scutel3@scut.edu.cn

营销部电话: 020 - 87113487 87111048 (传真)

策划编辑: 潘宜玲 刘 锋

责任编辑: 刘 锋 詹志青

印 刷 者: 佛山市浩文彩色印刷有限公司

开 本: 787mm × 1092mm 1/16 印张: 9.75 字数: 199 千

版 次: 2017 年 3 月第 1 版 2017 年 3 月第 1 次印刷

定 价: 35.00 元

版权所有 盗版必究 印装差错 负责调换

前 言

随着人类对太空或深海等未知领域探索的不断深入,以及某些高危工作的需求,机器人技术得到广泛关注和应用。由于当前人工智能技术的局限性和复杂的工作环境,当今机器人还无法实现完全的智能化和自动化,很多复杂的任务往往还需要人类的决策和参与。因此在机器人控制系统中,智能人机交互技术十分重要。它将操作员的意图智能地翻译为机器人的动作指令,从而实现操作员对机器人的引导和控制。

人机交互技术将认知科学、计算机科学、机械科学、人工智能、生命科学等多种学科结合起来,对人机交互中的交互行为、交互方式、机器辅助模型和人机协作模型等要素进行融合。人机交互充分发挥机器的智能性与自主性、人类的主导性和能动性,是一门实现人机优势互补、互相协作的技术。人机交互技术是随着计算机的兴起和广泛应用而发展起来的沟通工具,是人与计算机之间沟通的桥梁,消融了人与计算机之间通信和对话的界限,使得人与计算机的信息交流更加便捷和通畅。

早期的机器人编程控制方式采用示教再现的方式。这种人机交互方式属于受限人机交互方式。随着机器人应用越来越广泛,示教再现方式已经不能满足人类应用机器人的需求。20世纪70年代以来,一种新的控制方式使机器人技术发生了变革:用户可以用一种交互语言描述机器人的任务。美国斯坦福人工智能实验室于1973年开发出第一个机器人编程语言 WAVE。随后比较有代表性的是 IBM 公司的 T. J. Watson 研究中心开发的用于控制直角坐标系机械手的机器人语言: MAPLE、EMILY 和 ML。80 年代开始,研究者希望摆脱枯燥的机器人编程语言和冗长的编程代码,从而研究出以机器人图形仿真为基础的图形交互模型。90 年代开始,得益于精密电子设备的快速发展,出现了使用摇杆、力反馈器和外骨骼控制器等人机交互设备直接控制机器人的人机交互方式。这种交互方式可以使用户更加直观地控制机器人运动。这一类人机交互方式的特点是以机器为中心,主要表现为三个方面:一是人的交互基于离散事件;二是人的交互必须基于精确交互;三是用户饱受命令记忆和新功能学习之苦。这种方式限制了人机间的通信带宽,使得其既不能有效地利用当今计算机的强大计算和处理能力,也给用户控制机器人带来诸多不便。另一种是以用户为中心的非受限方式。这种方式强调将人类自然能力(尤其交流、运动和感知能力)与计算设

备及其感知和推理结合起来,通过多模态 (multimodal) 感知人类的自然行为,并以易于理解的多媒体 (multimedia) 形式实现多通道 (multichannel) 通信,建立“以人为中心”的感知用户界面。非受限人机交互方式包括:支持语音交互 (speech-based HCI)、支持笔迹交互 (pen-based/Calligraphic HCI)、支持视觉交互 (vision-based HCI)、支持情感交互 (affective-based HCI)、支持可穿戴交互 (wearable HCI)、支持人脑交互 (brain-computer Interaction)。这种非受限交互模式体现在以下三个方面:一是以用户为中心 (human-centered);二是多模态交互;三是智能感知 (multimedia perceptive)。

面向机器人的智能人机交互是一个非常具有挑战性的研究方向。其发展过程面临以下难题:人类行为在空间、时间和对象上的高维性;人类与机器存在异构性,两者之间的映射存在非线性、变参数、弱耦合等特性;机器人自身机构存在多自由度、变拓扑结构特点。随着机器人越来越深入人类生活的各个方面,对人机交互技术提出了越来越多的挑战。智能人机交互是未来人机交互的发展趋势,也是笔者一直坚持在这个方向上不懈地研究与探索的主要原因。回顾人机交互技术发展的历史和国内外目前对人机交互技术研究的现状,可以得到一个观点:单方向地让人类适应机器人或是让机器人适应人类都是一种不适当的做法。人类与机器人是两个不同本质的智能体,无论是两者的知识、决策行为还是思考行为都存在巨大的差异。与其让一方适应另一方,不如让双方互相适应。人机协调是人类与机器人相处的最好工作方式。只有进行分工合作,优势互补,才能尽可能地发挥人机交互技术的效能。鉴于笔者及机器人研究团队在智能人机交互领域取得的一些有益的积累和进展,深感应该将多年来在该领域的研究思想、理论方法、成果和心得做一系列的总结、归纳和提高,为从事智能人机交互技术研究的工程技术人员提供参考,促进智能人机交互技术的发展和应用。

本书涉及的研究工作,得到了国家“863”计划、国家自然科学基金项目(61403145)、中国博士后科学基金面上项目(2014M550436)以及华南理工大学的大力支持;在书稿撰写过程中得到了研究生刘欣、李备、罗锦聪、何子平、庞启超以及李迪教授等的积极帮助,在此表示衷心的感谢!

限于作者水平,书中疏漏和不足之处恳请读者和专家指正!

杜广龙 张 平
2016年10月

目 录

1 绪论	1
1.1 人机交互作用	1
1.2 人机交互的发展	2
1.3 人机交互元素	4
1.4 本章小结	6
2 人机特点	7
2.1 人的因素	7
2.2 机器的因素	8
2.3 人与机器的区别	9
2.4 本章小结	11
3 人机交互技术	12
3.1 传统人机交互技术	12
3.1.1 命令行交互	12
3.1.2 图形界面交互	13
3.1.3 直接操纵交互	15
3.2 新型人机交互技术	17
3.2.1 语音交互	17
3.2.2 视觉交互	19
3.2.3 眼球交互	20
3.2.4 肌肉感应交互	23
3.2.5 穿戴交互	26
3.2.6 情感交互	28
3.2.7 脑机交互	31
3.3 本章小结	33
4 人机协作技术	34
4.1 多智能体理论与技术	34
4.2 多通道融合交互技术	36
4.3 机器辅助技术	37
4.3.1 多体间碰撞检测模型	37
4.3.2 安全预警模型	44

4.3.3	具有预警功能的虚拟夹具模型	50
4.3.4	多自由度机器人路径规划模型	53
4.4	基于虚拟现实技术机器感知与反馈	55
4.4.1	虚拟现实技术	55
4.4.2	视觉反馈	58
4.4.3	听觉反馈	60
4.4.4	触觉反馈	63
4.4.5	多用户交互	67
4.5	三维重建	71
4.5.1	背景知识	71
4.5.2	算法实现	73
4.5.3	寻找匹配点集	75
4.5.4	计算旋转矩阵	76
4.5.5	体素更新	77
4.5.6	获取网格	78
4.5.7	实验结果及分析	80
4.6	本章小结	82
5	三维手势跟踪识别技术与动作信号处理	83
5.1	三维手势跟踪技术	83
5.1.1	基于立体视觉的手势跟踪技术	83
5.1.2	基于体感传感器的手势跟踪技术	85
5.1.3	基于穿戴式设备的手势跟踪系统	87
5.2	混合滤波器	90
5.2.1	基于粒子滤波的位置估算模型	90
5.2.2	基于粒子滤波的姿态估计模型	92
5.2.3	基于卡尔曼滤波的位置估算模型	95
5.2.4	基于卡尔曼滤波的姿态估算模型	97
5.3	多传感器坐标系配置技术	98
5.3.1	配准流程	99
5.3.2	配准算法介绍	100
5.4	精度优化算法	104
5.4.1	过阻尼过滤器	104
5.4.2	多空间变换技术	106
5.5	本章小结	108
6	基于非受限触觉反馈的交互方法	109
6.1	基于空气动力的触觉反馈方法	109

6.2 基于热感应的触觉反馈方法	112
6.3 基于电磁理论的触觉反馈方法	115
6.4 本章小结	121
7 智能人机交互系统的实践研究	122
7.1 面向多自由度机器人的人机交互界面设计	122
7.1.1 手势交互界面	122
7.1.2 视频反馈	123
7.1.3 虚拟场景	124
7.2 面向多自由度机器人的多智能体协作模型设计	124
7.2.1 多智能体协作系统结构	124
7.2.2 智能体设计	125
7.3 基于三维手势人机交互技术实现	127
7.3.1 有标记人机交互	127
7.3.2 穿戴式人机交互	129
7.3.3 混合传感器人机交互	132
7.3.4 基于电磁理论的力反馈人机交互	134
7.3.5 单人多机器人的人机交互	136
7.4 本章小结	137
参考文献	138

1 绪论

从一开始笨重的计算机到如今可以穿戴在身上的移动设备，人机交互界面发生了一次次的变革。人机交互界面是用户和计算机之间的通信接口，交互系统通常包括输入和输出两部分：计算机通过输出或显示设备向用户展示大量的信息和提示；用户通过输入设备向计算机输入相关指令。人机交互经过多年的发展和研究，已成为一门以计算机技术为主的跨多学科的技术，其研究涉及计算机科学、认知科学、心理学、机械学等领域^[1]。现如今，人机交互技术已成为从硬件转移到软件之后一个重要的、独立的研究领域，也是计算机行业的一大研究热点。

1.1 人机交互作用

从1946年第一台数字计算机诞生以来，计算机的应用越来越广泛。计算机已经渗透到我们生活工作的方方面面，它代替我们人类从事很多重复性的工作，处理大型的数据，在很多方面的表现甚至已经超越了人类。但是由于人工智能技术发展的程度和任务的复杂性，现今计算机的很多任务仍需人类的参与和协助，所以在很多情况下，人类都需要与计算机进行沟通。这种人类和计算机交流、通信的技术就称为人机交互技术，实现人机交互通信的软硬件系统就是人机交互系统^[2]。

在人机交互系统中，人仍然是一个关键的因素，因为人类有着计算机无法比拟的应急能力和决策能力，能够在计算机无法做出决策或者正常工作时协助计算机完成任务。

尽管人类有着计算机等机器无法企及的优点，但是计算机等机器也有着很明显的长处^[3]。首先，在一些极端危险的环境下，如深海、核辐射、太空、有毒气体等环境，机器人能代替人类完成危险的任务。第二，机器人可以完成重复率高的工作以解放人类的劳动力，从而提高生产效率。在很多的工厂车间中，使用自动化的调度和监控程序会大大地提高车间生产的效率，而且可以避免员工失职所导致的事故。第三，计算机强大的计算能力也使得它能比人类更快速高效地完成任务。

人机交互技术的诞生将人类和计算机的优势互补起来，人类的协助使得计

算机可以更好更快速地完成任任务。所以人机交互领域是一个越来越热门的研究领域,各种人机交互技术的研究层出不穷。

早期的计算机采用的是面板开关和穿孔纸带的交互技术,人通过开关、指示灯和卡片上的孔来表示数据和指令,通过专门的卡片输入机或者纸带输入机进行输入;计算机则通过控制面板显示寄存器的内容,通过打印机输出计算结果,操作非常不方便。之后,出现了交互命令语言和程序设计语言,它大大改善了上一代人机交互技术,提高了系统的效率。接着图形化界面和鼠标的产生使得计算机进入了千家万户,用户无需花费很长的时间学习如何与计算机进行交互,图形化界面以其直观形象的特点让用户可以在短时间内熟练地与计算机打交道,人机交互技术也从此进入了迅速发展的时期。后来,计算机硬件的突破成就使得人机交互技术的研究领域进入了新的时代,触摸屏、重力感应器、磁力感应器、体感设备等让用户与计算机交互时获得前所未有的体验,用户只要点击屏幕或者晃动手臂,计算机就会按照他们的要求工作,这大大提高了人机交互的效率,改善了用户体验。

从信息转换的角度来看,人机交互系统实际上就是实现“用户的感知”和“计算机处理空间”之间的信息转换。用户的感知是用户意识和多维信息的集合,所以通常是模糊的和不清的,而计算机处理空间则是一个规则的、精确的信息空间,两者的差异性使得之间的映射变得十分复杂。如果单纯使用规则精确的语言描述用户的感知,则会提高用户在完成其所需的交互行为所要付出的“认知努力”。因此,人机交互的重要任务就是选择一种合适的映射方式,将用户的感知映射到计算机的处理空间上^[4]。

1.2 人机交互的发展

由于人机交互界面直接决定了计算机软件的易用性和用户使用的体验,所以伴随着计算机发展,人机交互发展也十分迅速。在人机交互技术发展的历程中,人与计算机交互的方式发生了很多次的变革。人机交互界面的形式经历了手工操控、作业控制语言,到图形用户界面、多媒体多通道人机界面^[5]。其中,图形用户界面的出现更是推动了计算机的普及和发展。

1. 早期计算机的发展^[6]

1946年,世界上第一台数字计算机ENIAC诞生,它被用来计算弹道方程的积分。计算机数据和程序的输入主要由专门的纸带输入机和卡片输入机完成,用户通过计算机的开关、指示灯和控制面板对计算机程序进行调试和管理,计算机通过控制面板显示寄存器的内容,然后通过打印机在纸上输出计算结果^[7]。

在这个时期,计算机并未进行大批量生产,大多数的计算机只是在军事部门或者研究所中研制使用,而且通常用户就是设计者本身,所以彼时的人机交互方式过于笨拙和非人性化。

2. 作业控制语言和交互命令语言

随着计算机处理水平的提高,先前的卡片或者纸带式手工操控明显拖慢了计算机的处理速度,这种类似批生产的系统无法满足人们的需求。符号汇编语言、子程序库以及输入/输出控制程序的相继出现,提供了新型控制计算机的方式,渐渐改善了人们使用计算机的效率。

最早的一批程序设计语言,如 BASIC, COBOL, FORTRAN 等,为用户与计算机之间的通信提供了有力的工具^[8]。人们可以把各种问题以形式化的方法成批地向计算机输入,经过计算机处理后进行返回,这个过程不能被打断,即处理过程中用户不能干预,但是系统的吞吐率有了明显的提高。命令行界面是最早出现的人机界面。在这种界面中,机器等待用户的命令做出被动的反应,人通过键盘等设备输入命令信息,计算机处理后做出一定反应并以界面的形式返回结果,界面输出为静态的字符^[9-10]。在命令行界面中,界面和应用还没有分开。

在这个阶段,人机交互界面的自然性和效率都很低,由于每次操作都需要一些命令作为输入,操作员需要熟知计算机系统能接受的命令,否则无法使用,所以易用性比较差,但是命令行界面作为第一代人机交互界面,已经开始考虑如何方便用户与计算机进行通信。

3. 图形用户界面

1963 年,美国麻省理工学院的 Sketchpad 绘图系统引入了菜单、窗口、图标和存储图及存储图形符号的数据结构,开发了使用键盘和光笔进行绘制的交互技术,并提出了许多原理和图形化技术。这可以被认为是人机交互系统图形化界面时代的开端。不久后,美国加州一家公司发明了鼠标,从此人机界面进入了 GUI 时期,GUI 技术成为计算机技术中一个快速发展的分支。

1984 年,美国苹果公司推出了第一个通用的多窗口系统,它所提供的对话框、下拉菜单以及一些优秀的绘图软件都对用户界面设计起了重要的作用,这引发了人机交互界面的一次重大变革。1986 年,美国微软公司研制开发了面向普通用户的个人计算机系统——Windows 窗口系统,大大推动了计算机的普及。这种以窗口(windows)、图标(icon)、菜单(menu)和指向装置(pointing device)为基础的 WIMP 界面可看作是第二代人机交互方式^[11]。

对比第一代命令行人机交互界面,第二代 WIMP 界面的交互方式更加自然和高效。计算机的输出不再是单一的字符,人机间的通信带宽也大大提高了^[12]。但无论是第一代还是第二代人机交互方式,都属于单通道人机界面,用户使用鼠标和键盘在一维空间(命令行)或者二维空间(图形化界面)中完成人机交互,这种交互是一种静态的方式,难以自然地完在三维空间中的动态交互。

4. 多媒体多通道的智能用户界面

近些年来,由于计算机的输入/输出设备和多种技术的发展,如模式识别、

虚拟现实技术、新的传感技术、多媒体技术和可视化技术等,人们对人机交互技术提出了更高效、更直接的要求。新的人机交互技术利用人的多种感觉通道和动作通道(如声音、姿态、表情等输入),用户可以使用非精确的多通道并行的方式与计算机系统进行交互,提高人机交互的交互性和自然性。

多媒体界面是传统的 WIMP 界面在表现形式上做了改进的成果,它加入了多种媒体,可看作是 WIMP 界面的另一种风格。虽然它大大提高了人机交互中的通信带宽,但是人机交互效率依然很低。

多通道界面则是新的智能接口设备和技术不断出现的产物,它反映了用户想“自然地”与计算机进行交互的想法。它基于智能接口技术、手势识别、语音识别等技术,使得用户可以使用多种感觉通道和动作通道自然地与计算机进行交互^[13]。

多通道界面可以看作是人机交互的第三代界面,它允许用户使用自然的交互方式,如语音、手势、表情等。计算机通过对多个输入通道信息的整合,充分了解用户的交互意图。而且多通道界面的输出使用了三维、逼真的虚拟现实技术,使用户可以与计算机进行自然的交互。

1.3 人机交互元素

人机交互系统搭建了一座计算机和用户之间的桥梁,用户通过人机交互设备与计算机进行沟通,将自己的需求传达给特定的软件程序,以驱动特定的程序完成相应的操作。所以人机交互的过程主要涉及三个元素:用户、人机交互设备、交互软件^[14]。

虽然不同交互软件的使用目的不相同,但是它们都是通过同一个模式与用户进行交互。比如以一个绘画软件与一个成绩查询系统作对比,它们的功能,甚至是交互设备完全不一致,但是它们所使用的人机交互模式都是一致的。用户使用鼠标或者触摸屏在屏幕上勾勒一条直线,绘图软件便会按照用户的指令在屏幕的相应区域画出一条直线。对于成绩查询系统来说,用户通过键盘和鼠标输入自己的学号和查询的科目,点击查询后程序便根据输入的信息查询数据库,将科目成绩显示在软件界面上。这两个交互软件虽然操作、功能甚至交互设备都不一样,但是它们都有着相同的交互过程,都涉及三个元素。其中,人是人机交互系统的一个重要元素;交互设备是人机交互系统的物质基础;交互软件是人机交互系统的核心。

1. 用户

人机交互系统给用户提供了与计算机沟通交流的途径,其目的就是要让用户更加简便、直观地与计算机进行交流与通信。所以在人机交互系统里面,人是一个重要的元素,一个好的人机交互系统应该充分考虑到用户的需求以及良

好的用户体验。

人的因素与目标用户的各种特征有关。用户有许多类型,有精通计算机操作的专家和程序员,对于他们来说,软件的功能需求比较高,要求系统响应时间短,但是对于软件的易用性要求不是太高;有不擅长操作复杂的计算机软件的普通用户,对于他们来说,要求功能的介绍能够简洁易读,希望软件容易上手,各种需要的功能能够快速找到;还有不经常接触计算机的人群,如老年人等群体,他们希望字体够大,操作简便,最好只要点击一个按键可以实现他们想要的功能。所以设计一个人机交互系统时,需要找出目标用户的行为特征,以满足目标用户的需求。

2. 交互设备

在人机交互系统中,用户通过交互设备向计算机输入数据、指令等信息,计算机通过交互设备向用户输出处理结果、提示信息、出错消息等^[15]。交互设备是人机交互系统的物质基础,有交互设备才有人机交互的可能性。

人机交互设备从产生开始就注重自然直观的人机交互方式,也越来越呈现出多样化。现代的智能交互设备功能越来越强,如触摸屏、磁力感应器、重力感应器,甚至是体感设备,这些都改变了传统人机交互的方式,也使得人机交互从单纯的图形化界面往更多元的方向发展。

3. 交互软件

交互软件是人机交互系统的核心,用户通过交互软件实现与计算机的通信,并驱动计算机进行相应的操作。在人机交互系统中,人机交互界面不仅是主要的组成部分之一,也是衡量人机交互功能质量的一个重要指标^[16]。现如今,大部分的人机交互应用都是依据自身的软件需求和特性开发人机交互界面,因此设计和实现特有的人机交互界面成为一个交互软件重要的工作任务,占据了大部分的工作量。

在计算机的用户界面中,常见的人机交互界面类型有三种:

(1)命令行方式。即用户通过命令行界面向计算机输入指令,计算机收到指令后对其进行分析,并执行相应的操作,再将结果通过命令行界面反馈给用户。

(2)图形化交互界面。图形交互界面是当前人机交互的主流,具有很多成熟的商业化软件。图形界面使用键盘和鼠标作为输入设备,而且广泛使用了事件驱动技术。不过其主要缺点在于图形需要占用较多的屏幕空间,并且难以表达和支持抽象信息。

(3)直接操作。用户根据计算机输出设备显示的要求和提示,通过输入设备向计算机输入信息和指令,计算机收到输入信息和指令后做出相应的处理,输出设备显示的信息也会相应改变。由于这种方式有直观形象的特点,因此得到十分广泛的应用。

1.4 本章小结

计算机的发展,不仅包括计算机本身存储容量的不断扩大、处理速度的飞速提升,也包括计算机与用户之间交互界面的不断改善。计算机的用户界面直接关系到人们的使用体验,有关人机交互的研究一直是一大热点,所以这一领域的技术发展十分迅速。特别是随着硬件价格的不断降低、人的生产效率的提高,用户界面已成为计算机十分关键的一个部分。

人机交互技术是计算机应用领域的核心技术,其以用户为中心,是新一代的人机交互界面。人机交互界面作为人机通信的手段,其效果的优劣直接影响到人机交互软件的功能、用户的体验甚至人们的工作和生活^[17-18]。

回顾人机界面发展的历史,是一个从繁杂到简单的过程。第一代的计算机只是应用在军事等特定领域,而且需要专家才能与计算机进行交互。到了命令行交互界面阶段,计算机可以被更多用户所使用,这时的系统设计开始考虑如何方便用户使用。1963年美国麻省理工学院提出了图形用户界面技术,从此人机界面进入了GUI时期。1986年美国微软公司推出的Windows窗口环境应用于用户面更广的个人计算机,推动了计算机的普及。发展到如今的多通道人机交互界面,用户可以直接用语音、手势和表情等与计算机系统进行协同合作。

在整个计算机发展历史中,软件设计与开发一直扮演着十分重要的角色,而且其重要性还在逐步增加。在软件的设计与开发过程中,人机交互界面的设计和开发是整个人机研发过程的一个重要组成部分,占40%~60%。另外,由于不同人群对界面的要求不尽相同,人机交互界面也因此成为软件研发过程的一大难题。现如今,人机交互技术已成为世界各国软件研发人员最为关注的研究热点之一。

2 人机特点

人机交互系统是人类与机器有机的结合，所以寻求一种同时将人类和机器各自的优点有机结合，同时弥补双方各自缺点的交互方式是人机交互系统中的一个重要目标。在人机交互技术的研究中，认识了解人的因素、机器的因素和人机之间的区别十分重要，这有利于在人机交互系统的设计中充分发挥人和机各自的优势、弥补各自的缺点^[3]。

2.1 人的因素

人的因素所涉及的领域有很多，包括心理学、生理学、行为科学等，本节将着重讨论人机系统中人的作用以及在感知、决策等方面人的优势与缺点^[3]。

在第二次世界大战时，有些国家研究了许多新式武器和装备，但是设计者忽略了人的因素，因此设计的控制系统很多不适合人类操作，所以在实际应用场合派不上用场，造成了很大的损失。经过这次教训，设计师开始在设计中考虑人的因素，把人作为系统的一部分进行研究和分析。

随着自动化技术的发展，人的作用也在慢慢地发生改变。在人机系统中，人的基本功能有^[2]：①通过人自身的感应系统采集信息；②将收集到的信息进行分析、加工以及处理；③通过人的运动系统完成决策的执行。在现代的人机系统中，人从传统人机交互中的直接参与控制，变成了监督与管理。由于人的作用转化为决策和监督，所以人的行为对于系统的影响更为重要，一个错误的决策可能直接导致一个系统崩溃，人成为越来越重要的因素。而且，只有在人机交互系统的设计阶段充分考虑到人的性能和局限，才能规避风险，并且把人机协作效率最大化。

与机器不同，人的感知是生理分析器工作的结果^[3]。人在感知外界事物时，通常从事物的整体入手，将由多个部分组成的对象当成一个整体。比如当人观察一辆汽车，不会把车子分别感知为轮子和车门等，而是会首先将其视为一辆汽车。而且人类对于外界事物的感知都是不精确的。例如，在没有其他设备的协助下，人类无法直接说出一辆行驶的汽车的速度，只能使用“很快”或“有点慢”等模糊词汇来描述。人类可以同时感知视觉、听觉等多维的感觉信息，然后使用多维信息对事物做出综合的判断。但是，人类的感知结果可能会存在错误，

也就是存在错觉。所以,尽管有时外界的刺激相同,人类却可能做出不同的感知判断^[19]。而且不同的人会由于其自身不同的经历和知识水平,对同一信息有不同的感知结果,也就是说存在个体感知的差异性。

在决策方面,人脑处理信息的能力具有一定的界限。所谓的界限是指大脑皮层不能完全、准确地处理连续接收的各种信息,处理能力有一定的限度。科学实验表明,人很难同时处理两种以上的信息。如果时间充足,人可以准确无误地处理各种信息,不过如果几种信息互相交错时,人就很难处理,甚至会出现处理失误或者是忽略的现象。一方面人类的思维具有很强的创造性、自主性,具有较高的形象思维和抽象思维能力,还具有极强的学习能力;另一方面,人具有生理局限性,在知识的存储能力上有一定的限度,对知识的记忆具有时效性。在一定的程度上,记忆不是很可靠,人类的决策推理有时可能不是很严密。

在任务执行层面,人的动作行为具有灵活性和协调性,且能够适应环境的变化,随机应变能力强。但由于人的心理局限性,人执行任务的准确性常常受到心理因素的影响,人在害怕、紧张的情绪下的动作往往会出现颤抖、变形。

2.2 机器的因素

在人机交互领域,如何利用人类与计算机各自的优点,规避各自的不足,建立一种新型的人机协作关系,是组成高效、可靠性高的人机系统的关键。机器在很多方面有着人类无法企及的优点,但是由于人工智能等技术发展的局限性,机器在很多方面还是无法与人类比拟的^[3]。

在感知方面,机器人的长处在于精确,机器的传感器可以对相关的物理量进行定量检测,并可以根据传感器的精度对物理量做出精度范围内的判断。而且机器对环境有着高度的承受能力,它可以在极端的环境下工作,在一些人类无法到达的地方代替人类工作,比如深海、太空、核辐射环境下的工作,有毒放射性气体环境的监控、排爆工作等。目前,机器传感器的感知范围已大大超越了人类所能感知的范围,许多人类无法感知到的信息,机器的传感器都能准确探测出来,比如人类听不到的声波,看不到的红外线、紫外线等。然而,机器的强感知能力只体现在对单一信息的感知,当其身处复杂环境中时,往往有大量的环境信息需要感知,此时机器就无法像人一样对环境做出综合的感知。为了达到综合感知的目的,机器可以通过多传感器对环境进行探测,然后使用融合算法将多种传感信息合成起来。经过融合集成的多传感器信息可以比较完善、精确地反映环境特征。

在决策方面,机器人领域对机器是否具有智能的这一争论一直没有停止过。从目前来看,机器无法拥有人一样的智能决策能力,但是计算机可以模拟人脑思维的部分功能和结果。随着计算机技术发展,计算机可以模拟出人脑的更多

属性,而且能力会越来越强。目前常见的机器人模拟人类思维的方法有:基于符号机制的人工智能、基于连接机制的人工神经网络和模糊模拟技术等。虽然机器智能还不具有人类智能一样的能力,比如创造性思维和应变能力等,但是机器可以存储大量的知识,且存储的知识还具有永久性,做出决策时具有严密的逻辑性,这些都是机器智能的优势。机器能将人类解决问题时所采用的策略和与该问题有关的大量的实际知识和经验知识,不受个体和时间、空间的限制结合起来,协助人类做出决策。机器智能系统还可以利用已有的知识和经验进行推理和判断,具有智能上的启发性。另外,机器智能系统能够在人类的帮助下不断增长知识,修改原有规则,补充添加新的规则,具有一定的灵活性。除此之外,机器智能系统不受周围环境和人类心理所影响,甚至在一些极端的场合都可以高效地工作。

在执行方面,机器的执行功能有以下的特点:

- (1) 机器执行速度比较快、精度高、功率大,而且具有可扩展性;
- (2) 机器的操作范围广、耐久性强、不会疲劳;
- (3) 机器工作环境的适应性比较强,可以在人类无法工作的环境下代替人类工作;
- (4) 机器具有优秀的运动学、动力学特性;
- (5) 机器执行动作单一、动作柔性小、自由度少、不灵巧。

由此可见,在人机交互协作系统中,在功率大、速度快、精度要求比较高的任务中或者在恶劣的环境中,机器比人类更适合执行相应的任务。当双方联合共同完成一个任务时,通常是人类通过操纵小功率的人机交互系统来操纵机器执行大功率的任务^[4]。在这种主-从结构中,人类操作“主”部分给出任务的轨迹、位置和动作,而机器作为“从”部分,根据人类给出的信息,完成指定的任务。

2.3 人与机器的区别

在感知方面,人感知的信息类型是模糊信息,而机器人感知的是精确信息。人感知的范围很有限,且精度低,有时还会产生错觉。相比之下,机器人不仅感知的范围广、精度高,而且环境适应能力强。在处理信息方面,人对于信息的理解能力比机器高,而且人能在复杂的环境下感知多重环境信息,综合地感知周围环境特征。而对于机器来说,要把收集到的信息融合,需要一些多传感器信息融合算法对采集到的信息进行集成处理。针对在这方面人与机器的不同点,可以在感知层面上对人与机器进行如下的分工^[20]:

- (1) 由于机器的感知精度较高,因此对于精度要求较高的定量信息,如测量物体的速度、物体之间的距离时应使用机器传感器进行测量。