

“十三五”国家重点出版物出版规划项目

普通高等教育“十三五”人工智能与机器人规划教材


# 智能交互技术与应用

马楠 徐歆恺 张欢 编著

The Technology and Application  
of Intelligent Interaction



本书配有课件  
及教学视频

 机械工业出版社  
CHINA MACHINE PRESS

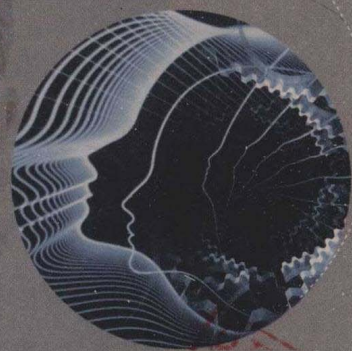
“十三五”国家重点出版物出版规划项目

普通高等教育“十三五”人工智能与机器人规划教材

# 智能交互技术 与应用

马楠 徐歆恺 张欢 编著

The Technology and Application  
of Intelligent Interaction



机械工业出版社  
CHINA MACHINE PRESS

本书深入浅出地探讨和解读了人工智能中的智能交互技术,包括交互原型系统设计、基于移动设备的智能交互软件开发、智能语音交互及交互技术设计与评价方法。第1章主要介绍智能交互技术的基本知识;第2章主要讨论了感知认知与交互技术的结合;第3章讲解了交互设计的基本准则;第4章讲解了交互系统的原型设计,介绍了面向软件应用的最基本的交互设计方法和基本流程;第5章针对移动设备,介绍了App软件交互设计方法;第6章针对当今热门的Android应用程序交互设计,详细解读了该操作系统的特点与架构、应用开发方法及手段,并通过App Inventor工具实现应用开发;第7章介绍了智能交互技术中的热门应用之一——语音交互技术的设计方法,并通过语音识别和合成、对话交互的实验使读者深入掌握智能语音交互技术;第8章介绍了智能交互技术的设计及评价方法;第9章重点介绍了智能交互技术的前沿问题及应用领域。

本书以自动驾驶作为主要应用场景介绍智能交互技术。书中各个章节中均设计了习题,大部分章节配有详实的实验指导。本书可作为智能科学与技术、人工智能计算机科学与技术、软件工程、机器人工程、数字媒体技术等专业本科生、研究生的教材,也可作为从事交互设计、用户界面设计、人工智能应用、移动软件开发等相关技术人员的参考书。

本书配有教学课件和实验素材,请选用本书作教材的老师登录 [www.cmpedu.com](http://www.cmpedu.com) 注册下载,或发邮件至 [jinacmp@163.com](mailto:jnacmp@163.com) 索取。

## 图书在版编目(CIP)数据

智能交互技术与应用/马楠,徐歆恺,张欢编著.

—北京:机械工业出版社,2019.8

普通高等教育“十三五”人工智能与机器人规划教材

“十三五”国家重点出版物出版规划项目

ISBN 978-7-111-63184-2

I. ①智… II. ①马… ②徐… ③张… III. ①人工智能—高等学校—教材 IV. ①TP18

中国版本图书馆CIP数据核字(2019)第140694号

机械工业出版社(北京市百万庄大街22号 邮政编码100037)

策划编辑:吉玲 责任编辑:吉玲 王康

责任校对:王欣 封面设计:鞠杨

责任印制:李昂

北京瑞禾彩色印刷有限公司印刷

2019年8月第1版第1次印刷

184mm×260mm·12.75印张·315千字

标准书号:ISBN 978-7-111-63184-2

定价:49.80元

电话服务

客服电话:010-88361066

010-88379833

010-68326294

封底无防伪标均为盗版

网络服务

机工官网:[www.cmpbook.com](http://www.cmpbook.com)

机工官博:[weibo.com/cmp1952](http://weibo.com/cmp1952)

金书网:[www.golden-book.com](http://www.golden-book.com)

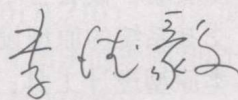
机工教育服务网:[www.cmpedu.com](http://www.cmpedu.com)

## 序

与研究物质和能量的许多科学（如物理学、化学）相比，人工智能是一门更富挑战性的科学，它使智能机器能够胜任一些通常需要人类的智慧才能完成的复杂工作。当前新一代人工智能技术正处在快速发展的黄金期，在生活、医疗、教育、金融、安全、制造等诸多领域，人们越来越多地体验到了它带来的便利和高效。在这类机器进行工作时，如何更好地实现人与机器的交流或信息交换，正是智能交互所研究的问题。

人与人的交互，体现了人的智商和情商；智能机器与人的交互、机器人之间的交互、机器人与人的交互亦然。本书以人工智能中的智能交互技术为核心，系统地讲述了智能交互技术的多样性和综合性，以及在各种各样的环节所涉及的技术及应用。作为人工智能的最普遍、最典型的应用场景，自动驾驶中的智能交互应用在书中被广泛提及。书中设计了大量深入浅出的实验，使读者能够更快地了解并掌握各类智能交互技术，相信有志于从事智能交互技术研究的读者，通过本书的学习一定会有所收获。

人的智能体现在记忆认知、计算认知和交互认知三大能力上，本书仍有进一步提升的空间。如果以交互认知为切入点，勇敢闯入全球人工智能研究的无人区，可望做出更大的贡献！



中国工程院院士 中国人工智能学会理事长

# 前 言

科学技术的进步推动了人工智能技术的飞速发展，人工智能技术已成为当下研究热点，并不断革新我们与机器的互动方式。纵观科学技术发展史，每一次变革都伴随着交互技术的革新，鼠标和键盘打开了PC时代的大门；触摸屏打开了移动互联网时代的大门；语音交互、人脸识别、手势识别等多模态智能交互技术带我们走进全新的AI时代。本书深入浅出地探讨和解读了人工智能中的智能交互技术，包括交互原型系统设计、基于移动设备的智能交互软件开发、智能语音交互及交互技术与评价方法，充分体现了“智能+交互”。

第1章主要介绍智能交互技术的基本知识，包括人机交互的起源与发展，以及各发展阶段的主要技术特征，感知智能在交互中的应用；介绍了智能交互技术的研究内容；另外，智能交互技术涉及多个学科领域，故介绍智能交互技术与相关学科的交叉融合。

第2章主要讨论了感知认知与交互技术的结合。由于人工智能要研究如何让计算机去做那些靠人的智力才能做的工作，模仿人的行为，因此智能交互就要模仿人的基本交互方式。这包括模拟人的感官体验、认知方法、知觉特征、认知过程及交互手段。读者可通过本章较好地了解智能交互领域目前研究的内容。

第3章讲解了交互设计的基本准则。首先从用户体验的角度介绍各种应用交互设计的流程和原则；然后，以桌面系统为例，介绍各种类型的应用特点及交互设计原则及相关技术；最后，通过思维导图实验，读者会尽快理解产品功能的交互设计思路。

第4章讲解了交互系统的原型设计，介绍了面向软件应用的最基本的交互设计方法和基本流程；介绍了各类原型设计工具；以手机App原型设计为实例，以实验的形式指导读者体验此类应用原型的设计思路。

第5章针对移动设备，介绍了App软件交互设计方法。目前移动应用是比较热门、使用率较高的应用形式，它们依托于移动操作系统，与桌面应用有不同的交互方式。本章在介绍不同移动平台的应用交互设计及规范的同时，通过实验读者可以亲身实践此类应用的设计方法及思路。

第6章针对当今热门的Android应用程序交互设计，详细解读了该操作系统的特点与架构、应用开发方法及手段，并通过App Inventor工具实现应用开发。通过本章实验，读者可以了解并掌握当前最热门的网络编程（通信接口开发）及简单的人工智能（人脸检测）编程的开发方法。

第7章介绍了智能交互技术中的热门应用之一——语音交互技术的设计方法，并通过语音识别和合成、对话交互实验，读者可以深入掌握语音交互技术。

第8章介绍了智能交互技术的设计及评价方法。如何从无到有设计一个完整的智能交互

产品，在设计中要考虑哪些因素，如何评估各类交互设计的质量以及对交互界面设计的评价都是本章重点讨论的问题。

第9章重点介绍了智能交互技术的前沿问题及应用领域。当前人工智能快速发展，智能交互技术已经应用到了各个领域，包括无人驾驶交互认知、智能车联网、云机器人等，本章通过多个方面对智能交互技术的应用进行了生动的介绍。

本书以自动驾驶作为主要应用场景介绍智能交互技术。书中各个章节中均设计了习题，大部分章节配有翔实的实验指导，帮助读者通过实验环节加深对相关内容的理解。本书可作为智能科学与技术、人工智能、计算机科学与技术、软件工程、机器人工程、数字媒体技术等专业本科生、研究生的教材，也可作为从事交互设计、用户界面设计、人工智能应用、移动软件开发等相关技术人员的参考书。

本书由北京联合大学机器人学院的马楠、工科中心的徐歆恺和智慧城市学院的张欢共同编写，马楠负责全书的统稿。第1、2章由马楠编写；第3章由徐歆恺、张欢、马楠共同编写；第4章由徐歆恺编写；第5、6章由张欢、徐歆恺、马楠共同编写；第7章由徐歆恺编写；第8章由张欢编写；第9章由马楠编写。在本书的编写过程中，北汽新技术研究院荣辉副院长对书中有关无人驾驶的内容进行了指导，吴修平、李佳洪、陈丽、孙慧荟、穆惠芳、马国栋对部分文字进行了整理，并负责了其他辅助工作。在本书的编写过程中，得到了中国工程院院士、中国人工智能学会理事长、北京联合大学机器人学院院长李德毅，中国人工智能学会智能交互专委会主任、北京航空航天大学计算机学院党委书记王蕴红教授，中国人工智能学会智能驾驶专委会副主任、北京联合大学鲍泓教授的悉心指导和帮助，在此表示最衷心的感谢；机械工业出版社的吉玲编辑在本书的编写过程中也对我们给予了热情的帮助与指导。

本书部分内容得到了国家自然科学基金面上项目“无人车多视视频信息获取与定位关键技术（项目编号：61871038）”、北京市自然科学基金项目“无人车多视目标识别（项目编号：4182022）”和北京联合大学“人才强校优选计划”百杰计划“车路协同环境下人机共驾操控负荷分析和交互机理研究（项目编号：BPHR2017CZ10）”的资助。

感谢为此书的编写无私付出的北京联合大学智能交互团队的老师和同学们。由于智能交互技术仍在不断发展，书中如有不妥之处，恳请广大读者批评指正。

编者电子邮箱：xxtmanan@ buu. edu. cn

编者

# 目 录

## 序

## 前言

<b>第 1 章 人机交互技术的发展</b> .....	<b>1</b>
1.1 人机交互概述 .....	2
1.1.1 人机交互的定义 .....	2
1.1.2 人机交互的起源与发展 .....	2
1.1.3 感知智能在交互技术中的作用 .....	5
1.2 智能交互的研究内容 .....	6
1.2.1 传统人机交互技术研究内容及发展 .....	6
1.2.2 智能交互技术研究内容 .....	7
1.3 智能交互技术与相关学科 .....	9
1.3.1 智能交互技术与计算机科学 .....	9
1.3.2 智能交互技术与软件工程 .....	10
1.3.3 智能交互技术与工业设计 .....	11
1.3.4 智能交互技术与生理学 .....	11
1.3.5 智能交互技术与认知心理学 .....	12
1.4 习题 .....	13
<b>第 2 章 感知认知和交互技术</b> .....	<b>15</b>
2.1 知觉和认知 .....	16
2.1.1 人的感知 .....	16
2.1.2 感官与交互体验 .....	17
2.1.3 知觉的特性 .....	19
2.1.4 认知过程及影响因素 .....	21
2.2 输入/输出设备 .....	22
2.3 交互技术 .....	24
2.3.1 基本交互技术 .....	25
2.3.2 语音交互技术 .....	26
2.3.3 多点触控交互技术 .....	27
2.4 基于感知技术的人车交互 .....	28

2.5	习题	29
<b>第3章</b>	<b>交互设计准则</b>	<b>30</b>
3.1	用户体验的定义	31
3.1.1	交互设计流程	31
3.1.2	交互设计的原则	32
3.1.3	用户体验的应用环境	37
3.2	桌面系统应用的交互界面设计相关类型	37
3.2.1	独占应用	37
3.2.2	轻应用	39
3.2.3	后台应用	40
3.2.4	Web 应用	41
3.3	桌面系统应用界面设计原则	44
3.4	桌面系统应用交互设计技术	46
3.5	实验1: 思维导图与产品功能设计	48
3.5.1	实验目的和类型	48
3.5.2	实验内容	48
3.5.3	实验环境	49
3.5.4	实验步骤	49
3.5.5	实验报告要求	55
3.5.6	实验注意事项	57
3.5.7	考核办法和成绩评定	57
3.6	习题	57
<b>第4章</b>	<b>交互原型设计</b>	<b>58</b>
4.1	交互设计体系	59
4.1.1	需求分析	59
4.1.2	架构与流程设计	59
4.1.3	方案设计	59
4.1.4	方案验证	59
4.1.5	设计跟踪	60
4.1.6	其他	60
4.2	交互原型设计工具简介	60
4.3	实验2: 手机 App 原型开发设计	65
4.3.1	实验目的和类型	65
4.3.2	实验内容	65
4.3.3	实验环境	66
4.3.4	实验步骤	66
4.3.5	实验报告要求	82
4.3.6	实验注意事项	82

4.3.7	考核办法和成绩评定	82
4.4	习题	82
<b>第5章</b>	<b>移动交互设计</b>	<b>83</b>
5.1	移动设备及交互方式	84
5.1.1	移动设备	84
5.1.2	交互方式比较	85
5.2	App 界面设计风格	87
5.2.1	界面配色	88
5.2.2	字体的选择	88
5.3	移动应用平台规范	89
5.3.1	Android 应用设计规范	89
5.3.2	iOS 应用设计规范	92
5.4	实验3: H5 轻应用交互	95
5.4.1	实验目的和类型	95
5.4.2	实验内容	95
5.4.3	实验环境	96
5.4.4	实验步骤	96
5.4.5	实验报告要求	102
5.4.6	实验注意事项	102
5.4.7	思考题	102
5.4.8	考核办法和成绩评定	102
5.5	习题	103
<b>第6章</b>	<b>移动终端的交互设计实战</b>	<b>104</b>
6.1	Android 操作系统	105
6.1.1	Android 系统特点与架构	105
6.1.2	Android 开发工具简介	108
6.2	图形化编程开发 Android 程序	110
6.2.1	App Inventor 开发环境	111
6.2.2	App Inventor 简单操作	111
6.3	实验4: 移动端网络接口调用的实现	117
6.3.1	实验目的和类型	117
6.3.2	实验内容	118
6.3.3	实验仪器、设备	118
6.3.4	实验原理	118
6.3.5	实验步骤及要求	120
6.3.6	实验报告要求	122
6.3.7	实验注意事项	122
6.3.8	思考题	122

6.3.9	考核办法和成绩评定 .....	122
<b>6.4</b>	<b>实验5: 移动端人脸检测实验 .....</b>	<b>123</b>
6.4.1	实验目的和类型 .....	123
6.4.2	实验内容 .....	123
6.4.3	实验仪器、设备 .....	123
6.4.4	实验原理 .....	123
6.4.5	实验步骤及要求 .....	126
6.4.6	实验报告要求 .....	130
6.4.7	实验注意事项 .....	131
6.4.8	思考题 .....	132
6.4.9	考核办法和成绩评定 .....	132
<b>6.5</b>	<b>习题 .....</b>	<b>132</b>
<b>第7章</b>	<b>语音交互设计 .....</b>	<b>133</b>
7.1	语音交互简史 .....	134
7.2	语音交互的优缺点 .....	135
7.2.1	语音交互的优点 .....	135
7.2.2	语音交互的痛点 .....	135
7.3	实验6: 语音识别和语音合成 .....	136
7.3.1	实验目的和类型 .....	136
7.3.2	实验内容 .....	136
7.3.3	实验仪器、设备 .....	136
7.3.4	实验原理 .....	137
7.3.5	实验步骤及要求 .....	137
7.3.6	实验注意事项 .....	144
7.3.7	思考题 .....	144
7.4	对话设计 .....	144
7.5	实验7: 对话交互 .....	146
7.5.1	实验目的和类型 .....	146
7.5.2	实验内容 .....	146
7.5.3	实验仪器、设备 .....	146
7.5.4	实验原理 .....	146
7.5.5	实验步骤及要求 .....	146
7.5.6	实验注意事项 .....	156
7.5.7	思考题 .....	156
7.6	习题 .....	157
<b>第8章</b>	<b>智能交互技术的设计与评价 .....</b>	<b>158</b>
8.1	智能交互技术设计思路 .....	159
8.2	智能交互产品设计的相关因素 .....	161

8.3	智能交互产品设计的质量评价 .....	163
8.3.1	交互需求评估 .....	164
8.3.2	交互设计评价 .....	164
8.3.3	服务任务评价 .....	165
8.3.4	交互性评价 .....	165
8.4	智能交互设计的测试 .....	167
8.5	智能交互界面设计的评价 .....	167
8.6	习题 .....	168
<b>第9章</b>	<b>智能交互技术前沿问题及应用领域 .....</b>	<b>169</b>
9.1	智能交互技术的前沿问题 .....	170
9.1.1	人工智能时代的智能交互技术 .....	170
9.1.2	标准化问题 .....	172
9.1.3	虚拟现实及网络用户界面 .....	174
9.2	无人驾驶中的交互认知 .....	179
9.2.1	基于自然语言交互认知 .....	179
9.2.2	基于多源数据融合的协同共驾模式多通道控制权交互研究简介 .....	180
9.3	智能交互技术在智能网联汽车环境下的应用 .....	182
9.3.1	车联网环境简介 .....	182
9.3.2	车联网环境与智能交互技术应用 .....	184
9.4	基于云机器人平台的智能交互及其应用 .....	185
9.4.1	云机器人平台简介 .....	186
9.4.2	智能交互技术在云机器人平台的应用 .....	187
9.5	习题 .....	188
<b>附录</b>	<b>交互设计快速检查清单 .....</b>	<b>189</b>
	<b>参考文献 .....</b>	<b>191</b>

# 第 1 章

## 人机交互技术的发展

## 1.1 人机交互概述

### 1.1.1 人机交互的定义

人机交互是研究人与计算机的交互，或者可以理解为人与“含有计算机的机器”的交互。系统可以是各种各样的机器，也可以是计算机化的系统和软件。在交互过程中，人通过和计算机界面的互动，产生一系列的输入和输出，然后完成具体的任务。在美国的 21 世纪信息技术计划中，将软件、人机交互、网络、高性能计算列为四大基础研究内容<sup>[1]</sup>。人机交互界面的设计要包含用户对系统的理解（即心智模型），实现系统的可用性和用户友好性。



人机交互（Human-Computer Interaction, HCI, 或 Human-Machine Interaction, HMI）是广受关注的交叉学科领域，涉及了计算机科学、心理学、社会学、人机工程学等多个研究和应用领域，是对人类使用的交互式计算系统进行设计、评估和实现，并对其所涉及的主要现象进行研究的学科，尤其在智能时代要研究人与机器人之间通过视觉、听觉、触觉和嗅觉等实现多模态信息智能交互功能。

### 1.1.2 人机交互的起源与发展

人机交互始终关注计算机系统提供给人类交互使用的设计、评价和实现方法。这包括人通过输入设备给计算机输入信息，计算机经过运算再通过输出设备给人提供信息反馈等内容。它与很多学科都有密切的关系，在机器方面，侧重于计算机图形学、操作系统、编程语言和开发环境等；在人的方面，需要考虑通信理论、工业设计、语言学、社会科学、认知心理学、社会心理学等因素。

人机交互是一个跨学科领域，也是用户界面设计过程中需要着重考虑的因素之一。人机交互对于用户界面的发展起着至关重要的推动作用，其发展过程大致可分为四个阶段<sup>[2]</sup>，见图 1-1。

#### 1. 萌芽期（1959~1969年）

1959年，美国科学家 Brian Shackel 发表了一篇名为“Skin-Drilling: A Method of Diminishing Galvanic Skin-Potentials”的论文，首次提出如何用人机工程学原理帮助用户减轻操作机器所带来的疲劳感。1960年，

美国心理学家和计算机科学家 J. C. R. Licklider 在论文“Man-Computer Symbiosis”中开创性地提出了人机紧密共栖的概念，被视为人机交互的启蒙观点。1969年，第一次人机系统国际会议在英国剑桥大学召开，同年第一份专业杂志《国际人机交互杂志》（International Jour-

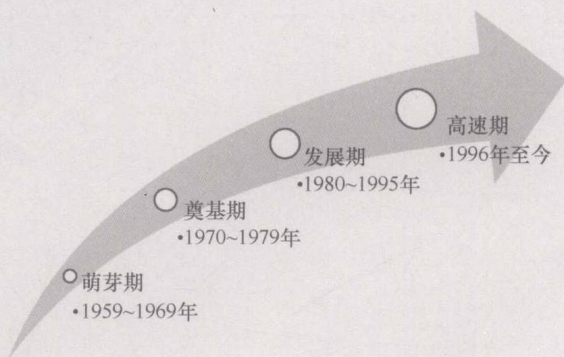


图 1-1 人机交互发展历程

nal Of Human-Computer Interaction) 创刊, 这一年可谓是人机交互发展史的里程碑。

## 2. 奠基期 (1970 ~ 1979 年)

1970 ~ 1973 年, 与计算机相关的人机工程学专著陆续出版, 为人机交互的发展指明了方向。1970 年, 成立了两个 HCI (Human-Computer Interface) 研究中心, 一个是英国的拉夫堡大学的 HUSAT 研究中心 (Loughborough University HUSAT), 另一个是美国施乐帕洛阿尔托研究中心 (Xerox PARC), 见图 1-2。



图 1-2 美国施乐帕洛阿尔托研究中心

## 3. 发展期 (1980 ~ 1995 年)

20 世纪 80 年代初期, 学术界相继出版的专著, 总结了当时最新的人机交互研究成果, 人机交互学科逐渐形成了自己的理论体系和架构。理论体系方面, 人机交互从人机工程学中独立出来, 更加强调认知心理学、行为学以及社会学等人文科学的理论指导。实践方面, 从人机界面 (人机接口) 延伸开来, 强调计算机对人的反馈作用, HCI 中的 I, 也由界面接口 (Interface) 变成了交互 (Interaction)。从词语表面的含义来看, 界面设计体现的是连接程序和用户之间的接口, 它为二者提供消息传递; 而交互设计则包含更广义的内容, 指的是功能、行为和最终的展示形式。

## 4. 高速期 (1996 年至今)

20 世纪 90 年代中期以来, 伴随计算机硬件性能的飞速提升和互联网技术的迅速发展普及, 人机交互研究的重心开始转移到多媒体交互、人机协同交互以及增强现实等方面, 人机交互技术更侧重以人为研究目标和中心。主要特点是基于语音、手写体、姿势、视线跟踪、表情等输入手段的多模态交互, 目的是使人能以声音、动作、表情等自然方式进行智能交互操作。Gartner 公司发布的 2018 年度新兴技术成熟度曲线涉及的部分重要技术是人机交互技术发展的源动力, 见图 1-3。

随着各种体感设备和可穿戴式计算设备的快速涌现和日趋成熟, 促使用户界面设计师们不断探索新的、更自然、更直观、更接近人类行为方式的人机交互界面, 并提出了人机交互的新思想: 自然用户界面逐渐成为主流。它使人机交互的过程更接近于人原来的自然交流形

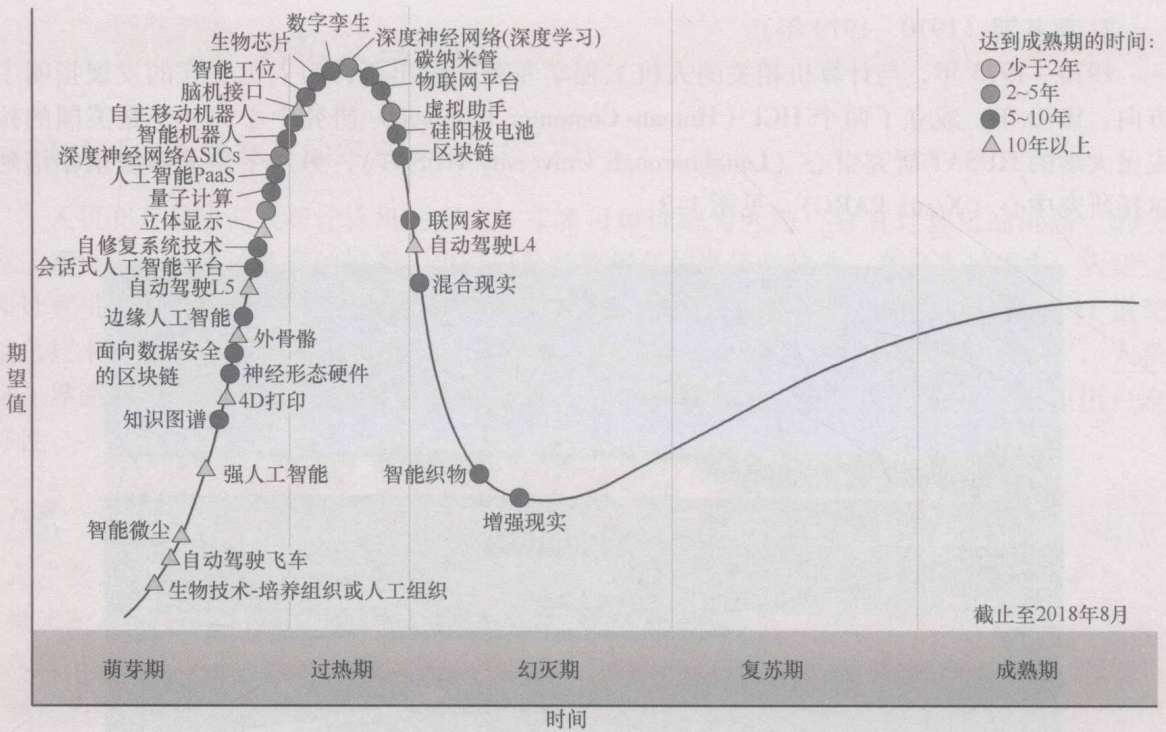


图 1-3 2018 年度新兴技术成熟度曲线<sup>[3]</sup>

式，用户只需要应用日常的自然技能，无须经过专门的适应与学习便可对计算机进行操作，减少了重新学习的认知负担，增加交互效率。丰富的输入和输出软硬件技术，也为应用设计学科和数码交互艺术学科提供了极其丰富的表现方法。

硬件包括：常见的硬件设备包括视觉交互设备、语音交互设备、触觉交互设备、笔式交互设备、虚拟环境中的交互设备，见图 1-4。输入设备有 Kinect、Leap Motion 等，输出设备有 3D 立体显示器，头戴显示器（如 Oculus Rift）等，输入输出设备有 Google Glass 等。

软件包括：界面部分有 3D 立体操作界面，交互方式有肢体动作、表情、语音等。基于




图 1-4 智能交互设备

大数据可以分析用户所处的环境，用户的爱好、习惯、心理、历史、社会关系，语境等。

### 1.1.3 感知智能在交互技术中的作用

人机交互技术与计算机始终相伴发展，随着 CPU、GPU 的运算能力日趋强大，网络和通信技术的快速发展，人工智能被进一步推动，显示技术的重大突破都为人机交互提供了新的起点与高度。近几年，在科技与需求的双轮驱动下，人机交互实现了多次革命，发展突飞猛进。如今，人机交互设备市场规模增长前景广阔，应用领域多种多样，市场引领者围绕多模式、个性化的技术进行设计认证和设备交互。触摸、对话、手势等，交互方式为我们开启一个全新的世界。

人工智能的飞速发展正在为人机交互带来巨大的推动作用，尤其在智能化信息处理方面，促进没有触摸屏、没有键盘的智能化时代交互技术飞速发展。在智能机器人、智能家居迅速普及时代，人机交互方式多种多样，比如手势识别、图像识别、体感识别等，都是未来智能交互技术发展的方向，人工智能正在从感知智能向认知智能发展。

 **感知智能**：利用图像识别、语音识别或者其他识别技术，把物理信号转化成数字信号，基于这些数字信号进行分析、判断、推理、规划、决策。例如，自动驾驶将感知和认知结合在一起，通过激光雷达、摄像头、超声波等各种传感器捕捉信号到数字世界，再通过相应的计算机算法根据当前车况路况实时做出规划决策。

语音交互是最自然的交互方式，在智能时代仍然是最重要的人机交互手段之一。基于智能语音的自然交互核心技术，首先要求能够完成语音识别，其次是以语境为基础的语音理解，直至高自然度语言的生成。满足人和设备之间自然交互的技术，应该是一种计算的人机交互框架，不光有字面的含义，还要有物理语境，因为不同的时间、地点、场所说的同样文字代表的含义是不一样的，结合上下文、人类的物理世界，构建出一个语境为中心的交互形态，才能实现真正的像人和人的自然对话。此外，还需要加入对个性化、一致化、情感化的考虑。相关研究还有很长的路要走，人工智能需要解决一系列商业化的落地问题。软硬一体化的云解决方案是一种解决思路，即使没有很好地解决认知的问题，自然语音技术同样能够为产品创造新的价值。不过，语音识别通过芯片的方式植入到 IoT（物联网）设备，还需要解决很多技术问题，例如远距离的多次声波反射和衰减，低功耗、低成本，高温、高湿环境的稳定性等。

微软小冰是聊天机器人的典型代表，见图 1-5。聊天机器人的结构，包括问话理解、回复（答案或

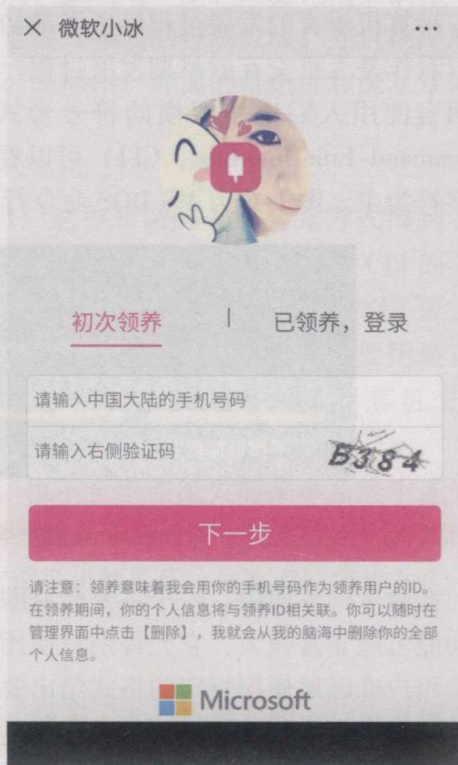


图 1-5 微软小冰

聊天)、根据上下文保证用户个性化信息以及语言风格一致性,最重要的是闲聊、信息问答、完成任务三个层面的引擎。闲聊:要能够满足多轮对话的一致性、个体化,基于搜索的回答需要针对特定主题提前建立聊天智库,实现多模型对话(包括图像理解)。信息问答:提高搜索引擎的水平,将搜索引擎的风格变换成聊天的风格,为每个知识源是建立问答。完成任务:不同任务的对话过程,把重要的信息通过对话设立出来,涉及信息抽取,以及文本之间的相互匹配。

此外,人机交互的应用都离不开深度学习或者是语音识别等人工智能核心技术的突破。百度深度学习实验室(IDL)希望将人工智能核心技术做到统治级别,通过深度学习技术,不仅要做好语音识别,还要实现图像识别基本技术(图像搜索、OCR、人脸识别)、细粒度图像识别(如菜品识别)、视频分析、AR、医学图像分析等方面的突破。

## 1.2 智能交互的研究内容

### 1.2.1 传统人机交互技术研究内容及发展

人机交互研究的具体表述方法是在实际生活场景中的应用,从最初的命令行界面,到图形用户界面,到当下的自然用户界面,随着语音交互、视觉图像交互、动作交互、脑电波交互等多模态人机交互技术的逐步发展和成熟,更加智能的人机交互方式将会深层次地改变我们日常生活的应用场景;一场互联网的主流终端模式和服务内容的竞争也在同步进行。

#### 1. 命令行界面交互阶段

计算机语言的发展过程经历了最初的机器语言、汇编语言,直至高级语言。这个过程也可以看作是人机交互的早期发展过程。机器语言和汇编语言对使用者提出了较高的要求,高级语言使用人们比较习惯的符号形式描述计算过程,降低了对人的要求。命令行界面(Command-Line Interface, CLI)可以看作第一代人机界面,输入信息为数据和命令信息,以字符为主。图1-6为MS DOS命令行被解析为命令和参数。



图1-6 一个MS DOS命令行,可被解析为命令和参数

在命令行界面中,一个主要研究问题是如何为各种命令制定恰当的名称。某些命令语言的功能可能非常强大,它允许用户使用大量修饰符和参数来构造非常复杂的命令序列。它们要求用户准确地使用规定的格式给出要完成的命令,且不能原谅用户可能犯下的任何形式的输入错误。命令行界面最大的弊端就是迫使用户不得不在没有多少计算机帮助的前提下,牢记复杂的命令和格式,这大大增加了用户的记忆负担,从而使大量入门者望而却步。尽管支持命令名称的缩写在一定程度上减轻了用户的使用负担,但并没有从根本上解决这一问题。