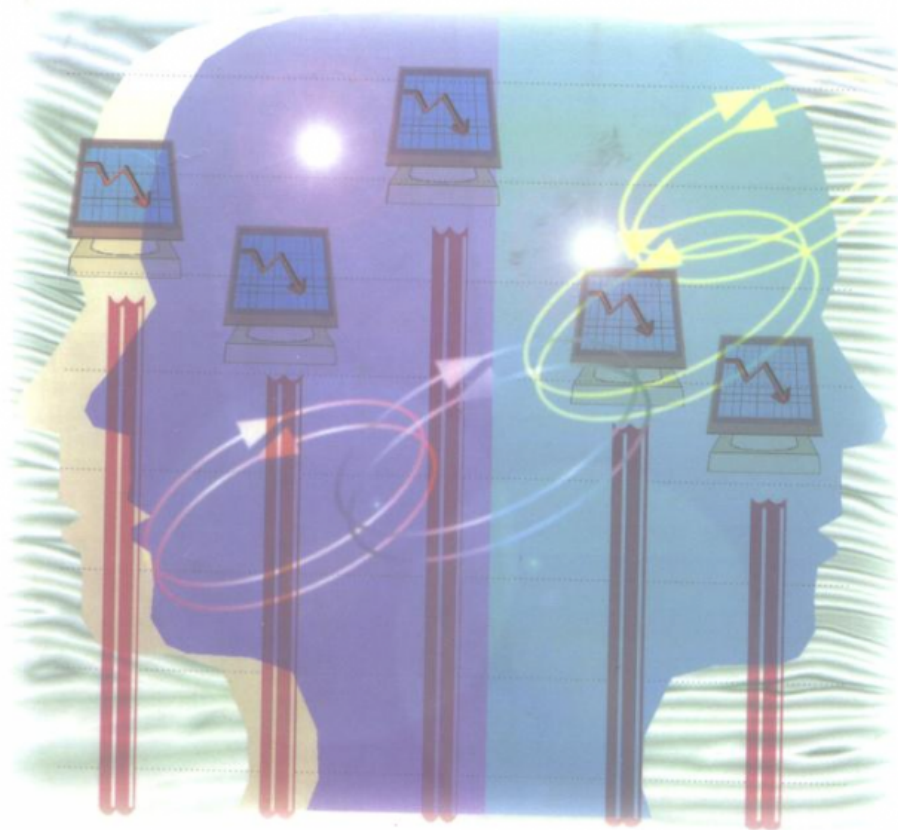


董士海 王 坚 戴国忠 等 著

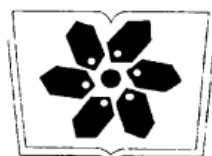
# 人机交互和 多通道用户界面



科学出版社

TP334.7  
D74

452330



中国科学院科学出版基金资助出版

# 人机交互和多通道用户界面

董士海 王 坚 戴国忠 等 著

科学出版社

1999

75185/26

## 内 容 简 介

本书依托国家自然科学基金重点项目“多通道用户界面研究”，详细介绍人机交互技术、交互设备与交互系统设计方法的 90 年代前沿进展，着重介绍了国内外在多通道用户界面方面的相关研究成果及实例，对下一代人机交互界面进行了展望，同时还提供了丰富的参考资料目录。

本书以人机界面及其相关领域的研究人员（包括研究生和高年级本科生）以及计算机界各层次工作人员为主要读者对象。

### 图书在版编目 (CIP) 数据

人机交互和多通道用户界面/董士海等著. - 北京: 科学出版社, 1999. 7

ISBN 7-03-007442-4

I. 人… II. 董… III. 用户接口 IV. TP334.7

中国版本图书馆 CIP 数据核字 (1999) 第 08786 号

科学出版社 出版

北京东黄城根北街 16 号  
邮政编码: 100717

新蕾印刷厂印刷

新华书店北京发行所发行 各地新华书店经售

\*

|                  |                   |
|------------------|-------------------|
| 1999 年 8 月 第 一 版 | 开本: 787×1092 1/16 |
| 1999 年 8 月 第一次印刷 | 印张: 16 1/4        |
| 印数: 1—3 000      | 字数: 369 000       |

定价: 26.00 元

(如有印装质量问题, 我社负责调换(环伟))

## 前 言

计算机人机界面是指计算机和它的使用者之间的对话接口,是计算机系统的重要组成部分。计算机的发展历史,不仅是计算机本身的处理速度和存储容量飞速提高的历史,也是不断改进计算机人机界面的历史。图形用户界面的广泛流行是当今计算机技术的重大成就之一,它极大地方便了非专业用户的使用,人们不再需要死记硬背大量命令,可通过窗口、菜单进行方便的操作。可以看到,人机界面已经从要求人去适应复杂的计算机,发展为计算机不断地适应人的方便使用,即以“用户”为中心的新阶段。

人机界面的重要性在于它极大地影响了最终用户的使用,影响了计算机的推广应用,以至于影响人们的工作和生活。由于开发工作量极大,加上不同人群对界面的要求不全相同,人机界面已成为计算机研制中的一个最困难部分。当前,Internet 发展异常迅猛,虚拟现实、科学计算可视化、多媒体技术对人机界面提出了更高的要求。面临新的挑战,人机界面将如何发展?图形用户界面会被新的界面淘汰吗?什么是多通道人机界面?如何设计多通道人机界面?新一代人机界面将是什么样?这些都是很多人感兴趣的问题。

人机交互领域是一个跨学科的领域,近年来它在欧美的发展迅速,相比之下,国内的研究水平较低,对它的重视程度也不够,在整个计算机研究与产业中的地位与其重要性很不相称。最近几年,这一情况开始有所改观,无论是在国家支持的科研项目(自然科学基金、863 计划、“九五”计划),还是在出版领域,人机交互逐渐成为人们关注的热点,而一本好书无疑会起到非常好的促进作用。为国内人机交互研究与应用推波助澜,正是编写本书的初衷。

本书是一本学术著作。它将具体介绍多种新的交互技术与设备,多通道人机交互系统的模型、设计问题、评估方法及开发环境,同时通过实例说明这些技术与方法的具体实现与运用。作为一本全面介绍人机交互的学术著作,它可以作为研究者的参考资料,也可作为人机交互或界面设计课程或培训班的教材。在科学技术发展的历史上,课程的开设与教材的出版经常作为一门学科正式确立的标志。

与此同时,本书一个同样重要的目标是使那些过去对人机交互不够了解的读者,特别是那些从事着与人机交互关系非常密切的工作(包括人工智能、模式识别、多媒体、因特网等等)的读者和研究者,更加清晰地了解当前人机交互的发展趋向,进一步把人机界面作为应用切入点,使之能够以新的视角来看待自己的工作,并将它在计算机系统和相应的市场中重新定位。这是一个科学技术转化为生产力的重要领域。此外,人机交互的发展,技术与设备的成熟必然意味着巨大的市场,引导企业决策人员在考虑自己的产品战略时,更加重视人机界面这一渗透各个产品的因素也是本书的意义所在。

本书旨在通过综述 90 年代中后期人机界面领域的最新研究进展,总结国家自然科学基金重点项目“多通道用户界面研究”的成果,深入讨论多通道人机界面模型、设计方法、评估方法、开发环境和应用实例等方面。本书试图从计算机科学、心理学等多学科,完善人机界面的可用性,从技术、应用等多个视角来介绍人机界面。

本书共分六个部分，十二章。第一部分是引言，介绍人机交互的发展历史、趋势及与虚拟现实技术的关系；第二部分介绍新的交互技术和设备；第三部分详细综述了近年来国际上人机交互，尤其是多通道用户界面的研究情况；第四部分是本书的重点，详细介绍了人机交互，尤其是多通道用户界面的理论和设计方法，多通道整合模型和算法，界面评估和可用性测试，软件结构和开发环境等；第五部分介绍了多通道用户界面的研究实例；第四、五部分的内容，总结了国家自然科学基金重点项目“多通道用户界面研究”（1995. 1~1997. 12）的主要进展；第六部分是对人机交互及用户界面的展望。附录中给出了详细的参考文献和因特网的网址，国内有关的书籍，国外的主要专业杂志和标准等，以供读者参考。

本书的材料来源主要包括两个方面。一是对国内外研究进展的跟踪与综述，具体来源包括因特网（WWW 和 FTP 网址）、各种相关中外文书籍、国外研究机构的技术报告、国际国内期刊和国际国内会议论文集等。二是国家自然科学基金重点项目“多通道用户界面研究”的部分总结，具体来源包括以此项目为背景发表的论文、未发表的技术报告、系统分析设计文档和与项目相关的讨论记录等等。

本书具有如下特点：

(1) 本书在消理解 90 年代国际国内人机界面的研究成果基础上，及时地反映了当前该领域中十分活跃的研究进展，具有突出的前沿性。

(2) 本书的讨论覆盖了交互技术与设备、系统设计方法与评估、交互系统实例分析等人机界面研究的各个方面，“引言”与“展望”分别讨论了人机界面发展的历史趋势和动人前景，具有相当的完整性。

(3) 本书在充分介绍国际研究进展的同时，以国家自然科学基金重点项目“多通道用户界面研究”为背景，介绍了国内研究成果，具有创新性。

(4) 本书为研究者提供了丰富的国内外技术与项目的参考资料，包括一个相关研究在因特网上的较全面的索引。

本书主要以人机界面及其各相关领域的研究人员（包括研究生和高年级本科生）为主要读者对象；同时兼顾普及性，适合计算机界多层次工作人员阅读，特别是作为企业决策、产品导向和研究支持等方面人员的参考。

1997 年本书被批准列为中国科学院科学出版基金资助项目，科学出版基金专家委员会对编著提纲提出了宝贵的意见。1997 年 2 月及 1998 年 6 月，我国著名的计算机专家：唐泽圣教授、陈由迪教授、赵沁平教授、高文教授、潘云鹤院士、吴立德教授、宣国荣教授、张福炎教授、怀进鹏教授等对国家自然科学基金“多通道用户界面研究”重点项目分别进行了评审或验收，提出了许多宝贵的意见。本书的出版始终得到国家自然科学基金委员会信息科学部的支持，在此谨向国家自然科学基金委员会、中国科学院科学出版基金委员会及上述所有的专家们表示衷心的感谢。

本书是北京大学计算机科学与技术系图形研究室、杭州大学工业心理学国家重点实验室和中国科学院软件研究所 MUIR 项目组师生共同劳动的结晶，特别是罗军、陈敏、孙宏晖、林应明、肖斌、蒋宇全、李茂贞、张高、关志伟、吕晨阳、胡睿琦、沈模卫、蒋成高、方志刚、符德江、乔永庆等同志做了大量的工作，没有他们的努力，本书是不可能及时完成的。被引用的国内研究成果只是部分杰出代表。在此谨向为本书作出贡献的所有同

志及被引用资料的作者致谢。

全书由我们统稿、修改和审定。由于时间仓促，作者水平有限，书中欠妥和纰漏之处恐难避免，恳请读者和同行不吝赐教。

董士海 王 坚 戴国忠

1998年8月

# 目 录

## 前 言

### 第一部分 引 言

|                        |        |
|------------------------|--------|
| 第一章 人机界面的发展历史 .....    | ( 1 )  |
| 1.1 引言 .....           | ( 1 )  |
| 1.2 人机界面的发展历史 .....    | ( 1 )  |
| 第二章 人机交互的发展趋势 .....    | ( 8 )  |
| 2.1 多媒体技术的支持 .....     | ( 8 )  |
| 2.2 采用多通道进行交互 .....    | ( 12 ) |
| 2.3 自然语言接口 .....       | ( 15 ) |
| 2.4 计算机支持的协同工作 .....   | ( 15 ) |
| 2.5 三维人机交互技术 .....     | ( 19 ) |
| 第三章 多通道人机交互与虚拟现实 ..... | ( 22 ) |
| 3.1 多通道人机交互 .....      | ( 22 ) |
| 3.2 虚拟现实技术 .....       | ( 25 ) |

### 第二部分 新交互技术

|                     |        |
|---------------------|--------|
| 第四章 交互设备与关键技术 ..... | ( 28 ) |
| 4.1 人机交互技术 .....    | ( 28 ) |
| 4.2 眼动跟踪 .....      | ( 32 ) |
| 4.3 姿势识别 .....      | ( 34 ) |
| 4.4 三维输入 .....      | ( 35 ) |
| 4.5 语音识别 .....      | ( 36 ) |
| 4.6 表情识别 .....      | ( 41 ) |
| 4.7 自然语言理解 .....    | ( 42 ) |
| 4.8 手写识别 .....      | ( 45 ) |
| 第五章 全息图像和听觉界面 ..... | ( 53 ) |
| 5.1 全息图像 .....      | ( 53 ) |
| 5.2 听觉界面 .....      | ( 59 ) |

### 第三部分 国外研究概况

|                      |        |
|----------------------|--------|
| 第六章 国外研究实例 .....     | ( 64 ) |
| 6.1 国际会议 .....       | ( 64 ) |
| 6.2 北美的多通道界面研究 ..... | ( 66 ) |
| 6.3 欧洲的多通道界面研究 ..... | ( 83 ) |

## 第四部分 理论与方法

|                                          |         |
|------------------------------------------|---------|
| <b>第七章 多通道用户界面模型及描述方法</b> .....          | ( 92 )  |
| 7.1 概述 .....                             | ( 92 )  |
| 7.2 交互原语 .....                           | ( 94 )  |
| 7.3 基于自然交互方式的多通道用户界面模型 (VisualMan) ..... | ( 96 )  |
| 7.4 基于事件-目标的多通道用户结构模型 .....              | ( 96 )  |
| 7.5 层次化的多通道用户界面描述方法 (HMIS) .....         | ( 103 ) |
| 7.6 分布式多通道用户界面的模型及原型 .....               | ( 109 ) |
| <b>第八章 多通道整合模型和算法研究</b> .....            | ( 113 ) |
| 8.1 多通道整合 .....                          | ( 113 ) |
| 8.2 面向任务的整合模型 (ATOM) .....               | ( 119 ) |
| 8.3 多通道分层整合模型和算法 .....                   | ( 127 ) |
| 8.4 基于概率模型的指称整合模型 .....                  | ( 131 ) |
| 8.5 基于模糊识别模型的视线交互整合模型 .....              | ( 138 ) |
| <b>第九章 用户界面评价与可用性测试</b> .....            | ( 145 ) |
| 9.1 用户界面评估的意义 .....                      | ( 146 ) |
| 9.2 多通道界面评估机制 .....                      | ( 147 ) |
| 9.3 界面表现评估 .....                         | ( 147 ) |
| 9.4 多通道界面人机交互评估 .....                    | ( 152 ) |
| 9.5 CAD 评估实例 .....                       | ( 155 ) |
| 9.6 多通道用户界面的可用性测试 .....                  | ( 156 ) |
| 9.7 界面评估清单：一种有效的可用性测试方法 .....            | ( 159 ) |
| <b>第十章 多通道用户界面的软件结构和开发环境</b> .....       | ( 167 ) |
| 10.1 软件结构 .....                          | ( 167 ) |
| 10.2 基于智能结构模型 .....                      | ( 170 ) |
| 10.3 基于智能体的多通道用户界面开发环境 MMID-KIT .....    | ( 173 ) |
| 10.4 采用多通道界面的 PCVRS 软件平台 .....           | ( 184 ) |
| 10.5 基于事件-目标的多通道用户界面的软件实现 .....          | ( 190 ) |

## 第五部分 国内研究实例

|                                 |         |
|---------------------------------|---------|
| <b>第十一章 多通道用户界面的应用与实例</b> ..... | ( 203 ) |
| 11.1 引言 .....                   | ( 203 ) |
| 11.2 虚拟座舱实例 .....               | ( 203 ) |
| 11.3 草图 CAD 实例 .....            | ( 206 ) |
| 11.4 VRoom 实例 .....             | ( 212 ) |
| 11.5 其他 .....                   | ( 215 ) |

## 第六部分 展望

|                                                 |         |
|-------------------------------------------------|---------|
| <b>第十二章 下一代人机界面展望——基于智能体虚拟现实化多通道因特网界面</b> ..... | ( 221 ) |
| 12.1 引言 .....                                   | ( 221 ) |
| 12.2 一种新的用户界面的范式——VIR .....                     | ( 222 ) |

|                        |       |
|------------------------|-------|
| 12.3 自然人机交互 .....      | (226) |
| 12.4 网络用户界面 .....      | (230) |
| 12.5 关于下一代界面的讨论 .....  | (232) |
| <b>附录</b> .....        | (237) |
| A. Internet 有关网址 ..... | (237) |
| B. 国内有关书籍 .....        | (239) |
| C. 国外有关杂志及标准 .....     | (240) |
| <b>参考文献</b> .....      | (241) |

# 第一部分 引言

## 第一章 人机界面的发展历史

### 1.1 引言

人机交互(Human-Computer Interaction)是研究人、计算机以及它们相互影响的技术。人机界面是计算机与人(使用者)之间通信和对话的接口,是计算机系统的重要组成部分。计算机的发展,推动着人机接口技术和人机界面的发展。从计算机早期的面板开关、显示灯和穿孔纸带等交互装置,发展到今天的视线跟踪、语音识别、手势输入、感觉反馈等具有多种感知能力的交互装置。人机界面经历了手工操作、命令语言和图形用户界面的三个阶段。图形用户界面是当前广泛流行的人机界面,它的发展和应用是计算机技术重大成就之一。人机界面影响到最终用户,影响到计算机的推广应用,甚至影响到人们的工作和生活。

近 20 年来计算机硬件技术以难以想象的速度发展,百亿次的运算速度、海量的存储容量以及智能的人机接口设备。另一方面随着制造工艺的发展,计算机的体积越来越小。其应用也越来越广。以虚拟现实为代表的计算机系统的拟人化和以掌上计算机为代表的计算机的微型化和随身化,将是计算机的两个重要的应用趋势。人机接口技术是适应这种趋势的瓶颈技术。人机结合以人为主,将是未来计算机系统的特点,实现人机高效合作将是新一代人机界面的主要目的。但是,我们目前仍然以键盘和鼠标为输入工具,输入输出严重不平衡;而新的智能接口设备不断出现,但是人们却不知如何使用它们,不能充分利用人的认知资源,人们长期追求的以自然的方法和计算机交流的理想仍然无法实现,因而许多新的计算机应用技术无法普及;用户界面开发周期太长,缺乏对用户界面分析、设计、评估和维护的科学方法和高级工具。

多通道用户界面就是在以上背景下发展起来的重要人机接口技术。它基于智能接口技术,充分利用人的多种感觉通道和运动通道,以并行、非精确方式与计算机系统进行交互,旨在提高人机交互的自然性和高效性。

人机通信问题将是 21 世纪信息领域中必须解决的重大课题,人机接口技术是计算机应用的核心技术,以用户为中心是下一代人机界面的设计思想,多通道用户界面是实现自然、高效的人机交互的关键技术。

### 1.2 人机界面的发展历史

#### 1.2.1 人机界面的定义

用户界面(User Interface)又称人机界面(Human-Computer Interface),它作为计算机系

统的一个重要组成部分,是计算机科学、心理学、认知科学和人素学(Human Factors)的交叉研究领域,也是计算机行业竞争的焦点从硬件转移到软件之后,又一个新的、重要的研究领域[Foley 1990]。

人机界面[Brown 1989a,董士海 1994a]是指人类用户与计算机系统之间的通信媒体或手段,它是人机双向信息交换的支持软件和硬件。这里“界面”定义为通信的媒体或手段,它的物化体现是有关的支持软件和硬件,如带有鼠标的图形显示终端。人机界面和人机对话(Human-Computer Dialogue)是两个不同的概念。人机对话是指用户与计算机系统之间的通信,它是人与计算机之间各种符号和动作的双向信息交换。这里的“对话”定义为一种通信,即信息交换,而且是一种双向的信息交换,可由人向计算机输入信息,也可由计算机向使用者反馈信息。这种信息交换的形式是以各种符号和动作的方式出现的,如键盘上的击键、鼠标的移动、显示屏幕上的符号或图形等。人机之间的对话是通过人机界面来实现的,在界面开发过程中,有时把它们作为同义词使用。

随着计算机技术的不断发展,软件在计算机中所占的比重越来越大。在软件研制中,人机界面的设计和开发在整个系统研制中所占的比重较大,约占 40%~60%。人机界面作为人机通信的途径,其效果好坏直接影响到一个软件的功能、使用及推广。因此,人机界面技术已成为世界各国软件工作者所着重研究的关键技术之一。

### 1.2.2 人机界面的发展过程

人机界面从命令行界面发展到 WIMP 界面,并最终发展为多通道界面。

#### 1. 命令行界面[李茂贞 1997]

人机之间的早期通信语言是机器语言,而后是汇编语言,直至高级语言。人机之间通过语言中的输入输出功能完成交互。这个时期用户对系统的运行很少进行干预,系统设计中基本不单独考虑人机界面的问题。这种状况大约持续到 1963 年。

60 年代中期出现了交互终端和分时系统。交互终端可以把各种输入输出信息直接显示在终端屏幕上,分时系统使用户可以分时共享计算机系统资源,这个时期的系统设计开始考虑如何方便用户的使用。例如,可以构成问答式对话、文本菜单或命令语言等方式来进行人机交互。这个时期的人机界面为命令行界面(Command Line Interface, CLI)。

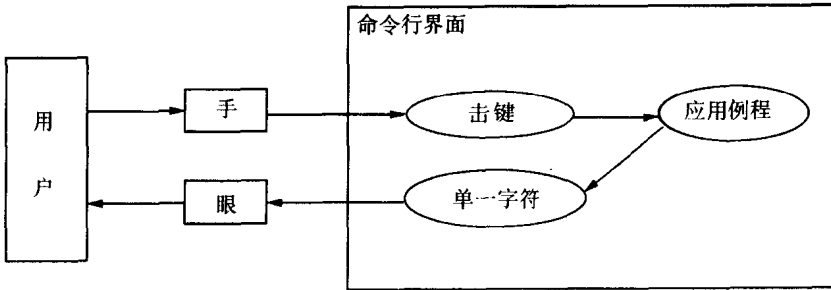


图 1.1 命令行界面概念模型

命令行界面是最早出现的人机界面。在这种界面中,人被看成操作员,机器只做出被动的反应,人只能使用手一种交互通道通过键盘输入信息,界面输出只能为静态单一字符。因

此,这种人机界面交互的自然性和效率都很差,命令行界面可以看作是第一代人机界面(图 1.1)。在命令行界面中,界面和应用还没有分开。

## 2. 图形用户界面[李茂贞 1997]

从 60 年代开始,特别是经过 70、80 年代,由于超大规模集成电路的发展,基于位图映象显示的高分辨率图形显示设备及鼠标定位设备的出现和应用、微机的普及、工作站的广泛使用,以及计算机图形学、软件工程、人工智能、窗口系统等软件技术的进步,都给图形用户界面技术的发展提供了强有力的支持,这时的人机界面进入了图形用户界面(Graphical User Interface, GUI)占统治地位的时期。

早在 60 年代,美国 MIT 的 Ivan Sutherland 就研究出了许多可以被认为是原始的 GUI 技术。Sutherland 在他发明的 Sketchpad 中,首次引入了菜单、不可遮盖的瓦片式窗口、图标,并采用光笔进行绘图操作。不久美国加州的 SRI International 公司的 Douglas Engelbart 发明了鼠标。70 年代施乐公司在 Alto 计算机上首次开发了位映象图形显示技术,即屏幕上显示的每个像素点都受计算机内存中某一位的控制。位映象图形显示技术为 GUI 的发展提供了许多新的可能性,如可重叠窗口、弹出式菜单、菜单条。这些工作奠定了目前图形用户界面的基础,形成了所谓的 WIMP 界面,即以窗口(Windows)、图标(Icon)、菜单(Menu)和指点装置(Pointing Device)为基础的人机界面技术。苹果公司于 80 年代将 WIMP 技术引入到微机领域,它推出的 Macintosh 以其全鼠标及下拉式菜单操作和直观的图形界面引发了微机人机界面的历史性变革。而后微软公司推出了 Windows 系统,并被应用于用户面更广的个人计算机平台,推动了计算机的普及。

WIMP 界面可看作是第二代人机界面,是基于图形方式的人机界面。在 WIMP 界面中,人被称为用户,人机通过对话进行工作,但用户也只能使用手这种交互通道输入信息。在 WIMP 界面中,界面输出为静态/动态二维图形/图像及其他多媒体信息。与命令行界面相比,WIMP 界面的人机交互自然性和效率都有较大的提高(图 1.2)。WIMP 界面很大程度上依赖于菜单选择和小装饰(Widget)。经常使用的命令大都通过鼠标来实现[David 1992]。鼠标驱动的人机界面使得初学者易于使用,但重复性的菜单选择会给有经验的用户造成不方便,他们有时倾向使用命令键而不是选择菜单[Iseki 1986]。WIMP 界面支持直接操纵(Direct Manipulation, DM)。Shneiderman[Schneiderman 1982]最早提出了 DM 的概念。DM 的 WIMP 界面具有如下特点[Ziegler 1988][董士海 1994a]:

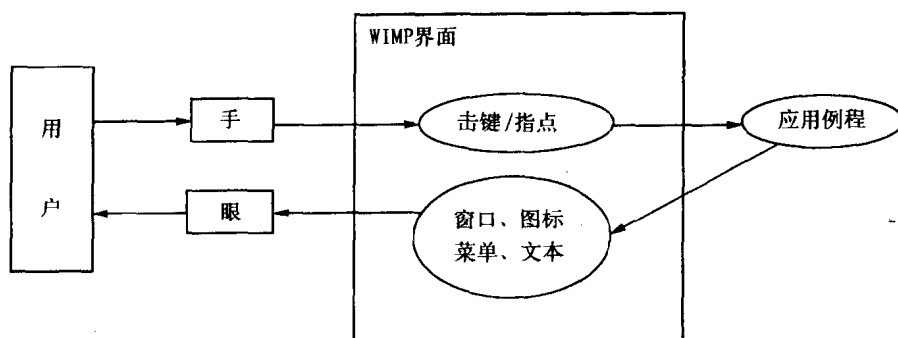


图 1.2 WIMP 界面概念模型

1) 对象的仿真表示。文字、符号、代码、图形及图像都可用来表示用户任务中的某一对象,这样易于人们的理解;

2) 实际动作代替复杂的语法;

3) 操作结果的立即应答和直观显示;

4) 动作的连续性和可逆性。

但 DM 也存在一些问题,如有些图标的含义不如文字清晰,其表示可能使用户操作时产生误动作或迷路,DM 要求显示屏大及响应时间快[Karl 1992]。

目前的多媒体界面(Multimedia Interface)可以看做 WIMP 界面的另一种风格。多媒体界面只是在界面信息的表现方式上进行了改进,采用了多种媒体,即计算机到用户的通信带宽提高。但是,用户到计算机的通信带宽却未提高,人机交互效率仍然很低。

从命令行界面发展到 WIMP 界面,计算机到用户的输出带宽大大提高,但从用户到计算机的通信带宽仍然受到限制,即计算机加工和呈现视觉、听觉的信息的能力与人机交互的能力是不相称的。这种不相称来自于硬件的限制和人的交互通道与对象操作间的不匹配。近年来,计算机输入输出装置在数量和能力上迅速增加,使得这种匹配成为可能。虚拟现实技术、多媒体和可视化对计算机系统的人机交互提出了高效、三维和非精确的要求,这些都是目前的 WIMP 技术所不能解决的问题。

无论是命令行界面,还是 WIMP 界面,实质上都属于单通道人机界面。在这种界面中,用户使用精确的信息在一维(命令行界面)和二维(图形用户界面)空间中完成人机通信,这是一种静态的人机界面,这种界面不具有自然进行三维直接操作的交互能力。为适应目前和未来的计算机系统要求,人机界面应能支持时变媒体(Time-Varing Media)实现三维、非精确及隐含的人机交互,多通道人机界面(Multimodal User Interface, MMI)是达到这一目的的重要途径[Wilsson 1991]。

### 3. 多通道人机界面

首先从用户的心理学角度给出几个概念。

#### (1) 通道

通道(Modality)一词源于心理学,其词源为“Mode”(方式或模式)。在讨论视觉、嗅觉、触觉等感觉的“方式(非内容)”时,或者在涉及多种感觉之间的关系或讨论一种感觉不同于其他感觉的特点时,心理学中使用“感觉通道”一词。本文把“通道”视为“方式”的同义词,不仅表示用户的“感觉方式”,也指用户的“动作方式”。用户可以使用手动、语音、眼神等多种交互通道与计算机系统进行交互。

#### (2) 通道(Modality)和模式(Mode)

尽管 Modality 的词源是 Mode(模式),但在多通道的研究中,它们却有着不同的内涵。Modality 指传送或获得信息的通信通道的类型,它包含了信息表达、感知以及动作执行的方式,定义了交换数据的类型;而 Mode 则指一种状态(State)或上下文信息(Context),它决定了如何对信息解释以获取意义(Meaning)。虽然 Modality 和 Mode 代表了不同的含义,但在一个多通道系统中它们却共同发挥作用。广义上讲,多通道系统是这样一种系统,它允许用户通过各种不同的人体通道如语音、手势、身体语言等与之进行通信,并能从中自动提取和传送语义信息。在许多的研究文献中,Mode 作为 Modality 的同义词使用。

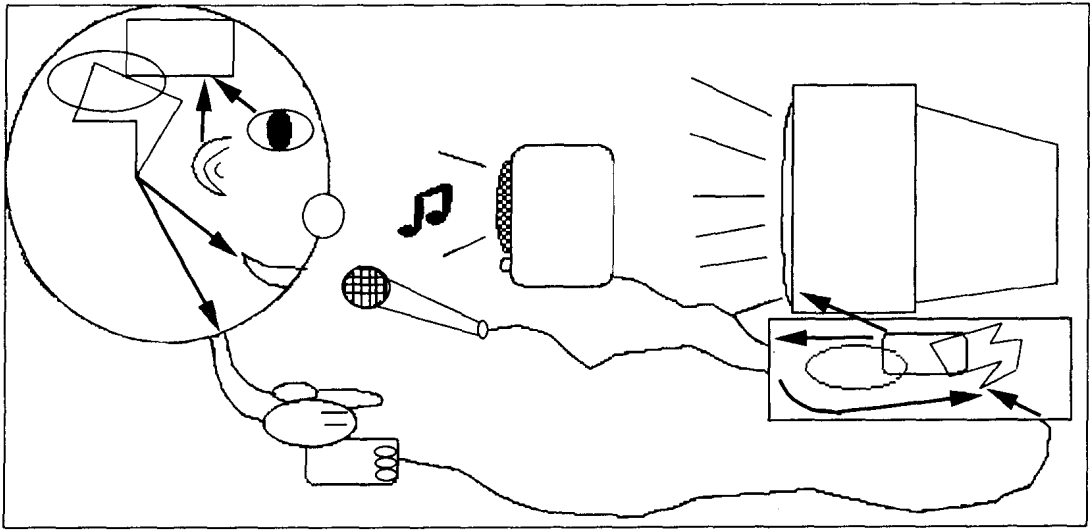


图 1.3 通道的概念

### (3) 多通道、多媒体与虚拟现实

多通道和多媒体都蕴含了多种通信通道的使用,区别在于系统内部信息所处的抽象层次的不同。多通道系统力求对语义的理解,它对用户的原始输入在较高的层次上(如任务领域)进行解释、分析和综合,提取语义内容,达到对用户的某种理解。然后,从较高的表现层次上传达给用户最恰当的反馈,它侧重于信息的理解和解释。而多媒体系统则表现为在计算机的控制下产生、操纵、表现、存储和传送一些独立的信息,这些信息至少是通过一种连续(时间有关)或离散(时间无关)媒体进行编码。多媒体侧重于信息的表现。

多通道与虚拟现实的主要区别在于两者研究的目的不同。虚拟现实(Virtual Reality, VR)依靠三维立体视觉、头盔式显示器、身体跟踪和立体音响等技术来模拟现实世界,旨在使用户获得一种沉浸式的多种感知通道的灵境体验。VR 中计算机生成的视听幻想(Audio-Visual Illusion)涉足人脑固有的感觉-效应通道的协调机制。而多通道研究则力求详尽地探索人体感知和控制行为中的各种并行和协作特性,并应用于计算机交互领域以获得一个谈话对象式的计算机(Computer-as-Dialogue-Partner)。从交互行为的研究上来说,VR 的研究与多通道研究之间存在交集,但它更侧重于应用系统。VR 系统是一个很好的多种交互设备综合使用的例子,而且多通道交互直接影响到 VR 的沉浸式体验,因此 VR 为多通道的研究提供了很好的契机和媒介。

### (4) 通道整合

通道整合(Modality Integration)是指用户在与计算机系统交互时,多个交互通道之间相互作用形成交互意图的过程。例如,在完成某个具体交互任务时,用户可能采用语音和手势两个通道,对任一通道而言,通道信息可能都不精确,但通过对两个通道进行整合,可以获取用户交互的意图。这样,语音和手势两个通道是互补的。

多通道人机界面(MMI)是基于视线跟踪、语音识别、手势输入、感觉反馈等新的交互技术。它允许用户利用多个交互通道以并行、非精确方式与计算机系统交互,旨在提高人机交互的自然性和高效性。MMI 就是要解决科学计算可视化、虚拟现实对计算机系统提出的高效、三维和非精确的人机交互要求。MMI 可以看作是第三代人机界面。

在 MMI 中,用户可以使用自然的交互方式,如语音、手势、眼神、表情等与计算机系统进行协同工作。交互通道之间有串行/并行、互补/独立等多种关系,因此人机交互方式向人与人交互靠拢,交互的自然性和高效性得到极大的提高(图 1.4)。

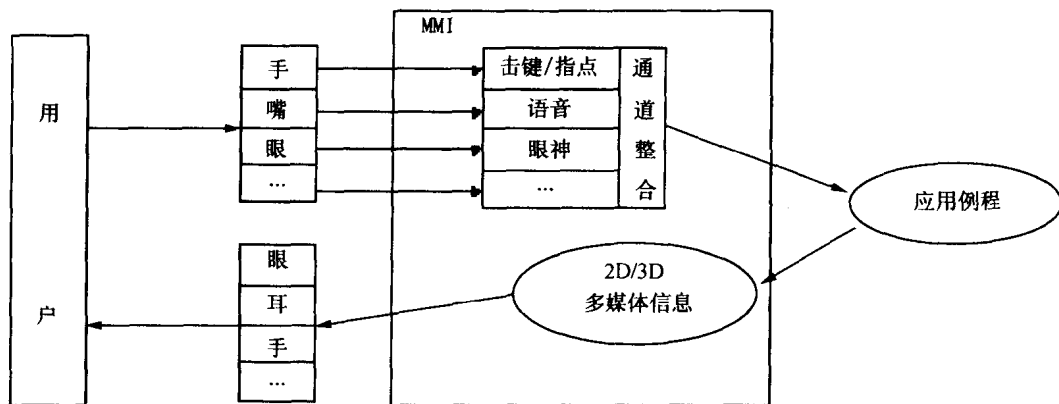


图 1.4 多通道人机界面概念模型

多通道交互研究中还存在一些特殊的问题,解决这些问题是多通道用户界面研究的主要工作。多通道人机界面研究中的主要问题是:

#### (1) 多通道整合

多通道用户界面的特点就在于,利用多个感觉和效应通道的并行和协作与计算机系统进行交互。由于多通道交互中各个通道不再局限于精确输入,因此,如何从多个并行、非精确的交互通道中获取用户要传达的信息,就成为多通道用户界面技术应该解决的主要问题,即多通道的整合问题。多通道整合是多通道用户界面研究的重点和难点,它的解决是实现多通道人机交互的首要前提。目前,有关多通道整合的研究主要集中于把语音与手势或手动的整合方面[Gourdol 1990,Hauptmann 1993,Kloosterman 1994]。

#### (2) 工效学问题

多通道人机交互并不意味着多个通道在所有场合都优于一个通道。必须对交互风格(涉及交互所用的感觉和效应通道、交互设备以及信息表示设计等)与交互任务之间的匹配加以具体分析。对多通道界面来说,存在着什么样的整合适用于哪种场合和任务的问题,这主要是工效学研究的内容[Karl 1994,Jacob 1994]。

#### (3) 软件结构

为了简化日益复杂的界面设计工作,人机交互的研究中提出了对话独立性原则,Seeheim 模型典型地表达了这样的思想。基于这种思想实现了许多用户界面管理系统(UIMS),但其不足之处是支持的语义反馈能力较弱[Zhong,Dong 1992]。多通道界面的重要目标是增强人机交互的语义反馈,而充分的语义反馈又是用户进行多通道协同操作的基础,只有在语义信息足够充分的情况下多通道整合才有可能进行。为了解决好这些问题,必须重新考虑多通道交互中软件结构的问题。

#### (4) 多通道界面描述方法

多通道用户界面基于多种交互通道,以用户使用的自然性和高效性为宗旨,需要对用户的工作负荷、视觉反应等人的因素作工效分析和评估,这就涉及到认知心理学知识。所有这些人机交互的特点无疑给多通道用户界面的描述方法提出了许多新的课题。传统的界面描

述方法如状态转换图及其扩充方法[Mark 1986; Bordegoni, Clive 1986; Jacob 1993c; Phillips 1986; Harel 1987; Harel 1988]、基于事件的描述[Mark 1986; Green 1985a, Green 1985b; Hill 1986; Flechia 1987]、基于上下文无关文法[Mark 1986, Ales 1994; Guest 1982; Hanau 1980]、基于 Petri 网[Philippe, 程景云 1994]等描述方法在描述多通道界面方面都存在一定的困难。具体来说,一个理想的多通道用户界面的描述方法应具备以下基本的表达能力:

- 1) 给设计者提供一种直观方法,使之觉得方便好用。

- 2) 描述多个通道的非精确输入/输出(I/O)事件。为实现多通道间的信息流整合,各个通道的 I/O 事件最好能采用统一的表示方法。

- 3) 方便地表示多个通道之间自然的并行、同步、选择等协作关系和约束关系。

- 4) 描述多通道非精确输入的整合。传统人机交互的输入操作中用户每次只使用一种交互设备来指定一个或一系列完全确定的命令或参数。在多通道界面中,可能需要整合多个交互通道的信息才能确定一个命令或动作语义,如何描述这种通道整合是多通道界面描述的难点。

- 5) 由于涉及到多通道人机交互的自然性,所以要对多通道用户界面模型进行评估和工效分析。这就要求描述方法应该显式地反映出用户模型[Norio 1989, Staggers 1993, Carroll 1988]和任务分析模型[Kieras 1988; Reisner 1981; Payne 1986; Barnard 1987; Johnson, Young 1989]的分析结果,以便进行认知心理学的评价。

#### (5) 多通道用户界面的评估

多通道用户界面的评估[Hekmatpour 1987, Sutcliffe 1991, Whitefield 1987, Moran 1981, Kieras 1985, Bennett 1987, Kellogg 1987]包括很多方面。针对多媒体,虚拟现实和可视化而言,MMI 比 WIMP 界面在自然性(Intuition)、易学性(Learnability)和高效性(Effectiveness)三个方面都要好。自然性是指最终用户使用计算机系统所要求的交互通道满足程度;易学性是指用户掌握系统功能和规则的难易程度;高效性指用户能成功地完成所规定任务的快慢程度。

为了使系统达到自然、易学和高效,用户界面的设计者必须有一个正确的用户模型[Card 1983, Carroll 1988, Gugerty 1993, Staggers 1993]。用户也必须建立起正确的心理模型。心理模型是指用户对系统及其成分的心理表达,而用户模型是指系统开发人员对用户特征的理解和表达,其中包括用户的物理特征以及认知特征。用户模型和心理模型的不一致会使整个人机界面的性能大大下降。

传统的人机交互主要利用了人的手与眼通过键盘、鼠标、显示器进行二维与精确方式的输入输出。多通道人机交互必须寻求新的交互手段,以充分利用人的眼、耳、嘴、手,视觉、听觉及触觉通道。交互技术的不断发展为多通道界面研究提供了技术基础。多通道界面研究可以考虑以下交互技术:

三维交互技术[董士海 1994b]、语音技术[Gourdol 1990, Dillon 1990]、视线跟踪技术[Hutchinson 1989, Jacob 1993a, Jacob 1993b]、姿势输入技术[Hauptmann 1993, Pigueiredo 1993]、反馈技术[董士海 1994a, Akamatus 1994]、自然语言界面[Wilsson 1991, Bos 1994]、其他交互技术如唇读、表情识别,用脚进行鼠标式输入等。

## 第二章 人机交互的发展趋势

人机交互的发展趋势将体现在多媒体表现方式、多通道交互、三维交互、计算机支持的协同工作(CSCW)几个方面。

### 2.1 多媒体技术的支持

计算机的用户界面是人机对话部分的接口。目前广泛使用的字符型和图形用户界面,是通过人输入字符串或点击图标等向计算机送入数据或命令,计算机也是通过输出字符串或图形告诉用户计算结果或各种信息。当前多媒体技术的迅速发展,为计算机用户界面技术注入了新的活力,它使用户可以采用十分自然的方式与计算机进行交互。计算机所表现的信息除文字、图形外,还包括声音、静止图像、动态图像、动画等。多媒体技术应用于计算机系统,使人们可得到更直观的信息,从而简化了用户的操作,扩展了应用范围,同时也对计算机的处理能力、存储能力、通信能力及编程技术等提出一系列新的要求。

#### 2.1.1 多媒体技术简述

什么是多媒体技术?什么是多媒体系统?对此人们各自有自己的理解。在1992年的SIGGRAPH会议上,当时SGI公司的董事长、1984年计算机图形学成就奖获得者James Clark阐述了他的观点:“多媒体技术是将正文、声音、图形、静止图像、动态图像等与计算集成在一起的技术。”从这个观点出发,他认为,目前是计算机图形图像产业飞速发展的极好机遇。其理由是消费类电子厂商正为高清晰度数字电视的标准争论不休,而三维图形硬件、图形图像通信、图像压缩等技术已成熟,因此,计算机工业有可能在原有电视、录像、音响的消费类电子产品基础上,迅速开发多媒体系统,这样会拓宽应用领域:电视、游戏、电子报刊、电视电话、电视会议、教育培训、文献数据库等。人们预计在不久的将来,在计算机统一管理下的视象、音响设备将广泛应用于家庭、办公室、研究部门等。

多媒体技术已经在大、中、小、微各种计算机硬件平台上开发,如SUN,SGI/Indigo;NeXT;IBM/PC;Macintosh等。由于微机易于普及,因此在微机平台上的多媒体系统竞争十分激烈,发展尤为迅速。多媒体数据库、多媒体网络通信、多媒体用户接口等研究领域十分活跃。发展多媒体技术需要解决的一系列关键技术有:超大规模集成电路、大容量光盘、图像压缩、实时信号处理、多媒体制作工具等。多媒体技术正是这些技术的综合。当前多媒体技术正处于广泛的实用化过程中,并向着进一步压缩、集成化、交互化等研究方向发展。

#### 2.1.2 多媒体技术的进展

下面分别就图像压缩及其标准、大容量光盘、多媒体系统、多媒体应用等方面讨论多媒体技术的进展。