




眼

动人机交互

Eye-movement
Based Human-Computer
Interaction

冯成志 著

 苏州大学出版社

眼动人机交互

Eye-movement Based Human-Computer Interaction

冯成志 著

T618
1439

苏州大学出版社

图书在版编目(CIP)数据

眼动人机交互/冯成志著. —苏州:苏州大学出版社,2010.4
ISBN 978-7-81137-466-7

I: ①眼… II. ①冯… III. ①眼动-人-机系统
IV. ①TB18

中国版本图书馆CIP数据核字(2010)第072456号

内 容 简 介

随着计算机硬件的发展、处理能力的增强,人机交互成为制约其发挥作用的关键因素。人与计算机的交互一直受到输入/输出之间信息不平衡的制约,使得用户到计算机的输入带宽不足,在此基础上催生了诸如视线追踪、语音输入等许多新的输入技术。本书首先回顾人机交互的重要性、眼动的概念与测量以及眼动人机交互的由来,然后阐述了人机交互中所涉及的工效学问题及相关的实验研究,以便为眼动人机界面的设计提供科学依据,并通过开发的视线人机交互系统验证了工效学参数的可行性。

本书可以作为高等院校计算机、心理学、工程等专业的本科生和研究生的教材,同时亦可作为从事眼动研究的科研工作者的读本。

眼动人机交互

Eye-movement Based Human-Computer Interaction

冯成志 著

责任编辑 马德芳

苏州大学出版社出版发行

(地址:苏州市十梓街1号 邮编:215006)

常熟高专印刷有限公司印装

(地址:常熟市元和路98号 邮编:215500)

开本 787×960 1/16 印张 11.75 字数 261 千

2010年4月第1版 2010年4月第1次印刷

ISBN 978-7-81137-466-7 定价:23.00元

苏州大学版图书若有印装错误,本社负责调换
苏州大学出版社营销部 电话:0512-65225020
苏州大学出版社网址 <http://www.sudapress.com>

前 言

人机交互(Human-Computer Interaction, HCI)是研究人、计算机以及它们之间相互影响、相互作用的技术,是人与计算机之间传递、交换信息的媒介和接口,用户通过它向计算机系统输入指令,计算机系统再通过它向用户反馈信息。但人与计算机的交互一直受到输入/输出之间信息不平衡的制约,使得用户到计算机的输入带宽不足。视线追踪技术是在“以用户为中心”、自然和谐而高效的人机交互思想指导下研究开发出的输入/输出新技术。借助于视线追踪技术,通过把捕获的眼动信号作为用户向计算机输入的一种信息,以此达到操作和控制计算机的目的。将视线追踪技术应用于人机交互能够既服务于普通大众,又服务于特定用户和场合。

尽管有关眼动的研究已有一百多年的历史,但将视线追踪技术应用于人机交互领域却是近二三十年来的事情。在国外,眼动人机交互的研究成果主要集中在20世纪90年代到21世纪初期,国内的相关研究相当匮乏。尽管有些人机交互的书籍也谈及把视线追踪技术作为一种交互方式,但大多缺乏实质性的研究内容。因此,有必要结合本人所做的相关研究工作将眼动人机交互整理成册,以便让更多的用户了解这一新兴领域,同时让对眼动人机交互感兴趣的学生和研究者比较系统地掌握眼动人机交互及相关的研究内容。

全书主要包括四个部分,共十一章内容。首先介绍人机交互的重要性及基本概念和最新成果的运用,为眼动人机交互做铺垫;然后通过眼动的概念和测量技术让读者了解视线追踪的基本原理;其次是眼动技术在人机交互领域中的工效学参数研究,主要包括眼动人机交互的时间和空间特性参数,眼动人机交互中边框、视标及空间排列方式和“米达斯接触”问题;最后是对工效学参数的验证和应用研究,通过系统实例(单通道眼控汉字输入系统和多通道家电管理系统)来考察工效学参数在实际的眼动人机交互中的可应用性。

本书的完成了却了自己多年的心愿。本书的出版要感谢苏州大学出版社的全力支持,他们饱满的工作热情和认真负责的工作态度给本人留下了美好的回忆。本书写作过程中还得到了我的师兄张光强的帮助,我的研究生潘壮志、尹鸾、胡彬和许晓璐等在书稿的校订工作中做了大量的工作,我的爱人贾凤芹女士的默默付出使得本人能够顺利完成此作,在此一并表示诚挚的感谢。

由于本人学识浅陋和时间所限,书中的缺点与错误在所难免,恳请各位读者不吝指教。

冯成志

2010年4月

目 录

| | |
|--------------------------------|----|
| 第 1 章 引言 | |
| 1.1 人机交互的重要性 | 2 |
| 1.2 人机交互的构成要素 | 3 |
| 1.3 本书研究的主要问题 | 9 |
| 参考文献 | 10 |
| 第 2 章 眼动的概念及测量 | |
| 2.1 眼动的生理机制 | 12 |
| 2.2 眼动及测量技术 | 13 |
| 参考文献 | 20 |
| 第 3 章 眼动人机交互的由来 | |
| 3.1 眼动用于人机交互的原因 | 24 |
| 3.2 眼动信号的提取和转换 | 27 |
| 3.3 眼动信号的利用 | 28 |
| 3.4 视线追踪技术在人机交互中的应用 | 31 |
| 3.5 多通道人机交互:视觉通道与其他通道的整合 | 34 |
| 参考文献 | 39 |
| 第 4 章 眼动人机交互中的时空特性 | |
| 4.1 实验方法 | 46 |
| 4.2 结果与分析 | 49 |
| 4.3 讨论 | 55 |
| 4.4 小结 | 57 |
| 参考文献 | 57 |
| 第 5 章 边框和视标对作业绩效的影响 | |
| 5.1 实验方法 | 60 |

| | | |
|-----------------------------|----------------------|-----|
| 5.2 | 结果与分析 | 61 |
| 5.3 | 讨论 | 64 |
| 5.4 | 小结 | 65 |
| | 参考文献 | 65 |
| 第6章 空间排列方式对作业绩效的影响 | | |
| 6.1 | 实验方法 | 66 |
| 6.2 | 结果与分析 | 68 |
| 6.3 | 讨论 | 71 |
| 6.4 | 小结 | 73 |
| | 参考文献 | 73 |
| 第7章 “米达斯接触”的解决方案 | | |
| 7.1 | 实验方法 | 75 |
| 7.2 | 结果与分析 | 77 |
| 7.3 | 讨论 | 86 |
| 7.4 | 小结 | 87 |
| | 参考文献 | 88 |
| 第8章 单通道眼控汉字输入系统 | | |
| 8.1 | 眼控汉字输入系统的实现 | 90 |
| 8.2 | 基于眼控汉字输入的屏幕软键盘布局研究 | 95 |
| 8.3 | 小结 | 103 |
| | 参考文献 | 103 |
| 第9章 多通道家电管理系统 | | |
| 9.1 | 基于视线和语音的多通道家电管理系统的实现 | 105 |
| 9.2 | 基于多通道家电管理系统的工效学研究 | 112 |
| 9.3 | 小结 | 117 |
| | 参考文献 | 118 |
| 第10章 视线追踪技术在界面评估中的作用 | | |
| 10.1 | 界面评估的重要性 | 121 |
| 10.2 | 界面评估的维度 | 123 |
| 10.3 | 传统的界面评估方法 | 123 |
| 10.4 | 利用视线追踪技术的界面评估 | 126 |
| | 参考文献 | 130 |

| | |
|------------------------------------|-----|
| 第 11 章 进一步的眼动交互研究 | |
| 11.1 眉目传情 | 132 |
| 11.2 望眼欲穿 | 141 |
| 参考文献 | 153 |
| 附录 | |
| 附录 1 数据表格 | 155 |
| 附录 2 实验指导语 | 161 |
| 附录 3 实验材料 | 165 |
| 附录 4 自编评价量表 | 169 |
| 附录 5 Levenshtein Distance 算法 | 173 |
| 附录 6 术语索引表 | 175 |
| 后记 | 177 |

第 1 章

引 言

人机交互是对人(用户)和计算机之间的交互进行研究的学科,它主要涉及设计和评估供人们使用的交互式计算机系统,其目的是提高人与计算机之间交互的效率,如今人机交互已经成为非常重要的理论和应用学科。随着人类社会的发展和科学技术的进步,计算机的功能越来越强大,所能完成的工作越来越多,甚至可以模仿人类思维。这充分说明了 Mark Weiser 所提出的“普适计算”(Ubiquitous Computing)的真正涵义。在人类的发展史上,“人类的进化特别是智能的进化,不仅依靠天然器官的变化,而且主要依靠人造器官——工具的变化。人类要过河,可以造船;人类要快跑,可以造出汽车、火车;人类要上天,可以造出飞机和飞船,等等”(方宗熙,江乃萼,1976)。早期的人类一方面通过改进自己的身体去适应环境去求得生存,另一方面也借助于制造出来的工具去适应环境,只有在后一种情况下才显现出人类对环境改造的积极性和主动性。人类之所以借助于工具,是因为人类自身有许多局限性。例如,我们的视听受到时间和空间的限制,我们的短时记忆容量(7 ± 2 个单元)有限,我们的运算速度不够快等。通过工具就可以弥补人类自身的缺陷,进而提高工作效率,因此广义的人机交互(Human-Machine Interaction, HMI)不只局限于计算机,还包括人类所使用的各种工具系统,但其最终目的都是使用户在使用系统时达到最高的效率和满意度。



图 1-1 人机交互的多学科性

如图 1-1 所示,正因为人机交互所涉及的多学科性,所以 HCI 有时也被称为 MMI(Man-Machine Interaction)或 CHI(Computer-Human Interaction)。

1.1 人机交互的重要性

作为一种自动的高速计算工具,计算机首先被广泛应用于工程计算和测量中,由于计算机所具有的自动执行存储程序的能力和超强的存储容量,使其成为取代人脑的“电脑”,从而把人类从繁琐的劳动中解放出来,投身到更具创造性的活动中去。在提高生产效率和生产力的同时,计算机还提高了工作质量,借助于计算机(及网络)人们可以快速获取所需的网络资讯;借助于文字排版系统人们可以编排出更加光彩夺目的作品;借助于计算机辅助设计系统人们可以快速设计出更加合理、更加美观的汽车和服装;借助于计算机应用程序人们可以规划自己的日程时间安排、建立自己的电子账本等;借助于计算机教学软件人们可以学习和掌握不同的知识和技能;借助于计算机游戏人们可以模拟各种活动或角色,丰富自己的业余生活等。总之,计算机的发展不仅可以极大地促进科学的探索,还可以丰富个人的文化娱乐生活。上述所提到的各种应用系统都离不开人机交互这一话题,而人机界面的设计和开发在整个系统研制中所占的比重约为 40% ~ 60%,再加上不同人群对界面的要求不尽相同,因此人机界面技术已经成为世界各国软件工作者着重研究的关键技术之一。

人机交互的研究在计算机科学相关领域引起越来越大的兴趣,逐渐成为计算机系统的重要组成部分。人机交互在软件中的地位日益突出,其重要性得到学术界、产业界的认同。在以人为本的理念下,自然和谐而高效的人机交互相关理论和技术,将成为重要的研究课题。从其发展就可以看出人机交互对于新技术的普及和应用具有重要的影响,如在互联网普及之前,目前我们已经熟知的电子邮件、即时通信和数据库管理系统从技术上讲早已比较成熟,但直到浏览器的出现,才真正使这些新技术更贴近人们的生活,可以说互联网的迅速普及与人机界面的突破密不可分,或者说人机界面的革新能够给人类的生产和生活带来巨大的变革。

随着物联网(Internet of Things)、信息物理系统(Cyber Physical Systems, CPS)等新概念、新技术的出现,人机交互技术的重要性日益凸显。设想一下如果能够设计出不用学习(或者只需极少的学习)就可以极大地方便用户使用的人机交互界面,此举对于新技术和新产品的推广和市场的占有具有多么深远的意义,这也是微软在人机交互领域的资金投入力度非常大的原因,这一目标也成为计算机相关领

域的科学家共同关注的焦点。人机交互是最接近大众的信息技术,可以说所有技术都无法摆脱人机交互这一“紧箍咒”,谁拥有解开这一“紧箍咒”的钥匙,谁就拥有巨大的市场和发展前景。

目前,国内诸如中科院软件所、中科院心理所、清华大学、北京大学、浙江大学、大连海事大学、北京理工大学、北京师范大学等已将人机交互作为一个重要的科研方向,亦显示出人机交互的重要性。由于人机交互涉及多个学科,具有多学科性(图 1-1),因而是所有学科里最需要跨学科参与的研究之一。例如,在微软亚洲研究院有 200 多名具有不同教育背景的博士科研人员从事相关的研究工作。人不但要与机器(计算机、机床)打交道,还要与物品(手机、智能汽车)打交道,这些都离不开人机交互。在现代信息社会中,可以说没有人机交互的产品是不存在的。

1.2 人机交互的构成要素

从人机交互的信息加工模型(图 1-2)或者人机交互的概念中可以看出,在人机交互中存在两个主体:一个是用户,一个是计算机。人机交互就是两个主体之间的交流与对话,而要保证双方对话的自然、高效、和谐,双方应该建立一个有效的用于沟通和交流的平台,这个平台就是人机界面(Human-Computer Interface, HCI 或 Man-Machine Interface, MMI),由于其是面向用户的,因此也称之为用户界面(User Interface, UI),它是用户与系统进行交互的各种手段的总称,包括输入(Input, 用户通过它来操纵系统)和输出(Output, 系统向用户展示用户操纵所带来的变化或结果)。例如,在驾驶汽车时,驾驶员(用户)通过方向盘来控制汽车(系统)行驶的方向,通过加速踏板、刹车踏板和变速杆来控制汽车(系统)的速度;同时通过前挡风玻璃、后视镜来判断汽车的位置,以及通过读取仪表盘上的速度计来获知目前的行驶速度;驾驶员正是通过这些设备完成驾驶任务。一般来说,用户界面常用于指人与计算机或电子产品交互时的界面,而人与机械系统、交通工具或工业设备交互时的界面常称作人机界面,不论何种称谓,都是指将用户与机器本身(系统)分隔开来的“层”。因此,人机交互的构成包括三个要素:用户、界面和计算机。

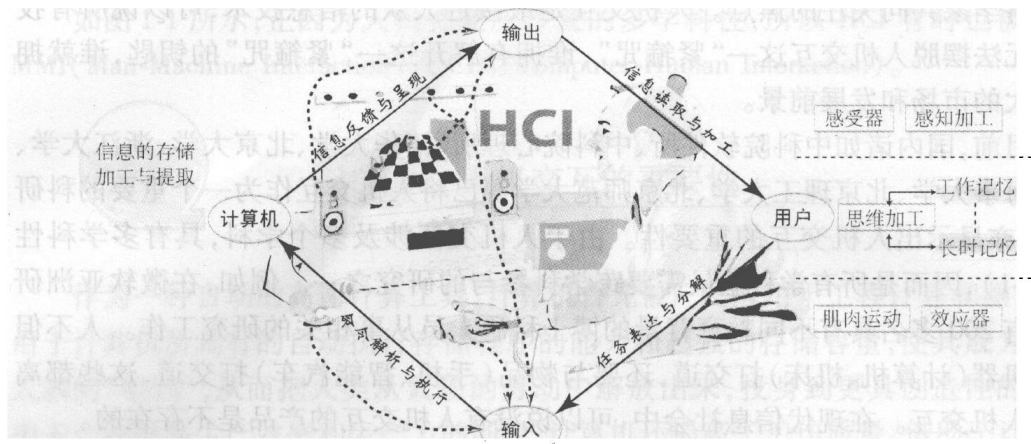


图 1-2 人机交互的信息加工模型

目前常用的用户界面包括：

1. 图形用户界面 (Graphical User Interface, GUI)

图形用户界面允许用户以图形的方式而非文本命令[早期的命令行界面,如 DOS(Disk Operating System)系统]与系统进行交互(图 1-3)。

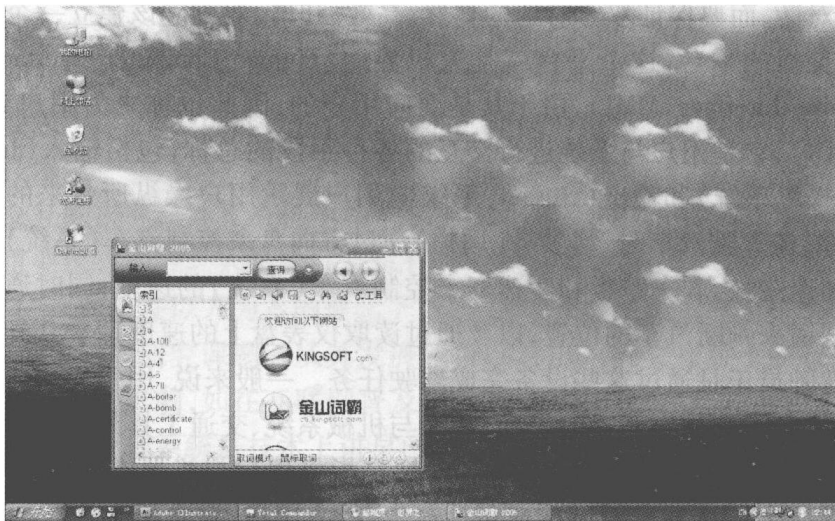


图 1-3 图形用户界面 (GUI)

2. 网络用户界面 (Web-based User Interface, WUI)

以超文本标识语言 HTML 及超文本传输协议 HTTP 为基础的网络浏览器是网络用户界面的代表,从世界上第一个网络浏览器 Mosaic 到网景公司的 Communicator,再到目前占据市场份额达 70% 的微软的 Internet Explorer,在互联网时代,成为

人们传递、交换资讯和获取信息不可或缺的人机界面(图 1-4)。目前我们可以利用 Java Applet、AJAX(Asynchronous JavaScript and XML)、Adobe Flex(大家熟知的 Flash 动画)、Microsoft .NET Framework 等开发新一代交互式网页应用程序,它们能够独立于网络浏览器,不需要通过刷新页面而实时地与用户进行交互,从而提高网络应用程序的交互效率,降低对通信带宽的占用。

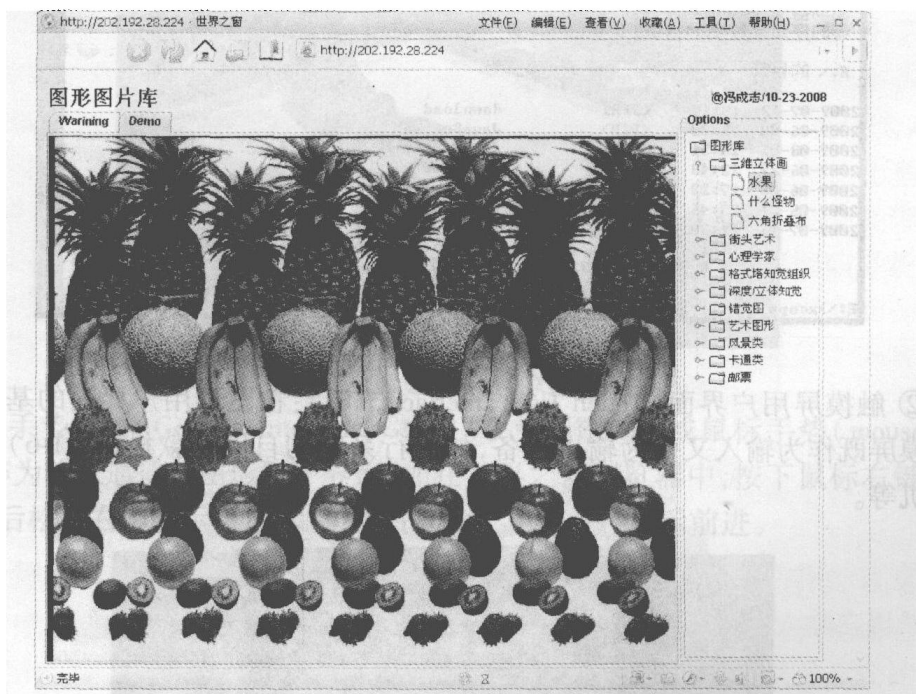


图 1-4 网络用户界面(WUI)

除此之外,还有其他用户界面,有些是在上述两种常用的界面的基础上发展起来的。例如:

① 命令行界面(Command Line Interface, CLI):用户通过键盘输入命令字符串,系统通过显示屏进行信息反馈(图 1-5)。

```
C:\WINDOWS\system32\cmd.exe
2009-08-12 14:35 <DIR>          文章
2009-07-07 11:27 <DIR>          课题
2009-05-15 21:20          728,928 输出.doc
                45 个文件      17,508,040 字节
                18 个目录  61,677,780,992 可用字节

D:\My Documents>e:
E:\>dir
驱动器 E 中的卷没有标签。
卷的序列号是 FCFE-6A29

E:\ 的目录
2009-07-12 18:10 <DIR>          download
2009-06-02 13:20 <DIR>          JavaSources
2009-08-01 14:31 <DIR>          Javaworkspace
2009-06-27 22:48 <DIR>          PQMETHOD
2009-06-08 17:23 <DIR>          temp
2009-05-21 23:45          921,654 test.bmp
2009-07-13 21:30 <DIR>          课题项目
                1 个文件      921,654 字节
                6 个目录  72,321,355,776 可用字节

E:\>xcopy *.* I:\
```

图 1-5 命令行用户界面(cmd.exe)

② 触摸屏用户界面(Touch User Interface, TUI):在图形用户界面的基础上利用触摸屏既作为输入又作为输出设备,如银行系统的自助取款机(图 1-6)和触摸屏手机等。



图 1-6 触摸屏用户界面(自助取款机)

③ 触觉界面(Tactile Interface, TI):能够通过触觉反馈机制向用户输出信息,多用于计算机模拟领域(图 1-7)。

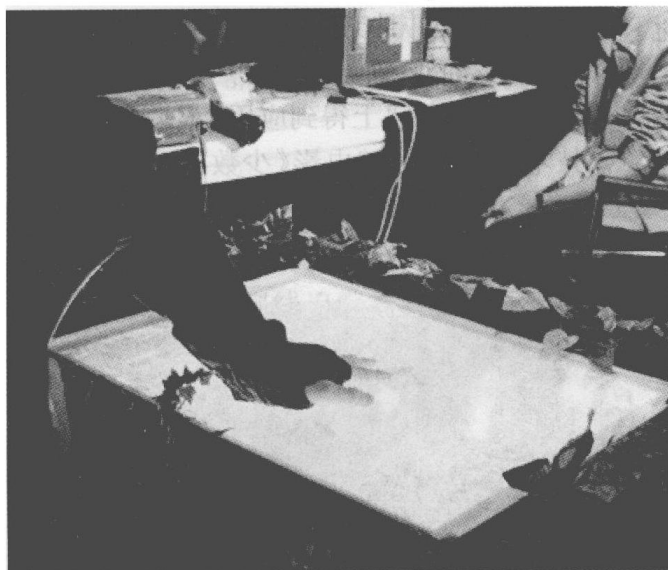


图 1-7 触觉界面(可穿戴式)

④ 手势界面(Gesture Interface, GI):以手势(图 1-8)或鼠标手势(mouse gestures)作为输入通道的图形用户界面,如在世界之窗浏览器中,按下鼠标右键向左拖动然后松开右键表示后退,如果向右拖动再松开右键表示前进。

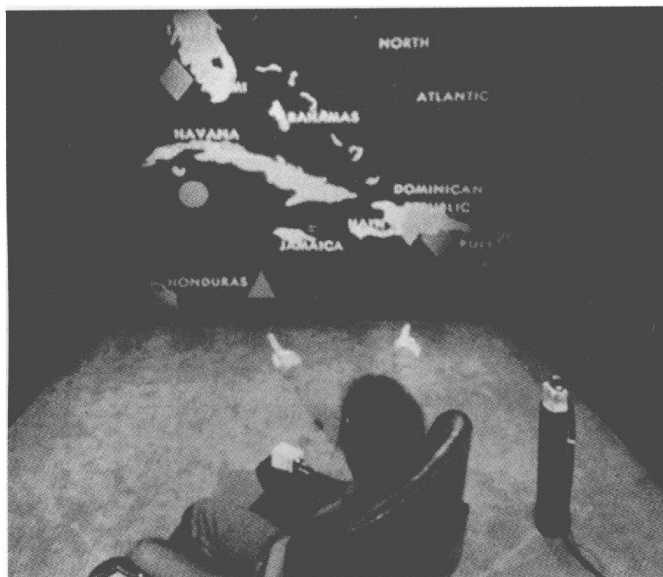


图 1-8 手势界面(Bolt,1980)

早期的触摸技术以单点触摸技术为主,在 2006 年的 SIGGRAPH 大会上,纽约大学的 Jefferson Y. Han 教授演示了运用在显示屏上的多点触摸技术,其领导研发

的新型触摸屏可由双手同时操作,并且支持多人同时操作(其演示视频可至<http://mrl.nyu.edu/~jhan/ftirtouch/multitouchreel.mpg> 下载)。如今多点触摸技术(Multi-Touch)在苹果公司 iPhone 手机上得到应用(苹果公司拥有多点触摸技术的专利)。对于多点触摸技术,看过好莱坞电影《少数派报告》的应该不陌生,其中男主角 Tom Cruise 戴上手套之后,运用多种手部姿势对屏幕上的照片、数据及影像进行缩放、移动及旋转等操作就是多点触摸技术的华丽应用(图 1-9)。



图 1-9 电影《少数派报告》里的场面

⑤ 自然语言界面(Natural Language Interface, NLI):用于搜索引擎或网页中,利用自然语言输入欲查找的问题(图 1-10)。



图 1-10 语音用户界面

从上面的几种界面可以看出,它们建立在人机交互技术的基础上,特别是输入技术的发展,给人机交互带来了无限生机,使人与计算机之间的交互越来越自然。

1.3 本书研究的主要问题

随着眼动测量技术日臻完善,视线追踪技术作为近年来出现的一种全新输入方式,视线追踪系统的性价比迅速提升,其对操作者的使用限制和干扰大为降低;同时由于视线追踪技术在人机交互中所表现出的潜在优势,使开发和优化基于视线追踪的人机实用系统成了人机交互领域中新兴的热点课题,在人机交互领域中备受青睐。然而若要将该技术成功地运用到人机交互中,有赖于对影响系统交互的时间参数、空间参数、反馈方式、对象空间排列及“米达斯接触”等多方面因素的工效学研究。这些研究目前尚属少见,仅有的研究只单独考察上述某个因素对系统效率的影响(Jacob, 1990; Stampe & Reingold, 1995; Yamato, Monden, et al., 2000),而忽视了各因素间的关系;在研究思路上也囿于根据现有的系统原型,比较执行特定的交互任务时的工效指标,以改进界面设计,而不是以工效学研究为先导发展系统界面。

本书在对人机交互的构成要素和一些基本概念有了基本了解后,接下来后续的章节中将介绍一些眼动的概念及测量技术和利用视线追踪技术特别是在心理学领域所进行的相关研究,以便对视线追踪技术有充分的理解,在此基础上论述将视线追踪技术运用于人机交互领域的缘由。本书将主要围绕以下几个方面的研究进行展开,构思如图 1-11 所示。

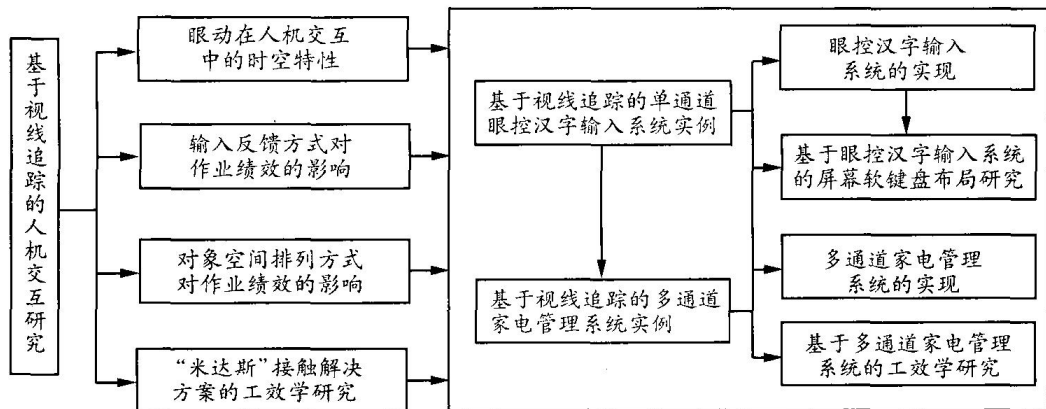


图 1-11 研究构思示意图