

第 3 章 人机交互界面

导读案例：芯片植入猪脑，实时读取信息

人们熟知科技大亨埃隆·马斯克(图 3-1),主要是因为他拥有几家知名公司,例如 SpaceX 和特斯拉,这两家公司分别参与了航天飞行和电动汽车的发展。马斯克以发表大胆、超前的技术宣言而著称,但从 SpaceX 和特斯拉的经验来看,其最终完成目标的时间远比计划设想的要长得多。不过,任何宏伟计划的实现都起源于最初的设想,关键是看你敢不敢想。2016 年,科技怪侠马斯克还创建了另一家研发脑机对接技术的 Neuralink 公司。

1. Neuralink 设备植入猪脑

2020 年 8 月 28 日,马斯克通过在线直播展示了几头小猪,其中一头是在两个月前大脑被植入脑机接口设备的小猪格特鲁德(图 3-2)。它状态良好,受到全球科技界的关注。马斯克称,“大约一枚硬币大小”的 Neuralink 设备(图 3-3)能够读取大脑活动,不会对大脑造成任何持久损害。该设备的电池续航时间为一整天,可以直接连接到相关的智能手机上。



图 3-1 埃隆·马斯克



图 3-2 小猪格特鲁德

2019 年展示的设备置于生物的左侧耳后,而新一代的 Neuralink 设备是完全无线的,使用感应充电,被置于头脑顶部的位置(图 3-4)。

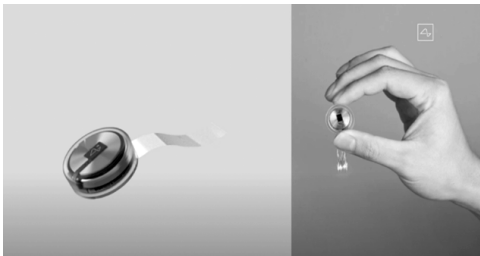


图 3-3 Neuralink 设备

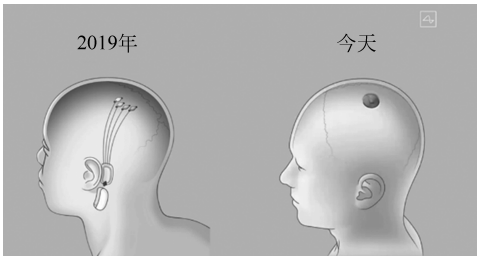


图 3-4 Neuralink 设备的安装位置

2. 新一代 Neuralink 机器人

Neuralink 机器人(图 3-5)是一个“神经外科手术机器人”,负责植入 Neuralink 设备,据说它每分钟可以将 192 个电极插入大脑。马斯克表示,植入过程可以在 1 个小时内完成,不需要全身麻醉。



图 3-5 Neuralink 机器人

直播现场展示了小猪格特鲁德的脑部活动信号可以被实时读取。当猪的鼻子触碰到物体时,实时无线传输的图像上出现了噪音(图 3-6)。

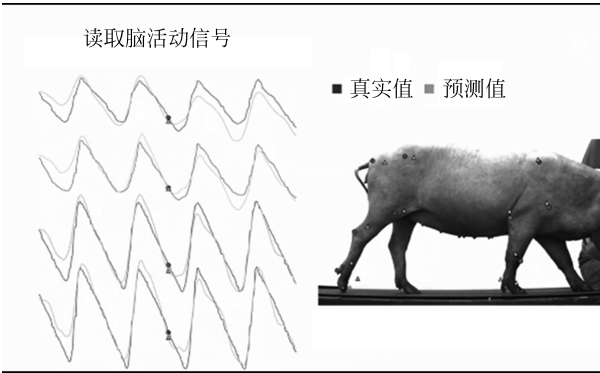


图 3-6 Neuralink 机器人读取的大脑活动信号

马斯克声称,Neuralink 的最终潜力几乎是无限的。例如,可以用心灵感应召唤一辆自动驾驶中的特斯拉,可以解决失明、瘫痪、听力障碍等问题。

3. 人脑与机器对接技术

这次直播活动意味着人类在利用植入设备治疗记忆力衰退、颈脊髓损伤、中风和成瘾等疾病方面又迈进了一步。患有神经疾病的人,未来有可能通过脑机连接技术来控制智能手机或计算机。马斯克的长远目标是迎接“超人认知”时代的到来,他认为人类需要与人工智能结合为一体,以避免未来 AI 变得过于强大,以致摧毁人类这一最糟情况的出现。

2019 年 7 月 17 日,Neuralink 首次对外宣布一款脑机接口系统,其原理是用长得像

缝纫机一样的机器人向大脑中植入超细柔性电极来监测神经元活动。系统包含一个微型探头,上面的 3000 多个电极与比头发丝还细的柔性细丝相连。Neuralink 定制了一款微小芯片,可以通过有线连接方式传输数据。据 Neuralink 称,该装置一次可以监测 1000 多个神经元的活动。在计算机视觉软件的帮助下,手术机器人可以避免撞击血管,减少大脑损伤和疤痕组织的形成。

Neuralink 曾对一只猴子进行过脑机对接试验,试验表明猴子可以通过大脑来控制电脑。Neuralink 寻求美国食品药品监督管理局批准的人类临床试验,其目标是将电极植入因脊髓上部损伤而完全瘫痪的患者的头骨上,通过脑机接口将大脑信号传递给植入在耳后的一个小装置,再将数据传输到计算机,使患者能够用意念来控制智能手机和计算机。

4. 任重道远

脑机接口技术被称作人脑与外界沟通交流的“信息高速公路”,它为未来恢复感觉和运动功能以及治疗神经疾病提供了希望。此外,它还能让人类大脑“升级”,也会使人类在面临未来 AI 威胁时更具竞争力。Neuralink 拥有一个由科学家、工程师以及临床医生组成的团队。

美国匹兹堡大学物理医学与康复学助理教授科林格形容,马斯克试图做的是在医疗技术这一困境领域中开展真正的“颠覆性创新”。宾夕法尼亚大学的本杰明表示,该技术真正的难点可能取决于人脑的复杂性,关键是如何对脑机对接后所记录下来的信息进行解码。

阅读上文,请思考、分析并简单记录:

(1) “将芯片植入猪脑,实时读取信息”的 Neuralink 设备由哪几部分组成? 各自的功能是什么?

答: _____

(2) 有人把马斯克形容为科技怪侠。马斯克身上有什么值得你羡慕或学习的吗?

☐ 富裕的长辈 ☐ 创新的头脑 ☐ 赚钱的本事 ☐ 超高的情商

分析: _____

(3) “大脑植入芯片技术”的前景如何? 会有哪些主要应用?

答: _____

(4) 请简单记述你所知道的上一周内发生的国际、国内或者身边的大事。

答: _____

3.1 界面的分类

计算机系统的人机交互功能主要依靠输入输出设备和相应的软件来完成,可供人机交互使用的设备主要有键盘、显示器、鼠标和各种模式识别设备等。模式识别,如语音识别、汉字识别等输入设备的发展,使操作者和计算机在类自然语言层级上进行交互成为可能。此外,通过图形进行的智能化人机交互也吸引着人们进行研究。

虽然通用的 QWERTY 键盘布局目前仍是主要的文本输入设备,但新的键盘策略已经出现,从而满足移动设备用户的需要;指点设备,特别是鼠标和触摸屏,使用户脱离键盘来完成任务;未来的计算可能包含更多的手势输入、三维指点、语音输入输出设备、可穿戴设备、多模式设备等,也许人的全身都会参与到某些输入输出任务之中。

一些研究者为植入设备提出了更加奇异的想法。创新的输入设备、传感器和效应器,还有把计算机集成到物理环境当中,都向各种应用敞开了大门。持续改进的语音识别器已经加入了较平凡但使用广泛的语音存储转发技术,更加强调基于电话的应用系统和非语音听觉界面。

智能手机应用产业无处不在,而网络的服务、内容、资源和信息也在不断涌现。一个主要问题是如何设计,以使人们能在不同设备和浏览器之间实现互操作,这些设备包括具有不同组成因素、尺寸和形状的智能手表、智能手机、笔记本电脑、智能电视以及计算机屏幕等。

技术的飞速发展鼓励人们从多角度对交互设计和用户体验进行思考。例如,输入可以通过鼠标、触摸板、笔、遥控器、操纵杆、RFID 阅读器、手势、多模式,甚至人脑-计算机交互进行。输出形式同样是多样化的,如图形界面、语音、混合现实、增强现实、触觉界面、可穿戴计算设备等。

3.2 实用性界面

所谓实用性界面,包括命令行界面、图形用户界面、移动设备、触觉设备、可穿戴计算和智能界面等。

3.2.1 命令行界面

早期的界面要求用户在计算机显示的提示符处输入缩写命令(如 1s),系统会对其做

出响应(如列出当前文件)。另一种输入命令的方式是按组合键(如 Shift+Alt+Ctrl),有些命令基于键盘的固定按键,如删除、输入和撤销,而其他功能键可依据用户的特定命令来设置(如按 F11 键控制打印)。

如今,命令行界面在很大程度上已经被图形界面取代,后者将菜单、图标、键盘快捷键和弹出/可预测文本命令等命令合并为应用程序的一部分。但命令行界面仍然有优点,一些系统管理员、程序员和高级用户认为命令语言更高效、更快捷。例如,使用一个命令一次性删除 10 000 个文件,比滚动浏览该数量的文件并选中再删除要容易得多。

3.2.2 图形用户界面

XeroxStar(施乐之星)界面推动了图形用户界面(GUI)的诞生。最初的 GUI 叫作 WIMP(即窗口、图标、菜单、鼠标指针),第一代 WIMP 界面主要采用盒状设计。用户交互发生在窗口、滚动条、复选框、面板、调色板和以各种形式出现在屏幕的对话框中。如今,GUI 已适用于移动设备和触屏设备,大多数用户的默认动作是在浏览和交互时使用单手指滑动和触摸,而不是使用鼠标和键盘。

WIMP 的基本构建单元仍然是现在 GUI 的一部分,并作为界面显示的一部分,且已经演变成多种不同的形式和类型,如音频图标和音频菜单、3D 动画图标,以及可以放入智能手表屏幕的基于微小图标的菜单。此外,窗口也大大扩展了使用方式和用途,如各种对话框、交互式菜单和反馈/错误消息框已经变得普遍。

窗口的发明克服了计算机显示器的物理限制,使用户能够在统一屏幕上观看更多的信息,并执行任务。用户可以随时打开多个窗口,如网页、文档、照片和幻灯片,在需要查看或处理不同的文档、文件和应用程序时切换。在一个应用程序下也可以打开多个窗口,如 Web 浏览器。

窗口中垂直或水平放置的滚动条使用户可以查看超过一个屏幕范围内的更多信息,使文档向上、向下或向侧面移动,并可以通过触摸板、鼠标或方向键控制滚动条的移动。可触摸的屏幕让用户可以简单地通过滑动来达到滚动条的效果。

图形界面中最常用的特定窗口是对话框,基本上所有的对话、信息、错误、清单和表单都通过它们来呈现。对话框中的信息通常被设计用于指导用户交互,用户遵循对话框所提供的一系列选项来操作。

3.2.3 移动设备

移动设备已经普及,人们越来越多地在日常生活和工作的各个方面使用它们——如手机、手环或手表。此外,人们会在不同环境中使用定制的移动设备,如在餐馆里点餐,在超市中检查库存以及用于多人游戏的移动设备等。许多航空公司为乘务员提供平板电脑,以便他们可以在空中和机场使用定制的航班应用程序。

智能手机和智能手表中嵌入了各种传感器,如用于检测运动速度的加速度计、用于测量温度的温度计,以及用于测量人体皮肤上汗液水平变化的皮肤电反应计,也有些应用程

序可能只是为了好玩。如早期由魔术师史蒂夫·喜来登开发的一个有趣的应用程序 iBeer(图 3-7),其成功的一部分原因在于巧妙地使用了手机内部的加速度计。它检测 iPhone 手机的倾斜度,模拟正在不断减少的一杯啤酒,其中的啤酒颜色及泡沫还有声音效果,给人一种啤酒在玻璃杯中晃动的错觉。如果手机足够倾斜,啤酒会被喝完,然后发出打嗝声。

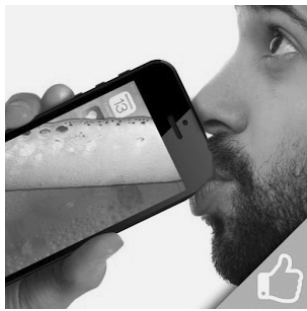


图 3-7 智能手机应用程序 iBeer

智能手机还可以用于通过扫描条形码来下载语境信息。消费者可以在逛超市时使用手机扫描条形码,下载产品信息。另一种提供快速访问信息的方法是使用存储 URL 的二维码,手机扫描二维码后,会将用户带到特定的网站。

移动界面通常使用小屏幕和有限的控制空间。设计人员必须仔细考虑包含什么类型的专用硬件控件,将它们放在什么位置,以及如何将它们映射到软件上。为移动界面设计的应用程序需要考虑内容导航能力,因为使用移动显示器时呈现的内容是有限的。人们还开发了许多移动浏览器,允许用户以更简化的方式浏览互联网、杂志或其他媒体。

3.2.4 触觉设备

触觉界面通过使用嵌入用户衣服或佩戴设备的振动器向人体提供振动反馈。游戏机也采用振动来提供丰富的体验。例如,驾驶模拟器的汽车方向盘可以通过各种方式的振动提供在道路上行驶的感觉。当驾驶员转弯时,用户可以感受到方向盘旋转的阻力——就像真正的方向盘一样。触觉振动反馈也可用于模拟远程人员沟通时的触觉传递。嵌入衣服中的振动器可以通过在身体的不同部位产生不同的力来重现拥抱或挤压的感觉。

另一种形式的反馈称为超触觉(图 3-8),即在空中创造出触觉的幻觉。它通过使用超声波来制造用户可以感觉但看不到的三维形状和纹理。这种技术可使用户感受到出现在空中的按钮和滑块的错觉。在汽车行业,超触觉的一个潜在用途是替代现有的物理按



图 3-8 超触觉

钮、旋钮和触摸屏。通过设计,超触觉按钮和旋钮可以在需要的时候出现在驾驶员旁边,如系统检测到驾驶员想要调低音量或切换无线电台时。触觉反馈也被嵌入衣服,有时这被称为外骨骼。

不同种类的振动会传递不同类型的触觉体验。触觉反馈的关键设计问题是如何找到振动器在身体中的最佳放置位置,应使用单点振动还是多点振动,什么时候振动,以及什么样的振动强度和频率可以使振动更具有说服力。

3.2.5 可穿戴计算

可穿戴计算泛指可以穿戴在身体上的设备,包括智能手表、健身追踪器、时尚科技穿戴和智能眼镜。新的柔性显示技术、电子纺织品等让人们想象中的可穿戴物品变成了现实。珠宝、帽子、眼镜、鞋子和夹克都是实验的主题,旨在为用户提供在现实世界中移动时与数字信息交互的方法。早期的可穿戴设备专注于便利性,人们无须取出和控制手持设备即可执行任务(如选择音乐)。如带有集成音乐播放器控件的滑雪夹克,穿戴者只需用手套触摸手臂上的按钮,即可更换音乐曲目。还有一些应用主要关注如何结合纺织品、电子产品和触觉技术,来创造新的通信形式。例如,有研究者开发了一款内嵌传感器的运动服,用来捕捉穿戴者的动作和与他人的互动,然后通过覆盖在裙子外部的电致发光刺绣来展示。它会根据穿戴者的运动量和速度改变模式,向别人展示穿戴者的心情,并在其周围创造一个神奇的光环。

外骨骼服装(图 3-9)也是一个将时尚与技术相结合的例子。它结合了触觉与可穿戴设备,可以帮助走路困难的人行走或帮助人们锻炼。在建筑行业,外骨骼服装帮助工人提供额外的动力——这使他们有点像超人——其金属框架上安装了机械肌肉,能增加穿着者的力量。重的物体因此感觉更轻,使人免受一定的身体伤害。



(a) 外骨骼服装



(b) 人类穿戴外骨骼服装

图 3-9 带有触觉反馈的外骨骼服装

2014 年开始发售的谷歌眼镜(图 3-10)是一种可穿戴设备,它具有各种时尚的风格。谷歌眼镜外表看起来像一副眼镜,但其中一个镜片是带有嵌入式摄像头的交互式显示器,可以通过语音输入进行控制。佩戴者可以通过它在移动中拍摄照片和视频,并查看如电子邮件、文本和地图等数字内容。佩戴者还可以使用语音命令完成网络搜索,其结果将出

现在屏幕上。除了日常功能之外,它还有很多额外的功能。

然而,很多人认为当和戴着谷歌眼镜的人在一起时,佩戴者会抬眼看向右边的屏幕,而不是看着他们的眼睛,这让他们感到不安。还有人担心戴着谷歌眼镜的人正在记录他们面前发生的一切。作为回应,美国的一些酒吧和餐馆甚至实施了“禁止使用谷歌眼镜”的政策。第一代谷歌眼镜在几年后就下市了。



图 3-10 谷歌眼镜

可穿戴设备的一个核心设计问题是舒适性。嵌入了技术的衣服同样需要让用户保持舒适。嵌入设备需要轻便、小巧、时尚,而且(除了显示器外)最好藏在衣服里。另一个问题是卫生。穿过的衣服可以清洗吗?拆下和更换电子装置会很麻烦吗?电池应该放在哪儿?其寿命有多长?一个关键的可用性问题是用户如何控制这些可穿戴设备,是通过触摸、语音还是更传统的按钮和刻度盘。可穿戴设备可以和更多的技术组合,包括LED、传感器、振动器、实体交互和AR。

3.2.6 智能界面

许多新技术的动机是让设备更加智能,无论是智能手机、智能手表、智能建筑、智能家居,还是智能家电。更宽泛地说,智能设备可以与用户和其他联网设备进行交互,其中许多是自动化的,不需要用户与它们直接交互。智能的目标是感知情境,也就是说,根据周围的情境做出适当的操作。为了实现这一目标,一些设备使用了人工智能技术,这样它们就可以学习环境和用户的行为。这种智能技术可以根据用户的偏好更改设置或控制开关。

智能建筑变得更加节能、高效、低成本。建筑师使用最先进的传感器技术来控制建筑系统,如控制通风、照明、安全和供暖功能。虽然智能建筑和智能家居改善了管理方式,但它们也会让用户感到沮丧,因为用户有时希望窗户能够打开,新鲜空气和阳光能进来。但是把人排除在自动化系统之外意味着人不再能决定这些操作。相比简单地引入自动化,并将人类排除出自动化系统,另一种方法是在考虑居民需要的同时引入智能技术。例如,该领域的一个新方法称为“人-建筑交互”(HBI),它关注的是理解与塑造人们在建筑环境中的体验。在解决人与“智能”环境交互问题的过程中,它关注的是人的价值、需求和优先级。

3.3 输入输出界面

目前,输入文本数据的主要方式仍然是键盘。指点设备已经历了数百次改进,以适应不同用户,并做出进一步的性能改进。更不寻常的设备,包括眼球跟踪器、数据手套和触觉或力反馈装置已经应用于特定的应用,如远程医疗。

3.3.1 键盘

很多移动设备,如苹果的 iPhone,已经完全放弃了机械键盘,而依赖在触摸屏上的指点、绘图和手势进行所有交互(图 3-11)。如果屏幕大到足以显示一个键盘,用户就能轻敲虚拟键盘。在对 7cm 和 25cm 宽的触摸屏键盘的使用中,用户经过培训后每分钟能输入 20~30 个英文单词,在输入的文本长度有限时,这个速度是可接受的。

另一种方法是在触敏表面上手写输入,通常使用输入笔,但字符识别仍容易出错。使用上下文线索、为击打速度加上方向能够提高识别率,而成功的手势数据输入方法的识别效果也相当好,且大多数用户很快就能学会编码,但培训对于新用户和间歇用户可能是个障碍。另一个有前途的方法是允许使用与轻击模式匹配的形状,用键盘上的速记手势来替代触摸屏键盘上的轻击。长期的研究确认了使用这种技术实现良好的文本输入性能的可能性。对于汉语,手写体识别技术戏剧性地增加了用户的潜能。



图 3-11 苹果 iPhone 的虚拟键盘

3.3.2 指点设备

对于复杂的信息显示,诸如计算机辅助设计工具、绘图工具或空中交通管制系统中的信息显示,指点和选择项通常是方便的。这种直接操纵方法之所以有吸引力,是因为用户能够避免学习命令,减少在键盘上的打字出错率,把注意力集中在显示效果上。其结果是执行得更快、错误更少、学习更容易和满意度更高。指点设备对小设备和大的墙面显示设备也是重要的,因为这些设备使得键盘交互不太实用。

多种多样的任务、各种各样的设备和使用它们的策略创造了丰富的设计空间。物理设备的属性(旋转或线性移动),移动的维数和定位(相对的或绝对的)是给设备分类的有用方式。

指点设备可以分为提供屏幕表面直接控制(如光笔、触摸屏或输入笔)和提供脱离屏幕表面的间接控制(如鼠标、轨迹球、操纵杆、指点杆、图形输入板、触摸板或数字纸)两类。用于专门用途的非标准设备和策略包括多点触控板和显示器、双手输入、眼球跟踪器、传

感器、3D跟踪器、数据手套、触觉反馈、脚踏控制和有形用户界面等。指点设备的成功标准是速度和精确性、任务的功效、学习时间、成本和可靠性、大小和质量。

眼球跟踪器是一种凝视检测控制器,使用瞳孔位置摄像机进行图像识别(图3-12)。200~600ms的定影时间用于做出选择。遗憾的是,每次凝视都有激活非有意命令的可能性。因此,需要把眼动跟踪与手动输入相结合,以处理此问题。目前,眼动跟踪主要仍是研究和评估工具,是一种运动残疾用户的可能辅助工具。

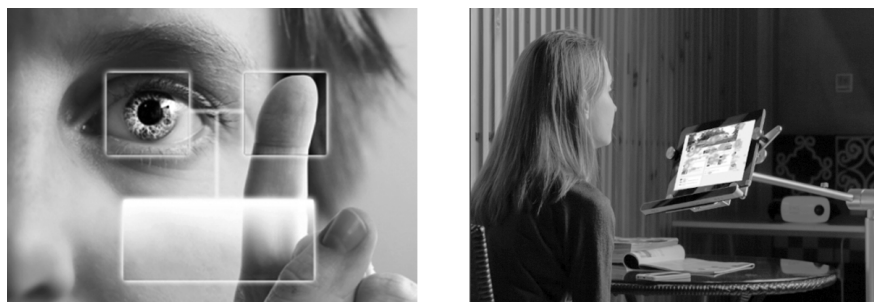


图 3-12 眼动跟踪器

3.3.3 显示器

显示器是从计算机到用户的主要反馈源,它具有很多重要特征,包括:

- (1) 物理尺寸(通常是对角线尺寸和深度)。
- (2) 分辨率(可用像素数)。
- (3) 可用颜色数和颜色的正确性。
- (4) 亮度、对比度和眩光。
- (5) 能耗。
- (6) 刷新率(足以允许动画和视频)。
- (7) 价格。
- (8) 可靠性。

采用的显示器尺寸是设计中需要的特殊策略。在数码相机的小液晶显示屏上的即时查看功能以及带触摸屏的移动电话,都已经是成功的应用案例,而墙面大小的高分辨率显示器也在创造着新的机会。如今,除了改进单个输入输出设备,多模态界面也做了一些工作,这种界面把若干个输入输出方式结合起来。研究人员最初相信,同时使用多种方式可以改进性能,但这些方法的应用系统数量还很有限。也存在着同步多模态界面的成功例子,如把语音命令与对于对象应用动作的指点结合起来。然而,更大的回报似乎是给予用户按需的方式之间切换的能力。例如,允许司机通过触摸动作或语音输入来操作导航系统。多模态界面的开发将使残疾用户受益,他们可能需要视频字幕、音频转录或图像描述。多模态界面的进步将有助于实现普遍可用性的目标。

另一个活跃的研究方向是情境感知计算。移动设备能够使用来自全球定位系统的卫

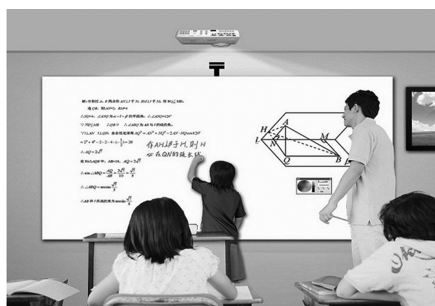
星、手机、无线连接或其他传感器的位置信息。这类信息允许用户接收附近的饭店或加油站的信息,使博物馆参观者或游客能够访问关于他们周围环境的详细信息。

按使用特征也能区分显示设备。可移动性、私密性、显著性(需要吸引注意力)、普适性(能够放置和使用显示器的可能性)和同时性(同时使用的用户数)能够用于描述显示器。

较简单的数字白板系统(白板显示器)允许协作者共享信息、进行头脑风暴和做出决策(图 3-13)。和台式机一样,白板使用用户的手指作为指点设备,还有彩色笔和数字橡皮擦,并增加了注释记录和软键盘。



(a) 使用数字白板场景1



(b) 使用数字白板场景2

图 3-13 数字白板的使用

个人显示技术包括小的便携式监视器,通常用黑白或彩色 LCD 制作。例如,抬头显示器把信息投射到部分镀银的飞机或汽车的挡风玻璃上,以便飞行员或驾驶员在接受计算机产生的信息的同时能够把注意力集中于周围。

另一选择是虚拟现实或增强现实应用系统中使用的头盔或头戴式显示器(图 3-14),这种显示器让用户甚至在转头时也能看到信息。实际上,如果该显示器配备了跟踪传感器,就能为用户提供不同级别的视野、音频性能和分辨率。可穿戴计算机的早期例子关注于小的便携式设备,人们能够在移动或完成其他任务时使用这种设备,诸如喷气发动机修理或库存控制,但当前的技术仍要求硬件在背包里或用户待在基础计算机附近。

产生 3D 显示器的尝试包括振动表面、全息图、偏振眼镜、红/蓝眼镜和同步的快门眼镜,给予用户强烈的 3D 立体视觉感。



图 3-14 头戴式显示器

3.3.4 笔设备

纸也能用作输入设备。基于笔的设备能够在纸面上书写,绘制、选择、移动对象,支持手绘草图(图 3-15),这充分利用了人们已经养成的良好的绘画和书写技能。早期的应用

证明了用摄像机捕获蓝图或实验记录册等大文档上注解的好处。



图 3-15 写在纸面上的墨水笔画,其数据被无线传输给计算机

数字墨水(如 Anoto,图 3-16)使用普通墨水笔和数码相机相组合,可以记录在特殊纸张上书写的所有内容。这种笔将一个小照相机装在笔尖上,记录写在特殊纸上的笔画,通过识别打印在纸张上的特殊非重复圆点图案来工作。图案的非重复性意味着笔能够确定正在写入哪个页面,以及笔指向的是页面上的哪个位置。使用数字笔在数字纸上书写时,笔中的红外光会照亮圆点图案,然后由微型传感器拾取。当笔在纸上移动时,笔会对点图案进行解码,并将数据临时存储在笔中。可以通过蓝牙或 USB 端口将存储在数字笔中的数据传输到计算机,因此手写笔记也可以转换并保存为标准字体文本。

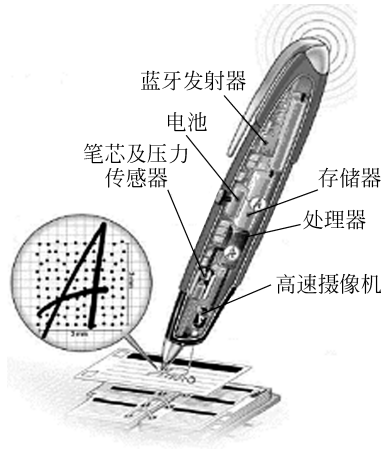


图 3-16 Anoto 笔及其内部组件示意图

数字笔的另一个优点是允许用户通过或使用纸质材料相同的方式快速并轻松地注释现有文档(如电子表格、演示文稿和图表)。这对于成员处在不同地点的团队来说非常有用。但是在小屏幕上使用基于笔的交互的一个问题是,有时在屏幕上阅读选项可能是困难的,因为书写时用户的手可能会遮挡其中一部分内容。

数字墨水和电子墨水(e-ink)不一样。电子墨水是一种用于电子阅读器(如 Kindle)的显示技术,旨在模仿纸上普通墨水的外观。这种显示器的反射效果就像普通的纸一样。

3.3.5 触摸屏

单点触摸屏已经有了很多应用,多用于自助服务终端(如售票机、博物馆导游),ATM 和排号机器。它们通过检测人在显示器上触摸的存在和位置来工作,人们通过单击屏幕选择选项。此外,多点触摸屏支持一系列更动态的指尖动作,如滑动、轻击、捏合、推动和敲击,通过栅格系统在多个位置定位触摸来实现这些功能(图 3-17)。这种多点触控方法

使智能手机和桌面等设备能够同时识别和响应多个触摸,使用户可以使用多个手指执行各种操作,如放大和缩小地图、移动照片、在写作时从虚拟键盘中选择字母以及滚动列表。也可以使用两只手在桌面上拉伸和移动物体。



图 3-17 触摸屏及其应用

手指手势所带来的交互灵活性产生了许多数字内容的体验方式,包括阅读、浏览、缩放和搜索平板电脑上的交互式内容,以及创建新的数字内容。

触摸屏与 GUI 不同,其关键设计问题是如何最好地使用不同类型的交互技术来支持不同场景下的活动。在多点触控界面上使用双手操作能够放大、缩小或旋转数字对象。长按并保持手指始终按在屏幕上能够执行拖动对象操作或调出菜单。一个或多个手指也可以与长按动作一起使用,以提供更广泛使用的手势操作。

3.4 功能性界面

功能性界面包括多媒体界面、基于手势系统、语音界面、虚拟现实、增强现实、机器人和无人机、多模式界面和脑机交互等方面。

3.4.1 多媒体界面

顾名思义,多媒体是在单个界面中组合不同的媒体,如图形、文本、视频、声音和动画,并将它们与各种形式的交互相连接。用户可以单击图像或文本中的链接,从而触发动画或视频等其他媒体。人们假设,与单个媒体相比,媒体和交互性的组合可以提供更好的呈现信息的方式,比如文字与视频结合会产生一加一大于二的效果。多媒体的附加价值在于它更容易学习和理解,更吸引人并令人愉快。

多媒体的一个显著特征是其快速访问多个信息的能力。一些多媒体百科全书和数字图书馆基于这种多样性原则设计,为给定主题提供各种音频和视频材料。例如,如果了解心脏,一个典型的多媒体百科全书将为你提供以下信息。

(1) 一个或多个真正的活的的心脏泵送或心脏移植手术的视频剪辑。

- (2) 心脏跳动的录音,也许还有一位著名的医生谈论心脏病的病因的录音。
- (3) 循环系统的静态图和动画,有时还带有叙述。
- (4) 几列超文本,描述心脏的结构和功能。

交互式模拟已经成为多媒体学习环境的一部分。一个早期的例子是向学生演示心脏复苏术,要求学生从计算机屏幕显示项中选择正确的选项,并以正确的顺序设置程序来救治病人。此外,还有其他类型的多媒体叙事和游戏,通过热点或其他类型的链接来引起学生的注意,并鼓励他们在显示屏上操作来实现探索学习。多媒体在很大程度上是为培训、教育和娱乐目的开发的。

3.4.2 基于手势系统

手势涉及移动手臂和手进行交流(如挥手告别或在课堂上举手发言),或向别人传递信息(如两手张开以表示某物的大小)。通过使用相机跟踪手势,然后使用机器学习算法进行分析,人们对使用技术来捕获和识别用户的手势进行了很多尝试。

大卫·罗斯创作了一个视频,描绘在各种场景中使用手势的一些灵感来源,包括由板球裁判员、音乐会中为聋人准备的手势示意者、说唱歌手、查理·卓别林、哑剧艺术家和意大利人制作的手势。他的团队开发了一个手势系统来识别一小部分手势,他们发现手势需要由“名词、动词和对象及对其的操作”这种特定的顺序组成才能被理解。例如,为了表达“扬声器,开启”,要使用一只手的手势来指定名词,用另一只手的手势指定动词。因此,如果要改变音量,用户需要用左手指向扬声器,同时抬起右手,以指示音量调高。

手势交互的一个应用领域是手术室(图 3-18)。外科医生需要在手术期间保持双手无菌,但也需要能够在手术期间观看 X 射线和扫描结果。然而,清洗手部和戴手套后,他们需要避免用手指接触任何键盘、手机和其他可能有菌的表面。为此开发了一种基于手势的系统,使用微软的 Kinect 技术可以识别外科医生通过手势进行的交互和操作,其手势包括用于向前或向后移动图像的单手手势,以及用于缩放和平移的双手手势。



图 3-18 手术室

使用手势输入的关键设计问题是计算机系统如何识别和描述用户的手势。特别地,

如何确定手或手臂运动的起点和终点,以及如何区分有意的手势(经思考的指向动作)和无意挥手之间的差别。

3.4.3 语音界面

语音用户界面(VUI)涉及与口语应用程序交谈,如搜索引擎、火车时刻表、旅行规划器或电话服务。它通常用于查询特定信息(如航班时间或天气)或向机器发出命令(如要求智能电视选择某一部动作电影或要求智能扬声器播放欢快的音乐)。因此,VUI是命令或对话类型的交互,其中用户通过听和说而不是单击或触摸与界面交互。有时,系统会主动提问,而用户只需要做出回答,如询问用户是否想要停止观看电影或收听最新的突发新闻。

语音系统现在变得更加复杂,并且具有更高的识别准确度。机器学习算法不断提高其识别说话内容的能力。对于语音输出,一些演员通常会为答案、信息或提示配音,这些通常比早期系统中使用的人工合成语音更友好、更有说服力且更令人愉快。

有许多基于语音的手机应用程序,使人们可以在移动的时候使用它们。例如,用户可以使用谷歌语音助手或苹果的 Siri 向手机说出自己想要查询的内容,而不必手动输入文字。移动翻译软件让人们在说话的同时利用手机上的应用程序(如谷歌翻译)进行翻译,这样人们可以与使用不同语言的人实时交流。人们对着手机说自己的语言,而另一个人将会听到软件翻译后的语言。从某种意义上说,这意味着世界各地的人可以彼此交谈,而不必学习母语以外的其他语言。

3.4.4 虚拟现实

虚拟现实(Virtual Reality,VR)(图 3-19)也称灵境技术或人工环境,自 20 世纪 70 年代左右开始出现,是利用电脑模拟产生一个三度空间的虚拟世界,提供给使用者关于视觉、听觉、触觉等感官的模拟,让使用者如同身历其境一般,可以及时、没有限制地观察三度空间内的事物。使用者进行位置移动时,电脑立即进行复杂的运算,将精确的 3D 影像传回,产生临场感。



图 3-19 虚拟现实

3D图形可以投影到CAVE(“洞穴”自动虚拟环境)地板和墙壁表面、桌面、3D电视、头盔或共享显示器(如IMAX屏幕)上。VR的主要吸引力之一,是可以为新的身临其境的体验提供机会,使用户能够与对象交互,并在3D空间中导航,这在物理世界或2D界面中是不可能的。用户除了被360°虚拟世界环绕,还能感受到声音和触觉反馈,由此产生的体验是高度沉浸的,让人感觉是真的在飞行。这里的存在意味着“意识状态,在虚拟环境中的(心理)感觉”,其中某人的行动方式与其在等效的真实事件中的行动方式类似。

早期的VR是使用头戴式显示器开发的,现在有了很多舒适、便宜、更加精准的VR头戴设备。开发人员创建更多具有吸引力的游戏、电影和虚拟环境,开发了支持许多技能的学习和培训的VR,设计了一系列应用程序,旨在帮助人们学习驾驶车辆或飞机,或执行精细的外科手术。

3.4.5 增强现实

随着2016年游戏《精灵宝可梦 Go》的问世,增强现实(Augmented Reality, AR)一举成名。AR是通过计算机系统提供的信息增加用户对现实世界感知的技术,将虚拟的信息应用到真实世界,并将计算机生成的虚拟物体、场景或系统提示信息叠加到真实场景中,从而实现对现实的增强。在视觉化的增强现实中,用户利用头盔显示器,把真实世界与电脑图形多重合成在一起,便可以看到真实的世界围绕着它(图3-20)。



图 3-20 增强现实

与AR密切相关的是混合现实,即将现实世界的视角与虚拟环境的视角结合在一起。AR起初主要是医学中的一个实验主题,其中虚拟物体(如X射线和扫描)叠加在患者身体的某部分之上,帮助医生理解正在检查或操作的内容。后来,人们利用AR帮助控制员和操作员快速做出决策。如在空中交通管制中,管制员能看到系统提供的飞机的动态信息,这些信息叠加在显示真实飞机着陆、起飞和滑行的视频屏幕上。这些附加信息使管制员能够轻松识别难以辨认的飞机——这在恶劣天气条件下特别有用。同样,平视显示器(HUD)用于军用和民用飞机,以便在恶劣天气中帮助飞行员着陆。HUD在折叠显示器上提供电子方向标记,且直接出现在飞行员的视野中。许多高端汽车提供具有AR技术

的挡风玻璃,其中导航就像真实地出现在路面上一样。

AR 技术同样取代了建造或修理复杂设备(如复印机和汽车发动机)的纸质手册,它直接把图纸叠加在机器上,告诉机械师该做什么以及在哪里做。AR 应用程序可用于从教育到汽车导航的各种环境,其数字内容直接叠加在实体地理位置和对象上。为了显示数字信息,用户可以在智能手机或平板电脑上打开 AR 应用程序,内容就会直接叠加在当前屏幕显示器上。

多数 AR 应用程序使用智能手机或平板电脑上的后置摄像头,然后将虚拟内容叠加在其拍摄的现实世界中。另一种方法是使用前置摄像头,将数字内容叠加到用户的面部或身体上。零售行业通过 AR 镜子可以让购物者“试用”太阳镜、珠宝和化妆品,其目的是让他们尽可能“试用”更多的产品,看看它们用在自己身上是什么样子。显然,这种虚拟试用有很多优势:与真实的试用相比,虚拟试用更方便、更吸引人、更容易。但是,它也有缺点:你只能看到自己试用它们是什么样子,但无法感受到头上虚拟配件的重量,也无法感受到脸上虚拟化妆品的质感。

3.4.6 机器人和无人机

机器人已经存在了很长时间,它们常常作为科幻小说中的角色出现,但也在其他方面起着重要的作用:作为制造装配生产线的一部分、作为危险环境下的远程调查员(如在核电站和拆弹环境中)、作为灾害(如火灾)或远程(如火星)调查和搜救人员。研究人员开发了控制台界面,使得人们能够使用操纵杆、键盘、摄像机和基于传感器的交互组合控制和导航偏远地形中的机器人。其中的重点是界面设计,能帮助用户通过实时视频和动态地图有效地操纵和移动远程机器人。无人机是远程控制的无人驾驶飞机。它首先由爱好者使用,然后被军队利用。后来它们变得更便宜、更大众化、更容易飞行,因此得以在更广泛的背景下应用。

3.4.7 多模式界面

多模式界面旨在通过使用不同的模式(如触摸、视觉、声音和语音)增加用户体验和控制信息的方式,从而丰富用户体验。为此,组合的交互技术包括语音和手势、眼睛注视和手势、触觉和音频输出,以及笔输入和语音。假设多模式界面可以使人机交互方式更灵活、更有效且更富有表现力,这种人机交互方式更类似于人类在物理世界中遇到的多模式体验。不同的输入输出方式会同时应用,例如,同时使用语音命令和手势在虚拟环境中移动,或者先使用语音命令,接着进行手势操作。用于多模式界面的最常见的技术组合是语音和视觉处理的组合。多模式界面还可以与多传感器输入组合,以追踪人体其他方面的数据。例如,通过追踪眼睛注视、面部表情和嘴唇的运动得到有关用户的注意力或其他行为的数据。这种方法可以根据感知到的需求、想法或兴趣级别来为定制用户界面和体验提供输入。

多模式系统依赖于识别用户行为的各个方面,包括手写、语音、手势、眼睛运动或其他

身体运动。在许多方面,多模式系统都比单模式系统更难以实现和校准,因为后者仅识别用户行为的单个方面。当今研究最多的交互模式是语音、手势和眼睛注视追踪。关键的研究问题是,将不同的输入和输出组合在一起,最终得到的究竟是什么,以及把人之间的交谈和手势交流作为与电脑的交互方式是否是自然的。

3.4.8 脑机交互

脑机交互提供了人的脑电波与外部设备(如屏幕上的光标或通过气流移动的冰球)之间的通信通道。一些项目研究了这种技术如何帮助增强人类的认知或感觉运动功能。脑机交互的工作方式是检测大脑神经功能的变化。树突和轴突相互连接成单个神经细胞,大脑充满了由这些神经细胞组成的神经元。每当人类思考、移动、感觉或记忆某些事物时,这些神经元就会变得活跃。小的电信号从一个神经元快速地传到另一个神经元,放置在人头皮上的电极在一定程度上可以探测到这种变化。这些电极可以被嵌入专门的耳机、发网或帽子中。

脑机交互也可应用于游戏控制。如通过脑机交互控制机器人和驾驶虚拟飞机。布朗大学的 BrainGate 研究小组进行了开创性医学研究,即使用脑机交互界面使瘫痪者能够通过脑机交互控制机器手臂自己进食(图 3-21)。

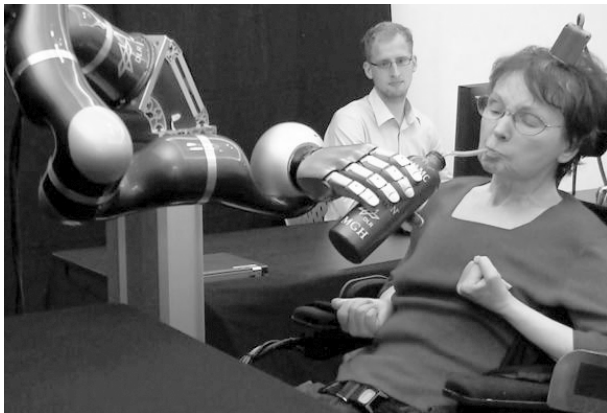


图 3-21 通过脑机交互控制机器手臂自己进食

3.5 平台性界面

平台性界面包括家用电器、可共享界面、实体用户界面以及自然用户界面。

3.5.1 家用电器

家用电器包括家中日常使用的机器(如洗衣机、微波炉、冰箱等)。大多数使用它们的人都会尝试在短时间内完成特定的操作,如启动洗衣机、看节目、买票或做饮料。他们不

太可能有兴趣花时间探索其界面或仔细翻阅学习手册后使用设备。现在很多家用电器都有 LED 显示屏,提供多种功能和反馈(如温度、剩余时间等)。其中一些可以连接到互联网,用户能够通过远程应用程序控制。

设计者需要将设备界面视为瞬态界面,交互时间较短。然而,设计人员常常提供全屏控制面板或不必要的物理按钮阵列,这会让用户感到沮丧和困惑,而只包含少数且结构化呈现的按钮界面会更好。在这里,简单性和可见性这两个基本设计原则是至关重要的。状态信息(如复印机正在做什么,售票机正在做什么以及要花多长时间清洗)应该以非常简单的形式在界面的显著位置展示。

3.5.2 可共享界面

可共享界面是为多人使用设备而设计的。与面向单个用户的个人计算机、笔记本电脑和移动设备不同,可共享界面通常提供多个输入,有时允许一个群组同时输入。具有这种界面的设备包括大型的墙壁显示器,人们可以使用笔或手势、交互式桌面,进行信息交互。交互式桌面可以区分同时触摸表面的不同用户。触摸表面下面嵌入一组天线,每个天线都发送一个独特的信号。每个用户都有自己的接收器,它们被嵌入用户坐着的垫子或椅子中。当用户触摸桌面时,接收器会识别其中的微小信号,以识别出被触摸的天线,并将其发送到计算机。因此,多用户可以使用手指同时与数字内容进行交互。

可共享界面的一个优点是它提供了一个大的交互空间,可以支持团队灵活工作,允许团队在同一时间内共同创建内容。用户可以一边指向和触摸显示的信息,一边查看别人交互的信息,并拥有相同的共享参考点。现在有许多为博物馆和画廊开发的桌面应用程序,旨在使游客了解其所在环境的各个方面。

另一种可共享界面的形式是软件平台,它可以让一组人同时工作,即使他们身处不同的位置。现在有多种商业产品可以让多个远程人员同时处理同一个文档。有些软件可以让多达 50 人在同一时间编辑同一份文档,同时会有更多的人观看。这些软件程序提供各种功能,如同步编辑、跟踪更改、注释和评论。

3.5.3 实体用户界面

实体用户界面是指用户通过物理环境与数字信息发生交互行为,开发的目的是通过赋予无形的数字信息以可触摸的实体形式来增强协作、学习和设计的能力,从而充分利用和挖掘人类对于物理对象和材料的掌握和操纵能力。

实体界面使用基于传感器的交互,其中物理对象(如砖块、球和立方体)与其数字表征一一对应。当一个人操纵物理对象时,计算机系统通过嵌入物理对象中的传感机制检测到其动作,从而产生数字效果,如声音、动画或振动。效果发生在多种媒体和环境中,也可以嵌入物理对象本身。例如,早期的流块原型描述了嵌在其中的数字和灯光的变化,这种变化取决于它们之间的连接方式。设计流块就是为了模拟现实生活中的动态行为,并按一定顺序做出反应(图 3-22)。

还有一种实体界面,是将物理模型(如一个冰球、一块黏土或一个模型)叠加在数字桌面上。在桌面上移动实体部件会导致桌面上发生数字事件。如 Urp(图 3-23)是最早的实体界面之一,用来进行城市规划:建筑的微型物理模型可以在桌面上移动,与数字化的风和阴影生成工具结合使用,阴影会随着时间而变化,空气的流动也会发生改变,根据建筑物物理模型的位置和方向,可以在桌面上进行气流、阴影、反射和其他数据的数字模拟。



图 3-22 流块:一种会思考的玩具



图 3-23 增强现实城市规划工作台 Urp

实体界面不同于其他方法。因为表征本身是实际存在的,因此用户可以直接操作、提动,重新排列、整理和操作。现有许多实体交互系统的目标是鼓励学习、辅助设计活动、增加趣味性和合作,如针对景观和城市规划的规划工具。

3.5.4 自然用户界面

针对用户体验设计的界面类型有很多,那么,更自然的界面类型会变成主流吗?

在自然用户界面(Natural User Interface, NUI)中,“自然”一词是相对图形用户界面(Graphic User Interface, GUI)而言的。GUI 要求用户必须先学习软件开发者预先设置好的操作,而 NUI 则只需要人们以最自然的交流方式(如语言和文字)与机器互动。直观地说,使用 NUI 的计算机不需要键盘或鼠标。特别是触控技术将使人机交互变得更加自然直观,更为人性化。

NUI 允许人们像与真实世界交互一样与计算机交互——使用他们自己的声音、手和身体。NUI 使用户能够与机器对话、触摸和对它们做手势、在检测脚部运动的垫子上跳舞、对它们微笑以获得反应,等等。从理论上讲,与学习使用 GUI 相比, NUI 更容易掌握,并且更容易映射到人们与世界的交互。

相比于记住按下哪个键来打开文件, NUI 意味着人们只需要抬起手臂或说“打开”就可以了。但 NUI 是否自然取决于许多因素,包括需要多少学习成本、应用程序或设备界面的复杂性,以及是否有准确性和速度的要求。有时候,一个手势胜过千言万语;还有时候,一个字抵得上一千种手势。这取决于系统支持多少功能。

在不同的界面类型中,手势、语音和其他类型的 NUI 使得控制输入及与数字内容交互变得更容易、更令人愉快,尽管有时它们可能并不完美。例如,研究表明,使用手势和全身动作作为电脑游戏和体育锻炼的输入形式是令人非常愉快的。此外,新型的手势、语音