

基于视线输入法的人机交互系统研究^①

The HCI System Based on Eye - Gazed Input

黄 乔 (武汉理工大学自动化学院 湖北武汉 430070)

摘 要: 在人与计算机的交互过程中,视觉通常作为人接收信息的感觉通道。基于视线输入的各种人机交互系统正被国内外学者广泛研究。本文分析和研究了最近国际成果,并着重对视线输入系统实现中的软件设计深入讨论,为今后实现更优性能的视觉输入系统打下了一定的基础。

关键词: 视线输入 眼动检测 人机交互

人与计算机交互是研究人与计算机及其相互作用的技术,其目的在于利用所有可能的信息通道进行人机交流,提高交互的自然性和高效性。目前人与计算机交互由于一直受到基于键盘及鼠标的窄带宽输入界面的限制,人机交互中输入与输出效率变得越来越不平衡,由此激发了基于视线跟踪、语音识别、手势术输入、感觉反馈等新的交互技术的多通道界面的研究,其中视线追踪技术倍受该研究领域的重视。

1 引言

视线跟踪的目的是根据眼睛的特征和位置以及其他一些脸部线索来决定用户注视的方向,把它作为一种可选的或互补的计算机输入通道,为人机交互研究提供一种分析工具。随着视线追踪技术日臻成熟,目前眼动测量达到的时间与空间精度为以视线追踪技术取代传统的指点设备,成为一种全新的输入通道提供了可能。

2 视觉输入法

视觉输入法是建立在对视线跟踪的基础之上,视线跟踪的实现可分为硬件检测、数据提取、数据综合三个步骤。视线跟踪的理论基础是:人在看清某物体的过程中总会伴随有眼球的转动,因此,一个人眼球的位置正好提供了一个很好的指示——眼球位置能反映出他所观察对象是所在场景的哪个部分。通常人眼球的

运动是近乎瞬时和快速发生的。而且随后就会有一个 200 - 600ms 的相对静止盯视过程,以便能看清物体。此时眼球是相对完全静止,并伴有微小的抖动,但是在这个盯视过程中,此抖动一般小于 1°。如果眼球有平稳的转动,这种情况也只是发生在人观察运动的物体的时候。其他眼球活动,比如旋转,分散等,在人机对话中显得相对不重要,可以忽略不计。计算机屏幕前方固定眼睛定位装置和视线跟踪装置,当用红外辅助光源(图 1 中 LED 阵列)照射人的脸部时,在眼睛角膜表面形成反射像,此像称为普尔钦(Purkinje)点。随着人眼在盯视计算机屏幕上的不同位置时,眼球会发生相应的转动。眼睛定位装置(见图 1)记录下眼球的初始位置。假定实验者的头在不动的情况下,并且红外发光二极管的位置是固定的,而眼球是一个近似的球体,所以当眼球转动时,可以认为普尔钦点的绝对位置不改变,而虹膜与瞳孔的位置要发生相应的变化,这样红外辅助光源在虹膜上形成的普尔钦点和瞳孔与虹膜的相对位置关系也将发生变化,红外摄像机(见图 1)能实时确定其位置,这种相对位置关系的确定可以通过图像处理来实现。然后由它们之间的相对位置关系,得出视线输入的方向。

对视线完全实时跟踪的同时,就可以着手利用视线对计算机进行操作,即所谓视觉输入法。确定视线的输入方向后,进而确定视线盯视目标在屏幕中的相对位置,将盯视目标设计成对话框形式,如屏幕软面

① 基金项目 基金项目(编号) 科研项目(编号)

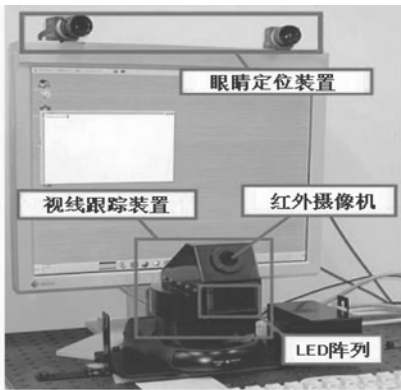


图1 系统主要硬件组成

板,把眼睛对对话框的操作设定一系列参数,如用眼球注视时间长短作为判断是否“点击”对应区域的条件,就可以形成了一个视觉输入系统。

3 视觉输入系统的相关软件设计

早在1993年 Baluja 和 Pomerleau 于提出了一种由人眼的图像来推测用户在计算机屏幕上注视点的方法,来推断眼睛在计算机屏幕上的注视位置,即以眼睛的方位代表注视方向。研究表明,把视线作为系统的输入方式,其难点在于能否精确定位眼球的相对位置以及当眼球转动时能否快速反馈给计算机其变化后位置。而在确定视线跟踪后眼球的实时位置后,利用眼球对计算机如何准确操作成为另一个最大的挑战。正是因为如此,在设计该系统时,往往分开去寻求难题的解决方案。

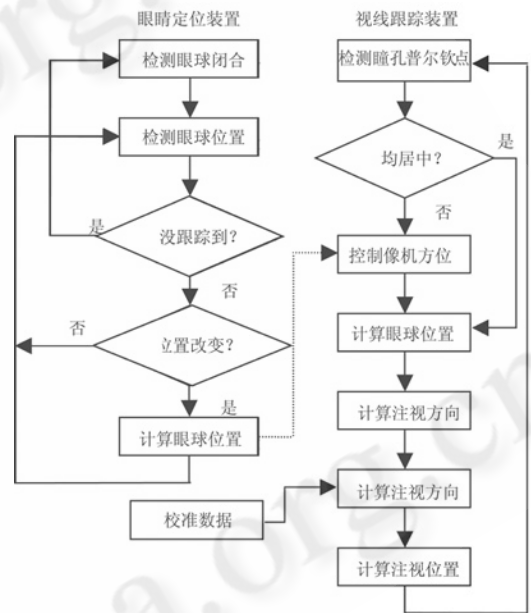
3.1 视线跟踪的软件设计方法

对视线的跟踪步骤可以分为对人眼球位置的跟踪和对眼球注视方向的跟踪。因此,在系统设计时,可利用不同的装置分别完成眼球位置跟踪和眼球注视方向跟踪。两者之间又有紧密联系,即眼睛定位装置中检测的眼球位置作为输入数据传送到视线跟踪装置,以确定实时的眼球注视位置。

在如何利用眼球跟踪数据上,我们可以把处理眼球运动的数据的过程可以划分为两个阶段:首先,我们应该明白处理这些大量的眼球运动的原始数据是为了过滤噪音,识别眼球的具体位置,补偿局部的抖动误差,以及更广泛地从所得信息中重构或复现用户的真

正目的,即眼球所注视的实际位置。这个处理阶段就是将那些连续的,比较杂乱的眼球位置信息流,也就是原始的,待处理的图象,转换成具有一定含义的特定标记,使它们最大程度上接近用户的真正意图,依靠这些标记就可以实现更高级更有效的人机对话。其次,我们就可以着手开发设计常见的人机交互技术,这些都是建立在把那些具有一定含义的特定标记作为系统输入的基础之上。

视线跟踪程序的主要流程图如下所示:



3.2 利用视线作为输入的软件设计

当前由于视线跟踪技术还没有完全成熟,另外眼动本身存在固有的抖动,以及眨眼等所造成的数据中断,会存在许多干扰信号,使得我们在把注视焦点与屏幕元素相关联时存在困难,另一方面还没有能够把视觉通道与其它通道整合,实现无缝连接的成熟、高效的融合算法。然而对以视线输入来实现文字操作的研究越来越受到更广泛的关注。剑桥大学卡文迪许实验室研发的名字为 Dasher 的软件(见图2)能够接受用户眼睛的盯视,并对可能输入的词汇进行预测。输入的字母会在屏幕内形成一个句子,从而实现利用视线的输入法。

对于在文字输入方面的应用,一种直接且常用的利用视线输入来实现文字输入的方法就是设计一种屏

幕中的软键盘,即类似人们常用的 QWERTY 键盘,把它作为文字输入的工具,而实现对键盘按键的“点击”就需要依靠文字操作者对该屏幕中该按键的盯视时间来决定。就像我们平时点击自己的键盘一样,一段预先设计的盯视时间后,眼睛就能击活屏幕中软键盘上不同的各种符号。这些符号就对应计算机的各种不同的操作。通过用户短暂的注视过程,目光的焦点就可以从一个目标转移到另一个目标。这个持续的注视某目标的过程称为盯视时间。熟练的操作者一般比普通用户需要的盯视时间少。实验已经证明,常见的盯视时间为 600ms 至 1000ms 之间,通过盯视时间来确定该操作是否进行,这个过程可以让用户很方便地决定,即当用户在盯视时间还未到来前将视线从按键处移开,该按键就不会被击活,任何操作将不会发生。如果设定的盯视时间太短,就会造成操作者在只是观察或者寻找目标但还没有确定要输入的时候系统已经发出了指令。如果盯视时间设定太长,操作者也会在使用的時候因为需要等待很长一段系统响应时间而感到不适,设定一段较长的盯视时间可以有效地防止用户因无意识的视线而击活某按键,然而,这样就必然会破坏利用视线作为输入在速度上的有利优势,同时也增加了整个系统对界面的响应周期。因此,对目标的盯视时间既要保证输入尽可能快又要保证输入错误尽可能少。

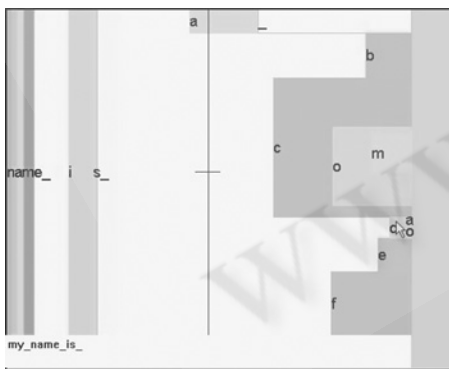


图2 操作时输入文字

对于利用视线作为人机交互系统的输入,人们最初的设计思路是试图在形式上将系统的输入结构设计成类似我们现在常见的鼠标结构。从理论上来说,尽

管在视线移动速度要比鼠标移动要快,但是在利用视线作为系统输入上,仍然有很多障碍需要正视和解决:

整个系统中没有类似鼠标按键的结构,虽然目前在很多场合可以利用视线盯视技术,但是操作前需要学习怎么使用,而且对于用户学习起来还存在一定困难。

点击和拖动选中目标目前还不太可能实现。

操作者对视线的控制一直都是存在的,即视线时时刻刻都在对系统构成输入,这一点和鼠标很不一样,因此,在用户操作的时候对其有意识和无意识的目标选择活动无法有效区分。

视线输入系统需要一小段盯视时间,但实际使用时很容易使用户产生视觉疲劳,从而限制了该系统的推广和运用。

4 结束语

随着科技的不断进步和社会的不断发展,视线盯视输入法作为对键盘、鼠标等传统人机交互方式的补充和挑战,有广泛的潜在应用价值,最为直接的应用在于助残、护理机器人人机交互平台等。在该领域更进一步的深入研究基础上,我们相信其应用范围更加广泛。

参考文献

- 1 刘伟,袁修干.人的视觉—眼动系统的研究.人类工效学,2000,64:41-44.
- 2 陈信,袁修干.人-机-环境系统工程总论(飞行器-人-环境系统工程系列专著)M.北京:北京航空航天大学出版社,1996.
- 3 E. Thomas eds. Human computer interaction using eye gaze input. IEEE Trans on Systems, Man and Cybernetics, 1989, 196: 1527-1533.
- 4 Tanr. Iverdiv, Jacob J. K. Interacting with eye movements in virtual environments. (A CHI Conference Proceedings), C. Netherlands, ACM Press. 2000, 15-21.
- 5 Hansen. D. W, Nielsen. M, Johansen eds. Eye Typing using Markov and Active Appearance Models. IEEE Workshop on Applications on Computer Vision, 2002, 132-136.