

面向大屏幕目标选择的眼动跟踪与手势交互*

程时伟, 朱安杰, 范菁

(浙江工业大学 计算机科学与技术学院, 浙江 杭州 310023)

通讯作者: 程时伟, E-mail: swc@zjut.edu.cn



摘要: 眼动跟踪方法具有很强的视觉指向性,可以将其应用于面向大屏幕的目标选择,进而避免鼠标操作方式在空间上的远距离移动.然而,仅仅利用眼动跟踪进行选择操作,也会产生选择精度降低、容易产生误操作等问题.因此,为了实现大屏幕上快速、准确的目标选择,提出一种融合眼动跟踪与手势的多通道交互方法,即通过眼动跟踪选择目标,利用手势进行选择确认.在目标尺寸小、目标间距较小时,通过光标稳定和二次选择机制进一步对交互过程进行优化.用户测试结果表明,该方法可以在大屏幕上针对不同尺寸和间距的目标完成有效的选择操作,与仅使用眼动跟踪的目标选择方法相比,任务完成速度提升了16%,任务完成正确率提升了82.6%.此外,针对层级菜单的具体选择任务,该方法与仅使用眼动跟踪的方法相比,任务完成速度提升了13.6%,任务完成正确率提升了55.7%.此外,该方法总体性能接近传统的鼠标操作方式,进一步验证了该方法在实际应用中的有效性.

关键词: 眼动跟踪;注视点;用户界面;多通道人机交互

中文引用格式: 程时伟,朱安杰,范菁.面向大屏幕目标选择的眼动跟踪与手势交互.软件学报,2018,29(Suppl.(2)):75-85. <http://www.jos.org.cn/1000-9825/18019.htm>

英文引用格式: Cheng SW, Zhu AJ, Fan J. Eye tracking and gesture based interaction for target selection on large display. Ruan Jian Xue Bao/Journal of Software, 2018,29(Suppl.(2)):75-85 (in Chinese). <http://www.jos.org.cn/1000-9825/18019.htm>

Eye Tracking and Gesture Based Interaction for Target Selection on Large Display

CHENG Shi-Wei, ZHU An-Jie, FAN Jing

(School of Computer Science & Technology, Zhejiang University of Technology, Hangzhou 310023, China)

Abstract: Mouse based target selection will require much movement when locating target across long distance on large display. On the other hand, eye tracking technique can locate target more easily and quickly across long distance. Hence, eye tracking has a high potential for fast targets selection on large displays. However, eye tracking still faces the challenges, such as low accuracy, and high error rate of selection operation, especially for gaze-only interaction. This study proposed a multimodal interaction method that combined gaze with gestures. This method utilized gaze for rough selection first, and then utilized hand gesture to confirm accurate selection. Furthermore, in order to keep the selection accuracy when targets are small and crowded, authors used semi-fixed gaze cursor and secondary selection mechanism to optimize the selection process. Finally, the method conducted a user study in different levels of target sizes and distance among targets. The results show that the selection speed and accuracy rate of proposed method are higher than those of the method only using gaze with 16% and 82.6% respectively. In addition, for the selection of the hierarchical menu items, the selection speed and accuracy rate of proposed method are higher than the those of method only using gaze with 13.6% and 55.7% respectively. In addition, the overall performance of proposed method is similar to the mouse based selection method, and it also validates the effectiveness of the proposed method.

Key words: eye tracking; fixation; user interface; multimodal human-computer interaction

* 基金项目: 国家重点研发计划(2016YFB1001403); 国家自然科学基金(61772468)

Foundation item: National Key Research and Development Plan (2016YFB1001403); National Natural Science Foundation of China (61772468)

收稿时间: 2018-06-15; 采用时间: 2018-11-19

近年来,随着投影屏幕、大型 LED 显示器等在人们工作生活中逐渐得到普及,面向大屏幕显示器的人机交互也得到了迅速发展^[1].目标选择是大屏幕交互场景下的重要操作,通常分为两步:定位要选择的目標和确认当前所定位的目标^[2].传统输入设备,例如鼠标在选择操作时的效率和精度虽然很高,但用户在使用时需要频繁移动手部,容易引起疲劳甚至导致“鼠标手”等肌肉劳损,而在大屏幕上这一问题将更加突出,因为用户往往需要操作鼠标从大屏幕的一端移动到另一端,产生大幅度的手部移动,使得交互过程中的不自然性更加明显^[3].

在这种背景下,基于眼动跟踪的交互方法也逐步应用于面向大屏幕的目标选择操作^[4,5].眼动跟踪是一种提取人眼运动特征,跟踪用户视线的交互技术,其具备速度快、对用户干扰小、易于使用等特点^[6].在目标选择的应用场景中,眼动跟踪具有很强的指向性,可以较快完成目标的定位和选择^[7].此外,视线方向在一定程度上表征了用户的交互意图,进一步提高了隐式交互的自然性^[8].然而,基于眼动跟踪的目标选择仍然面临着以下几个问题:1) 米达斯接触(Midas touch).眼动跟踪作为输入通道完成交互任务时,常常会因为无意识的眼跳运动,导致误操作的产生^[9].2) 眼动跟踪精度有限.当单个目标的尺寸或多个目标之间的间距小于眼动跟踪精度时,用户将难以选中正确的目标^[10].3) 用户体验较差.在人机交互过程中,如果长时间使用注视、眨眼、眼势等基于眼动跟踪的操作,将会导致用户疲劳度增加,严重降低主观体验^[11].

针对上述问题,本文将眼动跟踪与手势相结合,实现多通道融合驱动的目标选择.具体采用 LeapMotion 体感控制器进行手势交互,LeapMotion 不需要用户穿戴设备,并支持三维空间中的手势识别.通过眼动跟踪指向目标,手势确认目标的交互机制实现大屏幕上的目标选择^[12],特别是当目标的个体尺寸较小、目标之间间距较小时,通过手势改变眼动跟踪光标的选择范围,不仅可以提高眼动跟踪指向精度,还可以降低操作难度.同时,结合二次选择机制,实现准确的目标选择.此外,进一步验证了本文方法在层级菜单目标选择中的有效性.层级菜单各栏之间的间距很小,且通常需要从父级菜单开始,连续完成多层选择,操作难度较大,因此已有研究认为眼动跟踪不适用于层级菜单的选择^[13],而本文研究发现,通过眼动跟踪光标进行目标定位,并结合手势进行二次选择,可以在层级菜单的选择应用中达到接近鼠标操作的速度和准确率.

1 研究现状

1.1 面向大屏幕的交互

Keefe 等人进行了多人协同场景中大屏幕数据可视化的研究,验证了大屏幕实现自然交互的优越性和必要性^[14].Dachselt 和 Buchholz 利用手机上的陀螺仪,将手机屏幕投射在大屏幕上^[15].Boring 等人利用手机摄像头将手机屏幕上的触控操作映射到大屏幕^[16].除了将手机作为输入设备,Nancel 等人还设计了一种手势、按键、触摸相结合的技术,用于完成大屏幕上目标的缩放操作^[17].此外,研究表明,眼动跟踪在大屏幕上的操作速度比鼠标要快,因此将眼动跟踪交互技术应用于大屏幕具有很大的潜力^[18,19].然而 Kammerer 等人的研究表明,虽然在大屏幕上利用眼动跟踪完成任务的速度要很快,但其错误率也较高^[13].Yoo 等人将眼动跟踪和手势结合,应用于大屏幕的交互中,他们设计了“拉-推”的手势来控制一个目标的缩放,然而并没有考虑眼动跟踪在选择目标时精度不足的问题^[20].Turner 等人对大屏幕上的目标操作问题进行了更深入的研究,结合眼动跟踪和移动设备上的操作实现了大屏幕上目标的旋转、缩放、平移^[21].

1.2 基于眼动跟踪的多通道交互

早在 20 世纪 80 年代,Bolt 就认为眼动跟踪是一种快速且自然的输入方式,可以将眼动跟踪作为一个额外的输入通道来提升交互系统的性能^[22].Jacob 等人将眼动跟踪与其他输入通道结合起来,可以加强眼动跟踪在交互中的可靠性^[3].进一步,Zhai 等人提出了名为 MAGIC 的目标选择技术,结合眼动跟踪和鼠标实现目标的快速选择,他们的工作启发了后来的研究人员,为眼动跟踪在目标选择中的应用提供了新思路^[23].针对眼动跟踪中存在的米达斯接触问题,Bieg 等人提出了一种利用注视时间触发操作的交互技术,即当用户在某一目标上的注视时间超过阈值时才会触发相应的操作^[24].Lutteroth 等人通过确认机制来降低误操作,例如,当用户注视目标后,屏幕上显示确认按钮,只有当用户注视该确认按钮,才会触发后续相应的操作^[25].Kammerer 则提出了用户在选

择目标时首先用眼动跟踪定位目标,然后再利用语音进行确认的方法,从而避免米达斯接触问题^[13].Ware 和 Mikaelian 设计了 3 种基于眼动跟踪的选择技术,分别用物理按键、触屏按键、注视时间来触发选择,他们的研究表明用注视时间来触发选择将导致选择准确率明显降低^[26].类似地,Anderson 也将眼动跟踪与按键操作相结合,应用于目标选择,并发现导致目标误选的最大原因是眼动跟踪和按键确认的不同步性,即用户确认选择时眼睛已看向了下一个目标^[27].针对眼动跟踪精度有限的问题,常见的解决方案有:放大目标尺寸^[28]、与其他通道融合进行多次确认^[29]、优化目标匹配算法^[30,31].此外,Monden 等人将基于眼动跟踪的选择方法分成了 4 类:(1) 基础选择,即眼动跟踪选择目标后,第 2 输入通道确认选择;(2) 自动选择,即在确认选择后,眼动跟踪光标会自动定位到最近的目标;(3) 手动选择,即眼动跟踪进行初步选择,辅助输入通道再做进一步选择;(4) 半自动选择,即结合自动和手动的选择^[32].

综上所述,手势、语音、触控、眼动跟踪等都可以有效应用于大屏幕交互.其中,眼动跟踪作为一种自然、高效的输入技术,通过与其他交互通道相结合,具备较好的应用前景.然而,基于眼动跟踪的多通道交互仍然存在选择精度相对较低、用户体验有待提高等问题.因此本文融合眼动跟踪和手势交互,提出一种新的面向大屏幕的目标选择方法,重点解决单个目标尺寸小、多个目标之间间距小等导致的目标选择困难、选择错误率高等问题.

2 融合眼动跟踪与手势的目标选择

2.1 总体框架

本文方法的总体框架如图 1 所示,由眼动跟踪、手势识别、选择优化这 3 部分组成.眼动跟踪是通过头戴式眼动跟踪仪提取人眼图像特征,基于瞳孔中心-角膜反光点(pupil center cornea reflection,简称 PCCR)方法实时计算用户在屏幕上的注视点坐标^[33],这里,角膜反光点也称普尔钦斑.手势识别则基于体感控制器 LeapMotion 设计了 8 种手势,平均识别准确率为 95%,具体选取其中的抬起、缩放、挥手 3 种手势辅助完成目标选择,如图 2 所示.在选择目标时首先用眼动跟踪光标定位目标,然后通过抬起手势进行确认,避免单独使用眼动跟踪时产生的米达斯接触问题.选择优化则针对单个目标小、多个目标之间间距小的问题,通过提高眼动跟踪光标的稳定性和采取二次选择操作,提高目标选择的准确率.

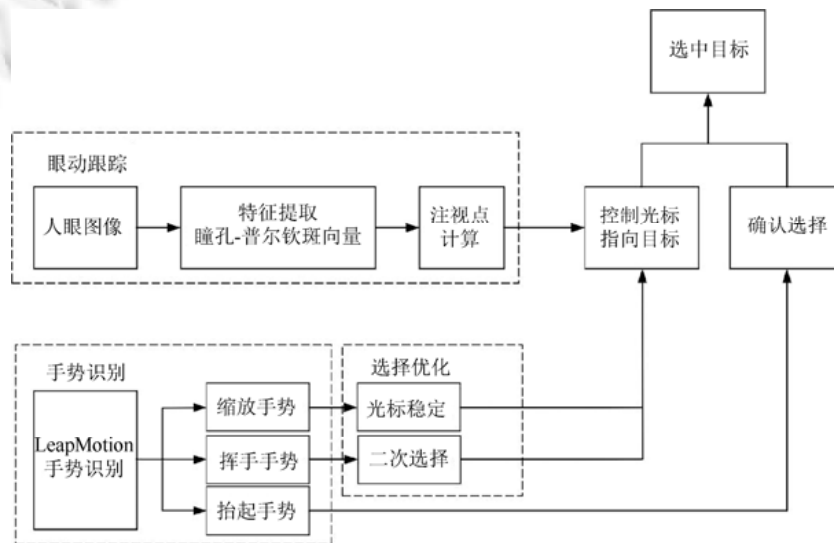


Fig.1 Framework

图 1 总体框架



Fig.2 Gestures

图 2 手势设计

2.2 眼动跟踪

眼动跟踪采用基于 PCCR 的方法计算注视点坐标.首先,通过附带红外光源的摄像头获取人眼图像,并对人眼图像进行高斯滤波、灰度转换等预处理操作;然后识别图像中瞳孔和普尔钦斑的位置,将它们的几何中心组成一个 PCCR 特征向量.为了提高识别效率,先识别出包含瞳孔的大致区域,在此基础上进一步对瞳孔轮廓进行拟合,最终精确识别瞳孔中心的位置.普尔钦斑是红外光源照射在角膜上产生的反光斑点,采用模板匹配法来识别,这样可以在普尔钦斑部分丢失或发生形变时仍然可以保证较高的精确性.当用户的眼球运动时,PCCR 向量会随之改变,因此可以利用此特征向量建立映射模型,计算用户在大屏幕上的注视点坐标.

2.3 选择优化

针对目标尺寸小或目标间隔紧密时选择困难的情况进行了优化,由光标稳定、二次选择两个部分组成.

光标稳定:本文为用户提供了一种眼动跟踪控制的圆形光标(下文简称光标).如图 3(a)所示,光标的圆心 a 为用户当前的注视点,当用户的注视点改变时,光标的位置也会随之移动.当光标覆盖目标后,用户可通过抬起手势进行选择确认,最终选中目标.研究表明,选择尺寸较小的目标时,采用面积更大的光标可以降低选择难度^[34].因此本文中的光标尺寸可以通过手势来控制,当目标的尺寸太小而光标难以覆盖时,用户可以通过缩放手势扩大光标半径 r .如图 3(b)所示,虚线表示半径扩大后的光标,其成功覆盖了小尺寸的灰色矩形目标,有效降低了对眼动跟踪的精度需求.

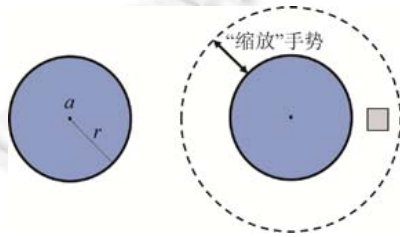


Fig.3 Scalable eye tracking cursor

图 3 可缩放的眼动跟踪光标示意图

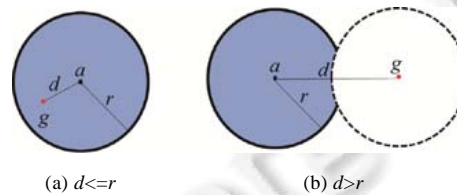


Fig.4 Improved eye tracking cursor

图 4 改进的眼动跟踪光标示意图

另一方面,在实际应用中,由于眼动跟踪设备存在一定的跟踪误差,导致眼动跟踪设备记录的注视点位置会在用户实际注视点位置附近产生扰动,从而导致目标选择时光标的位置也发生相应的扰动,特别是当目标尺寸较小且分布较为密集时,目标的选择会产生较大误差,出错率增加.例如,当用户确认选择的瞬间,光标发生轻微的跳动,导致错误的目标被确认选中.尽管已有研究人员针对眼动跟踪光标的稳定性问题进行了研究,并提出了相应的光标稳定方法,但是仍然存在一定的局限性.为此,本文提出一种改进的眼动跟踪光标方法,进一步提高光标的稳定性.如图 4(a)所示,用户当前实际注视点为点 g ,光标圆心为点 a ,光标半径为 r ,实际注视点与光标圆心之间的距离为 d .当用户的注视点在光标半径范围内($d \leq r$)时,光标的位置不会发生改变.只有当用户的视线发生较大范围的转移,即注视点超出光标范围($d > r$)时,才会相应改变光标的位置,如图 4(b)所示.这种光标一方面消除了眼动跟踪精度误差导致的不稳定性(即用户在选择目标时则侧重于光标的稳定性),另一方面又能根据幅度较大的眼跳及时调整光标位置(用户在移动光标时侧重于光标的灵活性),能有效帮助用户更好地选择小尺寸目标.这种光标的稳定性和灵活性与光标的尺寸相关,光标尺寸越大稳定性就越高,光标尺寸越小

灵活性就越高.用户可以通过缩放手势自由改变光标的尺寸,实现光标在稳定性和灵活性之间的切换,满足不同情况下的实际需求.

二次选择:采用上述尺寸可控制的眼动跟踪光标可以帮助用户完成小尺寸目标的选择,但是当目标的分布十分密集甚至发生重叠时,光标会同时指向多个目标,无法实现精确的选择.如图 5 所示,大屏幕上有着大量的小尺寸目标紧密排布,要求用户选择其中粗边框的既定目标.为了稳定光标,用户可以通过缩放手势扩大光标的尺寸,但也导致光标同时覆盖了相邻的 4 个目标,无法准确指向既定目标.针对这一问题,可以采用光标预选择方法,即所有被光标覆盖的目标都会被添加到预选列表中(如图 5 中被光标覆盖的 4 个目标).预选列表中存在一个选择指针,当用户通过抬起手势确认选择时会选中指针指向的目标,该目标会变成绿色.可通过挥手手势控制选择指针的移动,当其指向用户想要选择的目标时,用户可以最终确认选择,从而对既定目标完成准确选择.

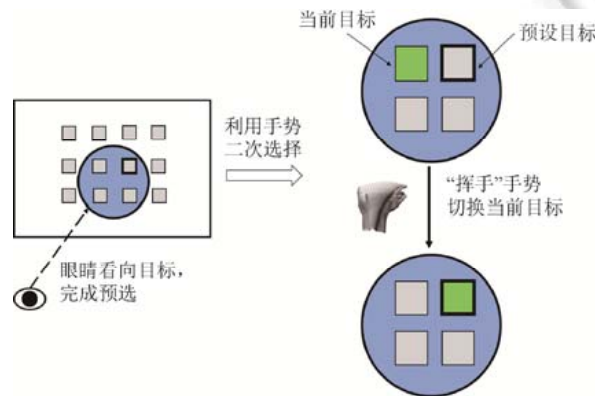


Fig.5 Secondary selection

图 5 目标二次选择

然而,二次选择仍存在一定的局限性.例如,目标分布过于密集时,将导致同时覆盖的目标过多,使得筛选时间过长,选择效率明显降低.为此,本文进一步根据眼动跟踪数据预测用户意图,通过目标被选择的可能性为其设置优先级,优先级高的目标将更容易被选中.具体的方法是:用户使用眼动跟踪光标进行选择时,眼动跟踪设备会记录用户过去一段时间窗口(例如 2s)内的注视点坐标,并将这些注视点的平均坐标作为当前交互意图下的关注点(下文简称为关注点).目标位置与关注点越近,其优先级就越高;此外,预选列表中目标优先级顺序还将根据用户的关注点动态改变.

如图 6(a)所示,如果要想选择指针指向最右侧的目标(粗边框),最多需要通过手势切换 7 次,效率较低.如图 6(b)所示,设置预选列表后,可以得到关注点 i ,根据目标与关注点 i 之间的距离大小设置优先级,因此最右侧的目标优先级最高,其动态调整为预选列表中的第 1 位,此时用户不需要挥手进行多次切换,而直接用抬起手势确认选中即可.这种方法虽然也存在一定误差,但可以明显减少用户在二次选择过程中的手势操作次数,显著提高选择效率.

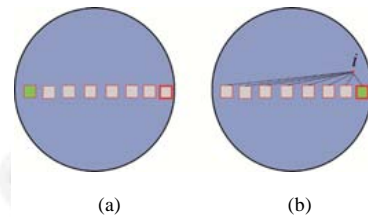


Fig.6 Target selection based on pre-selected list

图 6 基于预选列表的目标选择

3 用户实验

为了验证本文方法,设计开发了 Gaze+Gesture 原型系统,并进行了相应的用户实验.实验主要验证 2 个问题:(1) 眼动跟踪与手势融合的目标选择方法(下文简称眼动跟踪+手势)在大屏幕上是否有效,以及目标的尺寸、间隔对选择效果有何影响?(2) 眼动跟踪+手势的目标选择方法能否有效应用于具体的用户界面的交互,例如,层级菜单的菜单项选择.

实验设备:为了获取眼动跟踪数据,采用头戴式眼动跟踪仪^[33],如图 7(a)所示,该眼动跟踪仪包含两个摄像头,分别是场景摄像头和眼部摄像头.在眼部摄像头上加装红外光源以及滤光片,放置于人眼平视方向的下方,从而能够减少对用户的干扰并且拍摄清晰的人眼图像.场景摄像头与人眼平视方向平行,用于拍摄用户当前注视的场景.采用体感控制器 LeapMotion 实时识别用户手势,最高感知精度可达 1mm,如图 7(b)所示.



Fig.7 Experimental equipments

图 7 实验设备示意图

具体的眼动跟踪仪的跟踪精度和图像采样帧率见表 1.

Table 1 Accuracy and frame rate of eye tracking

表 1 眼动跟踪精度和帧率表

视距(m)	精度(cm)	帧率(帧/s)
1	5.53	30
1.5	8.47	30
2	9.55	30

将眼动跟踪仪和 LeapMotion 通过 USB 连接线接入同一台电脑(CPU 主频为 4GHz,内存大小为 16G).

实验对象:共招募 15 名用户(8 男、7 女)参与本次实验,年龄为 22 岁~28 岁,裸眼视力良好,无色盲,所有实验用户都有日常使用电脑的经历.

实验设计:采用组内实验方式,即每一位用户都要分别使用鼠标、眼动跟踪、眼动跟踪+手势这 3 种操作方式完成目标选择任务(见表 2).为了排除学习效应对实验结果的干扰,随机指定每一位用户使用 3 种不同操作方式的顺序.

Table 2 Operation description

表 2 操作方式说明

操作方式	具体操作步骤	姿势
鼠标	鼠标移动光标,单击左键确认选择	坐
眼动跟踪	视线移动光标,视线停留确认选择	站立
眼动跟踪+手势	视线移动光标,手势确认选择	站立

实验开始前先向用户简单介绍实验内容以及注意事项,并设置用户与大屏幕之间的距离为 1.5m.用户除了在鼠标操作方式下可以坐下,其他操作方式均为站立.实验开始后,首先让用户在大屏幕上完成眼动跟踪的九点标定(鼠标操作方式下不需要标定).然后让用户按照随机的顺序使用不同操作方式进行实验.每一种操作方式下,用户都拥有不超过 5 分钟的时间进行练习.当用户完成练习后,分别完成基础目标选择和层级菜单选择两个实验任务,实验的场景如图 8 所示.

在基础目标选择任务中,用户需要选择大屏幕上粗边框的既定目标.在实验场景中设定用户与大屏幕之间的距离为 1.5m,此时眼动跟踪的精度为 8.47cm.目标的尺寸和间隔这两个变量分别用 S 和 D 来表示,其中 S 的值为 $\{S_1=5, S_2=10, S_3=15\}$ (cm),分别对应于目标尺寸小于眼动跟踪精度、目标尺寸约等于眼动跟踪精度、目标尺寸大于眼动跟踪精度这 3 种情况; D 的值为 $\{D_1=0, D_2=5, D_3=10\}$ (cm),分别对应为目标间隔为 0、目标间隔小于眼动跟踪精度、目标间隔约等于眼动跟踪精度这 3 种情况.

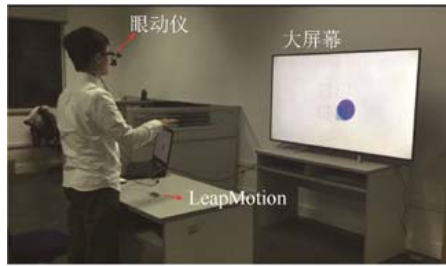


Fig.8 Experimental scene

图 8 实验场景

在层级菜单选择任务中,为了研究跨层连续选择对选择时长和出错率的影响,并保持眼动选择方式的一定优势,设置菜单项之间的距离为 D_1 ,菜单项短边尺寸为 S_2 (接近眼动跟踪精度),长边尺寸为短边的2.5倍(即 $2.5 \times S_2$),用户需要按照顺序选择屏幕上粗边框的菜单项.如图9所示,开始只显示顶部的父级菜单项,只有选择正确的父级菜单后,子菜单项才会出现.按照需要选择的菜单层数,将层级选择任务分为3种难度:简单(完成一级菜单的选择,如选中“文件”)、中等(完成二级菜单的选择,如先选中“文件”,再选中“另存为”)、困难(完成三级菜单的选择,如先选中“文件”,再选中“另存为”,最后选中“JPG”).



Fig.9 Experimental design

图 9 实验任务设计

评估指标:记录实验中选择目标的时长以及选择目标的正确率作为实验评估指标.将一个目标出现到目标被选中的时间记录为一个选择时长,若粗边框的目标没有被选中(即选择时长超出预设的极限阈值),则认为任务失败.在层级菜单选择任务中,用户需要在每一层菜单中都选择正确才算成功完成任务.实验结束后,用户还需要填写一份调查问卷,对不同的选择操作方式反馈主观满意度.

4 实验结果分析

在基础目标选择任务中,3种操作方式的平均选择时长和平均出错率分别如图10和图11所示.

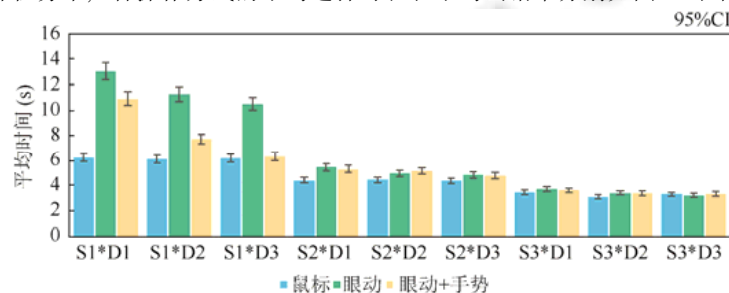


Fig.10 Average selection time of target selection

图 10 基础目标选择任务的平均选择时长

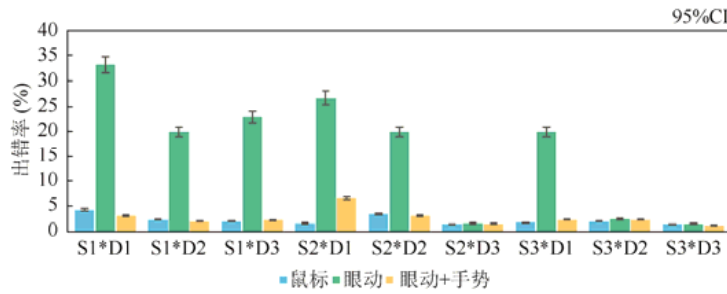


Fig.11 Average error rate of target selection

图 11 基础目标选择任务的平均出错率

目标尺寸 S 和间距 D 对平均选择时长的影响显著性见表 3.

Table 3 Effect of target size (S) and spacing (D) on average selection time

表 3 目标尺寸(S)、间距(D)对平均选择时间的影响

操作方式	假设检验结果	
鼠标	$F_D=1.380, p=0.255$	$F_S=107.356, p<0.001$
眼动跟踪	$F_D=0.133, p=0.875$	$F_S=42.106, p<0.001$
眼动跟踪+手势	$F_D=7.320, p=0.001$	$F_S=38.762, p<0.001$

对基础目标选择任务的实验结果做进一步分析后可以得出以下结论.

(1) 目标尺寸 S 对鼠标、眼动跟踪、眼动跟踪+手势这三种方式的平均选择时长都有显著的影响.从图 10 中可以看出,小尺寸目标(S_1)上的选择时长要明显长于较大尺寸目标(S_2 和 S_3)的选择时长.在选择 S_2 、 S_3 尺寸的目标时,3 种操作方式选择时长的差别很小,当尺寸为 S_1 时,眼动跟踪选择目标的时长要长于其他方式;眼动跟踪+手势选择时长虽然要长于鼠标操作方式,但明显小于眼动跟踪操作方式.

(2) 目标间距 D 对眼动跟踪+手势的平均选择时长有显著的影响,而对鼠标和眼动跟踪的平均选择时长的影响则不明显.如图 10 所示,在目标尺寸为 S_1 时,目标间距越小,眼动跟踪+手势的选择时长显著变长.这是因为在利用眼动跟踪+手势选择小尺寸目标时,用户会使用缩放手势扩张光标的大小,特别是在目标间距较小时,扩张后的光标就会同时覆盖多个目标,使得用户需要进行挥手手势进行二次选择,这些操作都会导致选择时长增加.另一方面,在不同的目标间距下,鼠标和眼动跟踪的选择时长的差异都很小,这是因为鼠标选择方式的精度较高,即使目标之间紧密排列也能较快地完成选择.而眼动跟踪的精度有限,因此当目标间距很小时,用户虽然选择得较快,但经常产生错误的选择.

(3) 不论目标尺寸和间距如何变化,鼠标操作方式的出错率都是最低的;眼动跟踪+手势的出错率也很低,跟鼠标操作方式的出错率很接近;而眼动跟踪的出错率是最高的,尤其是当目标的尺寸较小时(S_1),出错率要远远高于其他两种操作方式,而且即使在尺寸为 S_2 、 S_3 的情况下,也仍然无法较准确的选中间距较小的目标(D_1),如图 11 所示.这是因为眼动跟踪的精度有限,当目标的尺寸小于眼动跟踪精度时,就无法单纯利用眼动跟踪准确地选中目标.而眼动跟踪+手势中可以通过手势进行二次选择,即使目标尺寸小、间距近,也能更准确地选中.

在层级菜单的选择任务中,3 种操作方式的平均选择时长和出错率分别如图 12、图 13 所示.

可见,随着任务难度的增加,3 种操作方式的平均选择时长和出错率都有明显上升.而在所有任务难度下,鼠标操作方式的选择速度都是最快的,出错率也最小.眼动跟踪+手势的选择时长与鼠标操作方式没有显著性差异($F=0.120, p=0.730$),且出错率只在简单任务中稍高于鼠标操作方式.此外,眼动跟踪操作方式的选择时长最长,并且其出错率也远远高于其他两种操作方式.这与前人关于层级菜单的研究相一致,即单独使用眼动跟踪技术不适合用于层级菜单的选择.而本文提出的眼动跟踪+手势则在操作性能上已接近鼠标操作方式,具有较好的可行性和有效性.

实验结束后,每用户填写一份调查问卷,其结果如图 14 所示.

此外,用户对眼动跟踪操作方式在速度、易学性上的主观评价较高,但在精度和易用性上要明显低于其他

两种方式;同时,用户对眼动跟踪操作方式的总体满意度也较低.而另一方面,用户对鼠标、眼动跟踪+手势这两种操作方式的主观评价比较接近,其中鼠标操作方式在精度和易学性上的主观评价得分要略高于眼动跟踪+手势操作方式,这是因为他们对鼠标操作十分熟练,因此操作准确度较高.但用户对眼动跟踪+手势操作方式的总体满意度要高于鼠标操作方式.事实上,66.7%的用户最喜欢使用眼动跟踪+手势操作方式,因为他们认为在大屏幕前站立时,使用眼动跟踪+手势操作方式更加舒适;而坐着使用鼠标操作时(本文没有要求用户站立使用鼠标,因为如果这样,用户的手部操作舒适感会更差),抬头观看大屏幕会引起不适感,长时间操作易于产生疲劳.

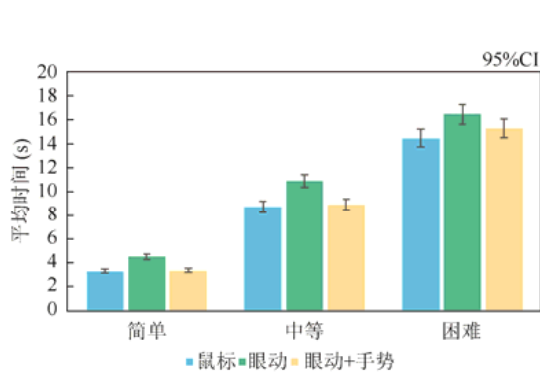


Fig.12 Average selection time of target selection on hierarchical menu

图 12 层级菜单选择任务的平均选择时长

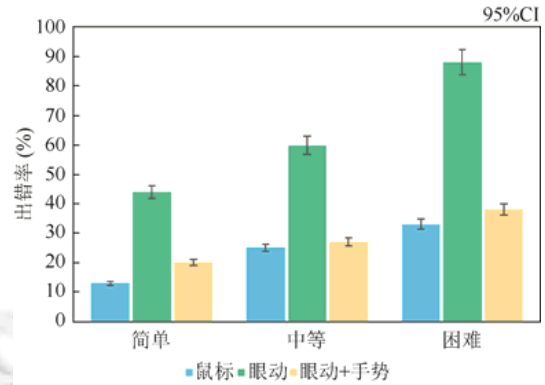


Fig.13 Average error rate of target selection on hierarchical menu

图 13 层级菜单目标选择任务的平均出错率

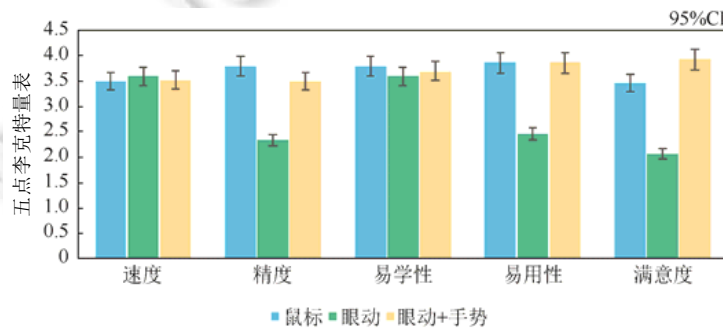


Fig.14 Users feedback

图 14 用户主观反馈

5 结 论

本文提出了一种在大屏幕上结合眼动跟踪和手势识别的目标选择方法.通过两个输入通道确认选择的方式,避免了眼动跟踪在选择目标时的米达斯接触问题.并通过优化光标、二次选择,解决了目标尺寸小、间距近时选择困难的问题,提高了选择精准度.在与鼠标、眼动跟踪这两种操作方式的对比实验中,眼动跟踪+手势的操作方式不仅保留了眼动跟踪速度快的优势,并且大幅度降低了出错率.实验结果表明,本文方法在大屏幕上选择目标的性能与鼠标操作方式下十分接近,且用户主观体验要优于鼠标操作,是一种有效的大屏幕目标选择方法.

下一步的研究将利用眼动跟踪和手势实现大屏幕上的目标移动、缩放、旋转等操作,并将本文方法与 VR 结合,实现虚拟环境下针对三维目标的快速与精确交互.

References:

- [1] Vogel D, Balakrishnan R. Distant freehand pointing and clicking on very large, high resolution displays. In: Proc. of the 18th Annual ACM Sym. on User Interface Software and Technology. 2005. 33–42.
- [2] Stellmach S, Dachzelt R. Still looking: Investigating seamless gaze-supported selection, positioning, and manipulation of distant targets. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2013. 285–294.
- [3] Fares R, Fang S, Komogortsev O. Can we beat the mouse with MAGIC? In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2013. 1387–1390.
- [4] Stellmach S, Dachzelt R. Look & touch: Gaze-Supported target acquisition. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2012. 2981–2990.
- [5] Schenk S, Dreiser M, Rigoll G, *et al.* GazeEverywhere: Enabling gaze-only user interaction on an unmodified desktop PC in everyday scenarios. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2017. 3034–3044.
- [6] Stellmach S, Stober S, Dachzelt R. Designing gaze-supported multimodal interactions for the exploration of large image collections. In: Proc. of the Conf. on Novel Gaze-Controlled Applications. New York: ACM, 2011. 1–8.
- [7] Ware C, Mikaelian HH. An evaluation of an eye tracker as a device for computer input. ACM SIGCHI Bulletin, 1987,17(SI): 183–188.
- [8] Lankford C. Effective eye-gaze input into windows. In: Proc. of the Symp. on Eye Tracking Research and Applications. New York: ACM, 2000. 23–27.
- [9] Horodniczy D, Cooperstock JR. Free the hands! Enhanced target selection via a variable-friction shoe. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2017. 255–259.
- [10] Land MF. Eye movements and the control of actions in everyday life. Progress in Retinal and Eye Research, 2006,25(3):296.
- [11] Sibert LE, Jacob RJK. Evaluation of eye gaze interaction. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2000. 281–288.
- [12] Smith BA, Ho J, Ark W, *et al.* Hand eye coordination patterns in target selection. In: Proc. of the Symp. on Eye Tracking Research and Applications. New York: ACM, 2000. 117–122.
- [13] Kammerer Y, Scheiter K, Beinhauer W. Looking my way through the menu: the impact of menu design and multimodal input on gaze-based menu selection. In: Proc. of the Symp. on Eye Tracking Research and Applications. New York: ACM, 2008. 213–220.
- [14] Keefe D F, Gupta A, Feldman D, *et al.* Scaling up multi-touch selection and querying: Interfaces and applications for combining mobile multi-touch input with large-scale visualization displays. International Journal of Human-Computer Studies, 2012, 70(10):703–713.
- [15] Dachzelt R, Buchholz R. Natural throw and tilt interaction between mobile phones and distant displays. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2009. 3253–3258.
- [16] Boring S, Baur D, Butz A, *et al.* Touch projector: Mobile interaction through video. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2010. 2287–2296.
- [17] Nancel M, Wagner J, Pietriga E, *et al.* Mid-Air pan-and-zoom on wall-sized displays. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2011. 177–186.
- [18] Adams N, Witkowski M, Spence R. The inspection of very large images by eye-gaze control. In: Proc. of the Working Conf. on Advanced Visual Interfaces, 2008. 111–118.
- [19] Gori J, Rioul O, Guiard Y. To miss is human: Information-Theoretic rationale for target misses in Fitts' law. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2017. 260–264.
- [20] Yoo BI, Han JJ, Choi C. 3D user interface combining gaze and hand gestures for large-scale display. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2010. 3709–3714.
- [21] Turner J, Alexander J, Bulling A. Gaze+RST: Integrating gaze and multitouch for remote rotate-scale-translate tasks. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2015. 4179–4188.
- [22] Bolt RA. Gaze-Orchestrated dynamic windows. Computer Graphics and Interactive Techniques, 1981, 109–119.
- [23] Zhai S, Morimoto C, Ihde S. Manual and gaze input cascaded (MAGIC) pointing. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 1999. 246–253.

- [24] Bieg HJ, Chuang LL, Fleming RW. Eye and pointer coordination in search and selection tasks. In: Proc. of the Symp. on Eye Tracking Research and Applications. New York: ACM, 2010. 89–92.
- [25] Lutteroth C, Penkar M, Weber G. Gaze VS. Mouse: A fast and accurate gaze-only click alternative. In: Proc. of the Annual ACM Symp. on User Interface Software & Technology. New York: ACM, 2015. 385–394.
- [26] Isokoski P. Look and lean: Accurate head-assisted eye pointing. In: Proc. of the Symp. on Eye Tracking Research and Applications. New York: ACM, 2014. 35–42.
- [27] Salvucci DD, Anderson JR. Intelligent gaze-added interfaces. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2000. 273–280.
- [28] Kumar M, Paepcke A, Winograd T. EyePoint: Practical pointing and selection using gaze and keyboard. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 2007. 421–430.
- [29] Drewes H, Schmidt A. The MAGIC touch: Combining MAGIC-pointing with a touch-sensitive mouse. In: Proc. of the IFIP Conf. on Human-Computer Interaction. Berlin: Springer-Verlag, 2009. 415–428.
- [30] Špakov O. Comparison of gaze-to-objects mapping algorithms. In: Proc. of the 1st Conf. on Novel Gaze-Controlled Applications. New York: ACM, 2011. 1–8.
- [31] Skovsgaard H, Mateo JC, Flach JM. Small-Target selection with gaze alone. In: Proc. of the Symp. on Eye Tracking Research and Applications. New York: ACM, 2010. 145–148.
- [32] Monden A, Matsumoto KI, Yamato M. Evaluation of gaze-added target selection methods suitable for general GUIs. In: Proc. of the Int'l Journal of Computer Applications in Technology. 2005,24(1):17–24.
- [33] Cheng SW, Sun ZQ, Lu YH. An eye tracking approach to cross-device interaction. Journal of Computer-Aided Design and Computer Graphics, 2016,28(7):1094–1104 (in Chinese with English Abstract).
- [34] Kabbash P, Buxton W A S. The “prince” technique: Fitts’ law and selection using area cursors. In: Proc. of the Conf. on Human Factors in Computing Systems. New York: ACM, 1995. 273–279.

附中文参考文献:

- [33] 程时伟,孙志强,陆煜华.面向多设备交互的眼动跟踪方法.计算机辅助设计与图形学学报,2016,28(7):1094–1104.



程时伟(1981—),男,湖北黄石人,博士,副教授,博士生导师,CCF 专业会员,主要研究领域为人机交互,普适计算,协同计算.



范菁(1969—),女,博士,教授,博士生导师,CCF 杰出会员,主要研究领域为服务计算,虚拟现实,人机交互.



朱安杰(1994—),男,CCF 学生会员,主要研究领域为人机交互.