

基于结构光源的大屏幕交互技术

秦永强, 喻 纯, 史元春

(清华大学计算机科学与技术系, 普适计算教育部重点实验室, 北京 100084)

摘 要: 本文提出采用结构光源充分表达用户手部运动的想法, 设计了结构光源和识别软件, 从光点序列中提取出光点运动轨迹和用户手势的旋转信息, 搜索光点位置序列的平坦区段锁定目标位置, 将旋转手势解释成交互原语以扩充激光笔的交互语义. 最后的性能评估结果表明, 结构光源在丰富交互语义和加速目标选择方面都有着较大的改进.

关键词: 结构光源; 旋转手势; 平坦区段; 性能评估

中图分类号: TP39 **文献标识码:** A **文章编号:** 0372-2112 (2009) 4A-075-04

Interaction on Large Display Using Structured Laser Pointer

QIN Yong-qiang, YU Chun, SHI Yuan-chun

(Department of Computer Science and Technology, Tsinghua University, Beijing 10084, China)

Abstract: In this paper, structured laser pointer (SLP) is proposed to sufficiently express the movement of user's hand. Firstly, the SLP is designed as input device, and then the corresponding signal processing application is programmed to filtrate the sequence of laser spots, and extract the trajectory of the laser spots and user gesture from the sequence. And the laser trajectory and user gesture are translated into original interactive language to increase the semantic interaction. Using a SLP can acquire a target quickly because the dwell of laser spots are not required. Instead, we search the plain section in the sequence of laser spots to quickly and accurately select the target that people want to point to. Finally, according to Fitts' law, a user study is designed that requires participants to achieve several tasks. It proved that SLP performs much better than traditional laser pen in speed, and using SLP can also express more semantic interaction.

Key words: structured laser pointer; rotation gesture; plain section; user study

1 引言

在基于激光笔的大屏幕交互技术中, 一般采用一个摄像机实时拍摄大屏幕, 检测出激光笔投射到大屏幕上的光点位置以引导光标, 并根据激光点的轨迹定制交互语义, 从而实现通过激光笔直接操控大屏幕内容的目的, 但传统的激光笔交互技术存在着交互语义少的缺陷: 传统的激光笔只有一个激光点, 只能表示平移和指向等简单的操作, 表达较为复杂的交互语义时, 只能通过绘制一定路径的图形, 或者通过光点驻停在目标位置一定时间和在激光笔上安装按键的方式来实现, 但是这种交互方式比较复杂, 用户难以掌握, 而且前两种方式触发交互命令的时间也较长.

为此, Sergey V Matveyev 等人的研究中^[1, 2]提到了利用多个激光点进行交互的想法. 但是, 他们是利用激光点排列形状的变化来表达交互信息, 而光点排列形状的

变化是靠用户手按激光笔上的按键触发机械装置来实现, 不能完全利用手部自然运动来进行交互.



图1 手部的运动

我们分析一下人手部可能的运动(如图1): 手部除了可以绕腕关节摆动外, 还可以以前臂为轴转动. 传统的激光笔只能表达手部绕腕关节的摆动(反映到屏幕上是在二维平面上的光点位置变化), 而无法表示手部绕前臂为轴的旋转(表现在屏幕上是在单个激光点绕自身中心的旋转). 具有固定排列形状的光点组可以清楚地表达出这种手部的旋转运动(表现为光点组绕其中心有旋转角度, 即 *Rotate* 分量). 本文提出的利用结构光源进行交互的方法能够充分利用用户的手势来表达交互信息, 将手势原语映射为命令原语, 并搜索光点位置序列中的平坦区段以确定目标位置, 从而实现快捷自然的人机交互.

2 结构光源的设计

结构光源是用户与计算机进行交互的硬件基础,要求能够清楚准确地表达手势信息;同时,为了保证结构光源具有更好的扩展性(例如可根据大屏幕上的光点位置求解用户所在的空间位置等信息),必须对结构光源的构造进行分析,结构光源应当遵循以下设计原则:

(1)能够定位大屏幕上的目标:这是基本要素.

(2)能够准确表达用户手势的旋转信息:当用户的手部以前臂为轴旋转时,从投射在大屏幕上的光点信息中提取出的旋转手势旋转分量应当与用户手旋转角度成线性关系,以便于准确表达用户意图.

(3)不会信息过冗余:以最少的激光头数目达到要求即可,过多的激光头会增加计算量,增大系统的误差,还有可能在求解参数时导致方程组无解.

在满足以上原则的前提下可以着手设计结构光源了.两个激光点能表达的信息量有限(只能表示手势中的旋转运动、光点移动),结构光源上应该配备的激光头数目要多于两个;而为了便于计算,激光点的数目又不能过多,否则会信息冗余.

最终我们设计的结构光源的激光头的数目是4个,激光头组的3D构造和投射到大屏幕上的光点分布如图2所示,外围三个激光头均与中心激光头有固定的夹角,所以用户到大屏幕的距离变化时,光点间的距离也会随之变化.此时,结构光源不仅能够满足前面所提的设计原则,所有激光点的坐标信息还为计算用户在与交互环境中的位置提供了可能.

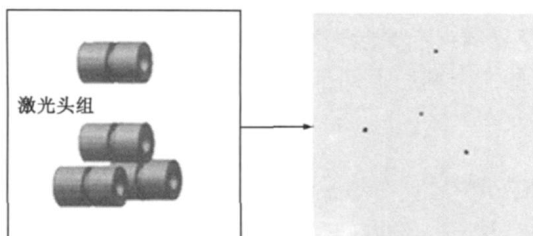


图2 激光头组的3D模型和光点分布

实验用的结构光源笔实物如图3所示(尺寸:25mm×25mm×150mm;质量:110g),为了便于调试,我们采用650nm的红色激光头作为光源,在实际应用中可以采用红外光.



图3 结构光源实物图

3 获取光点信号

结构光源是本文中实现激光笔交互的硬件基础,而准确地获取光点信号则是交互的前提.提取图像中激光点的经典算法是差分图像法^[3,4]:将当前帧的图像与背景图做差值并去噪后,灰度值大于指定阈值的区域就是激光点.

首先,采用FPGA处理图像,获取所有图像中的亮点位置,并根据最邻近原则合并为4个激光点区域,取每个光点区域的质心作为光点的坐标,跟踪连续帧的中心激光点位置可获得激光点的运动轨迹(图4).

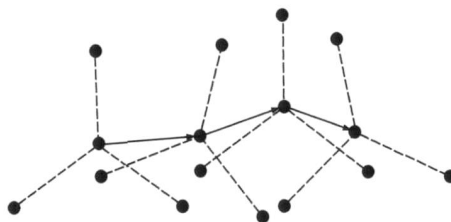


图4 跟踪激光点轨迹

同时,当用户的手绕前臂为轴旋转时,激光点组也会表现出相应的旋转,累加外围三个激光点相对于中心激光点的偏转角并求均值可提取出用户手势中的旋转分量(*Rotate*),如图5所示.

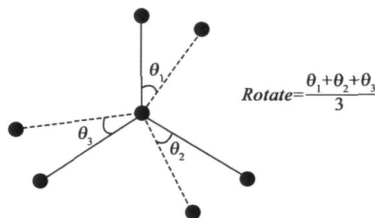


图5 *Rotate*分量提取

4 增强手势交互语义

准确地获取激光点信息并提取出光点运动轨迹和旋转手势后,结构光源能够表示的信息就很明确了:大屏幕上各个光点的坐标(x_i, y_i)、光点组的旋转信息(即用户手势中的*Rotate*分量),*Rotate*分量可以用来增强手势的交互语义.在接下来的工作中,我们主要用结构光源来实现鼠标的功能,从人类的习惯考虑,可定制以下交互语义.

空状态:初始状态,没有激光点或者激光点已经消失一段时间;

光标移动:中心激光点在大屏幕上的位置发生变化;

光标静止:中心激光点静止或者小幅度地来回移动;

左键按下:向左的*Rotate*分量连续*N*次大于的

角度阈值 $LeftBtnDTh$;

左键弹起:在左键按下的情况下,向左的 $Rotate$ 分量变为小于阈值 $LeftBtnUTh$;

右键按下:向右的 $Rotate$ 分量连续 N 次大于 $RightBtnDTh$;

右键弹起:在右键按下的情况下,向右的 $Rotate$ 分量变为小于阈值 $RightBtnUTh$;

单击左键:在指定的时间 t 内,向左的 $Rotate$ 分量完成一次从小于 $LeftBtnDTh$ 到大于 $LeftBtnDTh$ 再到小于 $LeftBtnUTh$ 的变化过程。

单击右键:在指定的时间 t 内,向右的 $Rotate$ 分量完成一次从小于 $RightBtnDTh$ 到大于 $RightBtnDTh$ 再到小于 $RightBtnUTh$ 的变化过程。

其中 $和$ 并不能由 $组合而成$ (实际上这四种状态主要用于拖拽)。可以看到, $Rotate$ 分量的加入,极大地扩展了手势的交互语义。

人机交互的实现利用有限状态机,定义 6 种激光点的状态:

(a) 激光点出现; (b) 激光点静止; (c) 激光点移动; (d) 激光点组向左旋转; (e) 激光点组向右旋转; (f) 激光点消失。

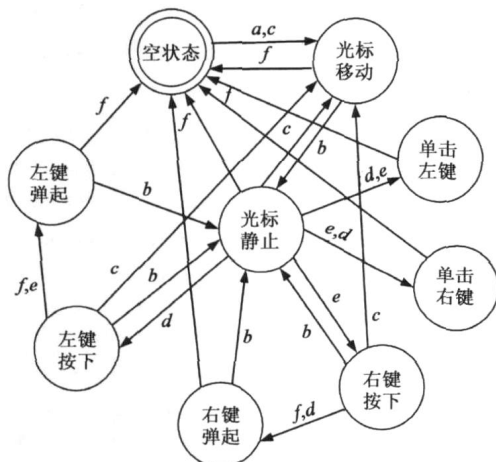


图6 有限状态机及输出

图6显示了有限状态机的状态以及输出,系统从空状态开始,当激光点出现就会进入光标静止或者是光标移动的状态;此后随着 $Rotate$ 分量的变化,系统可以在空状态、光标移动、光标静止、鼠标左右键按下和弹起等状态之间不断地切换,在规定时间内 t 内如果激光点状态 (d) (e) 依次出现,就会触发左键或者右键的单击消息,同时系统回到空状态。当激光点消失时,系统会自动进入空状态,所以说本系统始终是稳定的。

5 目标点的确定

在所有的基于激光笔的交互模式中,准确地确定要触发消息的目标点一直是个难题,我们提出了检测

光标位置序列中平坦区段的方法确定目标点:

当光标接近或到达目标点时,用户会慢下来确定是否已经指向目标,光标会在某个平衡位置附近小范围地抖动,表现在光标位置的时间序列上就是一个平坦区段,如图7所示。检测到平坦区段后,识别程序会自动记录当前光标位置并在光标的位置上画一红色圆圈提醒用户,如果用户目测坐标位置无误,只需要执行相应的手势操作即可,此时无论光标位置距离目标点有多大偏移,只要在摄像机的采集范围内,就能够正确地目标点触发消息。

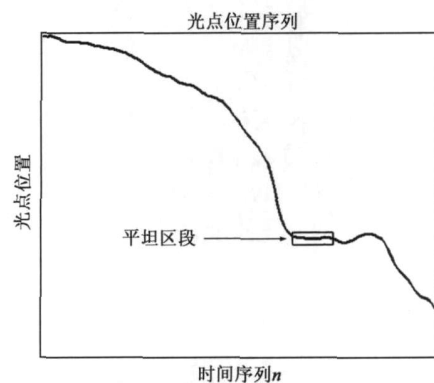


图7 光标位置序列中的平坦区段

6 用户体验测试

设计出结构光源的交互模式后,我们还设计了交互任务并请测试人群进行用户体验测试,以评估它在交互方面的性能,判断结构光源是否适合大屏幕交互。

6.1 测试程序

根据 Fitts 律的“dotting task”理论^[5],参考了应用程序^[6],我们编写测试程序的界面如图8所示。屏幕上会随机出现宽度为 30、50、80 像素(在大屏幕上的宽度为 4.3cm、7.3cm、11.7cm)的一对不同颜色的矩形区域,它们之间的距离有三种可能:240、360、480 像素(在大屏幕上的距离为 35.1cm、52.7cm、70.3cm)。测试者需要点击其中的绿色矩形条,测试程序会自动记录从目标出现到光标第一次移动到绿色矩形区域所花费的时间和移动到矩形条区域后到触发左键单击消息所花费的时间,然后继续出现新的矩形条以供测试。

6.2 测试人群

测试者站在距离大屏幕 4 米的地方,用三种交互设备进行测试,分别是鼠标、结构光源、安装有按键的激光笔,参与测试人员要服从以下要求:

(1) 站立操作。

鼠标并不是一个完整的输入设备,需要桌面的支持,所以可以放置在桌面上,但是测试者必须站立使用。

(2) 操作时要注意大屏幕。

在使用三种交互设备的过程中,用户不允许靠注视显示器完成任务,而应当一直注视大屏幕。

(3) 学习新型交互设备的使用方法。

每位用户都要事先接受结构光源和激光笔的使用方法培训至可自主操作,而不用参加鼠标的使用方法培训。

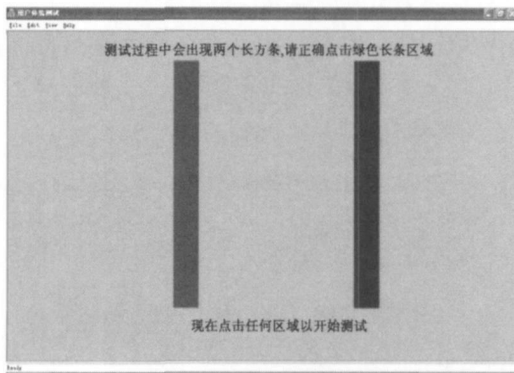


图8 用户体验测试程序

6.3 测试结果

经过测试,三种设备所花费的时间以及错误率如表1所示。

表1 用户体验测试结果

交互设备	移动到目标 时间/ms	正确点击目标的 时间/ms	错误率/%
鼠标	363	400	2.43
LP	546	538	5.90
SLP	540	590	7.23

从时间性能上来说,在将光标移动到目标位置时,使用结构光源与使用激光笔没有区别,而且都比鼠标要慢,这主要因为用户打开激光笔电源时激光束的指向是不可预知的;触发消息(表现在正确点击目标需要的速度)比使用鼠标和装有按键的激光笔也都要慢一些,但是比单点激光笔利用驻停光点来触发消息(至少需要1秒^[7])的速度要快了很多。

从错误率上来讲,使用结构光源的错误率仍然要高于其它两种设备。这主要是因为利用旋转手势触发消息时,光点会偏离目标点较大的距离,虽然检测平坦区域的方法能够在很大程度上提高选定目标的准确度,但仍然会存在误识别的情况。

测试者在最后的调查问卷中一致认为:使用结构光源进行交互,方式比较自然,不必要去记忆按键的位置,也不像站立使用鼠标那样不自然。

7 结束语

结构光源能够表达手部绕前臂的旋转运动,将旋转手势解释成交互原语并触发系统消息,可比普通的激光笔提供更多的交互手段。本文将旋转手势映射成鼠标左右键消息并进行了用户体验测试,结果表明,使

用结构光源进行交互,从时间性能和错误率上都可以接受,并且其交互方式更符合人类的习惯。结构光源还具有良好的可扩展性,将旋转手势解释成不同的命令原语就可以实现更多的交互功能。

参考文献:

- [1] Sergey V Matveyev, Martin Göbel. Direct interaction based on a two-point laser pointer technique [A]. Proceedings of The SIGGRAPH 2003 Conference on Sketches & Applications: in Conjunction with The 30th Annual Conference on Computer Graphics and Interactive Techniques [C]. New York, NY, USA: ACM, 2003. 1 - 1.
- [2] Sergey V Matveyev, Martin Göbel. The optical tweezers: multiple-point interaction technique [A]. Virtual Reality Software and Technology & Proceedings of The ACM Symposium on Virtual Reality Software and Technology, SESSION: Devices [C]. New York, NY, USA: ACM, 2003. 184 - 187.
- [3] 刘芳,等. 基于激光笔的远程人机交互技术[J]. 中国图像图形学报, 2003, 8(11): 1356 - 1360.
- [4] 清华大学. 嵌入式屏幕光点定位装置[P]. 中国专利: CN101025656, 2007-08-29.
- [5] Paul M. Fitts, The information capacity of the human motor system in controlling amplitude of movement [J]. Journal of Experimental Psychology: General, 1992, 121(3): 262 - 269.
- [6] A Web-based Test of fitts' law [CP]. <http://www.tele-actor.net/cgi-bin/fitts/applet1.pl>
- [7] Brad A Myers, Rishi Bhatnagar, Jeffrey Nichols, Choon Hong Peck, Dave Kong, Robert Miller, A Chris Long. Interacting at a distance: measuring the performance of laser pointers and other devices [A]. Proceedings of the SIGCHI conference on Human factors in computing systems [C]. New York, NY, USA: ACM, 2002. 33 - 40.

作者简介:



秦永强 男, 1986年出生于河北保定, 2008年毕业于北京邮电大学自动化学院, 现在清华大学攻读博士学位, 从事普适计算中的人机交互研究。

E-mail: qinyongqiang@gmail.com



喻纯 男, 1984年出生于湖北荆州, 现为清华大学博士生, 从事普适计算中的人机交互研究。