

文章编号: 1001- 2486(2010) 01- 0127- 06

基于多点触摸的自然手势识别方法研究^{*}

凌云翔, 张国华, 李 锐, 叶 挺
(国防科技大学 信息系统与管理学院, 湖南 长沙 410073)

摘 要: 基于指挥所中指挥作业应用需求, 定义了适于指挥所中人机交互的触摸手势; 并在对完成触摸手势过程中触摸点相互位置关系进行统计分析的基础上, 设计了一套对硬件设备特性依赖较小的多点触摸交互手势识别方法。在一种红外多点触摸交互平台上对整个交互手势的定义与识别进行了实验测试, 验证了该交互方法的可靠性、有效性和可扩展性。

关键词: 多点触摸; 人机交互; 指挥所
中图分类号: TP393 **文献标识码:** A

Research on Natural Gesture Recognition Method Based on Multi-touch

LING Yun-xiang, ZHANG Guo-hua, LI Rui, YE Ting
(College of Information System and Management, National Univ. of Defense Technology, Changsha 410073, China)

Abstract: A set of interaction handed-gestures were defined for Human-Machine Interaction (HMI) in command post according to the requirement of command tasks. Based on the statistical analysis of coordination relationship between touch points in the course of accomplishing the touch gestures, the current research designed a recognition method which is independent of hardware for these interaction handed-gestures. Finally, aimed at the definition and identification of interaction handed-gestures, some experiments had been tested on infrared multi-point touch table to verify the stability, validity and expansibility of this method.

Key words: multi-touch; HMI(Human-Machine Interaction); command post

目前, 触摸屏作为一种代替或补充普通键盘和鼠标的输入设备已经在许多场合和领域得到了使用, 尤其是新一代支持多点触摸交互的触摸屏的出现, 使触摸屏的应用出现了新的变化, 触摸屏可以为用户提供更多、更自然的交互方式。目前业界主流技术和产品有 DiamondTouch、FTIR-Touch 和 Microsoft Surface 等。DiamondTouch^[1] 是三菱电子研究实验室(MERL) 2004 年设计开发的一个支持多用户并发输入和手势交互的软硬件平台, 其采用了电感应原理, 触摸精度较低, 屏幕显示面积也受到局限。FTIR (Frustrated Total Internal Reflection) Touch^[2] 是纽约大学 2006 设计的一种支持多点触摸的硬件平台, 其采用了受抑内全反射技术, 凹凸不平的手指表面导致光束产生散射, 散射光透过触摸屏后到达光电传感器, 光电传感器将光信号转变为电信号, 系统由此获得相应的触摸信息, 目前只能使用在投影显示系统上。Microsoft Surface^[3] 是微软公司 2007 年推出的支持多点触摸和手势输入的智能桌面系统, 它采用了图像处理技术实现多点触摸, 在 Surface 平台内部有五个照相机拍摄手指在桌面触摸产生的红外反射光, 根据所拍摄图像中光点的位置对触摸点进行定位, 这些相机还能识别一些放在桌面上带有特殊标志的物体。

多点触摸技术的引入使得以触摸屏为代表的人机交互方式更加便利和自然, 但同时也带来了如何识别基于多点触摸双手交互手势的问题, 多点触摸原理的不同也导致了对手势识别方法的不同。此外, 常见商用多点触摸设备提供的手势交互语义主要面向日常生活使用^[4], 缺乏针对军用, 特别是针对指挥所应用的交互语义。在指挥所应用中, 需要多点触摸自然手势的环境主要是基于地理信息系统的图上操作^[5], 如何将自然手势映射到地理信息系统的图上操作, 是自然手势识别的一个重要问题。

^{*} 收稿日期: 2009- 05- 19
基金项目: 国家 863 计划资助项目(2007 AA01Z193); 国家自然科学基金资助项目(60875048)
作者简介: 凌云翔(1972—), 男, 教授, 博士。

针对指挥所应用背景,利用自然交互手势完成轨迹相对比较单调的特点,基于指挥所图上作业应用的交互手势语义定义和触摸信号采集,本文提出了一整套简捷可适应指挥所作业的多点触摸自然手势定义与识别方法,并基于红外多点触摸设备实现了上述方法,进行了实验验证。

1 指挥所自然手势交互实现结构

如图 1 所示,指挥所中基于多点触摸的自然手势交互的实现主要包括指挥所自然手势的定义、触摸信号采集、手势识别三部分。

自然手势的定义是根据指挥作业中人机交互任务定义静态和动态两类触摸手势,建立手势集合数据库。

触摸信息采集部分利用多点触摸设备采集触摸点信号,检测触摸信号并判断触摸信号变化情况。我们构建的触摸点信号采集的硬件环境是將一组成 90 度扇形排列的红外发光二极管和 CCD 摄像机组成一个红外光感应器件(图 2),将两个红外光感应器件分别固定于显示屏幕外框任意相邻的两角;除两个感应器件共用的一边外,在显示屏幕外框的其他三边安装条形的回归反射材料。当触摸行为发生时,对光感应器件发射的红外光产生遮挡,这些遮挡导致部分光线不能被回归反射材料反射,从而在 CCD 摄像机捕捉的条状红外图像上出现阴影,多个遮挡物将产生多个阴影;对 CCD 摄像机捕捉的红外图像利用图像处理技术中的边缘检测、区域探测、区域合并和过滤等方法消除图像干扰因素;然后,对 CCD 摄像机捕捉的红外图像进行阴影探测,获得图像中阴影的位置和数目。由于两个 CCD 摄像机之间的距离和相对位置已知,故可以通过三角定位算法进行各触摸点的屏幕坐标定位,如图 3 所示。

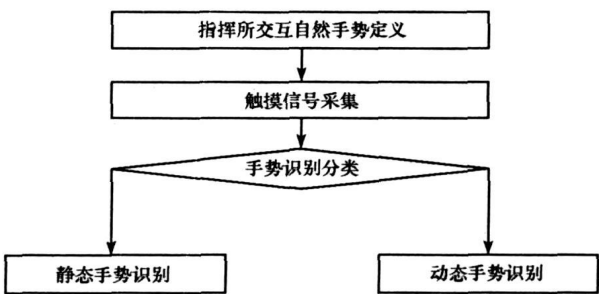


图 1 自然手势交互实现结构
Fig. 1 Natural gesture interaction method

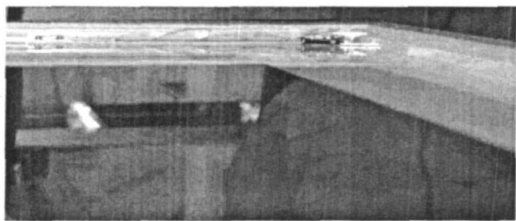


图 2 红外发光二极管与 CCD
Fig. 2 Infrared radiant diode and CCD

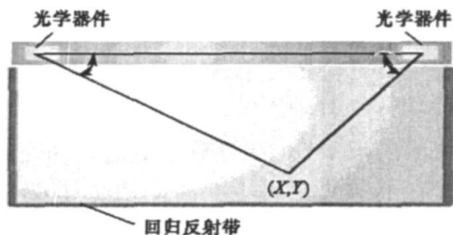


图 3 信号采集原理图
Fig. 3 Signal collection principle

自然交互手势的识别则是根据一个计算周期内(一个计算周期依据常人习惯的一次触摸平均停留时间而定,一般为 800~1200ms)采集到的各触摸点坐标是否有变化来确定。若无变化,则将该手势判为静态手势;若有变化,则将该手势判为动态手势,进而再根据不同静态手势和动态手势触点间相对位置及其变化情况决定最终识别结果。

2 指挥所自然交互手势集定义

根据指挥所作业中的人机交互任务及对手手自然张开与平面触摸时的触点位置关系特征,定义静态和动态两类触摸手势,建立手势集合数据库,包括静态触摸手势集合和动态触摸手势集合。

静态触摸手势集合包括单指静态手势、二指静态手势、三指静态手势、四指静态手势和五指静态手势。各静态手势的手指动作及操作语义如表 1 所示。为了丰富交互手势集的交互语义,对同一触点数目的交互手势又可定义不同的具体操作手势。图 4 所示为静态手势子集中的四指静态手势,四指静态手势在保证识别率的基础上又可分为单手四指和双手四指静态手势。为了便于描述,各种手势用英文

字母 L 、 R 和 A 、 B 、 C 、 D 、 E 的组合来表达, 其中 L 、 R 分别代表左手和右手, A 、 B 、 C 、 D 、 E 分别代表一只手的大拇指、食指、中指、无名指和小拇指。

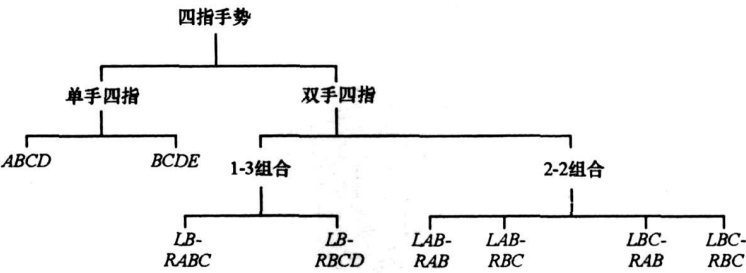


图 4 四指手势集合

Fig. 4 The collection of four fingers

表 1 静态手势语义

Tab. 1 Static gesture semantic

静态手势名称	手指动作	操作语义
单指	静态单指触摸	对象选取, 模拟鼠标左键单击等
二指	静态二指触摸	界面底色调整, 模拟鼠标左键双击等
三指	静态三指触摸	图层变换, 模拟鼠标右键单击等
四指	静态四指触摸	图像浏览, 模拟鼠标右键双击等
五指	静态五指触摸	军事目标影像图叠加, 模拟鼠标滚轮滚动等

动态触摸手势集合包括平移手势、弧移手势、双手旋转手势、单手旋转手势、缩放手势、捏取释放手势和撮取手势, 其中平移手势、弧移手势、捏取释放手势和撮取手势默认为单手手势, 缩放手势默认为双手手势, 旋转手势区分单手和双手旋转手势(表 2)。

表 2 动态手势语义

Tab. 2 Dynamic gesture semantic

动态手势名称	手指动作	操作语义
平移	手指水平或倾斜沿直线运动	作战指挥图的拖动
弧移	手指呈半圆弧运动	作战指挥图的水平、垂直翻转
双手旋转	以左手手指为中心, 右手手指绕其运动	作战指挥图在平面内的小角度旋转
单手旋转	以大拇指为中心, 食指等手指绕其运动	作战指挥图在平面内的大角度旋转
缩放	双手手指同时相向或背离运动	图像放大、缩小, 三维视点远近调整
捏取释放	模拟手指自然捏取张开动作	橡皮、画笔等对象拾取, 军标符号调整
撮取	模拟手指自然撮取动作	对象抓取

3 识别方法

识别方法的设计需基于上述定义的交互手势集合中手势的特点进行, 这样才能保证较高的识别正确率和识别效率。图 5 为静态手势识别流程。对于静态手势而言, 人手自然张开触摸时, 由于手指长度和位置的不同, 触点两两之间的距离有所不同且触点所构成多边形各顶角的角度也呈现一定的特征。经过对不同身高个体手指自然张开触摸平面时触点间距离的测量数据进行统计, 确定单手触点间的最大距离, 以此为依据将触点分为左右手两个集合。以图 4 中所定义的四指手势为例来介绍静态手势的识别方法。

以屏幕左侧第一个触摸点为基准点, 依次计算从基准点到其余各个触摸点间的距离, 当该距离小于

对于触点间有相对移位的触摸手势, 若触点数为 3 且相对距离逐渐减小, 则将其判为撮取手势(即模拟单手指 A、B、C 自然收拢抓取动作); 若触点数为 2, 相对作直线运动且连线与屏幕水平方向的夹角大于阈值 t_7 (范围为 $18\sim 20^\circ$), 则判为捏取释放手势(即模拟单手指 A、B 自然捏取和张开动作); 若触点数为 2, 相对作直线运动且连线与屏幕水平方向夹角小于等于阈值 t_7 , 则将其判为缩放手势(即双手手指 LB 和 RB 同时相向或背离做水平运动, 模拟对地图等操作对象的缩小或放大)。

判断运动触点无相对移动的手势是否有未发生移动的触点。若有, 则进一步判断模糊后的移动触点与静止触点间的连线长度变化是否小于阈值 t_5 , 同时该连线与屏幕水平方向的夹角是否大于阈值 t_6 (范围为 $8\sim 10^\circ$), 具备这些特征, 则将该手势初步判断为旋转手势。接下来, 以屏幕左侧第一个触摸点为基准点, 依次计算从基准点到其余各个触摸点间的距离, 若存在该距离大于阈值 t_1 的触点, 则进一步判断该手势为双手旋转手势, 否则, 判断为单手旋转手势(该距离小于等于阈值 t_1 的触摸点为左手触点, 大于阈值 t_1 的触摸点为右手触点)。

判断运动触点无相对移动的手势是否有未发生移动的触点。若无, 判断移动触点间无相对位移且无静止触点的手势的模糊点运动特征, 若该模糊点的运动轨迹从移动起始点至当前点的矢量与屏幕水平方向(X 轴)的夹角递减而矢量长度递增(即呈弧线运动轨迹), 则判为弧移手势, 反之判为平移手势。

图 7 表示旋转手势运动特征, 旋转过程中有一个手指(图中为 A)作为旋转中心静止, 另一个手指(图中为 B)作为运动手指绕旋转中心旋转。旋转过程中旋转半径同屏幕水平、垂直方向的夹角 δ 和 γ 会不断变化, 而旋转半径基本不变。图 8 表示单指弧移过程运动特征, 从移动起始点 O 至当前点 P、 P_1 、 P_2 、 P_3 的矢量与屏幕水平方向(X 轴)的夹角 α 逐渐减小而矢量长度逐渐增加。图 9 表示撮取动作完成过程运动特征, 三个手指 ABC 自然合拢, 手指间的距离渐渐缩小, 当触摸平台无法将三指进行区分时, 三指合拢为一个触点 G。图 10 表示捏取、释放动作完成过程的运动特征, 双指间距离(A 与 B 的距离)渐渐缩小或增大而双指间连线 AB 与屏幕水平方向(X 轴)的夹角 β 基本保持一定角度不变。图 11 表示缩放动作完成过程运动特征, 左手触点(LB)与右手触点(RB)进行相向或背离运动, 运动方向始终与屏幕水平方向(X 轴)基本保持水平。

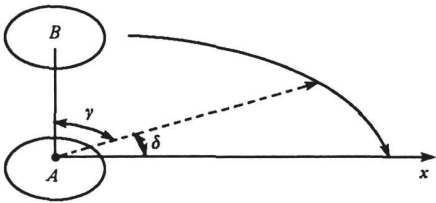


图 7 旋转手势运动特征

Fig. 7 Movement character of circumgyration gesture

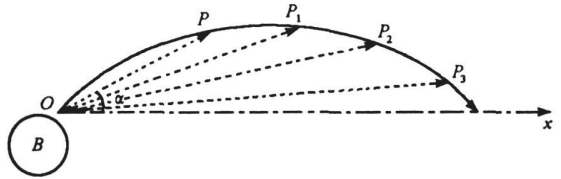


图 8 弧移手势运动特征

Fig. 8 Movement character of arc gesture

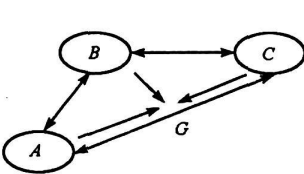


图 9 撮取手势运动特征

Fig. 9 Movement character of pinch gesture

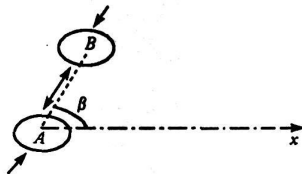


图 10 捏取释放手势运动特征

Fig. 10 Movement character of nip gesture

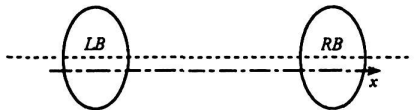


图 11 缩放手势运动特征

Fig. 11 Movement character of zoom gesture

4 识别与测试

使用 65 红外多点触摸屏作为实验设备, 对图 4 定义的四指手势、双手旋转手势、双手缩放手势、弧移手势、捏取手势、三指撮取手势进行 500 次识别试验, 结果见表 3(其中四指手势中的数据取不同手势识别结果的均值)。

表 3 实验结果

Tab.3 Experiment results

手势	四指	双手缩放	弧移	捏取	撮取	双手旋转
平均时间(s)	0.026	0.893	0.533	0.663	0.623	0.967
正确率	99.96%	99.87%	99.74%	99.78%	99.80%	99.87%

表 3 中的数据与触摸信号检测周期和识别过程中的阈值的设置等因素有很大关系,且由于动手势的识别要等触点运动趋势明确后才可进行,其识别效率明显低于静态手势。但从上述数据还是可以看到本文所述的手势定义与识别方法具有如下优点:

(1) 识别准确度高。本文基于指挥所作业应用定义触摸手势,根据指挥所作业特点区分静态手势和动态手势,手势集合定义明确,且基于解剖生理学数据设置距离阈值和运动特征参量分别对静态手势和动态手势进行识别,识别所需特征参数简单明确。因此,识别方法可以获得较高的识别准确率。

(2) 识别速度快。相比传统的方法而言,本方法没有复杂的电信号处理,资源和时间消耗相对较少,加之指挥所操作语义相对稳定,识别所需特征参数简单明确,因此手势识别响应时间较短。

(3) 可扩展性好。本方法所需特征参数较少,与硬件耦合性不紧密,基于触摸信号进行采集,记录触摸点的屏幕位置坐标,适用于各种不同的光感应多点触摸屏装置,因此通用性和可扩展性好,实现成本低。

参 考 文 献:

[1] Dietz P, Leigh D. DiamondTouch: A Multi-user Touch Technology[C]//Proceedings of ACM UIST' 01, Orlando FL, November 2001: 118– 124.

[2] Jefferson Y H. Low-cost Multi-touch Sensing Through Frustrated Total internal reflection[R]. ACM UIST, 2005: 115– 118.

[3] Ringel M, Ryall K, Shen C, et al. Release, Relocate, Reorient, Resize: Fluid Techniques for Document Sharing on Multi-user Interactive Surface [C]//CHI 2004 Extended Abstracts, 2004: 1441– 1444.

[4] 付永刚,张风军,戴国忠. 双手交互界面研究进展[J]. 计算机研究与发展, 2005, 42(4): 604– 613.

[5] 吴振峰,赵克俭. 未来指挥所发展展望[J]. 火力与指挥控制, 2005, 26(4): 12– 14.

[6] 任雅祥. 基于手势识别的人机交互发展研究[J]. 计算机工程与设计, 2006, 27(7): 1201– 1204.

(上接第 115 页)

4 结 论

在合理选择控制信号和反馈信号的情况下, DRLG 的抖动控制回路可以解耦为抖频和抖幅控制两个独立的控制回路,其中抖幅控制支路可以简化为一个一阶惯性控制回路。

参 考 文 献:

[1] Killpatrick J E. Laser Angular Rate Sensor[P]. U.S. Patent, 3,373,650, 1968.

[2] Killpatrick J E. Random Bias for Laser Angular Rate Sensor[P]. U.S. Patent, 3,467,472, 1969.

[3] McNair F, Hills W. Laser Gyro Coupling System[P]. U.S. Patent, 4,321,557,1982.

[4] Robert C, Oaks T. Dither Controller for Ring Laser Angular Rotation Sensor[P]. U.S. Patent, 4,597,667, 1986.

[5] Benoist R W, Calif M. Simplified Ring Laser Gyroscope Dither Control And Method[P]. U.S. Patent, 4,801,206,1989.

[6] Tazartes D A, Hills W, Mark J G. Ring Laser Gyroscope Dither Drive System and Method[P]. U.S. Patent, 4,981,359, 1991.

[7] 罗兵,等. 石英音叉陀螺不同谐振频率点性能分析[J]. 压电与声光, 2009(1): 312– 314.

[8] 潘献飞,等. 激光陀螺数字抖动控制方法与特性[J]. 国防科技大学学报, 2006, 28(5): 99– 103.

[9] 汤建勋,等. 用锁相环跟踪抖动机构自然频率的抖动控制方法[J]. 测控技术, 2000, 19(8): 28– 29.

[10] 朱春云,等. 机抖激光陀螺方波驱动方式的谐波分析[J]. 中国惯性技术学报, 2007, 15(6): 743– 746.