

# 基于 Kinect 骨骼跟踪技术的手势识别研究

樊景超, 周国民\* (中国农业科学院农业信息研究所, 北京 100081)

**摘要** 为了实现自然方便的人机交互, 提高用户体验, 尝试引入微软公司的一款体感周边外设 Kinect 设备, 利用 Kinect 传感器所获取的人体骨骼跟踪识别点, 通过对双手在三维空间中的位置坐标转换, 对缩放、旋转的手势进行匹配, 将合法的手势信息转换为识别结果, 反馈给系统并完成相应的控制指令输出, 实现与系统的交互控制。结果表明, 该系统具有较高的精度和稳定性, 具有一定的借鉴意义。  
**关键词** 骨骼跟踪; 手势识别; Kinect  
**中图分类号** S126 **文献标识码** A **文章编号** 0517-6611(2014)11-03444-03

**The Research of Gesture Recognition Based on Kinect Skeleton Tracking Technology**  
**FAN Jing-chao, ZHOU Guo-min** (Agricultural Information Institute, Chinese Academy of Agricultural Sciences, Beijing 100081)  
**Abstract** In order to achieve a natural human-computer interaction to facilitate and improve the user experience, this research attempt to introduce a Microsoft's peripheral Kinect somatosensory peripheral equipment. Using Kinect sensor acquired identification of the human skeleton tracking point, by the hands of the position in three dimensional space coordinate conversion for scaling, rotation gesture match. The information is converted into a valid gesture recognition result, the system and complete the feedback control command output, controlling interaction with the system. Experimental results show that the system has higher precision and stability, with a certain significance.  
**Key words** Skeleton tracking; Gesture recognition; Kinect

鼠标作为传统的图形人机交互界面的主要设备, 长期以来在计算机操作中处于主导地位, 但是鼠标只能用单手操作, 无法将人自然生活的双手操作技能应用于人机交互之中<sup>[1]</sup>。这种操作系统需要操作由图标和菜单构成的基于 GUI 的应用程序界面, 导致这种界面通常具有陡峭的学习曲线<sup>[2]</sup>。理想的自然交互界面的发起者认为交互界面的设计应该对用户非常直观, 使用用户先天就会的行为来进行交互是操作理想化的状态, 用户走到应用程序前面就能够开始使用它<sup>[3]</sup>。在过去的几年里, 随着触摸功能在智能手机和平板电脑的流行, 轻点(Taps)已经成为触摸平台的核心, 逐渐取代了键盘鼠标<sup>[4]</sup>。当看到孩子们开始走到任何触摸屏设备面前, 用手去触摸它, 期待它的响应时, 在这一点上看这一理念已经逐步实现。

为此, 笔者利用 Kinect 传感器的骨骼跟踪技术, 捕获用户手势运动数据作为控制信号, 并反馈给相应的控制软件, 再通过对手势的特征匹配从而转换为相应的控制指令并完成操作, 从而实现双手精准的手势操作来完成相应的放大、缩小和旋转功能。

## 1 Kinect 简介

**1.1 Kinect 设备** Kinect 是微软在 2010 年 6 月 14 日对 XBOX360 体感周边外设正式发布的名字, 其设备如图 1 所示。Kinect 共有 3 个摄像头, 中间一个是 RGB 摄像头用来获取 640×480 的彩色图像, 每秒钟最多获取 30 帧图像; 两边的是深度传感器, 左侧的是红外线发射器, 右侧的是红外线接收器, 用来检测玩家的相对位置。Kinect 的两侧是一组四元麦克风阵列, 用于生源定位和语音识别; 下方还有一个内置马达的底座, 可以调整俯仰角<sup>[5]</sup>。

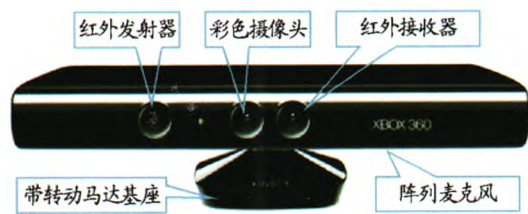


图 1 Kinect 设备

**1.2 Kinect 骨骼跟踪技术** Kinect 采集三维图像的光学部分主要包括两个部件: 红外线发射器和红外线/VGA 摄像头组。其工作原理<sup>[6]</sup>是: 红外线发射器发出红外线覆盖整个 Kinect 的可视范围, 摄像头组接收反射回光线来识别玩家的距离和特征。红外摄像头可以识别到的是物体的“深度场”(Depth Field), 其中每一像素的颜色代表了那一点物体到摄像头的距离。比如离摄像头近的身体呈亮红色、绿色等, 而离摄像头远的物体则呈暗灰色。

这是一种不同于传统的 TOF 或者结构光测量技术的光编码(light coding)技术, 使用的是连续的照明(而非脉冲), 也不需要特制的感光芯片, 而只需要普通的 CMOS 感光芯片, 这也大大降低了 Kinect 的成本<sup>[7]</sup>。光编码<sup>[8]</sup>是用光源照明给需要测量的空间编上码的一种结构光技术。但与传统的结构光方法不同, 它投射出去的光源并不是一副周期性变化的二维的图像编码, 而是一个具有三维纵深的“体编码”, 这种光源叫做激光散斑(laser speckle)。当激光照射到粗糙物体或穿透毛玻璃后会形成随机衍射斑点, 这些散斑具有高度的随机性, 而且会随着距离的不同变换图案, 所以空间中任意两处的散斑图案都是不同的。通过这种激光散斑可以对整个空间进行标识, 然后根据物体上的散斑图案就可以推测计算出物体的深度位。在进行物体散斑识别前要对整个空间的散斑图案进行记录和标定。

骨骼追踪技术通过处理景深数据来建立人体各个关节的坐标, 骨骼追踪能够确定人体的各个部分, 如哪部分是手、

**基金项目** “863”项目(2013AA102505)。  
**作者简介** 樊景超(1980-), 男, 辽宁沈阳人, 助理研究员, 博士, 从事农业信息技术的科研工作。\* 通讯作者, 研究员, 博士, 博士生导师, 从事农业信息技术的科研工作。  
**收稿日期** 2014-03-09



头部以及身体。骨骼追踪产生 X、Y、Z 数据来确定这些骨骼点。图 2 是 Kinect 骨骼跟踪技术对笔者进行跟踪显示的 20 个关节骨骼信息。

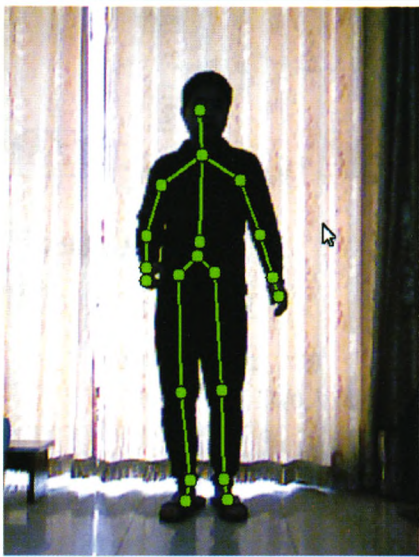


图 2 骨骼跟踪技术

2 多点触控系统设计

2.1 缩放手势定义 骨骼追踪技术的初衷:通过人体骨骼的关节点来控制人体运动<sup>[10]</sup>。在多点触控系统中缩放手势包括放大、缩小和停止 3 个。3 个手势的共同特征是双臂手关节都在肘关节以上,当双手不符合这一基本条件时认为是非缩放手势。3 个手势的主要差别是双臂与肘关节垂直中线的夹角的不同,这里的夹角根据左右是对称的关系,左手夹角的值要乘以 -1 来判定手臂运动的方向。据此定义多点触控系统中缩放手势的识别算法如图 3 所示。

步骤 1:系统在检测到人体的骨骼信息之后,首先判断双手节点是否在肘关节之上,如果是则进入步骤 2;否则重复步骤 1。

步骤 2:判断当前双臂是否垂直作为缩放手势的启动操作,是则进入步骤 3;否则退回到步骤 1。

步骤 3:用户挥动双手手臂进行缩放操作,系统根据双臂与肘关节垂直中线的夹角判断用户的指令信息。当  $15^{\circ} \leq$  夹角度数  $\leq 45^{\circ}$  时,认为是缩小手势;当  $-15^{\circ} \leq$  夹角度数  $\leq 15^{\circ}$  时,认为是停止手势;当  $-45^{\circ} \leq$  夹角度数  $\leq -15^{\circ}$  时,认为是放大手势。

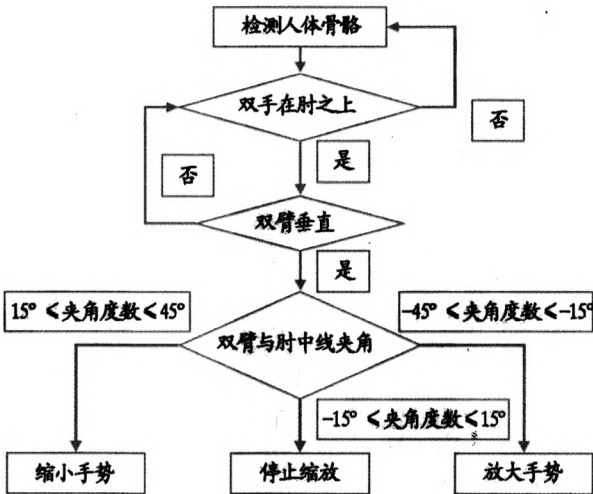


图 3 缩放手势识别算法

图 4、5 和 6 是笔者在室内利用 Kinect SDK 对自身在运动状态下进行的手势识别测试截图。笔者对双手的放大、缩小和停止进行了测试,结果显示,可以系统良好地识别并跟踪人的双手手势变换。



图 4 放大手势识别

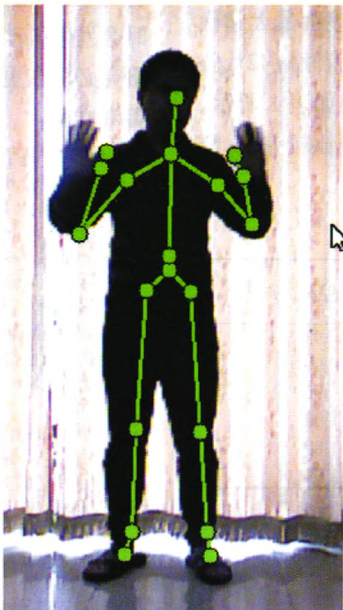


图 5 停止手势识别

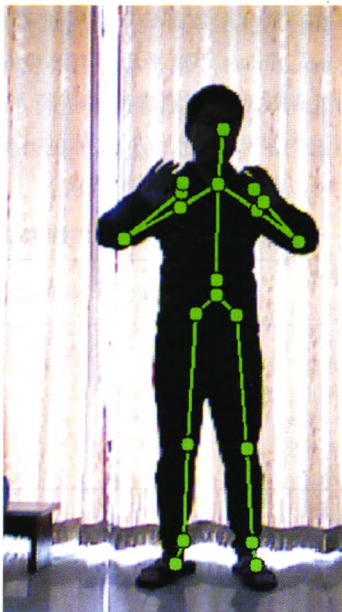


图 6 缩小手势识别

2.2 旋转手势定义 在多点触控系统中旋转手势包括顺时针旋转、逆时针旋转和停止旋转 3 个。3 个手势的共同特征是双臂手关节都在肘关节的外侧,当双手不符合这一基本条

件时认为是非旋转手势。3 个手势的主要差别是双臂与水平线的夹角的不同,这里的夹角根据左右是对称的关系,左手夹角的值要乘以 -1 来判定手臂运动的方向。据此定义多点



触控系统中旋转手势的识别算法如图 7 所示。

步骤 1:系统在检测到人体的骨骼信息之后,首先判断双手关节是否在肘关节外侧,如果是则进入步骤 2;否则重复步骤 1。

步骤 2:判断当前双手是否水平作为旋转手势的启动操作,是则进入步骤 3;否则退回到步骤 1。

步骤 3:用户挥动双手手臂进行选择操作,系统根据双手与水平线的夹角判断用户的指令信息。当  $15^{\circ} \leq \text{夹角度数} \leq 45^{\circ}$  时,认为是逆时针旋转手势;当  $-15^{\circ} \leq \text{夹角度数} \leq 15^{\circ}$  时,认为是停止旋转手势;当  $-45^{\circ} \leq \text{夹角度数} \leq -15^{\circ}$  时,认为是顺时针旋转手势。

图 8、9 和 10 是笔者在室内利用 Kinect SDK 对自身在运动状态下进行的手势识别测试截图。笔者对双手的顺时针、逆时针和停止进行旋转手势测试,结果显示,可以系统

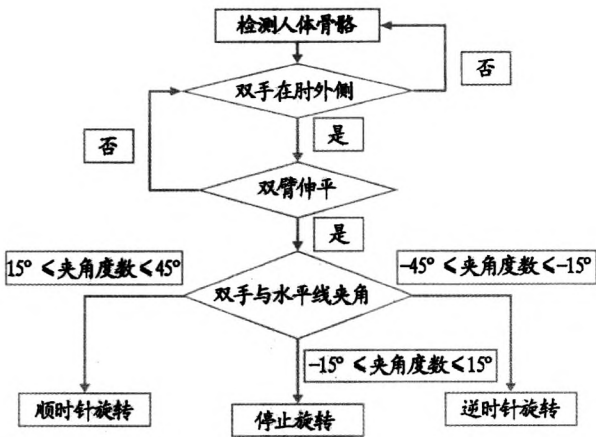


图 7 旋转手势识别算法  
良好地识别并跟踪人的双手手势变换。



图 8 顺时针旋转手势识别



图 9 停止旋转手势识别



图 10 面后退手势识别

3 测试结果

在普通实验室环境下进行试验,测试平面表面最大起伏  $\leq 3\text{ mm}$ , Kinect 距离平面垂直高度  $650\text{ mm}$ ,垂直倾角  $42^{\circ}$ ,分别对处理效率和识别准确率进行了试验,结果如表 1 所示。从测试的结果看,系统对肢体的几何信息识别比较完整,精确度可以满足手势识别的需求。

表 1 多点触控测试结果

多点触控 动作类型	测试次数	识别次数	识别率/%
放大手势	100	93	93
缩小手势	100	90	90
缩放停止	100	95	95
逆时针旋转	100	92	92
顺时针旋转	100	89	89
旋转停止	100	91	91

4 结论

该研究基于 Kinect 精确捕获被测对象骨骼信息空间坐标数据的方法,实现了对人体手势运动特征的识别。该设计所采用的硬件设备具有低成本、高效率、非依赖性等特点。随着虚拟现实在实际应用的进一步推广,自然交互技术在软件中的应用必将越来越深入,交互方式也会更直观、便捷。同时该研究还有一定的不足,由于 Kinect 本身在环境光线较

弱时,对识别对象的检测没有较好的防抖处理,所以在操作时会发生抖动产生的误差干扰,有待于进一步完善。相信该研究必将在更广阔的领域得到更深层次的应用。

参考文献

[1] 蔡睿妍. 基于 kinect 多点触控系统的研究与实现[J]. 计算机工程应用技术, 2012, 8(16): 3987 - 3989.

[2] [译] Kinect for Windows SDK 入门( + ): 手势识别上: 基本概念[EB/OL]. [http://www.cnblogs.com/yangcnu/archive/2012/04/21/KinectSDK\\_GesturesDetection\\_part1\\_BasicConception.html](http://www.cnblogs.com/yangcnu/archive/2012/04/21/KinectSDK_GesturesDetection_part1_BasicConception.html).

[3] 吕开阳. Kinect 体感技术在动物外科手术教学中的应用及展望[J]. 中国医学教育技术, 2012, 26(2): 171 - 173.

[4] 黄康泉, 陈碧金, 郑博, 等. Kinect 在视频会议系统中的应用[J]. 广西大学学报: 自然科学版, 2011, 36(S1): 308 - 314.

[5] 吴国斌. Kinect 人机交互开发实践[M]. 北京: 人民邮电出版社, 2013: 4 - 5.

[6] 余涛. Kinect 应用开发实践[M]. 北京: 机械工业出版社, 2012: 41 - 45.

[7] 韩旭. 应用 Kinect 的人体行为识别方法研究与系统设计[D]. 济南: 山东大学, 2013: 21 - 22.

[8] 邓小园. 基于 Kinect 运动捕捉的高尔夫挥杆分析与辅助训练系统的研制[D]. 北京: 北京邮电大学, 2013: 7 - 9.

[9] 杨林. 基于 Kinect 的人体目标检测与跟踪[D]. 大连: 大连海事大学, 2013: 8 - 15.


[10] 马庆. 基于 Kinect 的实时人体动画合成及重定向技术研究[D]. 杭州: 浙江工业大学, 2012: 11 - 15.

[11] 杨晓敏, 张奇志, 周亚丽. 基于 Kinect 深度信息的人体运动跟踪算法[J]. 北京信息科技大学学报, 2013, 28(1): 33 - 36.

[12] 陶丽君, 李翠华, 张希婧, 等. 基于 Kinect 传感器深度信息的动态手势识别[J]. 厦门大学学报: 自然科学版, 2013, 52(4): 493.

[13] 黄露丹, 严利民. 基于 Kinect 深度数据的人物检测[J]. 计算机技术与发展, 2013, 23(4): 121 - 124.

# 基于Kinect骨骼跟踪技术的手势识别研究

作者：[樊景超](#)，[周国民](#)，[FAN Jing-chao](#)，[ZHOU Guo-min](#)  
作者单位：[中国农业科学院农业信息研究所, 北京, 100081](#)  
刊名：[安徽农业科学](#)   
英文刊名：[Journal of Anhui Agricultural Sciences](#)  
年，卷(期)：2014(11)

本文链接：[http://d.g.wanfangdata.com.cn/Periodical\\_ahnykx201411107.aspx](http://d.g.wanfangdata.com.cn/Periodical_ahnykx201411107.aspx)