

手势识别技术在物联网课程教学中的应用

薛莹

(上海电子工业学校, 上海 201411)

摘要: 人体手势是一种自然并且直观的人际交流模式, 最新的Kinect设备可提供一种新的人机交互的方式, 能够捕捉、跟踪以及解密人体的动作、手势以及声音。文章利用Kinect进行人体手势识别, 在物联网课程的日常教学上提供无接触式互动。

关键词: Kinect; 隐马尔科夫模型; 手势识别; 教学应用

在物联网课程教学过程中信息化使用越来越广泛, 大多数情况是使用鼠标和键盘来控制计算机中的文字、图片、视频等为载体的信息。通常会让教师固定在某一个特定位置上课, 容易让学生感到枯燥, 单板, 不感兴趣。如果教师在上课过程中面对面地跟学生交流, 结合信息化使用加入人体手势识别并控制这些信息, 达到人机合一效果, 使教师摆脱键盘、鼠标等硬件设施的限制, 可以让学生更加真实地享受虚拟化课堂带来的乐趣, 体验物联网最新的技术。

1 手势识别控制原理

手势识别以手势识别代替鼠标、键盘来控制计算机。鼠标控制可以通过骨骼节点信息进行控制。键盘控制有4个过程: 手势存储, 手势获得, 手势识别, 转化为控制命令。Kinect for windows SDK是微软提供的软件开发工具包, 获取双手骨骼节点信息和深度信息流, 经过计算机预处理后加入学习库中, 通过隐马尔科夫模型(HMM)进行学习, 当人进行运动时实时识别。操作系统根据已获得识别手势得到控制指令, 达到控制键盘的效果(见图1)。

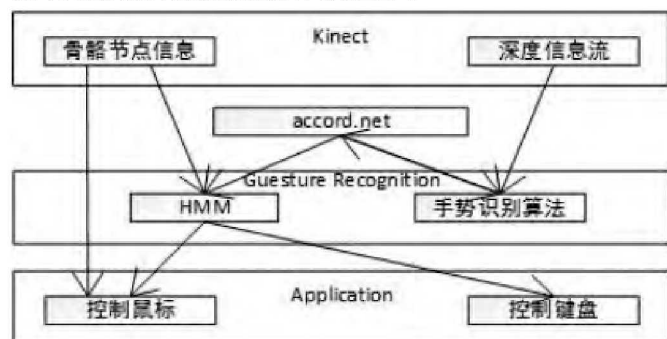


图1 人体手势识别控制结构

1.1 静态手势实现

深度信息流对手势进行分割, 从8个方向依次判断是否为边界, 将每个角落的边界提取出来进行处理, 边界点绘制成蓝色。对于不同手形以及不同距离, 对分割后的手势都能进行较好的轮廓绘制。使用K-curvature算法对指尖进行判断。轮廓边界点个数可以自己定义K参数的值, 现实使用中需调试该值到最佳, 对识别效率影响很大。计算机识别过程中将背景过滤, 深度信息将手臂从身体中剥离出来, 进行手势识别算法。获得的信息通过accord.net和骨骼节点信息获得后由隐马尔科夫模型中的已存储的手势进行判断, 控制鼠标

和键盘。

在使用过程中人跟Kinect设备之间的距离会对识别效果产生影响。静态手势的测试过程中随着手和设备之间距离越远, 获得图片宽度会越小, 距离越近, 获得图片宽度越大, 识别率上升。Kinect for windows第二版提供在0.5~4.5m范围内深度信息, 跟踪人数多达6人, 每人25个关节, 还包括新的关节部位如手指尖、拇指、肩膀中部, 它能提高软结缔组织和身体定位。

1.2 动态手势实现

在使用Kinect上课之前需要进行手势存储, 由训练得到自己需要的手势库, 将手势库以.dat的格式存储。使用耦合隐马尔科夫模型手势识别, 根据统计方法实现, 当有足够多得手势样本时, 程序不需要每次重新获得。当添加新的动作之后, 需对原有.dat文件更新, 操作系统才能对新手势进行识别。

1.3 软件实现

在PC机上安装Windows 8以上64位操作系统、.NET Framework 4.0框架、Visual Studio 2010以上版本和Kinect for Windows SDK。

软件实现步骤可分为:

步骤1: 打开Visual Studio 2010, 创建一个WPF工程。

添加Kinect程序集的引用在Solution Explorer中, 右键单击Kinect Wpf, 在右键菜单中选择“Add Reference...”。在弹出的对话框中, 在.NET标签页里, 选择“Microsoft.Research.Kinect”程序集。然后添加Coding Fun Kinect Toolkit。

步骤2: 打开MainWindow.xaml, 添加2个Image控件在设计器中, 调用Runtime对象获得获取RGB Camera数据。一个用于显示获得实时的图像信息, 还有一个用于显示景深信息。在程序中声明一个事件处理方法, video Frame Ready事件会传递一个Image Frame Ready Event Args参数, 其目的是获取实时的图像数据。有以下重要变量, Image Frame变量包含图片的各种信息, Type指定图像是来自实时的图像还是景深, Resolution指定分辨率, Image保存了图像的真实数据以byte[]数组的方式。根据Planar Image中数据创建一个图像对象, 传递给Image控件, 并显示到程序的界面上。还有一个控件获取景深信息, 过程与上面一个过程

作者简介: 薛莹(1981-), 男, 上海, 工程硕士, 讲师, 工程师; 研究方向: 物联网技术, 计算机网络。

类似。当Runtime对象被初始化后,添加了Runtime Options. Use Depth属性,添加获取景深信数据的事件处理,打开景深信的数据流,并处理信息。

步骤3: 实时获取硬件设备传来的骨骼数据流,将人体骨骼数据转换为骨骼点功能,该类识别人物骨骼数据,可以实现相关的内容浏览操作。在初始化时,创建Runtime对象指定Use Skeletal Tracking的Runtime Options,然后在Skeleton Frame Ready事件中添加处理函数。定义具体如下: 通过获取Kinect最前面的人物身体并且活动手的骨骼数据,调用系统动态库(user32. dll),在屏幕上模拟鼠标移动,当手在指定时间内完成从身体右侧挥动到身体左侧时,通过系统库模拟鼠标向左事件,当手在指定时间内完成从身体左侧挥动到身体右侧时,通过系统库模拟鼠标向右事件。

2 手势识别在物联网课程中的应用

首先定义一些基本手势,并将这些手势加入手势库中,键盘功能通过手势识别获得。鼠标功能通过手的位置移动来控制。通过两者的有效结合达到对计算机的无接触式控制。

在教室或实验室中的布置有2种方式实现。第1种电脑在讲台周边, Kinect连接到讲台的略上方或者下方正对教师,教师通过观察电脑屏幕来进行这操作,操作画面通过投影投射到背后的屏幕中,如图2所示。笔者不推荐使用这种方式教学,它不能体现出Kinect的优势,并且学生互动较差,每次演示还需指定特定区域。

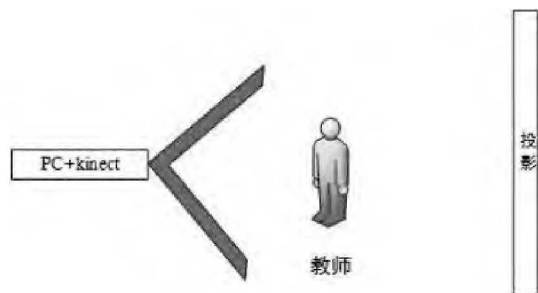


图2 背向投影屏幕操作示意

第2种情况, Kinect放置在投影屏幕或电子白板下方或者正上方,教师是背向投影屏幕或电子白板,面对学生,可捕捉到教师和大部分同学的手势,如图3所示。笔者建议使用这种方式教学,学生更容易接受,并且学生起立后也可完成对计算机的操作,形成互动。

在日常教学中的具体体现在对操作系统的控制,如: 控制鼠标、键盘。应用程序的控制,如: 幻灯片的播放、文档演示、视频播放等。

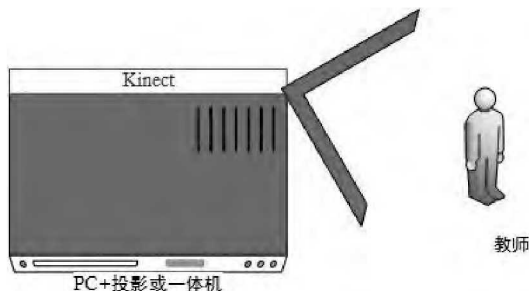


图3 面向投影屏幕操作示意

2.1 预设键盘操作及应用

NUI接口定义人体手势操作,以右手动态手势为例。先完成字母、数字手势,主要字符有a~z,数字手势0~9。然后完成方向型手势,上、下、左、右和前推、后拉。其中由于字母“1”、数字“1”手势与向下方向键重合,所以在“1”和“1”的下方加了个不同方向的勾以作区分。最后是动态双手手势,有放大、缩小、顺时针旋转和逆时针旋转等。其中部分手势如图4所示,预定义的手势图中的箭头都是标记手的运动方向向量,其中用标记的为手势为起始特征方向向量。

右	左	上
数字 0	数字 1	数字2
双手缩小	双手放大	双手顺时针寻转

图4 部分预定义的手势

2.2 鼠标操作应用

根据Kinect提供的内容定义鼠标可以通过右手节点的位置移动进行调节光标的上下左右,右手固定在一个位置保持不动达到一定时间,鼠标定义为右击和双击。左手举起高于肩膀时,定义为点击鼠标左键,手臂放下,左键释放。

2.3 幻灯片控制应用

幻灯片常用控制功有下一页,选择指针选项等,如表1所示。在打开PPT以后,按F5播放幻灯片。如果手向右挥动一定距离后,后右手停止,系统会识别并判断手势,比较HMM模型内手势库内容,找出最接近的手势,并发出并“下一页”指令,PPT会自动向下翻一页。

表1 动态手势控制PPT的功能对照表

功能	下一页	上一页	激光笔	笔	银光笔	跳入指定页
手势	向右	向左	字符L	字符P	字符H	数字手势

2.4 对于其他程序的应用

其他应用程序主要是针对鼠标和键盘的控制,再额外增加放大和缩小等手势,这些都需要操作系统的支持。如使用SmoothDraw软件结合手势,直接可以在屏幕上书写,具有板书功能,可替代电子白板,更接近于自然的三维的操作。如某些看图软件,可以增加放大、缩小手势等双手动态手势对图片进行操作。如某些文档上可以“n”手势为创建一个新文件的手势,以“x”手势为关闭文件手势。

3 结语

在实际使用过程中,由于深度信息、前排学生干扰等因素导致人体手势识别率并不高,经笔者测试识别率进为83%。在上课使用过程中,可提高教学的趣味性,学生互动踊跃,

(下转第124页)

方法^[3]的 $3K^2-K$ 相差不大。当 $3 < i < 5$ 时, SSOR预编码复杂度分别为 $6K^2+6K$, $8K^2+8K$, $10K^2+10K$ 相比较Neumann预编码的 K_3 , $2K^3-K^2$, $3K^3-2K^2$ 有一个量级的提升。

3 结论

本文将 $\omega_{SS} = 10^{-4}$ 设定为阈值, 当迭代次数为5次时, 停止迭代, 比较了迭代次数为2到5次时2种方法的区别, 由上述数据可得, 当迭代次数 $i > 2$ 时, 基于Neumann的预编码计算复杂度

为 $O(K^3)$, 对RZF预编码的复杂度改善并不明显。而对任意的迭代次数来说, 基于SSOR的预编码的计算复杂度都是 $O(K^2)$ 。当迭代次数 i 越来越大, 本文提出的基于SSOR的预编码方法通过迭代的方式, 避免了大尺寸矩阵的求逆, 相比基于Neumann的预编码方法降低了一个量级的复杂度。同时, 自适应的迭代方法, 也避免了过多迭代带来的计算开销以及迭代不够造成的性能缺失。

[参考文献]

- [1] J. Hoydis, S. ten Brink, and M. Debbah. Massive MIMO in the UL/DL of Cellular Networks: How Many Antennas Do We Need? [J]. IEEE J. Sel. Areas Commun., 2013 (2): 160-171.
- [2] A. Björck. Numerical methods in matrix computations [J]. Texts in Applied Mathematics, 2015 (6): 910-916.
- [3] S. Wagner, R. Couillet, M. Debbah, et al. Large System Analysis of Linear Precoding in MISO Broadcast Channels with Limited Feedback [J]. IEEE Trans. Inf. Theory, 2012 (7): 4509-4537.
- [4] M. Wu, B. Yin, G. Wang, et al. Large-scale MIMO detection for 3GPP LTE: Algorithms and FPGA implementations [J]. IEEE J. Sel. Topics Signal Process, 2014 (5): 916-929.

A Linear Precoding Scheme for Massive MIMO Systems Based on SSOR Method

Long Ken, Qing Ruiqiang, Tu Siyu

(Broadband Mobile Communication Mobilization Center,

Chongqing University of Posts and Telecommunications, Chongqing 400065, China)

Abstract: Massive multiple-input multiple-output (MIMO) has a large number of antennas and users, which leads to high computational complexity in precoding matrix. To solve this problem, this paper proposed a linear precoding scheme based on symmetric successive overrelaxation (SSOR). Compared to conventional RZF scheme, it can reduce one order of magnitude with negligible performance loss. To guarantee the performance of SSOR-based precoding, we also proposed a simple quantified relaxation parameter for SSOR-based scheme, which only depends on the MIMO system configuration.

Key words: massive mimo; RZF; symmetric successive overrelaxation; computational complexity

(上接第105页)

提高了学生的参与度, 只需在座位上起立就可以完成控制示教机。这种教学方式用于和学生互动性强的课程, 如物联网工程概论、英语、德育课程等, 其他课程有待考证。如体感技

术与其他技术相结合, 如人脸识别技术可以自动点名等, 从而带来全新的教学体验。

[参考文献]

- [1] 狄海进. 基于三维视觉的手势跟踪及人机交互中的应用 [D]. 南京: 南京大学, 2011.
- [2] 吴国斌, 李斌, 阎骥洲. KINECT 人机交互开发实践 [M]. 北京: 人民邮电出版社, 2013.
- [3] 倪晨, 邱鹏, 王锋, 等. Kinect 体感技术在人机交互中的应用研究 [J]. 信息技术与信息化, 2013 (1): 87-90.
- [4] 陈静, 陈泽宇, 王敏娟, 等. 基于 Kinect 的体感技术在教学领域中的应用 [A]. 第五届教育教学改革与管理工程学术年会 [C]. 2012.
- [5] 钱鹤庆. 应用 Kinect 与手势识别的增强现实教育辅助系统 [D]. 上海: 上海交通大学, 2011.

The Application of Gesture Recognition Technology in the Internet of Things Courses

Xue Ying

(Shanghai Electronics Industry School, Shanghai 201411, China)

Abstract: Human gesture is a natural and intuitive communication mode. The newest Kinect equipment can provide a new human-computer interaction way which can capture, track and decrypt the body movements, gestures and voice. This paper uses Kinect equipment for human gesture recognition, with no contact interaction in the Internet of things in the course of daily teaching.

Key words: Kinect; HMM; gesture recognition; teaching application