

基于Kinect的展馆交互系统设计与应用

牟 莉, 史政坤, 陈逍遙

(西安工程大学计算机科学学院, 陕西 西安 710000)

摘 要: 当前, 很多城市展馆的展览方式具有单调、缺乏科技感和交互性不强的缺点。为了使智慧城市展馆具有更高的科技性、更强的交互性以及更绚烂的界面观赏性, 利用Kinect2.0的骨骼追踪, 以及动作识别的特性和WPF界面编程, 设计出一种可以用手势进行人机交互的城市展馆展览系统。参观者和讲解员可以通过手势与系统进行人机交互达到参观和讲解的目的, 扩展了展览的灵活性, 提高了展馆的科技性, 加深了观众的参观感受。

关键词: 手势识别; 展览系统; Kinect2.0; 人机交互

中图分类号: TP311.1 **文献标识码:** A

Design and Application of Pavilion Interactive System Based on Kinect

MU Li, SHI Zhengkun, CHEN Xiaoyao

(College of Computer Science, Xi'an Polytechnic University, Xi'an 710000, China)

Abstract: At present, the exhibition styles of many city pavilions have the disadvantages of monotony, lack the scientific and technological feeling and lack of interactivity. In order to make the smart city pavilion more scientific, interactive and more user-friendly, this paper uses Kinect2.0's bone tracking and motion recognition features and WPF interface programming to design a pavilion exhibition system that can be used for human-computer interaction through gestures. Visitors and lecturers can achieve the purpose of visit and explanation by interacting with the system through gestures, which expands the flexibility of the exhibition, improves the scientific and technological sense of the exhibition hall, and enhances the visit experience.

Keywords: gesture recognition; exhibition system; kinect2.0; Human-computer interaction

1 引言(Introduction)

许多城市展馆仍旧以传统展览方式为主, 这种展览和讲解方式单调、科技性低、数据更新慢且人机交互困难, 使得参观者的收获少甚至把观看展览当成负担^[1]。当前的界面交互技术多以鼠标点击进行交互, 这种交互方式将人牢牢的束缚在了电脑屏幕前, 移动范围无法超过鼠标的使用范围, 这对于智慧城市展馆的讲解员来说是十分低效的, 严重束缚了讲解员的身体和能动性, 机械教条的进行讲解工作^[2]。AR技术的出现使得人们可以把虚拟世界套在现实世界中并进行互动^[3,4]。Kinect 2.0是AR技术的具体应用, 它可以把人体动作作为交互方式与计算机进行交互。参观者在摄像头面前做出相应的动作, 当Kinect 2.0检测到动作时, 触发相应的响应事件, 实现与展览系统的人机交互。这种交互方式极大地提高了讲解员的自由度和能动性, 比传统的展览方式更加让人舒服。

2 系统构架(System architecture)

2.1 系统架构

基于Kinect2.0的展览系统的结构综合上而下分为三个独立

的层级, 分别是UI层、控制层和数据层。

(1)数据层位于系统最底层, 提供系统正常运行的基本数据。包括要展示的诸如工商数据、教育数据等基本展览数据, 也包括Kinect实时控制数据。

(2)控制层: 控制层的功能是为整个平台的正常运行提供各种运算和控制, 主要包括点击按钮, 并实现要展示的详情页, 触发Kinect手势实现触发按钮点击功能, 以及屏幕切换功能。

(3)UI层: UI层是对系统功能的最终表示, 是直接跟讲解员交互以及供用户观看。

2.2 展示方式

为了方便对展览内容的展示和人机交互, 系统以三分屏的方式进行展示, 即左右两个屏幕是进行交互的系统选项, 中间屏幕是进行系统详细信息的展示, 采用这种方式的原因是, 系统所展示的项目多, 三个屏幕可以让讲解员和参观者对于要展示的内容一目了然, 而不至于出现需要滚动条进行滚动才能看到全部内容的情况, 而且也增加了人员在讲解时

的移动范围，不至于交互过程中过于死板。

2.3 动作识别

系统利用Kinect2.0获取用户的骨骼数据和关节的运动轨迹来识别用户的动作。Kinect2.0的硬件原理是由红外投影仪主动投射近红外光谱，照射到粗糙物体。光谱发生扭曲，会形成随机的反射斑点，进而被深度摄像头读取，深度摄像头分析读到的红外光谱得到红外光谱并生成深度图^[5,6]。也由于此原理，Kinect2.0的骨骼识别与动作获取不受光纤的影响，即使在很暗的地方，依然可以工作。

Kinect2.0的软件编译原理由五部分组成，分别是sensor、source、reader、frame、data^[7]。

经过硬件和软件的共同作用可以生成人体的骨骼图，kinect2.0一共可以绘制人体的25个关节，通过25个关节的位置和角度的变化识别人体所作的动作，并触发系统相应的响应事件达到人机交互的目的。

3 系统设计与实现(System design and implementation)

3.1 系统功能概述及目标

Kinect2.0作为在展馆中交互的工具，需提前定义好合法的动作，以体感人机交互为主，达到在展馆中沉浸式参观的目的。如果在参观过程中，参观者只能看一些单调的图片、文字和演示文档或是视频，会降低参观的体验。Kinect2.0的骨骼追踪和动作识别功能有效地克服了以上弊端，通过定义相关的手势和动作来对系统进行控制，把用鼠标和键盘完成的动作全部用手势来完成。参观者和讲解员只需用一些日常生活中的手势就可以完成这些操作，其活动空间和便捷度明显增加。系统定义三个屏幕，左右两个屏幕放置控制面板，用户通过手往前伸的操作来完成对按钮的点击，就像现实世界中，对按钮的按压一样，将这个动作定义位“按压”操作。反之，与用鼠标一样，点击按钮之后需要将鼠标抬起来一样，做完按压动作之后，需要把手回到原来的状态，这个动作是手往后缩，就像现实世界中，按压按钮之后抬起手一样，定义这个动作为“抬起”。同时，因为系统由三个屏幕，左右两边分别为控制区域，这就需要使用者在使用途中进行控制区域的切换，需要将手的位置移出现在所在屏幕位置，并将手一如待切换屏的位置，就好像现实世界中拿东西，需要将手移动到待拿的物体身上，定义这个操作为“移动操作”。考虑到参观人数众多，设定手超过头顶为获得交互的动作，就像现实中点名一样，点到名字的人举手，把这个动作定义为“举手”动作。当某操作者放弃交互，需要将手放置下垂状态，就像现实生活中，放下某物品后手恢复到自然状态，把这个动作定义为“放下”。系统核心的部分就是对这五个动作的准确识别，达到一个比较高的准确率。

3.2 动作识别算法

Kinect2.0在进行骨骼识别的过程中由三个坐标系和两个映射，三个坐标系分别是现实空间坐标、相机空间坐标和屏幕坐标；两个映射是现实空间到相机空间的映射，相机空间到屏幕空间的映射。第一个映射将空间世界转换为以相机为中心原点的坐标，并建立以相机为中心原点的坐标系，其坐标为相对于相机中心的(X,Y,Z)坐标值^[6]。第二个映射将相机坐标转换为只有(X,Y)的屏幕坐标并在屏幕上输出。

人在做动作时是一个三维空间的运动，包括角度和距

离，距离可用相机坐标的位置进行减法运算得到，角度是三维旋转，但是也可看作是在某二维平面的旋转，即三维空间在二维坐标系的投影，来计算关节在横向、纵向，以及前后的投影。也由此假定在t时刻的手的坐标是(x₁,y₁,z₁)则转化公式如式(1)所示^[8]。

$$\begin{cases} x_w = (x_i - \frac{w}{2})(z_w + D)F(\frac{w}{h}) \\ y_w = (y_i - \frac{w}{2})(z_w + D)F \\ z_w = K \tan(Hz_i + L) - O \end{cases} \quad (1)$$

其中，H=3.5*10-4rad，K=12.36cm，L=1.18rad，O=3.7cm，D=-10，F=0.0021。经过坐标公式的转化得到所需关节的坐标。

利用关节的坐标计算各关节移动后的角度，则关节的移动角度的向量a(a₁,a₂,a₃)与b(b₁,b₂,b₃)之间的角度计算^[6,9]公式如式(2)所示。

$$\theta_{i-j} = \frac{v_{i1} \times v_{j1} + v_{i2} \times v_{j2} + v_{i3} \times v_{j3}}{\sqrt{v_{i1}^2 + v_{i2}^2 + v_{i3}^2} \times \sqrt{v_{j1}^2 + v_{j2}^2 + v_{j3}^2}} \quad (2)$$

将坐标的距离值和角度值与标准数据进行对比，做相似检测，如果相似值在容错范围内，则表示动作发生，动作捕捉系统捕捉到动作发生后触发相应的响应事件，并将结果显示在相应的界面上。

3.3 窗口界面

窗口界面总体分为三个，每一个窗口占用一个屏幕，左屏幕和右屏幕分别展示选项按钮，按钮风格是由WPF的样式风格和模板风格自定义。中间屏幕是对应按钮的详情展示页，当用户选中要展示的内容时，中间屏幕会将选中内容的详细信息显示出来，供讲解员讲解和用户观赏。三个屏幕布局的基本样式如图1所示。



图1 屏幕布局

Fig.1 Screen layout

4 关键技术及解决方案(Key technologies and solutions)

系统需要支撑大量的来自Kinect 2.0的数据进行采集计算和应用，这些数据的应用和判断方式多种多样，为了保证系统的一致性、未来的扩展性和易维护性，通过关键技术解决方案，最大的提高动作识别的准确度和系统的流畅性，保证每一个过程数据的一致性。

4.1 屏幕范围设定

Kinect的工作原理是发射红外线，并探测红外光反射从而计算每一个关节的位置，根据位置的变化来捕捉人体的动作，但是在系统运行的大环境下，每个场景所运用的显示器的大小和分辨率不同，而且人体在移动时并不像操控鼠标一样有很大的限制，这就要求在不同环境和不同人使用的过程中，都应在人的手移动很短的距离内使指针像鼠标一样在人手很小的移动范围内，指针可以遍布整个屏幕，而且要满足人体工程学，以人手臂的半径为移动极限的范围内可以使指针移动整个屏幕，要想达到这个效果，要将人体肩关节的旋转进行欧拉角的计算，同时将手每移动一厘米的距离与屏

幕指针移动的距离相适应，其具体方法是通过Kinect2.0来计算人肩膀到手的距离和两肩膀的距离，从而得到人手的最大移动距离，以右手为例，右手向左移动最大距离是两肩距离，向右移动距离是手臂的长度，再通过获得屏幕分辨率从而得到手每移动一厘米的距离指针所移动的像素单位，从而覆盖整个屏幕区域。

4.2 屏幕切换设计

在三分屏中，因为有两个控制屏幕，所以需要进行控制窗口的切换，以达到选择不同功能的目的。Kinect 2.0是面向用户的，与三个屏幕是在同一屏幕内的，因此无法通过图像识别的方法来进行屏幕的切换，系统中使用的方法是位置标定的方法，虽然Kinect 2.0无法获取屏幕的位置，但是用户是清楚屏幕的位置以及是否要进行切换的。系统中对手进行位置判定角度判定以及肘部的角度判定，当手的位置已经离开左屏幕的区域，当手的X值达到系统设定的阈值时，同时肘的角度变化也达到阈值时，即进行控制界面的切换，为了保证系统进行屏幕切换的稳定性，还应进行肩部欧拉角的旋转角度，在实验中，肘部角度的阈值由勾股定理进行计算，一般值在120度认为是右切换，30度认为是从右往左切换。

4.3 动作识别判定

(1) 按压动作的判定

按压动作是用户选中某一选项，想查看该选项具体内容和动作，该动作的判定由手的位置的Z值的变化和肘部关节的角度变化共同约束和决定，该动作在充分考虑人体工程基础上进行设计，一般认为人体小臂的向前舒展在10cm时最感觉舒适，因此在进行位置移动的判定时也以此为基础，即当RightHand/LeftHand.Postion.Z>10时进行按压动作的判定。同时，为了提高准确度，进行肘关节的勾股定理的判定，当肘关节的变化值达到30度时认为按压操作的完成，而为了避免手在按压操作时的抖动，设置前后左右的容差值距离在5以内时，认为时对该按钮进行按压，当系统检测到按压状态之后触发相应的事件，详情页上显示该按钮的详细信息。

(2) 抬起动作的判定

抬起动作是用户按压完成后人本身进行回收的动作，是人体的一个习惯性动作，回收动作的设计同样考虑人体工程学进行设计，虽然人在按压动作时的距离是10cm但是在手臂回收的过程中总是习惯的回收至自己最舒适的位置，因为标准的不确定性，但最少也有5cm的移动，为了使用户的人机交互能够得到最佳的体验状态，因此抬起的识别应该比较宽松，将RightHand/LeftHand.Postion.Z>3时同时肘部的角度发生变化为该动作的判定依据，当满足条件时就认为抬起动作已经发生，从而系统结束按钮的响应事件。

(3) 举手动作的判定

举手动作是用户获得人机交互的动作，其判定依据是手的高度大于头的高度，即RightHand/LeftHand.Postion.Z>Head.Potion.Y时发生^[6]，在考虑人体工程学，举手的高度不应太高，一般认为整个手的底端超多头的高度最为适宜，所以阈值设置为15cm。同时，认为该动作应该是动态的而不是静态的，即为了防止有人挠头等行为的误判，将高度值的变化检测设置逐步递增以及肘部关节角度同时发生变化，当满足全部条件时，系统检测进行人机交互，出现交互

指针。

(4) 放下动作的判定

放下动作是用户放弃人机交互的动作，在考虑人体工程学的基础上，认为用户放弃人机交互即进入人站立的基本状态，即双臂自然下垂，其判定依据是屏幕指针首先是进入窗口的底端，其次是肘部的角度大约是180度而且手的Y值为肩关节的Y值加上手臂的长度。当条件满足时，系统停止人机交互并为下一个用户的交互做准备。

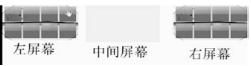
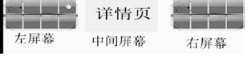
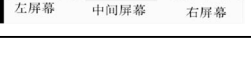
5 系统测试及评估(System testing and evaluation)

5.1 动作识别测试

为了评估系统的性能和动作识别的准确度，由对六个人对四组动作分别进行测试。在为了使环境与展馆环境相似，在一人进行测试时，其他人进行动作的干扰，将测试结果尽可能地接近展馆的真实环境，为了使动作更具有一般性和全面性，要求每个测试人员从开始至结束进行动作30次。测试内容从分别进行举手动作、指针全屏幕覆盖、按压动作、抬起动作、切换动作、放下动作的随机动作的测试，以检测系统对动作识别的准确度和灵敏度。其具体动作的表示和效果如表1所示。

表1 动作效果

Tab.1 Action effect

动作名称	动作效果		
举手			
按压			
抬起			
切换			
放下			

5.2 测试结果分析

在实验中，测试人员的主要工作是对动作识别灵敏度和准确度进行检测，测试要求测试人员想要触发的事件和系统实际触发的事件是否一致，动作测试结果如表2所示。

表2 动作测试结果

Tab.2 Action test result

序号	测试动作(准确次数/总数)				
	举手	按压	抬起	切换	放下
1	29/30	30/30	30/30	29/30	30/30
2	30/30	30/30	30/30	28/30	30/30
3	29/30	30/30	29/30	28/30	30/30
4	28/30	29/30	29/30	30/30	30/30
5	30/30	29/30	29/30	30/30	30/30
6	30/30	29/30	30/30	29/30	30/30
准确率	97%	95%	95%	96%	100%

5.3 系统评估

实验结果表明,动作识别率,除切换动作外都在95%以上,针对测试结果,对切换动作进行调整,加大两个屏幕切换的边界位置阈值,将切换动作的识别率提高到了95%,而且实验结果证明,即使在有人故意干扰的情况下系统仍可以正确运行,证明了系统有一定的抗干扰性,没有明显波动和跳帧。从而实现了系统的设计目标。

6 结论(Conclusion)

本文描述了利用Kinect进行动作识别,并应用于展馆系统的人机交互中,最后实验数据充分验证了该方法的有效性。方便了观众的参观和讲解员的讲解,提高了展馆的科技性,增加了展览的趣味性,而且建设成本也比较低,进一步扩大了展馆进行展览的方式。然而,本系统也存在缺陷,如参观者不能随自己的喜好随意进行动作的编排和设计,这也是下一步工作的目标之一。

参考文献(References)

- [1] F Febriansyah, N A suwastika, H Fakhurroja. Patient necessity notification system based on gesture recognition (Kinect V2) and internet of things using selection frame method[J]. Journal of Physics: Conference Series, 2019, 1192(1): 12051.
- [2] Ohashi Isamu. Demonstration Experiment of AR Exhibition System Inducing User Behavior with Visual Incompatibility[C]. Lecture Notes in Computer Science: Springer International Publishing, 2019: 174-186.

(上接第18页)

特点,大幅削减了数据采集和分析决策时间,数据分析平均仅需10—20min即可完成,基于多因素分析的综合评价指数更准确、全方位地反映了地区广告投放带来的损益等效果,减轻了员工与BI流量部门数据汇总操作人员的工作负担。

系统部署和实行多因素指标优化方案一个月内,每日订单数量平均增加167单,估计相较于优化前每日挽回流失客户价值达83.5万元。案例所实施的优化方案成本低廉,实施效果显著,能够较好地简化该公司业务推广方案决策流程和提升决策质量。

参考文献(References)

- [1] Sung Wook Ji, Young-jun Choi, Min Ho Ryu. The economic effects of domestic search engines on the development of the online advertising market[J]. Telecommunications Policy, 2016, 40(10-11): 982-995.
- [2] Amar Budhiraja, Akhil Ralla, P. Krishna Reddy. Coverage pattern based framework to improve search engine advertising[J]. International Journal of Data Science and Analytics, 2019, 8(2): 199-211.
- [3] David M. HTML5: Designing Rich Internet Applications[M].

- [3] Ahmed Mohammed-Hussein, Tahir sabir Azhin, Maghdid Halgur-Sarhang. Kinect-based Human Gait Recognition using Triangular Gird[C]. 1st International Conference on Advanced Research in Engineering Sciences, Institute of Electrical and Electronics Engineers Inc., 2018: 8723293.
- [4] 吴鹏. 基于Kinect2_0手势控制在多媒体教学中的应用[J]. 软件导刊, 2017, 16(4): 84-85.
- [5] 薛智宏. 基于Kinect的原地太极拳辅助训练系统[J]. 河北科技大学学报, 2017, 38(2): 183-188.
- [6] 李红波. 基于Kinect骨骼数据的静态三维手势识别[J]. 计算机应用与软件, 2015, 32(9): 161-165.
- [7] Toby Sharp. Accurate, Robust, and Flexible Real-time Hand Tracking: Supplementary Material[DB/OL]. <http://www.k4w.cn/news/1#>.
- [8] 王辉柏. 基于Kinect的射击游戏动作的设计方法研究[J]. 计算机仿真, 2018, 35(12): 388-391.
- [9] 吴晓雨. 基于Kinect的手势识别算法研究及应用[J]. 计算机应用与软件, 2015, 32(7): 173-176.

作者简介:

牟莉(1972-), 女, 硕士, 副教授. 研究领域: 智能化信息系统, 嵌入式系统应用。

史政坤(1994-), 男, 硕士生. 研究领域: 智能化信息系统。

陈逍遥(1994-), 男, 硕士生. 研究领域: 图像识别。

HTML5 Designing Rich Internet Applications, 2010: 9-98.

- [4] 黄德, 吕婷. 基于AHP桂台农业合作影响因素分析[J]. 现代商贸工业, 2019, 40(31): 17-18.
- [5] 张凤荣. 工业工程在企业信息化中的应用[J]. 网络与信息, 2012, 26(08): 15.
- [6] 罗英涛. 信息技术在工业工程中的应用分析[J]. 科学技术创新, 2019(22): 82-83.
- [7] 陈晗. 百度竞价推广业务营销策略研究[D]. 山东大学, 2018: 12-22.
- [8] 李晓薇. Web前端开发技术以及优化方向[J]. 电子技术与软件工程, 2019(19): 48-49.
- [9] 国家统计局. 中国统计年鉴-2018[Z]. 中国统计出版社, 2018: 29-106.

作者简介:

林之博(1997-), 男, 硕士生. 研究领域: 信息系统工程。

刘媛华(1974-), 女, 博士, 副教授. 研究领域: 系统工程, 经济与管理统计, 复杂系统理论方法及应用, 社会经济复杂系统。