

基于 Kinect 的手术辅助系统研究

陈燕军,汪 地,杨 浩,卢泉奇,苗同升

(上海大学 上海市机械自动化及机器人重点实验室,上海 200072)

摘 要:针对手术室无菌环境要求非常高的现实需求,以及传统鼠标键盘等人机交互方式存在的需要定期消毒且无法在手术过程中进行操作等客观限制,对目前手术室内人机交互存在的问题进行了分析,提出了一套基于 Kinect 自然人机交互技术的手术辅助系统。该系统包括一个病人资料库、一套 Windows 系统下基于 Kinect 自然人机交互技术的控制程序及一套 Kinect 手术辅助系统硬件。实验结果表明通过 Kinect 可以收集操作者的手势、动作、声音等,并将其转化为操作系统可以理解的命令,对手术辅助系统进行相关操作,从而辅助手术进行。该系统对于改善手术条件、节省手术时间、缓解病患痛苦具有一定帮助。

关键词:人机交互;Kinect;手术辅助系统

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2014)09-0081-03

doi:10.3969/j.issn.1673-629X.2014.09.018

Research on Operation Auxiliary System Based on Kinect

CHEN Yan-jun, WANG Di, YANG Hao, LU Quan-qi, MIAO Tong-sheng

(Shanghai Key Laboratory of Mechanical Automation and Robotics, Shanghai University,
Shanghai 200072, China)

Abstract: Aiming at the high demand for sterile operating room environment and the objective constraints of regular disinfection and not operating during the operation for traditional interactive ways such as mouse and keyboard, analyze the human-computer interaction problems in the operating room, and put forward a set of operation auxiliary system based on natural Kinect human computer interaction technology. The system includes a patient database, a set of control procedure based on natural Kinect human-computer interaction technology under Windows system and a set of Kinect operation auxiliary system hardware. The experimental results show that the Kinect can collect the operator gesture, movement, sound and so on, that change to the command which the operating system could understand, carrying out corresponding operation for the system to support the operation. The system is beneficial for improving the operation conditions, saving the operation time and alleviating patients pain.

Key words: human-computer interaction; Kinect; operation auxiliary system

0 引言

自然人机交互技术是人机交互技术的一个重要研究方向,主要研究人和计算机之间如何用自然的方式进行交互^[1]。2008年,比尔盖茨提出“自然用户界面(Nature User Interface, NUI)”的概念,人机交互在近几年发生了很大的改观,鼠标和键盘将逐渐地被更为自然的触摸式、视觉式及语音控制界面所代替。

近年来,计算机断层扫描技术(CT)、核磁共振(MRT)等非侵入性医学成像技术大量应用于医疗诊断中,这些技术的应用可以有力地辅助医生精确诊

断^[2]。然而,大量的影像资料也加大了医生在手术中的工作量,医生需要经常查阅影像资料,这样就束缚了医生的手术操作。

目前,尽管计算机辅助系统被大量地应用于手术室中,但是其主要是基于鼠标、键盘等人机交互技术。由于手术室中对无菌要求非常高,传统的鼠标键盘等人机交互方式存在需要定期消毒,且医生无法在手术过程中进行直接操作等限制。其需要护士为医生查找相关的影像资料,这样既增加了不必要的交流,也耗费掉了宝贵的手术时间。为了使医护人员从繁琐的重复

收稿日期:2013-11-20

修回日期:2014-02-21

网络出版时间:2014-07-17

基金项目:上海市机械自动化及机器人重点实验室开放基金项目(ZZ1001)

作者简介:陈燕军(1989-),男,江西丰城人,硕士,研究方向为嵌入式系统和NAO机器人;汪 地,博士,副教授,研究生导师,研究方向为嵌入式系统开发、机器人控制。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20140717.1226.007.html>

劳动中解脱出来,文中提出了一种基于 Kinect 的自然人机交互技术应用于辅助手术当中。Kinect 体感操作是一种非接触式的操作,其利用人体的手势、动作、声音就可以实现在手术进行过程中调阅患者的病灶影像,通过手势放大、缩小图,翻阅病例等操作。

1 Kinect 人机交互技术

Kinect for Windows 是在 Xbox 360 的基础上进行了改进而来的,其固件做了升级,支持“近景模式”、提升了骨骼跟踪等 API 性能、更好地兼容各种 Windows 计算机、采用更为先进的声学模型,并且是微软授权在 Windows 平台下进行 Kinect 开发的传感器^[1-3]。

图 1 所示为 Kinect 的关键部件,主要包括红外投影机、RGB 摄像头、麦克风阵列、仰角控制马达。红外投影机主要是通过发射和接收红外光谱来创建可视范围内的人体、物体的深度图像;RGB 摄像头用于拍摄视角范围内的彩色视频图像;麦克风阵列包括 4 个麦克风,可以通过 4 个麦克风采集声音,同时过滤背景噪音,定位声源;仰角控制马达可以通过编程来控制 Kinect 的仰角,从而获得最佳的视角。Kinect 成像系统自身的视角大小为垂直 43°,水平视角为 57°;仰角驱动马达可以驱动 Kinect 在±28°垂直方向上调整。



图 1 Kinect 的关键部件

Kinect 摄像头最高分辨率为 640×480,并且无论在任何照明条件下,都可以利用 Kinect 的 CMOS 红外探测器来感知空间。Kinect 通过发射黑色和白色的光谱来感知环境,黑色代表无穷远,白色代表无穷近。介于黑色和白色的灰度区域代表了 Kinect 对于物体的物理距离的探测结果。它可以收集到视野里的每一个点的信息,并将其生成一张深度图像用以代表环境信息。Kinect 的 COMS 红外探测器的探测频率为 30 帧/s,可以实时地探测环境中的深度。而 Kinect 中间的彩色摄像头可以用于获取人体的二维 RGB 彩色图像。

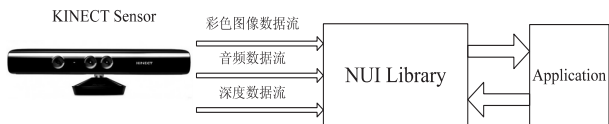


图 2 Kinect 数据传输架构

图 2 显示的是 Kinect 数据传输架构示意图。Kinect 传感器提供原始数据流给 SDK,这些数据包括彩

色图像数据流、音频数据流、深度数据流,通过 Kinect SDK 可以直接获取这些原始传感器数据流,并在此基础上进行底层应用开发。

2 系统总体方案设计

图 3 所示为基于 Kinect 人机交互技术的手术辅助系统结构示意图。其主要由系统标定、登录验证、系统自检、手势控制模块、语言控制模块组成。其中手势控制模块包括悬停选择,空气鼠标,划动,放大、缩小,垂直摆动,旋转等手势的识别。

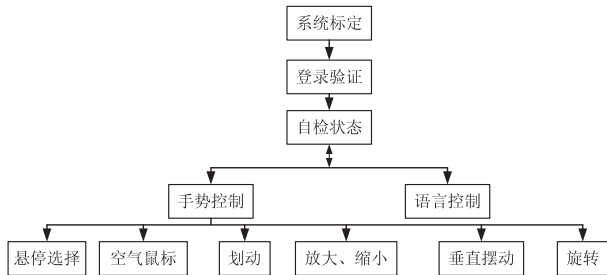


图 3 系统总体方案示意图

为了保证系统的安全性及可靠性,在这里引入人脸识别登录验证技术。Kinect 中提供了人脸识别的 SDK,应用程序可以获得实时的 3D 网格化面部特征,根据网格化的多边形信息,能实时追踪双眼的间距、眉毛的位置以及嘴的形状等。

悬停选择的功能主要是实现系统功能导航,通过将手置于手势辅助系统的界面上停留 1.5 s 左右,系统将进入功能导航进程。

空气鼠标的功能主要是通过非接触的方法来实现对系统界面的操作功能,用户通过挥手激活空气鼠标。空气鼠标激活后可以跟随用户的手移动,进入悬停选择之后鼠标将自动隐藏,系统进入手势控制、声音控制状态,这样可以避免鼠标在界面上乱动的现象。

划动操作用来实现翻阅医学影像资料的功能。用户用手在 Kinect 视角范围内自右向左水平划动可以实现医学影像在屏幕上切换;反之,当用手在视角范围内自左向右划动则可以控制医学影像的反向切换。

放大、缩小能够实现影像资料的放大和缩小。在手术室环境下,当医生选择好了一张医学影像之后,可以通过双手手掌靠近和伸展来实现缩小和放大的功能,从而能够准确定位病灶部位。

垂直摆动功能可以实现翻阅病例的功能。在大多数情况下,病例的内容会超过一页,这样就会超出屏幕的范围,为了翻阅病例,系统通过设计垂直摆动手臂来实现翻页。当用户抬起或垂下手臂时病例将在屏幕垂直方向上滚动。

语言控制可以大大减少操作的时间。用户可以通过语言识别迅速搜索信息,这样在某些场景下可以大

大提升自然交互的体验。

3 关键技术分析

3.1 骨骼跟踪技术

骨骼跟踪技术是通过处理深度数据来创建人体各个关节的坐标,骨骼跟踪能够确定人手、头部及身体的位置^[4-5]。在 Kinect 当中,是通过 20 个关节点的位置来表示人体的骨架,如图 4 所示。当用户进入 Kinect 视野范围内就可以找到用户的 20 个关节点并通过 (x,y,z) 坐标来表示。

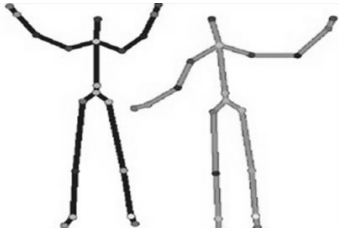


图4 人体关节示意图

NUI 骨骼跟踪分为主动和被动二种模式,最多能提供二套完整的骨骼跟踪数据^[6]。骨骼数据采用基于事件的方式通过 SkeletonStream 传输,在初始化并打开骨骼跟踪后,就可以从 SkeletonStream 获得骨骼数据, SkeletonStream 产生的每一帧数据 SkeletonFrame 都是一个骨骼对象集合。它包含一个骨架数据结构的数组,其中每一个元素代表着一个被骨架追踪系统所识别的骨架信息。每一个骨架信息包含有描述骨骼位置以及骨骼关节的数据。每一个关节有一个唯一的标识符,如头(head)、肩(shoulder)、肘(dllbow)等信息和对应的三维坐标数据。

3.2 语音识别

语音识别包括对待命令的识别和对自由形式的语音识别二类,自由形式的语音识别需要训练软件来识别特定的声音以提高识别精度,而命令识别则是利用了不同的策略来提高识别精度的^[7-8]。Kinect 采用的是通过 Microsoft. Speech 类库来识别命令语音的。

Kinect 中音频前置处理主要是通过 KinectAudioSource 这个对象来完成的^[9]。首先通过 KinectAudioSource 配置 Kinect 麦克风阵列的工作模式,然后通过 KinectAudioSource 类从麦克风阵列中提取原始音频流。音频流经过一系列的算法来处理降噪、自动增益控制和回声消除,从而提高音频质量。

经过 KinectAudioSource 处理的音频数据需要通过 Microsoft. Speech 类库来识别语音。Microsoft. Speech 类库的语音识别功能是通过 SpeechRecognitionEngine 对象来实现的,SpeechRecognitionEngine 从 Kinect 传感器中获取经过处理的音频数据流,通过分析和解译,从而匹配出最合适的语音命令。

4 系统具体实现

图 5 所示为手术辅助系统的应用界面,它包括病例档案、医学影像和化验结果三个功能键,骨骼跟踪视图;中间区域是病例信息显示区域,可以通过界面左侧的功能键实现病例档案、医学影像和化验结果之间的切换。

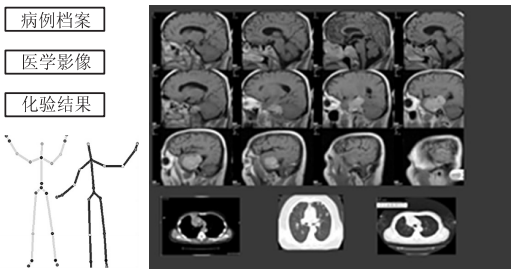


图5 手术辅助系统的应用界面

手术辅助系统是通过 C++语言来实现跨平台库的接口扩展功能,从而实现影像处理和人机交互。在 VS2010 环境下结合 OpenCV 和 OpenNI 来实现手势识别算法。OpenCV 是一个开源的跨平台计算机视觉库,可以实现图像处理 and 计算机视觉方面的很多通用算法。OpenNI(开放的自然交互)是一个多语言、跨平台的架构。OpenNI 主要目的是要形成一个标准的 API,其由一组可用来编写通用自然交互应用的接口组成,可以实现视觉和音频传感器的通信功能;NITE 是 OpenNI 的一个中间件,可以调用它来实现人手运动的跟踪和手势识别^[10-12]。

如图 6 所示,系统需要调用 DCMTK、VTK、ITK、wxWidgets 库函数来实现。DCMTK 是由德国 Offis 公司开发的,其提供了实现 DICOM 协议的一个平台,使得编程人员可以轻松实现 DICOM 标准的医学图像的处理^[13-14]。VTK 是一个开源的视觉化工具函数库,其可以实现跨平台的图像处理。ITK 是一款用 C++实现的开源的、跨平台的医学图像处理软件包。wxWidgets 是一个开源的跨平台的 C++架构库,它可以提供 GUI 和其他工具。

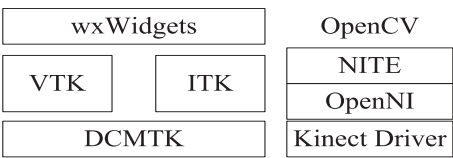


图6 系统调用函数库示意图

5 结束语

Kinect 作为一种新颖的输入方式,其可以通过手势、身体姿态及语音作为输入方式来实现控制功能。文中基于这一思路提出了一套基于 Kinect 自然交互技术的手术辅助系统,其可以在不适合徒手操作的手术

4 结束语

话题检测和跟踪是网络舆情系统最重要的应用之一,几乎每个舆情系统都会应用话题检测和跟踪分析网络舆情热点和突发事件。由于互联网每天发布的信息量大,如何快速准确地发现热点话题和突发性事件是近几年来舆情应用研究的热点。大多数研究者的共识就是应用统计模型,在自动聚类的基础上进行话题检测,将计算正确率的提高归结于聚类算法的改进上,很少关注聚类时的策略和方法的改进。但实际上策略和方法上的改进在一定应用范围内比算法改进更有效。文中就是运用语法分析抽取文中所有句子的主干,并根据主干提取 VSM 向量。基于双向量的话题检测大大提高了检测的正确率,在网络舆情热点监测方面得到了较好的验证和应用。下一步的工作将继续深入研究特征向量的提取方法和中心向量提炼策略,在语义分析的基础上进一步提高话题检测的正确率,从而得到更好的应用效果。

参考文献:

- [1] 张小明,李舟军,巢文涵. 基于增量型聚类的自动话题检测研究[J]. 软件学报,2012,23(6):1578-1587.
- [2] 王振宇,吴泽衡,唐远华. 基于多向量和二次聚类的话题检测[J]. 计算机工程与设计,2012,33(8):3214-3218.
- [3] 米海涛,熊德意,刘 群. 中文词法分析与句法分析融合策

(上接第 83 页)

室环境下协助医生进行手术以提高手术效率,减轻医护人员劳动强度。医生可以在手术过程中随时使用自己的手势、身体姿态及语音作为输入元素来实现病例资料、医学影像和化验报告的切换;同时可以实现医学影像的放大缩小、翻页、旋转等功能。

参考文献:

- [1] 余 涛. Kinect 应用开发实战[M]. 北京:机械工业出版社,2012.
- [2] 董士海. 人机交互的进展及面临的挑战[J]. 计算机辅助设计与图形学学报,2004,16(1):1-13.
- [3] Hernandez-Lopez J J, Quintanilla-Olvera A L, Lopez-Ramirez J L, et al. Detecting objects using color and depth segmentation with Kinect sensor[J]. Procedia Technology, 2012,3:196-204.
- [4] 李 斌,吴国斌. Kinect 引领人机交互变革[J]. 程序员,2011(9):100-103.
- [5] 邓 瑞,周玲玲,应忍冬. 基于 Kinect 深度信息的手势提取与识别研究[J]. 计算机应用研究,2013,30(4):1263-1265.
- [6] 何 贝,王贵锦,林行刚. 结合 Kinect 深度图的快速视频抠图算法[J]. 清华大学学报:自然科学版,2012,52(4):561

略研究[J]. 中文信息学报,2008,22(2):10-17.

- [4] 马 彬,洪 宇,陆剑江,等. 基于线索树双层聚类的微博话题检测[J]. 中文信息学报,2012,26(6):121-128.
- [5] 洪 宇. 基于语义结构和时序特征的话题检测与跟踪技术研究[D]. 哈尔滨:哈尔滨工业大学,2009.
- [6] 常 鹏,马 辉. 高效的短文本主题词抽取方法[J]. 计算机工程与应用,2011,47(20):126-128.
- [7] Hofmann T. Probabilistic latent semantic analysis[C]//Proc of UAI99. [s. l.]:[s. n.],1999.
- [8] Li Hang, Yamanishi K. Topic analysis using a finite mixture model[J]. Information Processing and Management,2003,39(4):521-541.
- [9] Landauer T, Foltz P, Laham D. Introduction to latent semantic analysis[J]. Discourse Processes,1998,25:259-284.
- [10] 唐籍涛,李 飞,郭昌松. 网络舆情监控中新词识别问题的研究[J]. 计算机技术与发展,2012,22(1):119-121.
- [11] 林鸿飞,高 天,姚天顺. 中文文本的可视化表示[J]. 东北大学学报:自然科学版,2000,21(5):501-504.
- [12] 金 珠,林鸿飞,赵 晶. 基于 HowNet 的话题跟踪及倾向性分类研究[J]. 情报学报,2005,24(5):555-561.
- [13] 赵 华,赵铁军,张 姝,等. 基于内容分析的话题检测研究[J]. 哈尔滨工业大学学报,2006,38(10):1740-1743.
- [14] Hotho A, Stumme G. Conceptual clustering of text clusters[C]//Proceedings of FGML workshop. Piscataway, NJ, USA: IEEE,2002:1-9.
- [1] 吕开阳,叶华茂,李晓光,等. Kinect 体感技术在动物外科实验教学中的应用及展望[J]. 中国医学教育技术,2012,26(2):171-173.
- [2] Kuehn T. The Kinect sensor platform[J]. Advance in Media Technology,2011(6):2192-2198.
- [3] Gallo L, Minutolo A, de Pietro G. A user interface for VR-ready 3D medical imaging by off-the-shelf input devices[J]. Computers in Biology and Medicine,2010,40(3):350-358.
- [4] 陈一民,张云华. 基于手势识别的机器人人机交互技术研究[J]. 机器人,2009,31(4):351-356.
- [5] 郝颖明,朱 枫. 外科手术计算机辅助导航技术[J]. 生物医学工程杂志,2004,21(2):306-310.
- [6] Chang Y J, Chen S F, Huang J D. A Kinect-based system for physical rehabilitation: a pilot study for young adults with motor disabilities[J]. Res Dev Disabil, 2011, 32(6): 2566-2570.
- [7] 罗 元,谢 彧,张 毅. 基于 Kinect 传感器的智能轮椅手势控制系统的设计与实现[J]. 机器人,2012,34(1):110-113.
- [8] Michael N, Fink J, Kumar V. Cooperative manipulation and transportation with aerial robots[J]. Autonomous Robots, 2011,30(1):73-86.

基于Kinect的手术辅助系统研究

作者：[陈燕军](#)，[汪地](#)，[杨浩](#)，[卢泉奇](#)，[苗同升](#)，[CHEN Yan-jun](#)，[WANG Di](#)，[YANG Hao](#)，
[LU Quan-qi](#)，[MIAO Tong-sheng](#)
作者单位：[上海大学 上海市机械自动化及机器人重点实验室, 上海, 200072](#)
刊名：[计算机技术与发展](#)[ISTIC](#)
英文刊名：[Computer Technology and Development](#)
年，卷(期)：2014 (9)

本文链接：http://d.wanfangdata.com.cn/Periodical_wjfz201409018.aspx