

基于 Kinect 的摔倒行为研究

杨张振

(南京邮电大学 通信与信息工程学院,江苏 南京 210003)

摘要:针对传统视频检测技术识别效率低和实时性差的问题,提出基于 Kinect 体感设备对人体摔倒行为进行判断识别。其中人体位于 Kinect 的检测范围之内,通过对 Kinect 设备获取到的深度图像进行处理,得到人体骨骼图像及人体关节点的位置信息;利用 Kinect 骨骼追踪技术,参考人体左肩、右肩 2 个骨骼点,以两肩中心点为目标,实时计算两肩中心关节点的空间位置、相对位置等参数,计算出不同帧之间两肩中心点位置之间的位移变化,并结合该位移向量与 $O-X,Y,Z$ 坐标系中的 Y 轴方向夹角,以二者相结合为条件来判断人体是否出现摔倒事件。经过实验验证,在室内环境中,该方法能够实现人体摔倒的自动实时检测,并且利用深度信息和骨骼信息对摔倒行为进行检测判断,能够有效地保护监测环境内当事人的个人隐私。

关键词:Kinect;深度图像;骨骼图像;摔倒;向量

中图分类号:TP39

文献标识码:A

文章编号:1673-629X(2018)04-0179-04

doi:10.3969/j.issn.1673-629X.2018.04.038

Research on Fall Detection Based on Kinect

YANG Zhang-zhen

(School of Telecommunications & Information Engineering, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China)

Abstract: Aiming at the problem of low efficiency and poor real-time performance of traditional video detection technology, we propose to judge and identify the human body's fall based on Kinect somatosensory equipment. In this paper, the human body is located within the detection range of Kinect, and the location information of human skeleton images and human joint can be acquired by the processing of depth images from the Kinect. Referred to the two bone points on the left and right shoulder, with the center of them as the object, we calculate the parameters like spatial position, relative position in the center node of two shoulders in real-time by Kinect bone tracking technology and displacement variation between the center points of the two shoulders out of different frames. Combined the angle between the displacement vector and the axis Y in the coordinate system of $O-X,Y,Z$ and the displacement variation as the conditions, it is necessary to judge whether the human body has a fall event. Experiments show that in the indoor environment, the proposed method can realize the automatic real-time detection of the human body's fall, and uses the depth information and the skeletal information to judge the fall behavior, which can effectively protect the personal privacy of the parties in the monitoring environment.

Key words: Kinect; depth image; skeleton image; fall; vector

0 引言

随着人类社会的发展和进步,人们的物质生活水平得到了大幅提高,同时人们对身体健康水平和生活环境的安全稳定水平的要求也在不断增长。出于对健康和安全的考虑,越来越多的监控设备被应用于室内环境以检测室内可能出现的意外摔倒情况^[1]。

摔倒行为可能会对人们的健康带来严重的后果,尤其是出现这一意外情况但没有及时发现处理,将会

加重事件带来的不利影响,比如老年人出现摔倒情况却没有得到及时救助,后果可能不堪设想。据全国伤害监测系统(NISS)2013年的收集统计,65岁及以上老年人跌倒死亡率为45.72/10万,因跌倒死亡是65岁及以上人群因伤害致死的第一死因^[2]。除导致死亡外,跌倒/坠落造成更多的是残疾、功能受限、活动受限等非致死性后果。因此对室内可能出现的摔倒行为进行检测是一项很有实际意义的工作。

收稿日期:2017-03-28

修回日期:2017-07-31

网络出版时间:2017-12-05

基金项目:国家自然科学基金(61172118)

作者简介:杨张振(1992-),男,硕士研究生,研究方向为图像处理与计算机视觉。

网络出版地址: <http://www.cnki.net/kcms/detail/61.1450.TP.20171205.0903.002.html>

传统的对摔倒行为进行检测的技术手段主要有两种:一种是利用穿戴式传感器对摔倒进行检测,当检测到摔倒行为时,与人体连接在一起的传感设备会发出摔倒信号^[3-5];另一种是依据视频装置对摔倒进行检测,用一个或多个摄像机对某个场景进行连续监控,通过对视频进行处理对摔倒进行检测^[6]。传统方法对摔倒行为的检测虽然取得了一定的成果,但是局限性也很大。利用传感器对摔倒行为进行检测会对人体行动产生不便,利用传统视频监控设备对摔倒行为检测不仅难以做到及时的检测识别,又要面临外界条件变化所带来的不利影响。随着计算机视觉技术的不断发展,科研人员开始采用深度图像对室内异常行为的检测进行研究^[7-8],而且随着采集 RGB-D 图像设备的不断出现,RGB-D 图像得到了越来越多的研究者的青睐。微软公司于 2009 年 6 月推出了 Kinect 体感设备, Kinect 的突出功能就是处理深度图像,它的骨骼追踪技术就是通过处理深度图像来建立人体各个关节的坐标,能够通过骨骼追踪技术确定人体的各个部分以及它们的位置。

在传统的检测方案中,需要人来对监控图像进行不断观察,这种方法局限性太强,不仅耗费了大量的人力物力,不能及时发现可能出现的异常行为,而且传统的检测方案对检测到的情况不能进行自动智能化的处理。因此,文中基于 Kinect 设备,对室内环境进行检测,利用深度图像能够克服外界光线条件变化和遮挡等不利因素影响的优点,提高对摔倒行为检测的准确率,并且能够对异常行为进行自动化处理,实现对异常行为的实时性检测,达到保护人体安全的目的^[9-10]。

1 Kinect 简介

Kinect^[11]是微软公司于 2009 年 6 月推出的具有里程碑意义的一种体感设备。它具有 3 个 Camera,中间的是 RGB Camera,用于获取 640×480 的彩色图像,每秒钟最多获取 30 帧彩色图像;两侧是两个深度(3D Depth)传感器,使用红外线检测玩家的相对位置,其原理和人眼立体成像原理一样;Kinect 两侧各有两组麦克风用来做 3D 语音识别,能够捕捉到声源附近各种各样的信息;下面底座内有一个马达可以调整 Kinect 的仰角,其仰角可调整范围大小为 43°。具体如图 1 所示。

2 骨骼图像获取方法

文中利用 Kinect 来获取用户的骨骼和关节点。Kinect 左右两边的红外线发射器和红外线 CMOS 摄影机构成了 3D 结构光深度传感器。该传感器通过黑白光谱的方式来感知环境:纯黑代表无穷远,纯白代表

无穷近。黑白间的灰色部分对应物体到传感器之间的物理距离。Kinect 设备运行时,红外线发射器发出激光覆盖整个 Kinect 的可视范围,激光在散射体表而产生漫反射,红外线/VGA 摄像头组接收漫反射光线,通过激光散斑的照明,利用光编码技术,对可视范围内的空间进行编码。然后, Kinect 对获取的具有三维纵深的编码进行解码运算,获得相应的彩色图像数据、深度图像数据以及骨架模型数据。 Kinect 与计算机通过 USB 接口相连,所获数据通过 USB 端口输入计算机,通过应用程序编程接口(application program interface, API)读取相关数据,按程序中规定的算法进行运算。 Kinect 提供了人体骨架中的 20 个关节点(见图 2),所有关节点依靠深度图像技术都可通过空间三维坐标(x, y, z)进行定义,其中 VGA 摄像头提供 x 和 y 坐标值,红外摄像头提供 z 向深度值。



图 1 Kinect 传感器

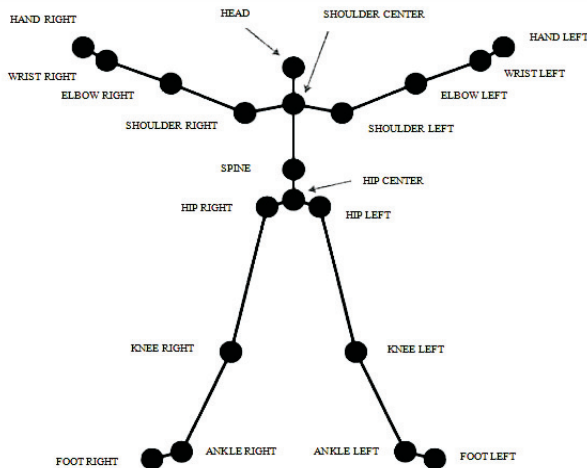


图 2 Kinect 识别到的 20 个人体关节点

与深度图像空间坐标不同的是,这些坐标单位是 m。坐标轴 x, y, z 是深度感应器实体的空间 x, y, z 坐标轴。这个坐标系是右手螺旋的, Kinect 感应器处于原点位置, z 坐标轴则与 Kinect 感应的朝向一致。 y 轴正半轴向上延伸, x 轴正半轴(从 Kinect 感应器的视角来看)向左延伸,如图 3 所示。

3 检测过程

当人体做出不同的动作时,由于人的动作和骨骼关节点空间位置、相对位置信息的变化有直接的关系,

所以当这些关节位置的数据发生变化时,可以根据对这些参数的处理来判断人体的动作^[12-14]。人体从站立到摔倒是一个非常短暂的过程,所用时间很少。因此在实验中,对当前骨骼图像帧和之前 15 帧骨骼图像中的关节位置信息进行对比,按照 Kinect 30 帧/s 的更新速率,文中所选取的两帧图像之间的时间差别仅为 0.5 s。摔倒事件自动检测系统主要选用 Kinect 所获取的 20 个关节点中的 2 个用于算法判别^[15],分别为左肩、右肩,2 个关节点的三维坐标 (x, y, z) 信息返回到 Kinect For Windows SDK (software development kit) 开发包中,形成骨架信息流,供摔倒检测算法调用。

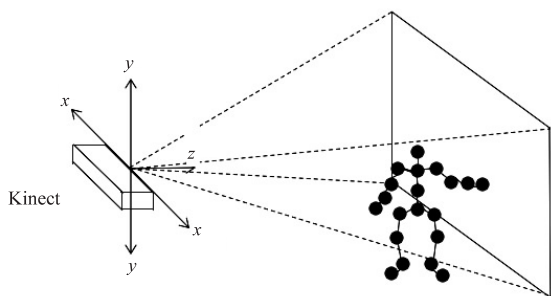


图3 人体骨骼图像坐标示意图

Kinect 位于离地面高度 1.2 m 的位置,以 Kinect 在空间中所处位置为原点 O ,建立空间三维坐标系 $O-X, Y, Z$ 。 $O-X, Y, Z$ 为定义的 Kinect 空间坐标系, Z 方向为 Kinect 的深度方向,与摄像头所在平面 XOY 垂直。以两肩中心骨架关节点为例,在 API 中可获得人体两肩中心在坐标系 $O-X, Y, Z$ 内的坐标 (x, y, z) ,并以 30 帧/s 的速度进行更新。人体两肩中心点在测试场景内第 f_i 帧的坐标为 $(x_{sc}^{(i)}, y_{sc}^{(i)}, z_{sc}^{(i)})$,第 f_{i+15} 帧的坐标为 $(x_{sc}^{(i+15)}, y_{sc}^{(i+15)}, z_{sc}^{(i+15)})$,经在 Kinect 设备上多次实验后,对实验结果进行统计,最终将阈值定为 $P=1.2$ m。计算所选取的两帧图像的两肩中心关节点在这两帧时间差内的位移大小,计算位移大小的原理是三维空间坐标系中的欧氏距离计算公式:

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2} \quad (1)$$

通过式(1)和阈值 P ,有如下不等式:

$$\sqrt{x^2 + y^2 + z^2} \geq P \quad (2)$$

式(2)中, $x = x_{sc}^{(i+15)} - x_{sc}^{(i)}$, $y = y_{sc}^{(i+15)} - y_{sc}^{(i)}$, $z = z_{sc}^{(i+15)} - z_{sc}^{(i)}$ 。

与此同时,根据第 f_i 帧和第 f_{i+15} 帧的两肩中心坐标点 $(x_{sc}^{(i)}, y_{sc}^{(i)}, z_{sc}^{(i)})$ 和 $(x_{sc}^{(i+15)}, y_{sc}^{(i+15)}, z_{sc}^{(i+15)})$ 计算该位移的空间向量,根据式(3)可以得到此位移向量。

$$\vec{AB} = (x_{sc}^{(i+15)} - x_{sc}^{(i)}, y_{sc}^{(i+15)} - y_{sc}^{(i)}, z_{sc}^{(i+15)} - z_{sc}^{(i)}) \quad (3)$$

然后根据摔倒时前后帧之间的位移向量与 $O-X, Y, Z$ 坐标系中的 Y 轴方向夹角 θ 加以区分摔倒和坐

下、下蹲等可能会出现误判的情况,当人体做出坐下或者下蹲动作时, θ 为 $0^\circ \sim 20^\circ$,而摔倒情况出现时, θ 为 $45^\circ \sim 75^\circ$ 。加上此判断条件后,经过实验验证,达到了预期,识别率达到了 92.3%。

4 实验

4.1 实验平台

实验采用 Intel (R) Core (TM) i3 CPU M350 @ 2.27 GHz 4 核。显卡为 AMD Radeon HD 5000。操作系统为 32 位 Windows 7 旗舰版,开发环境为 Visual Studio 2010, OpenCV 环境为 opencv-2.4.10。

4.2 实验结果与分析

实验人员站立于 Kinect 视场范围内,可以在计算机终端显示 4 幅图像和一个检测输出窗口,其中四幅图像分别为 color image, depth image, mask image 和 skeleton image,其中 mask image 是把 depth image 中的背景用黑色来填充,检测到的人体用白色填充的图像。图 4 即为站立时的 depth image, mask image 和 skeleton image。

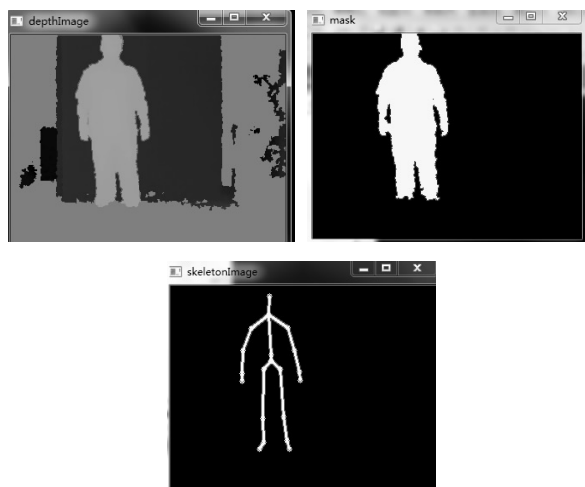


图4 站立时的深度图像及骨骼图像

当实验人员出现摔倒行为后,摔倒后的 depth image, mask image 和 skeleton image 如图 5 所示。

从摔倒后的骨骼图像看出,人体骨骼关节点信息发生了明显变化,其中左肩、右肩和两肩中心关节点位置在空间中的变化较为显著,结合实验结果的数据可以证明文中采用方法选取该关节点数据的科学性。

5 结束语

在现有的视频检测技术中,多数依赖于 RGB 图像的监控,这种方法不仅实时性差,而且不利于保护监控环境内的个人隐私。而采用基于 Kinect 的骨骼图像对摔倒行为进行监控,不仅能够实时地检测人体行为,而且精度高,同时有利于保护被检测人的隐私。但是由于 Kinect 拍摄的图像中包含了大量的数据信息,目

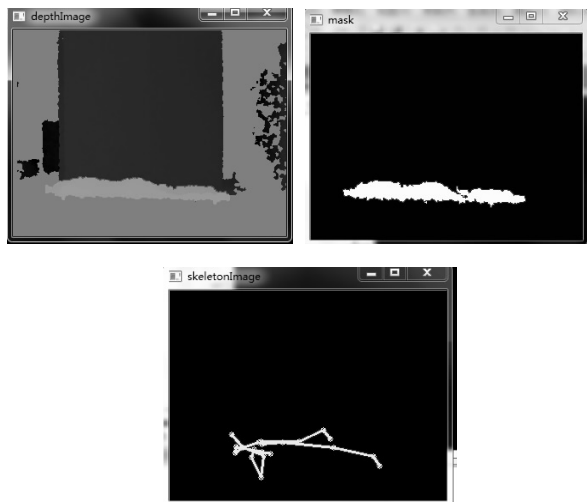


图 5 摔倒后的深度图像及骨骼图像

前计算机对如此数据量的信息进行提取计算的速度还不够高,还需要对程序的算法进行改进,提高计算效率。同时采用的设备为 Kinect1.0 产品,该产品能获取的视觉场景范围有限,并不能完整覆盖常见的家庭等室内环境。下一步要做的工作是在 Kinect2.0 设备的基础上对该方法进行开发研究,并且尝试采用双设备甚至多设备进行开发实验,进一步扩大该方案的应用场景并且提高检测率。

参考文献:

- [1] VAIDEHI V, GANAPATHY K, MOHAN K, et al. Video based automatic fall detection in indoor environment[C]// International conference on recent trends in information technology. [s. l.]: IEEE, 2011: 1016–1020.
- [2] 耳玉亮,段蕾蕾,叶鹏鹏,等. 2014 年全国伤害监测系统老

(上接第 178 页)

参考文献:

- [1] 李建荣,孔素真. 虚拟现实技术在教育中的应用研究[J]. 实验室科学, 2014, 17(3): 98–100.
- [2] 赵蔚,段红. 虚拟现实软件研究[J]. 计算机技术与发展, 2012, 22(2): 229–233.
- [3] 王文韬,谢阳群,李力. 虚拟现实技术在图书馆中的应用前景分析[J]. 图书馆, 2016(5): 10–14.
- [4] 马利亚. 基于虚拟现实的漫游系统关键技术的研究[D]. 哈尔滨: 哈尔滨工程大学, 2010.
- [5] 刘小英,张健. 三维校园虚拟现实平台的设计与实现[J]. 湖北农业科学, 2014, 53(12): 2909–2912.
- [6] 李清艳,傅自钢. 基于 OpenGL 的三维时钟设计[J]. 电脑与信息技术, 2016, 24(3): 52–54.
- [7] 肖健,魏雄,王仁波. 基于 OpenGL 的大型建筑三维场景模拟的实现与意义[J]. 电子质量, 2016(11): 88–92.
- [8] 马萍. 三维虚拟校园立体场景的设计与实现[D]. 济南: 山东师范大学, 2013.
- [9] 王琰萍. 山东信息职业技术学院虚拟校园的设计与实现

年跌倒/坠落病例特征分析[J]. 中华流行病学杂志, 2016, 37(1): 24–28.

- [3] 孔令富,李海涛. 基于位置和运动轨迹的老年人异常行为检测[J]. 计算机工程与设计, 2012, 33(2): 735–739.
- [4] 刘鹏,卢潭城,吕愿愿,等. 基于 MEMS 三轴加速度传感器的摔倒检测[J]. 传感技术学报, 2014, 27(4): 570–574.
- [5] 罗坚,唐璘,毛芳,等. 基于云计算的可穿戴式老年人异常行为检测系统研究[J]. 传感技术学报, 2015, 28(8): 1108–1114.
- [6] 韩磊. 图像序列中人的行为分析和识别方法[D]. 北京: 北京理工大学, 2009.
- [7] YANG S W, LIN S K. Fall detection for multiple pedestrians using depth image processing technique[J]. Computer Methods & Programs in Biomedicine, 2014, 114(2): 172–182.
- [8] 曹维清,李瑞峰,赵立军. 基于深度图像技术的手势识别方法[J]. 计算机工程, 2012, 38(8): 16–18.
- [9] LIU C L, LEE C H, LIN P M. A fall detection system using k-nearest neighbor classifier[J]. Expert Systems with Applications, 2010, 37(10): 7174–7181.
- [10] LIU Hong, ZUO Changling. An improved algorithm of automatic fall detection[J]. AASRI Procedia, 2012, 1: 353–358.
- [11] 石曼银. Kinect 技术与工作原理的研究[J]. 哈尔滨师范大学自然科学学报, 2013, 29(3): 83–86.
- [12] 黄国范,李亚. 人体动作姿态识别综述[J]. 电脑知识与技术, 2013, 9(1): 133–135.
- [13] 胡琼,秦磊,黄庆明. 基于视觉的人体动作识别综述[J]. 计算机学报, 2013, 36(12): 2512–2524.
- [14] 刘博,安建成. 基于关键姿势的人体动作识别[J]. 电视技术, 2014, 38(5): 38–41.
- [15] 谢亮,廖宏建,杨玉宝. 基于 Kinect 的姿势识别与应用研究[J]. 计算机技术与发展, 2013, 23(5): 258–260.

[D]. 青岛: 中国海洋大学, 2011.

- [10] 姜启业,程效军,谭凯. 基于 AutoCAD 和 3DMax 的建筑物三维建模[J]. 工程勘察, 2013(11): 71–74.
- [11] 田文波. 成都理工大学虚拟校园建设的研究[D]. 成都: 成都理工大学, 2013.
- [12] ZHOU Z, FENG Y, GANG R, et al. Virtual reality based process integrated simulation platform in refinery: virtual refinery and its application[J]. China Petroleum Processing & Petrochemical Technology, 2011, 13(3): 74–84.
- [13] NOVAKMARCINCIN J. Selected applications of virtual reality in manufacturing[J]. Journal for Technology of Plasticity, 2011, 36(1): 25–34.
- [14] KAMARAJ D C, DICIANNO B E, MAHAJAN H P, et al. Stability and workload of the virtual reality-based simulator-2[J]. Archives of Physical Medicine & Rehabilitation, 2016, 97(7): 1085–1092.
- [15] BAMODU O, YE X M. Virtual reality and virtual reality system components[J]. Advanced Materials Research, 2013, 765–767: 1169–1172.