

基于姿态估计的实时跌倒检测算法

于乃功[†], 柏德国

(1. 北京工业大学 信息学部, 北京 100124; 2. 计算智能与智能系统北京重点
实验室, 北京 100124; 3. 数字社区教育部工程研究中心, 北京 100124)

摘要: 为了快速、准确地检测老年人跌倒事件的发生, 给出一种基于姿态估计的实时跌倒检测算法. 首先利用基于深度学习的人体姿态估计算法得到人体关节点的坐标; 然后, 通过计算人体质心点的下降速度、跌倒后颈部关节点的纵坐标值是否大于阈值、以及肩部和腰部关节点的相对位置关系来判断跌倒是否发生. 所给出的跌倒检测算法利用单目相机进行检测, 便于以嵌入式方式应用于机器人. 实验结果表明, 所提出算法与当前先进方法相比取得了较好的效果.

关键词: 跌倒实时检测; 深度学习; 人体姿态估计; 单目相机; 机器人; 嵌入式

中图分类号: TP391

文献标志码: A

DOI: 10.13195/j.kzyjc.2019.0382

开放科学(资源服务)标识码(OSID):



引用格式: 于乃功, 柏德国. 基于姿态估计的实时跌倒检测算法[J]. 控制与决策, 2020, 35(11): 2761-2766.

Real-time fall detection algorithm based on pose estimation

YU Nai-gong[†], BAI De-guo

(1. Faculty of Information Technology, Beijing University of Technology, Beijing 100124, China; 2. Beijing Key Laboratory of Computational Intelligence and Intelligent System, Beijing 100124, China; 3. Digital Community Ministry of Education Engineering Research Center, Beijing 100124, China)

Abstract: In order to quickly and accurately detect the occurrence of falls in the elderly, this paper presents a real-time fall detection algorithm based on pose estimation. Firstly, the human pose estimation algorithm based on deep learning is used to obtain the coordinates of the joint point. Then, by calculating the falling speed of the centroid point when the human body falls, whether the ordinate value of the neck joint point after the fall is greater than the threshold, and the relative positional relationship of the shoulder-waist joint point in the image, whether the fall occurs is determined. The algorithm uses a monocular camera to detect, which is easily used in an embedded way for robots. The experimental results show that compared with the current advanced methods, the proposed algorithm has achieved good results.

Keywords: fall detection in real-time; deep learning; human pose estimation algorithm; monocular camera; robots; embedding

0 引言

在美国, 从 2007 年到 2016 年老年人因跌倒而死亡的人数呈上升趋势, 并且在 2016 年间, 年龄大于 64 岁的美国老年人中有 29 668 人因跌倒而死亡^[1]. 在我国, 2014 年全国伤害监测系统数据显示, 造成老年人因非故意伤害就诊的首要原因是跌倒/坠落^[2]. 如果老年人跌倒后能够及时得到治疗, 则可以降低 80% 的死亡风险和 26% 的长期住院治疗风险^[3], 所以根据上述问题设计一套老年人跌倒实时检测系统很有意义.

目前, 最常见的跌倒检测方法主要有 3 种: 基于

环境设备的检测方法、基于穿戴式传感器的检测方法以及基于视觉识别的检测方法. 基于环境设备的检测方法^[4-6]根据人体跌倒时形成的环境噪声进行检测, 这种方法通过感知物体压力和声音的变化检测跌倒, 误报率较高, 实际生活中很少采用此方法. 基于穿戴式传感器的方法^[7-10]利用加速度计和陀螺仪来进行跌倒检测. 文献[9]中利用加速度计和陀螺仪获取信号, 通过 SVM 分类器判断是否跌倒, 获得了 93% 的准确率. 文献[11]利用三轴加速度计获取信号, 通过多种机器学习算法进行分类判断是否跌倒. 此方法的缺点是长时间配戴传感器影响人的生活舒适度,

收稿日期: 2019-03-29; 修回日期: 2019-07-27.

基金项目: 国家自然科学基金项目(61573029); 国家重点研发计划项目(2017YFC0209905).

责任编辑: 薛建儒.

[†]通讯作者. E-mail: yunaigong@bjut.edu.cn.

并且由于老年人记忆力不好容易忘记佩戴,从事复杂活动时误报率较高.基于视觉识别的检测方法从利用传感器种类的角度可分为两类:一类是使用RGB相机提供的二维信息进行检测^[12-13];另一类使用深度相机Kinect提供的骨架关节点^[14-15]进行检测.文献[16]利用摄像头捕获视频,然后通过帧差法、卡尔曼滤波、光流法提取有效的背景信息并作为KNN分类器的输入,实现了96%的精确率.文献[17]为了保护老年人的隐私,提出了一种光学匿名图像传感系统,然后用卷积神经网络和自动编码器进行特征提取和分类.王君泽等^[15]利用Kinect提供的6个骨架点,实时计算6个骨架点的空间位置等参数来判定人体摔倒,但相对于单目相机而言,它的体积较大,价格昂贵,不适合嵌入式集成和批量使用.

当前,使机器人具备跌倒检测功能的研究工作较少.文献[18]中利用电脑控制机器人并用Kinect进行跌倒检测,精确率为80%.文献[19]中利用单目相机通过检测人体的外部轮廓获取轮廓比例,进行姿势标注以判断跌倒事件的发生.

根据上述分析,为了使跌倒检测方法具备实时性,并且做到跌倒检测设备能够嵌入式应用,本文利用普通单目相机,从一个新的角度进行跌倒检测.首先利用人体姿态估计算法,实时计算得到人体的骨架图和关节点坐标;然后将得到的坐标结合提出的跌倒检测算法进行检测.实验结果表明了该方法的有效性.

1 跌倒检测算法

算法首先利用Cao等^[20]提出的使用部分亲和域的实时多人2D姿态估计算法,实时获取人体的骨架图,得到的18个关节点如图1所示.算法用到了0(鼻)、1(颈)、2(左肩)、5(右肩)、8(左腰)、11(右腰)这6个关节点.其次,通过对获取到的骨架图遍历每个关节点,实时计算得到各个关节点在图像中的坐标 (x_i, y_i) ,下标 i 代表第 i 个关节点.

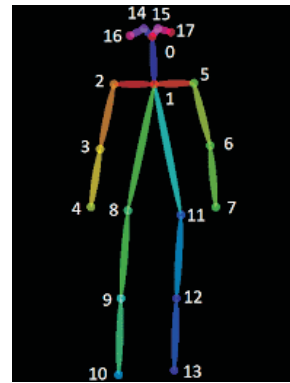


图1 姿态估计算法获取到的18个关节点

为了提高跌倒检测算法的通用性,分别研究单目相机安装在固定位置处的跌倒检测和单目相机安装在机器人上的跌倒检测方法.

1.1 单目相机安装在固定位置处的跌倒检测

本文提出的跌倒检测算法包括3个判定条件,其判定步骤如图2所示.

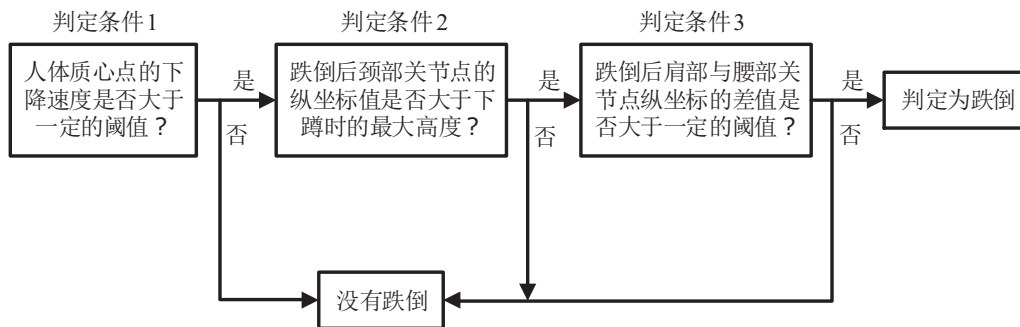


图2 跌倒检测算法判定步骤

1.1.1 判定条件1

为了保证检测的稳定性,同时取0、1、8这3个关节点的坐标,计算其坐标的平均值并作为人体质心点的坐标值.算法每10帧计算一次质心点纵坐标的移动速度 v .如果 v 大于一定的阈值 v_{th} ,则满足判定条件1. v 的表达式为

$$v = \frac{\sum_i y_{i_{10}} - \sum_i y_{i_1}}{3(t_{10} - t_1)}, \quad i = 0, 1, 8. \quad (1)$$

其中: $y_{i_{10}}$ 代表第10帧时第 i 关节点的纵坐标, y_{i_1} 代表第1帧时第 i 关节点的纵坐标, $t_{10} - t_1$ 代表从第1帧到第10帧所用的时间.判定条件1的表达式为

$$v \geq v_{th}. \quad (2)$$

为了确定阈值 v_{th} ,5个实验者分别模拟15次人体缓慢跌倒,然后求平均值得到 v_{th} ,即

$$v_{th} = \frac{1}{75} \sum_{i=1}^5 \sum_{j=1}^{15} v_{ij}. \quad (3)$$

其中: i 代表第 i 个人, j 代表第 j 次实验. 经过实验,将 v_{th} 设定为125 pixel/s,表示跌倒时每1s下降125个像素距离,其中pixel代表像素.

1.1.2 判定条件2

以 $v \geq v_{th}$ 作为判定跌倒是不充分的,这时会有很大的误判情况,例如人体下蹲系鞋带、捡物品等. 因此,为了排除上述情况,在人体下蹲时获取1关节点在图像中的最大高度 H_{th} . 判定条件1满足后每10帧获取1次 y_1 ,如果在一定时间内 y_1 始终大于 H_{th} ,则说明可能发生了跌倒. 判定条件2的表达式为

$$y_1 \geq H_{th}. \quad (4)$$

人体与相机不同距离时,关节点在图像中的高度 y_i 是不同的. 人体与相机较近时, H_{th} 比较大. 随着人距离相机逐渐变远, H_{th} 会逐渐减小,这样也会导致不同的 H_{th} . 所以首先需要判断人与相机的距离,然后确定此距离下的 H_{th} . 判断人与相机的距离,本文给出一种利用单目相机测距的新方法.

人与相机较近时,1关节点与8关节点两点之间的欧氏距离 l 较大. 随着人与相机越来越远, l 会逐渐减小,因此可用来判断人与相机的距离 L . l 的具体计算公式为

$$l = \sqrt{(x_1 - x_8)^2 + (y_1 - y_8)^2}. \quad (5)$$

由上述分析可知, l 、 H_{th} 和 L 之间呈线性关系,因此,三者之间可建立线性回归模型. L 与 l 之间的表达式为

$$L(l_i) = wl_i + b, \quad (6)$$

其中 i 代表第 i 次. 确定 w 与 b 的最优值可用最小二乘估计算法求解,即

$$(w^*, b^*) = \arg \min_{(w, b)} \sum_{i=1}^m (L(l_i) - l_i)^2, \quad (7)$$

$$E_{(w, b)} = \sum_{i=1}^m (L_i - wl_i - b)^2. \quad (8)$$

最优解如下式所示:

$$w = \frac{\sum_{i=1}^m L_i(l_i - \bar{l})}{\sum_{i=1}^m l_i^2 - \frac{1}{m} \left(\sum_{i=1}^m l_i \right)^2}, \quad (9)$$

$$b = \frac{1}{m} \left(\sum_{i=1}^m L_i - l_i \right), \quad (10)$$

其中 $\bar{l} = \frac{1}{m} \sum_{i=1}^m l_i$.

同理可得 L 与 H_{th} 之间的关系,由此确定了 H_{th} . 但是,当人快速坐在地上或下蹲长时间不起时,

判定条件2会造成误判,因为这时1关节点的纵坐标始终大于 H_{th} ,所以需要进一步判定.

1.1.3 判定条件3

如图3所示,人体跌倒后,左右肩与左右腰的关节点的纵坐标的差值是非常小的,因此,可用此条件判定人体是否跌倒. 判定条件3的表达式为

$$|y_2 - y_8| < h. \quad (11)$$

式(11)中 h 是5个实验者分别模拟20种跌倒后不同的姿势,20种姿势包括前倒姿势6种,后倒姿势6种,随机侧倒姿势8种. h 的表达式为

$$h = \frac{\sum_{i=1}^5 \sum_{j=1}^{20} |y_{ij2} - y_{ij8}|}{100}. \quad (12)$$

其中: i 代表第 i 个人, j 代表第 j 次实验. 经过实验, h 设定为140 pixel.

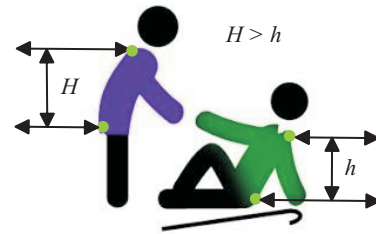


图3 判定条件3示意

1.2 单目相机安装在机器人设备上的跌倒检测

由于机器人高度的限制,相机安装位置过低有时会获取不到所用的6个关节点. 为了获取到全身的骨格点图,这里增加了第2个相机以便联合进行判断. 第2个相机比第1个相机高出一定的距离,第1个相机为仰视角度,第2个相机为俯视角度. 设置的具体参数将在实验结果中说明. 运用两个相机联合判断的方法是:首先利用第1个相机检验是否满足前两个判定条件;如果前两个判定条件满足,则开启第2个相机检测是否满足第3个判定条件;如果这3个条件全都满足,则说明跌倒发生了.

2 实验结果与分析

测试的摔倒情况分为前倒、后倒、随机地向各方向侧倒、下蹲系鞋带、缓慢或者快速坐在椅子上. 实验者一共5人,为了增强实验说服力,选取的实验者身高分别为1.60 m、1.72 m、1.78 m、1.83 m、1.9 m,所提出的算法只检测人体的骨架图,算法效果与实验者的体型、性别、年龄因素无关. 每人分别对上述5种情况做10次实验. 为了增强算法的鲁棒性,实验模仿家居环境,实验室的环境下摆放书桌、杂物、椅子等. 实验中设置的视频的分辨率为1920×1080.

2.1 相机安装在固定位置处的实验

第1个实验,将相机安装在距地面 1.55 m 的高度,俯视角度为 20°. 图4表示的是相机安装在固定位置时成功检测跌倒的效果. 表1为相机安装在固定位置时具体实验数据结果.



图4 相机在固定位置时的检测结果

表1 相机安装在固定位置时的实验数据

实验场景	实验者(每人进行10次成功检测出的次数)					成功率/%
	1	2	3	4	5	
前倒	8	9	9	8	9	86
后倒	9	7	8	8	7	78
侧倒	9	8	9	10	8	88
下蹲	9	9	10	9	10	94
坐在座位上	10	10	10	10	10	100

由表1的结果可知,检测后倒时的成功率最低,因为后倒有时会得不到身体的关节点. 为此,在同一高度处的对面方向即人体后方安装第2个相机,用两个相机同时检测,如果有一个相机检测到跌倒,则判定为跌倒. 表2为安装两个相机时的测试结果. 实验数据表明,用两个相机联合检测,提高了检测的精度.

表2 两个相机安装在固定位置时的实验数据

实验场景	实验者(每人进行10次成功检测出的次数)					成功率/%
	1	2	3	4	5	
前倒	9	10	9	9	9	92
后倒	10	9	9	9	8	90
侧倒	10	8	9	10	10	94
下蹲	9	10	10	10	10	98
坐在座位上	10	10	10	10	10	100

2.2 相机安装在机器人上的实验

第2个实验,模拟相机安装在机器人设备上. 将第1个相机置于距地面 30 cm 高度处,仰视角度为 40°. 第2个相机置于距地面 75 cm 高度处,俯视角度为 25°. 图5表示的是相机安装在机器人设备上成功

检测跌倒的效果. 表3为相机安装在机器人上的实验数据结果.



图5 相机在机器人设备上的检测结果

表3 相机安装在机器人上的实验数据

实验场景	实验者(每人进行10次成功检测出的次数)					成功率/%
	1	2	3	4	5	
前倒	10	9	9	8	9	90
后倒	9	9	8	8	9	86
侧倒	8	9	9	10	9	90
下蹲	9	10	10	10	10	98
坐在座位上	10	10	10	10	10	100

由以上实验结果可知:相机在固定位置时实现了 94.8% 的平均检测精度;相机在机器人设备上时,实现了 92.8% 的平均检测精度.

为了验证本文算法的有效性,将本文算法与其他基于视觉传感器的跌倒检测算法的性能进行对比. 评价的准则采用文献[16]给出的度量标准:

- 1) 真阳性(true positive):是跌倒的实验中检测到为跌倒的次数;
- 2) 假阳性(false positive):不是跌倒的实验中却被检测为跌倒的次数;
- 3) 真阴性(true negative):不是跌倒的实验中没有被检测为跌倒的次数;
- 4) 假阴性(false negative):是跌倒的实验中却没有被检测为跌倒的次数.

根据上述指标可得到算法的敏感度(sensitivity)、特异度(specificity)和精确率(accuracy)分别为

$$\text{sensitivity} = \frac{TP}{TP + FN}, \quad (13)$$

$$\text{specificity} = \frac{TN}{TN + FP}, \quad (14)$$

$$\text{accuracy} = \frac{TP + TN}{TP + FP + FN + TN}. \quad (15)$$

基于以上性能指标,各算法的对比结果如表 4 所示,显然,本文方法的特异性均优于其他方法.虽然本文方法有两个指标均次于文献[16]的方法 1,但是文献[16]中使用帧差法提取人体背景,这种方法容易受到周围运动物体的影响,例如人的周围有动物等,所以其没有做这种干扰性的实验;而本文方法只检测人类的骨架图,不受周围运动物体的影响.另外,基于

帧差法的背景提取,如果运动目标在相邻帧之间位置变化缓慢,则目标很难检测出来,且此种方法相机只能安装在固定位置处.文献[22, 24]是用深度相机实现的,其成本相较于普通 RGB 相机高昂,体积大,不便于嵌入式.综合成本和性能指标考虑,本文方法有一定优越性.另外,上述方法都不能用于机器人设备,而本文方法则适用.

表 4 本文方法与其他方法的对比

方法	敏感度 / %	特异度 / %	精确率 / %	传感器
帧差法 + 卡尔曼滤波 + 光流 ^[16]	96	97.6	96.9	RGB
多特征组合 ^[21]	未知	未知	90.9	RGB
HMM 模型 ^[22]	未知	93.3	95.24	Kinect
3D CNN + 自动编码器 ^[17]	93.3	92.8	未知	RGB + 热成像
光流 + 声音 ^[23]	90	92	未知	RGB
深度图像特征 + SVM ^[24]	100	97.5	未知	Kinect
本文方法	92	99	94.8	RGB

本文给出的跌倒检测算法的缺点一是在遮挡严重的情况下效果不是很理想,二是相机所处的位置太低时有时会获取不到人体骨架图从而导致检测失败.为了改善系统效果,提升检测精度,下一步的工作将在以下方面进行:

- 1) 改善人体姿态估计算法,增加人体骨架图检测的精度,进一步降低误报率;
- 2) 单个相机视野不开阔,且在有障碍物的情况下有时检测不到人体骨架图,因此可尝试用多个相机从多角度检测,以提高跌倒检测的成功率;
- 3) 模拟跌倒场景种类较为简单,需增加各种场景下的跌倒实验.

3 结 论

为了检测到老年人在居家生活条件下跌倒事件的发生,本文提出了基于人体姿态估计的跌倒检测算法.通过实时检测人体骨架得到关节在图像中的坐标,然后结合 3 个判定条件来确定跌倒事件的发生.实验表明,该方法适用于相机安装在固定位置处和相机安装在机器人设备上,所使用的检测设备便于嵌入式应用,提升了实用性.

参考文献(References)

[1] Burns E R, Kakara R. Deaths from falls among persons aged ≥ 65 years — United States, 2007-2016[J]. *Morbidity and Mortality Weekly Report*, 2018, 67(18): 509-514.

[2] 耳玉亮, 段蕾蕾, 叶鹏鹏, 等. 2014 年全国伤害监测系统老年人非故意伤害病例特征分析[J]. *中国健康教育*, 2016, 32(4): 312-317.

(Er Y L, Duan L L, Ye P P, et al. Analysis on the characteristics of unintentional injuries among the elderly from Chinese national injury surveillance system, 2014[J]. *Chinese Journal of Health Education*, 2016, 32(4): 312-317.)

[3] Noury N, Rrmeau P, Bourke A K, et al. A proposal for the classification and evaluation of fall detectors[J]. *IRBM*, 2008, 29(6): 340-349.

[4] Mohamad Daher, Maan El Badaoui El Najjar, Ahmad Diab, et al. Automatic fall detection system using sensing floors[J]. *International Journal of Computing & Information Sciences*, 2016, 12(1): 75-82.

[5] Minvielle L, Atiq M, Serra R, et al. Fall detection using smart floor sensor and supervised learning[C]. *The 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). Seogwipo*, 2017: 3445-3448.

[6] Li Y, Ho K C, Popescu M. A microphone array system for automatic fall detection[J]. *IEEE Transactions on Biomedical Engineering*, 2012, 59(5): 1291-1301.

[7] Wang F T, Chan H, Hsu M, et al. Threshold-based fall detection using a hybrid of tri-axial accelerometer and gyroscope[J]. *Physiological Measurement*, 2018, 39(10): 105002.

[8] Ranakoti S, Arora S, Chaudhary S, et al. Human fall detection system over IMU sensors using triaxial accelerometer[J]. *Computational Intelligence: Theories, Applications and Future Directions: Volume I*, 2018, 798: 495-507.

[9] Kumar V S, Acharya K G, Sandeep B, et al. Wearable sensor-based human fall detection wireless system[J].

- Wireless Communication Networks and Internet of Things, 2018, 493: 217-234.
- [10] Mehmood A, Nadeem A, Ashraf M, et al. A novel fall detection algorithm for elderly using SHIMMER wearable sensors[J]. Health and Technology, 2019, 9(4): 631-646..
- [11] Sara Zulj, Goran Seketa, Igor Lackovic. Accuracy comparison of ML-based fall detection algorithms using two different acceleration derived feature vectors[C]. IFMBE Proceedings of World Congress on Medical Physics and Biomedical Engineering 2018. Singapore: Springer, 2018: 481-485.
- [12] 宋菲, 薛质. 基于 Open CV 的老人跌倒检测的设计和实现[J]. 信息技术, 2015, 39(11): 137-139.
(Song F, Xue Z. Design and implementation of fall detection for elder person based on Open CV[J]. Information Technology, 2015, 39(11): 137-139.)
- [13] Yu M, Gong L, Clifford R, et al. A monocular camera based person-specific fall detection system exploiting deep neural network aided unsupervised[C]. IEEE International Conference on Acoustics, Speech, and Signal Processing. Brighton, 2019.
- [14] Patsadu O, Watanapa B, Dajpratham P, et al. Fall motion detection with fall severity level estimation by mining kinect 3D data stream[J]. The International Arab Journal of Information Technology, 2018, 15(3): 378-388.
- [15] 王君泽, 朱小龙, 瞿畅. 基于 Kinect 骨架追踪的人体摔倒自动检测[J]. 上海交通大学学报, 2015, 49(9): 1359-1365.
(Wang J Z, Zhu X L, Qu C. Automatic fall detection using human skeleton tracking algorithm based on kinect sensor[J]. Journal of Shanghai Jiaotong University, 2015, 49(9): 1359-1365.)
- [16] De Miguel K, Brunete A, Hernando M, et al. Home camera-based fall detection system for the elderly[J]. Sensors, 2017, 17(12): 2864.
- [17] Ma C, Shimada A, Uchiyama H, et al. Fall detection using optical level anonymous image sensing system[J]. Optics & Laser Technology, 2019, 110: 44-61.
- [18] Sumiya T, Matsubara Y, Nakano M, et al. A Mobile robot for fall detection for elderly-care[J]. Procedia Computer Science, 2015, 60: 870-880.
- [19] Sun P F, Ming A, Yao C, et al. Fall detection system based on mobile robot[J]. Image and Graphics Technologies and Applications, 2018, 875: 263-271.
- [20] Cao Z, Simon T, Wei S E, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]. The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Hawaii, 2017: 7291-7299.
- [21] Glen Debard, Peter Karsmakers, Mieke Deschodt, et al. Camera based fall detection using multiple features validated with real life video[C]. Workshop Proceedings of the 7th International Conference on Intelligent Environments. Nottingham, 2011: 441-450.
- [22] Triantafylou D, Krinidis S, Ioannidis D, et al. A real-time fall detection system for maintenance activities in indoor environments[J]. IFAC-PapersOnLine, 2016, 49(28): 286-290.
- [23] Geertsema E E, Visser G H, Viergever M A, et al. Automated remote fall detection using impact features from video and audio[J]. Journal of Biomechanics, 2019, 88: 25-32.
- [24] Panahi L, Ghods V. Human fall detection using machine vision techniques on RGB-D images[J]. Biomedical Signal Processing and Control, 2018, 44: 146-153.

作者简介

于乃功(1966—), 男, 教授, 博士生导师, 从事计算智能与智能系统、机器人学与机器人技术、机器视觉等研究, E-mail: yunaigong@bjut.edu.cn;

柏德国(1995—), 男, 硕士生, 从事智能机器人、计算机视觉的研究, E-mail: baideguo@emails.bjut.edu.cn.

(责任编辑: 李君玲)