

适用于家庭服务机器人的倒地人体检测方法^①

刘松, 陈小平

(中国科学技术大学 计算机科学与技术学院, 合肥 230027)

摘要: 人体检测是家庭服务机器人的一项基本功能。本文针对复杂家庭环境, 倒地人体面临地上杂物的干扰、遮挡等情况下, 提出一种结合三维点云分割和局部特征匹配的倒地人体检测方法。该方法对点云进行分割之后将每个物体横向切分成若干段, 对每段点云采用局部特征匹配并分类, 并根据匹配段数来判断是否为倒地人体。实验结果表明, 该方法在 0.3 秒的检测时间内, 实现平均误识别率低于 10% 的高检测率, 满足服务机器人实时性要求的同时具有良好的鲁棒性, 即使人体部分被遮挡, 依然可以检测到各种倒地姿态的人。

关键词: 家庭服务机器人; 三维点云分割; 倒地人体检测; 特征匹配

Detecting Lying Person for Home Service Robots

LIU Song, CHEN Xiao-Ping

(School of Computer Science and Technology, University of Science and Technology of China, Hefei 230027, China)

Abstract: Human detection is a basic functionality for home service robots. For complex family environments where lying person is partially occluded or in cluster, this paper proposes a lying person detection approach integrates 3D point cloud segmentation and local feature matching. Our approach segments the point cloud of each object into several pieces, matches local features of each object piece, and classifies them to detect lying person. Experiments show that our approach can achieve high detection accuracy with average recognition time less than 0.3s. Our approach meets the human detection requirements for service robots and is demonstrated to be practical and reliable, even when parts of human body is occluded.

Key words: home service robot; lying person detection; 3D point cloud segmentation; feature matching

家庭生活中人们经常因为各种原因摔倒在地, 导致骨折或其他损伤, 很难通过自身努力站立起来, 然而长时间躺在地上容易引发安全事故。及时检测倒地人体并呼救可以将死亡率降低 80%, 送医治疗的次数降低 30%^[1]。因此, 一种可信赖的方法来检测倒地人体并呼救就显得非常重要。

人体检测是机器视觉中重要研究方向。一部分的人体检测研究是基于人脸识别的, 如^[2-4]。另一部分是基于行人检测的, 如^[5-8]。然而, 目前人体检测往往聚焦于人体处于直立状态, 当人倒地状态时, 情况更加复杂。人跌倒后躺在地上时往往是背对机器人, 机器人无法看到人脸。不仅如此, 与站立相比, 人处于倒地

状态时, 人体姿态千变万化, 机器人无法通过先验知识深度信息来判断人腿的位置, 实现人腿的跟踪。因此, 使用传统的人体检测方法来检测倒地人体时往往很难获得较好的效果。

对于倒地人体的检测, 也有一些方法可以适用于服务机器人上。基于 2D 的方法^[9-11], 使用多模型来处理不同视角, 不同姿势。但是这些方法起源于人处于半直立状态的人体检测, 方法特别复杂, 由于此时不能依赖经典的人体上半身模型和一些对称信息, 故检测率较低。随着深度摄像头的普及, 家庭服务机器人人体检测中使用 3D 方法越来越多。除了文献[12]运用了 Kinect 骨架追踪的方法, 其他方法主要分为两类, 基

^① 基金项目: 国家自然科学基金(61175057)

收稿时间: 2016-01-31; 收到修改稿时间: 2016-03-10 [doi:10.15888/j.cnki.csa.005376]

于特征匹配的方法和基于建模的方法^[13]。由于倒地人体很难获得完整的骨骼信息,姿态千变万化,暂时也无法获得一个优良的模型,故采用基于特征匹配的较多。而基于特征匹配的方法又可以细分为基于柱状图特征^[14-17]和基于几何统计特征^[17-18]两种,3D 物体识别所应用的特征主要有^[19-22]。^[17]在跌倒检测中分别应用四种特征和四种监督学习算法比较了各自的优劣性,但并没有考虑到倒地人体与直立人体检测环境的差异。

在倒地人体的检测中,人与地面接触面积较大,点云高度相近,易混淆。因此,地面点云的拟合分割需要更加精确。如果地面分割效果欠佳,会遗留大量杂点,影响物体点云的拟合;如果出现过拟合,则人体部分点云被当做地面剔除,导致人体形状发生较大变化,影响检测结果。另外,与人处于直立状态相比,人体旁边杂物更多,易发生重叠,因此,对物体分割中点云集聚也要求很高,若将两个物体重叠在一起,人体本身的特性被覆盖,区分度减小。而且家庭环境比较复杂,家具等杂物较多,可能只能观察到人体一部分信息。倒地人体自身也容易产生遮挡,例如手和腿被身体压住,身体弯曲在一起等,无法获得完整的人体信息,而这些情况在人处于直立状态时是不存在的。

本文针对以上提到的倒地人体检测的特殊性,提出一种适用于安装有深度摄像头的家庭服务移动机器人的倒地人体检测方法。该方法将三维点云分割和局部特征匹配结合在一起,使用点云信息与一般视觉图像方法相比,更易进行地面模拟和物体分割。为了解决倒地状态出现的遮挡问题,该方法在物体分割之后,对每个物体进行横向切分成若干段,然后对每一段提取特征训练分类,当正面分类满足一定段数时,即可判断该物体为倒地人体,有效解决了遮挡时部分信息缺失的问题。人体检测常用的图像特征有 HOG, FPFH, SURE、HLSN 等。在提取特征时,本文构造每一点云段的局部表面法向量直方图。与其他特征相比,使用 HLSN^[17]的主要优势在于能获得物体外表形状特征,而人倒地的外形特征与家里其他物体有较大的差异,一般家具如桌椅沙发等,都具有棱角,而人体类似圆柱状,故应用该特征可以较好的区分人体和其他家具。实验结果表明,该方法在满足移动机器人实时性要求的同时具有较好的鲁棒性,适用于各种人体倒地姿势

和机器人不同观测视角,即使在人体部分被自身或其他物体遮挡下,依然可以检测到各种姿态的人。在多次比赛中,“可佳”服务机器人运用该方法的取得了较好的成绩。

1 方法概述

本文所述方法总体框图如图 1 所示,总体可以分成五部分:预处理、点云分割、点云分段、特征提取、识别判断。输入为深度摄像头采集的深度图像,首先对点云数据进行预处理,之后进行地面拟合,然后进一步分割平面上的物体,再对每个物体进行横向切分成若干段,对每一段点云提取特征训练分类,最后根据点云段正面分类段数是否达到固定值判断是否为倒地人体,输出检测结果。在检测之前,该方法需要建立样例模板库。提前标定不同性别、年龄的人以不同姿势躺在地上作为正面模板,同时标定一些家中地上常见物体,比如家居,宠物等作为负面模板。

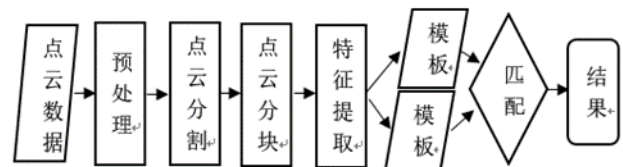


图 1 方法总体框图

2 各模段的算法设计与实现

2.1 预处理

预处理首先需要将深度图像数据转换为 3D 点云,由于采集的 3D 点云数据较多,在后面进行特征提取时计算量过大,影响该方法的实时性,因此,需要对点云进行稀疏化。本文采用 VoxelGrid 算法对点云体素网格进行向下采样,箱体大小设置为 $1\text{cm}\times 1\text{cm}\times 1\text{cm}$,将箱体内存所有点以重心代替。运用向下采样优于直接使用体素中心,虽然降低了时间效率,但是可以更好的保留物体表面特征。

根据坐标变换及内定参数平移点云,并以地面建立 $x-z$ 水平面坐标。由于倒地人体高度很低,不会高于 0.5m ,故可以过滤高于 0.55m 的物体,利用点云 ROI 技术过滤高于 0.55m 的点云,只考虑高度在 $0-0.55\text{m}$ 范围内的点云。

2.2 点云分割

点云数据经过预处理后,本文实现了一种快速高

效的点云分割方法,包括地面分割和物体分割两部分,分割之后进一步利用点云高度信息进行除杂,提高检测效率。

2.2.1 地面分割

在提取物体特征之前,首先需要进行平面分割将物体与地面分割开来。本文使用区域增长法代替基于RANSAC来分割水平面^[23]。由于RANSAC方法分割出的水平面容易包含较多法向量不垂直与水平面的杂点,且时间效率太低。而区域增长法则是一种比较普遍的方法,不需要先验知识也能取得最佳的性能。

由于采用地面作为水平面,故地面点云法向量都垂直与水平面,因此可采用这一准则,从一随机点云出发,将所有符合的点云进行聚类,分割出所有的水平面并根据阈值进行过滤。

2.2.2 物体分割

对平面上的物体进行分割,构成一个物体集合 $C(c_1, \dots, c_N)$ 。

物体分割同样采用区域增长法^[23],将地上不同物体分割开来,可以避免人体与地面上其他物体混淆在一起,减少杂点的影响。

2.2.3 除杂

在预处理阶段,我们已将高于0.55m以上的点云过滤掉,剩下的都是高度低于0.55m的点云,但是属于该物体的0.55m高度以下的部分还在,故该物体剩余部分高度近似0.55m,而倒地人体高度一般低于0.5m,故以0.05m为最大偏差阈值,进一步从集合C过滤掉高度在0.5~0.55m之间的物体。

2.3 点云分段

在复杂的家庭环境下,容易产生遮挡现象,机器人很难获得一个完整的倒地人体点云信息,故本文对人体点云进行横向分段检测。将每个物体 c_i 分成若干段,与^[15,16]不同的是,本文固定每段的横向长度,而层数不限。人一般高度集中于1.5-2m之间,故我们将每段固定长度定为0.25m,段与段之间允许0.05m的重叠,这样,正常倒地人体点云图像可以分成6-8段。由此,每个物体 c_i 生成了一个点云段序列 $S_i(s_{i1}, \dots, s_{ij})$ 。

基于每段固定0.25m的横向点云图像长度,故可以进一步除杂,过滤x横向长度小于0.25m的物体。

2.4 特征提取

人体检测常用的特征有Hog特征、局部表面法向量、快速特征点、SURE特征等,由于人体特征为圆柱

状,与家庭环境其他物体墙、桌椅、沙发等具有棱角特征具有较大差异,局部表面法向量特征可以利用这一特性较好区分室内倒地人体与其他物体,故本文在局部表面法向量直方图的基础上,结合一些2d和3d统计特征,组成特征向量。

经过分割之后,得到了物体集合C,每一物体 c_i 又包含若干段 S_i 。提取特征时,对每个物体的每一点云段 s_{ij} 上的所有点云计算其表面法向量,分别对每一极坐标(x, y, z)上建立一7个小区间的直方图,并加上每个点云的高度和深度信息组成特征向量,以此来检测倒地人体。

2.5 分类

本文采用支持向量机(SVM)来做分类器,内核函数为多项式。分别对每个物体的每一点云段 s_{ij} 进行分类,结果标记为1(人体点云段)和-1(非人体点云段)。如果一个物体的点云段标记结果大于设定阈值段数,即该物体有三段以上的点云段符合人体特征,则可判断该物体为倒地人体,否则没有发现倒地人体。该阈值可以根据容错度进行调节,如该方法应用于报警急救领域,则可以相应提高相似度阈值和增加最小段数大小,以减少误识别率。

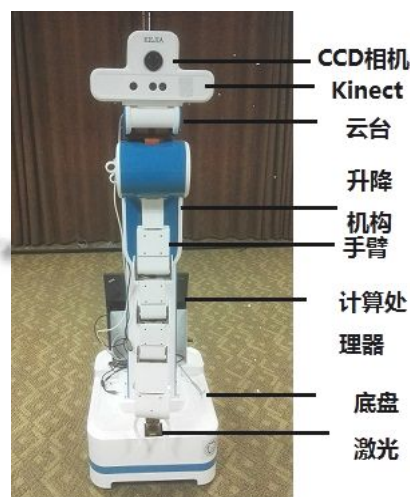


图2 移动家庭服务机器人“可佳”

3 实验

3.1 实验平台

本文的实验平台为移动家庭服务机器人“可佳”(如上图2)。其主要硬件结构有:最上面为视觉传感器方法包含一个高清相机Point Grey CCD和一个Kinect传感器,由一个具两自由度的云台承载,一个

双向差动轮底盘、一个六自由度的手臂等。计算中心部件为一台笔记本，配置为 Intel Core i7-2760QM 处理器，内存为 8GB，系统为 64 位 Ubuntu14.04。

3.2 实验数据



图 3 第一行图片为负面样本，下面两行图片为正面样本

目前已公开的人体数据库大部分是基于站立姿势的，倒地人体相关数据很少，故本文的实验数据均由家庭服务机器人“可佳”采集，包含正面模板数据和负面模板数据。由于手动采集和标注每一副图像数据非常耗时，故我们引入一个自动标注方法来采集数据。负面模板数据不包含人体，由人标记室内常见家具附近若干点，然后让机器人自动导航到每一个点，分别采集该点附近家具的图像，这样得到的模板数据更真实。而采集正面模板数据时，我们提前划定一 ROI 区域，该区域不含其它物体，然后让实验人员在区域自然跌倒，确保人体不被遮挡，能获取完整人体特征，保证在训练数据中人体每一部分特征数据的均衡，而不是基于某一特殊部分特征。而测试数据中，倒地人体部分被遮挡，保证其真实性。实验人员包含男生 16 人和女生 4 人，每个人采集 20 种不同的倒地姿势图像，共 400 副图像，负面物体包含沙发，健身器，床，椅子，货架等(部分样例如图 3)，分别采集不同角度的图像，共 100 副图像，正负样例共 500 例。如上图 3。

当该方法实际应用时，对识别率、误识别以及实时性有很高的要求，必须保证在各种环境下满足性能要求。测试时，我们分别测试了人跌倒后在无遮挡，有遮挡，有一般家具在附近，有形体类人物体出现等

(样例如图 4 所示)四种情况下的检测效果，我们将最小检测段数 n 设为 3，每种情况测试 30 次。测试结果如表 1。



图 4 有遮挡，椅子，类人物体出现场景

表 1 不同环境下方法检测性能

	无遮挡	有遮挡	椅子	类人物体
识别率(%)	90	86.67	83.33	73.33
误识别率(%)	7.4	7.69	16	20
检测时间(s)	0.252	0.252	0.256	0.263

从表 1 中可以看出，在以上各种情况下，方法识别率基本保持在 80% 以上，误识别率也在 20% 以下，检测时间在毫秒级，满足应用方法的基本要求。当出现类人物体时，误识别率有所增加，针对该问题，可以增加检测次数，累计 5 次或 10 次检测结果，然后取平均值判断，减小误识别率。

实时性和低误识别率是检测方法的基本要求，但同时保证两者往往有些困难。当应用该方法作为监测报警方法时，人跌倒后容易产生严重后果，高效的检测率很重要，但其低误识别率更重要，否则高频率的误识别，容易导致使用者失去信任而关掉该方法。因此，可以相应提高判断的最小点云段数 n ，而提高该阈值后，检测方法在检测倒地人体时往往需要更长的时间。在发生跌倒或危险状况后，本方法不需要机器人监测到人倒地的瞬间，如果跌倒者受伤比较严重自己难以站立长时间倒在地上，机器人还是有很多机会可以检测到这一状况。而如果跌倒者不是非常严重，跌倒者可以立即自己站立起来，虽然时间较短，这时就算机器人没有检测到这一状况也不要紧。故而我们可以允许该检测方法有一定的延时，来保证低误识别率。但也并不意味着该方法没有时间要求，跌倒者长时间不被发现也会错过黄金救治时间，造成严重后果，故该方法对实时性有一定要求。下图就是最小段数、误识别率、检测时间数据表，每种情况也是测试 30 次。测试结果如下表：

表2 阈值 n 取不同数值时的性能

最小段数 d	检测率(%)	误识别率(%)	检测时间(s)
2	93.33	14.28	0.247
3	90	7.4	0.252
4	76.67	4.3	0.267
5	63.3	5.2	0.284
6	46.67	7.1	0.311

从表2中可以看出,当 n 增加到3时,误识别率已有较大减小, n 再增大时,误识别率减小幅度有限,当 n 大于4时,误识别率反而出现小幅上升,主要由于检测率降低,误识别次数基本不变,导致误识别率不降反升.从中可以看出,一般 n 设为3比较合适.

4 结语

高效而可信赖的倒地人体检测是目前机器人视觉的一项难点.本文提出一种结合三维点云分割和局部特征匹配的倒地人体检测方法.该方法针对人体姿态复杂多样,部分人体容易被自身或其他物体遮挡的情况,通过对三维点云的分割、分段检测的方法,并采用基于局部特征匹配的方法实现倒地人体的检测.本文介绍的倒地人体检测方法在多智能体实验室家庭环境中的实验结果以及2015国际、国内RoboCup@Home人体检测测试结果表明了该方法的实时性和鲁棒性,能适应人体各种倒地姿势.然而本文人体点云特征匹配仅限于局部表面法向量,下一步可以综合其他特征来实现更精准的匹配,提高检测准确率.进一步我们可以研究机器人在家庭环境中的各个方向查找检测策略.

参考文献

- 1 Noury N, Rumeau P, Bourke A, Laighin G, Lundy J. A proposal for the classification and evaluation of fall detectors. IRBM. 2008. 29. 340-349.
- 2 Viola P, Jones M. Robust real-time object detection. International Journal of Computer Vision, 2001, 4.
- 3 Wu B, Ai H, Huang C, et al. Fast rotation invariant multi-view face detection based on real adaboost. Proc. of Sixth IEEE International Conference on Automatic Face and Gesture Recognition, 2004. IEEE. 2004. 79-84.
- 4 Li Y, Ai H, Yamashita T, et al. Tracking in low frame rate video: A cascade particle filter with discriminative observers

of different life spans. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2008, 30(10): 1728-1740.

- 5 Dollar P, Wojek C, Schiele B, et al. Pedestrian detection: an evaluation of the state of the art. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2012, 34(4): 743-761.
- 6 Enzweiler M, Gavrila D. Monocular pedestrian detection: Survey and experiments. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2009, 31(12): 2179-2195.
- 7 Geronimo D, Lopez AM, Sappa AD, et al. Survey of pedestrian detection for advanced driver assistance systems. IEEE Trans. on Pattern Analysis and Machine Intelligence, 2010, 32(7): 1239-1258.
- 8 Nian X, Xie K, Yang W, et al. A pedestrian detection method based on MB_LBP features and intersection kernel SVM. Proc. of the 2015 Chinese Intelligent Automation Conference. Springer Berlin Heidelberg. 2015. 361-369.
- 9 Wang S, Zabir S, Leibe B. Lying pose recognition for elderly fall detection. Robotics: Science and Systems VII, 2012, (1): 345-353.
- 10 Lv Q. A poselet-based approach for fall detection. 2011 IEEE International Symposium on IT in Medicine and Education. 2011. 209-212.
- 11 刘德建. 移动机器人平台下基于视觉的摔倒检测方法. 电子技术与软件工程, 2015, (18): 76-79.
- 12 王君泽, 朱小龙, 瞿畅. 基于 Kinect 骨架追踪的人体摔倒自动检测. 上海交通大学学报, 2015, 49(9).
- 13 Xia L, Chen CC, Aggarwal JK. Human detection using depth information by kinect. 2011 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE. 2011. 15-22.
- 14 Spinello L, Arras KO. People detection in RGB-D data. 2011 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). IEEE. 2011. 3838-3843.
- 15 Wu S, Yu S, Chen W. An attempt to pedestrian detection in depth images. 2011 Third Chinese Conference on Intelligent Visual Surveillance (IVS). IEEE. 2011. 97-100.
- 16 Hegger F, Hochgeschwender N, Kraetzschmar GK, et al. People detection in 3D point clouds using local surface normals. RoboCup 2012: Robot Soccer World Cup XVI. Springer Berlin Heidelberg, 2013: 154-165.
- 17 Spinello L, Spinello L, Arras K. A layered approach to

- people detection in 3D range data. National Conference on Artificial Intelligence. 2010.
- 18 Volkhardt M, Schneemann F, Gross HM. Fallen person detection for mobile robots using 3D depth data. 2013 IEEE International Conference on Systems, Man, and Cybernetics (SMC). IEEE. 2013. 3573–3578.
- 19 Spinello, Luciano, et al. A layered approach to people detection in 3D range data. AAAI. 2010.
- 20 Fiolka T, Stückler J, Klein DA, et al. Sure: Surface entropy for distinctive 3d features. Spatial Cognition VIII. Springer Berlin Heidelberg, 2012: 74–93.
- 21 Steder B, Rusu RB, Konolige K, et al. NARF: 3D range image features for object recognition. Workshop on Defining and Solving Realistic Perception Problems in Personal Robotics at the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS). 2010. 44.
- 22 Wang Y, Weinacker H, Koch B. A LiDAR point cloud based procedure for vertical canopy structure analysis and 3D single tree modelling in forest. Sensors, 2008, 8: 3938–3951.
- 23 柯翔, 陈小平, 靳国强, 等. 一种适用于室内服务机器人的实时物体识别方法. 计算机系统应用, 2013, 22(10): 84–89.