

DOI: 10.19650/j.cnki.cjsi.J2006065

基于模型约束的人体姿态视觉识别算法研究^{*}

刘今越^{1,2}, 刘彦开^{1,2}, 贾晓辉^{1,2}, 郭士杰^{1,2}

(1. 河北工业大学机械工程学院 天津 300130; 2. 河北省机器人感知与人机融合重点实验室 天津 300132)

摘要:针对移乘服务机器人对人体姿态识别高精度性的要求,以及现有的人体姿态识别方法在关节遮挡情况下识别精度低的问题,提出了一种基于模型约束的人体姿态识别算法,以解决移乘操作前机器人系统对人体关节空间坐标的精确提取。首先采用 OpenPose 算法识别彩色图像中未遮挡关节的像素坐标,通过对 RGB-D 相机的彩色图像与深度图像进行对齐,将关节像素坐标转换为 3D 坐标。然后依据人体模型相关参数以及未被遮挡关节坐标计算与之相连接的被遮挡关节的空间坐标,用于提高遮挡关节的识别精度。实验结果表明,所提出的算法在关节未遮挡时识别精度为 92%,在关节遮挡时达到了 90%。单帧计算平均用时约为 190 ms,满足移动服务机器人操作的实时性要求。

关键词: 移乘服务机器人; 人体姿态; 遮挡识别; RGB-D; 人体模型

中图分类号: TP242 TH789 **文献标识码:** A **国家标准学科分类代码:** 510.80 460.40

Research on human pose visual recognition algorithm based on model constraints

Liu Jinyue^{1,2}, Liu Yankai^{1,2}, Jia Xiaohui^{1,2}, Guo Shijie^{1,2}

(1. School of Mechanical Engineering, Hebei University of Technology, Tianjin 300130, China;

2. Hebei Key Laboratory of Smart Sensing and Human-Robot Interaction, Tianjin 300132, China)

Abstract: Aiming at the requirement of transfer service robots for high precision of human pose recognition and the problem of low recognition accuracy of existing human pose recognition methods under joint occlusion, this paper proposes a model constraint-based human pose recognition algorithm to solve the precise extraction of human body joint space coordinates of the robot system before the transfer operation. Firstly, the OpenPose algorithm is used to identify the pixel coordinates of the unoccluded joints in color image. Through aligning the color image and depth image of the RGB-D camera, the joint pixel coordinates are converted into 3D coordinates. Then, the spatial coordinates of the occluded joint connected with the unoccluded joint are calculated according to the relevant parameters of the human model and the unoccluded joint coordinates, which are used to improve the recognition accuracy of the occluded joints. The experiment results show that the recognition accuracy of the proposed algorithm is 92% when the joint is unoccluded, and reaches 90% when the joint is occluded. The average time for a single frame calculation is about 190 ms, which meets the real-time requirements of transfer service robot operation.

Keywords: transfer service robot; human pose; occlusion recognition; RGB-D; human body model

0 引言

随着机器人应用从工业领域向医疗、服务等行业不断地拓展,对机器人的安全性也提出了更高的要求^[1]。近些年,老年人、残疾人等弱势群体的行走、移动问题受到了各国科研团队的重视,服务机器人应运而生。如早

稻田大学推出的 Twendy-One 机器人^[2],可以为老年人和残疾人提供家庭护理工作。西安交通大学研制的助老伴行机器人^[3]可以辅助老年人行走,来解决老年人的行走问题。针对家庭中下肢失能、卧病在床的老年人,推出了一款既能托举也能抱人的移乘服务机器人,机器人若想实现上述协助老人移动的动作,最重要的是准确估计出人体关节的位置,即机器人要具备人体姿态识别的功能。

收稿日期:2020-02-09 Received Date:2020-02-09

^{*} 基金项目:国家重点研发计划(2019YFB1312103)、国家自然科学基金(U183222)、河北省教育厅重点项目(ZD2018246)资助

目前,人体姿态识别面临的问题是对硬件的依赖性较高以及对关节出现遮挡情况下的识别精度不足。

人体姿态估计主要是用来识别自然场景下人体关节的位置。人体姿态估计算法分为传统方法和基于深度学习的方法。传统方法大多是在深度图像中利用图像分割和图像形态学分析的手段定位人体骨架线的拐点作为关节点,这类方法受制于人体灵活运动的影响,模型可扩展性较差。李斌等^[4]应用图像分割的方法把人从背景中分割出来,从深度图中提取的人体骨架线的最大弯曲点视为关节点,此方法操作简单,但是精度低,对复杂环境以及人体姿势的适应性较差。近几年,大多数研究使用基于深度学习的方法从彩色图像中估计人体的 2D 关节坐标^[5-6]。Carreira 等^[7]创建了基于迭代误差反馈的人体姿态估计模型,使用 Coordinate 作为迭代目标,利用分层特征提取器提取关节热图,取关节热图中的得分最大值对应的像素坐标作为预测关节坐标,由于 Coordinate Net 没有对人体各个关节之间的结构信息建模,导致模型对多尺度的人体姿态估计泛化性较差。Yang 等^[8]使用树状结构构建人体全关节模型,得到一颗人体关节点关系树,并将树结构模型与深度卷积神经网络一起进行端到端训练,输出彩色图像中关节点像素坐标,因深度神经网络参数较多,导致计算效率低,实时性较差。Martinez 等^[9]利用深度前馈网络处理彩色图计算人体 3D 姿态,因受限于尺度与旋转归一化,无法提供全局坐标系下的人体姿态。Wei 等^[10]于 2016 年提出的卷积姿态机方法可以估计全身关节点,具有很强的鲁棒性,缺点是无法做到实时估计。Newell 等^[11]提出了堆叠沙漏网络,输入单张 RGB 图像,经过全卷积网络得到多尺度特征,再通过多个堆叠的沙漏,得到最终的关节像素坐标,劣势在于庞大的运算开销,实时性较差。Fan 等^[12]利用双源卷积网络(dual-source deep neural network, DS-CNN)方法构建了关节点检测网络和关节点定位回归网络,处理整幅图像时会因大量背景特征影响人体姿态特征的处理,故只适合简单背景的人体姿态识别。唐新宇等^[13]提出了一种基于 OpenPose 和 Kinect 的三维人体姿态估计方法,再通过霍特双参数指数平滑法对关节点运动轨迹进行平滑和预测,并成功应用到康复机器人领域,缺点是识别遮挡关节效果较差。以上提到的有关人体关节识别的方法都存在一个共性问题,即对于关节遮挡情况的适应性不强。

本文采用深度学习的方法从彩色图像中估计人体关节点像素 2D 坐标,将彩色图像与深度图像对齐,结合坐标变换计算人体关节 3D 坐标。利用 Vicon 测量关节坐标得到人体骨架信息来构建人体模型,通过约束条件的设定,求解相邻关节间的方向向量,来解决关节遮挡或者关节误识别的情况,提高被遮挡关节的识别精度。

1 移乘服务机器人功能简介

图 1 所示为本文自主研发的移乘服务机器人,此款机器人主要通过两种方式对被护理人实现移乘功能,即起到搀扶作用的腋下托举方式(图 1(a))和将人横向抱起的“公主抱”方式(图 1(b))。

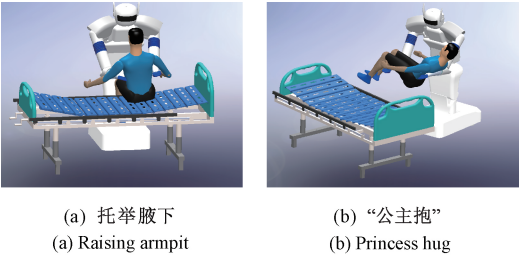


图 1 移乘服务机器人
Fig.1 Transfer service robot

该机器人工作过程为:首先利用声源定位模块识别被护理人的语言指令并进行判断,计算出被护理人相对于机器人的具体方位后调整方向使机器人朝向被护理人;而后利用导航模块规划合理路径并自主避障运动到距离被护理人较近的适宜地方;进一步利用视觉识别模块辨识人体各个关节的三维坐标,根据动作方式需要确定相应的人体关节位置;最后,合理规划两个机械臂动作,在保证安全性的前提下将被护理人“托举”或者“抱”起来。

2 人体 3D 关节识别

2.1 OpenPose 人体 2D 关节识别

OpenPose 网络结构人体 2D 关节识别方法^[14]由卡内基·梅隆大学感知计算实验室于 2017 年提出,该方法是基于大数据集,利用深度学习的方法从彩色图中估计人体部位的 2D 关节点,然后将这些关节点进行连接得到人体姿态骨架,具有高精度、高鲁棒性等特点。本文将人体关节定义为 14 个关节,如图 2 所示。

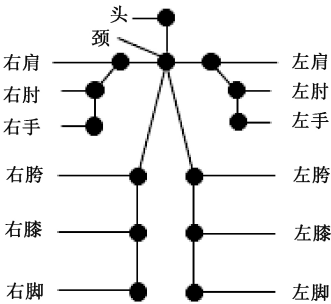


图 2 关节分布
Fig.2 Joint distribution

基于 OpenPose 的人体 2D 关节识别整体流程如图 3 所示。

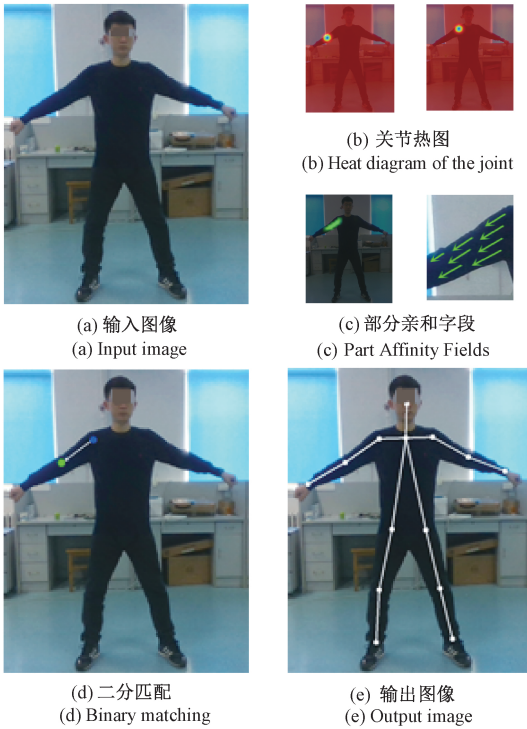


图 3 整体流程
Fig.3 Overall process

首先,利用 RealSense D435 相机获取受试者彩色图像和深度图像。如图 3(a)所示,输入图像为 $W \times H$ 大小的彩色图像,采用 VGG-19 网络对输入的彩色图像进行分类处理,得到特征图像(feature map);将特征图像输入到一个双分支多级卷积神经网络中,第 1 个支路中的每一级预测关节热图,生成关节热图 $S, S = (S_0, S_1, \dots, S_i, \dots, S_{13}), i$ 为关节标号 $i \in (0, 1, \dots, 13), S_i \in \mathbf{R}^{W \times H}$ (W 和 H 分别为彩色图像的宽度和高度)如图 3(b)所示,令 $p_i = \operatorname{argmax} S_i(u, v)$,其中 u, v 表示关节热图中的像素坐标, p_i 代表的是关节热图中得分最高的像素坐标(得分越高表示关节预测越准确),即预测的关节 i 的像素坐标;第 2 个分支中的每一级预测部分亲和字段(part affinity fields, PAFs),部分亲和片段用来检测肢体的走向,表达了人体相邻关节之间的关联程度,如图 3(c)所示,得到一组身体部位的 2D 矢量场 $L, L = (L_1, L_2, \dots, L_j, \dots, L_c), j \in (0, 1, \dots, c)$,具有 c 个矢量场,每个矢量场代表一个肢体;得到关节热图和部分亲和字段之后,使用二分匹配将同一人的相邻关节点进行连接,如图 3(d)所示;图 3(e)所示作为网络最终的输出图像,将人体的关节点的 2D 像素坐标输出显示。

为了增强该人体姿态识别算法在家庭环境下的鲁棒性,分别从网上下载并结合个人采集获取的包含被护理

人在内的家庭环境图像,并对这些图像采取旋转以及向内向外缩放的方式,将家庭环境数据集进行扩充,而后将文献[14]中已有的权重作为初始权重,用该数据集对文献[14]的网络进行迁移学习。

该方法只能识别出人体的 2D 关节像素坐标,移乘服务机器人若想实现“抱人”或者是“搀扶”功能,尚缺少相机与被护理人之间的距离信息,即深度信息,故需要将 2D 关节像素坐标转换成 3D 全局坐标。

2.2 人体 3D 关节计算

1) 深度图预处理

深度值的测量来源于深度图,因原始深度图中存在椒盐噪声以及深度无效值的空洞,需对深度图进行预处理。对椒盐噪声进行二维中值滤波。

$$M'(u, v) = \operatorname{med}\{M(u - k, v - l), (k, l \in W)\} \quad (1)$$

式中: $M(u, v)$ 为原始深度图像; $M'(u, v)$ 为滤波后深度图像; W 为滤波器的模板,尺寸大小为 3×3 ; k 和 l 为模板 W 中的数值。

通过最近邻插值算法对深度图中的空洞进行修复,其核心是将空洞处的深度值与其邻域的相邻点深度值进行对比,选择附近邻域中值最接近的非零点深度值作为所求点的深度值。深度图修复前后的对比如图 4 所示。

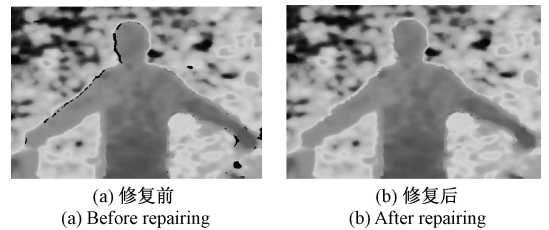


图 4 深度图修复

Fig.4 Repairing of the depth image

2) 彩色图像与深度图像对齐

由于彩色相机与两个红外相机构成的深度相机存在位置偏差,由 RealSense D435 相机获取的彩色图像和深度图像中同一像素点的坐标并不相同,因此,需要对彩色图像与深度图像进行对齐,步骤如下。

(1) 根据张氏标定法,对 RealSenseD435 的 RGB 相机和两个红外相机进行内参和外参的标定。得到 RGB 相机的内参矩阵为 M_{RGB} ,两个红外相机的内参矩阵分别为 M_L 和 M_R ,左侧红外相机坐标系与 RGB 相机坐标系的旋转矩阵为 R ,平移向量为 T 。内参矩阵为:

$$M_{\text{RGB}} = \begin{bmatrix} f_{x_{\text{RGB}}} & 0 & c_{x_{\text{RGB}}} \\ 0 & f_{y_{\text{RGB}}} & c_{y_{\text{RGB}}} \\ 0 & 0 & 1 \end{bmatrix},$$

$$\mathbf{M}_L = \begin{bmatrix} f_{xL} & 0 & c_{xL} \\ 0 & f_{yL} & c_{yL} \\ 0 & 0 & 1 \end{bmatrix}$$

其中, f_x 和 f_y 为相机的焦距; c_x 和 c_y 为相机光轴在图像坐标系中的偏移量。

(2) 在得到的深度图像的图像坐标系下有一点 P_L , 其像素坐标为 (u_L, v_L) , 深度值为 $Dep(u_L, v_L)$, 根据式(2)利用左侧红外相机的内参矩阵将深度图像中的二维像素坐标 P_L 转换成左侧红外相机坐标系下的三维坐标 $P'_L, P'_L = (X_L, Y_L, Z_L)$, 其中 $Z_L = Dep(u_L, v_L)$ 。

$$\begin{bmatrix} X_L \\ Y_L \\ Z_L \end{bmatrix} = Z_L \times \mathbf{M}_L^{-1} \times \begin{bmatrix} u_L \\ v_L \\ 1 \end{bmatrix} \quad (2)$$

(3) 接下来利用式(3)通过左侧红外相机与 RGB 相机之间的旋转矩阵 $\mathbf{R}$ 和平移向量 $\mathbf{T}$, 计算出 P'_L 在 RGB 相机坐标系下的坐标 P_{RGB} 。

$$P_{RGB} = \mathbf{R} \times P'_L + \mathbf{T} \quad (3)$$

在同一场景下, 分别得到棋盘格相对于左侧红外相机和 RGB 相机的旋转矩阵和平移向量, 分别为 $\mathbf{R}_{IR}, \mathbf{R}_{RGB}, \mathbf{T}_{IR}, \mathbf{T}_{RGB}$ 。

$$\begin{cases} P'_L = \mathbf{R}_{IR} \times P + \mathbf{T}_{IR} \\ P_{RGB} = \mathbf{R}_{RGB} \times P + \mathbf{T}_{RGB} \end{cases} \quad (4)$$

对式(4)中的两个公式进行消除 P 的操作可得:

$$P_{RGB} = \mathbf{R}_{RGB} \times \mathbf{R}_{IR}^{-1} \times P'_L + \mathbf{T}_{RGB} - \mathbf{R}_{RGB} \times \mathbf{R}_{IR}^{-1} \times \mathbf{T}_{IR} \quad (5)$$

联立式(3)和(5), 求出旋转矩阵 $\mathbf{R}$ 和平移向量 $\mathbf{T}$ 。

$$\begin{cases} \mathbf{R} = \mathbf{R}_{RGB} \times \mathbf{R}_{IR}^{-1} \\ \mathbf{T} = \mathbf{T}_{RGB} - \mathbf{R} \times \mathbf{T}_{IR} \end{cases}$$

将求出的旋转矩阵 $\mathbf{R}$ 和平移向量 $\mathbf{T}$ 代入式(3), 求出 P_{RGB} 的坐标为 $(X_{RGB}, Y_{RGB}, Z_{RGB})$ 。

(4) 进一步利用式(6)通过 RGB 相机的内参矩阵求出 P_{RGB} 投影在 RGB 图像上的二维像素坐标 $P_{rgb}, P_{rgb} = (u_{rgb}, v_{rgb})$ 。

$$\begin{bmatrix} u_{rgb} \\ v_{rgb} \\ 1 \end{bmatrix} = \frac{\mathbf{M}_{RGB}}{Z_{RGB}} \times \begin{bmatrix} X_{RGB} \\ Y_{RGB} \\ Z_{RGB} \end{bmatrix} \quad (6)$$

得出 (u_L, v_L) 与 (u_{rgb}, v_{rgb}) 的对应关系, 对深度图像中的每一个像素点重复上述操作, 直到遍历完整幅深度图像, 将彩色图像与深度图像对齐之后, 每个关节点的坐标形式为 $(X_{RGB}, Y_{RGB}, Z_{RGB})$, 如图 5(e) 所示。

流程如图 5 所示, RGB-D 相机同时获取受试者的彩色图像(图 5(b)) 和深度图像(图 5(c)), 将彩色图像通过 OpenPose 算法得到关节点 2D 像素坐标(图 5(d)), 彩色图像和深度图像对齐后, 结合坐标变换并引入深度值将 2D 像素坐标转换为 3D 全局坐标。

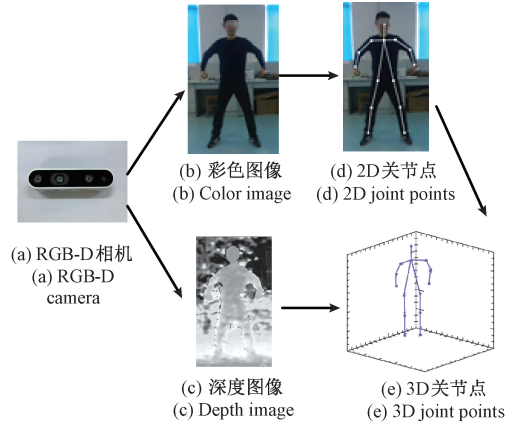


图 5 彩色图像与深度图像对齐

Fig.5 Colour image aligned with depth image

3 基于模型约束的人体姿态识别算法

3.1 人体模型的测量

家庭环境中存在相机视野无法完全捕捉人体所有关节的可能性, 即关节遮挡现象(自我遮挡或它物遮挡), 此时基于 OpenPose 算法的人体 2D 关节识别方法精度不足且极易出现关节误识别和位置波动, 如图 6 所示站姿时的右手腕关节(图 6(a)) 和卧姿时的左膝关节(图 6(b)), 为了提高关节估计的准确性, 本文通过引入人体模型作为姿态约束的方式来解决手臂、腿部关节的遮挡问题和误识别问题。

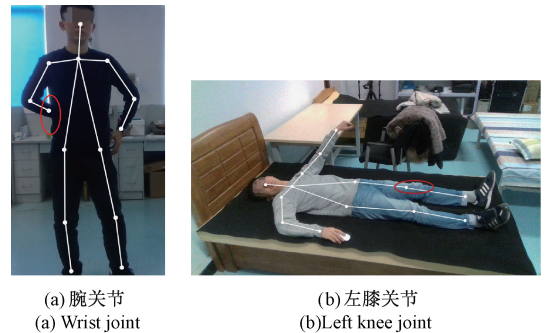


图 6 关节误识别与位置波动

Fig.6 Joint misidentification and position fluctuation

人体模型指的是人体各个相邻关节之间的长度信息(如肘、腕关节之间的长度, 膝、胯关节之间的长度), 人体模型的测量参照《中国成年人人体尺寸(GB/T 10000-1988)》中不同百分位数人体主要关节尺寸和身高的关系, 根据这一标准计算各相邻关节间的长度相对于身高的比例, 如表 1 和 2 所示, 使得输入受试者的身高, 判断输入身高与哪个百分位数对应身高更近, 就依据对应比

例关系得出受试者的相邻关节间的长度信息,以此作为人体模型参数。

表 1 人体主要尺寸(18~60 岁)
Table 1 Human main dimensions (18~60 years old) mm

性别	参数	百分位数						
		1	5	10	50	90	95	99
男	身高	1 543	1 583	1 604	1 678	1 754	1 775	1 814
	上臂长	279	289	294	313	333	338	349
	前臂长	206	216	220	237	253	258	268
	大腿长	413	428	436	465	496	505	523
	小腿长	324	338	344	369	396	403	419
女	身高	1 449	1 484	1 503	1 570	1 640	1 659	1 697
	上臂长	252	262	267	284	303	308	319
	前臂长	185	193	198	213	229	234	242
	大腿长	387	402	410	438	467	476	494
	小腿长	300	313	319	344	370	376	390

表 2 关节长度与身高的比值
Table 2 The ratio of joint length to height

性别	参数	百分位数						
		1	5	10	50	90	95	99
男	上臂长/身高	0. 181	0. 183	0. 183	0. 187	0. 190	0. 190	0. 192
	前臂长/身高	0. 134	0. 136	0. 137	0. 141	0. 144	0. 145	0. 148
	大腿长/身高	0. 268	0. 270	0. 272	0. 277	0. 283	0. 285	0. 288
	小腿长/身高	0. 210	0. 214	0. 214	0. 220	0. 226	0. 227	0. 231
女	上臂长/身高	0. 174	0. 177	0. 178	0. 181	0. 185	0. 186	0. 188
	前臂长/身高	0. 128	0. 130	0. 132	0. 136	0. 140	0. 141	0. 143
	大腿长/身高	0. 267	0. 271	0. 273	0. 279	0. 285	0. 287	0. 291
	小腿长/身高	0. 207	0. 211	0. 212	0. 219	0. 226	0. 227	0. 230

为了验证应用此方法得到的人体模型参数的可靠性,用动作捕捉器(motion capture)对 10 名不同身高的受试者逐个进行试验,按照图 3 中关节分布,用反光小球标记受试者各个关节,测量其关节的三维坐标,依次将这些三维关节点按照图 3 顺序进行连线,得到人体的骨架线图,如图 7 所示,并根据各个关节的三维坐标,计算相邻关节之间的长度,与相应人体模型参数进行对比,得出的结论为:用 motion capture 计算得到的关节间长度与人体模型参数的对应值相差不大,偏差在合理范围内。所以依据《中国成年人人体尺寸》标准中的身高与主要关节尺寸的关系,通过受试者身高计算相邻关节间的长度作为人体模型参数是可行的。

3.2 站姿、坐姿下被遮挡关节的识别

以站姿下被护理人右手臂的腕关节被遮挡为例,如

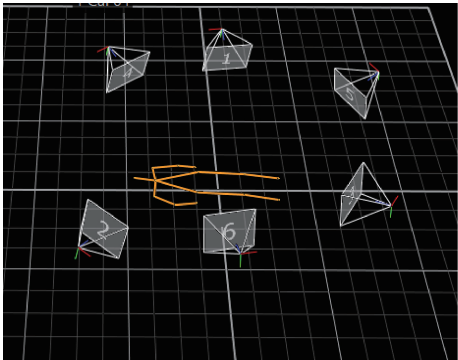


图 7 motion capture 测量结果
Fig.7 The measurement result with motion capture

图 8 所示,此类情况极为常见,关节误识别会造成机器人

进行腋下托举动作时机机械臂与被护理人右小臂的碰撞,从而影响安全性。图8(a)所示被护理人右手臂的腕关节出现自我遮挡,在相机的视野中不可见,但与腕关节相邻的肘关节可见,因此,通过彩色图像准确识别肘关节,将彩色图像与深度图像对齐后计算其三维坐标,由人体模型可知腕关节与肘关节间的长度信息,再求出腕关节与肘关节之间的方向,就可以计算腕关节的正确位置,流程如下。

1) 右手臂肘关节可见,令识别出的肘关节的像素坐标为 $(u_{\text{肘}}, v_{\text{肘}})$,根据内参矩阵计算出肘关节的三维坐标为 $(X_{\text{肘}}, Y_{\text{肘}}, Z_{\text{肘}})$;

2) 以肘关节的像素坐标为中心作边长为 L 个像素的矩形ROI(region of interest),其中: L 为颈关节与右肘关节之间的像素距离成正比,同时还要保证 $u_{\text{肘}}(v_{\text{肘}})$ 距离原始图像左侧(上侧)的像素距离和距离右侧(下侧)的像素距离均大于 L 。将此ROI剪切为一幅 $L \times L$ 的图像,如图8(b)所示;

3) 对此 $L \times L$ 的彩色图像进行灰度化处理,然后采用 3×3 的内核模糊化降噪,使用Canny算子对灰度图像提取轮廓,将提取的轮廓点像素坐标以数组的形式存储备用;

4) 以ROI中心像素坐标为起点,向左搜索图像,直至找到轮廓点,定义此轮廓点作为新的起点,用虫随法自上向下搜寻其8邻域,得到手臂的整个下轮廓,如图8(c)所示,将该小臂下轮廓视为圆弧和直线,采用文献[15]中的K-余弦曲率算法识别“直线-圆弧-直线”的切点的方法,找到直线与圆弧的切点 Q ,如图8(d)所示,将切点 Q 与小臂轮廓末端点 M 之间的轮廓点像素坐标存入新的数组备用;

5) 将步骤4)中得到的系列轮廓点 $((u_1, v_1), (u_2, v_2), \dots, (u_n, v_n))$, n 为轮廓点总数,通过内参矩阵计算得到三维坐标点集 $((X_1, Y_1, Z_1), (X_2, Y_2, Z_2), \dots, (X_n, Y_n, Z_n))$ 。

6) 由于深度图像的深度值的波动,导致轮廓点深度值 Z_n 存在一定的误差,在利用最小二乘法对轮廓点进行拟合时,需要所有的三维点的参与,运算量大且拟合直线存在明显误差,因此,使用随机采样一致性(random sample consensus, RANSAC)方法将深度值异常的“局外点”去除,算法过程为:随机从步骤5)得到的三维坐标点集中抽出两个点作为样本数据,计算两个点间的方向向量为 λ ,并以此构建直线方程,然后,以0.05 m为距离阈值,对剩余 $n-2$ 个点进行测试,测定各点与此直线方程的距离,小于该阈值的点作为内点,大于此阈值的点作为外点,对其进行舍弃,整个过程为迭代一次。迭代次数为:

$$K = \frac{\log(1-p)}{\log(1-\omega^m)} \quad (7)$$

式中: p 为迭代过程中从数据集内随机选取出的点都为“内点”的概率; ω 为“内点”占数据集总数的比例; m 为计算模型所需的最少样本数,故选 $m=2$;

7) 取迭代过程中,“内点”最多的一组为所需轮廓点,用最小二乘法将这些“内点”拟合为一条空间直线,并计算单位方向向量。

8) 计算右腕关节三维坐标:

$$\begin{bmatrix} X_{\text{右腕}} \\ Y_{\text{右腕}} \\ Z_{\text{右腕}} \end{bmatrix} = \begin{bmatrix} X_{\text{右肘}} \\ Y_{\text{右肘}} \\ Z_{\text{右肘}} \end{bmatrix} + L_{\text{右肘-右腕}} \begin{bmatrix} \alpha_{\text{右小臂}} \\ \beta_{\text{右小臂}} \\ \gamma_{\text{右小臂}} \end{bmatrix} \quad (8)$$

式中: $[\alpha_{\text{右小臂}}, \beta_{\text{右小臂}}, \gamma_{\text{右小臂}}]^T$ 为右肘关节到右腕关节的单位方向向量, $L_{\text{右肘-右腕}}$ 为通过人体模型测量的右肘关节到右腕关节的距离。然后,利用内参矩阵求解右腕关节在二维图像中的像素坐标 $[u_{\text{右腕}}, v_{\text{右腕}}]^T$,图8(e)所示为右腕关节在被遮挡情况下的识别结果。

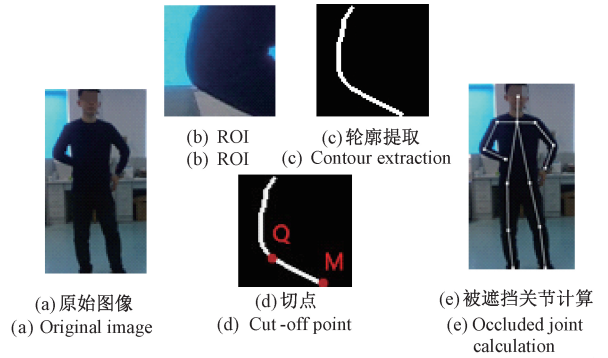


图8 站姿下被遮挡关节的识别

Fig.8 Recognition of occluded joint in standing position

3.3 卧姿下被遮挡关节的识别

在卧姿下,通过膝关节的三维坐标以及膝关节与膝关节的相连关系,计算膝关节的三维坐标,然后通过内参矩阵计算出膝关节在图像中的像素坐标。如图9所示,以左腿为例,依据颈关节与两膝关节中点的像素距离 J 确定一个 L' 像素长度, $L'=4J/5$,设两个膝关节中点像素坐标为 $(u_{\text{膝中}}, v_{\text{膝中}})$,以像素坐标点 $(u_{\text{膝中}} - L'/4, v_{\text{膝中}} - L'/4)$ 为左上角裁出边长为 L' 个像素长度的矩形ROI,如图9(b)所示;提取左腿的上轮廓,将轮廓点存入数组备用;求取上述轮廓点对应的深度值,通过内参矩阵得到上述轮廓点的三维坐标点集,如图9(c)所示;使用RANSAC算法剔除掉深度值误差较大的“外点”,用剩余三维坐标点集通过直线拟合计算三维直线的单位方向向量,通过式(9)得到左膝关节的三维坐标为:

$$\begin{bmatrix} X_{\text{左膝}} \\ Y_{\text{左膝}} \\ Z_{\text{左膝}} \end{bmatrix} = \begin{bmatrix} X_{\text{左髋}} \\ Y_{\text{左髋}} \\ Z_{\text{左髋}} \end{bmatrix} + L_{\text{左髋-左膝}} \begin{bmatrix} \alpha_{\text{左腿}} \\ \beta_{\text{左腿}} \\ \gamma_{\text{左腿}} \end{bmatrix} \quad (9)$$

式中: $[\alpha_{\text{左腿}}, \beta_{\text{左腿}}, \gamma_{\text{左腿}}]^T$ 为左膝关节到左膝关节的单位方向向量; $L_{\text{左膝-左膝}}$ 为人体模型中左膝关节到左膝关节距离。同理,得到右膝关节三维坐标,通过内参矩阵计算膝关节的像素坐标并标记到图像中,如图 9(d)所示。当出现右膝对左膝的遮挡时,左膝位置识别结果如图 9(e)所示。

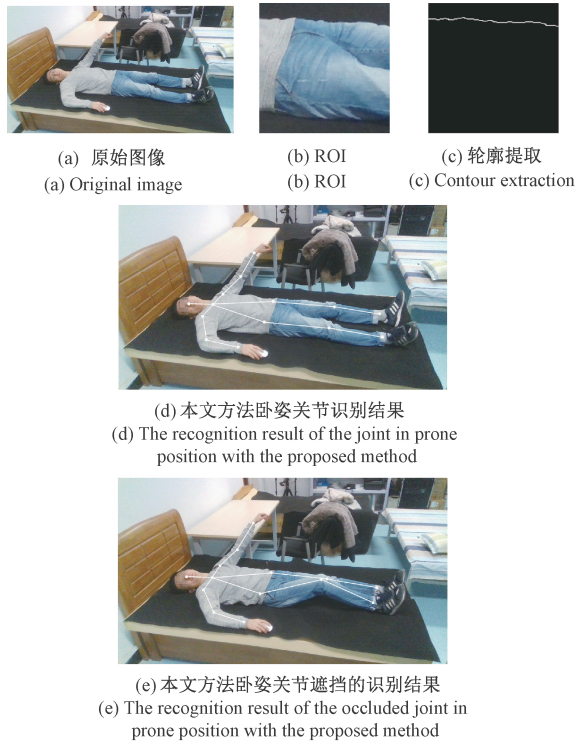


图 9 卧姿下的关节识别情况
Fig.9 Joint recognition in prone position

3.4 腋下点估计

为了完成“托举腋下”动作,需要提取人体腋下区域,进而确定适宜移乘机器人操作臂插入的腋下点。如图 10(a)所示,以右肩和右肘为顶点作一条直线,并计算此两点间的像素距离 $p_{\text{右肩-右肘}}$,然后分别过此两点垂直于该直线作两条直线,这两条直线方向由肩关节指向人体胸部,长度设定为 $p_{\text{右肩-右肘}}$,得到另外两个点 A 和 B,求右肩、右肘、A 点和 B 点组成的外接矩形。通过此方式得到的外接矩形通常不是水平的,需将得到的外接矩形旋转调整为水平方向,才能以矩阵的形式表示为图像,将调整为水平方向的外接矩形作为人体右侧腋下 ROI,如图 10(b)所示。同理,得出左侧腋下 ROI,如图 10(c)所示。当人体的手臂张开角度在 $10^\circ \sim 55^\circ$ 时,这两个 ROI 区域可以很好地展现出人体的腋下区域。

收缩 ROI 区域后剔除其边缘像素,避免将其误计算为腋下轮廓点,利用 Canny 算子在 ROI 中提取腋下轮廓

点,然后,将轮廓点拟合,连接起始点和终止点得到封闭曲线,即为腋下区域,取其中心点计为腋下点,如图 10(d)所示。记录腋下点在原图中的像素坐标,其深度值由所截取的 ROI 4 个顶点的深度平均值代替。

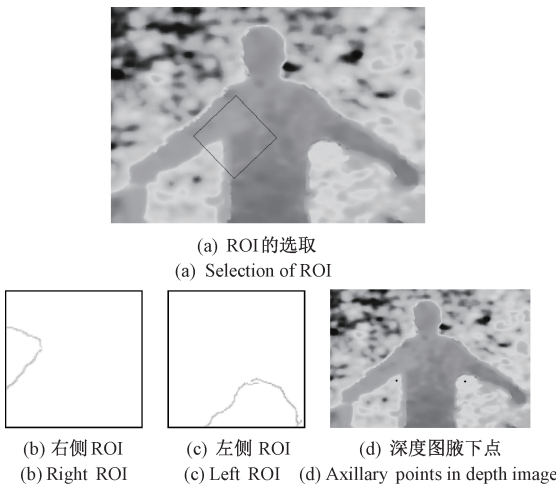


图 10 站姿下被遮挡关节的识别
Fig.10 Recognition of occluded joints in standing position

4 实验与结果分析

4.1 实验环境

在家庭环境下设计完成 3 个实验。实验 1 主要识别卧姿下人体的肩关节、膝关节;实验 2 主要识别站姿或者坐姿下人体的肩关节、肘关节以及腕关节,并提取腋下区域,进而确定一个合适的腋下点;实验 3 主要完成对被遮挡关节的识别。本文的实验距离,即受试者与相机之间的距离为 1.5~2 m。

4.2 实验结果精度评估

关节识别准确度的评估采用 PCKh^[16] 标准,即当预测的关节点在真实的关节点附近一定距离时,就判断该关节点识别准确。本文将这个距离设定为 3 cm,作为关节识别准确度的评价标准,即当预测的关节点落在以真实的关节点为球心、3 cm 为半径的球体内就视为识别准确。对 RealSense 相机和 motion capture 进行标定,得到二者之间的旋转矩阵和平移向量,以 motion capture 获得的关节三维坐标作为理论坐标,把应用本文姿态识别算法获得的关节三维坐标转换到 motion capture 的坐标系下作为实际坐标,求实际坐标和理论坐标之间的距离,与本文设定的 3 cm 距离进行比较,当求得的偏离距离 ≤ 3 cm 表示预测的关节点识别准确,否则就表示识别不准确。

在相同的家庭环境中,以卧姿和站姿(坐姿)下拍摄 300 帧图像作为测试数据。分别对 Kinect SDK、文

献[17]方法以及本文方法的关节识别准确度进行对比,结果如表 3 和 4 所示。

表 3 站姿、坐姿下关节识别准确度比较
Table 3 Comparison of joint recognition accuracy in standing and sitting positions

方法	Kinect SDK	文献[17]	本文方法
肩(左 & 右)	0.91	0.93	0.93
肘(左 & 右)	0.87	0.90	0.90
胯(左 & 右)	0.87	0.89	0.91
腕(遮挡)	0.21	0.35	0.90

表 4 卧姿下关节识别准确度比较
Table 4 Comparison of joint recognition accuracy in prone position

方法	Kinect SDK	文献[17]	本文方法
肩(左 & 右)	0.91	0.93	0.93
膝(左 & 右)	0.71	0.75	0.92
膝(遮挡)	0.22	0.39	0.91

由表 3 和 4 可知,应用 Kinect SDK 来估计家庭环境下的人体关节时,3 种姿态的关节识别准确度均较差,出于安全性的考量,此方法在移乘服务机器人对人体姿态识别方面尚存在局限性;人体关节未出现遮挡时,文献[17]的关节识别准确度均优于 Kinect SDK,但是当人体关节出现遮挡时,其识别能力明显不足。本文方法对家庭环境的适应性最好,尤其是对被遮挡关节和卧姿下膝关节的识别均明显优于另外两种方法。

在左侧膝关节被右侧膝关节遮挡的情况下,应用文献[17]和本文方法得到的结果对比,如图 11 所示,其中文献[17]方法错误地将左膝关节识别为右膝关节。

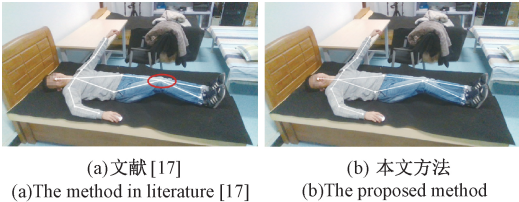


图 11 膝关节识别结果对比
Fig.11 Comparison of knee joint recognition results

在关节遮挡和关节未遮挡情况下分别采集 300 份含有受试者的图像,主要计算肩、肘、腕以及膝关节的实际坐标与其理论坐标之间的偏离距离,与本文人体姿态识

别的最大误差 3 cm 作比较。经计算,在关节(肩、肘、膝)未出现遮挡时,本文的关节识别准确度为 92%,如图 12 所示;当关节(腕、膝)出现遮挡时,关节识别准确度亦可达到 90%,如图 13 所示。

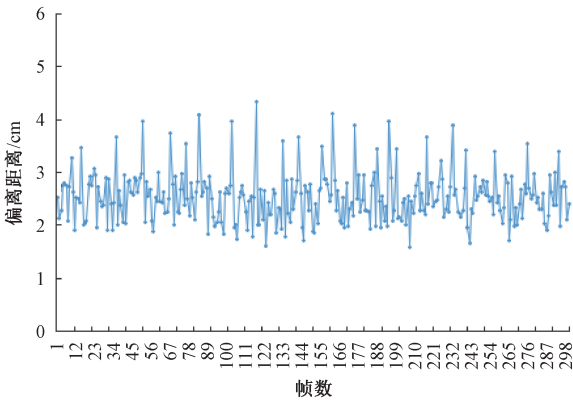


图 12 未遮挡关节实际坐标与理论坐标之间的距离
Fig.12 The distance between the actual and theoretical coordinates of the unoccluded joint

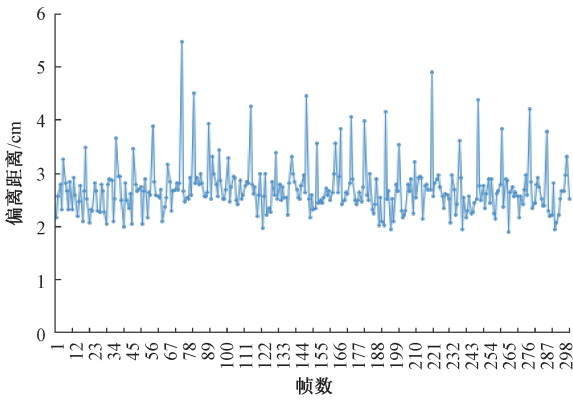


图 13 遮挡关节实际坐标与理论坐标之间的距离
Fig.13 The distance between the actual and theoretical coordinates of the occluded joint

对于遮挡关节的估计,误差主要来源于以下两点:1)对于遮挡关节的识别取决于与该遮挡关节相连接的未遮挡关节的识别精度,人体的不同姿势、人体关节与相机之间的不同角度都会影响未遮挡关节的识别精度,间接影响遮挡关节坐标的计算;2)在对遮挡关节所属肢体方向向量的求取存在较大误差时,会影响遮挡关节坐标的求取,如衣服的松紧程度、人体的体型(胖瘦程度)以及相邻肢体间过小的夹角会使衣服出现堆叠的现象,这些都会影响肢体轮廓的正常提取,从而影响肢体方向向量的计算。在图 13 中有偏差距离大于 5 cm 的情况,原因是未遮挡关节的估计出现了误识别,导致对遮挡关节的求取出现了较大偏差。

5 结 论

针对 OpenPose 算法在关节遮挡情况下出现误识别及关节波动的现象,本文提出了一种基于模型约束的人体姿态视觉识别算法并将其应用于移乘服务机器人系统用于精确获取人体关节的空间坐标。本文将彩色图像与深度图像对齐处理,从彩色图中识别人体的二维关节像素坐标,而后结合坐标变换方法计算相机坐标系中人体关节三维坐标。依据肩关节和肘关节提取腋下区域,从而找出合理的腋下点,实验表明,腋下点的识别精度达到了 90%。另外,依据人体模型相关参数以及未被遮挡关节坐标计算与之相连接的被遮挡关节的三维坐标。经实验,本文方法对关节的识别准确度,在未遮挡时达到了 92%;在被遮挡情况下达到了 90%。

参考文献

- [1] 王国彪,陈殿生,陈科位,等. 仿生机器人研究现状与发展趋势[J]. 机械工程学报,2015,51(13):27-44.
WANG G B, CHEN D SH, CHEN K W, et al. The current research status and development strategy on biomimetic robot[J]. Journal of Mechanical Engineering, 2015, 51(13): 27-44.
- [2] KOJIMA K, SATO T, SCHMITZ A, et al. Sensor prediction and grasp stability evaluation for in-hand manipulation [C]. IEEE /RSJ International Conference on Intelligent Robots and Systems, 2013: 2479-2484.
- [3] 王亚宾,张小栋,穆小奇,等. 用于助老伴行机器人的老年人摔倒预测方法研究[J]. 电子测量与仪器学报, 2018,32(7):1-7.
WANG Y B, ZHANG X D, MU X Q, et al. Research on falling prediction methods for elderly companion robots[J]. Journal of Electronic Measurement and Instrument, 2018;32(7):1-7.
- [4] 李斌. 单幅深度图人体关节点定位[D]. 武汉:华中科技大学,2012.
LI B. Key points of human body location based on single depth map[D]. Wuhan: Huazhong University of Science and Technology, 2012.
- [5] PISHCHULIN L, INSAFUTDINOV E, TANG S, et al. DeepCut: Joint subset partition and labeling for multi person pose estimation [C]. IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4929-4937.
- [6] 马森,李贻斌. 基于多级动态模型的 2 维人体姿态估计[J]. 机器人,2016,38(5):578-587.
MA M, LI Y B. 2D human pose estimation using multi-level dynamic model [J]. Robot, 2016, 38 (5): 578-587.
- [7] CARREIRA J, AGRAWAL P, FRAGKIADAKI K, et al. Human pose estimation with iterative error feedback[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4733-4742.
- [8] YANG W, OUYANG W, LI H, et al. End-to-end learning of deformable mixture of parts and deep convolutional neural networks for human pose estimation[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 3073-3082.
- [9] MARTINEZ J, HOSSAIN R, ROMERO J, et al. A simple yet effective baseline for 3D human pose estimation [C]. IEEE International Conference on Computer Vision, 2017: 2659-2668.
- [10] WEI S E, RAMAKRISHNA V, KANADE T, et al. Convolutional pose machines[C]. CVPR, IEEE, 2016: 4724-4732.
- [11] NEWELL A, YANG K, DENG J. Stacked hourglass networks for human pose estimation[C]. 14th European Conference on Computer Vision, Springer, 2016: 483-499.
- [12] FAN X C, ZHENG K, LIN Y W, et al. Combining local appearance and holistic view: Dual-source deep neural networks for human pose estimation [C]. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2015:1347-1348.
- [13] 唐心宇,宋爱国. 人体姿态估计及在康复训练情景交互中的应用 [J]. 仪器仪表学报, 2018, 39 (11): 198-206.
TANG X Y, SONG AI G. Human pose estimation and its application in scene interaction of rehabilitation training [J]. Chinese Journal of Scientific Instrument, 2018, 39 (11): 198-206.
- [14] CAO Z, SIMON T, WEI S E, et al. Realtime multi-person 2D pose estimation using part affinity fields[C]. IEEE Conference on Computer Vision and Pattern Recognition, 2017: 1302-1310.
- [15] 李洋. 基于视觉的数控刀具几何参数自动测量系统研究[D]. 天津:河北工业大学,2017:35-38.
LI Y. Research on vision-based automatic measurement system of NC tool geometric parameters [D]. Tianjin:

Hebei University of Technology, 2017;35-38.

- [16] HAQUE A, PENG B, LUO Z, et al. Towards viewpoint invariant 3D human pose estimation[C]. 14th European Conference on Computer Vision, Springer, 2016: 160-177.
- [17] 刘今越,李顺达,陈梦倩,等. 面向移乘搬运护理机器人的人体姿态视觉识别[J]. 机器人, 2019, 41(5): 601-608.
- LIU J Y, LI SH D, CHEN M Q, et al. Human pose visual recognition for transfer robots[J]. Robot, 2019, 41(5): 601-608.

作者简介



刘今越, 2000年和2005年分别于河北工业大学获得学士学位和硕士学位, 2008年于天津大学获得博士学位, 现为河北工业大学教授, 主要研究方向为智能检测与控制。

E-mail: ljy@ hebut.edu.cn

Liu Jinyue received his B. Sc. and M. Sc. degrees both from Hebei University of Technology in 2000 and 2005, respectively, and received Ph. D. degree from Tianjin University in 2008. He is now a professor at Hebei University of Technology. His main research interests include intelligent inspection and control.



刘彦开, 2017年于河南科技大学获得学士学位, 现为河北工业大学硕士研究生, 主要研究方向为计算机视觉。

E-mail: 609734496@qq.com

Liu Yankai received his B. Sc. degree in 2017 from Henan University of Science and Technology. Now, he is an M. Sc. candidate in Hebei University of Technology. His main research interests include computer vision.



贾晓辉(通信作者), 分别在2000年和2003年于河北工业大学获得学士学位和硕士学位, 2010年于天津大学获得博士学位, 现为河北工业大学讲师, 主要研究方向为力触觉感知、柔性精密定位系统设计。

E-mail: 2010081@ hebut.edu.cn

Jia Xiaohui (Corresponding author) received her B. Sc. and M. Sc. degrees both from Hebei University of Technology in 2000 and 2003, respectively, and received Ph. D. degree from Tianjin University in 2010. She is now a lecturer at Hebei University of Technology. Her main research interests includes force-tactile perception and compliant precision positioning system design.