

DOI:10.19651/j.cnki.emt.2314931

基于立体视觉与姿态识别的天轨机器人跟踪伺服控制

赵希禹 杨芳艳 李家龙

(上海理工大学机械工程学院 上海 200093)

摘要: 针对双足机器人步态训练平台自动化程度低,调试步态的过程中单人操作时辅助机器人行走和状态信息实时观测协调困难的问题,该文提出了一套基于立体视觉和姿态识别的天轨机器人跟踪伺服系统。首先利用双目相机取景,对左右目相机画面进行匹配后获取图像中每个像素点的深度信息,基于获取到的带有深度信息的图像,对双足机器人的姿态进行识别,获取每个关节的深度信息,根据关节深度信息判断双足机器人运动状态制定跟踪策略进行伺服控制,由于引入了姿态识别,可以根据双足机器人的姿态变化实现更程度的自动化跟踪保护。实验结果表现出高自动化程度和高动态的跟踪表现。

关键词: 立体视觉;姿态识别;跟踪伺服控制

中图分类号: TP391.7 **文献标识码:** A **国家标准学科分类代码:** 510.8050

Stereo vision and pose recognition-based tracking servo control for celestial orbital robot

Zhao Xiyu Yang Fangyan Li Jialong

(School of Mechanical Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China)

Abstract: To address the issues of low automation in the gait training platform for bipedal robots and the difficulty in coordinating real-time observation of robot walking and status information during gait debugging by a single operator, this paper proposes a stereo vision and pose recognition-based tracking servo system for celestial orbital robots. Firstly, using a stereo camera setup, the left and right camera images are matched to obtain depth information for each pixel. Based on the depth information obtained from the images, the pose of the bipedal robot is recognized, and the depth information of each joint is extracted. The motion state of the bipedal robot is determined based on the joint depth information, and a tracking strategy is formulated for servo control. By introducing pose recognition, a higher level of automation and tracking protection can be achieved based on the changes in the bipedal robot's pose. Experimental results demonstrate a high level of automation and dynamic tracking performance.

Keywords: stereo vision; pose recognition; tracking servo control

0 引言

人形机器人一直被机器人研究行业认为是机器人研究中的“皇冠”,也是人工智能领域发展的重要平台,而作为人形机器人,双足类人步态的研究也是机器人运动学中一个及其重要和复杂的领域^[1]。在研发双足步态过程中,机器人摔倒的情况是时常发生的,由于机器人双腿的结构、自重和电路板连带线路外置等因素,导致每一次摔倒都可能会造成控制电路板和线路的损坏,而一旦发生损坏,便会严重影响研发过程的损失成本和时间成本,所以一套完整的双足机器人辅助保护系统是十分有必要的。

目前该领域内缺少一套针对人形和双足机器人的精

准、完整的保护系统,传统的挂轨和移位器需要在调试训练期间实时有人在旁辅助,这便大大降低了调试人员的研发效率,给研究人员的独自调试带来了极大的不便,所以一套能实现自动跟踪定位的辅助系统便应运而生。

当前室内定位技术种类繁多,依赖于传感器的不同,大致可分为红外定位^[1]、蓝牙定位^[2]、超声波定位^[3]、无线网络定位^[4]、地磁导航定位^[5]和视觉定位^[6]、激光定位^[7]复杂环境下,机器视觉传感器凭借其强大的感知能力和图像处理技术所达到的高精度定位,引起了广泛关注^[8],使视觉定位方法备受瞩目。

目前,基于视觉的室内定位方法主要分为相对定位方法和绝对定位方法;相对定位方法借助于搭载在移动机器

人上的摄像头传感器,获取周边环境图像并进行处理,通过识别环境中的人工标记来实现机器人的定位。黄坤等^[9]设计了一种通过 YOLO 系列目标检测算法,对目标机器人身上颜色特征与形状特征相结合的人工信标进行检测匹配实现定位,再通过画面帧之间信标的像素移动,推算被跟踪主体的实际空间位移信息。该方法设计逻辑简单,定位精度较高,但易受空间大小和环境因素的限制,导致定位出现误差,当被跟踪主体失控,特征离开视野无法匹配,会导致跟踪系统的瘫痪和意外情况的发生。

相比之下,绝对定位方法使用相机传感器对目标物体及其周边环境进行图像采集,然后通过图像处理实现移动机器人的定位。这种方法能够有效减少机器人携带设备的复杂性,同时具备成本低、环境适应性强和算法简单等优势。田明锐等^[10]设计的大视场室内移动机器人高精度动态定位方法,利用单目相机,先采用多参量非线性标定方法提高定位静态定位精度,再用光信标简化图像处理算法提高定位速度实现对室内大范围下的移动机器人实时定位。该方法静态与动态定位精度高,但存在因采用单目相机多视角采样所导致的同名目标检测算法复杂,平面靶标标定过程工作量大的局限;

同时,在针对人形机器人的定位方法中,单一的空间定位缺乏表现力,无法全面反映人形机器人的运动姿态,而仅仅依赖于人形机器人关节传感器所反映的数据分析姿态存在数据传输延迟高和多关节角度误差大的问题。

针对以上问题和局限,本文提出了一种基于立体视觉和姿态识别的跟踪方法,移除搭载在天轨机器人上的摄像头,采用绝对定位方法的双目立体视觉来获取被跟踪主体在空间内的方位信息,并在此基础上利用机器视觉姿态识别算法来准确获取到被跟踪主体的姿态信息,便可以捕捉到其运动时的姿态,定制更精确的跟踪策略,在活动空间内不受特征点容易丢失的束缚,形成更强的实时保护,解放调试人员实现完全独立调试,给予了被跟踪主体在空间内更大的活动自由度以及完成更高难度动作的可能。

1 机器人控制系统总体设计

本文研究的是基于立体视觉和姿态识别用于辅助机器人双足行走的天轨机器人跟踪系统。图 1、2 为机器人的外形结构和框架模型图。包括三轴分立框架,X 轴为单轴挂载机器人本体,Y 轴为两侧并行齿槽导轨,Z 轴转盘升降机构,配备电子刹车防止电机卸载时负载掉落和大负载时电机持续工作过热。图 3 为本实验中采用的 Intel D435 深度相机^[11]。

在调试双足机器人的过程中,必须首先保证天轨机器人的鲁棒性和实时性,同时兼顾模块化和维护性。天轨机器人采用分布式体系和模块化设计,其主要组成包括主控模块、供电模块、信息采集模块和驱动模块,天轨机器人的系统框架如图 4 所示。

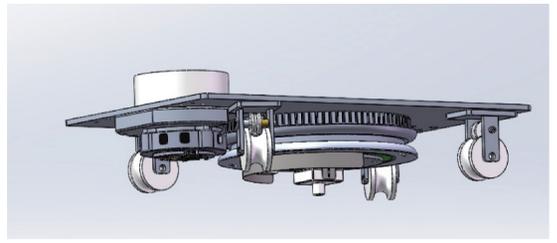


图 1 天轨机器人本体模型

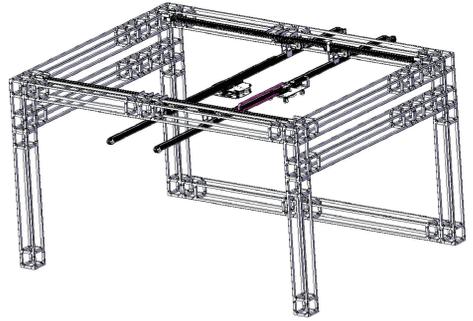


图 2 系统整体框架模型



图 3 Intel D435 深度相机

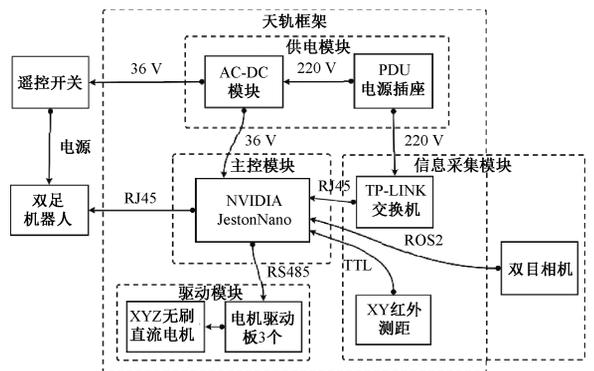


图 4 天轨机器人的系统框架

天轨机器人集成了三轴伺服电机驱动电路,完成了空间内 XYZ 全方位的移动控制,通过传感器在 XY 方向获取红外测距信息保证安全。主控模块采用具有强大的计算

性能和图形处理能力的小型计算机板 Jetson Nano^[12], 通过 USB 转 RS485 接口完成对天轨 XYZ 三轴的运动控制, 并通过局域网和 ROS2 (robot operating system 2, ROS2) 系统与第三人称角度的双目立体视觉图像采集进行数据传输。

本文研究基于立体视觉和姿态识别的伺服控制问题, 即利用双目相机获取到整个场景的深度图, 在深度图的基础上对被跟踪主体进行姿态识别获得各个关节点的深度信息, 将相机坐标系与天轨机器人的坐标系对齐, 获得准确的位置信息。跟踪过程中, 整个天轨机器人的控制过程如图 5 所示。

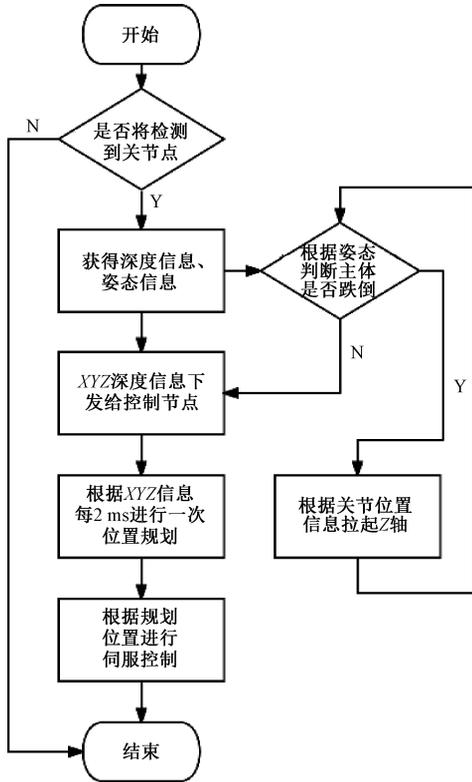


图5 天轨机器人的控制流程

2 立体视觉深度信息获取

2.1 双目相机成像原理

人眼识别现实世界中物体的位置时主要依靠双眼间存在一定距离, 面对同一物体时, 物体在两眼成像的位置有一定的差别, 这个差别, 称作为视差图像^[13]。

在双目相机中, 视差图像表现为两个相机图像之间像素位置的差异, 视差的存在是双目相机成像的基础, 如图 6 左目视图与图 7 右目视图对比产生的差异为视差。

图 8 为视差计算的原理图^[14], 其中: b 为基线, 即两个相机原点之间的距离; fl 和 fr 分别为左右相机的焦距, 理想的双目相机 $fl = fr = f$; cl 和 cr 分别为左右相机的光心, 理想的双目相机 $cl = cr$; xl 和 xr 分别为空间中的一点



图6 左目视图



图7 右目视图

P 投影在左右相机的图像中的横坐标。

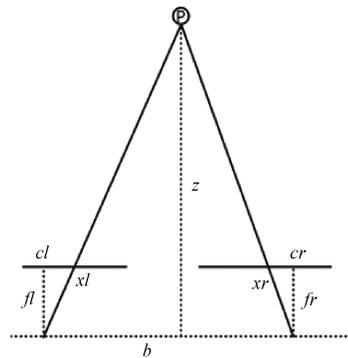


图8 视差原理

则根据几何关系, 可以得到:

$$\frac{b - (xl - cl) - (cr - xr)}{b} = \frac{z - f}{z} \quad (1)$$

整理得:

$$z = \frac{f \times b}{xr - xl} \quad (2)$$

右侧参数中相机焦距 f 、左右相机基线 b 为双目相机内固定参数, 若想求得 z , 需要知道左相机的每个像素点 (xl, yl) 和右相机中对应点 (xr, yr) 的对应关系, 这也是双目视觉的核心问题。

2.2 极线约束

如上所说, 只有知道 P 点在左右相机成像的像素差也

就是视差才能求得深度信息 z ，那么如何左右视图中匹配上 P 点，这里用到了极线约束^[15]。图 9 为极线约束原理图。

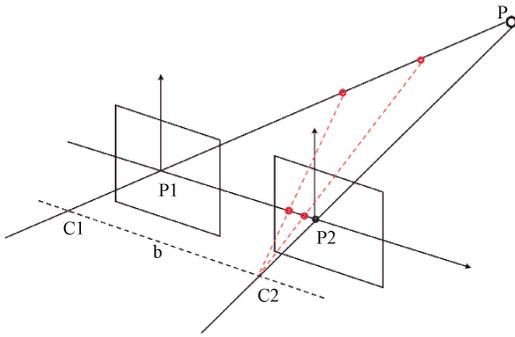


图 9 极线约束原理

极线约束是指当同一个空间点在两幅图像上分别成像时，已知左图投影点 P_1 ，那么对应右图投影点 P_2 一定在相对于 P_1 的极线上，这样可以极大的缩小匹配范围。根据极线约束的定义，我们可以在下图中直观的看到 P_2 一定在对极线上，所以只需要沿着极线搜索一定可以找到和 P_1 的对应点 P_2 。

2.3 立体匹配

极线约束的存在可以极大的缩小匹配范围，但在实际计算过程中也会产生误差累积左图的一个点，其在右图的对应点不一定恰好在极线上。但是应该是在极线附近，所以搜索范围需要适当放宽。

这里使用的立体匹配算法是一种区域匹配算法，常用的方法是基于块匹配(block matching)^[16-17]。该算法一般分为以下几个步骤：

- 1) 分割图像:将左视图和右视图图像分割成大小相同的块，通常是固定大小的正方形或矩形块^[17]。
- 2) 块匹配:对于左视图中的每个块，在右视图中搜索与之最匹配的块。匹配可以使用像素间的灰度差异或其他特征来衡量，图 10 为 640×480 图像的像素灰度值，式(3)、(4)为切割分配出的灰度块。常见的匹配度量方法包括：

(1) 差异绝对值和 (sum of absolute differences, SAD):计算块之间的像素差的绝对值的和。式(5)为对比计算后所得像素差绝对值。

(2) 差异平方和(sum of squared differences, SSD):计算块之间的像素差的平方和。

(3) 归一化互相关(normalized cross-correlation, NCC):计算块之间的互相关系数。

$$\begin{bmatrix} 47 & 87 & 90 & 71 & 38 & 25 & 24 & 23 & 23 \\ 37 & 63 & 77 & 75 & 44 & 26 & 24 & 23 & 23 \\ 30 & 46 & 57 & 58 & 39 & 26 & 25 & 24 & 24 \\ 34 & 60 & 79 & 68 & 42 & 27 & 26 & 25 & 24 \\ 50 & 80 & 98 & 87 & 48 & 28 & 26 & 25 & 25 \\ 63 & 85 & 90 & 95 & 52 & 28 & 26 & 25 & 25 \end{bmatrix} \quad (3)$$

1 Left Eye Gray Image:

2 Row 0:	19	19	19	19	19	19	19	19	19	19	19	20	21	23	31	49	58	58	57	59	60	59	59	
3 Row 1:	19	19	19	19	19	19	19	19	19	19	19	20	20	22	31	50	57	57	56	57	58	58	59	
4 Row 2:	19	19	19	19	19	19	19	19	19	19	19	20	20	21	28	44	50	49	48	52	55	57	58	
5 Row 3:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	23	30	32	31	30	30	35	42	49	54
6 Row 4:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	21	22	22	23	22	22	25	29	36	
7 Row 5:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	21	21	21	22	22	21	20	21	22	
8 Row 6:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	21	21	21	21	21	21	21	20	20	
9 Row 7:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	21	22	21	21	21	21	21	21	20	19
10 Row 8:	19	19	19	19	19	19	19	19	19	19	19	19	20	21	22	22	22	22	21	21	21	21	20	20
11 Row 9:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	21	22	22	22	22	21	21	21	20	20
12 Row 10:	19	19	19	19	19	19	19	19	19	19	19	19	20	20	21	22	22	22	22	21	21	20	20	19
13 Row 11:	20	19	19	19	19	19	20	19	19	19	20	20	21	22	22	22	22	22	21	21	20	20	19	
14 Row 12:	20	19	19	19	19	19	19	19	19	19	19	20	20	21	22	22	22	22	22	21	21	20	19	
15 Row 13:	19	19	19	19	19	19	19	19	19	19	19	20	20	21	22	22	22	22	22	21	21	20	20	
16 Row 14:	19	19	19	19	19	19	19	19	19	19	19	20	20	20	21	22	22	22	22	21	21	20	20	
17 Row 15:	19	19	19	19	19	19	19	19	19	19	19	20	20	20	21	22	22	22	22	21	21	20	20	
18 Row 16:	19	19	19	19	19	20	20	19	19	19	19	20	20	21	22	22	22	22	22	21	19	20	20	
19 Row 17:	19	19	19	19	19	19	19	19	19	19	19	20	21	21	22	22	22	22	21	21	20	20	20	
20 Row 18:	19	19	19	19	19	19	19	19	19	19	19	20	20	21	22	22	22	22	22	21	20	20	20	
21 Row 19:	20	19	19	19	19	19	19	19	19	20	19	19	20	20	21	22	22	22	22	21	20	20	19	
22 Row 20:	19	19	19	19	19	19	20	19	19	19	19	20	20	21	22	22	22	22	22	21	20	20	20	
23 Row 21:	19	19	19	19	19	19	19	19	19	20	19	20	20	21	22	22	22	22	22	21	20	20	19	
24 Row 22:	19	19	19	19	19	19	19	20	20	19	19	20	20	21	22	22	22	22	21	20	19	20	19	
25 Row 23:	19	19	19	19	19	19	19	19	20	20	19	19	20	20	21	22	22	22	22	21	20	19	20	

图 10 图像像素灰度值

$$\begin{bmatrix} 48 & 87 & 92 & 74 & 37 & 24 & 23 & 23 & 23 \\ 36 & 61 & 74 & 74 & 42 & 24 & 24 & 24 & 23 \\ 30 & 48 & 57 & 56 & 38 & 26 & 25 & 24 & 23 \\ 38 & 65 & 78 & 67 & 40 & 26 & 26 & 25 & 24 \\ 55 & 82 & 95 & 87 & 46 & 27 & 26 & 25 & 25 \\ 70 & 88 & 90 & 97 & 51 & 26 & 25 & 25 & 25 \end{bmatrix} \quad (4)$$

$$\begin{bmatrix} 1 & 0 & 2 & 3 & 1 & 1 & 1 & 0 & 0 \\ 1 & 2 & 3 & 1 & 2 & 2 & 0 & 1 & 0 \\ 0 & 2 & 0 & 2 & 1 & 0 & 0 & 0 & 1 \\ 4 & 5 & 1 & 1 & 2 & 1 & 0 & 0 & 0 \\ 5 & 2 & 3 & 0 & 2 & 1 & 0 & 0 & 0 \\ 7 & 3 & 0 & 2 & 1 & 2 & 1 & 0 & 0 \end{bmatrix} \quad (5)$$

3) 计算视差:对于每个左视图块，在右视图中找到最匹配的块后，计算它们之间的视差值。视差值表示了左视图块相对于右视图块的水平偏移量。

4) 生成深度图像:根据计算得到的视差值，可以根据摄像机的内参和基线距离等参数，将视差值转换为深度值，从而生成深度图像。

3 姿态深度信息识别

3.1 识别方法框架

人体姿态估计一直是计算机视觉领域的核心任务之一，姿态估计是指通过计算机视觉和图像处理技术，对人体或物体在二维或三维空间中的姿态进行推测或估计的过程，在人体的姿态估计中，主要关注的是对人体关节位置和角度进行估计，以推测人体的姿态和动作。本文研究的系统所服务的对象为人形机器人，与人体在身体结构和动作姿态上有高度的相似度，可以沿用人体的识别模型，但由于各类人形机器人在具体细节的结构设计与人体存在一定差异，模型的参数和权重需要重新训练。目前，该领域不存在通用的姿态识别模型，所以由此引出一套通用的训练方法。

实时人体姿态估计对人机交互、动作识别、运动分析和VTuber 技术等各种应用都很有吸引力。尽管在学术基准上取得了惊人的进展，但在计算能力有限的设备上稳健和实时的多人姿态估计仍然是一项具有挑战性的任务。目前主流的姿态识别平台主要包括以下几个:开源的姿态

估计库 OpenPose, Google 开发的跨平台的多媒体框架 MediaPipe, 微软研发的 Azure Kinect 提供的姿态识别应用 SDK (software development kit) 和开发工具, 被广泛使用的深度学习框架 PyTorch, YOLO 系列的 YOLOv7 姿态估计模型。

上述模型在公开数据集上获得了很出色的表现, 但是由于庞大的参数量及高延迟在工业领域还没有得到很好的应用。从 5 个方面实证研究了影响 2D 多人姿势估计框架性能和延迟的关键因素: Paradigm、骨干网络、定位方法、训练策略和部署。通过一系列优化, 本文引入 RTMPose (real-time multi-person pose)。

首先, RTMPose 采用了自上而下的方法, 即首先利用现有的检测器获取边界框, 然后独立估计每个人的姿势。尽管在过去, 由于额外的检测过程和不断增加的人群场景中的工作量, 检测部分曾是自上而下方法推理速度的瓶颈。然而, 由于实时检测器的卓越效率, 检测部分在大多数情况下 (每张图像 6 人以内) 不再成为速度瓶颈, 使得所提出的轻量级姿态估计网络能够实时执行所有实例的多次前向传递。

其次, RTMPose 选择了 CSPNeXt (cross stage partial network) 作为主干网络, 这是专门为对象检测而设计的。与为图像分类设计的主干网络相比, 后者对于密集预测任务 (例如对象检测、姿态估计和语义分割) 来说并不是最优的选择。一些主干网络虽然在公共姿态估计基准上取得了高精度, 但由于高计算成本、高推理延迟或部署困难而存在一些缺陷。CSPNeXt 在速度和准确性之间找到了良好的平衡, 并且易于部署。

第三, RTMPose 采用基于 SimCC (simple coordinate classification) 的算法来预测关键点, 该算法将关键点定位看作分类任务。相对于基于热图的算法, 基于 SimCC 的算法以更低的计算工作量实现了有竞争力的精度。此外, SimCC 采用了非常简单的预测架构, 只包含两个完全连接层, 使其易于在各种后端进行部署。

最后, 对以前的训练设置进行了回顾, 并根据经验提出了一组适用于姿态估计任务的训练策略。实验证明, 这组策略对于所提出的 RTMPose 姿势估计模型都带来显著的性能提升。

RTMPose 的网络结构图如图 11 所示, 该网络采用 CSPNeXt 作为主干网络用于学习图像底层特征, 然后经过 7×7 卷积层、全连接层和门控注意力单元, 结合软标签编码策略将关键点的横纵坐标分别转化为独立的一维向量表征。

3.2 模型训练数据集

数据集标记是模型训练的重要一步, 它直接影响模型的性能和准确性。本文采集实验室完整人形机器人图像数据集, 数据集共包括图像 1 350 幅。如图 12~14 所示, 对人体的数据采集主要包含 17 个关键点, 其中面部包含 5 个

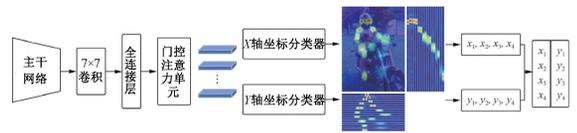


图 11 RTMPose 模型检测网络

关键点, 肩膀, 手肘, 手腕, 髋关节, 膝盖, 脚腕处各两个关键点, 由于实验选取的人形机器人可以省略面部表情, 所以图像标注简化为 13 个点。为了提高鲁棒性, 对原始数据进行数据增强, 主要目的方面增加样本数量, 缓解样本数量不足; 另一方面, 增加数据多样性, 提高模型泛化能力。所采用的数据增强操作: 随机选择一半的身体进行训练; 随机对包围框进行旋转、缩放、平移等变换; YOLOX 风格的颜色增强; 进行图像变换, 如模糊、中值模糊、随机遮挡等。

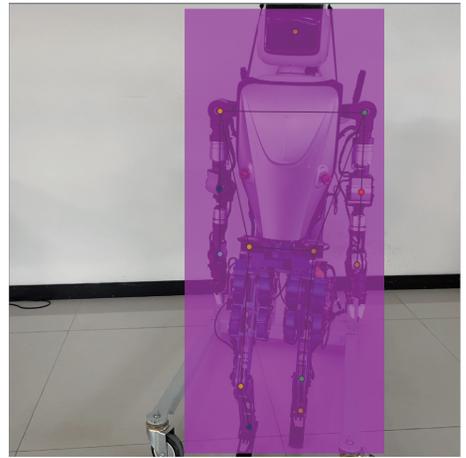


图 12 数据集标注(角度 1)

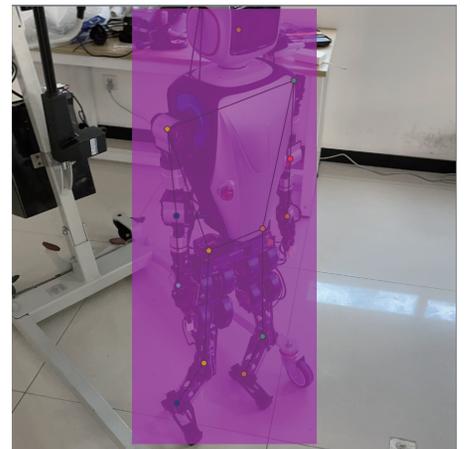


图 13 数据集标注(角度 2)

3.3 模型训练与评估

训练策略考虑了模型架构、数据处理、数据增强、学习率调整、优化器、日志和保存、分布式训练等多个方面, 以提高模型的性能和泛化能力。模型训练参数设置为: 基础批次 1 024; 初始基础学习率 0.004; 最大训练轮数 420。模型架构使用了基于 CSPNeXt 的骨干网络和 RTMCCHHead

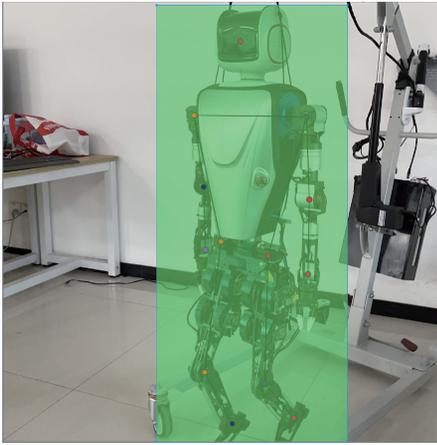


图 14 数据集标注(角度 3)

(real-time multi-person coordinate completion head)的姿态估计头部。骨干网络具有通道注意力和深度因子控制,而头部包含 KLDISCRETLOSS(kullback-leibler discretization loss)等损失函数。使用 PoseDataPreprocessor 进行数据预处理,包括 BGR 到 RGB 的转换和均值方差的归一化。通过多个数据增强阶段,包括随机缩放、旋转、模糊、CoarseDropout 等,以增加数据集的多样性,提高模型的鲁棒性。在训练过程中,通过 PipelineSwitchHook 实现了在第 390 轮切换数据处理 Pipeline,用于改变数据增强的方式。使用了分段调整学习率的策略,包括线性学习率衰减和余弦退火学习率调度器。使用了 AdamW(adam weight decay)优化器,并对梯度进行了裁剪,以稳定训练过程。设置了定期保存模型和记录日志的策略,以便在训练过程中进行监控和分析。设置了分布式训练的配置,使用了 NCCL(NVIDIA collective communication library)作为分布式后端,并指定了 fork 方式的启动。遍历训练集多个 epoch,通过 8 : 2 的比例训练和验证数据集进行模型的训练和评估。设置了可视化策略,如图 15 所示包括显示姿态估计结果和保存可视化结果的目录。



图 15 RTMpose 模型训练识别结果

如图 16 所示,训练结果表明模型在不同条件下都取得了较好的性能。

为了定量分析模型对姿态特征点识别的提取的效果,

Average Precision (AP)	@ IoU=0.50:0.95	area= all	maxDets= 20	l = 0.995
Average Precision (AP)	@ IoU=0.50	area= all	maxDets= 20	l = 1.000
Average Precision (AP)	@ IoU=0.75	area= all	maxDets= 20	l = 1.000
Average Precision (AP)	@ IoU=0.50:0.95	area=medium	maxDets= 20	l = -1.000
Average Precision (AP)	@ IoU=0.50:0.95	area= large	maxDets= 20	l = 0.995
Average Recall (AR)	@ IoU=0.50:0.95	area= all	maxDets= 20	l = 0.997
Average Recall (AR)	@ IoU=0.50	area= all	maxDets= 20	l = 1.000
Average Recall (AR)	@ IoU=0.75	area= all	maxDets= 20	l = 1.000
Average Recall (AR)	@ IoU=0.50:0.95	area=medium	maxDets= 20	l = -1.000
Average Recall (AR)	@ IoU=0.50:0.95	area= large	maxDets= 20	l = 0.997

图 16 训练评估指标

以数据集中人工标记的图为基准,计算验证集识别特征点的准确率,总计 270 张图,3 510 个特征点,结果达到了 86.89%的准确率。

4 运动伺服控制

天轨机器人的伺服系统由双目相机、三轴分立驱动、Jetson nano 处理器构成。核心包括深度信息获取,姿态关节检测及位置坐标对准。伺服校准系统利用深度位置信息与电机编码器记录位置的误差作为反馈量完成闭环控制,完成位置对准。运动控制算法采用五次多项式轨迹规划加 PID 算法,PID 是一种结合比例(proportional, P)、积分(integral, I)和微分(differential, D)3 种环节于一体的闭环控制算法,它是目前为止在连续控制系统中计数最为成熟的一种控制算法^[18]。基于 PID 控制的五次多项式轨迹规划在保证了电机准确运动到目标位置的前提下,生成平滑的轨迹,使得起点、终点以及中间点的位置、速度和加速度都满足连续性和平稳性,消除电机抖动^[19]。

五次多项式轨迹规划同时规划了位移方程、速度方程、加速度方程:

$$x(t) = c_0 + c_1t + c_2t^2 + c_3t^3 + c_4t^4 + c_5t^5 \quad (6)$$

$$v(t) = c_1 + 2c_2t + 3c_3t^2 + 4c_4t^3 + 5c_5t^4 \quad (7)$$

$$a(t) = 2c_2 + 6c_3t + 12c_4t^2 + 20c_5t^3 \quad (8)$$

带入起末位置初始条件:

$$x(t_s) = c_0 + c_1t_s + c_2t_s^2 + c_3t_s^3 + c_4t_s^4 + c_5t_s^5 \quad (9)$$

$$v(t_s) = c_1 + 2c_2t_s + 3c_3t_s^2 + 4c_4t_s^3 + 5c_5t_s^4 \quad (10)$$

$$a(t_s) = 2c_2 + 6c_3t_s + 12c_4t_s^2 + 20c_5t_s^3 \quad (11)$$

$$x(t_e) = c_0 + c_1t_e + c_2t_e^2 + c_3t_e^3 + c_4t_e^4 + c_5t_e^5 \quad (12)$$

$$v(t_e) = c_1 + 2c_2t_e + 3c_3t_e^2 + 4c_4t_e^3 + 5c_5t_e^4 \quad (13)$$

$$a(t_e) = 2c_2 + 6c_3t_e + 12c_4t_e^2 + 20c_5t_e^3 \quad (14)$$

变换为矩阵形式:

$$\begin{bmatrix} 1 & t_s & t_s^2 & t_s^3 & t_s^4 & t_s^5 \\ 1 & t_e & t_e^2 & t_e^3 & t_e^4 & t_e^5 \\ 0 & 1 & 2t_s & 3t_s^2 & 4t_s^3 & 5t_s^4 \\ 0 & 1 & 2t_e & 3t_e^2 & 4t_e^3 & 5t_e^4 \\ 0 & 0 & 2 & 6t_s & 12t_s^2 & 20t_s^3 \\ 0 & 0 & 2 & 6t_e & 12t_e^2 & 20t_e^3 \end{bmatrix} \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \end{bmatrix} = \begin{bmatrix} x_s \\ x_e \\ v_s \\ v_e \\ a_s \\ a_e \end{bmatrix} \quad (15)$$

图 17 为以 0~0.5 m 和 0~0.5 s 为起末位移和时间初始条件,通过五次多项式轨迹规划器规划出的轨迹波形:

图 18 为天轨机器人伺服控制框架,其中深度图和姿态检测负责提取关节位置信息,将其作为目标位置输入,反

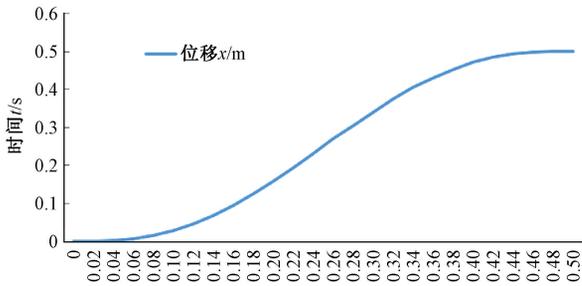


图 17 五次多项式规划轨迹

反馈每 2 ms 循环一次,每次循环计算目标位置与当前位置的偏差,将偏差传给五次多项式轨迹规划器,规划出每个循环的目标位置、速度和加速度。

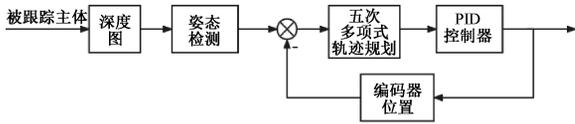


图 18 天轨机器人伺服控制框架

5 运行试验结果

本文设计的姿态估计天轨跟踪系统目前已在实验室投入应用,现场跟踪画面和特定关节特征点深度信息采集如图 19、20 所示。上一版本 MediaPipe 姿态识别模型识别图如图 21 所示,特征点存在明显的偏移,相比之下训练后的

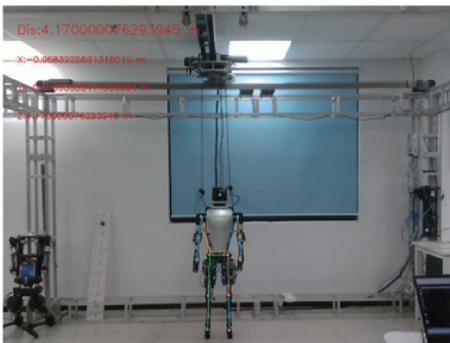


图 19 天轨实时跟踪识别画面

```
depth: 4.034000396728516
camera_coordinate: [-0.066094309091568, -0.0905824601650238, 4.034000396728516]
array('f', [4.034000396728516, -0.066094309091568, -0.0905824601650238, 4.034000396728516, 0.0])
depth: 3.9560000896453857
camera_coordinate: [-0.06481632590293884, -0.08883098512887955, 3.9560000896453857]
array('f', [3.9560000896453857, -0.06481632590293884, -0.08883098512887955, 3.9560000896453857, 0.0])
depth: 4.060000419616699
camera_coordinate: [-0.06652038348777771, -0.09116628766059875, 4.060000419616699]
array('f', [4.060000419616699, -0.06652038348777771, -0.09116628766059875, 4.060000419616699, 0.0])
depth: 4.006999969482422
camera_coordinate: [-0.06565192341804504, -0.08997617661952972, 4.006999969482422]
array('f', [4.006999969482422, -0.06565192341804504, -0.08997617661952972, 4.006999969482422, 0.0])
depth: 4.034000396728516
camera_coordinate: [-0.066094309091568, -0.0905824601650238, 4.034000396728516]
array('f', [4.034000396728516, -0.066094309091568, -0.0905824601650238, 4.034000396728516, 0.0])
```

图 20 机器人头部深度信息采集

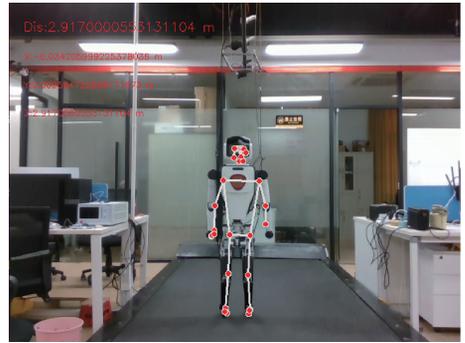


图 21 旧版本模型跟踪识别画面

RTMpose 模型特征点的识别精度有了显著的提高,帧数也从 104 提高到 143,大幅提高了实时性。

实际测试中,当机器人跌倒后,系统有较快的应对能力,能及时将机器人吊起,进行实时的保护,如图 22~24 所示运行效果表现良好。

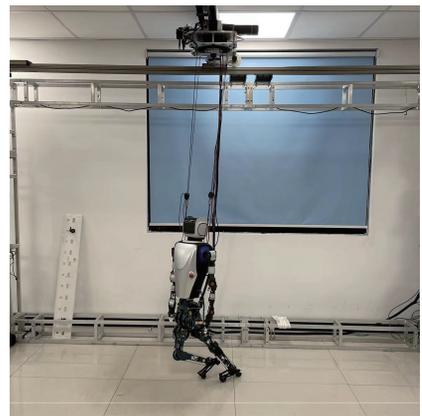


图 22 保护画面(一)

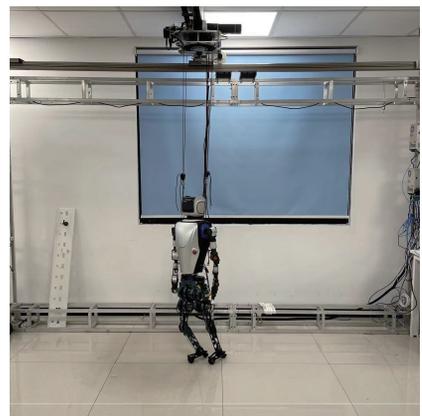


图 23 保护画面(二)

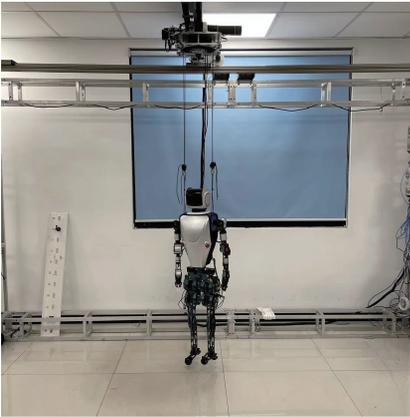


图 24 保护画面(三)

6 结 论

本文采用将双目深度视觉与姿态估计融合的方法,实现了天轨系统对人形机器人的实时跟踪保护。通过双目相机左右目画面匹配得到带有深度信息的图像,在图像的基础上使用训练的 RTMpose 姿态估计框架估计出机器人各个关节信息,得到每个关节的三维深度信息,最后通过关节三维信息进行 XYZ 3 轴方向上的实时位置跟踪。此外,通过在程序上实行高循环频率和五次多项式轨迹规划算法,保证了跟踪的实时性和平滑性,天轨机器人在实验室的运行结果证明了该系统设计方法有效。

现实调试中,本设计的应用在极大程度上提高了机器人调试效率,并对机器人实施了高度的保护,极大减少了跌倒意外的发生,具有十分先进的设计意义。

后续可以通过继续添加不同角度和距离的数据集,和对模型进行改进来增加模型对机器人姿态识别的准确性;并对机器人姿态进行建模,制定动作策略,以应对未来机器人更多样化的动作模式实现更高动态的跟踪功能。

参考文献

- [1] 宁静. 采用红外织网的室内定位技术[J]. 激光与红外, 2011, 41(7): 774-778.
- [2] 郝德华. 基于蓝牙 4.0 的室内定位算法与系统技术研究[D]. 锦州: 辽宁工业大学, 2019.
- [3] 代森. 基于超声波室内定位系统的设计与实现[D]. 成都: 西南交通大学, 2017.
- [4] 杨佳丽, 史恩秀, 王晔, 等. 基于无线网络的室内移动机器人定位方法研究[J]. 机械科学与技术, 2013, 32(3): 457-461, 468.
- [5] 林沂, 孙晶京, 闫旭. 地磁导航定位技术原理与方法综述[J]. 全球定位系统, 2023, 48(6): 32-41.
- [6] 杨春媚, 王峰, 晋博, 等. 基于计算机视觉的室内自主移动机器人导航综述[J]. 电脑知识与技术, 2017, 13(15): 141-143.

- [7] 危双丰, 庞帆, 刘振彬, 等. 基于激光雷达的同时定位与地图构建方法综述[J]. 计算机应用研究, 2020, 37(2): 327-332.
- [8] 高伟, 侯聪毅, 许万旸, 等. 室内导航定位技术研究进展与展望[J]. 导航定位学报, 2019, 7(1): 10-17.
- [9] 黄坤. 基于 Jetson Nano 人型机器人辅助行走天轨实时跟踪系统的实现[D]. 重庆: 重庆邮电大学, 2022.
- [10] 田明锐, 杨皓, 胡永彪. 大视场室内移动机器人高精度动态定位方法[J]. 中国机械工程, 2022, 33(2): 194-201.
- [11] 薛文喧. 基于 yolov3-tiny 的小尺度行人检测算法研究[D]. 北京: 北京邮电大学, 2020.
- [12] APATYA Y B A, SUBIANTORO A, YUSIVAR F. Design and prototyping of 3-phase BLDC motor[C]. 2017 15th International Conference on Quality in Research(QiR): International Symposium on Electrical and Computer Engineering, 2017: 209-214.
- [13] CHEN C, LAN G, CHEN C, et al. Stereo imaging using hardwired self-organizing object segmentation [J]. Sensors, 2020, 20(20): 5833-5833.
- [14] 张广军. 机器视觉[M]. 北京: 科学出版社, 2008: 96-128.
- [15] NIE J. Stereo matching method for non-coding circular reference points based on motion consistency [J]. Journal of Computational Science, 2018, 27: 454-461.
- [16] BARNES C, SHECHTMAN E, FINKELSTEIN A, et al. PatchMatch: A randomized correspondence algorithm for structural image editing[C]. Proceedings of SIGGRAPH, 2009: 1-11.
- [17] 刘昶, 姜楠. 基于双目视觉的棒材端面中心点定位[J]. 电子测量技术, 2023, 46(14): 174-181.
- [18] 张伟博. 基于改进 PID 控制的悬架机器人路径跟踪节能仿真研究[J]. 机械设计与制造工程, 2019, 48(8): 29-32.
- [19] 李胜琴, 张明瑞. 基于双五次多项式的智能汽车换道路径规划研究[J/OL]. 南京信息工程大学学报(自然科学版), 1-11 [2024-03-19]. <https://doi.org/10.13878/j.cnki.jnuist.20230614001>.

作者简介

赵希禹(通信作者), 硕士, 主要研究方向为天轨系统研究与姿态识别算法。

E-mail: 1532141002@qq.com

杨芳艳, 硕士, 副教授, 主要研究方向为机器人、人工智能、电路与系统。

E-mail: 379220890@qq.com

李家龙, 硕士, 主要研究方向为双足机器人运动控制算法研究。

E-mail: 920115574@qq.com