

DOI:10.19651/j.cnki.emt.2106361

基于强化学习的机器人人脸搜索和跟踪方法*

卞凯 乔铁柱 杨毅 张海涛

(太原理工大学 新型传感器与智能控制教育部重点实验室 太原 030024)

摘要: 巡检机器人在人员视觉检测中,人员位置呈现高动态特点。针对近距离下机器人只能捕捉到人体的局部信息导致被动视觉检测效率低下的情况,提出了一种基于深度强化学习的主动视觉方法。该方法使用深度卷积网络提取图像特征,利用强化学习策略训练动作决策网络控制云台相机转动。实验结果表明,该算法可使机器人根据图像中出现的人体局部信息转动云台相机实现人脸主动搜索,突破了传统方法被动式检测的局限性,提高了视觉检测算法在机器人应用中的适应性。

关键词: 主动视觉;强化学习;目标搜索;目标跟踪

中图分类号: TN911.73 **文献标识码:** A **国家标准学科分类代码:** 510.40

Robot visual face search and tracking method based on reinforcement learning

Bian Kai Qiao Tiezhu Yang Yi Zhang Haitao

(Key Laboratory of Advanced Transducers and Intelligent Control System Ministry of Education, Taiyuan University of Technology, Taiyuan 030024, China)

Abstract: In the visual inspection of personnel, the human position presents highly dynamic characteristics. For the situation that the robot can only capture the local information of the human body at close range leading to the inefficiency of passive vision detection, we propose an active vision method based on deep reinforcement learning. This method uses a deep convolutional network to extract image features and uses reinforcement learning strategies to train an action decision network to control the pan-tilt camera. Experimental results show that the algorithm enables the robot to turn the pan-tilt camera to achieve active face search based on the local information of the human body appearing in the image, which breaks through the limitations of the traditional method of passive detection and improves the adaptability of the visual detection algorithm in robot applications.

Keywords: active vision; reinforcement learning; object searching; object tracking

0 引言

近年来,巡检机器人获得了广泛的关注^[1-2]。巡检机器人主要利用相机来实现对待测的检测,并利用机器视觉算法来实现分析^[3-4]。在巡检机器人的应用场景中,经常需要对人员这种动态目标进行识别和检测。目前对于移动目标的机器视觉研究主要集中于目标跟踪^[5-8]。目标跟踪主要研究在一系列视频帧中对目标的定位问题^[9-11],这些研究都基于一个假设,即目标已经完全出现在图像中。然而对于移动机器人和移动目标来说,由于相对位置的不确定性,通常不能保证目标一开始就完整出现在相机视野中,因此这种假设通常是不成立的。此外,现有目标跟踪算法大多

基于固定相机拍摄的视频画面,实现目标在图像范围内的定位,难以直接运用于机器人云台相机的主动目标跟踪中。

本文提出了一种深度强化学习^[12-13]方法来解决上述问题。网络以相机捕捉的图像作为输入,以相机云台的俯仰、回转动作作为输出,来控制相机实现对面脸的主动跟踪。此外,该网络具备对目标关联结构的感知能力,即能够感知图像中出现的手臂、腿等人体局部信息,并据此对面脸方位做出估计,进而控制相机实现对面脸的主动搜索。

模型训练基于一个静态虚拟环境,并在其中引入一个预训练的人脸检测器来检测样本中的人脸,作为样本标签并生成奖励,来取代手工标记样本。虽然训练环境基于静态图像,但是网络在实际机器人上取得了不错的效果。理

收稿日期:2021-04-12

* 基金项目: NSFC-山西煤基低碳联合基金项目(U1810121)、2020年中央引导地方科技发展资金项目(YDZX20201400001796)资助

论上,用于生成奖励的人脸检测器性能会对网络性能产生限制,但是最终训练后的网络能对人脸检测器无法检测的人脸(侧面的脸等)实现感知和跟踪,展现出了较高的鲁棒性。

1 目标搜索

目前,大多数机器视觉研究都默认目标已经出现在相机视野中,但是对于移动机器人来说,寻找目标这一步骤是不容忽视的。传统的机器视觉算法大多基于固定相机,只能实现对画面中物体的被动目标检测。而移动机器人可以利用相机云台具备的自由度优势实现相机转动来获得更广的视野,从而实现对周围物体的搜索,以此提高机器人的视觉检测能力。

人脸作为人体的局部,其位置和身体其他部分存在很强的关联性。利用这种相对位置的关联性,即可实现人体任意局部出现在相机视野中对视野外人脸位置的准确估计。再借助对云台相机的角度调整,即可实现对人脸的主动搜索。由于搜索算法的目标是通过调整相机角度使人脸位于画面的中央,因此搜索算法自身就具备目标跟踪的能

力。并且只要画面中还存在目标相关局部信息,跟踪就不会丢失。相比单纯依靠目标自身特征实现的跟踪算法具备更高的鲁棒性。

2 方 法

2.1 网络结构

图 1 所示为网络的整体结构。网络由 16 层卷积层和 3 层全连接层组成。网络以 RGB 图像作为输入,以相机的俯仰和回转动作作为输出。网络的参数由 θ 表示。状态 s_t 代表大小为 $224 \times 224 \times 3$ 的输入图像。 a_t 为 t 时刻做出的动作。卷积层由预训练的 VGG-16 模型来初始化。3 层全连接层(fc1, fc2, fc3)都使用 ReLU 作为激活函数。输出层(fc3)输出每个可能动作的 Q 值 $Q(s, a; \theta)$ ^[14 15], 并且连接了一个 SoftMax 层来预测每个动作的概率 $p(a_i | s_t, \theta)$ 。动作 a_t 可以表示为:

$$a_t = \underset{a}{\operatorname{argmax}} p(a_i | s_t, \theta) \tag{1}$$

此外,为了提高在实际测试中相机运动的连续性,除了“停止”以外,动作集包括 4 个主运动方向“上、下、左、右”,和 4 个合成运动方向“右上、右下、左上、左下”。

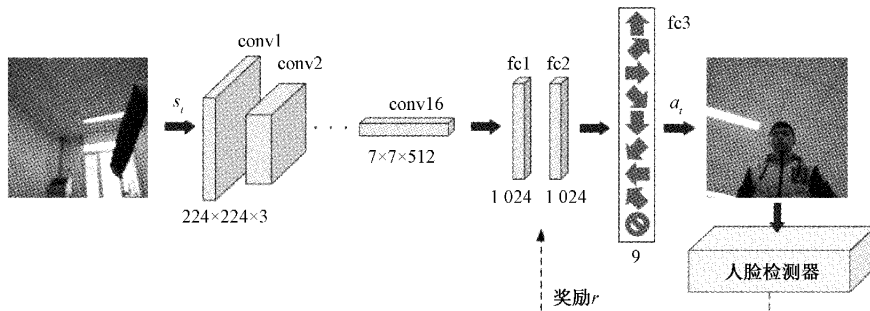


图 1 网络结构

人脸检测器是一个额外引入的预训练模型,用于代替人工来检测状态 s_{t+1} 下的人脸。这个人脸检测的结果仅仅在网络训练时用于生成奖励,并不会参与到网络的前向传播或者动作决策中。奖励函数 $r_t(s_{t+1})$ 的表达式为:

$$r_t(s_{t+1}) = \begin{cases} R_{\max} - \frac{\sqrt{(x_f - x_c)^2 + (y_f - y_c)^2}}{k}, & \text{有人脸} \\ 0, & \text{无人脸} \end{cases} \tag{2}$$

其中, R_{\max} 为当人脸刚好出现在视野中心时,返回的

最大奖励值。 $k > 0$ 是一个校正因子。 R_{\max} 和 k 都是可调的参数。 x_f, y_f 和 x_c, y_c 分别为图像和人脸中心坐标。

2.2 网络训练

深度强化学习网络需要通过不断地尝试来进行训练。而在实际环境中直接进行训练是很困难的,因为网络输出的动作是不确定的,需要监视和调整,因此训练效率很低。在实验中,深度强化学习模型由一个虚拟训练环境来实现训练。这个环境可以利用静态图像来模拟相机的转动和视野的变化如图 2 所示。图像中的大方框表示相机视野,它可以被网络的输出动作 a_t 控制。



图 2 虚拟训练环境中一次搜索过程的一系列关键帧

在训练过程中,系统会随机从训练集中抽取图像作为环境背景,同时随机初始化虚拟视野框位置。对于每张训练图像,网络会进行一定次数的搜索尝试,每次尝试都会得到相应的奖励反馈。训练过程可以由动作函数 $a_t(s_t)$, 状态转移函数 $s_{t+1}=f(s_t, a_t)$ 和奖励函数 $r_t(s_{t+1})$ 来描述。4 个关键参数 $\{s_t, a_t, s_{t+1}, r_t\}$ 会被存储到容量为 400 的记忆库中,并在误差计算和反向传播时被随机抽取。

3 实 验

算法被部署在一个装有云台相机的实际机器人上进行了测试如图 3 所示。机器人上的相机由算法控制进行转动来实现人脸搜索和跟踪。一个典型的完整过程如图 4 所示,图中标记表示了当前云台相机的转动方向。算法

的性能可由如下 3 个指标来衡量:搜索阈值、搜索效率和跟踪精度,分别描述了搜索的开始、搜索过程和跟踪过程 3 个阶段的性能。

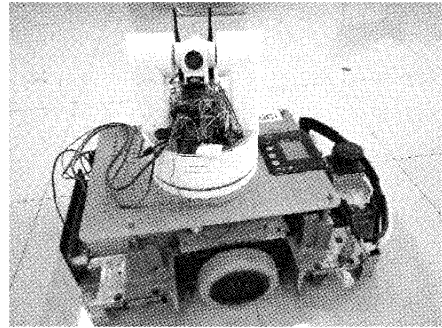


图 3 装有相机云台的机器人



图 4 典型搜索过程的一系列关键帧

3.1 搜索阈值

当画面中仅出现人体局部特征时,算法会根据对人体结构的先验知识来实现对人脸位置的估计和搜索。搜索阈值描述了算法对人体局部信息的敏感度。人体在画面

中的占比越大,越容易被算法捕捉。搜索阈值被定义为人体局部像素在图像中的占比。图 5(a)展示了一些搜索起始关键帧。图 5(b)标记出了相对应的人体局部的像素区域。搜索阈值越小,算法对人体局部信息就越敏感。

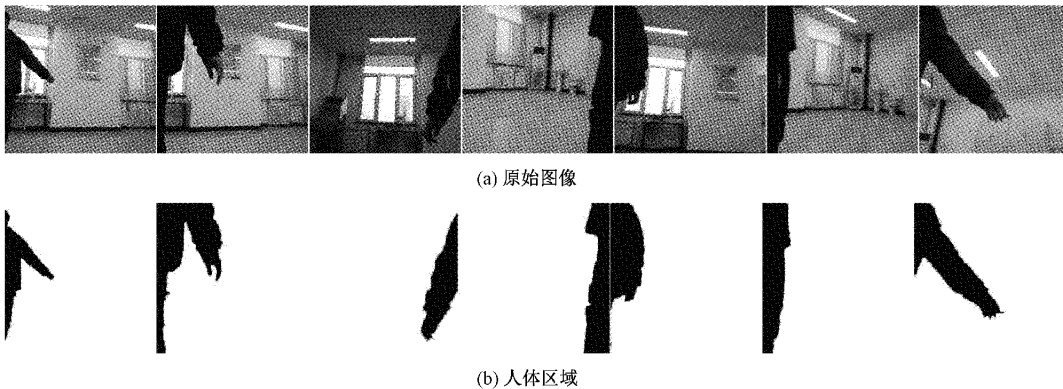


图 5 部分搜索起始帧中人体局部像素分布情况

受透视影响,在相同的像素占比下,远距离的物体在图像中的完整性会更高,这也会对算法产生影响。算法在不同距离下的搜索阈值如图 6 所示。每个距离下都做了 8 次独立测试。随着距离的增加,平均阈值会逐渐下降。

3.2 搜索效率

最理想的搜索路径是一条从起点到人脸的直线。但实际路径往往是一条曲线。图 7 所示为在一次测试种的相机角度曲线,实线代表实际路径,虚线代表最短路径。图中最短路径长度为 47.9,实际路径长度为 63.9,本次测试的搜索效率为 74.9%。

搜索效率用于衡量实际路径与最短路径的偏离程度。

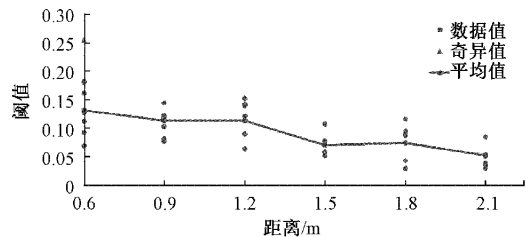


图 6 不同距离下的搜索阈值

搜索效率由 d/l 定义,其中 d 是最短路径长度, l 为实际路径长度。3 个不同的距离下的 30 次搜索效率如图 8 所示,其中方块、圆形和三角的数据点分布代表不同距离下的测

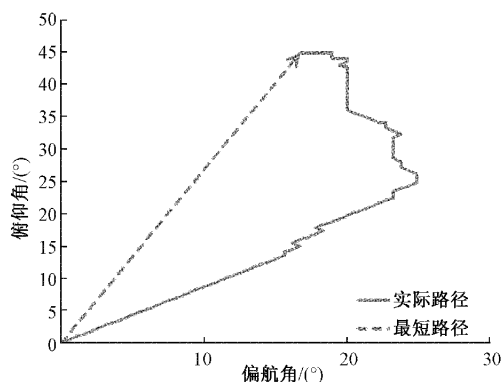


图7 一次测试中相机的角度曲线

试结果,他们基本服从相同的分布特征。数据表明,在该方法中,距离对搜索效率没有太大影响。该方法的平均搜索效率为75.9%。

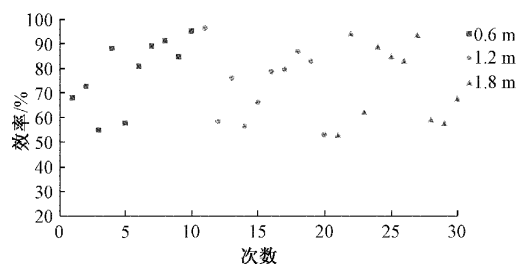


图8 30次独立测试的搜索效率

3.3 跟踪精度

当人脸完全出现在相机视野中时,搜索过程就结束了。将视野中心对准目标人脸的动作使搜索算法具备目标跟踪的能力。跟踪精度描述了跟踪状态下人脸与相机视野中心的重合程度。由于算法使用整幅图像作为输入直接得到相机动作,过小的目标会收到较大的背景噪声影响,因此目标在图像中的大小会影响跟踪精度。使用人脸像素在图像中的占比作为变量可以使跟踪精度在不同尺度的输入图像上得到一致的结果。跟踪精度与像素占比的关系曲线如图9所示。

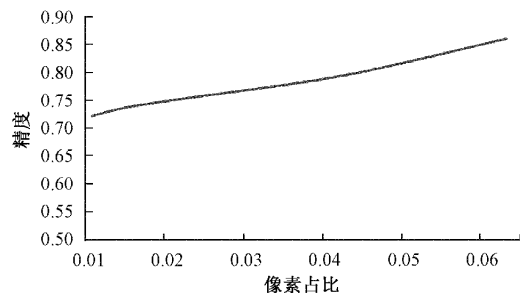


图9 跟踪精度与像素占比曲线

4 结论

针对传统被动式视觉检测在动态环境中效率低下的问题,采用深度强化学习技术提取图像中出现的目标相关

信息,结合机器人的灵活性实现了主动视觉搜索。该方法提高了巡检机器人捕获目标的能力,不再需要被动等待目标进入视野范围,提高了机器人在复杂动态环境中进行视觉检测的适应性,为后续的各种视觉诊断分析打下了基础。

参考文献

- [1] 黄山,吴振升,任志刚,等. 电力智能巡检机器人研究综述[J]. 电测与仪表, 2020, 57(2): 26-38.
- [2] 沈超. 矿用自动巡检机器人在黄陵一号煤矿的应用[J]. 陕西煤炭, 2020, 39(2): 118-120.
- [3] 郭丹蕊,车玉秋,何永涛. 基于智能巡检机器人的变电站双针仪表读数识别及其算法[J]. 数码世界, 2021(1): 3-4.
- [4] 王毅,何宁,王恺,等. 巡检机器人获取柑橘树上果实完整表面信息方法研究[J]. 农业机械学报, 2020, 51(4): 232-240.
- [5] YUN S, CHOI J, YOO Y, et al. Action-decision networks for visual tracking with deep reinforcement learning [C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, 2017: 1349-1358.
- [6] DANELLJAN M, BHAT G, SHAHBAZ KHAN F, et al. Eco: Efficient convolution operators for tracking[C]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017: 6931-6939.
- [7] WANG Q, ZHANG L, BERTINETTO L, et al. Fast online object tracking and segmentation: A unifying approach [C]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 1328-1338.
- [8] LUO W, SUN P, ZHONG F, et al. End-to-end active object tracking and its real-world deployment via reinforcement learning [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(6): 1317-1332.
- [9] FAN H, LIN L, YANG F, et al. LaSOT: A high-quality benchmark for large-scale single object tracking[J]. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019: 5369-5378.
- [10] 马素刚,赵祥模,侯志强,等. 一种基于ResNet网络特征的视觉目标跟踪算法[J]. 北京邮电大学学报, 2020, 43(2): 129-134.
- [11] 咎珊珊,李波. 融合改进YOLOv2网络的视觉多目标跟踪方法[J]. 小型微型计算机系统, 2020, 41(12): 2601-2606.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]. Communications of the ACM 60, 2017:

- 84-90.
- [13] MNIH V, KAVUKCUOGLU K, SILVER D, et al. Human-level control through deep reinforcement learning[J]. Nature, 2015, 518(7540): 529-533.
- [14] HASSELT H V. Double Q-learning[J]. Advances in Neural Information Processing Systems, 2010, 23: 2613-2621.
- [15] HASSELT H V, GUEZ A, SILVER D. Deep reinforcement learning with double q-learning [C].

AAAI' 16: Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, 2016: 2094-2100.

作者简介

卞凯, 硕士研究生, 主要研究方向为机器视觉和机器人技术。

E-mail: biankai0867@link.tyut.edu.cn

乔铁柱, 教授, 主要研究方向为无损检测与智能传感技术。

E-mail: qiaotiezhu@tyut.edu.cn