

文章编号: 2095-4980(2021)02-0217-07

## 基于强化学习的公共空间行人轨迹跟踪

许盛宇, 苏 婕, 卿 粼波\*, 牛 通

(四川大学 电子信息学院, 四川 成都 610065)

**摘 要:** 对公共空间中的多目标行人轨迹跟踪问题, 提出一种基于强化学习的多目标行人轨迹跟踪算法。首先采用高精度度的目标检测器检测公共空间视频中的行人目标, 并为每个目标分配一个独立的单目标跟踪器进行轨迹跟踪; 将每个目标作为独立智能体, 通过深度强化学习方式训练; 接下来结合跟踪轨迹与检测目标之间的表观和位置特征构建相似度代价矩阵; 最终通过匈牙利算法实现数据关联。实验表明, 在常用公开数据集上本文算法跟踪精确度达 76.1%, 表明算法对多目标轨迹跟踪的可行性与有效性。

**关键词:** 轨迹跟踪; 马尔科夫决策; 强化学习; 数据关联

**中图分类号:** TN820.4; TP391.41

**文献标志码:** A

**doi:** 10.11805/TKYDA2019412

## Pedestrian trajectory tracking in public space based on reinforcement learning

XU Shengyu, SU Jie, QING Linbo\*, NIU Tong

(School of Electronic Information Engineering, Sichuan University, Chengdu Sichuan 610065, China)

**Abstract:** Aiming at the multi-objective pedestrian trajectory tracking problem in public space, a multi-objective pedestrian trajectory tracking algorithm based on reinforcement learning is proposed. Firstly, a high-precision target detector is utilized to detect pedestrian targets in space, and an independent single-target tracker is assigned for each tracking target trajectory. Each target is trained as an agent by means of deep reinforcement learning, and combined with the appearance and position characteristics between tracking trajectory and detecting target, the tracking target is constructed. Similarity cost matrix is built to realize data association through Hungarian algorithm. Experiments show that the tracking accuracy of this algorithm is 76.1% on common open data sets. Good results have been achieved in multi-objective pedestrian trajectory tracking in public space.

**Keywords:** trajectory tracking; Markov decision; reinforcement learning; data association

近年来, 计算机视觉的研究被纳入了战略层面, 国家提出加强新一代视觉研发应用。作为计算机视觉领域的核心研究方向之一, 目标跟踪在人机交互、智能监控以及城市规划等领域具有重要研究价值。其中基于人的城市公共空间行人运动轨迹研究, 对城市结构、基础设施建设及交通管理预测等方面都有重要意义<sup>[1]</sup>。为获得公共空间中的行人运动轨迹, 需要对该场景下的行人目标进行持续准确的跟踪。考虑到实际公共空间场景中的人员流动性, 必然涉及多目标同时跟踪。相比于单目标跟踪, 多目标跟踪问题更为复杂, 因为不仅仅需要对每个目标进行有效的跟踪, 同时还需要解决不同目标之间的相互干扰问题。Wax 于 1955 年在文献[2]中首次提出多目标跟踪, 文献中提出的轨迹起始、轨迹维持以及轨迹终止等基本概念, 在目前亦是主要研究方向之一。在 2015 年 UCF Dr. Shah 等<sup>[3]</sup>提出 GMMCP 算法, 通过将多行人跟踪中的数据关联看作最大化多团问题, 利用最大二值整数规划构建无向图进行求解, 然而该算法的计算复杂度依赖于虚拟节点的建立方式, 对于复杂场景难以实施。同年, H Rezaatofighi 等<sup>[4]</sup>提出一种改进的联合概率数据关联 JPDA 算法, 该算法通过每个目标与轨迹之间的联合概率进行数据关联, 但对于新旧目标的更替情况效果并不理想。2018 年 Fang 等<sup>[5]</sup>提出一种基于循环

收稿日期: 2019-10-22; 修回日期: 2019-12-04

基金项目: 国家自然科学基金资助项目(61871278); 四川省科技计划项目资助项目(2018HH0143)

作者简介: 许盛宇(1996-), 男, 在读硕士研究生, 主要研究方向为多媒体通信。email:920518837@qq.com

\*通信作者: 卿粼波 email:qing\_lb@scu.edu.cn

自回归网络的多目标跟踪框架，该方法通过一个复合存储架构保留轨迹历史信息，再通过目标轨迹相似比进行数据关联，该算法关联方式受限于预测性能，对于轨迹丢失情况效果不理想。强化学习作为机器学习的核心方向之一，通过执行不同的行动在环境中获得最大化奖励，从而确定一种最优的决策方法，在智能控制等领域发挥着重要作用<sup>[6]</sup>。随着深度学习领域的快速发展，将深度学习方法与强化学习法相结合的深度强化学习算法在决策问题上取得了许多非常卓越的成果。Yun<sup>[7]</sup>等于 2017 年首次将深度强化学习运用于目标跟踪领域，实现了连续视频序列中特定目标的持续跟踪。Liangliang Ren<sup>[8]</sup>等于 2018 年提出一种基于协同强化学习的多目标跟踪框架，对视频中的跟踪目标进行全局性的协同状态更新。总体而言，目前基于深度强化学习的跟踪算法多用于传统的视觉单目标跟踪，在多目标跟踪领域研究较少。

然而，由于公共空间视频的背景较为复杂多变，多目标之间的遮挡交互现象频发，使得多目标跟踪愈发困难。因此，提升算法对于目标遮挡、误检漏检等问题的鲁棒性十分重要。基于此，本文提出一种基于强化学习的多行人目标跟踪算法。本文算法将目标轨迹跟踪转换成一个马尔科夫决策过程，用强化学习的方式实现复杂场景中的多目标行人轨迹跟踪。为了解决实际公共空间中出现的目标遮挡与轨迹漂移问题，本文算法结合了目标的表现特征以及位置特征因素进行数据关联。实验结果表明，本文算法能对复杂场景下的多行人目标进行有效跟踪，同时对目标遮挡、误检漏检等情况具有一定的鲁棒性。

## 1 基于强化学习的多目标轨迹跟踪

### 1.1 总体算法框架

本文提出的多目标跟踪算法总体框架如图 1 所示。该算法主要由目标检测器、多个单目标跟踪器以及数据关联模块三部分构成。其中目标检测部分使用基于回归的目标检测算法，该算法集多种基于深度学习的目标检测算法优势于一身，其中包括残差网络结构、多尺度预测、预选框策略等多种提高目标检测精确度的方法。本文方法为每一个跟踪目标分配一个由深度强化学习训练得到的单目标跟踪器，在每一帧中对多个目标同时进行跟踪，得到各个目标的轨迹，同时通过高精度度的目标检测器检测当前帧中目标的位置，并根据其当前的表现信息以及位置信息通过匈牙利算法将检测结果与跟踪轨迹进行匹配，实现检测与跟踪的数据关联，从而实现视频序列中连续的多目标行人轨迹跟踪。

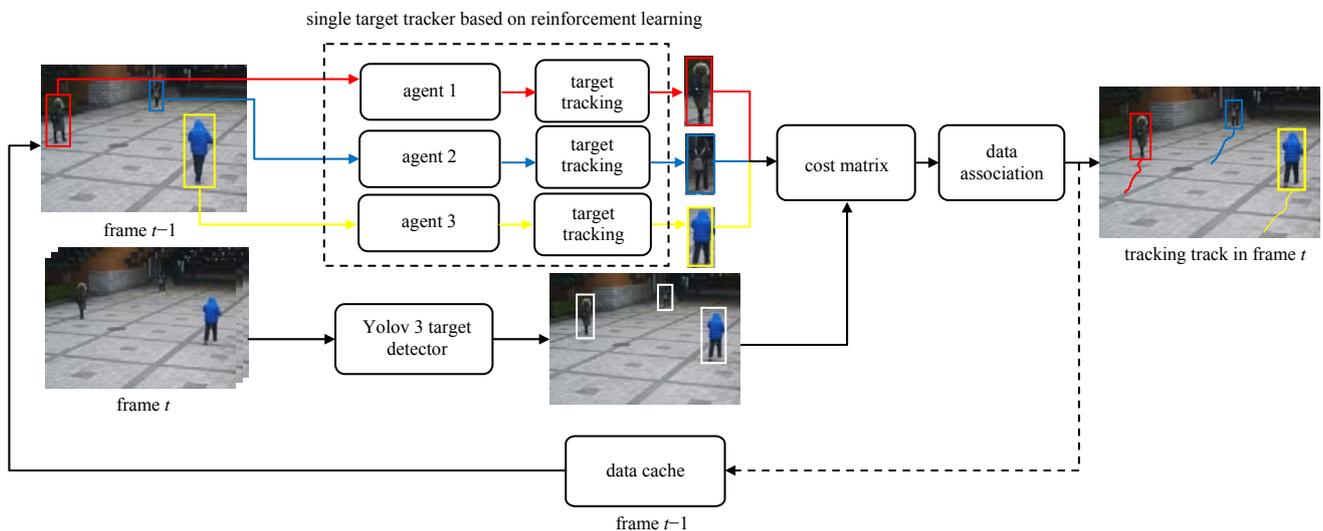


Fig.1 Framework of multi-target tracking algorithm  
图 1 多目标跟踪算法框架

### 1.2 基于强化学习的目标位置预测跟踪

#### 1.2.1 强化学习决策设置

强化学习<sup>[9]</sup>模型在每个状态下试图探索一个合适的行动，属于一个马尔科夫决策过程(Markov Decision Process, MDP)，每个状态下执行响应到的行动，该行动的具体实施方式没有标准样本，而是经过若干次试探之后，得到一个环境给出的奖励，强化学习通过最大化模型的奖励来学习状态和行动之间的映射函数。本文通过强化学习的方式训练每个单目标跟踪器，使得其有能力对其分配的目标进行持续的位置预测跟踪。

基于此，本文将跟踪任务转化为一个马尔科夫决策过程进行解决，该策略将每一个跟踪目标作为一个独立的智能体，包含目标当前的状态  $S$ ，目标执行的行动  $A$ ，状态变换函数  $S^* = f(S, A)$  以及根据行为环境予以的奖励  $R$ 。行动集合  $A$  中包含 {左移, 右移, 上移, 下移, 两倍左移, 两倍右移, 两倍上移, 两倍下移, 放大, 缩小, 停止} 共 11 种行动。目标当前状态  $S_t = ([x_t, y_t, w_t, h_t], d_t)$  定义为一个包含目标当前位置信息  $[x_t, y_t, w_t, h_t]$  以及目标在跟踪完成之前历史行动信息  $d_t$ ，其中  $x_t, y_t$  表示目标边界框中心点的坐标， $w_t, h_t$  表示边界框的宽度和高度。当视频逐帧送入本文网络中时，网络能够根据目标边界框在当前帧中的位置以及之前采取的历史行动  $d_t$ ，得到接下来智能体应该采取的行动。若选择等移位行动，则边界框在当前帧中进行相应的位置偏移；若选择范围变化行动，则边界框在当前帧中的大小进行相应的变化；若选择停止行动，则该目标在当前帧中的位置跟踪结束，得到新的目标状态。新的状态  $S^* = S_{t+1} = ([x_{t+1}, y_{t+1}, w_{t+1}, h_{t+1}], d_{t+1})$  中的位置信息即为目标在当前帧中的位置，并对行动的历史信息  $d_{t+1}$  进行同步更新。在当前帧跟踪完成时，环境将给出相应的奖励  $R$ ，奖励函数如式(1)所示：

$$R_t(S \rightarrow S^*) = \begin{cases} 1, & \text{if } IOU(F^*, G) \geq 0.75 \\ -1, & \text{otherwise} \end{cases} \quad (1)$$

式中： $F^*$  表示跟踪完成后边界框的位置； $G$  表示当前帧中目标的真实位置； $IOU$  是两个边界框间的交并比。

### 1.2.2 单目标跟踪器网络结构

图 2 是基于强化学习的单目标跟踪器的网络模型结构，该网络通过卷积层提取该智能体边界框中目标的特征信息，通过一个全连接层将该特征归一化为一个 512 维的特征向量。再分别经过两个全连接层得到该智能体接下来采取的行动，以及该行动的置信度。最后利用当前选择的行动更新行动历史状态，与特征向量融合，共同影响后续跟踪过程中的行动决策。

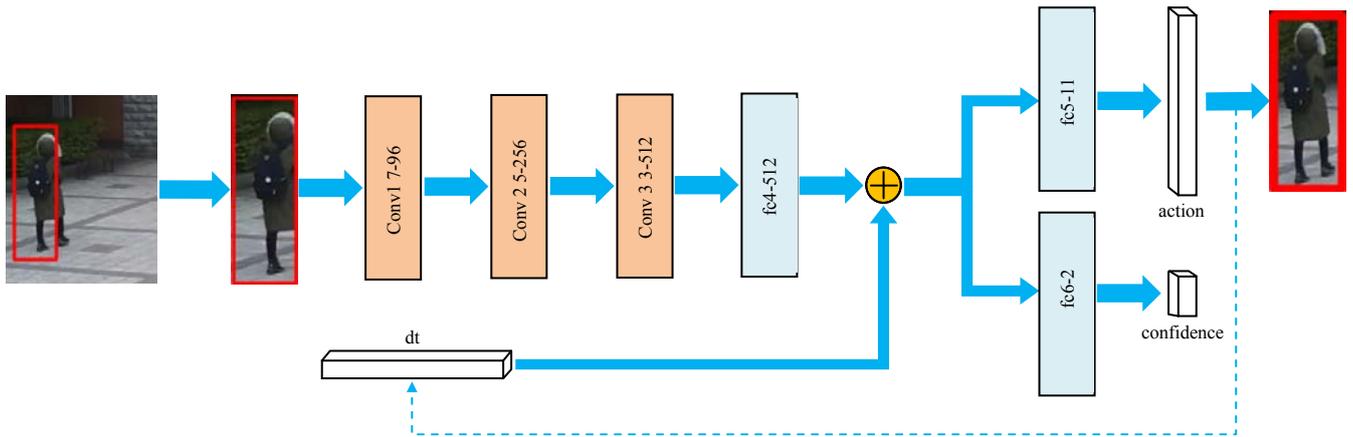


Fig.2 Network structure diagram of single target tracker  
图 2 单目标跟踪器网络结构图

鉴于需要学习的是多个离散的行动，本文采用策略梯度算法<sup>[10]</sup>对网络进行训练。本文强化学习网络在训练时首先选择若干个训练序列，使用  $S_{t,n}$  表示序列  $n$  中当前跟踪目标在  $t$  时刻下的状态， $A_{t,n}$  表示将要执行的行动，智能体当前选择的行动基于式(2)：

$$A_{t,n} = \arg \max_A \pi(A | S_{t,n}, \theta_{t,l}) \quad (2)$$

式中  $\pi(A | S_{t,n}, \theta_{t,l})$  表示在当前状态下行动  $A$  的概率分布。

为了使网络具有增大奖励值大的动作出现概率的能力，策略梯度算法需要学习一个基于梯度的策略性目标函数，接下来只需要通过对网络参数的优化，将目标函数最大化即可，网络参数更新方式如式(3)所示：

$$\Delta \theta_{t,l} \propto \sum_{n=1}^N \left( \sum_{t=1}^T \nabla_{\theta_{t,l}} \log \pi(A_{t,n} | S_{t,n}, \theta_{t,l}) R_t(S_{t,n} \rightarrow S_{t+1,n}^*) \right) \quad (3)$$

### 1.2.3 算法实现过程

为让本文网络具备一定选择正确行动的能力，首先通过监督学习的方式对网络进行预训练。通过给训练样本的真实标签位置添加高斯噪声进行混淆，生成大量虚拟边界框作为新的训练样本，将每一个生成的边界框移到原始真实标签位置，需要执行的行动和该动作的置信度作为新的样本标签，采用的损失函数如式(4)所示：

$$L_{\text{loss}}(F, L^{\text{action}}, L^{\text{confidence}}) = \frac{1}{l} \left( \sum H(L^{\text{action}}, \widehat{L^{\text{action}}}) + \sum H(L^{\text{confidence}}, \widehat{L^{\text{confidence}}}) \right) \quad (4)$$

式中： $F$  表示样本边界框； $L^{\text{action}}$  为行动标签； $L^{\text{confidence}}$  为置信度标签； $H$  表示交叉熵损失； $l$  表示一个批次训练样本的数量； $\widehat{L^{\text{action}}}$ 、 $\widehat{L^{\text{confidence}}}$  为网络预测的行动和执行度结果。

通过上述监督学习的训练方式，使得网络具备一定的选择合适行动的能力，此时网络中参数为  $\theta_{st}$ ，用此参数对网络进行初始化，用强化学习的方式进行二次训练得到  $\theta_{rl}$ 。强化学习网络在训练时首先选择若干个训练序列，在当前跟踪帧中，智能体通过当前行动状态的概率分布执行接下来的行动，并在选择到 {停止} 行动时终止跟踪过程，检查当前边界框的位置与对应的真实标签位置之间的交并比是否大于设置的阈值，并根据式(1)为其分配相应的奖励。最终通过对应的奖励值  $R$  以及每个状态下行动的概率，通过式(3)对网络参数  $\theta_{rl}$  进行优化。

综合上述监督学习与强化学习结合的方式，即可实现一个能对视频帧中的各个目标进行持续跟踪的单目标跟踪器系统。该系统既结合了目标的深度表观特征，又融合了各次跟踪过程中的历史行动信息，能够对行人目标轨迹进行持续的预测和跟踪，其对单个目标的跟踪效果如图 3 所示。



Fig.3 Effect of single object tracker  
图 3 单目标跟踪器效果图

### 1.3 基于度量学习的数据关联

为了将单目标跟踪器持续跟踪的目标轨迹与检测结果进行合理的关联，本文采用表观信息与位置信息结合的方式对二者之间的关联度进行度量。对于表观特征，本文通过行人重识别网络<sup>[11]</sup>对多目标行人表观信息深度特征进行提取，得到一个包含行人深度表观特征信息的向量。通过计算跟踪轨迹和检测目标表观特征向量之间的余弦距离，得到一个相似度代价矩阵  $S^{M \times N}$ ，余弦距离计算方式如式(5)所示：

$$d^{\text{appearance}} = \min \{ 1 - P_D^T P_T^i, i \in (1, k) \} \quad (5)$$

式中： $P$  为通过深度特征提取网络提取的表观特征向量； $k$  为当前目标轨迹中存储的历史  $k$  帧中的特征信息， $k$  设置为 50。

对于位置信息，本文通过轨迹集合与检测目标边界框之间的  $IOU$  进行相似度判别，若  $IOU(F_T, F_D) > GATE$ ，其中， $F_T, F_D$  分别表示跟踪轨迹与检测目标的边界框，则将代价矩阵  $S^{M \times N}$  中的对应值进行更新。设置  $GATE = 0.7$ ，更新值为  $10^5$ 。

综合跟踪目标轨迹与当前检测结果之间的位置与表观信息，得到相似度代价矩阵  $S^{M \times N}$ ，用于目标轨迹的数据关联，本文通过匈牙利算法<sup>[12]</sup>对该问题进行求解，如图 4 所示。在跟踪轨迹与检测集合进行关联的过程中，每个关联结果均对应一个关联代价，为找到使得总关联代价最小的分配方式，即利用匈牙利算法为所有关联的对象找到完全匹配图。

数据关联完成后，对于成功关联的跟踪轨迹与检测目标，将结合当前帧该目标的检测信息，对跟踪轨迹进行更新；对于未成功关联的跟踪轨迹，将认为其暂时处于遮挡状况，会保留其轨迹信息一段时间，若长时间未完成关联，则认为该跟踪轨迹已经消失；对于未成功关联的检测目标，认为这是当前帧中新出现的一个行人目标，会为其分配一个新的单目标跟踪器，并在后续帧中进行持续的跟踪。本文算法对于公共空间中短时间的目标遮挡问题，能够在其重新出现时利用位置和表观信息将遮挡目标关联至原轨迹。同时，本文算法在视频序列每一帧中都进行目标检测与数据关联，使得其对于可能产生的误检、漏检等情况具有一定鲁棒性。

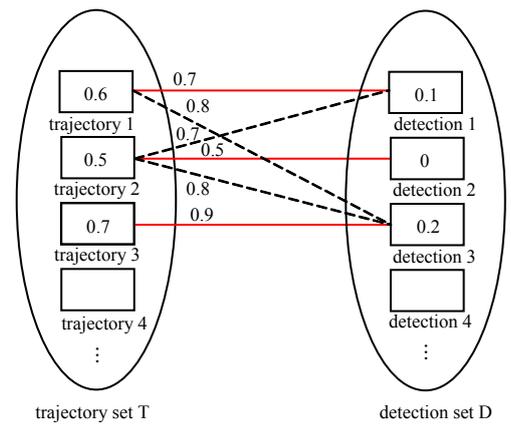


Fig.4 Matching of Hungarian algorithm  
图 4 匈牙利算法匹配图

## 2 实验与结果分析

本文基于深度强化学习的单目标跟踪器在 VOT2013<sup>[13]</sup>,VOT2014<sup>[14]</sup>,VOT2015<sup>[15]</sup>整合数据集上进行训练,为了让本文网络具备一定选择正确行动的能力,通过监督学习的方式对网络进行预训练。首先对训练样本的真实边界框赋予高斯噪声,生成大量虚拟的边界框作为新训练样本,将每一个生成的边界框移位到原始真实标签位置需要执行的行动和该动作的置信度作为新的样本标签,对网络进行预训练。再通过强化学习的方式让智能体通过当前行动状态的概率分布预测接下来的行动,并根据式(1)为其分配相应的奖励,通过最终的奖励值  $R$  以及每个状态下行动的概率,进一步优化网络参数。具体地,本文设置卷积神经网络学习率为 0.000 1,权重衰减为 0.000 5,每个批次大小为 128。

为说明本文算法的有效性,在公开数据集 2DMOT2015<sup>[16]</sup>和 MOT16<sup>[17]</sup>上进行测试,在对多目标跟踪结果进行评估时,往往考虑以下评价指标:跟踪准确度(Multiple Object Tracking Accuracy, MOTA)、跟踪精确度(Multiple Object Tracking Precision, MOTP)、完整跟踪轨迹的数量占比(Mostly Tracked, MT)和完全丢失跟踪的数量占比(Mostly Lost, ML),其中, MOTA:

$$MOTA = 1 - \frac{\sum_t (FN_t + FP_t + IDSW_t)}{\sum_t GT_t} \quad (6)$$

MOTP:

$$MOTP = \frac{\sum_{t,i} d_{t,i}}{\sum_t c_t} \quad (7)$$

表 1 和表 2 为本文算法与其他相关文献[18-23]中算法进行比较的结果。从表中可以看出,本文算法的 MOTP、MT 以及 ML 均优于其他几种算法。在 MOT16 数据集上的 MOTA 略低于文献[21]的 DCCRF 算法,可能的原因在于 DCCRF 算法中对行人目标进行特征提取时同时考虑先验帧中目标特征,并利用条件随机场将轨迹按置信度分类,提高了追踪准确度。本文算法对轨迹丢失情况 ML 有较大的改善,可能原因在于算法在每一帧追踪过程中均对行人目标进行检测关联,使得部分被遮挡轨迹在重新出现后仍然能重新连接到正确的追踪目标,有效改善了丢失情况 ML。

表 1 2DMOT2015 对比结果  
Table1 2DMOT 2015 results

method	MOTA ↑	MOTP ↑	MT ↑	ML ↓
NOMT <sup>[18]</sup>	33.7	71.9	12.2	44.0
JointMC <sup>[19]</sup>	35.6	71.9	23.2	39.3
AMIR <sup>[20]</sup>	37.6	71.7	15.8	26.8
DCCRF <sup>[21]</sup>	33.6	70.9	10.4	37.6
proposed	38.8	74.0	30.7	14.7

表 2 MOT16 对比结果  
Table2 MOT16 results

method	MOTA ↑	MOTP ↑	MT ↑	ML ↓
GM_PHD <sup>[22]</sup>	30.5	75.4	4.6	59.7
GMMCP <sup>[3]</sup>	38.1	75.8	8.6	50.9
MHT_bLSTM <sup>[23]</sup>	42.1	75.9	14.9	44.4
DCCRF <sup>[21]</sup>	44.8	75.6	14.1	42.3
proposed	43.2	76.1	15.9	29.5



Fig.5 Target occlusion in multi-target tracking  
图 5 多目标跟踪中目标遮挡情况



Fig.6 False detection and missed detection  
图 6 误检漏检情况

图 5、图 6 为本文跟踪算法对目标遮挡以及误检漏检情况的跟踪结果。图 5(a)展示了公共空间场景中目标跟踪情况,图 5(b)中 104 号跟踪目标被树木部分遮挡,本文算法依旧能够对其进行跟踪。由图 5(c)和图 5(d)可

知,当目标被完全遮挡时,本文算法对该目标的跟踪情况暂时丢失,但当目标在后续视频帧中重新出现时,算法能够重新对其进行关联,并不改变原有编号的情况下对其进行持续跟踪,表现出本文算法对于目标遮挡情况的鲁棒性。图 6 显示出当目标因遮挡或形变等因素被检测器漏检时,本文基于强化学习的单目标跟踪器依然能对其当前位置进行预测,并进行持续有效的跟踪。因此,本文算法能较好地实现多目标行人轨迹跟踪,并对目标遮挡、误检漏检等问题有一定的鲁棒性。该算法的优点在于,本文提出的多目标跟踪框架将每个单目标跟踪器作为一个独立的智能体,对公共空间视频中的目标进行分别跟踪,这使得即使因场景变化过大或目标被遮挡等原因导致检测器检测目标丢失,单目标跟踪器依旧能够对目标在当前帧中的位置进行预测,进行持续有效的跟踪,表明本文算法对目标遮挡、误检漏检等问题均具有一定的鲁棒性。然而,受限于相对复杂的网络结构以及深度强化学习算法计算速度的限制,本文提出的多目标跟踪框架暂时无法实现实时的在线多目标跟踪,但对离线场景下实际公共空间场景视频中的多行人目标表现出较好跟踪效果。将来,可以考虑进一步简化网络结构,提升算法速度,以满足可能存在的在线多目标跟踪需求。

### 3 结论

针对公共空间中的多目标行人轨迹跟踪问题,本文提出一种基于强化学习的多目标轨迹跟踪算法。本文算法将深度强化学习应用于多目标跟踪问题,将每个目标作为一个独立的智能体进行持续追踪,并通过表观和位置特征进行数据关联。实验表明,本文算法在跟踪精确度上达到 76.1%。对目标遮挡及误检漏检问题均具有一定的鲁棒性,且获得了较好的公共空间多目标行人轨迹跟踪效果,进一步为城市公共空间规划分析提供便利。

#### 参考文献:

- [1] 宋雪涛,蒲英霞,刘大伟,等. 利用行人轨迹挖掘城市区域功能属性[J]. 测绘学报, 2015,44(S1):82-88. (SONG Xuetao, PU Yingxia, LIU Dawei, et al. Mining the functional attributes of urban area by using pedestrian trajectory[J]. Journal of Surveying and Mapping, 2015,44(S1):82-88.)
- [2] WAX Nelson. Signal-to-noise improvement and the statistics of track populations[J]. Journal of Applied Physics, 1955, 26(5):586-595.
- [3] DEGHAN A, ASSARI S M, SHAH M. GMMCP tracker: globally optimal generalized maximum multi clique problem for multiple object tracking[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, MA, USA: IEEE, 2015:4091-4099.
- [4] REZATOFIGHI S H, MILAN A, ZHANG Z, et al. Joint probabilistic data association revisited[C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015:3047-3055.
- [5] FANG K, XIANG Y, LI X, et al. Recurrent autoregressive networks for online multi-object tracking[C]// 2018 IEEE Winter Conference on Applications of Computer Vision (WACV). Lake Tahoe, NV, USA: IEEE, 2018:466-475.
- [6] 周志华. 机器学习[M]. 北京:清华大学出版社, 2016. (ZHOU Zhihua. Machine learning[M]. Beijing: Tsinghua University Press, 2016.)
- [7] 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, HI, USA: IEEE, 2017:1349-1358.
- [8] REN L, LU J, WANG Z. Collaborative deep reinforcement learning for multi-object tracking[C]// European Conference on Computer Vision. Munich, Germany: Springer-Verlag, 2018:605-621.
- [9] RUMMERY G A, NIRANJAN M. On-line Q-learning using connectionist systems[M]. Cambridge, UK: University of Cambridge, 1994.
- [10] SUTTON R S, MC ALLESTER D A, SINGH S P, et al. Policy gradient methods for reinforcement learning with function approximation[C]// Advances in Neural Information Processing Systems. Cambridge, MA, USA: MIT Press, 2000:1057-1063.
- [11] WOJKE N, BEWLEY A, PAULUS D. Simple online and realtime tracking with a deep association metric[C]// 2017 IEEE International Conference on Image Processing (ICIP). Beijing: IEEE, 2017:3645-3649.
- [12] HUANG C, WU B, NEVATIA R. Robust object tracking by hierarchical association of detection responses[C]// European Conference on Computer Vision (ECCV). Marseille, France: Springer-Verlag, 2008:788-801.
- [13] KRISTAN M, PFLUGFELDER R, LEONARDIS A, et al. The visual object tracking VOT2013 challenge results[C]// 2013 IEEE International Conference on Computer Vision (ICCV). Sydney, NSW, Australia: IEEE, 2013:98-111.
- [14] KRISTAN M, PFLUGFELDER R, LEONARDIS A, et al. The visual object tracking VOT2014 challenge results[C]// European Conference on Computer Vision (ECCV). Zurich, Switzerland: Springer-Verlag, 2014:191-217.

- [15] KRISTAN M, MATAS J, LEONARDIS A, et al. The visual object tracking vot2015 challenge results[C]// 2015 IEEE International Conference on Computer Vision Workshop (ICCVW). Santiago, Chile: IEEE, 2015: 1–23.
- [16] LEAL-TAIXÉL, MILAN A, REID I, et al. MOTChallenge 2015: towards a benchmark for multi-target tracking[J/OL]. arXiv, 2015: 1504.01942v1-1–15.
- [17] MILAN A, LEAL-TAIXÉL, REID I, et al. Mot16: a benchmark for multi-object tracking[J/OL]. arXiv, 2016: 1603.00831v2-1–12.
- [18] CHOI W. Near-online multi-target tracking with aggregated local flow descriptor[C]// 2015 IEEE International Conference on Computer Vision (ICCV). Santiago, Chile: IEEE, 2015: 3029–3037.
- [19] EUPER M, TANG S, ANDRES B, et al. Motion segmentation & multiple object tracking by correlation co-clustering[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(1): 140–153.
- [20] SADEGHIAN A, ALAHI A, SAVARESE S. Tracking the untrackable: learning to track multiple cues with long-term dependencies[C]// 2017 IEEE International Conference on Computer Vision (ICCV). Venice, Italy: IEEE, 2017: 300–311.
- [21] ZHOU H, OUYANG W, CHENG J, et al. Deep continuous conditional random fields with asymmetric inter-object constraints for online multi-object tracking[J]. IEEE Transactions on Circuits and Systems for Video Technology, 2018, 29(4): 1011–1022.
- [22] SONG Y, JEON M. Online multiple object tracking with the hierarchically adopted gm-phd filter using motion and appearance[C]// 2016 IEEE International Conference on Consumer Electronics-Asia (ICCE-Asia). Seoul, Korea (South): IEEE, 2016: 1–4.
- [23] KIM Chanho, LI Fuxin, REHG James M. Multi-object tracking with neural gating using bilinear LSTM[C]// The European Conference on Computer Vision (ECCV). Munich, Germany: Springer-Verlag, 2018: 200–215.

-----

(上接第 209 页)

- [56] WANG J, STANTCHEV R I, SUN Q, et al. THz in vivo measurements: the effects of pressure on skin reflectivity[J]. Biomedical Optics Express, 2018, 9(12): 6467–6476.
- [57] LINDLEY-HATCHER H, HERNANDEZ-SERRANO A I, SUN Q, et al. A robust protocol for in vivo THz skin measurements[J]. Journal of Infrared Millimeter Terahertz Waves, 2019(40): 980–989.
- [58] CHEN X, PARROTT E P J, HUANG Z, et al. Robust and accurate terahertz time-domain spectroscopic ellipsometry[J]. Photonics Research, 2018, 6(8): 768–775.
- [59] LEE A W M, WILLIAMS B S, KUMAR S, et al. Real-time imaging using a 4.3-THz quantum cascade laser and a 320/spl times/240 microbolometer focal-plane array[J]. IEEE Photonics Technology Letters, 2006, 18(13): 1415–1417.
- [60] STANTCHEV R I, PHILLIPS D B, HOBSON P, et al. Compressed sensing with near-field THz radiation[J]. Optica, 2017, 4(8): 989–992.
- [61] YU X, YANG F, GAO B, et al. Deep compressive single pixel imaging by reordering Hadamard basis: a comparative study[J]. IEEE Access, 2020(8): 55773–55784.
- [62] 白佳俊, 陈强, 陈亮, 等. 超材料压缩感知成像技术[J]. 太赫兹科学与电子信息学报, 2005, 13(4): 569–573. (BAI Jiajun, CHEN Qiang, CHEN Liang, et al. Compressive sensing imaging technology based on metamaterial[J]. Journal of Terahertz Science and Electronic Information Technology, 2005, 13(4): 569–573.)
- [63] GUERBOUKHA H, NALLAPPAN K, SKOROBOGATIY M. Exploiting k-space/frequency duality toward real-time terahertz imaging[J]. Optica, 2018, 5(2): 109–116.
- [64] HICHEM Guerboukha, KATHIRVEL Nallappan, MAKSIM Skorobogatiy. Toward real-time terahertz imaging[J]. Advances in Optics and Photonics, 2018, 10(4): 843–938.