

面向移动单摄像机的多目标跟踪算法*

李兴玮, 陈慧敏, 吕林珏, 关少杰

(国防科技大学 智能科学学院, 湖南 长沙 410073)

摘要:针对移动单摄像机采集的视频序列中的运动多目标,重点研究了基于目标间的相对运动信息和数据关联策略的在线多目标自动跟踪器。利用目标间相对运动模型实现目标轨迹的恢复,减少目标轨迹碎片。运用事件匹配算法改进当前帧的检测响应与过去轨迹的分配,并降低跟踪过程中的目标身份转换次数。实验结果表明:该改进算法较原算法能够对序列中目标跟踪定位得更加精确,减少了轨迹碎片和身份转换次数,在TUD-Campus序列上达到了与国际前沿多目标跟踪算法相当的效果。

关键词:相对运动模型;事件匹配算法;数据关联;移动单摄像机;多目标跟踪器

中图分类号:TN95 **文献标志码:**A **开放科学(资源服务)标识码(OSID):**

文章编号:1001-2486(2020)06-120-07



听语音
与作者互动
聊科研

Multi-object tracking algorithm for mobile single camera

LI Xingwei, CHEN Huimin, LYU Linjue, GUAN Shaojie

(College of Intelligence Science and Technology, National University of Defense Technology, Changsha 410073, China)

Abstract: Aiming at the moving objects in the video sequences collected by the mobile single camera, an online multi-object automatic tracker was focused on research, which was based on the relative motion information and data association strategies. The recovery of the object trajectory was achieved by the relative motion model between the objects, and the object trajectory fragmentation was reduced. The assignment between detection of the current frame and the past trajectories was improved based on the event matching algorithm, which reduced the number of identity conversions during tracking. Experimental results show that the improved algorithm is more accurate than the original algorithm in tracking and positioning the object in the sequences, reducing the number of trajectory fragmentation and identity conversion, and our tracker achieve relative the same performance in the state-of-the-art on the TUD-Campus sequence in the international academic circle.

Keywords: relative motion model; event matching algorithm; data association; mobile single camera; multi-object tracker

基于检测的在线多目标跟踪策略是利用检测器处理当前采集的视频帧,根据反馈的检测响应进一步提取感兴趣区域的视觉信息并与过去帧中的轨迹找寻最优匹配,整个过程相互迭代、更新^[1]。目标跟踪往往被分为基于批量处理和基于当前帧处理(即离线模型和在线模型)^[1-2]。离线模型^[3]利用先验信息和未来视觉信息建立轨迹全局最优解,存在较大的计算量和复杂性,因此选择符合实际场合实时性要求的在线模型。一直以来,行人跟踪在目标跟踪领域是一项亟须解决的技术,同时行人也是机器人辅助自动驾驶、导航系统中最受关注的目标。因此,本文重点研究并改进一种基于检测的在线多行人跟踪器。基于单目移动相机拍摄的视频序列的行人跟踪技术中,运动模型是区分多个相似外观行人的重要线索。Kim和Bae等^[4-5]分别提出了一阶、二阶运

动模型对目标进行位置预测,之后提出了匀变速模型^[6]和非匀变速模型^[7]。关于目标间的运动模型大部分都是相互独立的,这种情况会随着摄像机的运动造成不可挽回的累积误差,最终导致跟踪失败。因此,研究人员开始尝试在不同的目标之间建立联系。Leal-Taixé等^[8]考虑了行人之间的速度约束、避让斥力和群组拉力。王江峰等^[9]使用双向时空连续性迭代关联。Yoon等^[10]首次提出相对运动网络模型的概念,实验结果可以看出该模型在摄像机轻微抖动和运动下较好地保持了目标原本的轨迹。上述各类基于目标间相对运动信息的跟踪器在一定程度上缓解了摄像机移动和相似外观造成的跟踪模糊。主要考虑高级驾驶辅助系统和机器人导航等在线采集图像系统中的目标跟踪问题,重点研究基于多目标的空间结构信息^[7]来描述目标之间相对运动关系,融合

* 收稿日期:2019-11-05

作者简介:李兴玮(1969—),男,湖南岳阳人,教授,博士,硕士生导师,E-mail:simu_kwx@nudt.edu.cn

最新提出的基点匹配事件策略^[11],重新定义相对运动网络中的损失函数,并改进门限算法应用过程中的相似性度量。为了缓解检测器的误检问题,利用结构约束对漏跟踪目标进行轨迹恢复,使得整个算法鲁棒性更强。

1 模型描述

将视频多目标跟踪问题转换为分层数据关联的搜索问题。整个算法步骤为:输入检测响应,提取目标区域的特征并初始化目标之间的相对运动模型,类似相对运动网络^[10];关联成功的轨迹和新的检测响应,利用事件匹配算法进行数据关联;利用相对运动恢复漏跟踪目标的轨迹;更新目标状态和相对运动模型。下面介绍算法中所用的模型。在下面公式中忽略时间参数, \mathcal{S} 表示视频序列所有目标状态集, \mathcal{D} 表示检测响应集, \mathcal{A} 表示目标轨迹和检测响应的分配事件集。其中,目标轨迹标签集 $i \in N$,检测响应标签集 $k \in M$,当前帧目标之间的相对运动集 \mathcal{R} 。

1.1 相对运动模型

以图像左下角作为原点,起始帧目标速度为0。假设在第 t 秒的目标 i 状态为 $s_t^i = [x_t^i, y_t^i, x_t^i, y_t^i, \omega_t^i, h_t^i]^T$, $s_t^i \in \mathcal{S}_t$ 。其中, (x_t^i, y_t^i) 和 (x_t^i, y_t^i) 代表目标 i 中心的位置坐标和速度, ω_t^i 和 h_t^i 是检测响应的宽和高。定义目标 i 和 j 之间相对运动模型为:

$$r_t^{<i,j>} = [x_t^i - x_t^j, y_t^i - y_t^j, x_t^i - x_t^j, y_t^i - y_t^j]^T \quad (1)$$

则在第 t 秒的目标 i 相对运动集合 $R_t^i = \{r_t^{<i,j>} \mid \forall j \in N_t\}$ 。序列中所有目标的相对运动集合用 $\mathcal{R} = \bigcup_{i \in N} R^i (r_t^{<i,j>} \in R^i)$ 表示。

1.2 分配事件模型

把检测响应和目标轨迹之间的匹配方案定义为 $a^{i,k} \in \mathcal{A}$, 即 $\{a^{i,k} = 1\}$ 表示检测响应 k 被分配给目标轨迹 i , 反之亦然。检测响应集 $\mathcal{D} \{d^k \in \mathcal{D}\}$ 的标签集 $k \in M$, 要求每一个检测响应仅分配一条轨迹或生成新轨迹; 每一条轨迹仅接受一个检测响应或结束跟踪。可以通过损失成本函数计算目标轨迹与检测响应之间的相似程度函数, 然后通过最小化总匹配损失成本来估计最佳匹配方案。

$$\mathcal{A} = \operatorname{argmin} \mathcal{C}(\mathcal{S}, \mathcal{R}, \mathcal{D})$$

$$\text{s.t.} \sum_{i \in N, k \neq 0} a^{i,k} \leq 1 \cap \sum_{k \in M \cup \{0\}} a^{i,k} = 1 \cap \sum_{i \in N} a^{i,0} \leq |N| \quad (2)$$

式中, $a^{i,0}$ 为一个误检目标的匹配作业。因此, 当视频序列中所有的目标 i 没有被检测器检测出来的

时候, 此时 $a^{i,0}$ 个数总和等于目标的个数 $|N|$ 。

1.3 损失成本模型

目标的轨迹是以目标在视频序列中按照时间变化上的位置、速度和大小的状态向量表示。在线多目标跟踪的目的是将当前帧的检测响应同当前的轨迹进行匹配, 或终止轨迹或生成新的轨迹, 即寻求最优匹配策略。在当前匹配事件 $a^{i,k}$ 下, 结构约束损失函数为:

$$\mathcal{C}(\mathcal{S}, \mathcal{A}, \mathcal{R}, \mathcal{D}) \big|_{a^{i,k}} = a^{i,k} \Omega_{i,k} + \sum_{j \in N, j \neq i} \sum_{q \in M \cup \{0\}, q \neq k} a^{j,q} \Theta_{i,k}^{j,q} \quad (3)$$

式中, j 是目标 i 相对运动模型中的联结目标, q 是检测响应。总损失成本包括基准成本 $\Omega_{i,k}$ 和连接项分配成本 $\Theta_{i,k}^{j,q}$ 。

$$\Omega_{i,k} = F_s(s^i, d^k) + F_a(s^i, d^k) \quad (4)$$

$$\Theta_{i,k}^{j,q} = \begin{cases} F_s(s^j, d^q) + F_a(s^j, d^q) + \\ F_c(s^j, r^{<j,i>}, d^k, d^q) & q \neq 0 \\ \tau & q = 0 \end{cases} \quad (5)$$

通过计算目标和检测响应的巴氏距离来估计尺寸损失成本 $F_s(s, d)$ 。采用颜色直方图的多色观测模型估计外观损失成本 $F_a(s, d)$ 。通过连接检测边界框和目标跟踪边界框的重叠率来估计位置损失成本 $F_c(s^i, r^{<j,i>}, d^k, d^q)$ 。

2 数据关联算法

2.1 结构约束事件匹配算法

当目标之间的分配事件 $a^{i,k} = 1$ 为真, 定义目标之间的分配事件集、结构信息集和先前帧的目标状态集分别为 \mathcal{A}^o 、 \mathcal{R}^o 和 \mathcal{S}^o 。在当前检测集 \mathcal{D} 条件下, 利用结构约束事件匹配算法将正确匹配的相关检测位置添加至新的匹配事件集 $\hat{\mathcal{A}}^o$ 中。其中, 目标的标签集 $i \in N^o$ 。

为了更好地进行下一步的匹配作业, 设置一个基准匹配, 基于结构约束对剩下的目标和检测结果进行匹配作业, 匹配过程详见文献[11]。由于可以根据相应基准匹配的数量多次测量每个分配作业的成本, 这个过程可以显著减少由于目标附近的误报、错误检测和约束错误导致的歧义。

2.2 漏跟踪目标恢复算法

使用结构约束事件匹配算法的跟踪目标 \mathcal{S}_t^o 结构约束信息进行漏跟踪目标的恢复。类似地定义漏跟踪的目标状态集和对应的相对运动集分别为 \mathcal{S}^m 和 \mathcal{R}^m , N^m 和 M^m 分别是漏跟踪目标的标签集和未匹配的漏跟踪虚拟检测集 \mathcal{D}^m 的标签集。

尺寸损失函数和外观损失函数与结构约束事件匹配算法^[11]一致。根据过去的运动一致性原则,位置损失函数中为了保证相对运动的可靠性,选择相对运动网络中运动速率和方向最相似的目标来计算重叠率。其他阈值定义都与相对运动事件匹配算法中的一致。

3 多目标跟踪框架

充分利用目标之间的相对运动信息改进数据关联框架。当目标匹配成功且出现新的检测响应时,采用结构约束事件匹配算法,否则采用原贝叶斯框架^[10]进行数据关联。在算法 1 中描述了在线多目标跟踪的整个过程。

算法 1 在线多目标跟踪框架

Alg. 1 Online MOT framework

输入: 视频序列和序列检测集 D

输出: 序列中每个目标的编号和轨迹信息

跟踪器 $tracker_{video} = \phi$,

For 视频序列帧 $f=1$ to F

- 预处理: 当前 f 帧的检测集 D_f , 跟踪目标状态集 S_f^o , 跟踪目标之间的结构约束 R_f , 漏跟踪目标状态集 S_f^m , 跟踪目标和漏跟踪目标之间的结构约束 R_f^m

• If (\sim isempty 检测响应 d) $\cap a^{i,k} = 1$

1. 利用结构约束事件匹配算法进行数据关联

$$\hat{A}^o = SCEA(S^o, R^o, D)$$

2. 初始化可关联的目标状态

$$S_1^o = \{s_1^i = [x_d^k, y_d^k]^T \mid a^{i,k} = 1\};$$

$$\forall i \in N^o, \forall k \in M$$

3. 当前帧检测响应和目标状态集利用卡尔曼滤波 $KF(\cdot)$ 更新目标状态集

$$S^o = \{s^i : = KF(s^i, d^k) \mid a^{i,k} = 1\},$$

$$\forall i \in N^o, \forall k \in M$$

4. 漏跟踪目标的恢复算法

$$\hat{A}^m = Recovery(S^m, R^m, S_1^o, D^m);$$

$$\hat{A} = \hat{A}^o \cup \hat{A}^m$$

Else

5. $\hat{A} = \text{Bayesian}(S, R, D)$

End

- 基于结构约束的相对运动网络的更新
使用转换矩阵和观测模型更新每个相对运动, 给定的检测响应通过卡尔曼滤波器进行更新预测
- 移除已正确关联的检测响应
- 轨迹目标和目标间结构运动模型的新建和终止

End

由于摄像机的移动, 目标间相对运动模型的

方向和速度是动态的, 采用分段线性恒速模型对时变运动进行建模。其中, 干扰假定为高斯白噪声。两个检测响应的位置差定义为观测量并使用卡尔曼滤波器^[12]更新相对运动集 R 。漏跟踪目标使用匀速运动模型进行简单预测。

目标的新建和终止与 Breitenstein 等^[13]提出的目标初始化方法类似, 即使用两次检测响应之间的重叠距离阈值判断方法进行新目标轨迹以及对应的相对运动模型的初始化和终止。

4 实验

4.1 实验平台

该实验在配置为 Inter (R) Core TM i5 - 3230M CPU @ 2.60GHz 处理器、4 GB 内存、64 位 Windows 操作系统的笔记本电脑上运行。下载 MOT2015 竞赛^[14]中的公开视频序列和 Xiang 等^[15]提供的检测响应进行测试, 采用的数据集的图像属性与相机配置等如表 1 所示。表中的密度表示序列中行人的稠密程度, 值越大代表行人稠密程度越大。

表 1 MOT2015 应用的序列介绍

Tab. 1 Overview of the sequences in MOT2015

编号	序列	帧数	摄像机	密度
1	TUD-Campus	71	静止	5.1
2	ETH-Bahnhof	1000	移动	5.4
3	ETH-Pedcross2	840	移动	7.5
4	KITTI-13	340	移动	2.2
5	Venice-2	600	静止	11.9

4.2 评价方法

采用了 Bernardin 等^[16]设计的针对多目标跟踪技术的 CLEAR MOT 评价体系, 是将实验结果与原图像的实际轨迹按照指标求解算法得到最终结果。参考的指标内容如表 2 所示。表 2 列举了文献[16]中的几个主要性能指标, 文中性能对比采用的指标根据实际情况进行必要的删减和添加。多目标跟踪准确度 (Multiple Object Tracking Accuracy, MOTA) 和多目标跟踪精确度 (Multiple Object Tracking Precision, MOTP) 是本文最为关注的两项指标。

MOTA 代表了多目标跟踪性能的总体指标。

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

式中, FN_t 、 FP_t 和 $IDsw_t$ 分别代表了轨迹在第 t 帧的错误匹配次数、错误检测次数和身份转化次数, GT_t 是地表实值中真实目标轨迹的总数。

$MOTP$ 描述了跟踪结果与地表实值之间的重叠比例。

$$MOTP = \frac{\sum_{i,t} d_i^i}{\sum_t c_t}$$

式中, c_t 是在第 t 帧的成功匹配目标数量, d_i^i 是目标 i 在 t 帧输出的目标跟踪框与文件注释边界框的面积重叠率, 求和代表了计算所有帧中目标的位置误差。

表 2 CLEAR MOT 主要评价指标

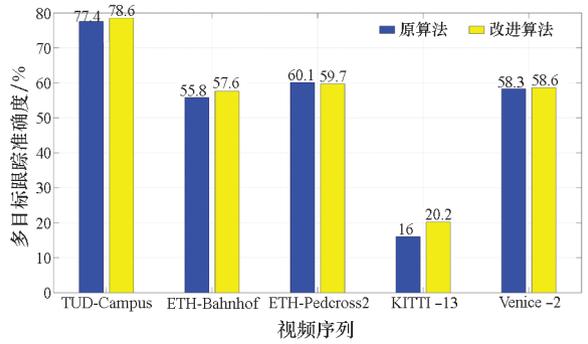
Tab.2 CLEAR MOT evaluation index

名称	含义	期望
跟踪准确度 $MOTA$	各类跟踪误差的整合指标	高
跟踪精确度 $MOTP$	位置估计的精确度	高
身份转换数 $IDsw$	目标标记转换次数	低
跟踪大轨迹数 MT	持续准确跟踪到的目标数	高
丢失轨迹数 ML	没有准确跟踪到的目标数	低

4.3 实验一

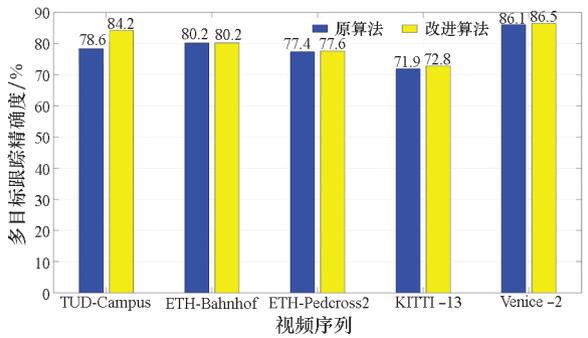
选取移动摄像机从不同视角拍摄的 ETH-Bahnhof、ETH-Pedcross2 和 KITTI - 13 视频序列^[17], 为了验证算法的普适性, 同样选择了两组由静止摄像机拍摄的 TUD-Campus 和 Venice - 2 视频序列。为了直观显示出改进算法框架较原来算法在总体性能上提高的效果, 对 $MOTA$ 、 $MOTP$ 和 $IDsw$ 三大性能指标进行分析, 如图 1 所示。从图 1 中可以看出, 改进算法可以在保持高跟踪精度的前提下, 提高跟踪准确度, 比如 TUD-Campus 序列 $MOTA$ 指标从 77.4% 提高至 78.6%; ETH-Bahnhof 序列的 $MOTA$ 指标从 55.8% 提高至 57.6%; 特别是镜头移动幅度大、遮挡严重的 KITTI - 13 序列的 $MOTA$ 指标提高了 26%。这也间接说明算法在静止摄像机拍摄的视频中具有优良性能。身份转换次数是衡量目标跟踪过程中保持目标身份完整性的指标。从图 1 (c) 可以看出, 改进算法大幅度地减少了 $IDsw$, 特别是在序列 ETH-Pedcross2 和 ETH-Bahnhof 中, $IDsw$ 值降低超过了 35%, 有利于提高跟踪识别的准确性。同时也可以看出, KITTI - 13 中 $IDsw$ 值是增加的。

为了进一步考察算法的实时性效率, 以算法执行总时长 (单位为 s) 和帧率 (单位为 帧/s) 来进行效率对比。基于一致的响应集, 多次测量取平均估计值, 最终结果如表 3 所示。



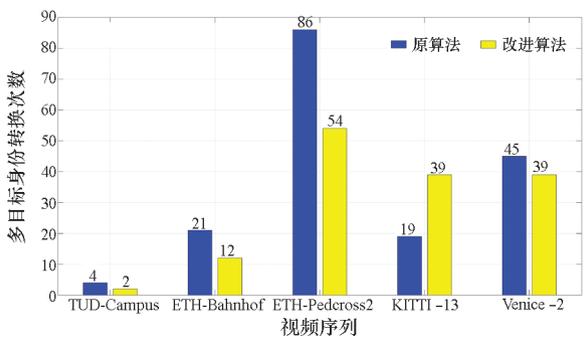
(a) 不同序列的跟踪准确度

(a) $MOTA$ for different sequences



(b) 不同序列的跟踪精确度

(b) $MOTP$ for different sequences



(c) 不同序列的身份转换次数

(c) $IDsw$ for different sequences

图 1 改进算法和原算法的对比

Fig.1 Comparison between improved algorithm and original algorithm

表 3 算法的运行效率分析

Tab.3 Analysis of the operation efficiency of the algorithm

视频序列	密度/(目标数/帧)	运行时间/s	帧率/(帧/s)
TUD-Campus	5.1	3.986 4	17.809 2
ETH-Bahnhof	5.4	22.915 3	43.638 4
ETH-Pedcross2	7.5	4.675 7	178.99 6
KITTI - 13	2.2	20.643 4	16.469 9
Venice - 2	11.9	0.8107 6	739.808 3

从表 3 的数据可以看出,算法在五段序列下的总运行时间达到 50 s 左右,平均帧率约为 25 帧/s。序列时长较短和硬件平台内存运行不稳定等因素导致个别序列与总体估计值存在较大误差。本节展示了各个序列的平均行人密度,可以大致认为密度越大的序列下,算法运行的速度越快。增大视频集的数据量,稳定算法的最终帧率,也是下一步工作重点。

4.4 实验二

针对应用最广的 TUD-Campus 视频序列,算法的最终序列跟踪结果如图 2 所示,对比实验的

量化指标参考表 4。为了更加直观地展示实验结果,一共选择 5 种离线算法和 3 种在线算法的最终结果进行对比。

针对序列 TUD-Campus,表 4 中的加粗部分是该指标下最好的结果。可以看出:文献[21]的离线算法的 *MOTA* 指标最优,利用未来帧信息的离线算法在数据关联中比在线方案更容易得到全局最优解。本文在线算法的 *MOTP* 指标最优。本文算法的其他指标与其他算法接近,缓解了序列中轨迹碎片的问题,较好地保持了目标的完整轨迹。图 2 显示了序列上的跟踪情况。目标 1 在整



图 2 在序列 TUD-Campus 的跟踪效果

Fig.2 Result of TUD-Campus sequence

表 4 在 TUD-Campus 序列的不同跟踪结果
Tab.4 Tracking result on TUD-Campus sequence

方法	<i>Rcll</i> /%	<i>Prcsn</i> /%	<i>FAR</i>	<i>GT</i>	<i>MT</i>	<i>PT</i>	<i>ML</i>	<i>FP</i>	<i>FN</i>	<i>ID_{sw}</i>	<i>FM</i>	<i>MOTA</i> /%	<i>MOTP</i> /%	<i>MOTAL</i>
文献[18]	50.4	57.5	1.89	8	3	2	3	134	178	3	11	12.3	70.1	12.9
文献[19]	32.6	82.4	0.35	8	1	3	4	25	242	0	1	25.6	72.9	25.6
文献[20]	82.5	93.7	0.28	8	5	2	1	20	63	3	2	76.0	77.4	76.7
文献[21]	87.5	98.4	0.07	8	5	2	1	5	45	1	0	85.8	77.0	86.0
文献[22]	66.6	95.5	0.15	8	3	4	1	11	118	10	13	60.6	78.2	63.2
文献[13]										2		67.0	73.0	
文献[23]					5					2	3	82.0	74.0	
文献[24]	83.8	99.3	0.03	8	5	2	1	2	58	0	1	83.3	76.9	83.3
本文	81.3	97.3	0.11	8	4	4	0	8	67	2	2	78.6	84.2	79.0

个序列中保持了完整的轨迹和唯一的标记。在目标4横穿目标3、5和6的过程中,目标3、4、5和6都很好地保持了原本的身份标记。特别是在第19帧,正确跟踪到了相互遮挡严重的目标1、4和5。对于一开始遮挡严重的行人,跟踪算法存在一定的问题,比如目标7开始时并没有被跟踪上,只有在第28帧的时候才开始跟踪。目标8在第55帧才开始跟踪。这一方面说明了算法对起始遮挡就很严重的目标不敏感,另一方面也说明了算法在后期处理遮挡、漏跟踪有很大的作用。

5 结论

视频中多目标间的运动模型受到摄像机的运动影响变得难以预测。一个鲁棒的多行人跟踪器除了需要估计未知数量的视频中感兴趣的目标及其各自独有的轨迹,还应在漏跟踪和误检的时候能够恢复,减少检测失败造成的跟踪失败。本文重点研究了由目标间的位置和速度搭建的相对运动网络,利用相对运动对漏跟踪目标进行恢复,采用分层数据关联的方法对分配事件进行管理。利用恢复漏跟踪目标算法对算法进行优化,提出改进的数据关联解决方案,以更详尽的方式估计目标与检测之间的分配。为了验证所提数据关联算法的有效性,利用了接受度最广、使用率最高的公开数据集进行实验验证并取得了较好的实验结果。该框架的跟踪性能较原始框架有了很大的提高并且可以和先进的离线跟踪算法媲美。

分析实验效果可知,本算法针对摄像机视野中的大目标具有良好的普适性,相对运动模型可以很好地解决部分遮挡和交叉问题。然而,也正由于该模型的存在,当出现误检目标时,如果没有即时消除该目标的相对运动模型会造成下一帧的其他目标跟踪失败。该算法容易丢失序列中尺度差距较大的目标。因此下一步的主要工作是修正相对运动模型,考虑过去帧的信息对跟踪模糊的目标进行更新,并继续完善损失函数中的相似度量,使得算法可以鲁棒地跟踪尺度差距明显的目标和小目标。

本文对算法的跟踪准确度和精确度做重点研究。受硬件平台影响未对实时性做重点评估。往往高精度需要降低运行效率作为代价。基于有限的平台下,根据实际需求评估算法的准确度和实时要求的平衡问题也是不容忽视的。因此实时性重要考察指标也将是下一步工作的重点。

参考文献 (References)

[1] 吴慧玲. 基于检测的在线多目标跟踪算法研究[D]. 合

肥: 中国科学技术大学, 2017.

WU Huiling. Research on detection based online multi-object tracking[D]. Hefei: University of Science and Technology of China, 2017. (in Chinese)

- [2] 罗鹏飞, 张文明. 一种多目标跟踪航迹起始新算法及其性能评估[J]. 国防科技大学学报, 1999, 21(6): 51-54.
- LUO Pengfei, ZHANG Wenming. A new algorithm for track start of multi-target tracking and its performance evaluation [J]. Journal of National University of Defense Technology, 1999, 21(6): 51-54. (in Chinese)
- [3] Bochinski E, Eiselein V, Sikora T. High-speed tracking-by-detection without using image information [C]// Proceedings of the 14th IEEE International Conference on Advanced Video and Signal based Surveillance (AVSS), 2017: 1-6.
- [4] Kim S, Kwak S, Feyereisl J, et al. Online multi-target tracking by large margin structured learning [C]// Proceedings of Asian Conference on Computer Vision, 2012: 98-111.
- [5] Bae S H, Yoon K J. Robust online multi-object tracking based on tracklet confidence and online discriminative appearance learning [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 1218-1225.
- [6] Milan A, Roth S, Schindler K, et al. Privacy preserving multi-target tracking [C]// Proceedings of Asian Conference on Computer Vision, 2014: 519-530.
- [7] Yang B, Nevatia R. An online learned CRF model for multi-target tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2012: 2034-2041.
- [8] Leal-Taixé L, Pons-Moll G, Rosenhahn B. Branch-and-price global optimization for multi-view multi-target tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2012: 1987-1994.
- [9] 王江峰, 张茂军, 包卫东, 等. 双向时空连续性轨迹片段关联的目标跟踪方法[J]. 国防科技大学学报, 2011, 33(2): 44-48.
- WANG Jiangfeng, ZHANG Maojun, BAO Weidong, et al. Object tracking using bidirectional spatio-temporal continuity tracklet association [J]. Journal of National University of Defense Technology, 2011, 33(2): 44-48. (in Chinese)
- [10] Yoon J H, Yang M H, Lim J, et al. Bayesian multi-object tracking using motion context from multiple objects [C]// Proceedings of the IEEE Winter Conference on Applications of Computer Vision, 2015: 33-40.
- [11] Yoon J H, Lee C R, Yang M H, et al. Computer Vision & Pattern Recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1392-1400.
- [12] Li X R, Jilkov V P. Survey of maneuvering target tracking. Part I. Dynamic models [J]. IEEE Transactions on Aerospace and Electronic Systems, 2003, 39(4): 1333-1364.
- [13] Breitenstein M D, Reichlin F, Leibe B, et al. Online multiperson tracking-by-detection from a single, uncalibrated

- camera [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(9): 1820 – 1833.
- [14] Leal-Taixé L, Milan A, Reid I, et al. MOT challenge 2015: towards a benchmark for multi-target tracking [J/OL]. Computer Science, 2015 [2019 - 09 - 28]. <https://arxiv.org/abs/1504.01942>.
- [15] Xiang Y, Choi W, Lin Y, et al. Subcategory-aware convolutional neural networks for object proposals and detection [C]// Proceedings of IEEE Winter Conference on Applications of Computer Vision, USA, 2017.
- [16] Bernardin K, Stiefelham R. Evaluating multiple object tracking performance: the CLEAR MOT metrics [J]. EURASIP Journal on Image and Video Processing, 2008(1): 246309.
- [17] Xiang Y, Alahi A, Savarese S. Learning to track: online multi-object tracking by decision making [C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 4705 – 4713.
- [18] Fragkiadaki K, Zhang W, Zhang G, et al. Two-granularity tracking: mediating trajectory and detection graphs for tracking under occlusions [C]// Proceedings of the European Conference on Computer Vision, 2012: 552 – 565.
- [19] Milan A, Leal-Taixé L, Schindler K, et al. Joint tracking and segmentation of multiple targets [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 5397 – 5406.
- [20] Keuper M, Andres B, Brox T. Motion trajectory segmentation via minimum cost multicut [C]// Proceedings of the IEEE International Conference on Computer Vision, 2015: 3271 – 3279.
- [21] Keuper M, Tang S, Yu Z, et al. A multi-cut formulation for joint segmentation and tracking of multiple objects [J/OL]. Computer Science, 2016 [2019 - 09 - 28]. <https://arxiv.org/abs/1607.06317>.
- [22] Pirsiavash H, Ramanan D, Fowlkes C C. Globally-optimal greedy algorithms for tracking a variable number of objects [C]// Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2011: 1201 – 1208.
- [23] Segal A V, Reid I. Latent data association: Bayesian model selection for multi-target tracking [C]// Proceedings of the IEEE International Conference on Computer Vision, 2013: 2904 – 2911.
- [24] Tang S, Andres B, Andriluka M, et al. Subgraph decomposition for multi-target tracking [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 5033 – 5041.