

基于多信息融合的多目标跟踪方法研究

张 静¹, 王文杰²

(1. 中北大学 软件学院, 太原 030051; 2. 湖北江山重工有限责任公司, 湖北 襄阳 441057)

摘要: 在线多目标跟踪作为计算机视觉和人工智能方面的一个研究热点, 随着深度学习的发展取得了较大的进展; 但是依然存在诸如复杂场景跟踪准确度低等亟待解决的问题; 针对多目标跟踪研究中存在的行人特征信息较少、跟踪目标被遮挡等问题, 提出了一种融合表观信息、轨迹历史信息和目标运动信息的多目标跟踪方法, 通过专门设计的双分支网络结构和损失函数使模型在学习时将三种信息相互融合; 改进相似性分数计算方法获得更多的特征信息, 提取更为鲁棒的特征; 多信息融合的多目标跟踪方法在计算方面开销较少, 能够在测试时达到实时的效果; 并且, 通过相关实验验证, 基于多信息融合的多目标跟踪方法能够在 MOT16 数据集上达到很好的性能, 可以更好地处理目标遮挡、目标误检及目标丢失等情况。

关键词: 计算机视觉; 深度学习; 多目标跟踪; 目标遮挡; 双分支网络

Multi-target Tracking Method Based on Multi Information Fusion

Zhang Jing¹, Wang Wenjie²

(1. School of Software, North University of China, Taiyuan 030051, China;

2. Hubei Jiangshan Heavy Industries Co., Ltd., Xiangyang 441057, China)

Abstract: As a research hotspot in computer vision and artificial intelligence, online multi-target tracking has made great progress with the development of deep learning in recent years. However, there are still many problems to be solved, such as low tracking accuracy of complex scenes. Aiming at the problems of pedestrian feature information in multi-target tracking research and occlusion of tracking targets, a multi-target tracking method is proposed to combine apparent information, trajectory historical information trajectory historical information and target motion information. The designed two branch network structure and loss function enable the model to fuse these information with each other during learning to obtain more feature information and extract more robust features, which improves the similarity score calculation method to obtain more feature information and extract more robust features. The proposed method has less computational overhead and can achieve realtime effects during testing. In addition, relevant experimental results show that the proposed method can achieve good performance on MOT16 and other data sets, which can better deal with target occlusion, target misdetection, and target loss.

Keywords: computer vision; deep learning; multi-target tracking; target occlusion; two branch network

0 引言

随着社会的发展, 多目标跟踪作为目标跟踪领域中重要的研究方向广泛应用于城市数字化管理、智能安防和自动驾驶等领域^[1]。多目标跟踪 (Multi-Object Tracking) 研究的主要目的在于通过对视频帧的检测分析, 找到多个感兴趣的目標并在整个视频中生成并维护这些目标的轨迹^[2]。常用的多目标跟踪方法可以分为离线方法和在线方法, 离线方法通过历史帧和当前帧的信息生成跟踪轨迹, 可以利用的信息更多, 效果更好, 但不适用于实时的计算机视觉任务。同时, 随着深度学习网络性能的提高, 目标检测方法也取得了很大的进步。因此, 在目前的研究大多是基于检测的跟踪 (Tracking by detection) 框架。

多目标跟踪中常用的基于检测的跟踪框架能够在大多数应用场景中取得很好的效果, 但是, 这种框架比较依赖于提取目标特征的质量, 当提取的目标特征鲁棒性较差或出现错误时, 就会出现 ID 交换或轨迹丢失的情况, 严重影响

最终得到的跟踪效果。在之前的研究中, 目标和轨迹的特征通常用传统方法的颜色直方图、HOG 特征、SIFT 特征等方法来描述。但是在处理遮挡、旋转、姿态变化等问题时都有一定的缺陷。如文献 [3] 中提到的 SORT 方法是一个应用较为广泛的多目标跟踪方法, 通过匈牙利算法和卡尔曼滤波器处理得到目标轨迹, 但是由于仅使用了传统特征作为目标特征, 所以跟踪效果较差。而文献 [4] 中提出的 Deep SORT 方法, 在 SORT 方法的基础上使用深度神经网络提取目标特征, 获得了较好的目标特征, 提升了跟踪效果。在文献 [5] 中提出的跟踪方法利用基本的深度学习网络提取特征, 同样极大地提升了跟踪效果。文献 [6] 中通过修改特征提取网络中的骨干网络, 提高了特征的鲁棒性, 在此基础上得到了更好的目标轨迹。文献 [7] 则使用 GoogleNet 的修改版网络, 并在自定义的目标重识别数据集上进行训练, 结合表观特征和空间特征, 也取得了很好的跟踪效果。但是, 在处理复杂场景下的多目标跟踪问题时, 仍然会发生轨迹偏移、轨迹丢失和 ID 交换等现象。主要是因为多目标跟踪处理的场景视频帧, 并且目标一般是在不断运动的, 而当前的研究主要关注于当前帧中目标的特征信息, 这样的信息是单独的、静态的并不适用于处

收稿日期: 2020-06-03; 修回日期: 2020-07-07。

作者简介: 张 静 (1980-), 女, 山西运城人, 博士, 讲师, 主要从事计算机视觉、数据挖掘方向的研究。

理复杂场景。同时在计算轨迹和目标的相似性分数时也没有考虑到轨迹中不同帧的图像和当前帧检测图像的相似程度是会受时间影响的。因此本文将对特征提取网络和相似性分数计算方法进行修改，并选用目前表现较好的 DenseNet^[8] 网络作为骨干网络，提取更为鲁棒的特征，更好的处理复杂场景中的跟踪问题。

同时检测方法作为跟踪框架的基础，影响着后续的跟踪质量。在文献 [9] 中提出的跟踪方法使用的是 YOLO^[10] 检测方法，虽然其检测速度较快，但是牺牲了很大一部分的检测精度，导致其在复杂场景中无法准确的检测到跟踪的目标。而在文献 [11] 中提出的多目标跟踪方法则使用了 Mask R-CNN^[12] 作为目标检测方法，牺牲了一部分的速度性能，但准确地检测出了跟踪的目标，提升了跟踪效果。说明在多目标跟踪场景中，并不需要速度很快的检测方法，而是需要能够准确检测目标的检测方法。因此本文使用 Mask R-CNN 检测方法作为框架中的目标检测方法。

因此，针对上面分析的问题，本文从实际监控场景出发，提出了一种融合了历史信息 and 运动信息的基于检测的多目标跟踪框架 (HM-TBD)，主要贡献点如下：1) 设计了一种双分支网络用于联合学习历史信息 and 运动信息；2) 为了得到更准确的相似性分数设计了一种新的相似性计算方式；3) 通过在 MOT16 数据集上的实验，本文提出的 HM-TBD 方法能够达到很好的效果。

1 目标跟踪方法

本文提出的基于检测的多目标跟踪框架主要包括四个部分：目标检测、目标特征提取、相似性分数计算和轨迹关联，如图 1 所示。



图 1 融合历史信息和运动信息的特征提取网络

在本文的研究过程中选择通过文献 [12] 中提出的 Mask R-CNN 检测方法对视频帧中的目标进行检测，在得到目标的检测结果之后，通过本文设计的网络提取更为鲁棒的特征描述符，之后计算检测和轨迹之间的相似性分数，构建相似度矩阵，最后利用匈牙利算法^[13] 进行轨迹更新，得到目标的轨迹信息。下面将对本文所使用的相关方法和提出的双分支网络进行详细说明。

1.1 目标检测方法

Mask R-CNN 是何凯明等人在 2017 年提出的一种新的目标检测算法，不仅训练速度快，并且能够获得很好的检测结果。Mask R-CNN 整体上沿用了 Faster R-CNN^[14] 的检测思想，采用相同的两步检测策略，并在其中添加了一个 Mask 预测分支。具体来说当视频帧输入后进行目标检测时，首先 Mask R-CNN 会通过其骨干网络 ResNet-FPN 提取用于检测的特征，之后通过 Mask R-CNN 中的 RPN 子网络初步得到目标的坐标信息和分类信息，再通过 RoI pooling 处理之后对目标进行包围框回归和分类，最后得到用于跟踪的目标的位置信息。Mask R-

CNN 的损失函数如式 (1) 所示：

$$L = L_{CLS} + L_{BOX} + L_{MASK} \tag{1}$$

其中： L_{CLS} 使用于约束目标分类的损失函数， L_{BOX} 使用于学习目标包围框回归的损失函数， L_{MASK} 则是用于学习目标 Mask 的损失函数。

1.2 特征提取方法

在得到视频帧中目标的位置信息后，就可以得到需要跟踪的目标图像，这时就需要对目标的特征进行提取了。在之前的大多数研究中都忽略了多目标跟踪处理的场景是一个不断运动的视频序列，因此只关注于单张图像中所包含的信息是不够的，这样提取出来的特征在面对复杂场景时的鲁棒性不高，将会导致跟踪效果较差。针对这一问题，本文根据多目标跟踪中常见的场景设计了一种融合历史信息 and 运动信息的双分支网络 (HM-Net)，用于提取更为鲁棒的特征计算相似性分数。本文提取特征的整体框架如图 2 所示。

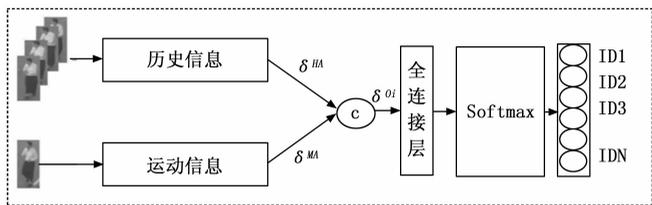


图 2 融合历史信息和运动信息的特征提取网络

从图 2 中可以看出提取特征的整体框架包括上下两个分支，其中上分支用于提取跟踪目标的历史信息，下分支用于提取跟踪目标的运动信息。上下两个分支独立生成用于向后传播的特征向量 δ^{HI} 和 δ^{MI} ，并将两个分支中得到的特征向量进行融合得到融合了历史信息 and 运动信息的目标特征向量 δ^O ，之后通过全连接层 (FC 层) 处理，将特征向量送入到 Softmax 层得到目标 ID 的预测概率。下面将详细说明提取历史信息 and 运动信息的两个分支网络的具体结构。

在得到跟踪目标的图像之后，为了提取跟踪目标中可以利用的历史信息，本文设计了基于 LSTM 的上分支网络。上分支网络的结构如图 3 所示。

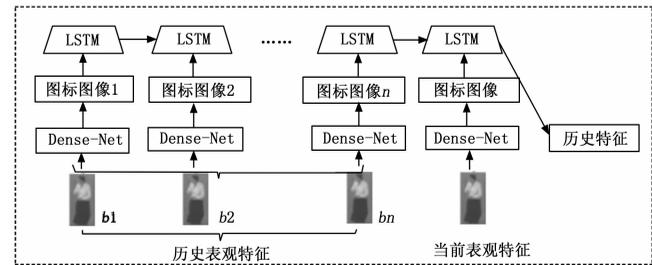


图 3 历史信息特征提取分支网络

以历史帧中的 N 张目标图像和当前帧中检测到的目标图像作为输入，通过骨干网络 Dense-Net 提取目标的表观特征，之后将提取到的 N 个历史表观特征和当前表观特征按照时间顺序输入到 LSTM 网络中用以提取目标具有区分性的运动特征，并将 LSTM 网络中最后一个隐藏层的特征向量作为目标的历史特征，用于和其他分支提取出的特征

进行融合预测目标的 ID, 进行学习。

提取目标的历史信息后, 为了提取目标的运动信息, 本文设计了基于目标关节点的下分支网络和运动提取器模块, 网络的结构如图 4 中所示。

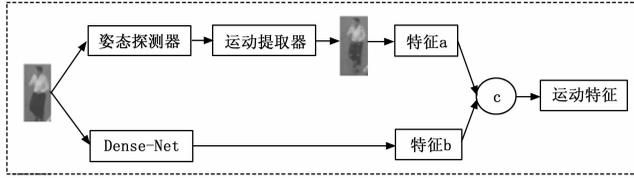


图 4 运动信息特征提取分支网络

提取运动特征时如图 4 中所示, 首先将当前帧检测到的目标图像输入到下分支网络中, 通过姿态探测器和运动提取器得到特征 a 将其作为目标的运动特征。具体来说, 本文选用 Alpha Pose 网络作为姿态探测器提取目标的关节点信息, 并用式 (2) 作为运动提取器计算目标的运动特征。

$$F_{ik}^{pm(t)} = (X_{ik}^{pm(t-1)} - X_{ik}^{pm(t)}, y_{ik}^{pm(t-1)} - y_{ik}^{pm(t)}) \quad (2)$$

其中: $F_{ik}^{pm(t)}$ 表示第 i 目标的第 k 个关节点在当前第 t 帧的特征信息, $(X_{ik}^{pm(t-1)}, y_{ik}^{pm(t-1)})$ 表示第 i 个目标的第 k 个关节点在第 $t-1$ 帧的坐标信息。之所以提取运动特征是因为在复杂场景中不同目标的运动信息差异较大, 这样提取到的特征区分性和鲁棒性较高, 计算得到的目标和轨迹间的相似性分数更准确, 可以更好地跟踪目标。最后将提取好的运动特征和通过 Dense-Net 提取出的表现特征进行融合, 作为目标最终的运动特征 δ^{MA} 。通过上面介绍的两个分支, 分别得到了目标的历史特征和运动特征, 并通过 Softmax 层计算得到目标的分类 ID。最后为了能够更好的训练网络以提取更好的特征, 本文选用交叉熵函数作为损失函数, 用来监督网络的学习, 如式 (3) 所示:

$$Loss = - \sum_k t_k \log y_k \quad (3)$$

其中: k 为训练时输入网络的目标图像数量, t 为目标图像的真实 ID, y 为目标图像的预测 ID。

1.3 相似性计算方法

多目标跟踪框架中的第三步为计算轨迹和当前帧检测图像之间的相似性分数。在之前的研究中, 计算相似性分数的方式是在轨迹中选取 n 张图像提取特征, 再分别计算它们与检测目标特征之间的距离进行相加, 将最后的距离作为相似性分数。这样做的缺点是没有考虑到轨迹是不断运动的, 时间间隔越短的轨迹图像会和目标图像越相似。因此, 本文提出增加权重的相似性计算方法。第一步计算轨迹中单个图像和检测目标之间的相似性分数, 本文选用余弦距离作为单个轨迹和图像的相似性分数, 计算方法如式 (4) 所示:

$$Score_{ji}^{track_k} = F_i^{track_k} \cdot F_j^{det} / \|F_i^{track_k}\| \cdot \|F_j^{det}\| \quad (4)$$

其中: $F_i^{track_k}$ 为第 k 个轨迹中第 i 个图像的特征向量, F_j^{det} 是当前帧中的第 j 个检测目标的特征向量, $Score_{ji}^{track_k}$ 表示第 k 个轨迹中的第 i 个图像和当前帧中第 j 个检测目标之间的相似性分数。假设一个轨迹有 n 张图像, 则在得到第 k 个轨迹中

所有图像和第 j 个检测目标之间的相似性分数之后, 先将分数按照视频帧的先后顺序进行排序 $Score_1, Score_2 \dots Score_n$, 之后再通过下面的式 (5) 计算带权重的相似性分数, 作为最终的相似性分数:

$$Score_j^{track_k} = \sum_i^n \frac{n * Score_i}{(1 + 2 + \dots + n)} \quad (5)$$

其中: $Score_j^{track_k}$ 表示第 k 个轨迹和第 j 个检测目标之间的最终相似性分数, $Score_i$ 则表示第 k 个轨迹中的第 i 个图像和第 j 个检测目标之间的相似性分数。从公式中可以看出当 n 越小时轨迹图像的帧号越接近于当前帧的帧号, 这时轨迹图像和检测目标之间的相似性分数所占的比重也越大, 通过这样的加权计算可以得到更为准确的相似性分数, 得到更好的跟踪结果。

1.4 轨迹关联方法

得到轨迹和检测目标的相似性分数之后, 需要通过轨迹关联方法得到轨迹和检测目标之间的匹配关系, 得到最终的目标轨迹。但是由于在跟踪过程中可能有新的目标进入场景或旧的轨迹离开消亡, 因此只是简单地得到匹配关系是不够的, 还需要对轨迹进行管理。

通过上面的分析, 本文的轨迹关联方法如下: 首先在得到了当前 m 个轨迹和 n 个检测目标之间的相似性分数之后, 基于这些分数构造一个 $m * n$ 的相似度矩阵, 之后通过匈牙利算法进行二分图匹配得到轨迹和检测目标之间的匹配关系。在得到匹配关系之后, 进行轨迹管理, 轨迹的状态包括三种: 生存、未知和消亡。首先设定 F_{MAX} (一般设为 3) 为帧数的阈值, 用于确定轨迹的状态; 当一个轨迹匹配和当前帧中的检测目标成功匹配或新的轨迹出现时, 将其这些轨迹的状态视为生存并维护其信息, 其中新轨迹是指没有匹配上的检测目标在其后连续 F_{MAX} 帧都成功匹配关联后才能确定为新的轨迹; 而在本帧中没有匹配上的检测目标和没有匹配上且连续未匹配上帧数小于 F_{MAX} 的轨迹都为未知状态, 并需要维护其信息; 最后到当前帧连续未匹配帧数超过 F_{MAX} 的轨迹将视为消亡状态, 停止维护其信息。这样就完成了当前帧的跟踪过程。

2 实验及分析

为了验证本文提出的融合历史信息 and 运动信息的多目标跟踪方法 (HM-TBD) 的有效性和可行性, 本文选用常见的几种多目标跟踪方法进行对比。并通过在常用的多目标跟踪数据集上进行实验, 根据跟踪评价的主要指标对比分析, 验证方法的有效性和可行性。下面将对实验过程中所用到的数据集和评价指标进行介绍, 并对实验结果进行分析。

2.1 实验数据集

本文多目标跟踪的对比验证实验选用了 MOT16 数据集^[15]进行验证, 训练集和测试集各包含 7 个相机, 共 14 个相机, 其中 6 个相机为静止拍摄相机, 8 个为运动拍摄相机, 共 10 000 多张图像。MOT16 中所包含的部分场景如图 5 所示。

图 5 展示了 MOT16 的部分场景, 可以观察到该数据集



图 5 MOT16 部分场景示例

所包含的场景，相机的视角多变、场景复杂，且具有固定和移动拍摄两种运动状态，因此能够对本文提出的多目标跟踪方法的泛化能力及鲁棒性进行充分的考验。

2.2 评测指标

本文选择使用常见的多目标跟踪评价指标来对各个方法进行评测，主要指标有：识别 F1 分数 (Identification F1Score, IDF1)、身份交换次数 (ID switches, IDs)、多目标跟踪准确度 (Multiple Object Tracking Accuracy, MOTA)、多数跟踪数 (Mostly tracked, MT)、多数丢失数 (Mostly lost, ML) 以及轨迹跳变数 (Fragmentation, Frag)^[15]。计算方法如下面的公式所示。

IDF1 需要由识别精确度 (Identification Precision, IDP) 和识别召回率 (Identification Recall, IDR) 求得，IDP 的计算方法如式 (6) 所示：

$$IDP = \frac{IDTP}{IDTP + IDFP} \quad (6)$$

式 (6) 所计算的 IDP 表示每个行人所记录的框中的 ID 识别的精确度，其中 IDTP 和 IDFP 分别代表真阳性 ID 数量和假阳性 ID 数量。IDR 的计算方法如式 (7) 所示：

$$IDR = \frac{IDTP}{IDTP + IDFN} \quad (7)$$

式 (7) 所计算的 IDR 表示每个行人所记录的框中的 ID 识别的召回率，其中 IDFN 表示假阴性 ID 数量。IDF1 的计算方法由 IDP 和 IDR 得到，计算方法如式 (8)：

$$IDF1 = \frac{2 \times IDP \times IDR}{IDP + IDR} \quad (8)$$

式 (8) 所计算的 IDF1 表示每个行人所记录的框中的 ID 识别的 F1 分数，一般作为评价跟踪好坏的首要评价指标。IDs 表示一个行人的跟踪轨迹中行人 ID 切换的次数，通常能反映跟踪的稳定性，越小越好。Frag 代表跳变数，轨迹只要在跟踪和不跟踪两种状态中跳变一次，则跳变数加一次，越小越好。根据以上所得到的信息可以计算 MOTA，如式 (9) 所示：

$$MOTA = 1 - \frac{FN + FP + Frag}{GT} \quad (9)$$

式 (9) 所计算的 MOTA 用于评价单相机下多目标跟踪的准确性，其中 FN 表示所有帧中的假阴性检测之和，FP 表示所有帧中假阳性检测之和，GT 则表示真实的跟踪目标，同时 MOTA 也是常用衡量多目标跟踪性能指标之一。

通过上述评价方法，可验证多目标跟踪方法的综合性能，评估方法的稳定性与精度。

2.3 实验结果与分析

为了验证本文提出的融合历史信息 and 运动信息的多目标跟踪方法的有效性，选取 MOT16 数据集进行评测，并和常见的 4 种多目标跟踪方法进行对比，实验中所用的计算机硬件配置为：Intel (R) Core 系列 CPU，4 核 8 线程，主频为 3.40 GHz；使用了两条 8.0 GB 共 16.0 GB 的内存；显卡为 NVIDIA GeForce GTX 1060，位宽为 192bit，显存大小为 6 GB。框架选用 Pytorch 框架。实验中所用的检测结果都为数据集中提供的 Mask R-CNN 检测结果，轨迹关联方法都使用匈牙利算法进行关联。实验结果如表 1 所示。

表 1 多目标跟踪方法效果对比表

Method	MOTA ↑	IDF1 ↑	MT ↑	ML ↓	IDs ↓	Frag ↓
SORT	59.3	49.3	11.7%	30.9%	1001	1764
MOTDT	61.9	50.9	15.2%	38.3%	792	1859
DMAN	60.2	54.8	17.4%	42.7%	532	1616
Deep SORT	61.4	52.2	34.0%	18.2%	781	2008
HM-TBD	63.0	65.8	36.8%	25.6%	549	902

表 1 中，HM-TBD 表示本文提出的融合历史信息 and 运动信息的多目标跟踪方法，为了方便观察，数值越大表示效果越好的指标后方标注了向上的箭头，数值越小表示效果越好的指标后方标注了向下的箭头。首先从表中可以看出本文提出的方法在 MOTA 指标方面相较于其他方法要更好，结合 MOTA 的计算方法可以说明本文提出的跟踪方法在使用相同的检测方法的情况下，发生的 Frag 次数更少，表中的 Frag 数据也印证了这一点，从这两个指标中可以看出本文提出的跟踪方法得到的目标轨迹更为稳定、准确。同时从表中也可以看出 HM-TBD 方法的 IDF1 指标相较于其他方法较好，从 IDF1 的计算方式中可以看出，IDF1 表示目标识别的准确性，这就说明 HM-TBD 方法中所提出的融合历史信息 and 运动信息的特征提取网络效果更好，能够提取出更有区分性、更为鲁棒的目标特征。最后在表中的其他三个指标中本文提出的方法也获得了较好的表现，能够准确跟踪较多的目标，并且发生轨迹丢失和 ID 交换的次数较少。图 6~8 是跟踪结果较好的三个跟踪方法的跟踪轨迹图。



图 6 DMAN 方法跟踪效果



图 7 MOTDT 方法跟踪效果

从图 6 可以看出当目标发生遮挡时, DMAN 方法无法准确的再次跟踪到目标, 如第 170 帧的两位行人在经过遮挡后, 重新出现在画面中时两位行人的 ID 都发生了变化, 发生了轨迹变跳。从图 7 中可以看出发生同样的情况后, MOTDT 方法同样无法再次跟踪到目标, 发生了轨迹变跳和 ID 交换。而从图 8 中可以看出, 经过遮挡后, 本文提出的 HM-TBD 方法依然可以准确的跟踪到目标, 并且保证了轨迹的准确, 没有发生轨迹变跳和 ID 交换。通过对实验结果的分析, 可以看出, 本文提出的融合历史信息 and 运动信息的多目标跟踪方法在复杂场景下依然能取得较好的跟踪效果, 证明了本文提出的方法的有效性和可行性。



图 8 HM-TBD 方法跟踪效果

3 结束语

通过对多目标跟踪实际应用场景的分析, 本文提出了在提取目标特征时融合历史信息 and 运动信息, 有效地提高了特征的区别性和鲁棒性; 并在相似性分数计算阶段加入了权重的影响, 使得相似性分数更能真实地反映轨迹和检测目标之间的相似程度; 通过这两处的改进能够提高多目标跟踪方法在复杂场景下跟踪目标的鲁棒性和准确性, 提高跟踪效果。但在本文中只针对特征提取和相似性分数进行了改进, 而目标检测方法和轨迹关联方法没有改进, 因此下一步的工作就是对多目标跟踪方法进行整体改进, 提高跟踪效果。

参考文献:

- [1] 陈 艺. 改进 Camshift 算法的多行人目标跟踪方法 [J]. 导航定位学报, 2019, 7 (4): 30 - 36.
- [2] 王志余. 基于特征融合的复杂场景多目标跟踪算法研究 [J]. 软件导刊, 2020 (4): 46 - 49.
- [3] Bewley A, Ge Z, Ott L, et al. Simple Online and Realtime Tracking [A]. 2016 IEEE International Conference on Image Processing (ICIP) [C]. 2016, 3464 - 3468.
- [4] Wojke N, Bewley A, Paulus D. Simple online and realtime tracking with a deep association metric [A]. IEEE International Conference on Image Processing (ICIP) [C]. 2017, 3645 - 3649.
- [5] Li Wang, Nam Trung Pham, Tian-Tsong Ng, et al. Learning deep features for multiple object tracking by using a multi-task learning strategy [A]. In 2014 IEEE International Conference on Image Processing (ICIP) [C]. 2014, 838 - 842.
- [6] Chanh Kim, Fuxin Li, Arridhana Ciptadi, et al. Multiple hypothesis tracking revisited [A]. In Proceedings of the IEEE International Conference on Computer Vision [C]. 2015, 4696 - 4704.
- [7] Yu F W, Li W B, Li Q Q, et al. Poi: Multiple object tracking with high performance detection and appearance feature [A]. In European Conference on Computer Vision [C]. 2016, 36 - 42.
- [8] Li S, Deng M, Lee J, et al. Imaging Through Glass Diffusers Using Densely Connected Convolutional Networks [J]. Optica, 2018, 5 (7): 803 - 813.
- [9] 杨金龙, 程小雪, 缪佳妮, 等. 卷积特征多伯努利视频多目标跟踪算法 [J]. 计算机科学与探索, 2019 (11): 1945 - 1957.
- [10] Redmon J, Farhadi A. YOLO9000: better, faster, stronger [A]. IEEE Conference on Computer Vision and Pattern Recognition [C]. 2017, 7263 - 7271.
- [11] Zhou Z W, Xing J L, Zhang M D, et al. Online multi-target tracking with tensorbased high-order graph matching [A]. In 2018 24th International Conference on Pattern Recognition (ICPR) [C]. 2018, 1809 - 1814.
- [12] He K M, Gkioxari G, Dollár P, et al. Mask r-cnn [A]. In Proceedings of the IEEE international conference on computer vision [C]. 2017, 2961 - 2969.
- [13] Bewley A, Ge Z Y, Ott L, et al. Simple online and realtime tracking [A]. In 2016 IEEE International Conference on Image Processing (ICIP) [C]. 2016, 3464 - 3468.
- [14] Ren S, He K M, Girshick R, et al. Faster r-cnn: Towards real-time object detection with region proposal networks [C]. In Advances in Neural Information Processing Systems, 2015, 91 - 99.
- [15] 赵开开. 基于循环神经网络的视频目标跟踪算法研究 [D]. 成都: 电子科技大学, 2019.
- [16] 黎 阳, 沈 焯, 刘 敏, 等. 融合运动信息与表现信息的多目标跟踪算法 [J]. 电子科技, 2020 (9): 1 - 5.
- [9] 李继宇, 兰玉彬, 施叶茵, 等. 基于状态预测的田间机-地传感器系统协同采集方式研究 [J]. 农业机械学报, 2018, 23 (6): 246 - 253.
- [10] 白玉川, 温志超, 徐海珏, 等. 三维海洋水动力计算 σ 坐标转换存在的问题及改进方法 [J]. 应用数学和力学, 2019, 40 (8): 840 - 855.
- [11] 韩逢庆, 肖 丹, 官礼和. 一种非视距环境下基于 TOA 三维定位的新方法研究 [J]. 应用数学和力学, 2019, 40 (11): 1259 - 1269.
- [12] 王兆年, 常 凝, 王文杰, 等. 抗干扰 GPS 接收机中角度、频偏和时延的联合估计及跟踪方法 [J]. 信号处理, 2019, 16 (4): 609 - 616.
- [13] 赵文晔, 高井祥, 李增科, 等. 地图匹配辅助的 KF-PF 室内定位算法模型 [J]. 武汉大学学报 (信息科学版), 2018, 43 (5): 806 - 812.

(上接第 222 页)