

基于深度学习的红外与可见光决策级融合跟踪

唐聪^{1,2*}, 凌永顺^{1,2}, 杨华^{1,2}, 杨星^{1,2}, 同武勤³

¹国防科技大学电子对抗学院, 安徽 合肥 230037;

²脉冲功率激光技术国家重点实验室, 安徽 合肥 230037;

³西南电子电信技术研究所, 四川 成都 610041

摘要 提出了一种基于深度学习的红外与可见光决策级融合跟踪方法。通过建立参数传递模型,从现有基于深度学习的检测模型中抽取指定对象的可见光检测模型,作为红外检测的预训练模型,在采集的红外图像数据集上进行微调训练,得到基于深度学习的红外检测模型。在此基础上,建立了基于深度学习的红外与可见光决策级融合跟踪模型,进行了单波段跟踪与双波段融合跟踪对比实验。结果表明,所提方法跟踪精度和成功率比单波段跟踪均有所提升,具有较好的稳健性。

关键词 机器视觉; 目标跟踪; 决策级融合; 双波段; 深度学习

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP56.071502

Decision-Level Fusion Tracking for Infrared and Visible Spectra Based on Deep Learning

Tang Cong^{1,2*}, Ling Yongshun^{1,2}, Yang Hua^{1,2}, Yang Xing^{1,2}, Tong Wuqin³

¹College of Electronic Countermeasures, National University of Defense Technology, Hefei, Anhui 230037, China;

²State Key Laboratory of Pulsed Power Laser Technology, Hefei, Anhui 230037, China;

³South-West Electron and Telecom Technology Institute, Chengdu, Sichuan 610041, China

Abstract A decision-level fusion tracking method based on deep learning for infrared and visible spectra is proposed. By building the parameter transfer model, the visible detection model of the specified objects is extracted from the existing deep-learning-based detection model. This visible detection model is used as the infrared detection pre-training model, and the fine-tuning training on a collected infrared image dataset is done to obtain the infrared detection model based on deep learning. On this basis, a decision-level fusion tracking model based on deep learning is built. An comparison experiment between single-band tracking and dual-band fusion tracking is carried out. The research results show that the proposed method improves the tracking accuracy and success rate compared with the single-band tracking, and has good robustness.

Key words machine vision; object tracking; decision-level fusion; dual-band; deep learning

OCIS codes 150.1135; 100.4999; 100.2000

1 引言

目标跟踪已经成为计算机视觉领域重要的研究方向和研究热点,可应用于精确制导武器、无人机侦察监视等军事领域,也可应用于机器人导航、人机交互,行人与车辆的视频监控等民用领域^[1-3]。将多源传感器进行数据/信息融合^[4]可以实现单一传感器

无法获得的目标跟踪效果,提升目标跟踪性能,而其中最典型的是红外与可见光的融合跟踪。在可见光条件下,传感器通常获取目标的反射特性,即目标的颜色、纹理等特征,采用高分辨成像器件更利于获取目标的细节信息;而在红外条件下,传感器通常获取目标的热辐射特性,即目标的温度特征(与环境对比度较明显),便于探测目标,可以避免光照不均及阴

收稿日期: 2018-09-18; 修回日期: 2018-10-17; 录用日期: 2018-10-22

基金项目: 国家自然科学基金(61503394,61405248)、安徽省自然科学基金面上项目(1708085MF137)

* E-mail: tangcong17@nudt.edu.cn

影等因素的影响,特别是在弱光条件下具有天然的优势。红外与可见光融合跟踪同时具备双波段各自的优势。

目前,红外与可见光融合跟踪大多仍采用传统方法,如基于联合稀疏表示^[5]、多特征融合^[6]、Kalman滤波和压缩感知等方法^[7],而基于深度学习的红外与可见光融合跟踪研究鲜有报道。基于深度学习的目标跟踪大多针对可见光条件下自然图像的目标跟踪,针对红外条件下的目标跟踪研究不多。近年来,基于深度学习的目标跟踪方法主要有:模型采用离线训练,在线微调的模式,如深度学习跟踪器(DLT)^[8];利用现有的基于深度卷积神经网络强化特征提取,促进目标跟踪模式,如层次化卷积特征跟踪(HCFT)^[9];采用跟踪序列视频集预训练,在线微调模式,如多域网络(MDNet)^[10]。这三种模式均需要在模型设计后进行离线训练,在线跟踪的同时进行在线训练。

深度学习技术广泛应用于语音、图像、自然语言和机器翻译等领域,而在图像领域,尤其以物体检测最成功。本文采用基于深度学习的物体检测算法进行候选目标提取,结合全局性的颜色直方图特征和局部性的方向梯度直方图特征进行目标选择,以实现跟踪,该方法不同于常规利用深度学习进行目标跟踪的模式,是作者课题组前期的一项工作,称为基于深度学习物体检测的视觉跟踪方法(TDLD)^[11]。这里采用某种信息融合方法与之结合,从而进行基

于深度学习的红外与可见光融合跟踪研究。

红外与可见光融合跟踪有三种方式:像素级融合跟踪、特征级融合跟踪和决策级融合跟踪^[12]。像素级融合跟踪需要先将图像进行像素级融合,再利用融合图像进行目标跟踪;特征级融合跟踪通过特征提取和特征融合,再基于融合后的特征进行目标跟踪;而决策级融合跟踪分别采用不同的传感器进行目标跟踪,而后对跟踪结果在决策层上进行融合。本文采用的基于深度学习物体检测的视觉跟踪方法关键在于深度学习物体检测模型。基于深度学习的模型属于数据驱动型,若采用前两种融合跟踪方法,需要准备大量的数据以进行物体检测模型训练。目前,可见光图像数据集可从公开渠道获得,红外图像数据集则较难获得,而且进行像素级融合跟踪和特征级融合跟踪的模型训练要求所采用的训练数据集中的可见光图像与红外图像来自同一场景,使得训练数据的获得更难。若采用决策级融合跟踪的方法,可见光检测模型可以从现有的模型中抽取得到,不需要额外训练,而红外检测模型可以从可见光检测模型上通过微调^[13]得到,极大地降低了对训练数据的要求。因此,本文采用决策级融合的方法开展基于深度学习的红外与可见光融合跟踪研究。

2 基于深度学习物体检测的视觉跟踪

利用经典跟踪算法进行视觉跟踪,结果中出现跟踪漂移的位置通常不可知,如图1所示。



图1 经典跟踪算法中的跟踪漂移。(a)文献[9]算法;(b)文献[14]算法;(c)文献[15]算法;(d)文献[16]算法

Fig. 1 Tracking drift using several classic tracking methods. (a) Algorithm in Ref. [9];

(b) algorithm in Ref. [14]; (c) algorithm in Ref. [15]; (d) algorithm in Ref. [16]

从图1可以发现,发生跟踪漂移时,经典跟踪算法所框选的目标与真实目标在视觉观感上差别较大,甚至在纹理、颜色和轮廓上看不出有任何关联。在视觉跟踪过程中,通常以物体(或物体块)作为跟

踪目标,若能从跟踪视场中检测出相应的物体,则将其作为候选目标进行目标选择,从而实现目标跟踪。

本课题组前期提出了一种基于深度学习物体检测的视觉跟踪算法——TDLD,该算法将基于深度

学习物体检测的方法引入目标跟踪,有效提取了候选目标,结合全局性的颜色直方图特征和局部性的方向梯度直方图特征进行目标选择进而实现跟踪,具体细节参见文献[11]。本文的出发点在于建立红外与可见光双波段融合跟踪,而目前,基于深度学习的目标检测模型大多是在可见光条件下的目标检测,红外目标检测模型比较少。因此,这里首先进行基于深度学习的红外目标检测模型设计与训练,为后续进行融合跟踪做好铺垫。

3 基于深度学习的红外目标检测模型

3.1 基于深度学习的红外目标检测预训练模型

通常可见光条件下基于深度学习的目标检测模型的预训练模型经过 ImageNet^[17]、Pascal VOC^[18]训练,单发多盒探测器(SSD)^[19]模型的训练也如此。SSD模型是一个用于物体检测的深度神经网络模型,其属于基于回归的目标检测模型的典型代表,具有检测精度高、时效性好等特点。但如果直接使用训练好的 SSD 模型(VGG_VOC0712_SSD_300×300_iter_240000.caffemodel),由于新模型分类层的分类类别与 SSD 模型不一致,SSD 模型分类卷积层的参数无法直接传给新模型。为了便于说明问题,选择交通检测与跟踪中常见的 5 类目标,包括行人、汽车、公交车、摩托车和自行车。而 VOC 数据集中包含自行车、公交车、汽车、摩托车和行人,因此,从 VOC 预训练模型中通过编程抽取 5 类目标和背景(共 6 类物体)的权值。同时,目标定位仅与 4 个坐标值信息相关,与类别数目无关,因此可以通过参数共享继承可见光模型定位目标的能力。为了得到性能较好的预训练模型,实现模型参数传递是一种有效的方式,所得模型既可以快速完成训练,又可以获得较好的检测性能。另外,模型分类层必须重新设计,与卷积层结合之后,得到预训练模型。该预训练模型也是针对 5 类目标物体的基于深度学习的可见光检测模型。

3.1.1 参数传递模型

设计红外预训练模型时,首先根据模型结构进行参数初始化得到初始模型,同时,为了充分利用可见光模型(pretrained SSD model)已具备的目标检测能力,将可见光模型中各卷积层参数传递给红外模型对应的卷积层,实现参数共享。卷积层传递的参数包括卷积核参数和偏置参数(均视为矩阵)。设可见光模型与红外模型的参数矩阵分别为 \mathbf{W}_1 、 \mathbf{W}_2 ,其中, \mathbf{W}_1 为 $(M \times r) \times c \times k \times k$ 维矩阵, \mathbf{W}_2 为 $(N \times r) \times c \times k \times k$

维矩阵, M 、 N 分别为可见光模型和红外模型的物体种类数, r 为默认框类别序号, c 为卷积层通道数, k 为卷积核大小。设可见光模型与红外模型的偏置矩阵为 \mathbf{b}_1 、 \mathbf{b}_2 (均为 $4 \times r \times 1$ 的一维矩阵)。

进行参数传递时, \mathbf{W}_1 与 \mathbf{W}_2 关系为

$$\mathbf{W}_2[j + N \times r, \dots] = \mathbf{W}_1[i + M \times r, \dots], \quad (1)$$

另外, \mathbf{b}_1 与 \mathbf{b}_2 关系为

$$\mathbf{b}_2[j + 6 \times r] = \mathbf{b}_1[i + 21 \times r], \quad (2)$$

式中 i 、 j 分别为可见光模型和红外模型的类别序号。

3.1.2 参数传递的实施

红外预训练模型以 SSD 模型为架构,需要进行参数传递的层为 conv4_3_norm_mbox_conf、fc7_mbox_conf、conv6_2_mbox_conf、conv7_2_mbox_conf、conv8_2_mbox_conf、conv9_2_mbox_conf,其中,conv4_3_norm_mbox_conf、conv8_2_mbox_conf、conv9_2_mbox_conf 三个卷积层的特征图默认框类别有 4 种,即 r 可取 0、1、2、3,则默认框索引集 $R = \{0, 1, 2, 3\}$,而 fc7_mbox_conf、conv6_2_mbox_conf、conv7_2_mbox_conf 三个卷积层的特征图默认框类别有 6 种,即 r 可取 0、1、2、3、4、5,则默认框索引集 $R = \{0, 1, 2, 3, 4, 5\}$ 。选取的目标类别为背景、自行车、公交车、汽车、摩托车、行人,其在可见光模型的分层中对应的类别序号为 0、2、6、7、14、15;而在红外模型中,这 6 类目标对应的类别序号为 0、1、2、3、4、5。 i 为物体类别在可见光模型中分类层的索引值,根据具体的物体而定,对于上述 5 种类别,外加背景, i 可取 0、2、6、7、14、15,则类别索引集 $I = \{0, 2, 6, 7, 14, 15\}$, j 为物体类别在红外模型中分类层的索引值,直接取 0、1、2、3、4、5,则类别索引集 $J = \{0, 1, 2, 3, 4, 5\}$ 。两个模型卷积层之间的参数传递过程按照参数传递算法进行,其具体过程如下。

1) 初始化参数矩阵 \mathbf{W}_2 和偏置矩阵 \mathbf{b}_2 ,分别为 $(M \times r) \times c \times k \times k$ 维全零矩阵和 $(4 \times r) \times 1$ 维全零矩阵。

2) 在默认框索引集 R 中选择当前未被遍历的首位索引 r 。

3) 从红外预训练模型的类别索引集 I 和可见光模型的类别索引集 J 中逐个抽取同一类别的序号 i 和 j ,并根据(1)式和(2)式从可见光模型的参数矩阵和偏置矩阵中获取当前类别下红外预训练模型的参数矩阵和偏置矩阵,直至所有类别均遍历该操作。若默认框索引集 R 遍历未结束,则跳至步骤 2),反之,则进入下一步。

4) 模型参数传递结束,得到红外预训练模型参数。

通过参数传递,红外预训练模型与可见光模型实现了目标定位和目标识别的参数共享,使其具备一定的目标检测能力。

3.2 红外数据集模型训练

得到预训练模型后,需要采集红外数据集进行模型训练。2017年12月,在校园、广场、十字路口、

人行下穿隧道等自行车、公交车、汽车、摩托车、行人较多的场所采集上述5类目标的红外图像1800张,其中,自行车334张,公交车205张,汽车430张,摩托车339张,行人492张。所采用的红外相机(T630型,FLIR公司,美国)的输出图像尺寸为480 pixel×640 pixel。

3.2.1 红外数据集采集

采集的红外图像数据集中部分结果如图2所示。

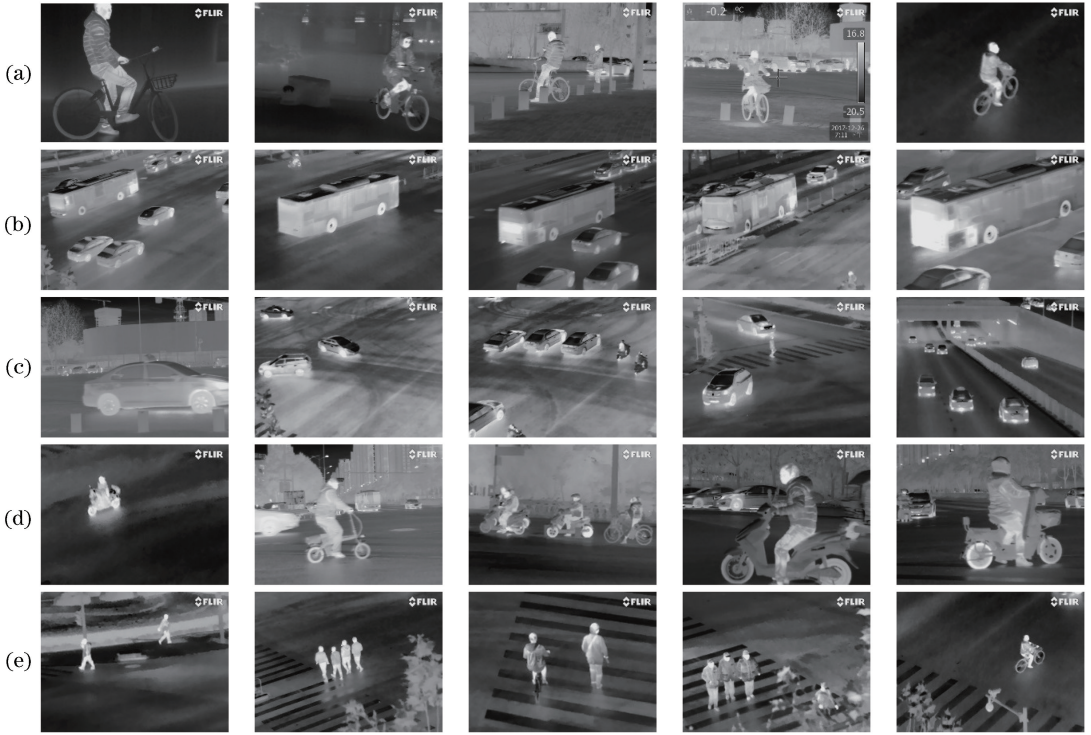


图2 红外图像数据集(示例)。(a)自行车;(b)公交车;(c)汽车;(d)摩托车;(e)行人

Fig. 2 Infrared images dataset (examples). (a) Bicycle; (b) bus; (c) car; (d) motorbike; (e) pedestrian

3.2.2 模型训练

经过3000次迭代训练,模型在红外图像数据集上的训练损失约为2.5,训练损失变化曲线如图3所示。

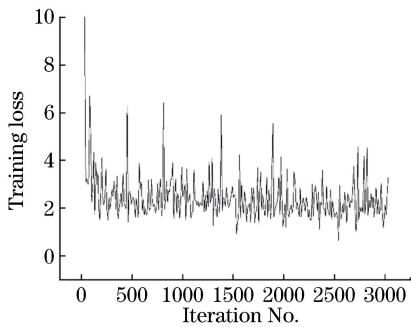


图3 训练损失值随迭代次数的变化曲线

Fig. 3 Training loss versus number of iterations

在训练过程中,每迭代1000次进行一次测试,测试精度与迭代次数的变化曲线如图4所示。

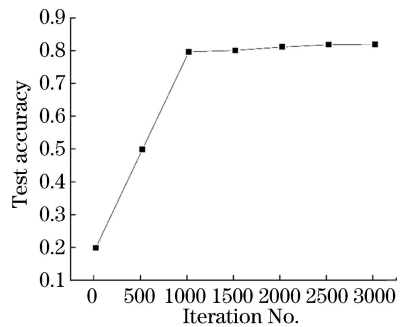


图4 测试精度随迭代次数的变化曲线

Fig. 4 Test accuracy versus number of iterations

通常目标检测的性能评采用平均精度均值(mAP)进行评价,mAP表示多类物体平均精度(AP)的平均值。而AP的计算依赖于准确率-召回

率($P-R$)曲线^[20],其中,检测准确率指检测正确的正样本在检测结果中的占比;召回率指检测正确的正样本在所有正样本中的占比。

利用微调得到的红外目标检测模型对整个红外测试集进行检测,计算自行车、公交车、汽车、摩托车、行人的 AP 及 mAP,结果如表 1 所示。

表 1 红外测试数据集上 5 种类别的 mAP

Table 1 mAP of five classes on infrared test datasets

mAP	AP				
	Bicycle	Bus	Car	Motorbike	Person
0.823	0.788	0.896	0.864	0.806	0.758

由图 3 可知,该模型在训练过程中较好地收敛,表明模型稳定可用,再结合图 4 和表 1 可知,该模型具备较好的目标检测能力,即所提方法可在已有的基于深度学习的可见光物体检测模型上获得性能较

好的基于深度学习的红外物体检测模型。

获得基于深度学习的红外物体检测模型和可见光物体检测模型后,可根据基于深度学习物体检测的视觉跟踪算法设计基于深度学习的红外和可见光视觉跟踪模型。在此基础上,便可开展基于深度学习的红外与可见光决策级融合跟踪研究。

4 基于深度学习的红外与可见光决策级融合跟踪

4.1 模型建立

基于深度学习的红外与可见光决策级融合跟踪采用红外与可见光分别进行目标跟踪,在后端进行决策级融合。其模型如图 5 所示。

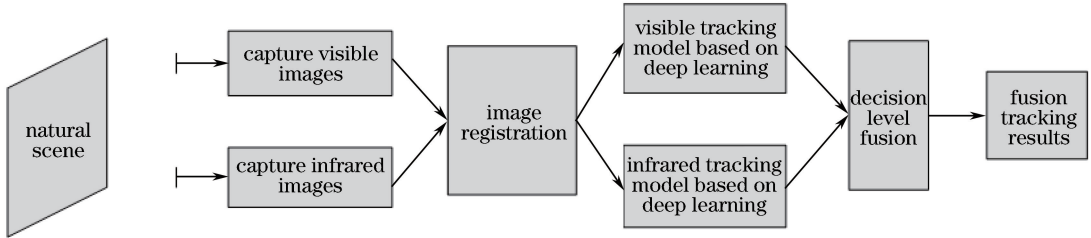


图 5 基于深度学习的红外与可见光决策级融合跟踪

Fig. 5 Decision-level fusion tracking for infrared and visible spectra based on deep learning

针对自然场景,采集红外图像与可见光图像并配准后,分别送入基于深度学习的可见光跟踪模型和基于深度学习的红外跟踪模型,对跟踪结果进行决策级融合,即可得到最终的跟踪结果。

4.2 决策级融合

进行双波段融合跟踪时,每个波段显示的图像可能会因为波段的不同而不同,但因跟踪同一目标,搜索视窗和模板大小一致,而跟踪模板为双波段,彼此起到互补作用。融合跟踪过程中选择置信度较大的跟踪目标作为跟踪结果,如图 6 所示。具体过程如下。

1) 利用 TDLT 在基于深度学习的红外跟踪模型与可见光跟踪模型上分别进行单波段目标跟踪,在每一帧上获得两个候选目标位置 B_{infrared} 和 B_{visible} ,为了甄选目标,同时输出两个候选目标作为物体的置信度,即分别输出 $(C_{\text{infrared}}, B_{\text{infrared}})$ 和 $(C_{\text{visible}}, B_{\text{visible}})$;

2) 比较红外跟踪目标物体置信度 C_{infrared} 和可见光跟踪目标物体置信度 C_{visible} 的大小,选择物体置信度较大的目标作为最终跟踪目标,即当

$C_{\text{infrared}} \geq C_{\text{visible}}$ 时,选择 B_{infrared} 标定的候选目标作为跟踪目标,同时,将红外跟踪模型的模板标定区域作为融合模型的模板区域,即可见光跟踪模型的模板根据红外跟踪的模板区域进行更新;否则,选择 B_{visible} 标定的候选目标作为跟踪目标,同时,将可见光跟踪模型的模板标定区域作为融合模型的模板区域,即红外跟踪模型的模板根据可见光跟踪的模板区域进行更新;

3) 红外跟踪模型和可见光跟踪模型分别根据当前的模板进行目标跟踪,各产生新的候选目标,重复步骤 1),完成下一帧的融合跟踪。

4.3 融合跟踪实验及分析

为验证双波段融合算法的可行性,在实验室条件下设计两种模拟环境,即透红外不透可见光和透可见光不透红外的实验环境。针对条件一,采用烟幕遮挡的方式进行模拟(烟幕具有透红外而不透可见光的特性);而针对条件二,采用玻璃遮挡的方式进行模拟(玻璃具有透可见光不透红外的特性)。

同时采集两段金属仿真装甲车通过上述实验环境的红外图像序列与可见光图像序列,在此基础上

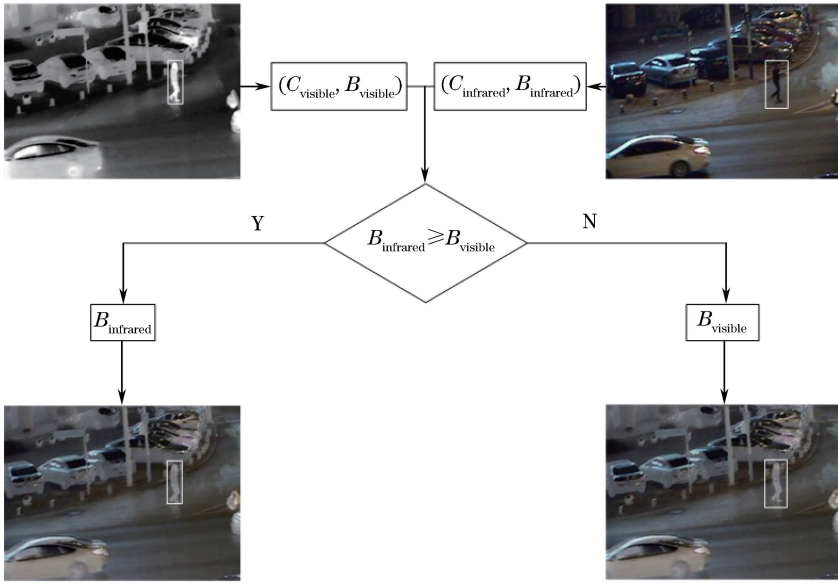


图 6 决策级融合跟踪过程

Fig. 6 Process of decision-level fusion tracking

进行基于深度学习的红外与可见光单波段和双波段融合跟踪实验。硬件设备有红外热像仪、可见光相机、双机位支架、三自由度云台脚架、HDMI 转 CVBS 转换器、图像采集卡、计算机；实验在 Ubuntu 操作系统平台上进行，GPU 采用 Quadro K4000，采

用 Python、C++ 混合编程。

4.3.1 定性分析

分别采用基于深度学习的红外跟踪模型、可见光跟踪模型、红外与可见光融合跟踪模型进行目标跟踪，实验结果如图 7 所示。

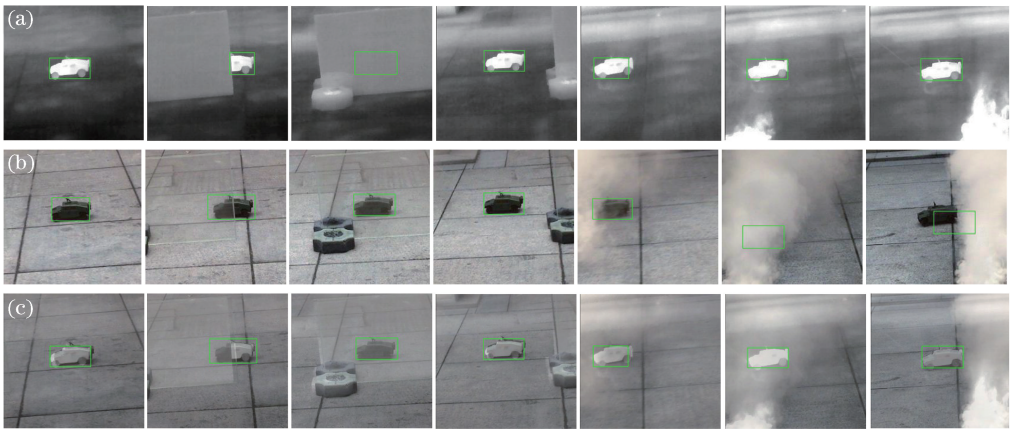


图 7 跟踪结果(帧序列号为 1,15,32,55,129,143,162)。(a)红外跟踪;(b)可见光跟踪;(c)红外与可见光融合跟踪

Fig. 7 Tracking results(frame sequence number is 1, 15, 32, 55, 129, 143, 162). (a) Infrared tracking; (b) visible tracking; (c) fusion tracking infrared and visible

图 7 给出了红外与可见光单波段跟踪与双波段融合跟踪的部分跟踪结果,对比可知,目标经过红外遮障(模拟红外隐身)时,可见光可有效对其实施跟踪,当目标经过烟幕(模拟可见光隐身)时,红外波段可有效对其实施跟踪。第 1 帧为开始进行跟踪的初始帧,红外跟踪与可见光跟踪均有效,融合跟踪采用红外跟踪的结果,由于红外图像中目标的对比度高于可见光图像中目标的对比度,红外跟踪的置信度

高于可见光跟踪;第 15 帧为目标前半部进入红外遮障,虽然红外跟踪与可见光跟踪仍有效,但融合跟踪采用可见光跟踪的结果,红外跟踪只能跟踪到仿真装甲车的局部,跟踪结果的置信度低于可见光跟踪结果的置信度;第 32 帧为目标完全进入红外遮障,仅可见光跟踪有效,融合跟踪必然采用可见光跟踪的结果;第 55 帧为目标完全穿出红外遮障,此时,红外跟踪与可见光跟踪同时有效,融合跟踪采用红外

跟踪的结果;第 129 帧为目标前半部分进入可见光烟幕,红外跟踪与可见光跟踪同时有效,融合跟踪采用红外跟踪的结果;第 143 帧为目标完全淹没在可见光烟幕里,仅红外跟踪有效,融合跟踪采用红外跟踪的结果;第 162 帧为目标完全穿出可见光烟幕后,红外跟踪与可见光跟踪均有效,融合跟踪采用红外跟踪的结果。因此,双波段的融合跟踪可有效利用红外与可见光之间的互补性,实现在红外隐身与可见光干扰条件下的目标跟踪。

4.3.2 定量分析

评价跟踪效果时常采用重叠率、中心位置误差^[21]两个评价指标,本文采用该指标对跟踪效果进行评价,在 Origin 绘图软件中呈现结果,如图 8 和图 9 所示。

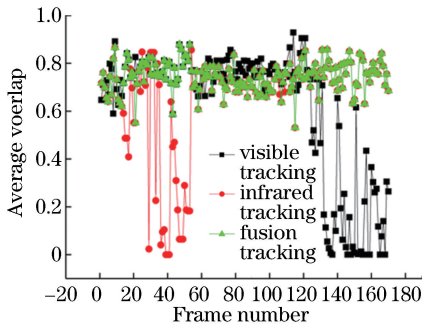


图 8 双波段融合跟踪与单波段跟踪在重叠率上对比

Fig. 8 Comparison of overlap score between dual-band fusion and single band tracking

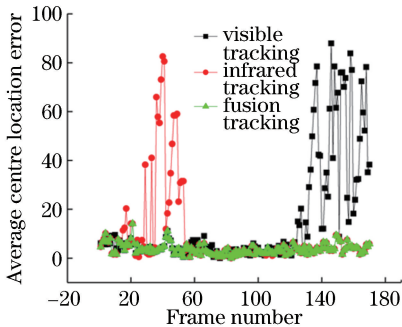


图 9 双波段融合跟踪与单波段跟踪在中心位置误差上的对比

Fig. 9 Comparison of centre location error between dual-band fusion and single band tracking

根据采集视频可知,仿真装甲车在第 9 帧开始经过红外遮障,在第 49 帧时完全穿出,在第 110 帧时开始经过可见光烟幕,在第 162 帧时完全穿出。结合图 8 和图 9 可以发现,从第 28 帧开始,红外跟踪出现大偏差漂移,此时红外遮障已将目标完全遮蔽,而从第 132 帧开始,可见光跟踪出现大偏差漂移,此时可见光烟幕已将目标完全遮蔽,而融合跟踪

则因为双波段的互补性,全程均具有较高的覆盖率和较低的中心位置误差。

为了增强对比效果,这里定义了目标丢失率,即跟踪序列中目标丢失的帧数占总帧数的比例,其表达式为

$$R = N_{\text{fail}} / N, \quad (3)$$

式中: N_{fail} 为跟踪序列中目标丢失的帧数;丢失的定义为目标跟踪框与地面真值物体框之间的交叉区域(IoU) >0.5 ; N 为目标跟踪总帧数。

定量计算后,融合跟踪与单波段跟踪的性能对比如表 2 所示。

表 2 双波段融合跟踪与单波段跟踪的性能对比

Table 2 Performance comparison between dual-band fusion tracking and single band tracking

Performance	Visible tracking	Infrared tracking	Fusion tracking
Average overlap score / %	61.8	66.4	74.0
Average centre location error / pixel	14.6	8.8	4.0
Target loss rate	0.23	0.11	0

由表 2 可知,相比于可见光跟踪和红外跟踪,融合跟踪在平均覆盖率和平均中心位置误差上均有明显提升,同时,跟踪过程中无目标丢失。而红外跟踪在丢失目标的情况下,通过红外遮障后能够再次跟踪目标,其原因在于其更新模板的物体框与可见光模板同步更新,使得在单波段目标丢失的情况下,依然能够获取较好的跟踪模板,便于后续目标跟踪。

5 结 论

将基于深度学习的目标检测模型迁移至目标跟踪,同时通过参数传递模型,从已有的深度学习物体检测模型中获得了特定对象的可见光条件下的物体检测模型,进而通过微调获得了红外条件下的物体检测模型,再采用决策级融合方式实现了红外与可见光融合跟踪,充分利用了双波段的信息互补性。实验结果表明,相比于单波段目标跟踪,所提出的基于深度学习的红外与可见光决策级融合跟踪方法的跟踪精度和成功率更高,稳健性更好。该方法不仅限于双波段的融合跟踪,可扩展为基于深度学习的多波段融合跟踪,应用更广泛,是将现有深度学习模型进行模型迁移学习的一个较好的范例,具有一定的理论和应用价值。不足之处在于,在目前的实验环境下,所提算法的速度为 2~3 frame/s,下一步将提升算法的实时性。

参 考 文 献

- [1] Sivanantham S, Paul N N, Iyer R S. Object tracking algorithm implementation for security applications [J]. Far East Journal of Electronics and Communications, 2016, 16(1): 1-13.
- [2] Tang C, Ling Y S, Zheng K D, *et al.* Object detection method of multi-view SSD based on deep learning[J]. Infrared and Laser Engineering, 2018, 47(1): 0126003.
唐聪, 凌永顺, 郑科栋, 等. 基于深度学习的多视窗 SSD 目标检测方法[J]. 红外与激光工程, 2018, 47(1): 0126003.
- [3] Luo H B, Xu L Y, Hui B, *et al.* Status and prospect of target tracking based on deep learning[J]. Infrared and Laser Engineering, 2017, 46(5): 0502002.
罗海波, 许凌云, 惠斌, 等. 基于深度学习的目标跟踪方法研究现状与展望[J]. 红外与激光工程, 2017, 46(5): 0502002.
- [4] Hall D L, Llinas J. An introduction to multisensor data fusion[J]. Proceedings of the IEEE, 1997, 85(1): 6-23.
- [5] Liu H P, Sun F C. Fusion tracking in color and infrared images using joint sparse representation[J]. Science China Information Sciences, 2012, 55(3): 590-599.
- [6] Zhao G P, Bo Y M, Chen Y. A method for tracking object in infrared and visible image based on multiple features[J]. Acta Armamentarii, 2011, 32(4): 445-451.
赵高鹏, 薄煜明, 陈益. 多特征提取的红外和可见光目标跟踪方法[J]. 兵工学报, 2011, 32(4): 445-451.
- [7] Yun X, Jing Z L, Xiao G, *et al.* A compressive tracking based on time-space Kalman fusion model [J]. Science China Information Sciences, 2016, 59(1): 012106.
- [8] Wang N Y, Yeung D Y. Learning a deep compact image representation for visual tracking [C] // International Conference on Neural Information Processing Systems, 2013: 809-817.
- [9] Ma C, Huang J B, Yang X K, *et al.* Hierarchical convolutional features for visual tracking[C] // IEEE International Conference on Computer Vision, 2015: 3074-3082.
- [10] Nam H, Han B. Learning multi-domain convolutional neural networks for visual tracking[C] // IEEE Conference on Computer Vision and Pattern Recognition, 2016: 4293-4302.
- [11] Tang C, Ling Y S, Yang H, *et al.* A visual tracking method via object detection based on deep learning [J]. Infrared and Laser Engineering, 2018, 47(5): 0526001.
唐聪, 凌永顺, 杨华, 等. 基于深度学习物体检测的视觉跟踪方法[J]. 红外与激光工程, 2018, 47(5): 0526001.
- [12] Fendri E, Boukhriss R R, Hammami M. Fusion of thermal infrared and visible spectra for robust moving object detection [J]. Pattern Analysis and Applications, 2017, 20(4): 907-926.
- [13] Guo Y M, Liu Y, Oerlemans A, *et al.* Deep learning for visual understanding: a review [J]. Neurocomputing, 2016, 187: 27-48.
- [14] Wang L, Liu T, Wang G, *et al.* Video tracking using learned hierarchical features [J]. IEEE Transactions on Image Processing, 2015, 24(4): 1424-1435.
- [15] Wang N, Li S, Gupta A, *et al.* Transferring rich feature hierarchies for robust visual tracking [J]. CarX iv preprint arX iv:1501.04587, , 2015.
- [16] Wang X, Hou Z Q, Yu W S, *et al.* Robust visual tracking via multiscale deep sparse networks [J]. Optical Engineering, 2017, 56(4): 043107.
- [17] Russakovsky O, Deng J, Su H, *et al.* ImageNet large scale visual recognition challenge [J]. International Journal of Computer Vision, 2015, 115(3): 211-252.
- [18] Vicente S, Carreira J, Agapito L, *et al.* Reconstructing PASCAL VOC [C] // IEEE Conference on Computer Vision and Pattern Recognition, 2014: 41-48.
- [19] Liu W, Anguelov D, Erhan D, *et al.* SSD: single shot multibox detector[C] // European Conference on Computer Vision, 2016: 21-37.
- [20] Davis J, Goadrich M. The relationship between precision-recall and ROC curves [C] // International Conference on Machine Learning, 2006: 233-240.
- [21] Wu Y, Lim J, Yang M H. Object tracking benchmark [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1834-1848.