



(12) 发明专利

(10) 授权公告号 CN 114627339 B

(45) 授权公告日 2024. 03. 29

(21) 申请号 202111317119.8
(22) 申请日 2021.11.09
(65) 同一申请的已公布的文献号
申请公布号 CN 114627339 A

G06T 7/269 (2017.01)
G06T 7/277 (2017.01)
G06V 10/74 (2022.01)
G06V 40/10 (2022.01)
G06V 40/20 (2022.01)

(43) 申请公布日 2022.06.14
(73) 专利权人 昆明物理研究所
地址 650221 云南省昆明市五华区教场东路31号

(56) 对比文件
CN 110399808 A, 2019.11.01
CN 112380997 A, 2021.02.19
CN 112381856 A, 2021.02.19
CN 112597877 A, 2021.04.02
WO 2020155873 A1, 2020.08.06
沈茂东;高宏;付新阳;周伟;张俊岭;公凡奎;冯志珍.基于改进SSD的电力检修多目标人员追踪方法.计算机系统应用.2020,(第08期),全文.

(72) 发明人 王元余 张晋 宋治杭 朱亮
张强 李伟 杜欣悦 林丹丹
林宇

(74) 专利代理机构 昆明今威专利商标代理有限公司 53115
专利代理师 赛晓刚

审查员 闫飞燕

(51) Int. Cl.
G06V 10/77 (2022.01)

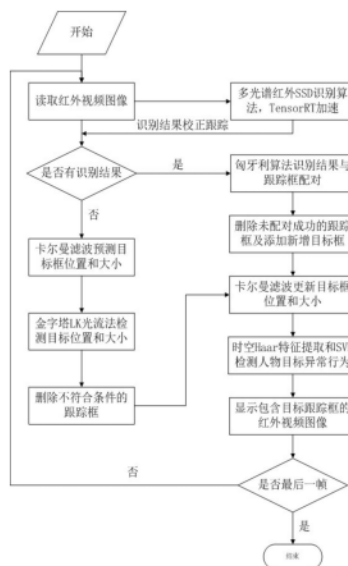
权利要求书4页 说明书6页 附图2页

(54) 发明名称

茂密丛林区域对越境人员的智能识别跟踪方法及存储介质

(57) 摘要

本发明公开了一种茂密丛林区域对越境人员的智能识别跟踪方法及存储介质,采用红外多光谱图像数据,利用动植物和人物之间红外多光谱特征显著不同的特点,图像数据经过降维处理并结合Single Shot MultiBox Detector神经网络模型后,能同时保证了人物识别的速度和精度。由于边境区域有正常生活的居民,本发明采用机器学习算法对人物进行多目标跟踪并进行异常行为分析,能够有效区分非法越境人员和正常生活的居民,提高了边防人员对边境区域中人物甄别的效率。



1. 一种茂密丛林区域对越境人员的智能识别跟踪方法,其特征在于,包括以下步骤:

步骤1,读取红外多光谱图像;

步骤2,利用主成分分析PCA对红外多光谱图像进行降维处理,保留特征值较大的前5个光谱特征,去掉冗余的光谱特征信息;

步骤3,调用SSD目标识别算法对经PCA降维后的光谱图像特征进行多目标识别,包括:

步骤3.1,基于深度学习框架训练SSD目标识别算法模型;

步骤3.2,利用TensorRT对模型进行序列化;

步骤4,在第一次对红外视频进行多目标识别并得到识别结果后,对多目标跟踪器进行初始化针对每个目标识别框建立一个跟踪器,同时使用FAST特征点检测器检测每个目标识别框的特征点信息,并存储相应目标识别框的位置、大小、目标类型、目标编号和特征点信息,同一帧图像中的目标编号没有重复,并且同一个目标在前后帧中的编号保持不变;

步骤5,在接收到新的一帧视频图像后,如果此时另外一个线程中的目标识别器没有产生新的识别结果,则执行下列子步骤:

步骤5.1,基于卡尔曼滤波算法和前一帧的目标识别框结果进行预测,得到对应目标框在当前帧中位置和大小预测结果;

步骤5.2,利用前一帧目标的特征点信息和金字塔LK光流法计算当前帧中与前一帧目标特征点相对应的特征点位置信息,分别计算前一帧和当前帧中不同特征点之间的距离,这样在前后两帧图像中便得到了相同数量的距离数据,接着计算相同两个特征点在前后两帧图像中距离的比值 $d_{\text{当前帧}}/d_{\text{前一帧}}$,计算距离比值的中位数,利用中位数更新当前帧中目标框的大小和目标框左上角坐标位置;

步骤5.3,基于步骤5.2得到当前帧中目标框位置和大小观测值,利用卡尔曼滤波算法,在均方误差最小的情况下,对当前的状态进行修正,从而得到目标框位置和大小最优估计值,并保证前后两帧中同一目标的编号不变;随后对当前帧中的目标框重新进行特征点检测,增加目标框的特征点数量以便在后续帧中更好的跟踪目标,同时对跟踪框进行筛选,剔除矩形框完全位于图像之外以及特征点数量过少的目标跟踪框,同时删除对应目标编号,在视频图像接收显示线程中将筛选后的跟踪框结果叠加到视频图像中;

步骤6,接收到新的一帧视频图像后,此时另外一个线程中的目标识别器产生了新的识别结果,执行下列子步骤:

步骤6.1,计算前一帧每个目标框和目标识别器最新结果中各个目标框之间的匹配分数,得到匹配分数矩阵,目标框之间的匹配分数计算中采用了目标不同方向上的相对位移比例和目标框的交并比相结合的方式,具体公式如下:

$$\text{score} = (1 - \text{IOU}) * (|x_k - x_{k-1}| / w_{k-1}) * (|y_k - y_{k-1}| / h_{k-1}) * 100$$

式中:IOU为两个矩形框的交并比, x_k 、 y_k 、 x_{k-1} 、 y_{k-1} 、 h_{k-1} 、 w_{k-1} 分别为当前帧中目标框左上角坐标、前一帧目标框左上角坐标、前一帧目标框的高度和宽度信息;针对匹配分数矩阵,利用Kuhn-Munkres匈牙利算法得到前后帧中目标框的最优匹配结果矩阵;

步骤6.2,根据步骤6.1得到的匹配结果对跟踪框进行更新,删除不满足条件的跟踪框以及添加新识别出来的目标框;具体步骤如下:

步骤6.2.1,如果当前帧的目标框能与前一帧的某个目标框能匹配上,则匹配计数器增加1,如果匹配计数器累计增加到或超过阈值3时,则确认接受该跟踪框结果,并利用卡尔曼

滤波算法对该跟踪框位置和大小进行修正,得到最优的跟踪框估计结果,并在视频图像接收显示线程中将跟踪框结果叠加到视频图像中;

步骤6.2.2,如果当前帧的目标框不能与前一帧的任何目标框匹配上,则非匹配计数器增加1,如果非匹配计数器累计增加到或超过阈值3时,则该目标跟踪框进入待删除状态;

步骤6.2.3,针对步骤6.2.2得到的待删除的跟踪框,将其分别与前一帧的所有目标框进行计算交并比IOU,如果IOU的值均小于等于阈值0.8,则说明该待删除的跟踪框是新识别出来的目标框,即识别出了新的目标,随后按照步骤3对该跟踪框建立一个新的目标跟踪器以及存储相应信息,并对该跟踪框进行编号,编号原则是在当前正在使用的最大编号基础上递增,递增后的编号如果超过了0-19的范围,则从0开始递增,使用当前所有目标框没有正在使用的编号,最后在视频图像显示线程中将跟踪框结果叠加到视频图像中;

步骤6.2.4,如果待删除的跟踪框与前一帧中某个目标框交并比IOU值大于0.8,则说明该跟踪框是真实无法与任何目标框匹配的而确认要删除的,随后删除该目标框的相关信息和目标编号;

步骤7,对已跟踪的人物目标进行提取运动人体行为的时空特征点,通过时空特征点映射信息建立描述符;以每个时空特征点为中心选取一个时空立方体,计算立方体内各点的Haar特征;

随后利用训练好的SVM支持向量机对提取的特征进行分类,判断人物行为是否存在异常,如奔跑、爬行、跳跃和/或攀爬,从而为初步区分越境人员和正常居民提供依据;

步骤8,在接收到新的一帧多光谱图像后,根据目标识别器线程中是否产生新的结果重复执行步骤5、步骤6和步骤7,直至视频最后一帧图像则结束。

2. 根据权利要求1所述的智能识别跟踪方法,其特征在于,所述步骤3.1中还包括:

基于深度学习框架训练SSD目标识别算法模型,该模型可以用于识别复杂背景下的人物;模型训练集使用的图像宽高尺寸为640x512像素,训练的迭代步数为100000步。

3. 根据权利要求1所述的智能识别跟踪方法,其特征在于,所述步骤3.2中还包括:

利用TensorRT对模型进行序列化,并加速模型推理过程,序列化后的模型对输入的红外多光谱图像进行目标识别,其中目标识别的置信度阈值为0.5,当识别框的置信度得分大于等于置信度阈值时,则会在视频帧上显示该识别框,同时输出结果还包含了识别框中目标的类别以及识别框的位置和大小和编号(x,y,w,h,id),其中x、y为该识别框的左上角坐标,w、h为该识别框的宽度和高度尺寸,id为该识别框的编号。

4. 根据权利要求1所述的智能识别跟踪方法,其特征在于,所述步骤5.1中还包括:

所述卡尔曼滤波算法的状态方程和观测方程分别为:

$$X_k = AX_{k-1} + W_k \quad (1)$$

$$Z_k = HX_k + V_k \quad (2)$$

式中A是状态转移矩阵, W_k 是均值为0,协方差矩阵为Q,且服从正态分布的过程噪声,H是测量矩阵, V_k 是均值为0,协方差矩阵为R,且服从正态分布的测量噪声,且 W_k 和 V_k 互不相关;考虑需要得到目标框的位置和大小信息,因此描述目标运动状态的状态向量X和观测向量Z分别为:

$$X = (c_x, c_y, v_x, v_y, h, w)^T \quad (3)$$

$$Z = (c_x, c_y, h, w)^T \quad (4)$$

式(3)和(4)中的 c_x 、 c_y 为目标框的中心点坐标值, v_x 、 v_y 为目标框的沿x和y方向移动的速度大小, h 、 w 为目标框的高度和宽度大小;

在确定状态向量 X 和观测向量 Z 后,则状态转移矩阵 A 和测量矩阵 H 分别为:

$$A = \begin{bmatrix} 1 & 0 & \Delta T & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

过程噪声的协方差矩阵 Q 和测量噪声的协方差矩阵 R 分别初始化为:

$$Q = \begin{bmatrix} 0.05 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.05 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.05 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.05 \end{bmatrix} \quad (7)$$

$$R = \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix} \quad (8)。$$

5. 根据权利要求1所述的智能识别跟踪方法,其特征在于,所述步骤5.1中的计算公式如下:

$$h_k = h_{k-1} * \text{median} \quad (9)$$

$$w_k = w_{k-1} * \text{median} \quad (10)$$

$$x_k = x_{k-1} + \text{median}_x - (w_k - w_{k-1}) / 2 \quad (11)$$

$$y_k = y_{k-1} + \text{median}_y - (h_k - h_{k-1}) / 2 \quad (12)$$

公式(9)-(12) h_k 、 w_k 、 x_k 、 y_k 分别表示当前帧中目标框的高度、宽度、左上角横坐标、左上角纵坐标值; median 为距离比值的中位数, median_x 、 median_y 则分别为前后两帧中对应特征点的横坐标和纵坐标差值的中位数。

6. 根据权利要求1所述的智能识别跟踪方法,其特征在于,所述步骤6.1中的具体公式如下:

$$\text{score} = (1 - \text{IOU}) * (|x_k - x_{k-1}| / w_{k-1}) * (|y_k - y_{k-1}| / h_{k-1}) * 100 \quad (13)$$

式(13)中IOU为两个矩形框的交并比, x_k 、 y_k 、 x_{k-1} 、 y_{k-1} 、 h_{k-1} 、 w_{k-1} 分别为当前帧中目标框左上角坐标、前一帧目标框左上角坐标、前一帧目标框的高度和宽度信息;针对匹配分数矩阵,利用Kuhn-Munkres匈牙利算法得到前后帧中目标框的最优匹配结果矩阵。

7. 根据权利要求1所述的智能识别跟踪方法,其特征在于,在步骤2中,所述目标编号为0-19的20个号码。

8. 根据权利要求1所述的智能识别跟踪方法,其特征在于,在步骤7中,所述描述符的构

建采用时空Haar特征,每个点均采用一个7维特征向量来表示,将所有点的特征向量进行求和则得到该点的描述向量,如下式所示:

$$D = (\Sigma L_x, \Sigma L_y, \Sigma L_t, \Sigma L_{xt}, \Sigma L_{yt}, \Sigma L_{xy}, \Sigma L_{xyt}) \quad (14)$$

式中: L_x 、 L_y 、 L_t 、 L_{xt} 、 L_{yt} 、 L_{xy} 、 L_{xyt} 为描述人物运动时空信息的7种时空Haar特征,能较好的表示一定范围内人物运动变化信息。

9. 根据权利要求1所述的智能识别跟踪方法,其特征在于,在步骤7中,所述SVM核函数采用径向基函数,如下式所示:

$$k(x, x') = \exp\left(\frac{-\|x-x'\|^2}{2\sigma^2}\right) \quad (15)$$

式中: x 和 x' 为需要进行高维空间映射的两个特征向量, $\|x-x'\|^2$ 为计算 x 和 x' 特征向量之间的欧几里得距离, σ 是一个自由设定的参数。

10. 一种计算机可读存储介质,其上存储有计算机程序,其特征在于,该计算机程序被处理器执行实现如权利要求1至9任一项所述的智能识别跟踪方法。

茂密丛林区域对越境人员的智能识别跟踪方法及存储介质

技术领域

[0001] 本发明属于红外多目标智能识别跟踪技术领域,涉及一种茂密丛林区域对越境人员的智能识别跟踪方法及存储介质,具体涉及一种基于深度学习的针对边境茂密丛林区域越境人员的红外多光谱智能识别跟踪方法及系统。

背景技术

[0002] 红外多目标识别与跟踪是计算机视觉领域的热点问题,由于红外图像的特殊性,其也广泛应用于全天候视频监控、夜间无人驾驶等领域,红外目标识别部分主要是从红外视频图像中确定目标的类别、位置和大小,红外目标跟踪主要是基于目标的初始状态,准确预测目标在后续视频帧中的位置和大小。随着深度学习相关技术的不断发展,卷积神经网络也被人们广泛的应用于红外目标识别与跟踪技术领域,并在该领域取得了极大的成功。

[0003] 目前主流的目标识别算法主要包括one-stage和two-stage两类方法。其中one-stage目标检测算法(也称one-shot object detectors),其特点是一步到位,均匀的在图片的各个位置进行密集抽样,速度相对较快。另外一类目标检测算法是two-stage的,如Faster R-CNN算法首先生成候选框(region proposals,可能包含目标的区域),然后再对每个候选框进行分类以及修正位置,这类算法相对较慢,因为它需要多次运行检测和分类流程。而one-stage检测方法,仅仅需要送入网络一次就可以预测出所有的边界框,因而速度较快,非常适合边缘计算的移动端。最典型的one-stage检测算法包括YOLO、SSD、SqueezeDet以及DetectNet等,其中SSD算法同时对浅层和深层特征进行识别,并减少了RPN层,相对于其它算法,其在算法速度和识别准确度上有更大的优势。

[0004] 目前的目标跟踪算法主要分为两类:Detection Free Tracking (DFT,无检测跟踪)和Detection Based Tracking (DBT,基于检测跟踪)。DFT算法需要人工标注第一帧图像中的目标,之后则目标检测和跟踪同时进行,相比于DBT算法,DFT算法需要事先人工标注目标的类型和位置,并且对于不是在第一帧中出现的目标或者在后续帧中消失的目标,DFT算法将无法处理。对于DBT算法则在多目标跟踪过程中包含了一个独立的检测过程,也就是说在跟踪之前每一幅图像中的目标都是已经检测出来的,因此这类算法的跟踪性能较大程度依赖于检测器目标识别能力的强弱,随着目标跟踪算法的发展,该类算法也成为了多目标跟踪算法的主流。

[0005] 在红外多目标的识别与跟踪算法中,红外目标识别起到极其重要的作用,由于帧差法和背景差法的局限性,当光线较差以及目标发生遮挡时,相应算法的识别效果极大的下降,从而极大的影响后续目标编号的确定和目标轨迹的跟踪,因此对红外目标识别算法的识别精度提出了极高的要求,不能出现误、漏检的情况,并且在红外多目标跟踪过程中,基于ARM架构的嵌入式终端设备对多目标跟踪算法的实时性提出了极高的要求,同时运动目标速度较大、运动目标短时遮挡等也是目前多目标跟踪的重难点。

[0006] 现有的目标检测使用方法通常为背景差法,即从静止的背景中分离出运动的目标,通常利用背景建模方式结合帧差法,检测出前景目标,从而实现目标的检测,但是由于

背景差法的局限性,在光线较差以及目标发生遮挡时检测的效果较差,从而为后续跟踪效果带来较大影响,并且现有检测算法无法针对检测结果中的人物进行进一步的行为分析,针对人物的异常行为给出报警提示。

[0007] 因此,研制一种针对边境茂密丛林区域越境人员的智能识别跟踪方法,将有利于边防人员及时发现追踪非法越境人员。

发明内容

[0008] 本发明要解决的技术问题在于针对边境茂密丛林区域的越境人员所具有的隐蔽性强的特点,解决传统方法较难准确、快速地发现 and 跟踪越境人员问题。

[0009] 本发明的总体构思是采用红外多光谱图像数据,利用动植物和人物之间红外多光谱特征显著不同的特点,图像数据经过降维处理并结合Single Shot MultiBox Detector神经网络模型后,能同时保证了人物识别的速度和精度。由于边境区域有正常生活的居民,本发明采用机器学习算法对人物进行多目标跟踪并进行异常行为分析,能够有效区分非法越境人员和正常生活的居民,提高了边防人员对边境区域中人物甄别的效率。整个算法架构采用GPU/CPU异构并行的方式,人物识别、跟踪、行为分析等多个线程同时运行,能够满足嵌入式系统实时性的要求。

[0010] 本发明解决上述技术问题所采用的技术方案为:

[0011] 一种茂密丛林区域对越境人员的智能识别跟踪方法,包括:针对多目标识别,采用基于深度学习的目标识别算法SSD。为了提高识别准确度和识别速度,输入采用多光谱图像并经过PCA进行特征降维,采用TensorRT对模型进行加速推理,同时将算法移植到基于ARM架构的linux系统,更便于实际嵌入式设备的推广应用。

[0012] 针对多目标跟踪,采用金字塔LK光流法和卡尔曼滤波相结合的算法,依据上一帧的目标跟踪框对当前帧的目标进行预测,通过计算匹配分数对前后两帧的目标框进行配对,根据配对结果对当前帧的目标框进行更新,删除不满足判定条件的目标框以及添加新增的目标框,同时在跟踪过程中利用多目标识别器的识别结果对跟踪结果进行校正。

[0013] 针对行人异常检测,采用时空Haar特征提取和SVM支持向量机分类的方法,有效的对越境人员的异常行为进行识别,支持向量机SVM是基于最大超平面间隔原则,是一个二次优化的问题,使最终的超平面是全局最优解;SVM在解决非线性样本数据时,通过将非线性变换到高维特征空间,在此空间构建核函数实现原空间的非线性判别函数,使学习机器有较好的推广能力,同时解决了维数问题,其算法的复杂度与维数无关。

[0014] 采用上述的识别跟踪方法,在整个过程中采用CPU/GPU异构并行处理策略,视频流接收显示、多目标识别、多目标跟踪、人物异常行为识别等多个线程同时进行,从而实现多目标的实时跟踪。

[0015] 该方法及系统对各种复杂场景、目标运动速度较大和目标遮挡等实现了实时的红外多目标识别与跟踪,具有较高的识别速度、识别准确度和鲁棒性。

[0016] 具体的,本发明的一种茂密丛林区域对越境人员的智能识别跟踪方法,包括以下步骤:

[0017] 1) 读取红外多光谱图像

[0018] 2) 利用主成分分析技术(PCA)对红外多光谱图像进行降维处理,保留特征值较大

的前5个光谱特征,去掉冗余的光谱特征信息。

[0019] 3) 调用SSD目标识别算法对经PCA降维后的光谱图像特征进行多目标识别,包括的子步骤如下:

[0020] 3.1) 基于深度学习框架训练SSD目标识别算法模型,该模型可以用于识别复杂背景下的人物。模型训练集使用的图像宽高尺寸为640x512像素,训练的迭代步数为100000步。

[0021] 3.2) 将目标识别的SSD代码转换为C++语言,并与训练后的模型一起移植到基于ARM架构的linux系统,系统环境配置为:ubuntu18.04系统、opencv3.4.6、cuda10.2、cudnn8.0、TensorRT7.1。

[0022] 3.3) 利用TensorRT对模型进行序列化,并加速模型推理过程,序列化后的模型对输入的红外多光谱图像进行目标识别,其中目标识别的置信度阈值为0.5,当识别框的置信度得分大于等于置信度阈值时,则会在视频帧上显示该识别框,同时输出结果还包含了识别框中目标的类别以及识别框的位置和大小和编号(x,y,w,h,id),其中x、y为该识别框的左上角坐标,w、h为该识别框的宽度和高度尺寸,id为该识别框的编号。

[0023] 4) 在第一次对红外视频进行多目标识别并得到识别结果后,对多目标跟踪器进行初始化针对每个目标识别框建立一个跟踪器,同时使用FAST特征点检测器检测每个目标识别框的特征点信息,并存储相应目标识别框的位置、大小、目标类型、目标编号和特征点信息,其中目标编号为0-19的20个号码,同一帧图像中的目标编号没有重复,并且同一个目标在前后帧中的编号保持不变。

[0024] 5) 在接收到新的一帧视频图像后,如果此时另外一个线程中的目标识别器没有产生新的识别结果,包括的子步骤如下:

[0025] 5.1) 基于卡尔曼滤波算法的公式(1)和前一帧的目标识别框结果进行预测,得到对应目标框在当前帧中位置和大小预测结果。其中卡尔曼滤波算法的状态方程和观测方程分别为:

$$[0026] \quad X_k = AX_{k-1} + W_k \quad (1)$$

$$[0027] \quad Z_k = HX_k + V_k \quad (2)$$

[0028] 式中A是状态转移矩阵, W_k 是均值为0,协方差矩阵为Q,且服从正态分布的过程噪声,H是测量矩阵, V_k 是均值为0,协方差矩阵为R,且服从正态分布的测量噪声,且 W_k 和 V_k 互不相关。考虑需要得到目标框的位置和大小信息,因此描述目标运动状态的状态向量X和观测向量Z分别为:

$$[0029] \quad X = (c_x, c_y, v_x, v_y, h, w)^T \quad (3)$$

$$[0030] \quad Z = (c_x, c_y, h, w)^T \quad (4)$$

[0031] 式(3)和(4)中的 c_x 、 c_y 为目标框的中心点坐标值, v_x 、 v_y 为目标框的沿x和y方向移动的速度大小,h、w为目标框的高度和宽度大小。

[0032] 在确定状态向量X和观测向量Z后,则状态转移矩阵A和测量矩阵H分别为:

$$[0033] \quad A = \begin{bmatrix} 1 & 0 & \Delta T & 0 & 0 & 0 \\ 0 & 1 & 0 & \Delta T & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (5)$$

$$[0034] \quad H = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \quad (6)$$

[0035] 过程噪声的协方差矩阵Q和测量噪声的协方差矩阵R分别初始化为:

$$[0036] \quad Q = \begin{bmatrix} 0.05 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.05 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.05 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0.05 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.05 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0.05 \end{bmatrix} \quad (7)$$

$$[0037] \quad R = \begin{bmatrix} 0.1 & 0 & 0 & 0 \\ 0 & 0.1 & 0 & 0 \\ 0 & 0 & 0.1 & 0 \\ 0 & 0 & 0 & 0.1 \end{bmatrix} \quad (8)$$

[0038] 5.2) 利用前一帧目标的特征点信息和金字塔LK光流法计算当前帧中与前一帧目标特征点相对应的特征点位置信息,分别计算前一帧和当前帧中不同特征点之间的距离,这样在前后两帧图像中便得到了相同数量的距离数据,接着计算相同两个特征点在前后两帧图像中距离的比值 $d_{\text{当前帧}}/d_{\text{前一帧}}$,计算距离比值的中位数(例如在前后两帧中分别有100个特征点,则分别得到4950个距离数据,以及4950距离比值数据,并计算这4950个距离比值的中位数),利用中位数更新当前帧中目标框的大小和目标框左上角坐标位置。具体计算公式如下:

$$[0039] \quad h_k = h_{k-1} * \text{median} \quad (9)$$

$$[0040] \quad w_k = w_{k-1} * \text{median} \quad (10)$$

$$[0041] \quad x_k = x_{k-1} + \text{median}_x - (w_k - w_{k-1}) / 2 \quad (11)$$

$$[0042] \quad y_k = y_{k-1} + \text{median}_y - (h_k - h_{k-1}) / 2 \quad (12)$$

[0043] 公式(9) - (12) h_k 、 w_k 、 x_k 、 y_k 分别表示当前帧中目标框的高度、宽度、左上角横坐标、左上角纵坐标值。 median 为距离比值的中位数, median_x 、 median_y 则分别为前后两帧中对应特征点的横坐标和纵坐标差值的中位数。

[0044] 5.3) 基于步骤5.2得到当前帧中目标框位置和大小观测值,利用卡尔曼观测方程(2),在均方误差最小的情况下,对当前的状态进行修正,从而得到目标框位置和大小最优估计值,并保证前后两帧中同一目标的编号不变。随后对当前帧中的目标框重新进行特征点检测,增加目标框的特征点数量以便在后续帧中更好的跟踪目标,同时对跟踪框进行筛选,剔除矩形框完全位于图像之外以及特征点数量过少的目标跟踪框(特征点个数小于2个),同时删除对应目标编号,在视频图像接收显示线程中将筛选后的跟踪框结果叠加到视频图像中。

[0045] 6) 在接收到新的一帧视频图像后,如果此时另外一个线程中的目标识别器产生了新的识别结果,包括的子步骤如下:

[0046] 6.1) 计算前一帧每个目标框和目标识别器最新结果中各个目标框之间的匹配分数,得到匹配分数矩阵,目标框之间的匹配分数计算中采用了目标不同方向上的相对位移比例和目标框的交并比相结合的方式,具体公式如下:

$$[0047] \quad \text{score} = (1 - \text{IOU}) * (|x_k - x_{k-1}| / w_{k-1}) * (|y_k - y_{k-1}| / h_{k-1}) * 100 \quad (13)$$

[0048] 式(13)中IOU为两个矩形框的交并比, x_k 、 y_k 、 x_{k-1} 、 y_{k-1} 、 h_{k-1} 、 w_{k-1} 分别为当前帧中目标框左上角坐标、前一帧目标框左上角坐标、前一帧目标框的高度和宽度信息。针对匹配分数矩阵,利用Kuhn-Munkres匈牙利算法得到前后帧中目标框的最优匹配结果矩阵。

[0049] 6.2) 根据6.1得到的匹配结果对跟踪框进行更新,删除不满足条件的跟踪框以及添加新识别出来的目标框。具体步骤如下:

[0050] 6.2.1) 如果当前帧的目标框能与前一帧的某个目标框能匹配上,则匹配计数器增加1,如果匹配计数器累计增加到或超过阈值3时,则确认接受该跟踪框结果,并利用卡尔曼滤波算法对该跟踪框位置和大小进行修正,得到最优的跟踪框估计结果,并在视频图像接收显示线程中将跟踪框结果叠加到视频图像中。

[0051] 6.2.2) 如果当前帧的目标框不能与前一帧的任何目标框匹配上,则非匹配计数器增加1,如果非匹配计数器累计增加到或超过阈值3时,则该目标跟踪框进入待删除状态。

[0052] 6.2.3) 针对6.2.2得到的待删除的跟踪框,将其分别与前一帧的所有目标框进行计算交并比IOU,如果IOU的值均小于等于阈值0.8,则说明该待删除的跟踪框是新识别出来的目标框,即识别出了新的目标,随后类似步骤3对该跟踪框建立一个新的目标跟踪器以及存储相应信息,并对该跟踪框进行编号,编号原则是在当前正在使用的最大编号基础上递增,递增后的编号如果超过了0-19的范围,则从0开始递增,使用当前所有目标框没有正在使用的编号,最后在视频图像显示线程中将跟踪框结果叠加到视频图像中。

[0053] 6.2.4) 如果待删除的跟踪框与前一帧中某个目标框交并比IOU值大于0.8,则说明该跟踪框是真实无法与任何目标框匹配的而确认要删除的,随后删除该目标框的相关信息和目标编号。

[0054] 7) 对已跟踪的人物目标进行提取运动人体行为的时空特征点,通过时空特征点映射信息建立描述符,描述符的构建采用维度相对较少的时空Haar特征。以每个时空特征点为中心选取一个时空立方体,计算立方体内各点的Haar特征,每个点均采用一个7维特征向量来表示,将所有点的特征向量进行求和则得到该点的描述向量,如式(14)所示:

$$[0055] \quad D = (\Sigma L_x, \Sigma L_y, \Sigma L_t, \Sigma L_{xt}, \Sigma L_{yt}, \Sigma L_{xy}, \Sigma L_{xyt}) \quad (14)$$

[0056] 式中: L_x 、 L_y 、 L_t 、 L_{xt} 、 L_{yt} 、 L_{xy} 、 L_{xyt} 为描述人物运动时空信息的7种时空Haar特征,能较好的表示一定范围内人物运动变化信息

[0057] 随后利用训练好的SVM支持向量机对提取的特征进行分类,判断人物行为是否存在异常,如奔跑、爬行、跳跃、攀爬等,从而为初步区分越境人员和正常居民提供依据,其中SVM核函数采用径向基函数,如式(15)所示。

$$[0058] \quad k(x, x') = \exp\left(\frac{-\|x-x'\|^2}{2\sigma^2}\right) \quad (15)$$

[0059] 式中: x 和 x' 为需要进行高维空间映射的两个特征向量, $\|x-x'\|^2$ 为计算 x 和 x' 特

征向量之间的欧几里得距离, σ 是一个自由设定的参数。

[0060] 8) 在接收到新的一帧多光谱图像后, 根据目标识别器线程中是否产生新的结果重复执行步骤5)、6)、7), 直至视频最后一帧图像则结束程序。

[0061] 本发明的有益效果:

[0062] (1) 对边境茂密丛林区域的红外多光谱图像数据进行主成分分析降维处理, 利用 Single Shot MultiBox Detector 神经网络模型对红外多光谱图像中的人物进行智能识别, 获得比传统方法更高的准确率。

[0063] (2) 结合LK光流法和卡尔曼滤波算法对边境茂密丛林区域中出现的人物进行多目标跟踪获取人物运动轨迹, 采用支持向量机算法对人物进行异常行为分析, 有效区分人物是否是越境人员或境内正常生活的居民。

[0064] (3) 算法采用GPU/CPU异构并行计算方式, 能够达到嵌入式系统实时应用的速度要求。

附图说明

[0065] 图1: 本发明的方法使用的SSD网络结构示意图。

[0066] 图2: 本发明的方法的整体架构示意图。

[0067] 图3: 本发明的方法的具体流程图。

具体实施方式

[0068] 按照如图3所示的方法流程图实现本发明的智能识别跟踪方法。

[0069] 将本发明的智能识别跟踪方法编制成计算机程序存储在计算机可读存储介质, 该计算机程序被处理器执行实现本发明的智能识别跟踪方法。

[0070] 本实施例中, 红外搜索跟踪设备或光电吊舱中, 其硬件架构为Xilinx FPGA+NVIDIA TX2异构模式, 包括有本发明的计算机可读存储介质。FPGA完成图像读取, 并将图像传输给TX2的GPU模块完成SSD识别算法与基于TensorRT模型的加速算法。接着完成识别结果判定, 包括:

[0071] a) 未成功识别结果

[0072] 将卡尔曼滤波置于FPGA内通过并行处理进行加速, 在对目标位置做出预测之后, 快速通过FPGA完成金字塔图像的重构, 接下来将结果传输至TX2模块利用L-K光流法标定目标具体位置与大小。在删除冗余跟踪框后, 将结果图像传输给FPGA, 再次硬件平台上利用卡尔曼滤波来对目标位置做出预测并持续跟踪。

[0073] b) 成功识别结果

[0074] 继续在TX2模块内利用匈牙利算法识别结果并将跟踪框与目标配对; 在完成对跟踪框的进一步迭代更新后, 将结果图像传输至FPGA, 在该硬件平台上利用卡尔曼滤波来对目标位置做出预测并持续跟踪。

[0075] 以上a) 或b) 步骤完成后, 继续依赖FPGA在并行处理方面的优势, 开展时空Haar特征提取和支持向量机来判别目标是否异常, 最终将异常目标剔除。

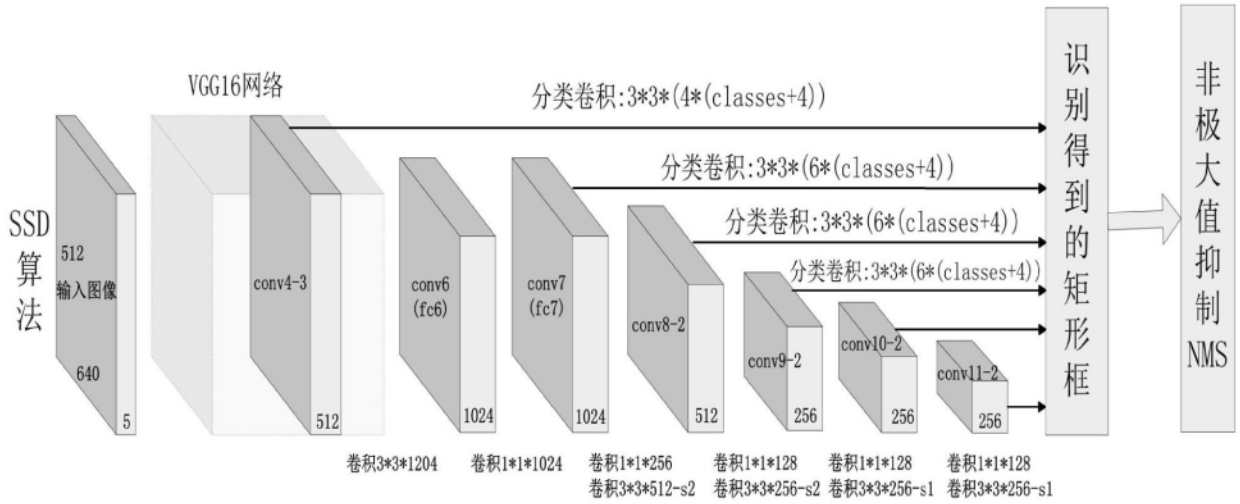


图1

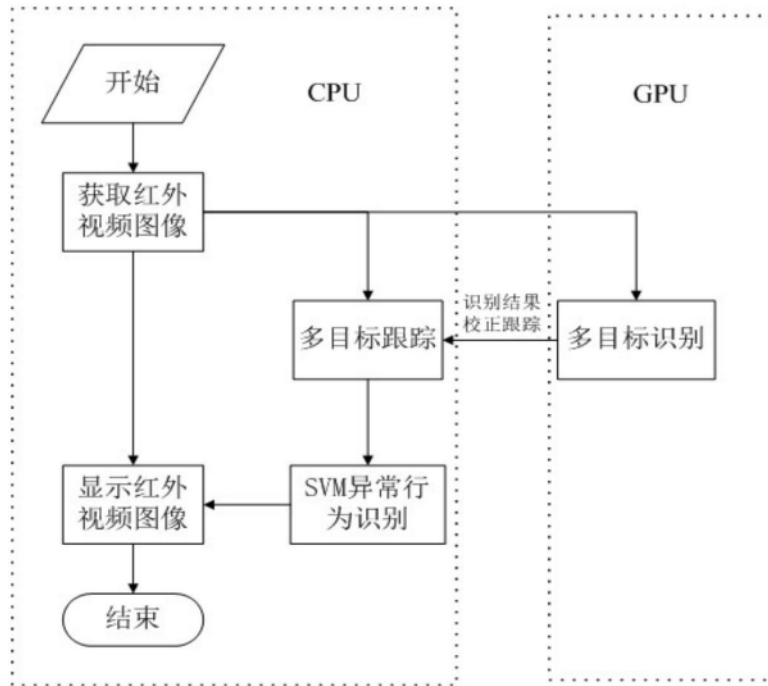


图2

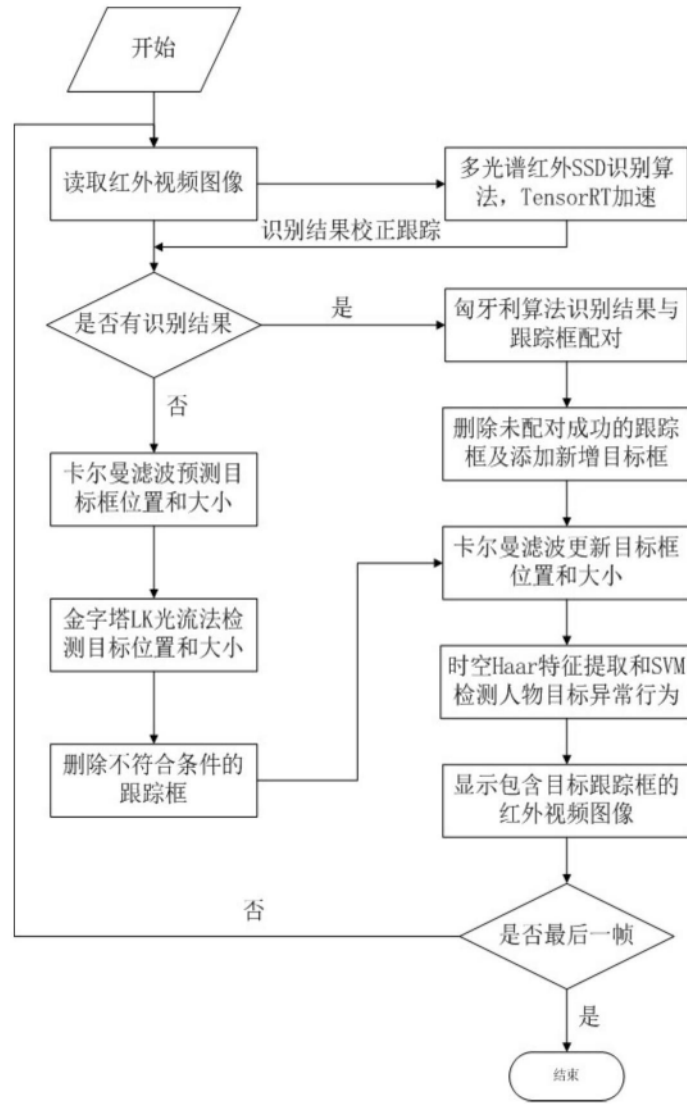


图3