



(12) 发明专利申请

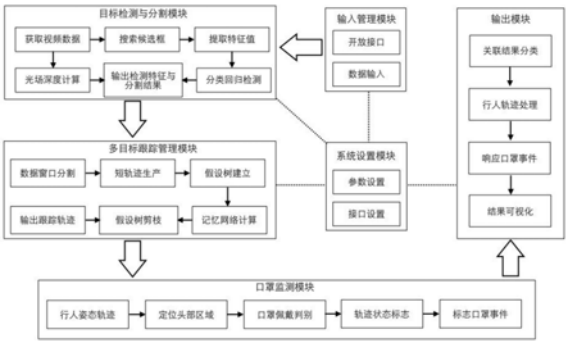
(10) 申请公布号 CN 111832400 A  
(43) 申请公布日 2020. 10. 27

(21) 申请号 202010497684.6  
(22) 申请日 2020.06.04  
(71) 申请人 北京航空航天大学  
地址 100191 北京市海淀区学院路37号  
(72) 发明人 盛浩 叶珍 张洋 王帅 吴玉彬  
(74) 专利代理机构 北京科迪生专利代理有限公司 11251  
代理人 安丽  
(51) Int. Cl.  
G06K 9/00 (2006.01)  
G06K 9/34 (2006.01)  
G06N 3/04 (2006.01)

权利要求书5页 说明书9页 附图6页

(54) 发明名称  
一种基于概率神经网络的口罩佩戴情况的  
监测系统及方法

(57) 摘要  
本发明涉及一种基于概率神经网络的口罩佩戴情况的监测系统及方法,包括:输入管理模块,负责处理输入视频的读取和预处理;目标检测与分割模块负责对于视频中的每个目标进行检测和识别,并分割出行人口罩部分进行检测;多目标跟踪关联模块对于每个视频中的检测结果在该视频内进行多对多关联;口罩监测模块,完成初步轨迹关联后对置信度较低和离散轨迹;输出模块,设立在统一的数据中心节点,支持多路跨摄像机数据输出,将分别对各个场景进行多目标跟踪,统一汇聚输出计算结果;系统设置模块用于配置多个模块中所用的网络模型的训练。本发明为视频监控、行为分析等领域提供了技术支持。



1. 一种基于概率神经网络的口罩佩戴情况的监测系统,其特征在于,包括:输入管理模块、目标检测与分割模块、多目标跟踪关联模块、口罩监测模块、输出模块和系统设置模块;

输入管理模块,为监控摄像机拍摄的输入视频上传和处理提供接口,也包含用户所使用的操作界面,是目标检测与分割模块、多目标跟踪关联模块、口罩监测模块、输出模块、系统设置模块的入口;

目标检测与分割模块,利用PCB卷积神经网络以及区域生成网络对于输入视频的每一帧的每一个行人进行检测和识别;首先利用选择搜索算法确定视频每一帧中的行人检测的候选框,所述候选框为能将单个行人完全包住的矩形框;然后利用卷积神经网络提取候选框的特征,利用区域生成网络对候选框进行筛选确定最终的行人检测框,最终的行人检测框即为能将单个行人完全包住的最小矩形框;由目标检测模块检测和识别处理后得到行人检测特征,行人检测特征包括检测后得到的每个检测框以及每个行人的特征,并利用分类器对行人检测框进行分类,根据分类结果对检测框进行对齐调整,并利用光场的深度计算辅助分类回归检测算法对输出的检测数据进行分割,实现行人的头部部分的分割得到分割结果;所述检测数据和分割结果将作为多目标跟踪关联模块、口罩监测模块和输出模块的数据基础;

多目标跟踪关联模块,负责对于每个视频得到的目标检测与分割模块的检测结果在该视频内进行关联;在多目标跟踪的约束条件下,利用基于多假设概率神经网络的多目标关联算法,完成对所有检测结果的关联任务;所述关联任务即如果为同一个行人则匹配,否则不匹配;所述多目标跟踪的约束条件包括:空间约束、时间约束、可允许的漏检数目约束和相似度约束;所述空间约束是指产生关联的两两检测框在视频中的位置距离不可超出一定阈值;所述时间约束是指产生关联的两两检测框在时间上的交集必须为空集;所述可允许的漏检数目约束是指关联后导致时间不连续轨迹的漏检检测的数目必须在设定阈值之内;所述相似度约束是指判断行人的两个检测是否应该产生关联关系的度量依据;所述目标关联模块包括多假设概率神经网络和剪枝算法;所述多假设概率神经网络提供基于多假设概率神经网络的多目标关联算法,将检测序列作为轨迹信息作为输入,基于多假设概率神经网络的多目标关联算法融合基于长短期记忆网络的概率神经网络来学习轨迹的时序特征从而拟合输入轨迹的未来变化趋势,从趋势特征上对短轨迹进行匹配得到相似度分数,根据相似度来对短轨迹进行关联;多假设树的节点生成取决于短轨迹间的相似度分数;所述剪枝算法为避免由于多假设树层数增长带来的计算时间指数增长和空间复杂度的指数增长,利用基于长短期记忆网络的概率神经网络对多假设概率神经网络生成的多假设树进行剪枝,该深度模型将多假设树的每个分支作为输入,输出结果为每个分支为同一个行人的置信度分数,此时多假设树的每个节点为行人检测的特征即外观特征,最后将置信度最高的分支保留,其余分支进行剪枝;多目标跟踪关联模块中所参考的历史特征信息为行人短轨迹的轨迹外观特征和运动特征,所述交互信息即产生交互的相行人轨迹之间的关系特征;最后得到结果全局最优的行人关联结果,即匹配的行人轨迹;

口罩监测模块,负责利用目标检测与分割模块的检测数据和分割结果使用姿态估计网络,判断行人的朝向,即背面、正面和侧面,对行人检测和分割结果进行头部面部定位和检测,对正面和侧面的行人面部,使用口罩检测算法进行口罩检测,得到口罩检测结果后将对面部口罩的检测结果进行实时在线关联跟踪,得到实时的面部口罩轨迹的跟踪状态的变

化,并定义三个口罩事件:正常、提示和警告,将得到的口罩轨迹的跟踪状态的变化按这三个口罩事件进行分类,得到分类结果;

输出模块,将多目标跟踪关联模块得到的结果全局最优的行人关联结果和口罩监测模块得到的口罩轨迹跟踪状态的口罩事件分类结果进行处理后将行人轨迹的跟踪结果可视化并实时对于跟踪过程中的口罩事件作出响应;输出模块将多目标跟踪关联模块输出的行人关联结果进行分类,同一个行人的所有检测框设置为同一颜色,不同行人其检测框颜色设为不同;然后对于行人关联结果所生成的轨迹进行处理,所生成的轨迹即同一行人的连续帧形成的检测序列;所述处理包括对于轨迹的插值处理和平滑处理;对于口罩检测模块的口罩事件分类结果根据优先级分别作出响应,所述响应的方式以及优先级均由系统设置模块的参数所控制,处理完毕后将结果进行保存并可视化;

系统设置模块,用于配置输入管理模块、目标检测与分割模块、多目标跟踪关联模块、以及输出模块的参数配置;对于输入管理模块,系统设置模块设置输入视频的最大限制参数,通道数限制参数以及输入目标视频的地址;对于目标检测与分割模块,系统设置模块设置目标检测生成的细节,包括检测框的大小、宽高以及是否需要直接使用预训练的模型进行计算,若需要则设置模型的根目录;用户根据实际需要设置模型输出根目录,以便在PCB卷积神经网络以及区域生成网络模型在使用结束后的复用;对于多目标跟踪关联模块在运行代码时,首先读取系统设置模块中的参数数据,完成PCB卷积神经网络以及区域生成网络算法的参数配置,由用户实现对PCB卷积神经网络以及区域生成网络算法的透明性操作,提升输出结果的可调节性,实现操作决策的可控;对于口罩监测模块,用户自行设置三个事件的响应级别,默认的优先级别为:“警告”大于“提示”,“提示”大于“正常”;系统设置模块提供这三个事件的响应方式设置,用户根据需求进行个性化设置,内容包括提示框的显示内容、颜色、辅助语音和提示音。

2. 根据权利要求1所述的基于概率神经网络的口罩佩戴情况的监测系统,其特征在于:所述多目标跟踪关联模块中多目标关联算法具体实现如下:

(1) 首先将目标检测和分割模块得到的行人检测特征以用户设置的窗口数为单位进行分组,设第一个窗口为当前窗口 $W_c$ ;

(2) 以当前窗口 $W_c$ 所有检测为横纵轴构造相似度矩阵,行人检测的相似度计算方式间公式(1),即为两两行人检测间的运动特征即坐标、宽高的欧式距离与外观特征即检测特征距离的加权和,其中 $\text{Dist}_{\text{mot}}^{i,j}$ 表示行人检测i和j之间的运动特征的距离, $\text{Dist}_{\text{app}}^{i,j}$ 表示行人检测i和j之间的外观特征的距离,公式(2)中 $x$ 、 $y$ 、 $w$ 、 $h$ 分别指代行人检测的坐标和宽高,公式(3)中的 $a_k$ 表示行人检测i的外观特征的第k个元素, $b_k$ 表示行人检测j的外观特征的第k个元素,将得到的相似度分数 $S^{i,j}$ 用于构造相似度矩阵,根据相似度矩阵,计算得到最佳匹配的短轨迹集合,所述短轨迹是指生成的属于同一个行人的检测序列,即轨迹长度不超过当前窗口长度的轨迹集合;

$$S^{i,j} = w_{\text{mot}} \text{Dist}_{\text{mot}}^{i,j} + w_{\text{app}} \text{Dist}_{\text{app}}^{i,j} \quad (1)$$

$$(\text{Dist}_{\text{mot}}^{i,j})^2 = (x^i - x^j)^2 + (y^i - y^j)^2 + (w^i - w^j)^2 + (h^i - h^j)^2 \quad (2)$$

$$(\text{Dist}_{\text{app}}^{i,j})^2 = \sum_k (a_k - b_k)^2 \quad (3)$$

(3) 判断当前窗口 $W_c$ 是否是最后一个窗口,若是则进入步骤(4),或不是则将 $W_{c+1}$ 赋值为 $W_c$ ,进入步骤(2);

(4) 设第一个窗口为当前窗口 $W_c$ ;

(5) 取当前窗口 $W_c$ 中一条短轨迹 $k$ 作为当前节点,并建立父节点,判断当前窗口 $W_c$ 是否为最后一个窗口,若不是最后一个窗口,则进入步骤(6),若是最后一个窗口,则进入步骤(12);

(6) 计算短轨迹 $k$ 与下一个窗口 $W_{c+1}$ 中所有轨迹的相似度,轨迹之间的相似度计算通过多假设概率神经网络计算得到,多假设概率神经网络将两条短轨迹的每一帧的检测特征作为输入,在内部进行高维拼接和筛选,输出两两短轨迹的相似度分数,得到相似度高于用户设置阈值的短轨迹集合 $m$ ;

(7) 判断子节点数量是否超出用户设置的阈值,若是,则进入步骤(5),反之,则将短轨迹集合 $m$ 中所有短轨迹设置为短轨迹 $k$ 的子节点,进入步骤(8);

(8) 判断是否遍历完当前窗口 $W_c$ 中所有短轨迹,若是则进入步骤(9),反之,则进入步骤(5);

(9) 计算所得多假设树的层数是否在用户设置阈值范围内,若是则进入步骤(3),反之则进入步骤(10);

(10) 将建立的多假设树的每个分支作为输入,输入到长短期记忆网络的概率神经网络,计算得到每个分支为同一个行人的置信度分数;

(11) 根据置信度分数对多假设树进行剪枝,进入步骤(3);

(12) 选则全局分数最高的分支作为匹配的行人轨迹。

3. 一种基于概率神经网络的口罩佩戴情况的监测方法,其特征在于,实现如下:

(1) 用户首先通过登陆步骤进入到输入管理模块,在登陆界面,通过输入框记录用户输入信息,界面提交后,数据库通过比对输入信息和数据库中缓存的用户信息并匹配用户ID,分为匹配成功和匹配失败两种状态:若匹配成功,数据库通过ID查找该用户并记录相关输入及查询缓存信息并将用户的相关信息返回前端;若匹配失败,则弹出消息框,显示用户不存在;登陆成功后,根据用户定位到视频所在的绝对路径,上传视频;

(2) 目标检测与分割模块步骤:首先对于输入视频的每一帧,在其中利用选择性搜索的算法确定多个检测框的候选框,然后将图像依次输入到一个卷积神经网络,得到候选框特征,包括行人的坐标、宽高以及初步的卷积特征,然后利用一个区域生成网络对于每个候选框进行筛选,得到最终的检测框;使用分类器将所有的检测框按照是否对齐进行分类,完成分类后,对于未完全对齐的检测框利用回归器进一步的调整检测框位置;利用光场的深度计算辅助分类回归检测算法基于卷积特征对检测数据进行分割,利用光场的深度信息并基于行人检测的卷积特征对于图像区域进行头部分割;检测与分割结束后,将得到的检测分割结果保存到系统设置模块设置的根目录位置;

(3) 获得所有检测的检测分割结果后,将进行多目标跟踪,采用多假设概率神经网络的多目标跟踪关联算法,该多目标跟踪关联算是基于多假设跟踪框架,所述多假设跟踪框架分成多假设树的生成、多假设树的剪枝和最佳路径的选择三个阶段;多假设树的生成阶段,系统将多假设树的生成和深度学习相结合,利用基于长短时记忆神经网络的网络结构在处理时序特征上的优势,辅助生成多假设树,具体的生成实现步骤是:首先将视频以系统设置模块设置的帧数为窗口长度单元分割成多个不重叠的窗口,在每个窗口内的所有检测框之间进行匹配,生成最初的短轨迹,这些短轨迹被视为原子节点,即不可分,多目标跟踪的关

联任务即在这些窗口之间,将两两相邻窗口之间的原子节点相互匹配,为每一个潜在的节点即目标生成一棵假设树,生成这些假设树需要同时满足的条件是:①子节点与父节点所在的窗口序号差的绝对值不可跨越系统设置模块给出的窗口阈值数;②子节点与父节点的相似度,即由基于长短时记忆神经网络的网络模型所评估的分数,不可低于系统设置模块给出的阈值;③一个父节点的子节点的数目不可超过系统设置模块给出的阈值,将深度学习融合到剪枝算法中,利用深度模型记忆目标的外观特征,为每只分支是否为同一个目标的自信度进行打分,使行人轨迹在经历遮挡、拥挤的信息丢失的场景后恢复时能够根据外观的记忆与已有的轨迹进行匹配;在跟踪完所有的检测目标后,根据该框架算法为每棵树每个分支的打分,取最高分的分支作为最佳的匹配结果,最后将结果输出到系统设置模块所保存的根目录;

(4) 口罩监测模块为三个步骤:①行人姿态估计,②头部面部定位和检测,③口罩检测,④监测口罩事件,获得所有行人的检测和分割结果后,利用目标姿态估计网络识别行人的状态,包括正面、侧面和背面,从而判断行人的朝向;再利用行人的分割结果对行人进行头部和面部定位,并对头部部分的检测框进行修正后提取特征,该特征利用目标检测与分割模块中的特征提取的卷积神经网络提取出的目标的头部深度外观特征;利用这些外观特征利用口罩检测网络对于一定朝向,包括正面和侧面的行人进行口罩检测,实时判断出行人是否佩戴口罩,并将检测结果赋予每个行人对象,与行人的ID相关联,检测结果包括:佩戴、未佩戴、未知,其中佩戴表示系统能通过网络识别出口罩的特征,未佩戴是指系统对于正面后侧面朝向的行人进行识别后未能检测出口罩,未知则是对于背面朝向的行人,系统本身无法识别行人脸部从而无法识别并判断口罩的佩戴情况;口罩监测的最后一步是监测口罩事件,利用口罩检测步骤所得到的检测结果,将每个ID的行人进行实时检测,并根据口罩检测结果的实时变化对行人的状态进行评估并反应口罩事件,口罩事件具体划分为三个:①正常,②提示,③警告;这三个事件也分别对应了三个颜色,即绿、黄、红,以及三个优先级,用户根据需求设置优先级,默认设置为:事件被判为正常事件,即为绿色事件,为三级优先,正常事件是指从该目标出现在视频开始到当前帧从始至终都未能检测到“未佩戴”情况,提示事件,即为黄色事件,二级优先,提示事件是指目标从出现在视频中到当前帧的口罩检测状态由“佩戴”转向“未佩戴”,而警告事件,即为红色事件,一级优先,警告事件是指从目标出现在视频中到当前帧从始至终都只有两个状态“未佩戴”或者“未知”,未检测出“佩戴”状态;

(5) 输出模块根据设置中的各项参数决定输出的细节,包括是否自动插值处理、是否平滑处理、输出检测框的样式;若是用户选择自动插值处理,则在得到查询输出的结果后再一次甄别输出目标,对于不连续的轨迹,标记产生不连续的所有间隔,并采用均值法在间隔的前后帧的检测框属性取均值,属性包括:检测框的坐标、宽高,计算得到的新的检测框作为插值数据插入到间隔中,若是用户选择平滑处理,则同样对于查询结果再一次甄别输出轨迹,对于每一条轨迹,画出其中每个检测的中心点的位置,用折线相连,并将中心点在视频中的位置分别为x、y轴呈现在坐标轴中,利用线性回归的算法对于每一条轨迹进行拟合,并将拟合误差最小的拟合线条作为新的平滑轨迹,原轨迹中的每个检测框根据新平滑轨迹中对应时间帧的检测坐标调整参数,最后根据用户设置的输出检测框的参数,包括颜色、检测框宽度和是否显示行人ID,在图像中画出行人的检测框,并随着视频一帧一帧显示;检测框

结果实时显示后,根据口罩监测模块所得到的口罩事件的优先级和系统设置模块对于不同事件的对象进行分类并实时对目标进行警告或提示。

## 一种基于概率神经网络的口罩佩戴情况的监测系统及方法

### 技术领域

#### 一种基于YOLOv5的口罩佩戴检测方法

[0001] 本发明涉及一种基于概率神经网络的口罩佩戴情况的监测系统及方法,属于计算机视觉视频监控侦察信息化建设方向。

### 背景技术

[0002] 近年来,计算机视觉技术发展迅猛,随着视频跟踪技术的越发成熟,其在实际应用方面,如智能无人机交互、视频监控监察、异常事件监测等任务中发挥了显著的作用,成为监控侦察方面重要的基础和核心技术之一。基于视频图像的多目标跟踪技术是目前计算机视觉领域一个非常具有挑战性和吸引力的基础研究方向,虽然该方向在众多领域有着广泛的应用,甚至是很多研究领域的技术前提,如行为分析、智能交通等等,然而受各方面限制,业界却缺乏相对集中的处理视频多目标跟踪问题的系统研究和智能化平台。

[0003] 视频多目标跟踪,任务一般可以描述为利用计算机对图像序列中感兴趣的目标(人)进行检测、提取和跟踪,将表示同一个目标的检测用全局唯一的ID表示,并从始至终将每个人的轨迹完整地呈现出来,从而使人们更好地理解 and 描述行为的目的。行人多目标跟踪是视频多目标跟踪的一个重要方向,由于行人是典型的非刚性物体,在跟踪过程中会伴随动作、形变等情况的发生,目标特征的表征面临挑战,此外,跟踪场景的拥挤程度、光线强弱、行人的数量变化、多目标的交互遮挡等各种特殊情况的发生都会给数据关联带来困难,并且跟踪目标的多变性使得用一个跟踪器及一套系统解决所有的多目标跟踪问题变得困难。

[0004] 多目标跟踪技术在应用方面的研究的目的在于其是对目标进行动作识别、行为分析等应用的实现基础,能够广泛应用于视频监控、无人机交互、医学图像处理等多个领域。

[0005] 在科学研究方面由于目标的多样性和多变性,场景的多样性和多变性以及目标间交互的复杂性使得多目标跟踪的研究成为难点,大多数跟踪方法在跨摄像机的多目标检测和跟踪上表现出低鲁棒性和高轨迹互换的指标,且往往效率低下,并且难以成系统地查询兴趣目标的跟踪结果,并对特定目标人群(未佩戴口罩)进行实时信息反馈与交互,大多多目标跟踪算法依托于检测,而检测器的好坏通常会决定多目标跟踪效果的上限,在这个有限的检测器技术的依托上,如何突破上线是多目标跟踪技术研究的突破口。多目标跟踪技术的发展会同时为视觉研究其他各个领域带来基础上实质性的突破,并会为视频行为分析、智慧城市、智慧交通等科研工作提供支持。

[0006] 本发明着眼于未来人工智能机器视觉领域的发展智能视频监控的发展需求,以提升视频多目标对象内容理解能力为总体目标,瞄准视频多目标跟踪技术这一问题,结合深度学习框架,根据多目标跟踪关联过程相关约束,设计合适的算法提升跟踪结果的精确性,并集成多目标跟踪与查询平台,并为视觉领域进一步的视频分析提供技术支持并做好基础。



## 发明内容

## 先说明发明目的

[0007] **发明技术解决问题**: 为了克服跨摄像机的多目标检测和跟踪难、效率低, 难以成系统地查询兴趣目标的跟踪结果, 并对特定目标人群(未佩戴口罩)进行实时信息反馈与交互等问题, 提供一种基于概率神经网络的口罩佩戴情况的监测系统及方法, 综合考虑了实际工程中试试结果显示地需要和科研设计中可设计性的需求, 既为用户提供了具体的多目标跟踪黑盒结果, 同时也为科研人员提供了全透明算法设计框架, 参数可调, 架构可设计。该系统是帮助用户快速高效获取多摄像机多目标跟踪结果查询、设计多目标跟踪框架的系统, 结果可靠性高, 系统实用性强、设计良好、运行稳定、易用性好、可修改、可保存, 并为视频监控、行为分析等领域提供了技术支持。

[0008] **本发明技术方案**: **接着说明发明技术方案**

[0009] 一种基于概率神经网络的口罩佩戴情况的监测系统, 包括: 输入管理模块、目标检测与分割模块、多目标跟踪关联模块、口罩监测模块、输出模块和系统设置模块;

[0010] 输入管理模块, 此部分封装入前端的监控摄像机中, 为监控摄像机拍摄的输入视频上传和处理提供接口, 也包含用户所使用的操作界面, 是目标检测与分割模块、多目标跟踪关联模块、口罩监测模块、输出模块、系统设置模块的入口;

[0011] 目标检测与分割模块, 利用PCB卷积神经网络以及区域生成网络对于输入视频的每一帧的每一个行人进行检测和识别; 首先利用选择搜索算法确定视频每一帧中的行人检测的候选框, 所述候选框为能将单个行人完全包住的矩形框, 然后利用卷积神经网络提取候选框的特征, 利用区域生成网络对于候选框进行筛选确定最终的行人检测框, 所述PCB卷积神经网络以及区域生成网络是否需要使用预训练的网络由系统设置的参数决定, 所述最终的行人检测框即为能将单个行人完全包住的最小矩形框, 由目标检测模块检测和识别处理后得到行人检测特征, 行人检测特征包括检测后得到的每个检测框以及每个行人的特征, 并利用分类器对行人检测框进行分类, 根据分类结果对检测框进行对齐调整, 并利用光场的深度计算辅助分类回归检测算法对输出的检测数据进行分割, 实现行人的头部部分的分割, 得到分割结果, 所述检测数据和分割结果将作为多目标跟踪关联模块、口罩监测模块和输出模块的数据基础;

[0012] 多目标跟踪关联模块, 负责对于每个视频得到的目标检测与分割模块的检测结果在该视频内进行关联; 在多目标跟踪的约束条件下, 利用基于多假设概率神经网络的多目标关联算法, 完成对所有检测结果的关联任务; 所述关联任务即如果为同一个行人则匹配, 否则不匹配; 所述多目标跟踪的约束条件包括: 空间约束、时间约束、可允许的漏检数目约束和相似度约束; 所述空间约束是指产生关联的两两检测框在视频中的位置距离不可超出一定阈值; 所述时间约束是指产生关联的两两检测框在时间上的交集必须为空集; 所述可允许的漏检数目约束是指关联后导致时间不连续轨迹的漏检检测的数目必须在设定阈值之内; 所述相似度约束是指判断行人的两个检测是否应该产生关联关系的度量依据; 所述目标关联模块包括多假设概率神经网络和剪枝算法; 所述多假设概率神经网络提供基于多假设概率神经网络的多目标关联算法, 将检测序列作为轨迹信息作为输入, 基于多假设概率神经网络的多目标关联算法融合基于长短期记忆网络的概率神经网络来学习轨迹的时序特征从而拟合输入轨迹的未来变化趋势, 从趋势特征上对短轨迹进行匹配得到相似度分数, 根据相似度来对短轨迹进行关联; 多假设树的节点生成取决于短轨迹间的相似度分数,



多假设概率神经网络有着突出的处理时序特征的能力,能充分利用行人轨迹的历史特征信息和交互信息,使得匹配分数能反映出轨迹之间的实际匹配程度,从而更好更精确地生成多假设树;所述剪枝算法为避免由于多假设树层数增长带来的计算时间指数增长和空间复杂度的指数增长,多假设概率神经网络的缺点是随着时间的增长其计算量会成指数增加,导致效率降低,由此剪枝算法尤为重要;多目标跟踪关联模块充分考虑由于拥挤场景和长期遮挡导致的行人轨迹断连和行人轨迹交换的问题,利用基于长短期记忆网络的概率神经网络对多假设概率神经网络生成的多假设树进行剪枝,该深度模型将多假设树的每个分支作为输入,输出结果为每个分支为同一个行人的置信度分数,此时多假设树的每个节点为行人检测的特征,即外观特征,最后将置信度最高的分支保留,其余分支进行剪枝;多目标跟踪关联模块中所参考的历史特征信息为行人短轨迹的轨迹外观特征和运动特征,所述交互信息即产生交互的相行人轨迹之间的关系特征;最后得到结果全局最优的行人关联结果,即匹配的行人轨迹;

[0013] 口罩监测模块,负责利用目标检测与分割模块的检测数据和分割结果使用姿态估计网络,判断行人的朝向,即背面、正面和侧面,对行人检测和分割结果进行头部面部定位和检测,对正面和侧面的行人面部,使用口罩检测算法进行口罩检测,得到口罩检测结果后将对面部口罩的检测结果进行实时在线关联跟踪,得到实时的面部口罩轨迹的跟踪状态的变化,并定义三个口罩事件:正常、提示和警告,将得到的口罩轨迹的跟踪状态的变化按这三个口罩事件进行分类,得到分类结果;

[0014] 输出模块,将多目标跟踪关联模块得到的结果全局最优的行人关联结果和口罩监测模块得到的口罩轨迹跟踪状态的口罩事件分类结果进行处理后将行人轨迹的跟踪结果可视化并实时对于跟踪过程中的口罩事件作出响应;输出模块将多目标跟踪关联模块输出的行人关联结果进行分类,同一个行人的所有检测框设置为同一颜色,不同行人其检测框颜色设为不同,所述颜色的分配情况由系统设置模块的参数控制;然后对于行人关联结果所生成的轨迹进行处理,所生成的轨迹即同一行人的连续帧形成的检测序列;所述处理包括对于轨迹的插值处理和平滑处理;所述插值处理是指遍历所有的轨迹,选出轨迹内存在不连续帧的轨迹,对于这些轨迹的不连续处插入一定帧数的检测信息,补充并缓解由于不连续带来的信息丢失的问题;所述平滑处理即再一次对于行人轨迹内的所有检测结果在感官层面做一些微调,例如坐标、宽、高等信息,使得用户有一个良好的视频感官体验;对于口罩检测模块的口罩事件分类结果根据优先级分别作出响应,所述响应的方式以及优先级均由系统设置模块的参数所控制,处理完毕后将结果进行保存并可视化;

[0015] 系统设置模块,用于配置输入管理模块、目标检测与分割模块、多目标跟踪关联模块、以及输出模块的参数配置;对于输入管理模块,系统设置模块设置输入视频的最大限制参数,通道数限制参数以及输入目标视频的地址;对于目标检测与分割模块,系统设置模块设置目标检测生成的细节,包括检测框的大小、宽高以及是否需要直接使用预训练的模型进行计算,若需要则设置模型的根目录;用户根据实际需要设置模型输出根目录,以便在PCB卷积神经网络以及区域生成网络模型在使用结束后的复用;对于多目标跟踪关联模块在运行代码时,首先读取系统设置模块中的参数数据,完成PCB卷积神经网络以及区域生成网络算法的参数配置,由用户实现对PCB卷积神经网络以及区域生成网络算法的透明性操作,提升输出结果的可调节性,实现操作决策的可控;对于口罩监测模块,用户自行设置三

个事件的响应级别,默认的优先级别为:“警告”大于“提示”,“提示”大于“正常”;系统设置模块提供这三个事件的响应方式设置,用户根据需求进行个性化设置,内容包括提示框的显示内容、颜色、辅助语音和提示音。

[0016] 所述多目标跟踪关联模块中多目标关联算法具体实现如下:

[0017] (1) 首先将目标检测和分割模块得到的行人检测特征以用户设置的窗口数为单位进行分组,设第一个窗口为当前窗口 $W_c$ ;

[0018] (2) 以当前窗口 $W_c$ 所有检测为横纵轴构造相似度矩阵,行人检测的相似度计算方式同公式(1),即为两两行人检测间的运动特征即坐标、宽高的欧式距离与外观特征即检测特征距离的加权和,其中 $\text{Dist}_{\text{mot}}^{i,j}$ 表示行人检测 $i$ 和 $j$ 之间的运动特征的距离, $\text{Dist}_{\text{app}}^{i,j}$ 表示行人检测 $i$ 和 $j$ 之间的外观特征的距离,公式(2)中 $x$ 、 $y$ 、 $w$ 、 $h$ 分别指代行人检测的坐标和宽高,公式(3)中的 $a_k$ 表示行人检测 $i$ 的外观特征的第 $k$ 个元素, $b_k$ 表示行人检测 $j$ 的外观特征的第 $k$ 个元素。将得到的相似度分数 $S^{i,j}$ 用于构造相似度矩阵,根据相似度矩阵,计算得到最佳匹配的短轨迹集合;所述短轨迹是指生成的属于同一个行人的检测序列,即轨迹长度不超过当前窗口长度的轨迹集合;

$$[0019] \quad S^{i,j} = w_{\text{mot}} \text{Dist}_{\text{mot}}^{i,j} + w_{\text{app}} \text{Dist}_{\text{app}}^{i,j} \quad (1)$$

$$[0020] \quad (\text{Dist}_{\text{mot}}^{i,j})^2 = (x^i - x^j)^2 + (y^i - y^j)^2 + (w^i - w^j)^2 + (h^i - h^j)^2 \quad (2)$$

$$[0021] \quad (\text{Dist}_{\text{app}}^{i,j})^2 = \sum_k (a_k - b_k)^2 \quad (3)$$

[0022] (3) 判断当前窗口 $W_c$ 是否是最后一个窗口,若是则进入步骤(4),或不是则将 $W_{c+1}$ 赋值为 $W_c$ ,进入步骤(2);

[0023] (4) 设第一个窗口为当前窗口 $W_c$ ;

[0024] (5) 取当前窗口 $W_c$ 中一条短轨迹 $k$ 作为当前节点,并建立父节点,判断当前窗口 $W_c$ 是否为最后一个窗口,若不是最后一个窗口,则进入步骤(6),若是最后一个窗口,则进入步骤(12);

[0025] (6) 计算短轨迹 $k$ 与下一个窗口 $W_{c+1}$ 中所有轨迹的相似度,轨迹之间的相似度计算通过多假设概率神经网络计算得到,多假设概率神经网络将两条短轨迹的每一帧的检测特征作为输入,在内部进行高维拼接和筛选,输出两两短轨迹的相似度分数,得到相似度高于用户设置阈值的短轨迹集合 $m$ ;

[0026] (7) 判断子节点数量是否超出用户设置的阈值,若是,则进入步骤(5),反之,则将短轨迹集合 $m$ 中所有短轨迹设置为短轨迹 $k$ 的子节点,进入步骤(8);

[0027] (8) 判断是否遍历完当前窗口 $W_c$ 中所有短轨迹,若是则进入步骤(9),反之,则进入步骤(5);

[0028] (9) 计算所得多假设树的层数是否在用户设置阈值范围内,若是则进入步骤(3),反之则进入步骤(10);

[0029] (10) 将建立的多假设树的每个分支作为输入,输入到长短期记忆网络的概率神经网络,计算得到每个分支为同一个行人的置信度分数;

[0030] (11) 根据置信度分数对多假设树进行剪枝,进入步骤(3);

[0031] (12) 选则全局分数最高的分支作为匹配的行人轨迹。

[0032] 本发明一种基于概率神经网络的口罩佩戴情况的监测方法,实现如下:

[0033] (1) 用户首先通过登陆步骤进入到输入管理模块,在登陆界面,通过输入框记录用

户输入信息,界面提交后,数据库通过比对输入信息和数据库中缓存的用户信息并匹配用户ID,分为匹配成功和匹配失败两种状态:若匹配成功,数据库通过ID查找该用户并记录相关输入及查询缓存信息并将用户的相关信息返回前端;若匹配失败,则弹出消息框,显示用户不存在;登陆成功后,根据用户定位到视频所在的绝对路径,上传视频;

[0034] (2) 目标检测与分割模块步骤:首先对于输入视频的每一帧,在其中利用选择性搜索的算法确定多个检测框的候选框,然后将图像依次输入到一个卷积神经网络,得到候选框特征,包括行人的坐标、宽高以及初步的卷积特征,然后利用一个区域生成网络对于每个候选框进行筛选,得到最终的检测框。使用分类器将所有的检测框按照是否对齐进行分类,完成分类后,对于未完全对齐的检测框利用回归器进一步的调整检测框位置。利用光场的深度计算辅助分类回归检测算法基于卷积特征对检测数据进行分割,主要利用光场的深度信息并基于行人检测的卷积特征对于图像区域进行头部分割;检测与分割结束后,将得到的检测分割结果保存到系统设置模块设置的根目录位置;

[0035] (3) 获得所有检测的检测分割结果后,将进行多目标跟踪,采用多假设概率神经网络的多目标跟踪关联算法,该多目标跟踪关联算是基于多假设跟踪框架,所述多假设跟踪框架分成多假设树的生成、多假设树的剪枝和最佳路径的选择三个阶段;多假设树的生成阶段,系统将多假设树的生成和深度学习相结合,利用基于长短时记忆神经网络的网络结构在处理时序特征上的优势,辅助生成多假设树,具体的生成实现步骤是:首先将视频以系统设置模块设置的帧数为窗口长度单元分割成多个不重叠的窗口,在每个窗口内的所有检测框之间进行匹配,生成最初的短轨迹,这些短轨迹被视为原子节点,即不可分,多目标跟踪的关联任务即在这些窗口之间,将两两相邻窗口之间的原子节点相互匹配,为每一个潜在的节点即目标生成一棵假设树,生成这些假设树需要同时满足的条件是:①子节点与父节点所在的窗口序号差的绝对值不可跨越系统设置模块给出的窗口阈值数;②子节点与父节点的相似度,即由基于长短时记忆神经网络的网络模型所评估的分数,不可低于系统设置模块给出的阈值;③一个父节点的子节点的数目不可超过系统设置模块给出的阈值,将深度学习融合到剪枝算法中,利用深度模型记忆目标的外观特征,为每只分支是否为同一个目标的自信度进行打分,使行人轨迹在经历遮挡、拥挤的信息丢失的场景后恢复时能够根据外观的记忆与已经有的轨迹进行匹配;在跟踪完所有的检测目标后,根据该框架算法为每棵树每个分支的打分,取最高分的分支作为最佳的匹配结果,最后将结果输出到系统设置模块所保存的根目录;

[0036] (4) 口罩监测模块为三个步骤:①行人姿态估计,②头部面部定位和检测,③口罩检测,④监测口罩事件,获得所有行人的检测和分割结果后,利用目标姿态估计网络识别行人的状态,包括正面、侧面和背面,从而判断行人的朝向;再利用行人的分割结果对行人进行头部和面部定位,并对头部部分的检测框进行修正后提取特征,该特征利用目标检测与分割模块中的特征提取的卷积神经网络提取出的目标的头部深度外观特征;利用这些外观特征利用口罩检测网络对于一定朝向,包括正面和侧面的行人进行口罩检测,实时判断出行人是否佩戴口罩,并将检测结果赋予每个行人对象,与行人的ID相关联,检测结果包括:佩戴、未佩戴、未知,其中佩戴表示系统能通过网络识别出口罩的特征,未佩戴是指系统对于正面后侧面朝向的行人进行识别后未能检测出口罩,未知则是对于背面朝向的行人,系统本身无法识别行人脸部从而无法识别并判断口罩的佩戴情况;口罩监测的最后一步是监

测口罩事件,利用口罩检测步骤所得到的检测结果,将每个ID的行人进行实时检测,并根据口罩检测结果的实时变化对行人的状态进行评估并反应口罩事件,口罩事件具体划分为三个:①正常,②提示,③警告;这三个事件也分别对应了三个颜色,即绿、黄、红,以及三个优先级,用户根据需求设置优先级,默认设置为:事件被判为正常事件,即为绿色事件,为三级优先,正常事件是指从该目标出现在视频开始到当前帧从始至终都未能检测到“未佩戴”情况,提示事件,即为黄色事件,二级优先,提示事件是指目标从出现在视频中到当前帧的口罩检测状态由“佩戴”转向“未佩戴”,而警告事件,即为红色事件,一级优先,警告事件是指从目标出现在视频中到当前帧从始至终都只有两个状态“未佩戴”或者“未知”,未检测出“佩戴”状态;

[0037] (5) 输出模块根据设置中的各项参数决定输出的细节,包括是否自动插值处理、是否平滑处理、输出检测框的样式;若是用户选择自动插值处理,则在得到查询输出的结果后再一次甄别输出目标,对于不连续的轨迹,标记产生不连续的所有间隔,并采用均值法在间隔的前后帧的检测框属性取均值,属性包括:检测框的坐标、宽高,计算得到的新的检测框作为插值数据插入到间隔中,若是用户选择平滑处理,则同样对于查询结果再一次甄别输出轨迹,对于每一条轨迹,画出其中每个检测的中心点的位置,用折线相连,并将中心点在视频中的位置分别为x、y轴呈现在坐标轴中,利用线性回归的算法对于每一条轨迹进行拟合,并将拟合误差最小的拟合线条作为新的平滑轨迹,原轨迹中的每个检测框根据新平滑轨迹中对应时间帧的检测坐标调整参数,最后根据用户设置的输出检测框的参数,包括颜色、检测框宽度和是否显示行人ID,在图像中画出行人的检测框,并随着视频一帧一帧显示;检测框结果实时显示后,根据口罩监测模块所得到的口罩事件的优先级和系统设置模块对于不同事件的对象进行分类并实时对目标进行警告或提示。

[0038] 本发明与现有技术相比的优点在于:

**最后说明发明的有益支出**

[0039] (1) 本发明设计了包括输入管理模块、目标检测与分割模块、多目标跟踪关联模块、口罩监测模块、输出模块、系统设置模块的一种基于多假设概率神经网络的多视频目标跟踪与查询系统。能根据相机前端上传的视频及设定的参数高效生成跟踪方案,同时实时监测监控人群佩戴口罩的情况,并对未佩戴口罩的行人目标进行警告或提示,并可自行在多目标跟踪框架中训练、修改、调整方案。提高了多目标跟踪设计的效率、节约了时间和人力成本、提供了集成的实时跟踪和交互系统;

[0040] (2) 本发明可实现不同场景、不同摄像机的多视频跨摄像机的多目标跟踪,跟踪结果可视化,空间利用率高、跟踪结果可读取可设计、多目标跟踪方案参数可读取、可修改、可调整、可保存,以供实际工程及科学研究使用。

[0041] (3) 本发明的多目标跟踪目标关联算法采用了多假设概率神经网络于传统多假设跟踪框架结合的算法,设定了相应的超参数,短轨迹窗口大小、短轨迹生成约束、轨迹间隔阈值、相似度阈值、剪枝置信度阈值的大小都可以根据用户需要进行调节,可以得到不同的适合用户需要的多目标跟踪结果。实现了跟踪决策的自动化,系统通过参数自动计算,可以对当前决策起到辅助作用,能够很好地完成计算机与人协同工作的任务。

[0042] (4) 本发明实现了不同算法的目标检测用于多目标的关联,并对各种算法的检测结果分别评估,可根据用户的需求选择最佳的检测结果,分为三种:均值、中位数、随机。若是选择均值那么在多个检测器相应的检测结果取均值,中位数则是选择在多个检测器结果



中检测框参数最为居中的检测结果,随机则是随机选择一个检测器。不同的策略适合不同的应用场景,极大地适应用户的需求。

[0043] (5) 在多目标跟踪关联模块中,系统提供了各种辅助操作以保证多目标跟踪关联的精确性,如动静摄像机参数分离操作、多视频参数高度聚合、边界超出设置、外观特征与运动特征加权比例设置等,具备易用性及精确性。

[0044] (6) 本发明提供了方便且完备的算法方案管理策略,用户在训练和调试多目标跟踪框架的过程中,可对得到的算法方案进行保存,在之后随时可以读取方案并进行方案的修改和另存,以供实际工程中的操作、训练和改进。

[0045] (7) 本发明设计了行人口罩佩戴实时监测环节,帮助从实时视频数据中监测并跟踪描绘出未佩戴目标的轨迹,并实时根据事件优先级别对相应事件对当事行人做出反馈,此外用户可以根据特征设置需要跟踪的具有这些特征的人群(如佩戴口罩、未佩戴口罩等)。

[0046] (8) 系统设计了可维护接口和可扩展接口。使得系统可以替换某个模块,而不影响其他模块的使用。模块都可以扩展,添加新的功能也不影响其他模块的工作。

## 附图说明

[0047] 图1为本发明的总流程图;

[0048] 图2为本发明的目标检测与分割模块流程图;

[0049] 图3为本发明的多目标跟踪算法流程图;

[0050] 图4为本发明的口罩监测算法流程图;

[0051] 图5为本发明的关键系统架构图;

[0052] 图6为本发明的多假设概率神经网络结构图。

## 具体实施方式

[0053] 下面结合附图及实施例对本发明进行详细说明。

[0054] 如图5所示,本发明包括:输入管理模块,负责处理输入视频的读取和预处理。目标检测与分割模块,负责对于视频中的每个目标即行人进行检测和识别,并分割出行人口罩部分进行检测。多目标跟踪关联模块,负责对于每个视频中的检测结果在该视频内进行多对多关联。口罩监测模块,完成初步轨迹关联后,对置信度较低和离散轨迹,使用戴口罩的检测进行辅助关联,提高轨迹完整性,并发现其中的戴口罩和摘下口罩两个事件。输出模块,平台设立在统一的数据中心节点,支持多路跨摄像机数据输出,将分别对各个场景进行多目标跟踪,统一汇聚输出计算结果,包括所有未佩戴口罩行人的移动轨迹,摘口罩和戴口罩事件,以及行人的多维度特征,即外观和运动特征。系统设置模块,用于配置多个模块中所用的网络模型的训练。本发明解决了跨摄像机的实时多目标检测和跟踪难、效率低等问题,设计良好,运行稳定,并为视频监控、行为分析等领域提供了技术支持。

[0055] 如图1所示,本发明的总流程图如下:

[0056] (1) 首先用户进入本发明的平台,输入其账号和密码;

[0057] (2) 系统比对与数据库信息是否一致,若不一致则(1),若一致则进行下一步(3);

[0058] (3) 数据库返回用户信息于平台并展示;

- [0059] (4) 系统开放视频的输入接口；
- [0060] (5) 用户选则上传视频；
- [0061] (6) 将检测结果返回用户；
- [0062] (7) 结束。
- [0063] 如图2所示,本发明的目标检测与分割模块,实现流程如下:
- [0064] (1) 系统读入视频帧；
- [0065] (2) 使用搜索算法确定行人候选框；
- [0066] (3) 使用利用卷积神经网络提取行人候选框特征；
- [0067] (4) 利用区域生成网络对候选框进行筛选,得到最终行人检测框；
- [0068] (5) 使用回归器调整并纠正检测框的位置；
- [0069] (6) 结合并利用光场计算完成深度估计分割行人检测；
- [0070] (7) 若未完成所有图像的计算,则返回(3)；
- [0071] (8) 保存得到的行人检测数据和分割结果；
- [0072] 如图3所示,本发明的多目标跟踪算法流程图：
- [0073] (1) 首先将所有视频以最小窗口长度为单位对视频进行分割,得到窗口序列  $W_1W_2...W_n$ ；
- [0074] (2) 对于每一个窗口  $W_i$ ,在窗口内以所有检测作为横纵轴构造相似度矩阵,计算得到最佳匹配的短轨迹集合；
- [0075] (3) 设当前窗口为  $W_i$ ,以其中一条短轨迹  $k$  为根结点建立父节点；
- [0076] (4) 判断下一个窗口  $W_{i+1}$  中是否存在短轨迹集合  $m$  使得  $m$  中所有短轨迹与短轨迹  $k$  的相似度小于阈值,若有则(5),反之则(3)；所述短轨迹间的相似度通过多假设概率神经网络获得,将  $m$  中的短轨迹分别与短轨迹  $k$  拼接作为多假设概率神经网络的输入,输出则为  $m$  中的短轨迹与短轨迹  $k$  的相似度；
- [0077] (5) 判断短轨迹  $k$  的子节点数量是否小于阈值,是则(6),反之则(3)；
- [0078] (6) 将  $m$  中所有短轨迹作为短轨迹  $k$  的子节点；
- [0079] (7) 判断是否所有轨迹都已经关联,若是则(8),反之则(3)；
- [0080] (8) 将生成的假设树的每条分支组成的轨迹序列中的所有检测的外观与运动信息输入到概率神经网络,计算得到每条分支的置信度分数；
- [0081] (9) 利用置信度分数来对于低于系统指定阈值的分支进行剪枝；
- [0082] (10) 判断是否计算了所有分支,若是则(11),反之则(8)；
- [0083] (11) 根据置信度分数选择假设树中的最佳匹配生成行人跟踪轨迹。
- [0084] 如图4所示,本发明的口罩监测算法,流程如下：
- [0085] (1) 输入目标检测与分割模块的检测数据和分割结果；
- [0086] (2) 利用姿态估计网络判断行人的朝向；
- [0087] (3) 根据检测数据和分割结果定位头部区域；
- [0088] (4) 利用深度神经网络提取头部特征；
- [0089] (5) 使用分类器计算行人佩是否戴口罩的置信度；
- [0090] (6) 标志轨迹中检测对应口罩佩戴状态；
- [0091] (7) 判断是否出现未佩戴口罩事件,若是,则(8),反之则(9)；



- [0092] (8) 标志警告事件；
- [0093] (9) 判断是否出现未知状态，若是，则 (11)，反之则 (10)；
- [0094] (10) 标志提示事件；
- [0095] (11) 标志正常事件；
- [0096] (12) 结束；
- [0097] 如图6所示，本发明的计算行人短轨迹间相似度的多假设概率神经网络结构如下：
- [0098] (1) 将两条长度为 $n$ 的短轨迹序列的行人检测输入PCB神经网络分别得到两条 $1536*n$ 的特征序列；
- [0099] (2) 将这两条特征序列在时间维度上进行拼接，得到 $1536*2n$ 的特征序列；
- [0100] (3) 将该特征序列输入到拥有256个神经元个数的LSTM网络，提取最后一个序列的输出作为下一层的输入，即得到 $256*1$ 的特征；
- [0101] (4) 通过 $128*256$ 的全连阶层FC，得到 $128*1$ 的特征；
- [0102] (5) 继续通过 $2*128$ 的全连接层FC，得到 $2*1$ 的特征；
- [0103] (6) 通过SOFTMAX层对输出结果进行归一化得到概率值；
- [0104] (7) 定义第一个值为两条行人短轨迹为同一个人的概率值，即相似度，输出相似度。

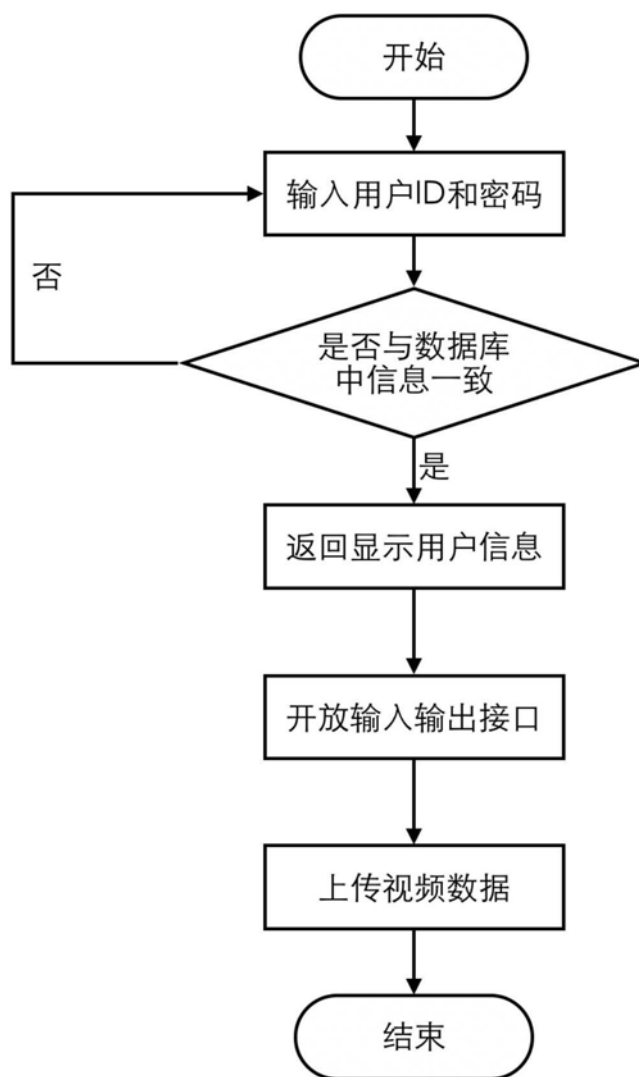


图1

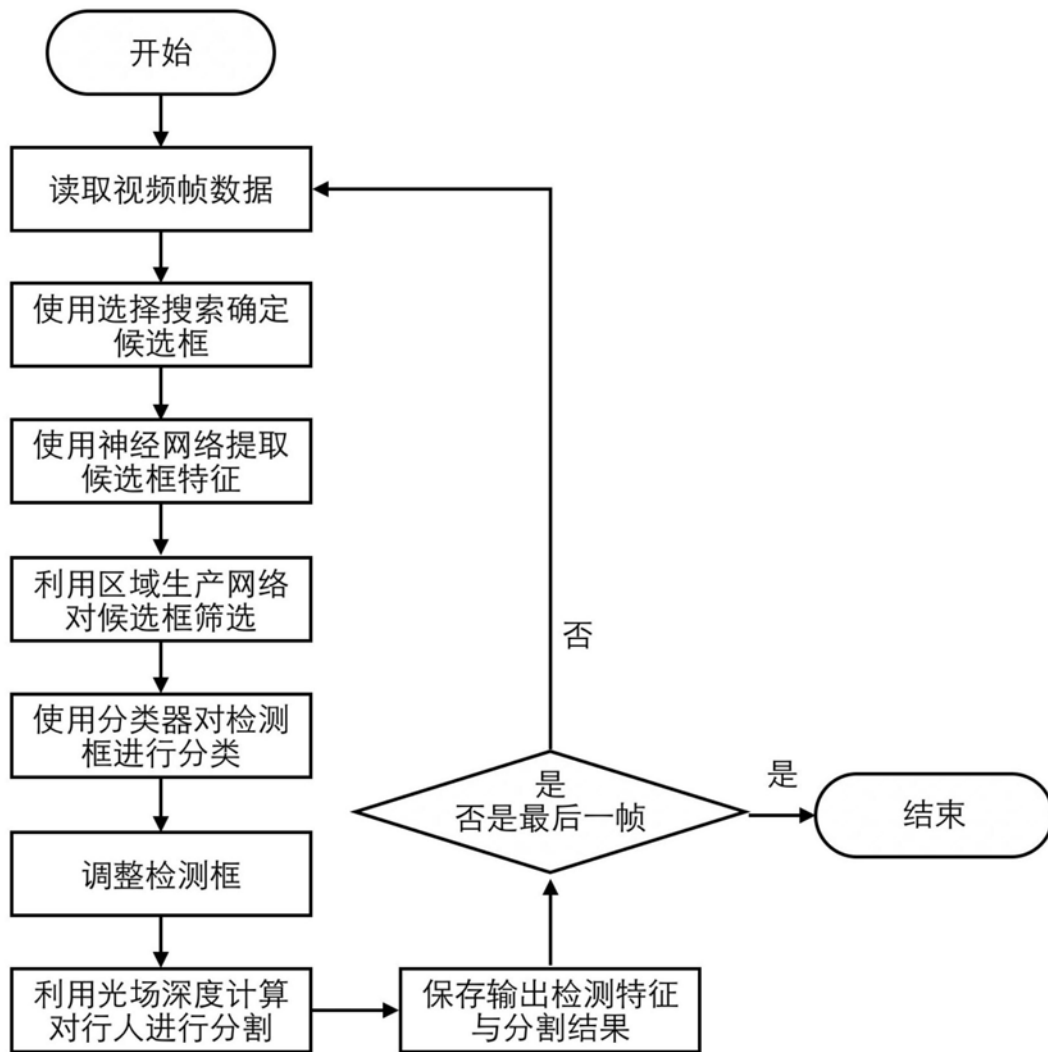


图2

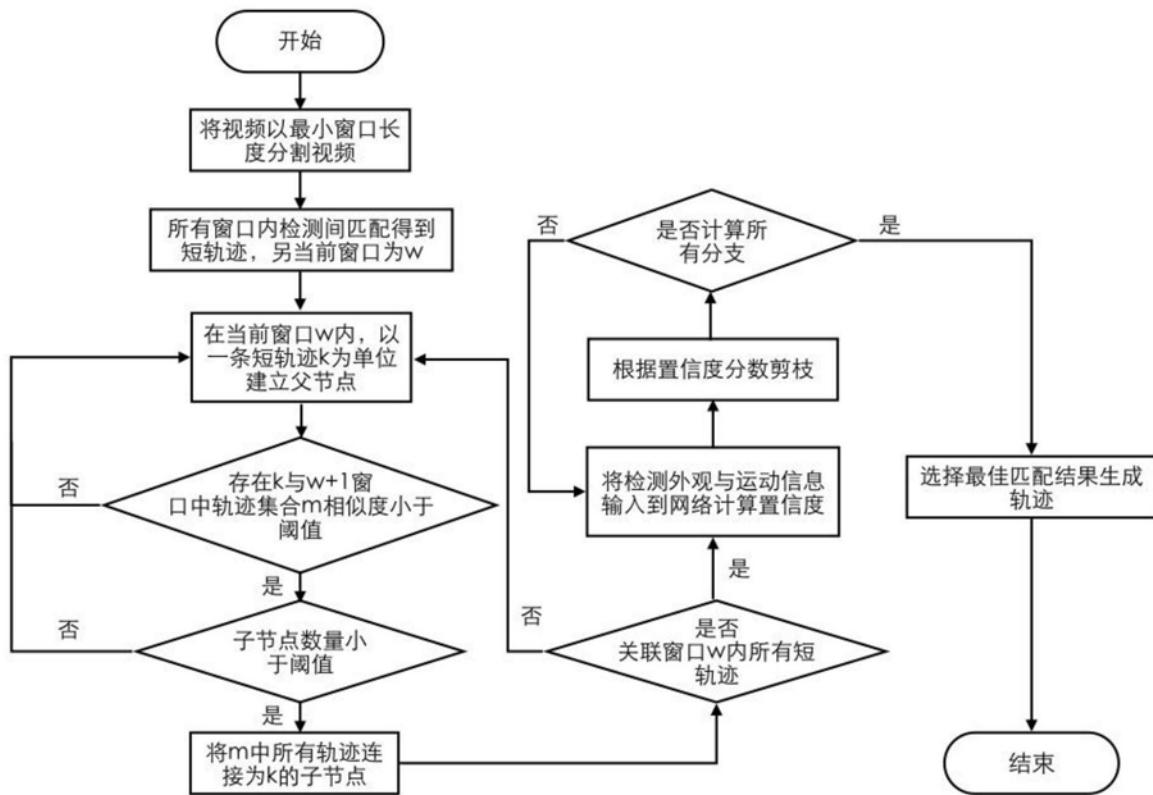


图3

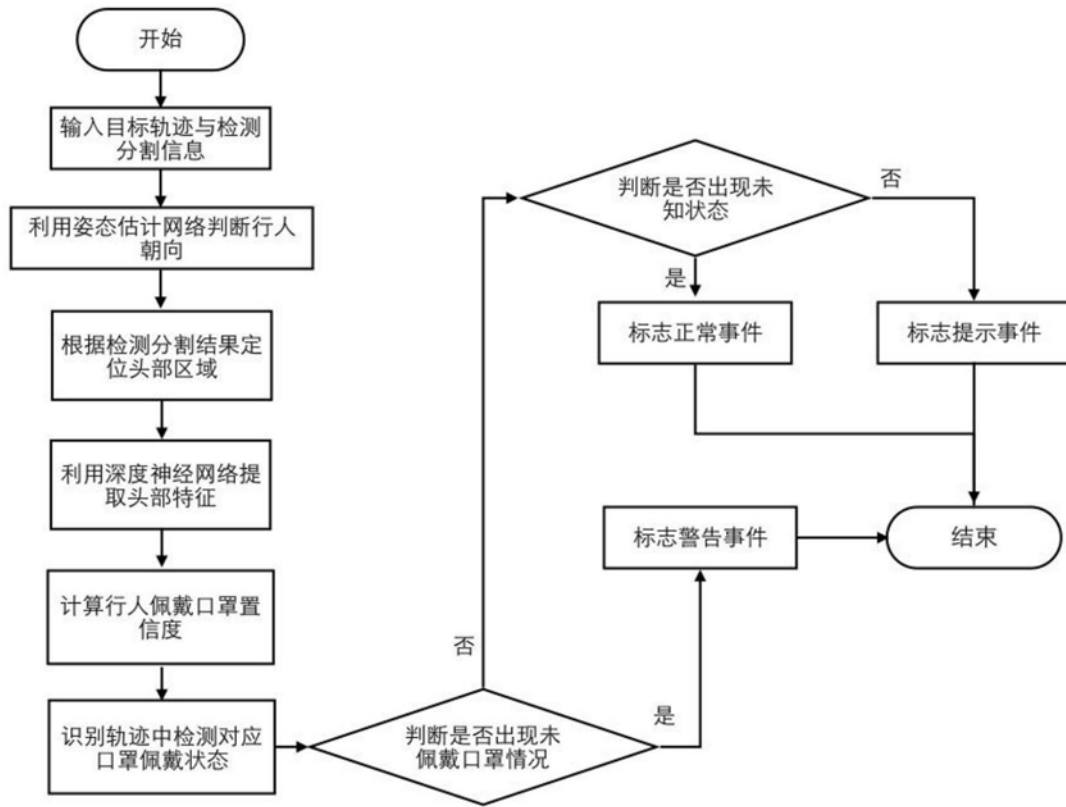


图4

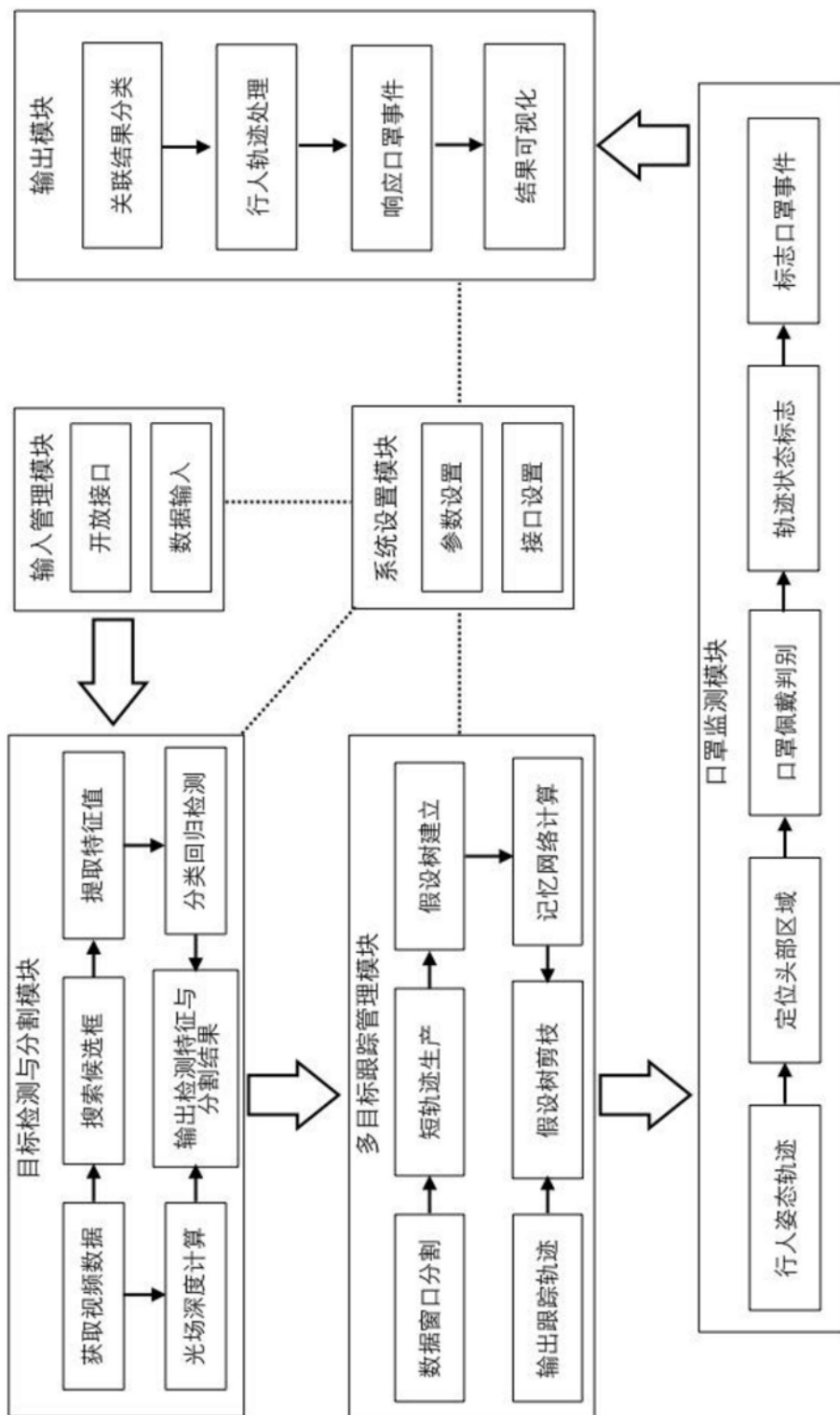


图5



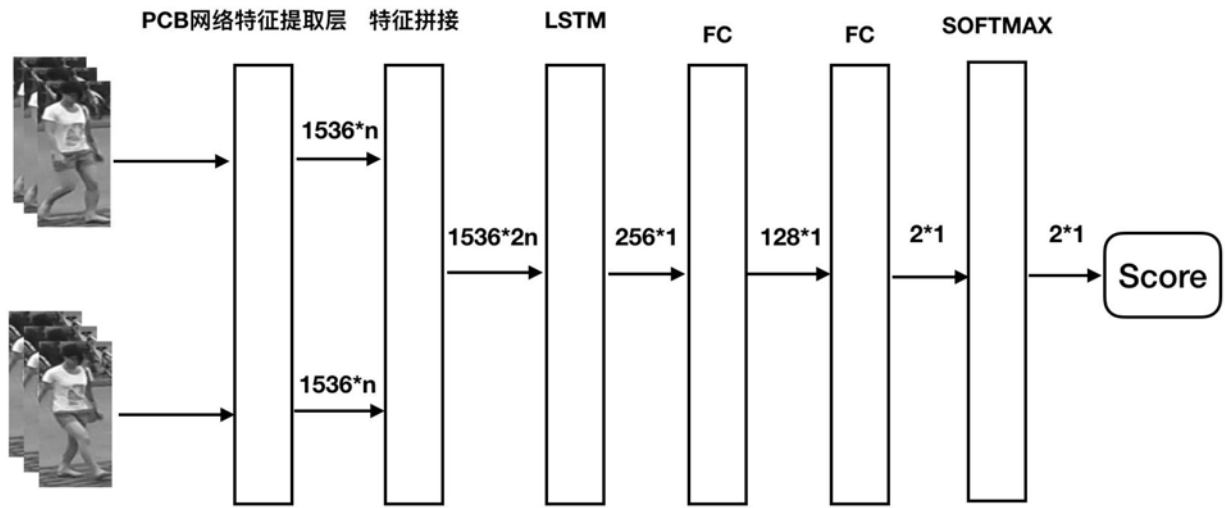


图6