



(12) 发明专利

(10) 授权公告号 CN 113092807 B

(45) 授权公告日 2024.05.14

(21) 申请号 202110429725.2

(22) 申请日 2021.04.21

(65) 同一申请的已公布的文献号

申请公布号 CN 113092807 A

(43) 申请公布日 2021.07.09

(73) 专利权人 上海浦江桥隧运营管理有限公司

地址 200120 上海市浦东新区中国(上海)

自由贸易试验区临港新片区环湖西二
路888号B楼388室

专利权人 上海大学

上海城建城市运营(集团)有限公
司

上海市政养护管理有限公司

(72) 发明人 封玲 王晓宇 王一明 胡珉

高新闻 薛子欣 赵鹏 陈莹锋

(74) 专利代理机构 上海京沪专利代理事务所

(普通合伙) 31235

专利代理师 杨金淑

(51) Int.Cl.

G01P 3/68 (2006.01)

G06T 7/207 (2017.01)

G06T 7/70 (2017.01)

G06N 3/0464 (2023.01)

G06N 3/08 (2023.01)

(56) 对比文件

AU 2020100371 A4, 2020.04.16

CN 110472496 A, 2019.11.19

CN 110738690 A, 2020.01.31

CN 111914664 A, 2020.11.10

CN 111932580 A, 2020.11.13

CN 112001948 A, 2020.11.27

CN 112037249 A, 2020.12.04

KR 102122859 B1, 2020.06.15

高新闻 等. 基于多目标跟踪的交通异常事
件检测. 计算机应用研究. 2020, 第38卷(第06
期), 第1879-1883页.

审查员 张翼

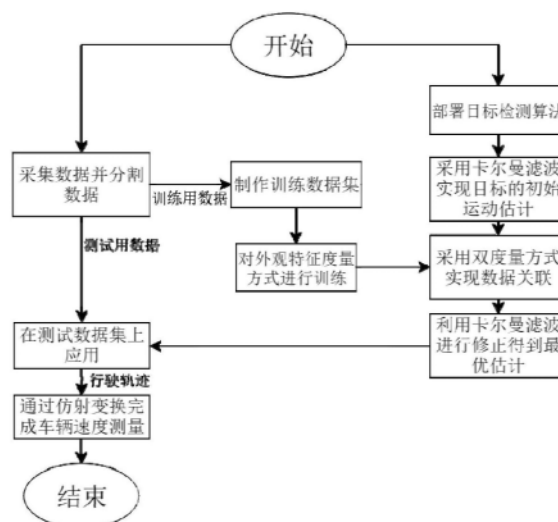
权利要求书4页 说明书8页 附图2页

(54) 发明名称

基于多目标跟踪算法的城市高架道路车辆
测速方法

(57) 摘要

本发明公开了基于多目标跟踪算法的城市
高架道路车辆测速方法, 本发明的有益效果: 本
发明采用先进的多目标跟踪算法, 实现对城市高
架道路中车辆行驶速度的精确测量, 能够提高道
路交通的安全性和效率, 同时本发明能够在无需
添加外部设备的情况下, 通过软件的方式, 在任
意已有的视频监控设备上实现车辆测速功能, 大
大降低了城市道路车辆测速的成本和复杂性。



1. 一种基于多目标跟踪算法的城市高架道路车辆测速方法,其特征在于,具体包括如下步骤:

步骤1)、采集待测区域的视频数据,并将其分为训练用数据和测试用数据;

步骤2)、对步骤1)采集的视频数据进行预处理,对训练用视频数据逐帧进行标记,制作标签;

步骤3)、部署通用多目标跟踪算法,并且利用步骤2)制作得到的训练数据集对初始模型进行特向训练;

步骤4)、采用步骤3)训练得到特向多目标跟踪算法,对步骤1)采集得到的视频数据进行跟踪,得到车辆的运动轨迹;

步骤5)、通过仿射变换完成图像空间到现实空间的映射,再根据车辆的运动轨迹对其行驶速度进行测定;

步骤5)通过仿射变换完成图像空间到现实空间的映射,再根据车辆的运动轨迹对其行驶速度进行测定,具体步骤如下:

步骤5.1)、事先通过测量和标定,将现实场景中,位于摄像机视角内的道路等距离分为若干区域,区域的形状近似为矩形;取区域的四个顶点,并且为其赋予坐标 (x'_1, y'_1) , (x'_2, y'_2) , (x'_3, y'_3) , (x'_4, y'_4) ;

步骤5.2)、通过测量和标定,找出每个区域四个顶点在图像中的像素坐标 (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) ;

步骤5.3)、通过仿射变换计算两者的变换参数,计算公式如下:

$$\begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

其中 $i \in [1, 4]$,分别代入实现标定的四个点,求得该区域所对应的仿射变换参数 a, b, c, d, e, f 的值;然后重复该步骤,计算获得所有候选区域从图像空间向现实空间的仿射变换参数值;

步骤5.4)、对于某一区域内的车辆,逐帧分解其行驶轨迹,得到车辆在 $k-1$ 帧和 k 帧时的图像坐标分别为 (x_{k-1}, y_{k-1}) 和 (x_k, y_k) ;通过该区域的仿射变换参数将坐标映射到现实空间中,得到现实坐标为 (x'_{k-1}, y'_{k-1}) 和 (x'_k, y'_k) ;则车辆在 k 帧的行驶速度表示为:

$$S_k = \sqrt{(x'_k - x'_{k-1})^2 + (y'_k - y'_{k-1})^2}$$

$$v_k = S_k * \text{Fps}$$

其中 S_k 为车辆在 $k-1$ 和 k 帧之间,在现实空间中行驶的距离;Fps为摄像机视频的帧率;

步骤5.5)重复步骤5.4),直至所有车辆在每一帧时的行驶速度,并在距离与时间二维坐标轴中画出其每一点的坐标,最终拟合得出车辆的行驶速度曲线;

所述的步骤2)对步骤1)采集的视频数据进行预处理,对训练用视频数据逐帧进行标记,制作标签,具体检测步骤如下:

步骤2.1)、逐帧切分视频数据,得到单帧图像数据;

步骤2.2)、标出首帧图像中所有的车辆的位置,并且为每一个车辆初始化独立的轨迹

ID;

步骤2.3)、对于在后续帧中出现的车辆,若其在先前帧中出现并已经获得相应的轨迹ID,则采用同样的轨迹ID在后续帧中标记其位置;若为首次出现,则为其初始化独立的轨迹ID,并利用该ID标记其位置;

步骤2.4)、生成相应的坐标文件,并和单帧图像数据相对应,组合成为训练数据集;

所述步骤3)部署通用多目标跟踪算法,并且利用步骤2)制作得到的训练数据集对初始模型进行特向训练,具体如下:

步骤3.1)、首先采用YOLOv3目标检测算法,检测出每一帧图像中的车辆;

步骤3.2)、部署通用多目标跟踪算法,通过相邻帧之间的运动估计和数据关联实现在多帧图像中多个目标的实时跟踪;

步骤3.3)、利用步骤2)制作得到的关于车辆的专项训练数据集,对通用多目标跟踪算法进行关于车辆跟踪的特向训练;

所述步骤3.1)采用YOLOv3目标检测算法,检测出每一帧图像中的车辆,具体如下:

步骤3.1.1)、对每帧输入的图像,首先送入YOLOv3模型,输出一组向量;其中包含了所有目标的位置信息及其类别向量,所述所有目标的位置信息为车辆在图片中的位置信息,具体为检测框的左上角坐标与宽高大小,格式如下公式:

$$p_i = [x, y, w, h]^T$$

步骤3.1.2)、考虑到每一帧中新目标的出现以及旧目标删除的特殊性,故在对检测结果进行运动估计和数据关联之前,设定关于新旧目标建立和删除的规则,如下公式:

$$Bool_i = (x < 20 \text{ or } y < 20 \text{ or } (x_{max} - x) < 20$$

$$\text{or } (y_{max} - y) < 20) \text{ and } \overline{Bool_{exist}^i}$$

上式为逻辑表达式,其中布尔型变量 $Bool_i$ 取值为1表示第i个新对象被建立,取值为0表示已存在的第i个对象被删除; $\overline{Bool_{exist}^i}$ 表示第i个对象是否已存在的取反; $(x < 20 \text{ or } y < 20 \text{ or } (x_{max} - x) < 20 \text{ or } (y_{max} - y) < 20)$ 表示目标位置是否处于图像边缘;即当目标出现在图像边缘时,若其已经存在则代表就目标即将消失,考虑将其删除;若不存在则代表新目标的出现,对其进行初始化;

所述步骤3.2)部署通用多目标跟踪算法,通过相邻帧之间的运动估计和数据关联实现在多帧图像中多个目标的实时跟踪,具体如下:

步骤3.2.1)、对于通过步骤2.1)成功检测和初始化的目标,将采用卡尔曼滤波器对其进行运动估计;首先对卡尔曼滤波器的目标状态参数进行初始化,由于应用场景的特点,模型采用了匀速线性运动假设;其中估计模型的参数如下所示:

$$x = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

其中 u 和 v 分别代表目标中心在图像中的坐标, s 和 r 分别代表目标边框占图像总面积的比例和长宽比; $\dot{u}, \dot{v}, \dot{s}$ 分别代表这三个量在时间上的差分;故该状态估计的观测变量直接写作下式:

$$Observation = [u, v, s, r]^T$$

基于上述的估计模型和观测变量,卡尔曼模型将进行迭代更新;针对车辆目标运动,引入一个离散控制过程系统;该系统用一个随机线性微分方程来描述,使用 $k-1$ 时刻的最优估

计预测得到k时刻车辆的先验状态估计,表示为 \hat{x}_k ;

$$\hat{x}_k = A\hat{x}_{k-1} + Bu_{k-1} + \omega_{k-1}$$

上式中, x_k 是k时刻的n维系统状态向量, u_k 是k时刻对系统的一维控制向量; A 是n*n维状态转移矩阵, B 是系统控制矩阵; ω_k 表示预测过程中的噪声,假设其为高斯白噪声;

步骤3.2.2)、当运动估计模型根据上一帧的状态,输出目标在当前状态的先验估计状态即预测值之后;需要对当前帧的预测值和检测值进行匹配;匹配成功之后才能进行下一步的修正,得到当前帧目标的最优估计;

采用加权方式融合运动相似度和外观相似度的方式来进行数据关联,完成对预测值和检测的匹配;其中运动相似度计算采用马哈拉诺比斯距离,外观相似度采用了余弦距离;马哈拉诺比斯距离的计算方式如下所示:

$$d_{(i,j)}^1 = (d_j - y_i)^T S_i^{-1} (d_j - y_i)$$

其中 (y_i, S_i) 表示第i条轨迹分布 y_i 在测量空间 S_i 上的投影, d_j 则表示第j个检测框结果;马哈拉诺比斯距离代表了检测结果的不确定性,即检测结果距离平均跟踪结果相差了多少个标准差,在物理上反映了物体运动的剧烈程度;因此马哈拉诺比斯距离用作物体运动特征的度量;

外观相似度的计算方式采用最小余弦距离,如下公式:

$$d_{(i,j)}^2 = \min\{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i\}$$

其中 r_i 代表了通过对每个检测结果截取出的局部图像数据计算得到的外观特征描述子;关于外观特征的计算,采用多层卷积神经网络,提取其深层外观特征,作为每个目标的外观特征描述子;

其中卷积神经网络的结构由两个卷积层和六个残差结构组成,最终的整体特征图通过最后的全连接层得到;同时通过标准化将外观特征投影到单位超球面上,以方便将外观特征代入到最小余弦公式的计算中;该网络结构将会利用步骤2)制作的车辆训练数据集进行训练,从而提高其对车辆这一物体外观特征提取的能力;

最终对于度量目标运动程度的马哈拉比斯距离和度量目标之间外观特征相似度的最小余弦距离这两种度量指标,采用加权求和的方式对两者进行融合,如下所示:

$$c_{i,j} = \lambda d_{(i,j)}^1 + (1 - \lambda) d_{(i,j)}^2$$

$c_{i,j}$ 代表了第i个目标检测结果与第j条轨迹之间的整体匹配相似度;其中 λ 代表了两种度量指标的累加权重;当目标的运动程度剧烈或遮挡情况较多时,适当调小 λ 的取值;由于应用对象在城市道路上的行驶车辆,目标的运动较为规律且可见光传感器的视角较高,故 λ 取值为0.5;

步骤3.2.3)、经过步骤3.2.2)完成当前帧的检测值 z_k 和目标的先验估计状态即预测值 \hat{x}_k 的匹配之后,利用当前状态的检测值对先验状态估计进行修正,得到当前状态的最优估计,即后验状态估计 \hat{x}_k 如下式:

$$\hat{x}_k = \hat{x}_k + K(z_k - H\hat{x}_k)$$

其中 z_k 是m维的测量向量,代表了k时刻的测量值; H 是m*n大小观测矩阵; $(z_k - H\hat{x}_k)$ 表示

测量向量与其预测值的差,也称为测量过程的残余,它反映了预测值与实际值之间的误差;残余为0表示预测准确; K 为 $m*n$ 大小矩阵,称为残余的增益,其计算方式如下式所示:

$$P_k = AP_{k-1}A^T + Q$$
$$K = P_k H^T (HP_k H^T + R)^{-1}$$

其中 P_k 为先验协方差,由 $k-1$ 帧后验协方差预测得到;基于获得的 K 值,更新当前状态的后验协方差,如下式所示:

$$P_k = (I - KH)P_k$$

上式计算获得的后验协方差 P_k 将会送入到下一步预测中,并用于先验协方差的预测;当上一帧的预测结果和当前帧的检测结果成功的进行数据关联之后,当前帧的检测结果将用于更新当前帧的目标状态,即更新卡尔曼滤波器迭代参数,并在该状态下通过卡尔曼滤波器最优地求解速度分量;如果数据关联失败,则不利用检测值对卡尔曼滤波器进行更新和修正,而是简单预测其状态;这表明旧目标在新帧中没有被找到,将通过其上一帧的状态简单的线性预测当前帧的状态,直到被找到或进入对象删除判断条件。

基于多目标跟踪算法的城市高架道路车辆测速方法

技术领域：

[0001] 本发明属于图像处理技术领域，具体涉及一种基于多目标跟踪算法的城市高架道路车辆测速方法，常用于城市高架道路，桥梁等设施。

背景技术：

[0002] 车辆超速行驶一直是引起交通事故较为普遍和直接的因素，若不进行严格的监测，将会对人们生命健康和社会秩序造成巨大的影响。同时对车辆行驶速度的监测也对减少交通拥堵和有效进行车辆调控有着重大帮助。

[0003] 目前最为普遍的车辆测速方式是采用超声波和激光测速仪，这种方式测速准确，但是由于其昂贵的设备费用以及较为严格的工作环境，目前无法像视频监控一样广泛的安装在城市道路的各个角落，虽有采用深度学习技术，通过Faster RCNN目标检测技术和SORT多目标跟踪技术实现了车辆测速，但是由于其采用的技术较为简单，检测的速度和准确度仍存在不足。

发明内容：

[0004] 本发明的目的就在于为了解决上述问题而提供基于多目标跟踪算法的城市高架道路车辆测速方法。

[0005] 为了解决上述问题，本发明提供了一种技术方案：

[0006] 一种基于多目标跟踪算法的城市高架道路车辆测速方法，具体包括如下步骤：

[0007] 步骤1) 采集待测区域的视频数据，并将其分为训练用数据和测试用数据；

[0008] 步骤2) 对步骤1) 采集的视频数据进行预处理，对训练用视频数据逐帧进行标记，制作标签。具体步骤如下：

[0009] 步骤2.1) 逐帧切分视频数据，得到单帧图像数据；

[0010] 步骤2.2) 标出首帧图像中所有的车辆的位置，并且为每一个车辆初始化独立的轨迹ID；

[0011] 步骤2.3) 对于在后续帧中出现的车辆，若其在先前帧中出现并已经获得相应的轨迹ID，则采用同样的轨迹ID在后续帧中标记其位置。若为首次出现，则为其初始化独立的轨迹ID，并利用该ID标记其位置；

[0012] 步骤2.4) 生成相应的坐标文件，并和单帧图像数据相对应，组合成为训练数据集；

[0013] 步骤3) 部署通用多目标跟踪算法，并且利用步骤2) 制作得到的训练数据集对初始模型进行特向训练。具体步骤如下：

[0014] 步骤3.1) 首先采用YOLOv3目标检测算法，检测出每一帧图像中的车辆。具体步骤如下：

[0015] 步骤3.1.1) 对每帧输入的图像，首先送入YOLOv3模型，输出一组向量。其中包含了所有目标的位置信息及其类别向量，对于本方法的应用场景，只需检测车辆，故只取其中的车辆检测类别。如此输出的结果只包含了车辆在图片中的位置信息，具体为检测框的左上

角坐标与宽高大小,格式如下公式:

$$[0016] \quad p_i = [x, y, w, h]^T$$

[0017] 步骤3.1.2) 考虑到每一帧中新目标的出现以及旧目标删除的特殊性,故在对检测结果进行运动估计和数据关联之前,设定关于新旧目标建立和删除的规则,如下公式:

$$[0018] \quad Bool_i = (x < 20 \text{ or } y < 20 \text{ or } (x_{max} - x) < 20$$

$$[0019] \quad \text{or } (y_{max} - y) < 20) \text{ and } \overline{Bool_{exist}^i}$$

[0020] 上式为逻辑表达式,其中布尔型变量 $Bool_i$ 取值为1表示第*i*个新对象被建立,取值为0表示已存在的第*i*个对象被删除。 $\overline{Bool_{exist}^i}$ 表示第*i*个对象是否已存在的取反。 $(x < 20 \text{ or } y < 20 \text{ or } (x_{max} - x) < 20 \text{ or } (y_{max} - y) < 20)$ 表示目标位置是否处于图像边缘。即当目标出现在图像边缘时,若其已经存在则代表就目标即将消失,考虑将其删除;若不存在则代表新目标的出现,对其进行初始化。

[0021] 步骤3.2) 部署通用多目标跟踪算法,通过相邻帧之间的运动估计和数据关联实现在多帧图像中多个目标的实时跟踪。具体步骤如下:

[0022] 步骤3.2.1) 对于通过步骤2.1) 成功检测和初始化的目标,将采用卡尔曼滤波器对其进行运动估计。首先对卡尔曼滤波器的目标状态参数进行初始化,由于专利应用场景的特点,模型采用了匀速线性运动假设。其中估计模型的参数如下所示:

$$[0023] \quad x = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

[0024] 其中u和v分别代表目标中心在图像中的坐标,s和r分别代表目标边框占图像总面积的比例和长宽比。 $\dot{u}, \dot{v}, \dot{s}$ 分别代表这三个量在时间上的差分。故该状态估计的观测变量可以直接写作下式:

$$[0025] \quad \text{Observation} = [u, v, s, r]^T$$

[0026] 基于上述的估计模型和观测变量,卡尔曼模型将进行迭代更新。针对车辆目标运动,引入一个离散控制过程系统。该系统可用一个随机线性微分方程来描述,使用*k*-1时刻的最优估计预测得到*k*时刻车辆的先验状态估计,表示为 \hat{x}_k :

$$[0027] \quad \hat{x}_k = A\hat{x}_{k-1} + Bu_{k-1} + \omega_{k-1}$$

[0028] 上式中, x_k 是*k*时刻的*n*维系统状态向量, u_k 是*k*时刻对系统的一维控制向量。*A*是*n***n*维状态转移矩阵,*B*是系统控制矩阵。 ω_k 表示预测过程中的噪声,假设其为高斯白噪声。

[0029] 步骤3.2.2) 当运动估计模型根据上一帧的状态,输出目标在当前状态的先验估计状态(以下简称预测值)之后。需要对当前帧的预测值和检测值进行匹配。匹配成功之后才能进行下一步的修正,得到当前帧目标的最优估计。

[0030] 本方法采用加权方式融合运动相似度和外观相似度的方式来进行数据关联,完成对预测值和检测的匹配。其中运动相似度计算采用马哈拉诺比斯距离,外观相似度采用了余弦距离。马哈拉诺比斯距离的计算方式如下所示:

$$[0031] \quad d_{(i,j)}^1 = (d_j - y_i)^T S_i^{-1} (d_j - y_i)$$

[0032] 其中 (y_i, S_i) 表示第*i*条轨迹分布 y_i 在测量空间 S_i 上的投影, d_j 则表示第*j*个检测框结果。马哈拉诺比斯距离代表了检测结果的不确定性,即检测结果距离平均跟踪结果相差

了多少个标准差,在物理上反映了物体运动的剧烈程度。因此该指标可以用作物体运动特征的度量。

[0033] 外观相似度的计算方式采用最小余弦距离,如下公式:

$$[0034] \quad d_{(i,j)}^2 = \min \{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i\}$$

[0035] 其中 r_i 代表了通过对每个检测结果截取出的局部图像数据计算得到的外观特征描述子。关于外观特征的计算,本方法采用多层卷积神经网络,提取其深层外观特征,作为每个目标的外观特征描述子。

[0036] 其中卷积神经网络的结构由两个卷积层和六个残差结构组成,最终的整体特征图通过最后的全连接层得到。同时通过标准化将外观特征投影到单位超球面上,以方便将外观特征代入到最小余弦公式的计算中。该网络结构将会利用步骤2)制作的车辆训练数据集进行训练,从而提高其对车辆这一物体外观特征提取的能力。

[0037] 最终对于度量目标运动程度的马哈拉比斯距离和度量目标之间外观特征相似度的最小余弦距离这两种度量指标,采用加权求和的方式对两者进行融合,如下所示:

$$[0038] \quad c_{i,j} = \lambda d_{(i,j)}^1 + (1 - \lambda) d_{(i,j)}^2$$

[0039] $c_{i,j}$ 代表了第i个目标检测结果与第j条轨迹之间的整体匹配相似度。其中 λ 代表了两种度量指标的累加权重。当目标的运动程度剧烈或者遮挡情况较多时,可以适当调小 λ 的取值。但是本方法主要的应用对象在城市道路上的行驶车辆,目标的运动较为规律且可见光传感器的视角较高,故 λ 取值为0.5;

[0040] 步骤3.2.3) 经过步骤3.2.2) 完成当前帧的检测值 z_k 和目标的先验估计状态(预测值) \hat{x}_k 的匹配之后,利用当前状态的检测值对先验状态估计进行修正,得到当前状态的最优估计,即后验状态估计 \hat{x}_k 如下式:

$$[0041] \quad \hat{x}_k = \hat{x}_k + K(z_k - H\hat{x}_k)$$

[0042] 其中 z_k 是m维的测量向量,代表了k时刻的测量值。H是m*n大小观测矩阵。 $(z_k - H\hat{x}_k)$ 表示测量向量与其预测量(先验估计)的差,也称为测量过程的残余,它反映了预测值与实际值之间的误差。残余为0表示预测准确。K为m*n大小矩阵,称为残余的增益,其计算方式如下式所示:

$$[0043] \quad P_k = AP_{k-1}A^T + Q$$

$$[0044] \quad K = P_k H^T (H P_k H^T + R)^{-1}$$

[0045] 其中 P_k 为先验协方差,由k-1帧后验协方差预测得到。基于获得的K值,更新当前状态的后验协方差,如下式所示:

$$[0046] \quad P_k = (I - KH)P_k$$

[0047] 上式计算获得的后验协方差 P_k 将会送入到下一步预测中,并用于先验协方差的预测。当上一帧的预测结果和当前帧的检测结果成功的进行数据关联之后,当前帧的检测结果将用于更新当前帧的目标状态,即更新卡尔曼滤波器迭代参数,并在该状态下通过卡尔曼滤波器框架最优地求解速度分量。如果数据关联失败,则不利用检测值对卡尔曼滤波器

进行更新和修正,而是简单预测其状态。这表明旧目标在新帧中没有被找到,将通过其上一帧的状态简单的线性预测当前帧的状态,直到被找到或者进入对象删除判断条件。

[0048] 步骤3.3) 利用步骤2) 制作得到的关于车辆的专项训练数据集,对通用多目标跟踪算法进行关于车辆跟踪的特向训练;

[0049] 步骤4) 采用步骤3) 训练得到特向多目标跟踪算法,对步骤1) 采集得到的视频数据进行跟踪,得到车辆的运动轨迹。

[0050] 步骤5) 通过仿射变换完成图像空间到现实空间的映射,再根据车辆的运动轨迹对其行驶速度进行测定。具体步骤如下:

[0051] 步骤5.1) 事先通过测量和标定,将现实场景中,位于摄像机视角内的道路等距离分为若干区域,区域的形状近似为矩形。取区域的四个顶点,并且为其赋予坐标 (x'_1, y'_1) , (x'_2, y'_2) , (x'_3, y'_3) , (x'_4, y'_4) ;

[0052] 步骤5.2) 通过测量和标定,找出每个区域四个顶点在图像中的像素坐标 (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) ;

[0053] 步骤5.3) 通过仿射变换计算两者的变换参数,计算公式如下:

$$[0054] \quad \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

[0055] 其中 $i \in [1, 4]$, 分别代入实现标定的四个点,求得该区域所对应的仿射变换参数 a, b, c, d, e, f 的值。然后重复该步骤,计算获得所有候选区域从图像空间向现实空间的仿射变换参数值。

[0056] 步骤5.4) 对于某一区域内的车辆,逐帧分解其行驶轨迹,得到车辆在 $k-1$ 帧和 k 帧时的图像坐标分别为 (x_{k-1}, y_{k-1}) 和 (x_k, y_k) 。通过该区域的仿射变换参数将改坐标映射到现实空间中,得到现实坐标为 (x'_{k-1}, y'_{k-1}) 和 (x'_k, y'_k) 。则车辆在 k 帧的行驶速度可以表示为:

$$[0057] \quad S_k = \sqrt{(x'_k - x'_{k-1})^2 + (y'_k - y'_{k-1})^2}$$

$$[0058] \quad v_k = S_k * \text{Fps}$$

[0059] 其中 S_k 为车辆在 $k-1$ 和 k 帧之间,在现实空间中行驶的距离。 Fps 为摄像机视频的帧率。

[0060] 步骤5.5) 重复步骤5.4), 直至所有车辆在每一帧时的行驶速度,并在距离与时间二维坐标轴中画出其每一点的坐标,最终拟合得出车辆的行驶速度曲线。

[0061] 本发明的有益效果:本发明采用先进的多目标跟踪算法,实现对城市高架道路中车辆行驶速度的精确测量,能够提高道路交通安全性和效率,同时本发明能够在无需添加外部设备的情况下,通过软件的方式,在任意已有的视频监控设备上实现车辆测速功能,大大降低了城市道路车辆测速的成本和复杂性。

附图说明:

[0062] 为了易于说明,本发明由下述的具体实施及附图作以详细描述。

[0063] 图1为本发明的整体结构流程图;

[0064] 图2为卡尔曼滤波预测和修正流程图;

[0065] 图3为速度曲线拟合示意图。

具体实施方式：

[0066] 如图1-3所示,本具体实施方式采用以下技术方案:一种基于多目标跟踪算法的城市高架道路车辆测速方法,具体包括如下步骤:

[0067] a.使用城市高架道路监控,采集待测区域的视频数据,并按照固定比例将其分为训练用数据和测试用数据;

[0068] b.对步骤a中采集的视频数据进行预处理,首先逐帧切分视频数据,得到单帧图像数据;

[0069] 然后标出首帧图像中所有的车辆的位置,并且为每一个车辆初始化独立的轨迹ID;对于在后续帧中出现的车辆,若其在先前帧中出现并已经获得相应的轨迹ID,则采用同样的轨迹ID在后续帧中标记其位置。若为首次出现,则为其初始化独立的轨迹ID,并利用该ID标记其位置;最后生成相应的坐标文件,并和单帧图像数据相对应,组合成为训练数据集;

[0070] c.部署通用目标检测算法。首先对于目标检测算法采用YOLOv3算法,检测出每一帧图像中的车辆。对每帧输入的图像,首先送入YOLOv3模型,输出一组向量。其中包含了所有目标的位置信息及其类别向量,对所用的COCO数据集训练的YOLOv3模型来说,类别向量 \bar{n} 是一组81维向量,表示81个物体类别。

[0071] 对于本专利的应用场景,只需检测车辆,故只取其中的车辆检测类别。如此输出的结果只包含了车辆在图片中的位置信息,具体为检测框的左上角坐标与宽高大小,格式如下公式:

$$[0072] \quad p_i = [x, y, w, h]^T$$

[0073] d.设定新旧目标初始化和删除的规则。考虑到每一帧中新目标的出现以及旧目标删除的特殊性,故在对检测结果进行运动估计和数据关联之前,设定关于新旧目标建立和删除的规则,如下公式:

$$[0074] \quad Bool_i = (x < 20 \text{ or } y < 20 \text{ or } (x_{max} - x) < 20$$

$$[0075] \quad \text{or } (y_{max} - y) < 20) \text{ and } \overline{Bool_{exist}^i}$$

[0076] 上式为逻辑表达式,其中布尔型变量Bool_i取值为1表示第i个新对象被建立,取值为0表示已存在的第i个对象被删除。 $\overline{Bool_{exist}^i}$ 表示第i个对象是否已存在的取反。(x < 20 or y < 20 or (x_{max} - x) < 20 or (y_{max} - y) < 20)表示目标位置是否处于图像边缘。即当目标出现在图像边缘时,若其已经存在则代表就目标即将消失,考虑将其删除;若不存在则代表新目标的出现,对其进行初始化。

[0077] e.采用卡尔曼滤波算法实现相邻帧之间的运动估计。对于通过步骤d成功检测和初始化的目标,将采用卡尔曼滤波器对其进行运动估计。首先对卡尔曼滤波器的目标状态参数进行初始化,由于专利应用场景的特点,模型采用了匀速线性运动假设。其中估计模型的参数如下所示:

$$[0078] \quad x = [u, v, s, r, \dot{u}, \dot{v}, \dot{s}]^T$$

[0079] 其中u和v分别代表目标中心在图像中的坐标,s和r分别代表目标边框占图像总面

积的比例和长宽比。 $\dot{u}, \dot{v}, \dot{s}$ 分别代表这三个量在时间上的差分。故该状态估计的观测变量可以直接写作下式:

$$[0080] \quad \text{Observation} = [u, v, s, r]^T$$

[0081] 基于上述的估计模型和观测变量,卡尔曼模型将进行迭代更新。针对车辆目标运动,引入一个离散控制过程系统。该系统可用一个随机线性微分方程来描述,使用k-1时刻的最优估计预测得到k时刻车辆的先验状态估计,表示为 \hat{x}_k^- :

$$[0082] \quad \hat{x}_k^- = A\hat{x}_{k-1} + Bu_{k-1} + \omega_{k-1}$$

[0083] 上式中, x_k 是k时刻的n维系统状态向量, u_k 是k时刻对系统的一维控制向量。A是n*n维状态转移矩阵,B是系统控制矩阵。 ω_k 表示预测过程中的噪声,假设其为高斯白噪声。

[0084] f.利用运动相似度和外观相似度实现预测值和检测之间的数据关联。当步骤e根据上一帧的状态,输出目标在当前状态的先验估计状态(以下简称预测值)之后。需要对当前帧的预测值和检测值进行匹配。匹配成功之后才能进行下一步的修正,得到当前帧目标的最优估计。

[0085] 本专利采用加权方式融合运动相似度和外观相似度的方式来进行数据关联,完成对预测值和检测的匹配。其中运动相似度计算采用马哈拉诺比斯距离,外观相似度采用了余弦距离。马哈拉诺比斯距离的计算方式如下所示:

$$[0086] \quad d_{(i,j)}^1 = (d_j - y_i)^T S_i^{-1} (d_j - y_i)$$

[0087] 其中 (y_i, S_i) 表示第i条轨迹分布 y_i 在测量空间 S_i 上的投影, d_j 则表示第j个检测框结果。马哈拉诺比斯距离代表了检测结果的不确定性,即检测结果距离平均跟踪结果相差了多少个标准差,在物理上反映了物体运动的剧烈程度。因此该指标可以用作物体运动特征的度量。

[0088] 当运动不确定性较低,即物体的运动程度较为平缓,马哈拉诺比斯距离的度量性能较为优秀。但是针对本专利的城市道路车辆行驶的应用场景,极易出现车辆相互遮挡的情况,此时卡尔曼滤波预测的状态分布将快速离散,马哈拉比斯距离的度量性能将大幅下降。针对于此,本专利引入目标的外观特征距离作为另一度量指标来应对此问题。

[0089] 外观相似度的计算方式采用最小余弦距离,如下公式:

$$[0090] \quad d_{(i,j)}^2 = \min \{1 - r_j^T r_k^{(i)} \mid r_k^{(i)} \in R_i\}$$

[0091] 其中 r_i 代表了通过对每个检测结果截取出的局部图像数据计算得到的外观特征描述子。关于外观特征的计算,本专利采用多层卷积神经网络,提取其深层外观特征,作为每个目标的外观特征描述子。

[0092] 其中卷积神经网络的结构由两个卷积层和六个残差结构组成,最终的整体特征图通过最后的全连接层得到。同时通过标准化将外观特征投影到单位超球面上,以方便将外观特征代入到最小余弦公式的计算中。该网络结构将会利用步骤b制作的车辆训练数据集进行训练,从而提高其对车辆这一物体外观特征提取的能力。

[0093] 最终对于度量目标运动程度的马哈拉比斯距离和度量目标之间外观特征相似度的最小余弦距离这两种度量指标,采用加权求和的方式对两者进行融合,如下所示:

$$[0094] \quad c_{i,j} = \lambda d_{(i,j)}^1 + (1 - \lambda) d_{(i,j)}^2$$

[0095] $c_{i,j}$ 代表了第i个目标检测结果与第j条轨迹之间的整体匹配相似度。其中 λ 代表了两种度量指标的累加权重。当目标的运动程度剧烈或者遮挡情况较多时,可以适当调小 λ 的取值。但是本专利主要的应用对象在城市道路上的行驶车辆,目标的运动较为规律且可见光传感器的视角较高,故 λ 取值为0.5;

[0096] g.再次利用卡尔曼滤波算法,实现当前帧最终跟踪结果的输出。经过步骤f完成当前帧的检测值 z_k 和目标的先验估计状态(预测值) \hat{x}_k 的匹配之后,利用当前状态的检测值对先验状态估计进行修正,得到当前状态的最优估计,即后验状态估计 \hat{x}_k 如下式:

$$[0097] \quad \hat{x}_k = \hat{x}_k + K(z_k - H\hat{x}_k)$$

[0098] 其中 z_k 是m维的测量向量,代表了k时刻的测量值。H是m*n大小观测矩阵。 $(z_k - H\hat{x}_k)$ 表示测量向量与其预测量(先验估计)的差,也称为测量过程的残余,它反映了预测值与实际值之间的误差。残余为0表示预测准确。K为m*n大小矩阵,称为残余的增益,其计算方式如下式所示:

$$[0099] \quad P_k = AP_{k-1}A^T + Q$$

$$[0100] \quad K = P_k H^T (H P_k H^T + R)^{-1}$$

[0101] 其中 P_k 为先验协方差,由k-1帧后验协方差预测得到。基于获得的K值,更新当前状态的后验协方差,如下式所示:

$$[0102] \quad P_k = (I - KH)P_k$$

[0103] 上式计算获得的后验协方差 P_k 将会送入到下一步预测中,并用于先验协方差的预测。当上一帧的预测结果和当前帧的检测结果成功的进行数据关联之后,当前帧的检测结果将用于更新当前帧的目标状态,即更新卡尔曼滤波器迭代参数,并在该状态下通过卡尔曼滤波器框架最优地求解速度分量。如果数据关联失败,则不利用检测值对卡尔曼滤波器进行更新和修正,而是简单预测其状态。这表明旧目标在新帧中没有被找到,将通过其上一帧的状态简单的线性预测当前帧的状态,直到被找到或者进入对象删除判断条件。

[0104] h.利用步骤b制作得到的关于车辆的专项训练数据集,对通用多目标跟踪算法进行关于车辆跟踪的特向训练。其中训练过程采用Adam optimizer对模型进行10个epoch的特向训练。软件采用pyTorch实现,硬件采用NVIDIA Tesla T4-16GB显卡。

[0105] i.采用步骤h训练得到特向多目标跟踪算法,对步骤a采集得到的视频数据进行跟踪,得到车辆的运动轨迹。

[0106] j.通过仿射变换完成图像空间到现实空间的映射,再根据车辆的运动轨迹对其行驶速度进行测定。事先通过测量和标定,将现实场景中,位于摄像机视角内的道路等距离分为若干区域,区域的形状近似为矩形。取区域的四个顶点,并且为其赋予坐标 (x'_1, y'_1) , (x'_2, y'_2) , (x'_3, y'_3) , (x'_4, y'_4) ;再通过测量和标定,找出每个区域四个顶点在图像中的像素坐标 (x_1, y_1) , (x_2, y_2) , (x_3, y_3) , (x_4, y_4) ;通过仿射变换计算两者的变换参数,计算公式如下:

$$[0107] \quad \begin{bmatrix} x'_i \\ y'_i \\ 1 \end{bmatrix} = \begin{bmatrix} a & b & c \\ d & e & f \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix}$$

[0108] 其中 $i \in [1, 4]$, 分别代入实现标定的四个点, 求得该区域所对应的仿射变换参数 a, b, c, d, e, f 的值。然后重复该步骤, 计算获得所有候选区域从图像空间向现实空间的仿射变换参数值;

[0109] 然后对于某一区域内的车辆, 逐帧分解其行驶轨迹, 得到车辆在 $k-1$ 帧和 k 帧时的图像坐标分别为 (x_{k-1}, y_{k-1}) 和 (x_k, y_k) 。通过该区域的仿射变换参数将改坐标映射到现实空间中, 得到现实坐标为 (x'_{k-1}, y'_{k-1}) 和 (x'_k, y'_k) 。则车辆在 k 帧的行驶速度可以表示为:

$$[0110] \quad S_k = \sqrt{(x'_k - x'_{k-1})^2 + (y'_k - y'_{k-1})^2}$$

$$[0111] \quad v_k = S_k * \text{Fps}$$

[0112] 其中 S_k 为车辆在 $k-1$ 和 k 帧之间, 在现实空间中行驶的距离。 Fps 为摄像机视频的帧率。

[0113] 最后重复上述操作, 直至所有车辆在每一帧时的行驶速度, 并在距离与时间二维坐标轴中画出其每一点的坐标, 最终拟合得出车辆的行驶速度曲线。

[0114] 以上显示和描述了本发明的基本原理和主要特征和本发明的优点, 本行业的技术人员应该了解, 本发明不受上述实施例的限制, 上述实施例和说明书中描述的只是说明本发明的原理, 在不脱离本发明精神和范围的前提下, 本发明还会有各种变化和改进, 这些变化和进步都落入要求保护的本发明范围内, 本发明要求保护范围由所附的权利要求书及其等效物界定。

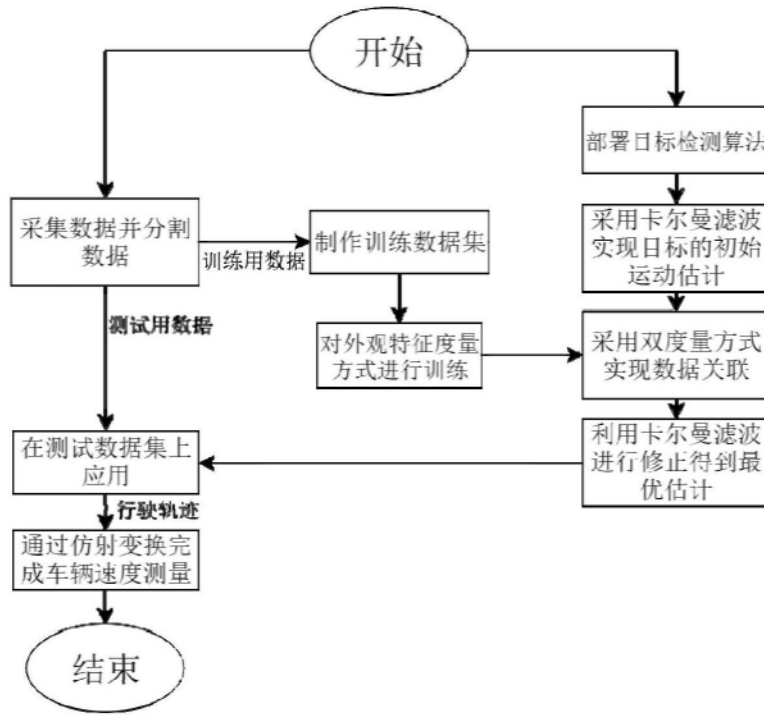


图1

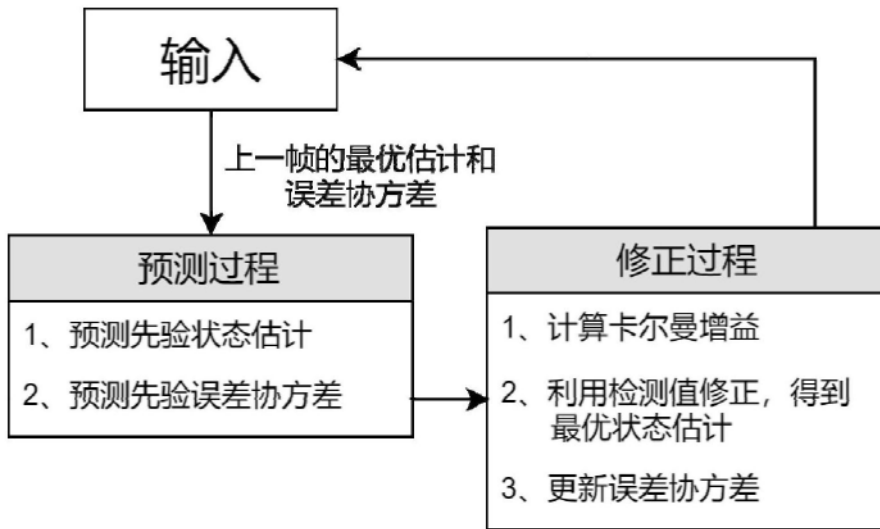


图2

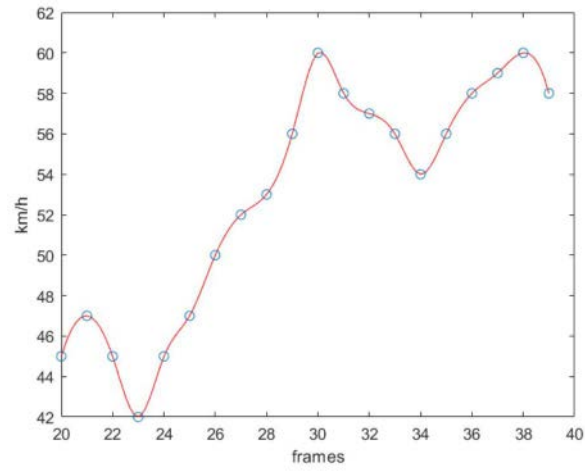


图3