

地铁应用场景中半监督群体异常行为检测模型^{*}

程媛¹ 吴帆² 张宁² 徐炜²

(1. 中铁第四勘察设计院集团有限公司, 430063, 武汉;

2. 东南大学智能运输系统研究中心轨道交通研究所, 210018, 南京//第一作者, 正高级工程师)

摘要 为了解决拥挤、奔跑及踩踏等地铁群体异常行为带来的公共安全风险问题, 针对地铁车站监控视频样本的种类繁多、均衡性差、背景环境复杂多样等问题, 提出一种可迁移性强、训练及运行效率高、样本量标注较少的半监督群体异常行为检测模型。结合卷积自编码器和卷积长短期记忆网络, 分别对图像样本进行空间维度特征压缩重构和时间维度运动特征叠加, 并通过自动检测实时监控视频图像, 及时发现不同场景下人群状态的异常行为, 以降低危害乘客财产安全的可能性。基于现有的视频监控系统, 所提模型可以完成对站内高发区域和场景的全覆盖, 将监控图像每个像素的重构误差以热图或散点图的形式叠加到原始图像中, 有助于监管人员迅速发现异常区域, 并及时响应处置。在 Subway、CHUK Avenue 及自建的无锡地铁数据集上验证了所提模型的准确性。研究结果表明, 与同类经典模型相比, 所提模型在提升检测精度的同时也能基本满足地铁车站场景下的实时性应用需求。

关键词 地铁; 半监督; 群体异常行为

中图分类号 U29-39

DOI:10.16037/j.1007-869x.2023.05.004

Detection Model of Semi-supervised Group Abnormal Behavior in Metro Application Scenario

CHENG Yuan, WU Fan, ZHANG Ning, XU Wei

Abstract To solve the public safety risk issues caused by group abnormal behavior in metro such as crowding, running and trampling, targeting the wide variety, poor balance, complex and diverse background environment problems of metro station monitoring video samples, a semi-supervised group abnormal behavior detection model with strong mobility, high training and operation efficiency, and less sample labeling is proposed. Combined with convolutional auto-encoder and convolutional long/short term memory network, the image samples are compressed and reconstructed with spatial dimension features, and superimposed with temporal dimension motion

features respectively. Meanwhile, real-time monitoring video images are automatically examined, thus group abnormal behaviors in different scenarios are identified in time, lowering the threat on public/private property and safety. Based on the existing video monitoring system, the proposed model can accomplish full coverage of high-incidence areas and scenes in the station, and the reconstruction error per pixel of the monitoring image is superimposed into the original image in the form of heat map or scatter plot, which is helpful for the supervisors to quickly locate abnormal areas and respond in time. Accuracy of the proposed model is verified by Subway, CHUK Avenue and the self-built Wuxi Metro dataset. Research results show that compared with similar classical models, the proposed one can elevate detection accuracy while generally meeting the real-time application requirements in metro station scenario.

Key words metro; semi-supervised; group abnormal behavior

First-author's address China Railway Siyuan Survey and Design Group Co., Ltd., 430063, Wuhan, China

0 引言

地铁车站具有场景复杂、客流流动性大、密度高等特征, 这对其运营管理提出了极大的挑战, 类似奔跑和踩踏等群体异常事件使得地铁的安防工作面临着严峻的考验。作为地铁车站监控的主要手段, 车站范围内的视频监控还停留在人工查验的初级阶段。这种以人力为主的查验方式在效率、成本及正确率等方面均有诸多不足之处, 容易出现错检、漏检等情况^[1-2]。因此基于现有视频监控系统, 开展智能化异常行为检测技术的研究是非常有必要的。智能化异常行为检测技术能够迅速检测人群状态、加快事件介入速度、减少群体异常行为带来的损失, 进一步提升地铁的运营管理

^{*} 交通运输部建设科技项目(2015318J33080); 江苏省重点研发计划(社会发展)项目(BE2016740)

水平和应急处置能力,从而保障地铁的正常运行。

近年来,智能化安防下的异常行为检测技术已经取得了显著进展。根据样本标签的可用性,群体异常行为检测模型可以分为 3 类:有监督学习检测模型、无监督学习检测模型和半监督学习检测模型。有监督学习检测模型的实质是将群体异常行为检测问题看作是一个多分类问题,利用标记好的训练样本训练获得评价规则下的最优模型,判定异常事件的种类。无监督学习检测模型通过聚类 and 关联等方法直接对所训练的样本集进行建模,学习出群体的行为种类。半监督学习检测模型基于已标记的群体正常行为运动特征,通过其概率、密度或样本域等方法构建模型,与模型不匹配的事件即判定为异常事件。文献[3-5]提出了一种稀疏重构代价的方法,首先通过正常样本来训练行为字典,再根据测试样本在行为字典上的稀疏重构损失与阈值之间的关系来判断异常事件是否发生。考虑到地铁监控视频中存在场景复杂、拥挤环境下对乘客遮挡严重、乘客群体异常行为种类多且定义不唯一、样本不均衡等特点,结合地铁的实际管理需求,有监督学习检测模型和无监督学习检测模型均对样本质量提出了较高的要求,不能很好地满足地铁安防工作的实际应用需求。

本文基于卷积自编码器和 LSTM(长短期记忆)模型,利用状态空间压缩还原机制构建可以自动提取行为特征的半监督群体异常行为检测模型(以下简称“检测模型”)。该检测模型重点分析了乘客的宏观运动特征,提升了判定异常行为是否发生的准确率,具有较强的可迁移性,能够广泛应用于不同应用场景下的地铁视频监控。

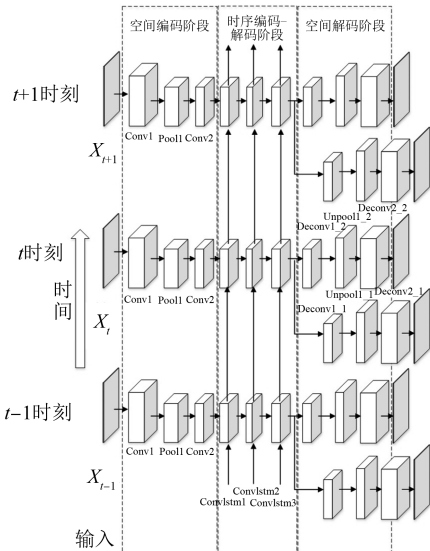
1 群体异常行为检测模型

检测模型以分辨率为 227 像素×227 像素的灰度视频图像数据作为输入,结合 LSTM 模型的卷积自编码器,运用历史回归检验等多种优化手段,提高其异常行为检测的准确性和时效性。

1.1 检测模型整体架构

检测模型整体结构示意图如图 1 所示,各层间以激活函数串联传递。由图 1 可知,检测模型的主干结构为卷积自编码器的空间编码和解码,以及两者间卷积 LSTM 模型的时序编解码环节。检测模型中的权重参数采用更为有效的 Xavier 算法进行随机初始化,保持每一层输出值的方差相等,尽可能

避免出现梯度消失或梯度饱和的情况。卷积层和反卷积层的激活函数均选用 ReLU(整流线性单元)函数,以提高检测模型的运算速度和收敛速度。池化层和反池化层均选用效果较好的最大化池化和反池化,以减少信息丢失的情况。受多任务学习理论的启发,空间解码阶段拆分为具有相同隐藏特征层结构的两个分支,分别重构当前的帧序列和过去的帧序列,对历史帧的重构可以进一步提高检测模型的检测效果。



注: X_t 为 t 时刻的输入视频帧; Conv 为卷积层网络; Deconv 为反卷积层; ConvLstm 为卷积长短期记忆层; Pool 为池化层; Unpool 为反池化层。

图 1 检测模型整体结构示意图

Fig. 1 Diagram of detection model overall framework

模型的训练阶段采用欧式距离衡量损失,通过不断优化检测模型中各层的权重参数,最终获得训练好的非线性分类(正常行为和其他异常行为)模型。模型训练的损失函数可以表示为:

$$\hat{f}_w = \argmin_w \frac{1}{2N} \sum_t \|X_t - f_w(X_t)\|_2^2 + \|X_{t-1} - g_w(X_t)\|_2^2 + \gamma \|W\|_2^2 \quad (1)$$

式中:

N ——一次性进入模型的样本个数;

$f_w(\cdot)$ ——参数为 W 的重构自身的非线性变换函数;

$g_w(\cdot)$ ——参数为 W 的重构前一时空片段的非线性变换函数;

γ ——L2 范数权重系数;

$\|W\|_2^2$ ——L2 范数,可以提高检测模型的泛化能力。

1.2 卷积自编码器

类似于传统的自编码器模型,作为检测模型主体结构的卷积自编码器模型同样可以分为编码和解码两个阶段。编码阶段通过卷积层和池化层交替处理,提取出图像的空间特征;解码阶段通过反卷积层和反池化层交替处理,将所提取的空间特征向量重构为与原始数据数据维度相同的图像。最后,通过计算重构图像与原始输入数据之间的误差,进行多次迭代训练后获得卷积自编码器参数。

1.3 卷积 LSTM 神经网络

Conv-LSTM (卷积长短期记忆)神经网络是在2015年首次提出的,是一种针对时空序列数据的神经网络新结构。不同于传统的通过全连接层连接的LSTM模型,Conv-LSTM神经网络将传统的全连接结构转换为卷积结构,不仅充分利用了卷积核的权值共享特点,还能在保证精度的前提下,提高检测模型的运行效率,更好地提取运动的空间特征。

本文通过在卷积自编码器的空间编码阶段与空间解码阶段插入时序编码阶段和解码阶段,将视频图像的空间特征作进一步解析和扩展。所提检测模型可以更好地检测视频中空间的运动特征信息,可在一定程度上提高检测模型的准确性和实用性。

2 群体异常行为标识

基于半监督学习理论,检测模型的训练阶段只采集群体正常行为事件视频作为模型的输入,通过不断优化损失函数来训练和学习检测模型的参数,当损失函数趋于收敛时停止训练检测模型。此时,根据重构误差即可判断出此刻的乘客群体行为是否正常。

2.1 异常行为判别

检测模型的作用主要是对群体异常行为进行判别。首先,通过滑动窗口获取连续视频帧,构成检测模型的时空片段输入,通过基于Conv-LSTM神经网络改进的自编码器模型对输入数据进行重构,以达到对群体异常行为进行检测的目的。当检测模型的输入为正常群体行为视频帧时,其对应的样本重构误差较小;而当检测模型的输入为异常群体行为视频帧时,则会存在较大的重构误差。 t 时刻下,视频帧的重构误差总和 $e(t)$ 可以表示为:

$$e(t) = \sum_{(x,y)} e(x,y,t) = \sum_{(x,y)}$$

$$\|I(x,y,t) - f_w(I(x,y,t))\|_2 \quad (2)$$

$$e(x,y,t) = \|I(x,y,t) - f_w(I(x,y,t))\|_2 \quad (3)$$

式中:

$I(x,y,t)$ —— 输入图像序列中 t 时刻所在帧空间位置处 (x,y) 点的像素值;

$e(x,y,t)$ —— 输入图像序列中 t 时刻所在帧空间位置处 (x,y) 点的像素重构误差。

对重构误差进行归一化,将异常得分限定在0~1之间,可以获得群体异常行为的评价函数。异常程度评分函数 $s_a(t)$ 和正常程度评分函数 $s_r(t)$ 可以表示为:

$$s_a(t) = \frac{e(t) - \min_t e(t)}{\max_t e(t) - \min_t e(t)} \quad (4)$$

$$s_r(t) = 1 - s_a(t) \quad (5)$$

式中:

$\min_t e(t)$ —— 所有时刻图像重构误差总和的最小值;

$\max_t e(t)$ —— 所有时刻图像重构误差总和的最大值。

越不符合正常群体行为模式的事件正常得分值就越低。在获得每一帧的正常程度得分后,利用Persistence1D算法获得局部最小值,并通过与设定好的阈值进行比较来判断此时是否发生异常。群体异常行为通常是一个持续的过程,持续时间一般至少为2~3s,本文将异常点前后2s设为同一件群体异常行为,相邻的2个异常点(前后间隔时间差小于2s)进行合并处理。当检测模型的测试结果与标注的重叠率超过50%时,认为异常行为被成功检测到。

2.2 异常区域定位

绝大多数异常行为检测模型局限于对异常行为的检测,而本文所提检测模型在检测异常行为的基础上,进一步提出了基于重构误差的异常定位方法。首先,对检测模型重构误差矩阵进行均值滤波,以减少场景扰动带来的干扰;接着进行阈值判定,只有像素点与其周围像素点的重构误差值均大于阈值时,才认定这个像素点是异常点,并通过热力图的形式表现出来。异常点的数量和聚集程度体现了异常发生区域的紧迫性。所提检测模型的异常区域定位功能能够进一步减轻监管人员的工作负担,有效提高检测模型的实用性。

3 试验分析

在 Subway、CUHK Avenue 及自建数据集上构建试验平台,对所提检测模型的准确性和运行时间等指标进行验证。试验平台的硬件配置选用 INTEL core i7 8750H 处理器、NVIDIA GTX 1080 显卡和 32 G 内存条。所提检测模型通过 Python 语言和 TensorFlow 深度学习框架实现,并进行异常行为检测,对照试验组采用经典的稀疏组合模型和 Hasan 模型(传统的卷积自编码器模型)进行异常行为检测。

3.1 Subway 数据集

Subway 数据集是由文献[6]于 2007 年提出并构建的基于以色列海法市某地铁站真实运营场景的异常行为检测数据集。该数据集记录有地铁入口和地铁出口 2 个具体的地铁场景下长达 139 min 的视频数据,包含走错方向、未付款、闲逛、人群不正常互动、急停、快跑和保洁作业等数十种类型的异常事件,很好地复现了地铁车站中较为常见的场景。Subway 数据集中的异常行为如图 2 所示。



a) 地铁入口翻越闸机



b) 地铁出口逆向通行

图 2 Subway 数据集中的异常行为

Fig. 2 Abnormal behaviors in Subway dataset

3.2 CUHK Avenue 数据集

CUHK Avenue 数据集中:训练集包含了 16 段香港大学校园监控视频片段的正常群体行为,共计 15 328 帧;测试集包含了 21 段香港大学校园监控视频片段的异常群体运动事件,共计 15 324 帧。该数据集主要用于验证不同背景环境和应用场景下,异常行为检测模型的鲁棒性和迁移能力。

3.3 自建数据集

为充分测试检测模型,本文选用 2018 年 11 月无锡地铁 1 号线和 2 号线的换乘站——三阳广场站 6 号口通道处、编号为 V0110081 监控摄像头的视频数据作为自建数据集。该监控场景包含商铺、电梯、立柱、栅栏、横幅和广告牌等,具有场景较为复杂、人流密集、人流方向不一及人流行为动作多变等特点,能够较好地反映国内地铁车站的真实监控场景。

自建数据集中:训练集共包含 30 个正常行为的视频片段,每个视频片段的时长约为 10 s(250 帧左右);测试集中共包含 10 个视频片段,每个视频片段的时长约为 20 s(500 帧左右),其中不仅包含正常行为,也包含了行人奔跑、人群异常停留及逆行等异常行为。

3.4 试验结果与对比分析

以 Subway 数据集的试验结果为例,由检测模型训练获得的 ROC(受试者工作特征)曲线如图 3 所示。ROC 曲线显示了选取不同检测阈值情况下 FPR(假正类率)和 TPR(真正类率)之间的变化关系。曲线越靠近左上角,表明模型的检测准确性越高、AUC(曲线下面积)值越大、EER(等错误率)值越小。经计算可知,Subway 数据集下验证检测模型的 AUC 值为 0.866,EER 值为 0.191。

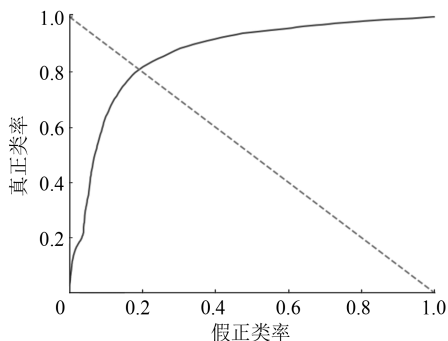


图 3 检测模型的 ROC 曲线

Fig. 3 ROC curve of the detection model

检测模型重构误差如图 4 所示。以 0.6 作为判别是否是异常行为的检测阈值,左侧灰色部分为正常行为所在帧,重构误差一直保持在较低水平,而右侧白色部分为异常行为所在帧,重构误差一直保持在较高水平,且不存在较大范围的误差波动。由图 3 和图 4 可知,所提检测模型具有较好的准确率 and 稳定性。异常行为检测结果的发生区域及其优先程度可以进一步通过热力图、异常点数量及聚集

程度来体现。Subway 数据集中的异常区域定位如图 5 所示,不同数据集下各模型的训练结果对比如表 1 所示。由表 1 可知,与稀疏组合模型和传统卷积自编码器 Hasan 模型相比,所提检测模型的查全率和准确性明显较高。

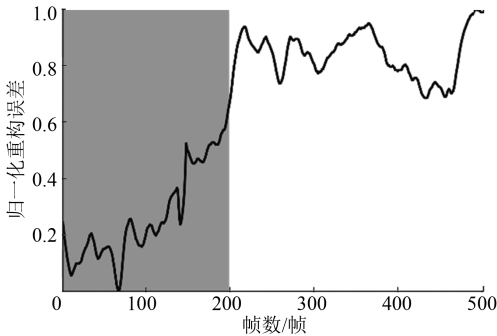


图 4 检测模型重构误差

Fig. 4 Reconstruction error of the detection model

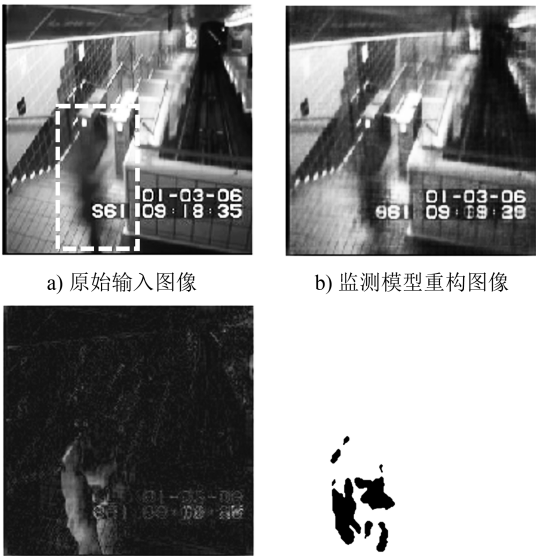


图 5 Subway 数据集中的异常区域定位

Fig. 5 Location of abnormal areas in Subway dataset

表 1 不同数据集下各模型的训练结果对比

Tab. 1 Comparison of training results of each model with different datasets														
数据集名称	总帧数/帧	异常事件数/件	正确报警			误报次数			AUC			EER		
			稀疏组合模型	Hasan模型	Conv-LSTM模型	稀疏组合模型	Hasan模型	Conv-LSTM模型	稀疏组合模型	Hasan模型	Conv-LSTM模型	稀疏组合模型	Hasan模型	Conv-LSTM模型
Subway	121 749	66	57		61	4		9	0.816	0.861	0.866	0.260	0.204	0.191
Avenue	15 324	47			45			6		0.702	0.838		0.251	0.270
自建数据集	5 000	21		18	19		5	3		0.843	0.909		0.231	0.169

在地铁站的实际应用场景中,异常行为检测模型的实用性还和时效性有着较大的关系。过慢的处理速度不仅会影响到突发事件的及时响应和处置速度,严重情况下还会导致后续二次事件的发生。在所构建的试验平台下,两种运行模式(GPU(图形处理器)和 CPU(中央处理器))下的图像平均检测时间如表 2 所示。由表 2 可知:在 GPU 运行模式下,所提检测模型基本可以保持 20 帧/s 以上的运行速度;在 CPU 运行模式下,所提检测模型依然可以保持 2 帧/s 左右的运行速度。该检测模型的检测速度基本能够符合地铁环境的实际应用需求。

表 2 两种运行模式下的图像平均检测时间

Tab. 2 Average examination time of images in two operation modes				
运行模式	预处理时间/s	重构时间/s	识别时间/s	总计时间/s
CPU	0.01	0.500	0.02	0.530
GPU	0.01	0.025	0.02	0.045

4 结语

本文针对地铁监控视频和车站内群体异常行为特征,基于半监督学习模型,提出一种准确、高效的群体异常行为检测和定位模型。相比于传统的异常行为检测方法,所提检测模型能够提取正常行为的空运动特征,通过对输入图像的重构进行异常行为的识别并定位,大幅提高了检测模型的查全率和检测效率,并在此基础上保证了其时效性,基本符合地铁应用场景下的实际应用需求,对地铁的安全管理和应急处置具有一定的应用价值。但目前,该检测模型还无法辨识出具体是哪种群体异常行为,这也是未来研究工作需要考虑和改进的地方。

参考文献

[1] 杨谱. 基于视频图像处理的地铁站拥挤风险预测与人群疏散 [D]. 北京:北京交通大学,2018.
YANG Pu. Risk prediction of crowd in subway station based on

video image processing and crowd evacuation[D]. Beijing:Beijing Jiaotong University, 2018.

- [2] 魏万旭,方勇,胡华,等. 基于视频数据挖掘的城市轨道交通车站行人交通行为特征提取系统研究[J]. 铁道运输与经济, 2021,43(8): 119.

WEI Wanxu, FANG Yong, HU Hua, et al. Study on characteristics extraction system of pedestrian traffic behavior for subway stations based on video data mining[J]. Railway Transport and Economy, 2021,43(8): 119.

- [3] CONG Y, YUAN J, LIU J. Sparse reconstruction cost for abnormal event detection[C]//2011 Computer Vision and Pattern Recognition (CVPR). Colorado Springs, CO: IEEE, 2011: 3449.

- [4] WANG C, LIU H. Unusual events detection based on multi-dictionary sparse representation using Kinect [C]//2013 IEEE International Conference on Image Processing. Melbourne, VIC: IEEE, 2014: 2968.

- [5] LI A, MIAO Z, CEN Y, et al. Abnormal event detection based on sparse reconstruction in crowded scenes[C]//2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Shanghai, China: IEEE, 2016: 1786.

- [6] ADAM A, RIVLIN E, SHIMSHONI I, et al. Robust real-time unusual event detection using multiple fixed-location monitors [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2008,30(3): 555.

(收稿日期:2020-11-21)

(上接第 11 页)

2) 加快现行标准的修订工作,以更好地满足科学技术发展水平和社会发展的需求。

3) 配套制定外文版标准。国家标准外文版是加强中外标准信息交流和促进对外合作高质量发展的重要支撑,因此应加快各外文版标准的编制工作。

4) 强化标准的实施应用。虽然现行的 3 个国家标准已获得了一定程度的推广应用,但实施效力仍需加强。因此,应进一步加强标准的应用示范工作,建设一些示范项目,既规范了标准的实施应用,又能在实践中进一步研究和完善标准体系,进而在落实标准的同时构建起标准验证技术支撑体系。

参考文献

- [1] 任海,陈燕申,秦国栋. 城市轨道交通的非传统安全威胁及技术防范[J]. 城市轨道交通研究, 2009,12(9): 1.

REN Hai, CHEN Yanshen, QIN Guodong. Non-traditional security threat and protective techniques in urban rail transit district [J]. Urban Mass Transit, 2009,12(9): 1.

- [2] 王伟奇,林鑫炜,任海. GB/T 38311—2019《城市轨道交通安全防范通信协议与接口》解读[J]. 城市轨道交通研究, 2021, 24(1): 84.

WANG Weiqi, LIN Xinwei, REN Hai. Interpretation of GB/T 38311—2019 communication protocols and interfaces of public security and protection system in urban rail transit[J]. Urban Mass Transit, 2021,24(1): 84.

- [3] GB 51151—2016 编制组. 城市轨道交通公共安全防范系统工程技术规范 GB 51151—2016 实施指南[M]. 北京:中国建筑工业出版社, 2020.

Compilation Committee of GB 51151—2016. Implementation guide for the Technical Code for Engineering of Public Security and Protection System of Urban Rail Transit (GB 51151—2016) [M]. Beijing: China Architecture and Building Press, 2020.

(收稿日期:2022-06-10)

(上接第 16 页)

WANG Deshan, XIE Jinrong. Calculation method of semi-trailer train steering characteristics [J]. Automobile Technology, 1990(10): 1.

- [2] 裴金. 半挂车复杂路线同轨迹转向的研究[J]. 专用汽车, 1992(2): 6.

PEI Jin. Study on same-trajectory steering of semi-trailer complex route[J]. Special Purpose Vehicle, 1992(2): 6.

- [3] 郑旺辉. 后轮转向半挂汽车列车的轨迹偏差分析[J]. 系统工程与电子技术, 1994,16(6): 72.

ZHENG Wanghui. Analysis of track bias of the automobile train with rear wheel steering[J]. Systems Engineering and Electron-

ics, 1994,16(6): 72.

- [4] 林熊熊. 半挂汽车列车弯路运动轨迹计算机仿真[J]. 汽车工程, 1997,19(1): 60.

LIN Xiongxiang. Computer simulation on trajectory of semi-trailer train on turning path [J]. Automotive Engineering, 1997, 19(1): 60.

- [5] 王佳栋. 汽车列车运动轨迹跟踪控制仿真研究[D]. 长春:吉林大学, 2005.

WANG Jiadong. The simulation study on the control of the moving track of truck-trailer [D]. Changchun: Jilin University, 2005.

(收稿日期:2020-11-12)