

自动驾驶合规性虚拟测试场景构建方法研究进展

徐家正, 陈星, 王景升*, 晏松

(中国人民公安大学 交通管理学院, 北京 100038)

摘要: 基于场景的测试推动了自动驾驶技术的开发与应用, 如何构建合规性虚拟测试场景已成为自动驾驶技术领域的关键问题。首先针对不同主体对于合规性的需求进行了详细分析, 基于场景交互性对测试场景构建方法进行分类, 回顾了一系列虚拟测试场景构建方法, 对各类场景构建方法从研究进展、应用情况和优缺点等方面进行归纳总结。此外, 为解决不同主体差异化测试需求的场景构建方法适用问题, 提出了覆盖度、边缘性、真实性三类测试场景构建方法评价指标, 并用三类指标对各类场景方法的优势和局限性进行了综合评估。最后, 基于评估结果详细分析了各类测试方法对差异化测试需求的适用性, 并对当前存在的不足和未来研究趋势进行了总结和分析, 以期自动驾驶测试场景构建研究提供有益参考。

关键词: 自动驾驶测试; 虚拟场景构建; 覆盖度; 边缘性; 真实性; 合规性测试

中图分类号: U467.5 **文献标志码:** A **文章编号:** 1673-0143(2024)02-0087-10

DOI: 10.16389/j.cnki.cn42-1737/n.2024.02.010

Research Progress on Automatic Driving Compliance-related Virtual Test Scene Construction Methods

XU Jiazheng, CHEN Xing, WANG Jingsheng*, YAN Song

(School of Traffic Management, People's Public Security University of China, Beijing 100038, China)

Abstract: Scenario-based testing has driven the development and application of autonomous driving technology. How to construct compliance-related virtual test scenarios has become a key issue in the field of autonomous driving technology. This article first analyzed the detailed requirements for compliance from different entities, classified the test scenario construction methods based on scenario interactivity, reviewed a series of virtual test scenario construction methods, and summarized them in terms of research progress, application status, advantages, and disadvantages. In addition, to address the applicability of scenario construction methods for different entities' differentiated testing needs, this article proposed evaluation indicators for three types of test scenario construction methods: coverage, edge cases, and realism. The advantages and limitations of various scenario

收稿日期: 2023-09-29

基金项目: 国家自然科学基金重点资助项目(6213000319); 中国人民公安大学基本科研新任教师启动基金资助项目(2022JKF434)

作者简介: 徐家正(1998—), 男, 硕士生, 研究方向: 智能交通、自动驾驶测试。

***通信作者:** 王景升(1970—), 男, 副教授, 硕士, 研究方向: 智能交通。E-mail: wjs1970@vip.163.com

methods were comprehensively evaluated using these indicators. Finally, based on the evaluation results, the applicability of various testing methods to differentiated testing needs was analyzed in detail. The shortcomings and future research trends were summarized and analyzed, aiming to provide a valuable reference for research on autonomous driving test scenario construction.

Key words: autonomous driving testing; virtual scene construction; coverage; edge cases; realism; compliance testing

0 引言

自动驾驶车辆(autonomous vehicles, AV)在减少交通事故、污染排放和出行拥堵等方面具有很好的发展前景,能够有效提高交通便利性及燃料利用率^[1]。但近年来特斯拉(Tesla)等自动驾驶公司发生的事事故^[2]证明,自动驾驶车辆在走向大规模商用前还面临许多严峻的挑战。美国高速公路安全管理局(NHTSA)于2021年7月至2023年5月共公开98起违规碰撞事故,其中就有高达11起导致重伤或死亡,未来每年也将发生数千起涉及L2级别以上的AV违规事故^[3]。为了预防类似事故的发生,自动驾驶车辆从实验室走向商业市场前必须进行科学完善的测试,如何模拟复杂多变的真实交通环境是自动驾驶测试面临的重要挑战,在此背景下,基于场景的虚拟测试对于该问题具有很大优势。

基于场景的虚拟测试通过重构真实交通环境构建测试场景,解决了开放道路测试中场景“无风险”占比高、随机且重复等问题,具有较好的安全性、经济性和可控性。目前,国内外学者针对虚拟测试场景构建已做了许多研究工作。徐向阳等^[4]对比分析了场景的定义及架构,总结归纳了现有虚拟场景构建体系,提出了关于场景构建工具链和方法链的研究展望。Riedmaier等^[5]认为基于场景的测试是一种很有前景的方法,并提出一种新的分类手段来比较现有虚拟场景构建方法。蒋拯民等^[6]围绕测试内容、测试方法、测试场景等主题开展前沿性研究,认为场景挖掘是AV测试的核心技术,并指出虚拟技术置信度不足的问题。

虚拟测试场景往往以测试系统安全能力边界为场景构建目标^[7],由于场景具有不可穷尽、无限丰富、极其复杂、难以预测等特点,导致场景很容易忽略合规性的测试,因而影响测试结果的置信度。尤其面对自动驾驶相关标准、法规、规则日益增长的复杂性和苛刻性,加强自动驾驶合规性测试研究势在必行。同时,现有测试由于存在测试需求与测试场景不适用的问题,因而影响测试效率及结果的可靠性,从而导致测试场景无法发挥其主要作用。因此,实现测试需求与测试场景构建方法的适用是一个很有挑战性且亟需解决的问题。

1 合规性测试需求分析

从自动驾驶测试应用的角度出发,可将测试主体分为汽车研发制造企业、政府机构及监管部门和第三方测试机构,明确测试目标是进行AV测试需求分析的关键,各类测试主体的合规性测试目标如下:

1) 汽车研发制造企业:开发、设计自动驾驶系统,根据市场发展及自身技术水平设计生产AV,进行内部测试验证,确保AV的合规性。

2) 政府机构及监管部门:制定标准法规以规范自动驾驶技术的开发应用,评估并验证车企生产AV的合规性,负责AV认证上路后的合规性监管。

3) 第三方测试机构:独立于车企及政府,以“测试者”的身份,对AV的性能及合规性进行全面评估,提供专业客观的测试结果及建议,验证AV能否达到预期的合规性。

不同测试主体开展自身业务的同时相互协作,通过数据共享、合作测试等方式共同促进自动驾驶测试技术的发展,现阶段各主体的虚拟测试需求如表1所示。

表1 不同测试主体需求分析

Tab. 1 Analysis of different test entities' requirements

测试主体	合规性虚拟测试需求
汽车研发制造企业	I-a: 模拟各种类型的道路、天气、环境条件
	I-b: 复现不同复杂程度的驾驶情形
	I-c: 支持进行大规模测试
	I-d: 确保测试符合真实遵守情况
政府机构及监管部门	II-a: 准确评估AV的合规能力范围
	II-b: 复现真实场景的高复杂与高交互性
	II-c: 评估法规政策的科学合理性
第三方测试机构	III-a: 发现潜在的违规危险
	III-b: 拓展更多测试方式,覆盖现行标准法规

2 虚拟测试场景构建方法分析

2.1 场景构建方法分类及定义

现阶段主要通过机理建模、数据驱动等方法构建虚拟场景,并倾向于构建有挑战性的边缘场景。交互程度是指被测车辆与场景中的动静态实体的交互关系,能很好地反映场景测试难度,进而实现针对不同等级AV的测试需求进行场景适用。因此,本文将场景构建方法分为高交互合规性测试、低交互合规性测试两类。低交互场景是根据需要预先设定好的测试场景,构建后不再变动;高交互场景中则存在多个智能参与者相互关联,涵盖更多可能的情况。

2.2 测试场景构建方法

测试场景能够验证AV在不同交通情形下的标准法规遵守情况,同时评估AV在应对不同交通情形下的行为合规性,重点检测AV在异常情况、紧急情况及各种危险情形下的应对能力,由于交通法规本身也存在一定的模糊性,因此对于在法规不明情况下的合理性也是其测试的目标之一。

2.2.1 低交互合规性测试 1) 标准法规

由相关标准、法规、评价规程等数据构建的场景称为标准法规场景,被用于AV的合规性测试,如国内外的ISO、E-NCAP、GB/T等标准法规中都构建此类场景,这些标准法规场景初步形成AV的合规性认证体系。国内一些学者开展过相关研究,如曹树星^[8]以自适应巡航控制系统(ACC)为研究对象分析了现行标准法规,基于《GB/T 20608—2006》中相关性能指标,历经逻辑场景设计、参数范围确定、具体场景批量生成三个步骤,分别在直道、弯道工况下构建目标识别和弯道适应能力测试的标准法规场景;周文帅^[9]考虑国内实际交通情况,在分析国内外自动紧急制动系统(AEB)测试标准法规的基础上,构建了在十字路口路况下车辆在左转弯时可能遇见的多种碰撞风险测试场景,并结合标准法规测试需求提出相应的测试方法,最后将所构建场景与法规中的直行场景结合,构建出AEB典型测试场景库。

标准法规场景基本覆盖AV的基本性能要求范围,但存在复杂性、边缘性低的局限性,由于该类场景基于特定工况和交通法规构建,而真实交通环境千变万化,因此难以覆盖各种情况,导

致无法确定AV的实际安全边界。

2) 组合方法

组合方法的核心是对被测目标的输入变量进行组合以构建场景。最初依据专家经验的分析结果进行变量的参数组合构建场景,现阶段基于本体论的组合逐渐流行。本体常被用于形式化表达领域知识,广泛应用于各种专家知识、标准指南等先验知识的存储和管理,相比于专家经验,基于本体论的组合能够以测试系统可读的格式储存、维护和更新,更具有普适性,因为该方法在知识储存方式上的优势保证了场景的可溯源性,也便于进行建模故障检查、进一步推断推理等行为,提高了测试场景的有效性。

Bagschik等^[10]将具有语义的信息与真实交通场景结合,使用基于本体的方法生成测试场景,该过程包括知识获取、知识建模和组合生成三个阶段。知识获取阶段确定如何组织基础设施的相关原则;知识建模阶段使用分层模型进行信息表示,最终以自然语言对实体进行聚类;组合生成阶段使用分层组合方式构建测试场景。本体提供了人类及计算机都能识别的知识库,可以基于形式化知识进行场景的自动创建,从而提高测试效率。还有学者经逻辑性分析后从技术性指标考虑进行组合构建场景,如舒红等^[11]针对特定的交通环境,通过分析主车与其他交通参与者的可能相对位置与驾驶行为组合确定复杂场景群,并使用PICT组合测试工具将主车可能的运动方向依次与干扰车的可能运动方向进行组合,设置约束条件构建出不同的组合场景群,最后通过筛选规则获得有测试价值的基础测试场景群,在此基础上可对各场景进行定量安全风险评估。本文对其整体流程进行梳理,场景流程如图1所示。

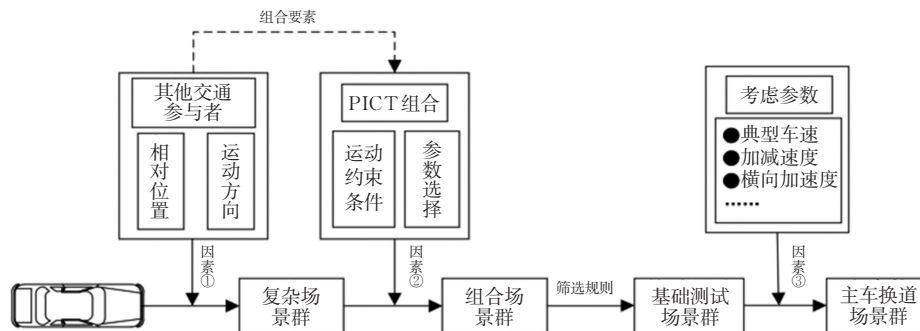


图1 PICT应用于构建测试场景流程

Fig. 1 Application of PICT in constructing test scenario processes

组合方法有效地解决了场景覆盖度低的问题,但未考虑交通参与者间的关系与交互,构建的模型宏观抽象,主要用于生成逻辑场景,有一定的局限性,并随着场景要素的增加,生成的场景数量呈指数级增长,易构建出无意义和低效率的冗余测试场景,同时该方法受限于专家知识框架,难以构建复杂的边缘场景。

2.2.2 高交互合规性测试 为了满足在复杂交通情况下的测试需求,现有研究采用聚类分析、重要性采样等方法来扩展测试场景库,不仅能够覆盖真实的交通环境,还在测试效率、真实性、覆盖度等方面存在优势。

1) 聚类分析

聚类分析能够利用自然驾驶数据、事故数据等结合合规性测试需求提取场景特征要素构建场景,并根据特征要素对数据集进行划分,从而构建典型高交互场景。现研究多从违规事故数据集中选取不同特征要素,经聚类分析得到不同类型的典型事故场景。如陈吉清等^[12]分析了641例与道路路段类型相关的真实违规事故数据,选取时间、路段类型、车辆驾驶动作、相对运动和交互对象类型五个要素作为聚类变量,通过对变量进行独热编码处理和聚类分析,构建了15个道路

路段类型相关的测试场景,采用的独热编码方法能有效处理非连续型数值特征。

还有部分学者根据我国实际情况,结合聚类分析进行合规性测试研究。胡林等^[13]针对我国在自动紧急制动(AEB)系统研究中缺少两轮车测试场景的问题,基于419例真实碰撞案例经聚类分析提取得到11类典型事故场景;并根据伤亡程度分析参数特征,获取不同场景下两轮车的运动状态、车速等详细参数,构建了在良好照明条件下汽车直行通过路口的主要测试场景。徐向阳等^[14]研究了适用于中国的AEB测试场景,基于多元Logistic回归分析违规事故影响程度并确定场景特征要素,采用层次聚类算法对测试场景进行聚类挖掘,基于499例真实事故,将场景特征要素作为聚类参数得到8类路口典型危险场景,并建立符合中国交通状况的AEB路口测试场景。

聚类分析可以充分利用真实数据构建更高交互程度的合规性测试场景,但聚类算法的选择对场景的分类结果影响较大,影响测试结果的可靠性。在数据集方面,该方法依赖数据集的真实性和完整性,且利用已有数据集构建的场景难以发掘AV的潜在危险特征,因此无法构建具有挑战难度的边缘场景。

2) 重要性采样(importance sampling, IS)

通过从自然驾驶数据中提取关键数据,可针对特定需求有效构建场景。根据重要性采样方法(IS)原理,通过生成新的采样分布代替原有自然分布来加速测试进程,可模拟背景车辆危险行为。该方法利用现有数据即可完成场景构建,节省时间成本,并且由于事件已经发生,构建场景具有高合理性和可追溯性。

Gelder等^[15]使用核密度估计(kernel density estimation)对真实场景参数分布进行拟合,利用蒙特卡罗模拟方法从分布中抽样,以确保评估的安全及合规水平与实际水平相对应,并使用IS生成新的关键测试场景。Wang等^[16]将IS的加速方法与可达性分析相结合,应用于构建人行横道的功能场景。Xu等^[17]提出了一种基于改进IS技术的AV测试方法,首先通过IS提取典型切入场景中的关键变量,对所提取变量进行统计拟合,并采用遗传算法计算IS中的最优参数,作者同时考虑到分布拟合可能存在误差,还采用误差修正参数对结果进行修正,以确保测试结果的准确性。本文对其整体流程进行梳理后的流程图如图2所示。

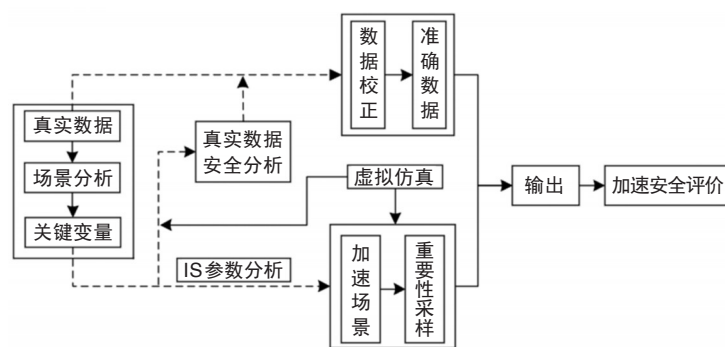


图2 基于改进的重要性采样测试方法流程图

Fig. 2 Flowchart of the improved importance sampling testing method

重要性采样方法有效地加速了AV测试进程,较好地提高了合规性测试场景的交互性,但其研究数据主要来源于安全合规情况,无法应对危险路况,构建场景质量也取决于数据质量及IS策略,具有很强的数据依赖性。此外,现阶段IS方法中构建的重要性函数都较简单,可能会导致合规性测试结果存在较大误差。

3) 深度学习(deep learning, DL)

深度学习具有强大的泛化、表达能力,能够处理AV测试中涉及的复杂、高维数据,如雷达数

据、GPS数据等,深度学习技术能够有效地发掘它们的复杂结构,并对数据进行正确的表达和表示。目前,深度学习模型对有限数据的集中泛化能力已被广泛应用于自动驾驶测试场景构建,常用的深度学习模型有循环神经网络(RNN)、生成对抗网络(GAN)等。例如,循环神经网络(RNN)模型可用于预测背景车辆运动轨迹,基于生成对抗网络(GAN)进行自我纠错可以构建更加符合真实交通环境的场景。

Demetriou等^[18]提出了一个基于真实数据驱动生成不同长度轨迹的深度学习框架,包括将轨迹长度作为循环条件使用RC-GAN生成更加灵活的驾驶轨迹,以及利用基于循环自编码器的GAN架构解决轨迹长度多变的问题,在此基础上,通过训练前馈神经网络来预测轨迹长度,并收集实际数据验证其性能,该方法可以灵活准确地生成不同轨迹长度的测试场景,步骤见图3。Jenkins等^[19]利用RNNs和LSTM单元对事故数据进行建模以构建潜在AV测试场景,由于是基于事故数据训练智能体,导致其与测试车辆交互时倾向于做出危险行为,并且场景可解释性较弱,同时受技术限制和收集数据成本的影响,现有的数据集只能提供有限的多车交互场景信息。此外,深度学习还可用于真实驾驶数据的分类,有利于构建更具针对性的场景。

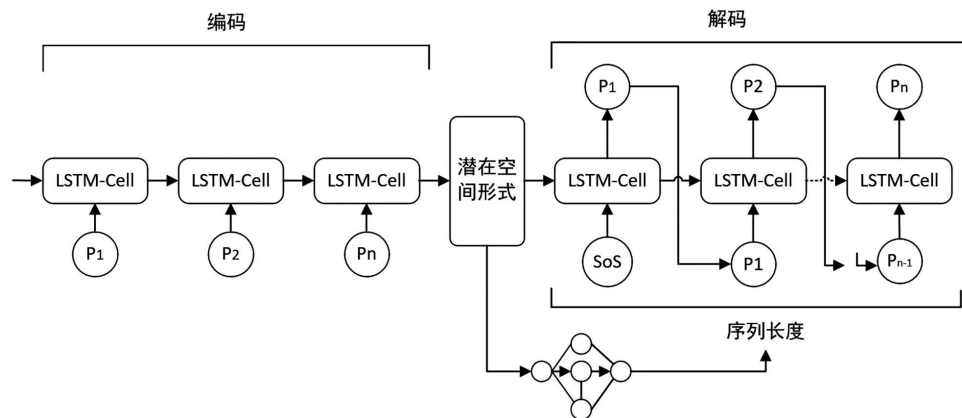


图3 含序列长度估计的循环自编码器结构示意图^[18]

Fig. 3 Illustration of a recurrent autoencoder structure with sequence length estimation^[18]

由于深度学习技术描述非训练车(NPC)的行为与描述训练数据一样困难,因此构建的测试场景可解释性较弱,可结合改进的场景要素提取方法弥补这些差距^[20]。此外,由于技术限制及成本问题,许多公开数据库^[21]并没有提供足够多的多车交互场景信息,而深度学习模型对高质量数据依赖性较强,导致所构建场景测试效果受限,因此为构建更准确的测试场景,需要不断增加训练数据集以覆盖所有可能的状态。

2.3 方法评价

本部分广泛地调查了虚拟测试场景构建方法相关文献及技术进展,按照不同的交互程度划分两类测试场景构建方法,并按分类对现有方法进行具体分析和总结。

1)基于标准法规的方法和组合方法是当前构建低交互合规性测试场景的主要手段,它们测试目的明确、场景组成简单,能够验证AV上路的基本要求,在测试初期使用,能够节省大量测试时间。

2)高交互性场景是为更全面、准确地评估AV综合合规性而构建的,具有较高的复杂性和多样性,适用于挖掘潜在的违规风险,能够针对性地构建更复杂、交通状况多变、涉及多决策的合规性测试场景。

3 场景构建方法评价及分析

自动驾驶测试评价相关标准成为各个国际标准化组织的重点工作方向^[22],但现有测试工具

及方法难以满足L3级及以上自动驾驶汽车测试需求,针对这一问题,本节以研究不同方法的适用性为目的进行方法评价分析。测试场景构建方法的评价角度是多元的,作为业内流行的评价指标,覆盖度与AV的鲁棒性、适应性关联;边缘性代表着各种异常突发或极端罕见的交通情形;真实性则保障了虚拟测试结果的意义,它们与AV的综合合规性能较为直接,更能证明AV的合规性。因此在评估过程中,选取覆盖度(coverage)、边缘性(edge)、真实性(fidelity)三个代表性指标。

3.1 评价指标介绍

1) 覆盖度:场景因要素种类、属性的不同而数量繁多,高覆盖度的测试场景指具有不同属性的实体种类多、数量多,能够全面满足不同被测主体的测试需求,合规性主要测试实体是AV与动态流量参与者的交互。因此,覆盖度主要体现在流量参与者种类和数量的多少。

2) 边缘性:具有开发探索能力的方法可以在现有范围之外生成一些边缘场景,用于发现AV更多潜在的问题,高边缘性场景一定是未来高水平AV产业化的必经之路。

3) 真实性:即场景的道路情况、交通组成、天气等方面与真实道路环境的相似程度,影响测试结果的可靠性和置信度。具有高真实度的测试场景与现实交通环境高度相似,具体指的是测试场景中交通参与者间因果关系的真实性,所构建的场景应贴近人类的认知。

3.2 构建方法分析

测试的需求是多元的,因此每类构建方法都存在研究价值。低交互性测试中的标准法规场景来源于现行标准法规,构建场景简单直观,常用于快速检测AV遵守法规能力。组合方法通过组合场景要素构建场景,有助于提高场景多样性,但易组合出“不真实”场景,因此在组合前宜进行相关分析。聚类分析是基于真实数据建立场景要素与条件之间的关系,而重要性采样是选取符合指定条件的要素以构建场景,两者均具有较好的真实性,测试结果置信度高。但前者易忽略难以被分类的场景,后者在边缘性方面欠缺考虑。深度学习在交互性方面表现较好,能让AV在场景中不断学习和优化,实现更加智能、自主性强的自动驾驶决策系统。为方便对各类方法进行评价,根据三个指标综合评估各类方法,同一类方法中的不同文献间也可能在标准上有所不同,依据其针对具体标准的相对体现程度,将其分为“高-•••”“中-••”“低-•”“无- /”四个档,最后根据选取的指标,结合前文讨论内容对各场景构建方法进行对比分析,得出具体评价结果如表2所示。

表2 各具体构建方法评价结果

Tab. 2 Evaluation results of each specific construction method

方法	文献	覆盖度	边缘性	真实性	优势	劣势																																																								
标准法规场景	[8]	•	/	•••	标准统一,提供更好的可比性和一致性	场景基础、覆盖度较低,缺乏创新																																																								
	[9]	•	/	•••			组合方法	[10]	••	/	••	覆盖度较高;可控性强	场景宏观抽象;需要大量计算资源进行数据处理、模型构建	[11]	•••	••	••	聚类分析	[12]	••	••	•••	代表性较强、真实性高;	数据依赖性强、场景较基础;	[13]	••	•	•••	多样性强,能够覆盖更广泛	不当的聚类参数选择影响有效性和真实性	[14]	••	•	•••	泛的驾驶情况		重要性采样	[15]	•••	•	•••	真实性高、合理性强;	依赖数据情况、采样函数质量;难以刻画复杂情形,复杂度低	[16]	••	/	•••	高效率,对数据进行加权	[17]	••	••	•••	抽样提高数据利用率	深度学习	[18]	••	•	•••	真实性高、交互性好;	依赖数据质量,可解释性差	[19]	•
组合方法	[10]	••	/	••	覆盖度较高;可控性强	场景宏观抽象;需要大量计算资源进行数据处理、模型构建																																																								
	[11]	•••	••	••			聚类分析	[12]	••	••	•••	代表性较强、真实性高;	数据依赖性强、场景较基础;	[13]	••	•	•••		多样性强,能够覆盖更广泛	不当的聚类参数选择影响有效性和真实性	[14]	••	•	•••	泛的驾驶情况		重要性采样	[15]	•••	•	•••	真实性高、合理性强;	依赖数据情况、采样函数质量;难以刻画复杂情形,复杂度低	[16]	••	/		•••	高效率,对数据进行加权	[17]	••	••		•••	抽样提高数据利用率	深度学习	[18]	••	•	•••	真实性高、交互性好;	依赖数据质量,可解释性差	[19]	•	••	••	能够加快测试速度					
聚类分析	[12]	••	••	•••	代表性较强、真实性高;	数据依赖性强、场景较基础;																																																								
	[13]	••	•	•••	多样性强,能够覆盖更广泛	不当的聚类参数选择影响有效性和真实性																																																								
	[14]	••	•	•••	泛的驾驶情况																																																									
重要性采样	[15]	•••	•	•••	真实性高、合理性强;	依赖数据情况、采样函数质量;难以刻画复杂情形,复杂度低																																																								
	[16]	••	/	•••	高效率,对数据进行加权																																																									
	[17]	••	••	•••	抽样提高数据利用率																																																									
深度学习	[18]	••	•	•••	真实性高、交互性好;	依赖数据质量,可解释性差																																																								
	[19]	•	••	••	能够加快测试速度																																																									

4 虚拟测试场景构建方法总结与启示

测试需求因测试主体不同存在差异,各虚拟场景构建方法的方式分类复杂,适用标准也不统一。针对这一问题,本文总结了各测试主体的合规性测试具体需求,并对其进行综合分析。本节针对每个主体的具体测试需求情况(见表1的I-a~III-b),选择合适的方法对相应需求进行分析,并基于方法的评价结果对虚拟测试场景构建方法现状及未来研究趋势进行总结。

4.1 现状总结

针对性的场景构建方法能较好地满足各主体测试需求,具体情况如下。

1) 汽车研发制造企业(开发):基于标准法规构建的场景可以快速检测AV的法规遵守情况,而基于组合的测试方法能够复现多种类型的道路、天气条件,这两种方法均支持企业进行大规模测试(见表1的I-a、I-c)。重要性采样从自然驾驶数据中挖掘的场景存在交通参与者间的安全关键交互,特别是一些转弯和危险行为,能够有效测试AV在高交互的复杂情形下的合规能力(见表1的I-b);基于深度学习的场景构建方法能够生成多样性行为和不确定性状态的背景车辆,使得AV在交互时发生事故,并且能在复杂多变的交通情形下验证AV的综合合规性能(见表1的I-a、I-b)。

2) 政府机构及监管部门(认证及监管):政府机构及监管部门应测试AV的行动路径能力,以判断它们是否能遵守现行标准法规。深度学习可以构建各种背景车辆的运动路径,准确测试它们的行动路径能力,评估AV的合规能力范围(见表1的II-a)。此外,基于组合、聚类的方法能够构建不同难度的合规性测试场景,二者能够协作全面评估AV的合规性,以判断企业所生产AV是否能在高复杂、交互情形下遵守现行标准法规(见表1的II-b)。

3) 第三方测试机构(“用户”测试):测试结果的可靠性和测试效率是测试机构的首要考虑因素,体现在场景的边缘性、覆盖度等指标中,可以通过深度学习、组合方法有效提高这些场景指标,包括生成动态和静态不规则障碍物、不规则道路,组合粗糙和陡峭的道路属性等,以发现潜在的违规风险(见表1的III-a)。此外,第三方测试机构能通过拓展虚拟测试方法,覆盖现行标准法规,全面评估AV的综合能力(见表1的III-b)。

4.2 未来研究趋势

与此同时,虚拟测试场景构建方法研究领域还存在一些不足与空白,需要未来着重进行研究。

1) 未来应继续提升测试场景的质量。真实性的提高需要兼顾数据与车辆动力学,测试过程中应充分利用高精度地图数据、传感器数据等信息,并引入多样性交通流及随机因素以增加模型的不确定与意外性,车辆动力学模型也要考虑惯性感受、风速气温等真实因素(见表1的I-d)。针对覆盖度不足的问题,与其他方法进行联合测试是一个有着较好前景的研究方向,通过场景片段的拼接、组合,能够有效提高测试场景的覆盖度(见表1的I-a)。以下研究有利于提高方法的边缘性,组合方法考虑组合要素的选取,如环境车辆的运行速度方向、状态、道路类型等;聚类分析可在数据选取阶段考虑危险道路、事故多发路段;重要性采样则通过研究IS函数构建偏向危险的测试场景(见表1的I-b、II-b)。随着真实性及边缘性的提高,测试场景能更准确地检测出AV潜在的违规风险,提高测试效率(见表1的III-a)。

2) 未来的场景构建方法可能更依赖于数据驱动方法。数值质量影响着聚类分析、重要性采样等数据驱动方法的效果,对于场景质量至关重要。此外,交互数据是测试场景生成的重点,然而由于没有专门的标准交互数据库,无法保证交互者间的统一比较。除了中国之外,其他国家的

数据也很重要,每个国家都有独特的交通环境特征,因此,有必要尽可能地整合更多国家的数据类型,以加快虚拟测试研究。

5 结语

测试需求在自动驾驶车辆的测试过程中起着至关重要的作用,针对不同需求的虚拟场景构建方法是必不可少的,但由于测试需求处于测试链的中间位置且复杂多样,使得测试场景的适用充满了挑战。本文通过总结近年来有关虚拟测试场景构建方法的代表性文献,从不同交互程度的角度将测试方法分为低交互性测试场景和高交互性测试场景。随后,提出三类指标分析比较了这些方法的优缺点。最后,在分析不同测试主体需求的基础上,提出了更适合各主体具体需求的测试场景构建方法,进一步总结该领域现阶段不足和未来的发展趋势。

参考文献 (References)

- [1] LI L, ZHENG L Y, WANG N N, et al. Artificial intelligence test: a case study of intelligent vehicles [J]. *Artificial Intelligence Review: An International Science and Engineering Journal*, 2018, 50(3):441-465.
- [2] BRODKIN J. Tesla on autopilot crashed into stopped truck during highway lane closure [EB/OL]. (2023-06-28)[2023-09-01]. <https://arstechnica.com/techpolicy/2023/06/tesla-on-autopilot-crashed-into-stopped-truck-during-highway-lane-closure/>.
- [3] STEENBURG J V. Incident reporting for automated driving systems and level 2 advanced driver assistance systems [EB/OL]. (2022-06-07)[2023-09-01]. <https://www.regulation.gov/document/NHTSA-2021-0070-001>.
- [4] 徐向阳,胡文浩,董红磊,等. 自动驾驶汽车测试场景构建关键技术综述[J]. *汽车工程*, 2021, 43(4): 610-619.
- [5] RIEDMAIER S, PONN T, LUDWING D, et al. Survey on scenario-based safety assessment of automated vehicles [J]. *IEEE Access*, 2020, 8: 87456-87477.
- [6] 蒋拯民,党少博,李慧云,等. 自动驾驶汽车场景测试研究进展综述[J]. *汽车技术*, 2022(8):10-22.
- [7] KOOPMAN P, WAGNER M. Challenges in autonomous vehicle testing and validation [J]. *SAE International Journal of Transportation Safety*, 2016, 4(1): 15-24.
- [8] 曹树星. 自动驾驶汽车测试场景库构建与应用研究[D]. 西安:长安大学,2022.
- [9] 周文帅. 基于典型测试场景的自动驾驶汽车紧急制动系统测试评价方法研究[D]. 西安:长安大学,2021.
- [10] BAGSCHIK G, MENZEL T, MAURER M. Ontology based scene creation for the development of automated vehicles [C]// 2018 IEEE Intelligent Vehicles Symposium. New York: IEEE, 2018: 1813-1820.
- [11] 舒红,袁康,修海林,等. 自动驾驶汽车基础测试场景构建研究[J]. *中国公路学报*, 2019, 32(11): 245-254.
- [12] 陈吉清,舒孝雄,兰凤崇,等. 典型危险事故特征的自动驾驶测试场景构建[J]. *华南理工大学学报(自然科学版)*, 2021, 49(5):1-8.
- [13] 胡林,易平,黄晶,等. 基于真实事故案例的自动紧急制动系统两轮车测试场景研究[J]. *汽车工程*, 2018, 40(12):1435-1446,1453.
- [14] 徐向阳,周兆辉,胡文浩,等. 基于事故数据挖掘的AEB路口测试场景[J]. *北京航空航天大学学报*, 2020, 46(10):1817-1825.
- [15] GELDER E D, PAARDEKOOPEP J P. Assessment of automated driving systems using real-life scenarios [C]//2017 IEEE Intelligent Vehicles Symposium (IV). DOI:10.1109/IVS.2017.7995782.

- [16] WANG X , PENG H , ZHAO D. Combining reachability analysis and importance sampling for accelerated evaluation of highway automated vehicles at pedestrian crossing [J]. ASME Letters in Dynamic Systems and Control, 2021, 1(1):11–17.
- [17] XU Y M, WANG J Q, ZOU Y T. Accelerated testing for automated vehicles safety evaluation in cut-in scenarios based on importance sampling, genetic algorithm and simulation applications [J]. Journal of Intelligent and Connected Vehicles, 2018, 1(2):28–38.
- [18] DEMETRIOU A, ALFSVG H, RAHROVANI S, et al. A deep learning framework for generation and analysis of driving scenario trajectories [J]. SN Computer Science, 2023, 4(3): 251–258.
- [19] JENKINS I R, KNAUSS A, YIN H, et al. Accident Scenario generation with recurrent neural networks [C]//2018 21st International Conference on Intelligent Transportation Systems (ITSC). IEEE, 2018:3340–3345.
- [20] 白雪松, 邓伟文, 任秉韬, 等. 一种自动驾驶仿真场景要素的提取方法[J]. 汽车工程, 2021, 43(7):1030–1036, 1065.
- [21] WANG W, ZHAO D. Extracting traffic primitives directly from naturalistically logged data for self-driving applications [J]. IEEE Robotics & Automation Letters, 2018, 3(2): 1223–1229.
- [22] 中华人民共和国工业和信息化部. 我国牵头的首个自动驾驶测试场景国际标准正式发布[EB/OL]. (2020–10–13)[2023–09–01]. https://www.miit.gov.cn/xwdt/gxdt/sjdt/art/2022/art_7a9f08fc0db14852a8a207ee5466dfb4.html.

(责任编辑:胡燕梅)