

DOI:10.19344/j.cnki.issn1671-5276.2024.04.052

自动驾驶仿真测试场景库体系建设方法综述

廖媛, 褚观耀, 程前

(中国汽车工程研究院股份有限公司, 重庆 401122)

摘要:基于自动驾驶仿真测试场景库可以有效加速测试进度,提高测试的安全性和覆盖度。对自动驾驶中的场景进行分析,论述了相关的场景定义、场景元素以及场景来源;对场景中的数据处理、特征提取、场景聚类进行描述,总结出自动驾驶测试仿真场景库的体系搭建方法,并对未来发展趋势作了相关预期分析,可以为自动驾驶测试开发提供参考和借鉴。

关键词:自动驾驶;仿真测试;场景库体系;综述

中图分类号:U461.91 **文献标志码:**A **文章编号:**1671-5276(2024)04-0271-05

Summary of Construction Method of Scene Library System Based on Automatic Driving Simulation Test

LIAO Yuan, CHU Guanyao, CHENG Qian

(China Automotive Engineering Research Institute Co., Ltd., Chongqing 401122, China)

Abstract:Since scene library based on automatic driving simulation test can effectively accelerate the test progress and improve the safety and coverage of the test, this paper analyzes the scene in automatic driving, discusses the relevant scene definition, scene elements and scene sources, describes the data processing, feature extraction and scene clustering in the scene, summarizes and proposes the system construction method of the simulation scene library of automatic driving test, and looks forward the trend of future development, which can be reference and lesson for the development of automatic driving test.

Keywords:autonomous driving; simulation test; scene library system; summarize

0 引言

自动驾驶技术发展日益精进,但其商业化落地需要经过大量的测试验证。传统的测试方法主要是利用搭载自动驾驶功能的实车进行道路测试。这种模式时间久、成本高、测试覆盖度低,并且受到较多的限制,无法进行极端边界方面的测试^[1]。由此可见,单纯依靠实车测试具有较大的局限性,不利于快速验证自动驾驶功能的可靠性。针对实车测试的不足之处,仿真测试可以在实车测试之前进行大量的虚拟测试,有效加速自动驾驶技术开发,这使得仿真测试技术日益受到业界的关注和重视,国内外科技公司、车企和高校均积极开展仿真测试场景方面的研究^[2]。

从目前仿真测试发展趋势来看,仿真测试逐渐向以下几个方面发展。1)基于云平台的高并发测试、加速测试。云平台具备海量数据存储、处理和管理的能,支持大规模仿真构建,覆盖海量驾驶场景;可有效建立城市级地理信息+虚拟交通流

信息,可实现跨专业多用户信息交互与共享。2)功能安全场景库、V2X 场景库以及预期功能安全场景库成为场景库建设的重点内容。3)不同厂商采用统一数据格式标准,共建基础场景库,形成通用的、可移植的模式。4)随着车辆网联化程度持续渗透,信息安全方向的验证测试与评价也必然是自动驾驶仿真测试的一个重要模块。

自动驾驶场景中人-车-路-环境之间的相互作用可以从时间、空间的维度来分析,这是一个复杂的动态关系模型,而基于这种关系模型形成的场景集是自动驾驶仿真测试开发的基础。但是,现实世界中的场景具有无限丰富、极其复杂、不可预测等特点,在虚拟环境中要完全复现这些场景是十分困难的。如何利用有限的测试场景去映射出无限丰富的世界是实现自动驾驶测试验证的关键。基于场景库的仿真测试是目前解决自动驾驶路测数据匮乏的重要路线。场景库中的测试场景对现实世界的覆盖率越高,仿真测试结果的准确性就越高。

第一作者简介:廖媛(1987—),男,重庆人,本科,研究方向为汽车主动安全、汽车整车性能测试、汽车整车主观评价等,486346459@qq.com。

为了解决传统道路测试的一系列局限性问 题,本文基于国内外自动驾驶仿真测试研究状况,在第 1 节中针对场景作为自动驾驶仿真测试场景库的主体进行了多维度的描述,介绍了场景定义、场景元素以及场景的数据来源等内容;在第 2 节中梳理了常见的数据处理方式,包括数据预处理、场景特征提取、场景聚类分析等;在第 3 节中总结归纳了自动驾驶仿真测试场景库体系的搭建方法、分布式架构以及基于场景的 V 字开发模型;在第 4 节中针对目前自动驾驶仿真测试场景库的开发提出了研究方向及展望。

1 自动驾驶仿真测试场景

1.1 场景定义

目前在自动驾驶领域场景的定义尚未统一,ULBRICH 等^[3]认为场景是按照时间序列进行发展的事件集合,全部场景均在同一个初始化时间序列的环境中进行。GO 等^[4]认为场景是由参与者、背景信息和行动或事件顺序构成的时间序列,简而言之,任何领域的场景要素都是一样的,但是场景的使用却大不相同。MENZEL 等^[5]将自动驾驶场景分为 3 类,即功能场景、逻辑场景、具体场景。朱冰等^[6]认为场景的本质是自动驾驶车辆在每个行驶时间戳与环境的动态组合。

上述研究中对场景的核心认知具有一致性,即场景包含测试车辆、其他交通参与者以及车辆环境,并且相互之间具有交互行为。由此可见,场景是时间序列中测试车辆和周围环境各元素动态组合的描述,其具有无限丰富、难以预测、不可穷尽等特点。

1.2 场景元素

对于场景元素的内容和类型,各研究主体提出的定义也有所差异。SAUERBIER 等^[7]认为场景元素包括测试车辆、交通环境元素、驾驶任务信息和特定驾驶行为。GEYER 等^[8]提出场景元素包括预定驾驶任务、静态场景元素和动态场景元素。GROH 等^[9]将场景元素分为静态元素、动态元素和环境元素。

自动驾驶测试车辆在行驶过程中会对周围场景元素产生影响,同样,周围场景元素也会反过来影响测试车辆,即车辆本身属性对场景构成具有重大影响。对此,本文梳理了图 1 所示的场景元素分类框架,将场景元素概括为两大类:车辆和环境。其中,环境包括气象、静态环境、动态环境和交通参与者;

车辆包括静态属性、动态属性、驾驶任务。

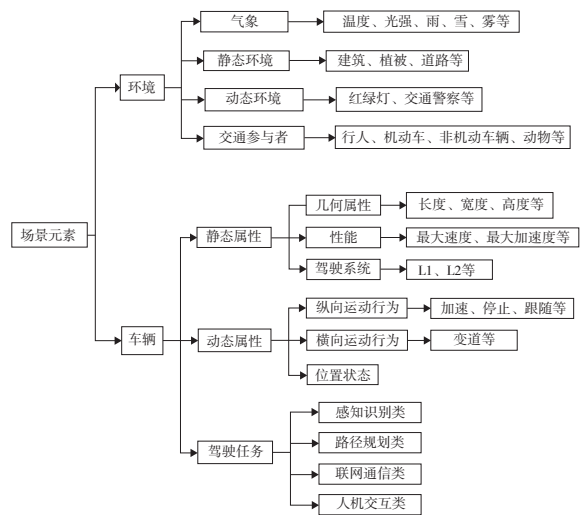


图 1 场景元素分类架构

1.3 场景数据来源

目前国内外已有部分机构收集了相关数据,形成了仿真测试场景库。美国 NHTSA、德国 KITTI、国内的腾讯 TAD Sim、百度 Apollo 等都为自动驾驶测试开发了专门的场景数据收集系统,为仿真场景搭建提供数据支撑。仿真测试场景就数据来源来说,主要有真实数据和仿真数据^[10],如图 2 所示。

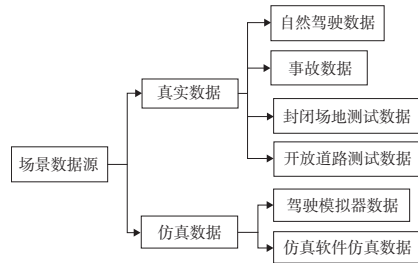


图 2 场景数据来源分类

1) 真实数据

自然驾驶数据、事故数据以及测试数据都可以作为真实数据来源。自然驾驶数据和事故数据作为真实数据的最常见形式,在研究中占有较大的比重。

自然驾驶数据一般是通过车载设备采集得到的^[11],比如摄像头、导航设备、雷达等传感器在车辆正常行驶时采集数据。常见的自动驾驶数据采集环境包括城市道路、高速公路等。

事故数据一般是从国家交通事故中进行人为分析筛选得到的。目前许多国家和相关机构都形成了一些交通事故数据库,比如美国国家公路交

通安全管理局的 GES 数据库、德国的 GIDAS 深度事故数据库、国内中汽技术研究中心的 CIDAS 事故数据库。

2) 仿真数据

仿真数据主要是指驾驶模拟器和软件仿真形成的数据。

驾驶模拟器可以在仿真软件中由真实测试人员驾驶来进行数据采集,既有安全性又能保证一定的真实性。仿真软件则是人为设定驾驶任务,让虚拟测试车辆在仿真环境下行驶,从而收集相关数据。仿真环境建立的核心是交通流建模和周围静态道路建模。

2 场景数据处理

要使用不同来源的数据进行仿真场景库建设的前提就是进行数据处理,场景数据处理的关键是提取体现目标场景的特征元素。

德国 PEGASUS 项目中形成了自动驾驶场景仿真数据处理的完整流程^[12]:数据生成—数据格式检查—信息标注—场景关联度分析—场景发生概率分析—数字化聚类分析—生成测试场景。

百度 Apollo 对场景聚类方法进行分析,提出了场景数据清洗、场景元素分析、数字化聚类等操作^[13]。基于国内外对自动驾驶场景数据的处理方法,可以得出数据处理过程主要包括数据预处理、特征提取、聚类分析。

2.1 数据预处理

不同数据来源的数据自然有格式、单位上的差异,并且原始数据也包括一些无效、错位的数据。因此有必要对传感器收集的数据进行清洗。

数据清洗主要是对冗余、缺失、异常的数据进行数据修复,目前的数据清洗一般是算法加人工辅助的方式,数据清洗应当在满足数据要求质量的前提下尽可能降低清洗代价。设 $C_{ost}(x)$ 和 $C_{ost}(s)$ 分别表示单个数据元组和整个数据集的清洗代价,其数据表达式如下:

$$C_{ost}(x) = \varphi(x) \sum_{i=1}^n D(t_i, t'_i) \quad (1)$$

$$C_{ost}(s) = \sum_{x \in s} C_{ost}(x) \quad (2)$$

式中: x 为分组后的单个数据元组; t_i 是数据要素; t'_i 是修复后数据要素; D 是 t_i 与 t'_i 之间的距离; s 是所有数据元组总和; $\varphi(x)$ 是 x 与全部数据元组总和的比。

2.2 特征提取

清洗后的数据要通过特征提取才能对场景进行解耦。文献[14]基于特征变量的统计分布,选择某些场景组成要素,比如道路类型、车速等。文献[15]选择光照、道路类型等特征要素。目前对场景特征元素的提取都是根据主观需求来确定,缺乏客观依据。文献[16]根据事故场景选择对事故有影响的要素进行多元 Logistics 回归模型,其模型如下。

$$\ln \left[\frac{P(Y = i | X)}{P(Y = A | X)} \right] = \alpha_i + \sum_{j=1}^N \beta_{ij} \dot{X}_j \quad (3)$$

式中: Y 表示事故场景的影响程度,其取值为 $i = 1, 2, 3, \dots, A-1$; α 和 β 分别为回归模型的截距和系数; X 为影响事故程度的变量, N 为变量 X 的数量。

2.3 聚类分析

由于各个特征要素之间单位各不相同,因此在聚类分析前需要进行特征变量设置^[17]。根据特征变量的表现形式可以进行静态和动态处理,静态处理用常量表示状态,比如天气的晴和雨可以分别用 0 和 1 表示;动态处理即元素状态时刻发生改变,无法以常数表示,比如测试车速和目标速度采用极差标准化来处理,可以使得动态变量取值在 0~1 之间,标准化后的变量 z_{xy} 的计算表达式如下^[18]。

$$z_{xy} = \frac{z_{xy} - \min_{1 \leq x \leq i} z_{xy}}{\max_{1 \leq x \leq i} z_{xy} - \min_{1 \leq x \leq i} z_{xy}} \quad (4)$$

式中: i 表示数据样本数量; y 表示单个数据样本特征变量个数。

对场景特征进行数字化后进行聚类分析。聚类分析的核心就是将相似度较高的对象或者元素分为一类,相似度较低的元素进行拆分^[19]。聚类算法的主要流程如图 3 所示。

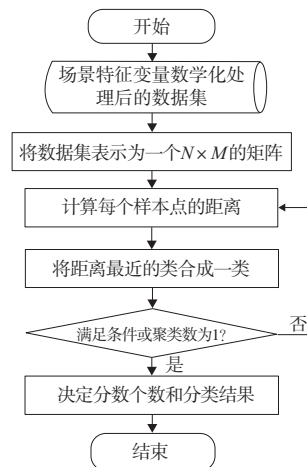


图 3 层次聚类算法的流程

3 仿真场景库建设

仿真场景库建设应当满足自动驾驶各个级别(L0—L5)测试需求^[20],并符合以下要求。1)覆盖面广,包含自然驾驶场景、标准法规场景、功能安全场景、危险场景、参数重组场景等场景,最大程度覆盖现实世界可能出现的驾驶场景。2)扩展性强,针对指定场景要素自动进行场景泛化处理,批量化生成测试用例。3)场景数量多,场景数量规模超过百万级。4)兼容性好,支持各大主流仿真平台和软件。

3.1 场景库体系搭建

自动驾驶场景搭建流程如图4所示,整个搭建过程由数据层贯穿到场景层^[21]。

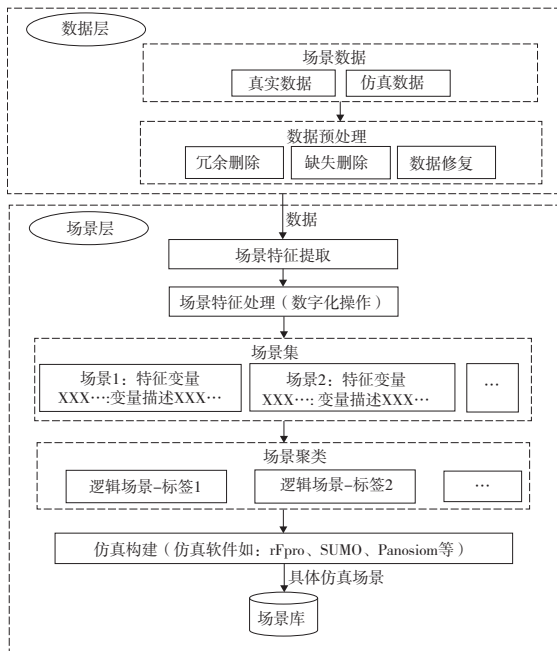


图4 仿真场景库搭建

1)数据层通过实车测试、仿真软件采集构建场景需要的数据源,然后对数据进行冗余、确实、异常等预处理操作,最后将数据格式和参数统一化。

2)场景层将接受到的数据层数据进行特征提取、数字化设置形成场景集,然后基于聚类处理形成多种逻辑场景,最后由仿真软件构建大量的具体场景形成场景库。

3.2 分布式架构

从数据—场景—场景库的流程来看,整个数据系统要满足数据处理、场景生成等要求^[22],这些核心功能的实现需要依赖海量存储和并行计算

技术,单节点无法承担海量数据和高性能并发的需求,需要将场景库设计为多节点共分担任务的架构。这就需要采用分布式架构,降低单个服务器的压力,实现低时延、高可靠,其整体架构设计如图5所示。

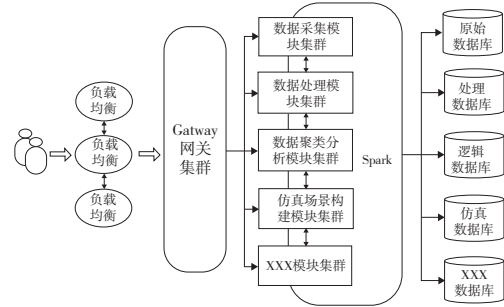


图5 分布式架构示意

3.3 基于场景库的V字开发模型

自动驾驶技术的演进使得测试场景变得无限丰富,在道路测试中无法做到全覆盖,如图6所示,基于场景库的开发模型具有更好的优势^[23],其涵盖虚拟测试和实车测试。虚拟测试包括模型在环测试(MIL)、驾驶员在环测试(DIL)、硬件在环测试(HIL)、车辆在环测试(VIL),实车测试包括封闭和开放道路测试^[24]。

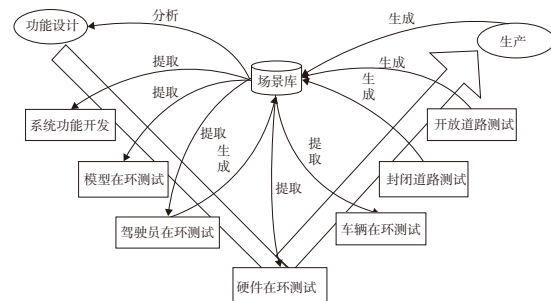


图6 基于场景库的V字开发模型

4 结论及其展望

面对自动驾驶技术的迭代演进,自动驾驶测试面临诸多挑战,开发仿真测试场景库是自动驾驶商业化落地的有效基石。目前国内的研究人员也对仿真测试做了大量的研究,许多国内公司正在逐步开发并完善自动驾驶仿真测试场景库,以期应对未来自动驾驶测试需求。从当前来看,仿真测试场景库开发具有潜在的挖掘研究空间,可以从以下几个方面着手。

1)开发国产仿真软件。仿真技术是场景库开发的重要环节,直接对仿真逼真度产生影响,但目

前场景仿真技术难度大、成本较高,缺乏对中国典型场景的仿真测试,国外仿真软件无法覆盖中国道路场景特色,未来可以基于国内仿真需求,开发中国特色的仿真软件。

2) 覆盖边缘极限场景。目前对正常驾驶场景的仿真较为常见,可以满足基本的测试需求,但由于边界场景数据源缺乏且构建困难,边缘极限场景丰富度尚有欠缺,而此类场景体现了整个仿真场景库的完整度。需要针对此挑战,以构建边界场景为目标,衍生开发多种边界场景。

3) 提高场景生成和提取效率。由数据库到场景库的过程核心是实现场景库的生成和测试过程的场景提取,由于场景生成过程较为复杂并且数据库的规模庞大,导致其耗时久、效率低,为了进一步提高测试效率,未来可以借助分布式加速引擎开发,加速场景生成和提取过程。

参考文献:

- [1] 马依宁,姜为,吴靖宇,等. 基于不同风格行驶模型的自动驾驶仿真测试自演绎场景研究[J]. 中国公路学报,2023,36(2):216-228.
- [2] 李平飞,金思雨,胡文浩,等. 用于自动驾驶仿真测试的车-车事故场景复杂度评价[J]. 汽车安全与节能学报,2022,13(4):697-704
- [3] ULBRICH S, MENZEL T, RESCHKA A, et al. Defining and substantiating the terms scene, situation, and scenario for automated driving [C]//2015 IEEE 18th International Conference on Intelligent Transportation Systems. Gran Canaria, Spain:IEEE,2015:982-988.
- [4] GO K, CARROLL J M. The blind men and the elephant[J]. Interactions,2004,11(6):44-53.
- [5] MENZEL T, BAGSCHIK G, MAURER M. Scenarios for development, test and validation of automated vehicles [C]//2018 IEEE Intelligent Vehicles Symposium (IV). Changshu, China:IEEE,2018:1821-1827.
- [6] 朱冰,张培兴,赵健,等. 基于场景的自动驾驶汽车虚拟测试研究进展[J]. 中国公路学报,2019,32(6):1-19.
- [7] SAUERBIER J, BOCK J, WEBER H, et al. Definition of scenarios for safety validation of automated driving functions[J]. ATZ Worldwide,2019,121(1):42-45.
- [8] GEYER S, BALTZER M, FRANZ B, et al. Concept and development of a unified ontology for generating test and use-case catalogues for assisted and automated vehicle guidance[J]. IET Intelligent Transport Systems,2014,8(3):183-189.
- [9] GROH K, KUEHBECK T, FLEISCHMANN B, et al. Towards a scenario-based assessment method for highly automated driving functions[Z]. [S.I.:s.n.],2017.
- [10] BATSCH F, KANARACHOS S, CHEAH M, et al. A taxonomy of validation strategies to ensure the safe operation of highly automated vehicles[J]. Journal of Intelligent Transportation Systems,2022,26(1):14-33.
- [11] 朱西产,刘智超,李霖. 基于自然驾驶数据的驾驶员紧急变道行为开环模型[J]. 汽车安全与节能学报,2015,6(4):328-332.
- [12] PARK M K, LEE S Y, KWON C K, et al. Design of pedestrian target selection with funnel map for pedestrian AEB system [J]. IEEE Transactions on Vehicular Technology,2017,66(5):3597-3609.
- [13] 刘杰. Apollo 开放平台的创新实践[J]. 软件和集成电路,2018(7):38-39.
- [14] 苏江平,陈君毅,王宏雁,等. 基于中国危险工况的行人交通冲突典型场景提取与分析[J]. 交通与运输,2017,33(增刊1):209-214.
- [15] 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.
- [16] 徐向阳,周兆辉,胡文浩,. 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.
- [17] 康诚,严欣,唐晓峰. 智能网联汽车自动驾驶仿真测试技术研究综述[J]. 时代汽车,2022(22):4-6.
- [18] 蒋燕. 自动驾驶仿真测试的关键场景生成工具设计与实现[D]. 南宁:广西大学,2022.
- [19] 鄢继仁. 面向自动驾驶仿真测试的动态场景描述语言设计与实现[D]. 南宁:广西大学,2022.
- [20] 罗钟林. 基于实时交通流映射的实车在环智能网联汽车仿真系统研究[D]. 成都:电子科技大学,2022.
- [21] 上官伟,李鑫,曹越,等. 基于虚实结合的自动驾驶仿真测试技术与应用[C]//第十六届中国智能交通年会科技论文集.长沙:[s.n.],2021:490-499.
- [22] 曾德松,高琛,谭北海,等. 自动驾驶仿真测试场景库构建研究[J]. 现代计算机,2021,27(31):1-12,32.
- [23] 王润民,赵祥模,徐志刚,等. 一种自动驾驶整车在环虚拟仿真测试平台设计[J]. 汽车技术,2022(4):1-7.
- [24] 王凤娇,宗岩,靳志刚,等. 自动驾驶仿真测试场景设计[J]. 科学技术创新,2020(29):100-101.

收稿日期:2023-02-09