

基于改进组合测试的自动驾驶仿真测试场景生成方法研究

欧洋辰, 陈凌珊

上海工程技术大学, 机械与汽车工程学院, 上海

收稿日期: 2022年12月7日; 录用日期: 2023年3月3日; 发布日期: 2023年3月10日

摘要

基于标准法规场景生成的自动驾驶虚拟仿真场景及功能场景库, 可以针对不同的ADAS功能进行仿真验证。本文选择功能场景的设计运行域(ODD), 并对ODD中的元素离散化取值计算其权重, 改进一种基于参数权重的组合测试方法, 针对ODD生成自动驾驶虚拟仿真测试场景。通过该场景应用于模型在环(MIL)测试平台上的自动紧急制动(AEB)系统, 验证了这种场景生成方法的有效性。

关键词

自动驾驶虚拟仿真, 主观赋权, 客观赋权, 组合测试, 测试用例生成

Research on the Scenario Generation Method of Autonomous Driving Simulation Test Based on Improved Combination Test

Yangchen Ou, Lingshan Chen

College of Mechanical and Automotive Engineering, Shanghai University of Engineering Science, Shanghai

Received: Dec. 7th, 2022; accepted: Mar. 3rd, 2023; published: Mar. 10th, 2023

Abstract

The virtual simulation scenarios and functional scenario libraries of autonomous driving generated based on standard regulatory scenarios can be simulated for different ADAS functions. In this paper, the design run domain (ODD) of the functional scenario is selected and the weights are calculated for the discretization values of elements in ODD. This paper improves a combined test

method based on parameter weights to generate virtual simulation test scenarios for autonomous driving for ODD. By applying this scenario to the automatic emergency braking (AEB) system on the model-in-the-loop (MIL) test platform, the effectiveness of the scenario generation method is verified.

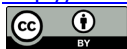
Keywords

Virtual Simulation of Autonomous Driving, Subjective Empowerment, Objective Empowerment, Combined Testing, Test Case Generation

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

随着自动驾驶汽车的商业化, 驾驶安全性变的越来越重要。为了保证安全, 自动驾驶汽车需要进行大量的测试, 在测试的早期阶段, 传统的封闭场地测试与真实道路测试存在测试风险性大、测试成本高、测试效率低的问题, 因此基于虚拟仿真技术的自动驾驶虚拟仿真测试已经成为目前市面上大多数汽车厂商选择的自动驾驶测试方法[1]。

虚拟仿真测试通过仿真软件进行物理建模与数值仿真, 通过较为真实的渲染技术构建与现实场景一致的仿真场景, 包括道路、天气、光照以及交通参与者等信息。虚拟仿真测试的基础是场景的生成, 即如何获得大批量的场景用于测试。仿真场景可以看作对现实的道路交通环境的再现, 而场景可以理解为具有特定任务的车辆行驶时所处环境的有限抽象与充分映射[2]。仿真场景生成的方法主要分为基于机理建模的生成方法, 即通过标准法规或专家经验, 参考实车测试的场景生成仿真场景, 以及基于数据驱动的场景自动生成方法, 即通过真实数据采集并转化生成的仿真场景和通过事故场景数据分析进行危险场景衍生研究。

参考软件测试领域, 系统的故障一般由其自身的因素及因素间的相互作用引起, 因此需要对因素间的组合进行全覆盖测试, 这就是组合测试, 又称 CT 测试。根据研究表明, 单一系统因素引发的软件故障约占 30%~40%, 由两个系统因素引发的故障约占 40%~50%, 由此 70%~80%的软件故障是由单一或两个参数的组合作用引起的[3]。

Zhang 等人研究了 ADAS 功能中的 LDW, 对其进行测试后发现, 被测功能系统的故障均是由道路形状、车道线模糊与缺失情况、光线的变化情况、天气情况等单一因素导致[4]。高锋等测试了车身控制器系统的功能, 发现系统共产生 21 个系统缺陷, 其中由单个系统参数的特定取值引发的有 16 个, 由 3 个系统参数的不同取值导致的系统缺陷有 5 个[5]。综上所述, 基于 ADAS 功能的自动驾驶虚拟仿真测试与软件测试具有很高的相似之处, 所以在自动驾驶仿真测试领域的测试用例生成方法, 可以参考软件测试领域中的 CT 方法。场景组合测试生成方法可以覆盖现实中更多的场景, 但会导致生成的场景数量过多, 测试效率降低。如何保证在覆盖度高的基础上, 有效的减少场景生成的数量, 成为了目前研究的重点。

论文的结构如下, 本文第二部分分析了 AEB 功能的 ODD 并对其元素进行了权重计算。第三部分介绍了改进组合测试算法的基本原理。第四部分通过仿真测试验证了该算法的有效性与优势, 最后一部分对全文进行了总结。

2. 功能场景 ODD 及其元素权重计算

自动驾驶虚拟仿真场景的来源包括参考标准法规场景构建、基于真实道路采集场景转化和基于事故场景重构等, 而仿真测试按照其性质可以分为功能测试、场景测试和任务测试, 其中功能测试是注重自动驾驶的单项或多项功能的实现, 是先确定待测功能再设计测试场景的测试方式, 为实现功能测试需要生成功能场景, 而功能场景是根据 ADAS 标准法规建立的仿真场景。本节选择 AEB 功能, 参考法规确定该功能的 ODD 元素及其取值, 并采用赋权方法计算元素的权重。

2.1. 功能场景 ODD 及其元素选择

ODD (Operational Design Domain)指运行设计域, 即设计时确定的驾驶自动化功能的外部环境条件。不同的汽车厂商会根据自身对于自动驾驶系统商业化应用、以及自身技术能力的情况, 对系统的 ODD 进行个性化的定义。仿真场景应从预期的 ODD 中选择, 进一步考虑所有场景序列对 ODD 的覆盖度。

《自动驾驶系统设计运行条件白皮书》中对 ODD 的主流构建框架有较为全面的整理与介绍。其中德国的 PEGASUS 项目确定场景中 ODD 有关的元素共有六层, 其中第一层道路层, 主要包括道路几何结构、路面质量、路面边界; 第二层基础设施层, 包括结构边界、标志牌、信号灯; 第三层暂时性操作层, 包括道路上临时搭建的设施, 例如因雨雪天气造成的临时封路、道路改造现场等; 第四层目标物, 包括道路参与者的自身状态、行为等; 第五层自然环境层, 包括天气、光照等其他环境信息; 第六层数字信息层, 包括 V2X 信息、数字地图信息等。而 SAE 工业技术协会建立的 ODD 通用术语与框架则由七个维度构成, 包括天气相关的维度条件、道路表面条件、道路设施、操作限制、道路使用者、非静止的道路目标以及连接性。

ASAM 的 openODD 是 openX 系列中的关于设计运行范围的安全标准。其中将 ODD 分为三类: 场景、环境和动态元素。其中场景包括区域、可行驶区域、路口、特殊结构、固定道路结构和临时道路结构; 环境包括天气、颗粒物、照明条件、连接性等; 动态元素包括交通和自身车辆。

由 PEGASUS 的分层模型及文中其他机构对 ODD 元素的定义可以得出, 考虑覆盖度这一评价指标, 应适当扩展传统 ODD 指标的参考范围, 不光要考虑车辆运行的外部环境条件, 同时也要考虑车辆自身及其他交通参与者的动态信息。

基于 ADAS 功能建立的功能场景需要满足其测试标准法规。欧盟新车安全评定委员会(Euro-NCAP)建立了 ADAS 功能中的 AEB 系统的测试评价标准, 该标准将 AEB 测试分为两种: 目标车为汽车以及目标车为弱势道路使用者, 包括行人、自动车, 非机动车等, 该标准还规定了 AEB 功能的测试方法和评价标准。根据我国国家标准, 乘用车自动紧急制动系统(AEB)性能要求及实验方法中的规定, AEB 测试条件应为: 在道路水平、路面干燥、具有良好附着能力的沥青路面或混凝土上进行, 附着系数应该为 0.8 以上; 试验时外部环境温度应为 0℃~45℃; 在整个试验中, 在水平可视范围内能够观察到目标, 因此能见度宜为 500 m 以上; 试验时风速不大于 5 m/s。

综上所述, 本文选择 AEB 场景, 参考测试标准法规并适当扩展环境条件, 将 ODD 分为环境条件与车辆运行状态两部分。其中环境部分包括道路类型、道路表面、道路形状、车道数量、道路坡度、道路使用者、天气、时间以及交通标志; 车辆运行状态包括自车车速、目标车车速、目标行人速度、目标二轮车速度以及两车相对距离, 并赋予每个元素不同的取值。其中车辆运行状态中的车速均为场景运行时的初始速度。基于 AEB 功能的 ODD 元素选取如图 1 所示。

ODD 中有些元素的取值具有连续、无界等性质, 使得无法对元素进行权重计算, 因此对于这些元素, 应对其取值进行离散化处理。以自车车速为例, 根据我国的道路法规可知, 自车车速取值范围为[0 km/h,

120 km/h], 同时参考 AEB 测试法规规定, AEB 测试自车车速取值范围主要在[20 km/h, 60 km/h]内, 因此离散化处理后得到的自车车速取值的如图 2 所示:

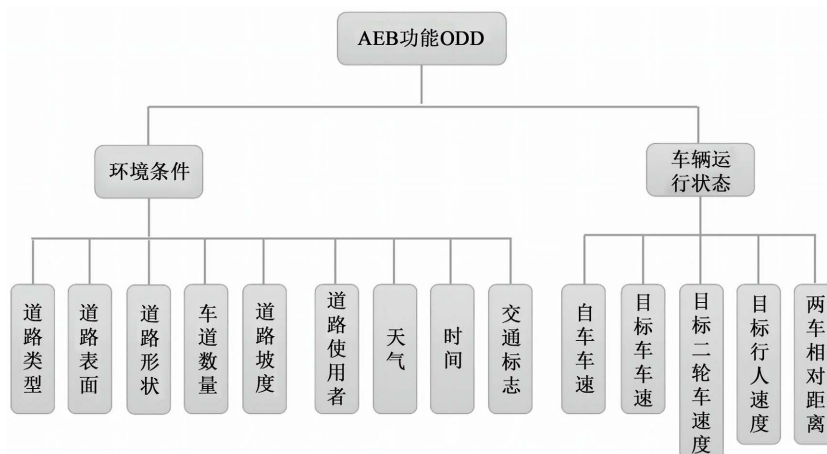


Figure 1. ODD element based on AEB functionality
图 1. 基于 AEB 功能的 ODD 元素

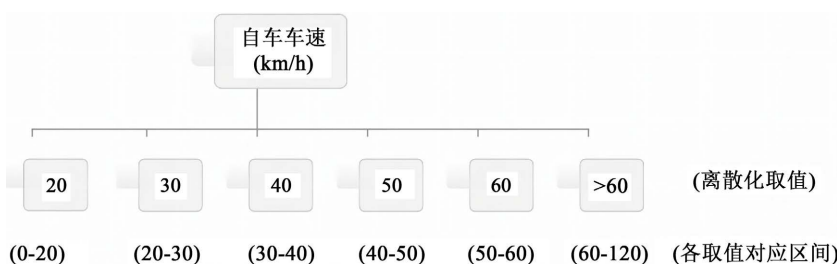


Figure 2. Discretization of self-driving speed value
图 2. 自车车速取值离散化处理

2.2. 主观权重赋权方法

主观赋权法是指评价者根据各风险因素的重要性不同进行主观对比, 通过计算得出对应权重的方法, 主要有专家打分法、层次分析法、灰色关联法等[6], 本文选择层次分析法(AHP)进行赋权。

AHP 法是指建立一个描述系统特征的树型结构, 除第一层和最有一层外, 任意一层与其上下层为隶属关系; 对每一层级的不同因素, 参考选择的比例尺度, 将任意两两因素进行相互比较, 得到重要性比较结果指标, 据此构成判断矩阵, 最终得到反应该因素重要性的权重。本文建立的 AEB 功能场景 ODD 模型, 符合 AHP 法的模型构建。

采用标度法对隶属于同一个上层元素节点的本层任意两个元素子节点 a_i 和 a_j 的相对影响程度赋予一个量化的标度 a_{ij} , 采用 9 分位比率标度法对标度 a_{ij} 进行排定[7]。比例标度如表 1 所示。

根据比例标度表, 可以得到如表 2 所示的重要度比较结果。

由于比较的可逆性, 标度 a_{ij} 还具有以下性质:

$$a_{ij} = \frac{1}{a_{ji}} > 0 \quad (1)$$

由上表可构造标度矩阵 A:

Table 1. Proportion scale table

表 1. 比例标度表

比例标度	标度含义
1	表示两个因素比较, 具有相同的重要性
3	表示两个因素比较, 一个因素相比另外一个因素稍微重要
5	表示两个因素比较, 一个因素相比另外一个因素一般重要
7	表示两个因素比较, 一个因素相比另外一个因素较为重要
9	表示两个因素比较, 一个因素相比另外一个因素极为重要
2, 4, 6, 8	介于上述两个相邻比较判断结果的中间值

Table 2. Importance comparison results

表 2. 重要度比较结果

	A_1	A_2	...	A_n
A_1	a_{11}	a_{12}	...	a_{1n}
A_2	a_{21}	a_{22}	...	a_{2n}
...
A_m	a_{m1}	a_{m2}	...	a_{mn}

$$A = [a_{ij}]_{mn} \tag{2}$$

然后, 需要计算出标度矩阵 A 的最大特征值所对应的特征向量, 即为权重向量 W_i 。

由于判断矩阵 A 是通过两两比较主观得到的比例标度, 因此不可避免的会出现结果的主观性, 为使计算的结果更合理, 需要检验矩阵的一致程度。因此, 引入一致性指标(Consistency index) CI 和平均随机一致性指标(Average random consistency index) RI 来衡量标度矩阵的一致性, 指标的计算公式如下:

$$CI = \frac{\lambda_{\max} - n}{n - 1} \tag{3}$$

RI 的取值会随着标度矩阵的阶数而变化, 见下表 3。

Table 3. The stochastic consistency indicator RI takes the value

表 3. 随机一致性指标 RI 取值

n	1	2	3	4	5	6	7	8	9	10	11	12
RI	0	0	0.58	0.9	1.12	1.24	1.32	1.41	1.45	1.49	1.51	1.54

最终, 可通过随机一致性比率(Random consistency ratio) CR 来反映标度矩阵一致性:

$$CR = \frac{CI}{RI} \tag{4}$$

当 $CR \leq 0.1$ 时, 可以认为标度矩阵 A 满足一致性要求, 此时通过矩阵 A 计算得到的最大特征值对应的特征向量就可以当作 ODD 元素的主观权重。

2.3. 客观权重赋权方法及主客观集成赋权方法

AEB 功能测试场景 ODD 中的一些元素, 比如自行车车速分布、行人速度分布以及环境条件中的部分

元素如道路类型、道路表面、天气等, 可以通过现有论文的统计数据以及国家统计局的数据得到其真实分布, 因此对于这些元素, 可以采用客观赋权法。

客观赋权相比较于主观赋权法, 不同的是计算权重时不通过评价者的主观态度进行赋权, 而是基于评价指标客观信息的差异确定。本文采用 CRITIC 法对 ODD 元素进行客观赋权, CRITIC 法(Criteria Importance Though Intercriteria Correlation)法是由一种基于数据波动性的客观权重赋权法, 由两个概念作为确定权重的理论基础, 即对比强度和冲突性[8]。

假设现有一组数据, 有 m 个待评价对象, n 个评价指标, 构成原始数据矩阵 X :

$$X = \begin{pmatrix} x_{11} & \cdots & x_{1n} \\ \cdots & \ddots & \cdots \\ x_{m1} & \cdots & x_{mn} \end{pmatrix} \quad (5)$$

第一步是进行无量纲化处理, 对于正向指标使用正向化处理, 对于负向指标使用负向化处理。

第二步是计算指标变异性, 指标变异性可以用方差、标准差或者熵来表现, 本文选择标准差:

$$\begin{cases} \bar{x}_j = \frac{1}{m} \sum_{i=1}^m x_{ij} \\ S_j = \sqrt{\frac{\sum_{i=1}^m (x_{ij} - \bar{x}_j)^2}{m-1}} \end{cases} \quad (6)$$

在 CRITIC 法中, 标准差可以看作反应各指标内取值的波动情况, 当标准差越大时, 表示该指标的数值间差异越大, 证明该指标越能够反映出更多的信息, 此时应该给该指标分配更高的权重。

第三步是计算指标冲突性, 指标冲突性可以用相关系数进行表示:

$$R_j = \sum_{i=1}^n (1 - r_{ij}) \quad (7)$$

使用相关系数来表示指标之间的相关性与冲突性, 相关性与冲突性成反比, 与其他指标的相关性越强, 代表该指标就与其他指标的冲突性就越小, 此时该指标反映的和其他指标相同的信息就会较多, 体现的评价内容和其他指标相比就有重复之处, 因此该指标的评价强度降低, 应该减少对该指标分配的权重。

第四步是计算信息量:

$$C_j = S_j \sum_{i=1}^n (1 - r_{ij}) = S_j \times R_j \quad (8)$$

其中 C_j 越大, 代表信息量较大, 那么该评价指标在整个评价指标体系中的作用就越大, 就应该给其分配更大的权重。

第五步即可计算出客观权重:

$$W_j = \frac{C_j}{\sum_{j=1}^n C_j} \quad (9)$$

主观赋权法的优点在于能够较好的体现专家经验的主观意识和判断, 但这同时也可能是缺点, 因为由于认知的偏好及经验的欠缺, 结果的可靠性会大大降低; 而客观赋权法虽然能够依据真实数据, 通过数学方法进行计算得到数据之间的重要程度, 但是缺少了主观意识的补充, 会导致判断结果片面单一。因此为了取长补短, 本文选择了一种主客观集成赋权方法, 以层次分析法和 CRITIC 法为例, 将通过主管赋权法与客观赋权法得到结果再次进行加权计算并得到权重结果。根据应用场景的不同, 选用不同的

主观赋权法、客观赋权法后, 均能采用下述步骤对主观和客观权重进行集成, 得到更合理的权重值。

设某个 ODD 元素有五个取值子节点, 通过层次分析法得出的权重向量为:

$$P = (w_{11}, w_{12}, w_{13}, w_{14}, w_{15}) \quad (10)$$

通过 CRITIC 法得出的权重向量为:

$$Q = (w_{21}, w_{22}, w_{23}, w_{24}, w_{25}) \quad (11)$$

第一步为减少计算量, 采用线性加权法构造多目标规划模型, 加权后的权重为:

$$W_j = k_1 P + k_2 Q \quad (12)$$

其中 $k_1^2 + k_2^2 = 1$, $k_1, k_2 \geq 0$ 。

第二步利用线性评价模型, 得到:

$$y_i = \sum_{j=1}^n x'_{ij} W_j \quad (13)$$

其中 x'_{ij} 为公式(5) CRITIC 法中无量纲化后得到的矩阵的第 i 行第 j 列值。显然, y_i 越大说明权重值越大, 因此, 可以构造多目标规划模型为:

$$\max \sum_{i=1}^m y_i = \sum_{i=1}^m \sum_{j=1}^n x_{ij} W_j = \sum_{i=1}^m \sum_{j=1}^n x_{ij} (k_1 W_{1j} + k_2 W_{2j}) \quad (14)$$

根据上述模型, 利用多目标规划求解方法[9]可得:

$$k_1 = \frac{a}{\sqrt{a^2 + b^2}}, \quad k_2 = \frac{b}{\sqrt{a^2 + b^2}} \quad (15)$$

式中: $a = \sum_{i=1}^m \sum_{j=1}^n x_{ij} W_{1j}$, $b = \sum_{i=1}^m \sum_{j=1}^n x_{ij} W_{2j}$ 。

将 k_1 和 k_2 带入公式(12), 最终可以得到主客观集成赋权后的权重向量 W_j 。

通过主客观集成赋权法得到的权重向量, 是用于衡量每个 ODD 最底层元素及其取值的相对权重指标。在 ODD 的模型中, 元素共分四层, 其中第四层元素节点是分属于不同的第三层元素节点, 而第三层元素节点又分属于环境条件与车辆运行状态两个第二层节点, 因此需要在计算完第四层元素节点各取值的权重以后, 回溯到第三层, 计算环境条件与车辆运行状态下各元素间的权重, 同理再计算第二层, 最后将第四层元素取值节点权重与其各层级父节点权重相乘, 得到最终的绝对权重。绝对权重的计算公式为:

$$\hat{W}_{ij} = \prod_{(a,b) \in \hat{\Psi}_{ij}} W_{ab} \quad (16)$$

其中 $\hat{\Psi}_{ij}$ 是指最底层元素的每个取值权重与最上层 ODD 元素节点间的回溯路径上经过的全部元素节点的下标。

3. 基于参数权重的改进组合测试方法

在得到 ODD 元素取值的权重指标后, 即可用于自动驾驶虚拟仿真测试场景的生成。针对现有的组合测试生成方法只能保证其生成的测试用例尽可能覆盖多的因素组合, 但是生成的测试用例数量较多, 无法保证其效率及有效性, 本章提出引入 ODD 元素权重对现有的组合测试算法进行改进, 从而使得通过改进算法生成的场景库测试用例集能够在保证覆盖度的同时, 减少测试用例的生成数量。

3.1. 测试用例与组合测试

测试用例(Test Case)是指根据不同的测试需求, 对测试任务的一组描述。针对自动驾驶虚拟仿真测试

而言, 测试用例就是为被测的自动驾驶系统设计一组输入条件, 从而验证被测系统能否在该条件下正常运行并获得预期的输出结果[10]。自动驾驶虚拟仿真测试的场景库即由一个个场景组成, 每个场景也可以看作是一个测试用例, 而每个测试用例是由多个能够对被测系统的相关功能产生影响的影响因素的取值组成, 在本文中, 影响因素即为 ODD 中的每一个元素。

组合测试(Combinatorial Testing, 简称 CT)方法, 在保证覆盖度较高的情况下, 可以减少测试用例的生成数量。组合测试多用于软件测试领域, 研究表明 ADAS 功能测试中的故障模式与系统缺陷, 大多数是由 1 到 3 个系统参数组合导致, 这与软件系统测试的故障原因相似, 因此在自动驾驶仿真测试领域的测试用例生成方法, 可以参考软件测试领域中的 CT 方法。

本文以两两组合测试为例, 生成自动驾驶仿真测试用例, 在介绍前需要了解 CT 算法的理论基础。

假设待测系统 SUT (Soft Under Testing)有 n 个输入参数, 可以用一个集合 X 来表示, $X = \{x_1, x_2, x_3, \dots, x_n\}$ 。集合中的参数 $x_j (j \in [1, n])$ 有不同的取值, 每个参数的取值集合可以看作 D_j , 即参数 x_j 有 $|D_j|$ 个不同的参数取值。本文的待测系统就是基于 ADAS 功能的自动驾驶 ODD。

定义 1. (测试用例集): 假设集合 $T = \{t_1, t_2, t_3, \dots, t_n\}$, 如果对 $\forall t_i \in T (i \in [1, n])$, 存在 $t_i = \{t_{i1}, t_{i2}, t_{i3}, \dots, t_{im}\}$, 且 $t_{ij} \in D_j (j \in [1, n])$ 。那么, 我们就称集合 T 是待测系统 SUT 的测试用例集, 集合 T 中的任一元素 t_i 则为待测系统 SUT 的一条测试用例[11]。本文的测试用例集就是基于 ADAS 功能的自动驾驶虚拟仿真测试场景库。

定义 2. (两两组合覆盖测试用例集): 假设集合 $T = \{t_1, t_2, t_3, \dots, t_n\}$ 是待测系统 SUT 的一个测试用例集, 若对 $x_i \in X, x_j \in X (i, j \in [1, n] \text{ 且 } i \neq j)$ 任意取值组合 $(y, z) (y \in D_i \text{ 且 } z \in D_j)$, 都至少在 T 中找到一条测试用例 t_m , 满足 $y \in t_m$ 且 $z \in t_m$ 。那么, 我们就称集合 T 是待测系统 SUT 的一个两两组合覆盖测试用例集。两两组合覆盖测试用例集也可称为二维组合覆盖测试用例集。

目前主要的研究在于通过 CT 算法找到具有最小规模的测试用例集, 使用的算法大致可以分为以下三类: 数学构造方法、贪心算法与元启发式搜索算法。本文选择贪心算法作为组合测试生成测试用例集的方法。贪心算法是指: 初始生成一个空的测试用例集或者一个存在部分参数的测试用例集 T 开始, 研究未覆盖的所有 n 维组合, 依次为依据, 进行逐行或者逐列的扩充, 将所有未覆盖的 n 维组合, 全部填充到测试用例集 T 中。该方法根据扩展方式不同主要分为: 基于 one-test-at-a-time 策略的逐条生成测试用例集方法和基于 in-parameter-order 策略的逐参数扩展生成测试用例集方法[11]。

在徐宝文等人的试验中表明, 在低维覆盖的组合测试中, 尤其是二维组合测试, 基于 in-parameter-order 的策略是明显劣于 one-test-at-a-time 策略的, 但是在高维组合测试中, in-parameter-order 策略在时间性能和测试用例集的规模上, 均优于基于 one-test-at-a-time 策略的算法[12]。

由于本文选择的待测系统是基于 ADAS 功能的自动驾驶 ODD, 且研究表明 ADAS 系统的故障均是由单一参数或 2~3 个参数诱发, 组合覆盖测试用例集的维度较低, 因此本文选择基于 one-test-at-a-time 策略作为待改进的组合测试算法, 以此生成测试用例集。

3.2. 基于参数权重的改进 ES 算法

ES 算法又称遍历搜索算法, 属于 one-test-at-a-time 策略, 主要设计思想为: 确定好所有参数的取值组合集, 遍历其中所有的测试用例, 寻找可以覆盖最多的未覆盖 n 维组合集的测试用例, 同时使得寻找的下一个测试用例集也可以覆盖尽可能多的未覆盖 n 维组合集。该算法可以简单概述为, 同时考虑当前的多覆盖与后续的多覆盖。

传统的 One-test-at-a-time 策略算法, 比如 AETG 算法, 在寻找下一个测试用例去补充测试用例集时, 选择的是下一个能覆盖最多未覆盖 n 维组合的测试用例。而梁凡等在改进的 AETG 算法中选择的是在未

能覆盖的 n 维组合中, 先统计各参数出现的次数, 将参数按出现的次数, 从大到小排序, 然后选择出现次数最大的参数进行扩展[13]。而本文使用 ES 算法考虑当前的多覆盖和后续的多覆盖时, 引入参数取值的权重进行计算。计算当前测试用例集中所有出现的参数对应取值的权重和, 记为 a , 则未覆盖的参数对应取值的权重和记为 b , 其中 $a \geq 0, b \geq 0, a + b = 1$, 因此可以称该算法为 $ES(a, b)$ 算法。传统的 $ES(a, b)$ 算法的 a, b 值为定值, 一般取 $(a = 1, b = 0)$ 、 $(a = 0.75, b = 0.25)$ 、 $(a = 0.5, b = 0.5)$ 、 $(a = 0, b = 1)$ 等几种情况进行计算, 而本文的 a, b 值随着测试用例集的增加而动态更新。在遍历参数的取值组合集时, 动态更新参数 a 和 b 的值, 使得每次计算的权重函数不同。

在测试用例集的扩充过程中, 考虑后续的多覆盖时也要考虑能够使用较少的测试用例覆盖较多的未覆盖组合集, 这就要求未覆盖的组合集中的参数要尽量分散, 即未覆盖的参数可以多, 但是每个参数未覆盖的取值要少。

下面对算法进行设计, 以覆盖维度 $n = 2$ 为例, 算法所设计的所有参数变量及其含义如表 4 所示。

Table 4. Variables involved in $ES(a, b)$ algorithms
表 4. $ES(a, b)$ 算法中涉及的变量

变量名称	变量含义
UT	所有参数的取值组合集
UC	所有未覆盖的二维取值组合集
T	生成的测试用例集
t	一个测试用例
$C(t)$	t 所覆盖 UC 中的取值组合的集合
$g(t)$	$C(t)$ 中的元素个数
$h(t)$	UC 未覆盖 $C(t)$ 中的组合集的最大的参数个数
Cover _{over} (T)	测试用例集 T 当前覆盖所有参数取值的权重和
Cover _{not} (T)	测试用例集 T 当前未覆盖所有参数取值的权重和
f	当前测试用例的权重函数

设计的 $ES(a, b)$ 算法具体步骤流程, 如下图 3 中所示。

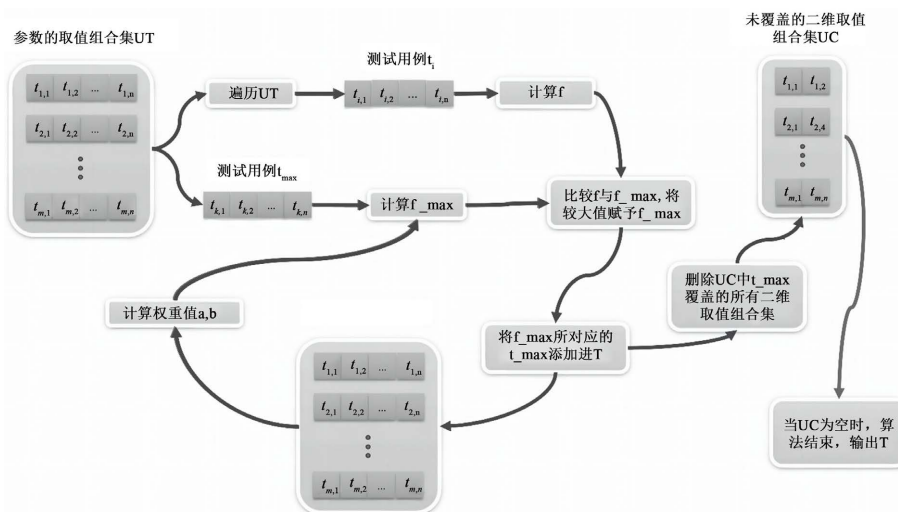


Figure 3. Flow diagram of test cases generation process in $ES(a, b)$ algorithm
图 3. $ES(a, b)$ 算法生成测试用例过程的方法流程图

算法的生成步骤如下:

步骤 1: 生成初始的 UT、UC、T, 其中 $T = \emptyset$ 。

步骤 2: 若 $UC \neq \emptyset$, 则跳转到步骤 3。否则, 结束算法, 测试用例集生成完成。

步骤 3: 从 UT 中删除 T 中存在的取值组合。将 UT 中的第一个取值组合定义为 t_{\max} , 遍历当前 T 中所有测试用例的参数取值的权重值进行求和, 若某个参数的取值出现多次, 则只算作一次, 最终得到参数 a , 并计算得到权重 b , 计算权重 a 乘 $g(t_{\max})$ 与权重 b 乘 $h(t_{\max})$ 之差, 称之为 f_{\max} 。

步骤 4: 遍历 UT 中的所有取值组合, 每一个取值组合定义为 t_i , 计算权重 a 乘 $g(t_i)$ 与权重 b 乘 $h(t_i)$ 之差, 称之为 f_i 。比较 f_{\max} 与 f_i , 倘若 $f_{\max} \leq f_i$, 则将 f_i 赋值给 f_{\max} , t_i 赋值给 t_{\max} 。

步骤 5: 遍历 UT 中的所有取值组合后, 将最终的 t_{\max} 添加进 T 中, 同时从 UC 中删除 $C(t_{\max})$ 。跳转至步骤 2。

最后, 给出 $ES(a, b)$ 算法实现的伪代码形式, 如表 5 中所示。

Table 5. $ES(a, b)$ algorithm pseudocode

表 5. $ES(a, b)$ 算法伪代码

Algorithm 1:

```

1: Input: UT, UC,  $C(t)$ ,  $a$ ,  $b$ ,  $g(t)$ ,  $h(t)$ 
2: Output: T
3:  $UC = UC_0$ ,  $T = []$ .
4: while  $UC \neq \emptyset$  do
5:  $t_{\max} = t_1$ ,  $t_1 = UT \setminus T$ .
6:  $a = Cover\_over(T)$ ,  $b = Cover\_not(T)$ .
7:  $f_{\max} = a \times g(t_{\max}) - b \times h(t_{\max})$ .
8: for all  $t_i$  in  $UT \setminus T$  do
9:  $f = a \times g(t_i) - b \times h(t_i)$ .
10: if  $f_{\max} \leq f$  then
11:  $f_{\max} = f$ 
12:  $t_{\max} = t_i$ 
13: end if
14: end for
15:  $T = T \cup \{t_{\max}\}$ .
16:  $UC = UC \setminus C(T_{\max})$ 
17: end while
18: return T

```

3.3. 算法应用与效果验证

本节以第二章中 AEB 功能的 ODD 作为待测系统, ODD 最底层元素作为输入参数, 将 $ES(a, b)$ 算法的实际应用效果进行展示。本文选取目前在工程应用领域较为广泛、且已有生成工具或者开源代码的三种 CT 算法 PICT、AllPairs、AETG 作为对比。当 $n = 2$ 时, 生成的两两组合覆盖测试用例集数量结果如下图 4 所示。

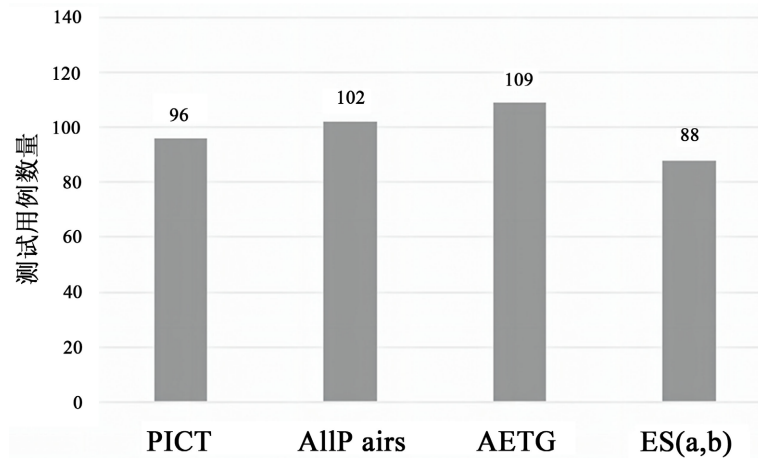


Figure 4. When $n = 2$, the combined test method generates a comparison of the number of test case sets

图 4. $n = 2$ 时组合测试方法生成测试用例集数量对比

从图 4 中可以看出, 和传统的 CT 算法 PICT、AllPairs 和 AETG 相比, 当组合覆盖强度 $n = 2$ 时, 由 $ES(a, b)$ 算法生成的用例个数相比于上述三种算法, 分别减少了 8.3%、13.7% 以及 19.3%。为展现 $ES(a, b)$ 算法的生成效果, 本文选择 AEB 功能的 ODD 中的环境条件下的参数共 8 个, 并对参数取值进行简化, 通过第二章的主客观集成赋权法重新计算其元素不同取值的权重值, 此时 UT 存在 2916 个取值组合集, UC 存在 211 个二维未覆盖取值组合集, 最终生成两两组合覆盖测试用例集, 其中包含用例个数 15 个。所有测试用例情况如表 6 所示。

Table 6. A two-by-two combination overlay test case set

表 6. 两两组合覆盖测试用例集

测试用例序号	道路类型	道路形状	道路表面	道路坡度	车道数	天气	时间	交通标志
1	乡村/国道	路口	其他路面	下坡	多	雪天	夜晚	交通灯
2	乡村/国道	弯道	沥青路面	上坡	单	雨天	黄昏	交通标志
3	高速	直道	混凝土路面	无坡度	多	晴天	白天	交通标志
4	城市主干道	路口	其他路面	上坡	单	晴天	白天	交通灯
5	乡村/国道	弯道	混凝土路面	无坡度	单	雪天	夜晚	交通灯
6	城市主干道	直道	沥青路面	下坡	多	雨天	黄昏	交通灯
7	高速	路口	其他路面	无坡度	单	雨天	黄昏	交通标志
8	高速	直道	沥青路面	上坡	多	雪天	夜晚	交通标志
9	乡村/国道	弯道	混凝土路面	下坡	多	雨天	白天	交通标志
10	高速	弯道	沥青路面	下坡	单	晴天	白天	交通灯
11	乡村/国道	直道	混凝土路面	上坡	单	晴天	黄昏	交通灯
12	乡村/国道	路口	沥青路面	无坡度	多	雪天	黄昏	交通灯
13	城市主干道	路口	混凝土路面	下坡	多	雪天	白天	交通标志
14	乡村/国道	弯道	其他路面	下坡	多	雨天	夜晚	交通灯
15	乡村/国道	直道	其他路面	下坡	多	晴天	夜晚	交通灯

综上所述, 本文提出的 $ES(a, b)$ 算法相对于现有的 CT 算法中的 one-test-at-a-time 策略算法的改进效果已经得到了有效验证。

4. 仿真测试与分析

为验证通过改进组合测试方法生成的场景是否可用于自动驾驶虚拟仿真测试, 本文使用 prescan 与 simulink 搭建仿真测试环境, 按照上文生成的 ODD 元素取值, 建立仿真测试场景, 进行 MIL 仿真测试。

4.1. 联合仿真环境搭建

Prescan 是一款用于驾驶员辅助系统和安全系统开发验证的仿真软件, 其主要功能有场景搭建、传感器建模、开发控制算法以及运行仿真[14]。在 prescan 中可以通过其自带软件库搭建静态路网, 包含了 ODD 元素中的所有道路信息; 可以通过 General Settings 中的 Weather 模块配置天气信息, 通过 Lighting 模块配置时间与光照信息。在添加车辆模型后, 可以配置车辆的运动状态信息。Prescan 的传感器建模主要针对摄像头、雷达/毫米波雷达以及车联网通讯传感器, 根据 ECU 软件可以对传感器进行仿真, Prescan 生成测试场景界面如图 5 所示。

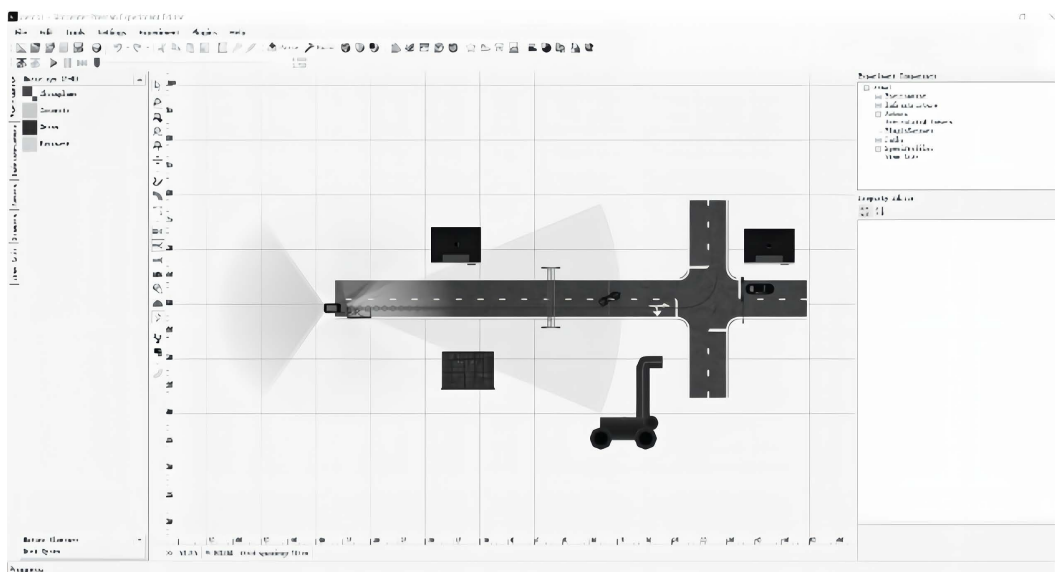


Figure 5. Prescan GUI interface audit generates test scenarios

图 5. Prescan GUI 界面生成测试场景

搭建完成后, 在 GUI 界面的工具栏中点击 Build 进行编译。编译通过后, Prescan 可以自动生成.slx 的可执行 Simulink 文件。

使用 Simulink 可以对动力学模型、传感器模型以及 AEB 控制算法进行设置与编辑。控制器作为输入端, 可以接收传感器模型通过仿真场景感知到的输入信号, 经过内部控制算法模块的计算, 把输出端发送请求指令给车辆动力学模型, 动力学通过接收的请求指令做出响应, 使得仿真场景中的主车驾驶行为变化, 形成一个闭合的在环测试, 从而达到验证控制器算法的目的, 为模型在环测试, 即 MIL 测试。

由于本文的研究重点在于仿真测试场景的生成, 使用仿真软件验证生成的场景可以用于仿真测试即可, 因此本文采用 Prescan 自带的动力学模型、传感器模型与 AEB 控制算法。Prescan 生成 simulink 中的 GPS 惯导模块、传感器模块、车辆动力学模型、AEB 控制算法模块如图 6 所示。

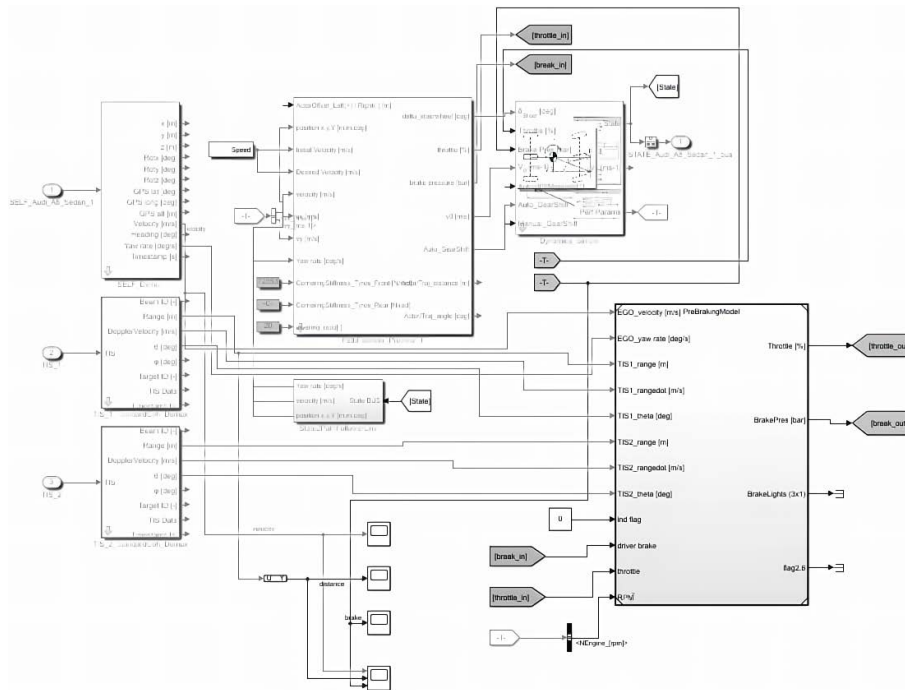


Figure 6. Prescan generates the status of each module in simulink
 图 6. Prescan 生成 simulink 中各模块情况

4.2. 仿真结果分析

本文选择 AEB 标准法规中的前车静止, 测试车辆即主车与前车追尾场景(Car-to-Car Stationary, CCRs), ODD 元素的取值为: 城市道路、直道、沥青路面、双向 4 车道、道路无坡度、白天、晴天、有交通标志, 目标车类型为小型汽车, 主车初始速度为 54 km/h, 目标车初始速度为 0 km/h 且一直静止不动, 两车相对距离为 150 m。根据改进组合测试生成的 AEB 功能 ODD 元素及其取值在 Prescan 中绘制场景并进行仿真, 仿真结果如图 7 所示。

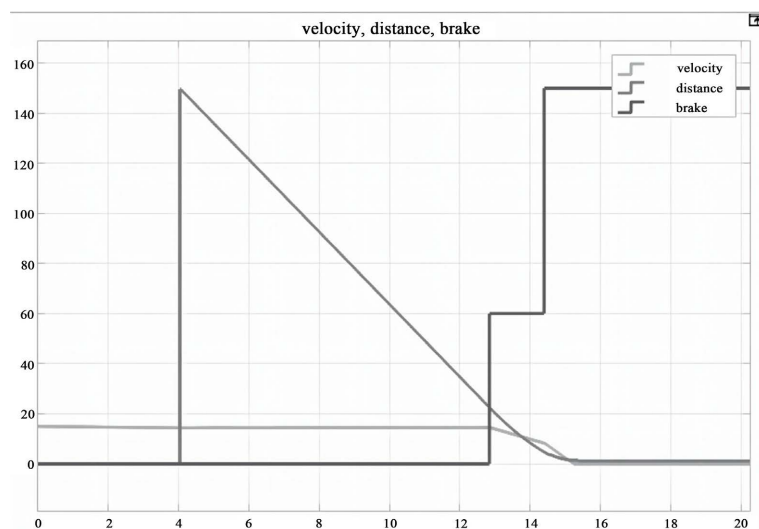


Figure 7. Prescan simulation results
 图 7. Prescan 仿真结果

图 7 中曲线分别代表主车与目标车的相对距离、主车车速以及驻车制动力曲线。主车减速分为两个阶段, 即轻微制动与全力制动至车辆刹停。该 AEB 算法使用碰撞事件(TTC)指标进行判断, $TTC < 2.6$ s 时系统给出报警, $TTC < 1.6$ s 系统给出 40% 的半程制动, $TTC < 0.6$ s 时系统全力制动。

综上, 通过本文提出的改进组合测试方法生成的场景用于 AEB 功能的 MIL 测试的可行性已得到验证。

5. 结语

本文根据 ADAS 功能测试的标准法规, 总结了功能场景 ODD 元素及其取值, 并使用主观赋权法层次分析法、客观赋权法 CRITIC 法以及主客观集成赋权法计算元素取值的权重。引入计算好的权重指标改进组合测试方法中的 $ES(a, b)$ 算法, 使得通过该算法生成的场景库中场景的数量明显少于其他组合测试算法。最后, 通过 simulink 和 Prescan 搭建联合仿真, 证明通过改进组合测试生成的算法可以用于自动驾驶 MIL 测试。但是, 在 Prescan 中手动搭建测试场景仍较为耗时, 后续的工作考虑依据仿真软件支持的 OpenX 标准的场景文件, 设计场景快速自定义生成工具, 实现场景的大批量、高效率测试。

参考文献

- [1] 邓伟文, 李江坤, 任秉韬, 等. 面向自动驾驶的仿真场景自动生成方法综述[J]. 中国公路学报, 2022, 35(1): 316-333.
- [2] 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.
- [3] Maity, S. and Nayak, A. (2005) Improved Test Generation Algorithms for Pair-Wise Testing. *16th IEEE International Symposium on Software Reliability Engineering (ISSRE'05)*, 235-224.
https://www.researchgate.net/publication/4195962_Improved_test_generation_algorithms_for_pair-wise_testing
- [4] Zhang, Q., Chen, D., Li, Y., et al. (2015) Research on Performance Test Method of Lane Departure Warning System with PreScan. *Proceedings of SAE-China Congress 2014: Selected Papers*, 328, 445-453.
https://doi.org/10.1007/978-3-662-45043-7_45
- [5] 高锋, 张强, 周晶洁, 等. 车身控制器功能逻辑测试技术的研究[J]. 汽车工程, 2009, 31(10): 966-970.
- [6] 王慧. 基于多智能体进化算法的 COA 方法研究[D]: [硕士学位论文]. 西安: 西安电子科技大学, 2010.
- [7] 郭金玉, 张忠彬, 孙庆云. 层次分析法的研究与应用[J]. 中国安全科学学报, 2008, 18(5): 148-153.
- [8] Diakoulaki, D., Mavrotas, G. and Papayannakis, L. (1995) Determining Objective Weights in Multiple Criteria Problems: The CRITIC Method. *Computers & Operations Research*, 22, 763-770.
[https://doi.org/10.1016/0305-0548\(94\)00059-H](https://doi.org/10.1016/0305-0548(94)00059-H)
- [9] 刘三明. 多目标规划的若干理论和方法[D]: [博士学位论文]. 大连: 大连理工大学, 2006.
- [10] 段剑犁. 面向智能驾驶系统的自动化仿真测试评价技术研究[D]: [博士学位论文]. 重庆: 重庆大学, 2020.
- [11] 王雅婷. 参数带权值的组合测试用例集生成研究[D]: [硕士学位论文]. 合肥: 安徽大学, 2015.
- [12] 戴丽. 组合测试用例生成技术的研究与应用[D]: [硕士学位论文]. 广州: 华南理工大学, 2011.
- [13] 宋晓秋, 梁凡. 两两组合测试用例生成的遍历搜索算法[J]. 计算机工程与设计, 2019, 40(2): 433-437.
- [14] 赵伊齐, 张引, 申成刚, 等. 基于 Prescan 的智能驾驶辅助系统在环研究[J]. 汽车实用技术, 2019(9): 47-49+59.