

基于代理遗传优化的智能驾驶系统加速测试方法

朱冰, 汤瑞, 赵健, 张培兴, 李文旭

(吉林大学汽车仿真与控制国家重点实验室, 吉林 长春 130022)

摘要: 提出了一种基于代理遗传优化的智能驾驶系统加速测试方法。首先, 通过场景要素层次分析权值与优解区域特征改进参数采样模块中的拉丁超立方采样区间, 实现了采样效率与优化效果的协同提升; 其次, 利用参数采样结果和重复度筛选机制增加遗传寻优模块的种群多样性, 克服了传统遗传算法的局部收敛难题; 然后, 利用基于循环更新机制的代理筛选模块对场景测试结果进行预测, 平衡了加速算法与代理模型应用之间的效率与精度矛盾; 最后, 搭建仿真平台在高维时序分解的前车变速场景下对待测智能驾驶系统进行加速测试与验证。结果表明, 本文提出的方法可有效搜寻大量关键场景并提升测试效率。

关键词: 汽车工程; 智能驾驶系统加速测试; 代理模型; 遗传算法; 拉丁超立方采样

中图分类号: U467.13

文献标志码: A

Accelerated Test Method of Intelligent Driving System Based on Surrogate Genetic Optimization Model

ZHU Bing, TANG Rui, ZHAO Jian, ZHANG Peixing, LI Wenxu

(State Key Laboratory of Automotive Simulation and Control, Jilin University, Changchun 130022, China)

Abstract: This paper proposed an accelerated test method for intelligent driving system based on the surrogate genetic optimization model. First, the Latin hypercube sampling interval in the parameter sampling module was improved by using the weights and optimal solution region features of the scenario element hierarchical analysis method, achieving a synergistic improvement in sampling efficiency and optimization effect. Next, by utilizing parameter sampling results and repeatability screening mechanism, the population

diversity of the genetic optimization module was increased, overcoming the local convergence problem of traditional genetic algorithms. Then, the surrogate filtering module based on cyclic update mechanism was used to predict the test results of the scenario, which balanced the contradiction between the efficiency and accuracy of the accelerated algorithm and the application of the surrogate model. Finally, a simulation platform was built to accelerate test process and verification of the intelligent driving system to be tested in the front vehicle speed change scenario of high-dimensional time series decomposition. The results indicate that the method proposed in this paper can effectively search for a large number of key scenarios and improve testing efficiency.

Keywords: automotive engineering; intelligent driving system accelerated testing; surrogate model; genetic algorithm; Latin hypercube sampling

在智能驾驶技术的发展过程中, 如何对其进行高效可信的测试是产业发展亟需解决的重要问题^[1]。基于场景的测试方法在测试安全、效率和成本方面均展现出巨大优势, 已成为一种重要的智能驾驶系统测试验证方式。

然而, 由于测试场景参数维度庞大, 遍历测试会产生“参数爆炸”问题^[2]。兰德公司研究^[3]表明, 在 95% 的置信水平下, 以不超过 20% 的相对误差构造置信区间时, 大约需要 88 亿英里的测试才能证明智能驾驶系统的安全性, “长尾效应”显著。针对关键场景进行加速测试可以极大地提升测试速度^[4], 有效解决“参数爆炸”与“长尾效应”问题, 减少测试资源消耗。

收稿日期: 2023-10-10

基金项目: 国家重点研发计划(2022YFB2503402); 国家自然科学基金(U22A20247, 52172386); 吉林省科技发展计划(20220201023GX)

第一作者: 朱冰, 教授, 博士生导师, 工学博士, 主要研究方向为智能驾驶系统测试评价技术。
E-mail: zhubing@jlu.edu.cn

通信作者: 赵健, 教授, 博士生导师, 工学博士, 主要研究方向为汽车地面系统分析与控制。
E-mail: zhaojian@jlu.edu.cn



论文
拓展
介绍

目前基于场景的加速测试方法主要有重要性采样与优化搜索两类。在重要性采样方法中,Zhao等^[5]通过改变概率分布函数,在变道场景中提升危险场景出现的概率。Feng等^[6]通过密集强化学习生成了专注于安全关键测试场景的仿真环境,实现了针对自然驾驶环境测试过程的等效加速测试。但重要性采样方法在拟合驾驶数据的过程中可能会出现误差,而且在相互作用的变量个数较多时,无法有效体现加速测试效果。优化搜索方法通过关注关键场景下的系统表现,将加速测试问题转换为场景参数空间中的优化搜索问题。邢星宇等^[7]使用遗传算法和贝叶斯算法,自动加速生成了决策规划系统的边界场景;Klück等^[8]使用遗传算法、模拟退火和随机搜索算法对智能驾驶系统进行测试,验证了遗传算法的最优性;张培兴等^[9]提出了一种并行优化搜索算法,将多节点并行仿真与优化方法高效结合。优化搜索方法可以提高测试效率,但现有算法仍需不断执行耗时的仿真过程,工程上广泛使用易于计算的代理模型解决此类黑箱问题^[10]。然而优化算法在高维参数空间上存在局部收敛特征,难以通过全面的场景数据训练高精度的代理模型,影响了代理模型的应用,而过于分散的场景数据获取会降低算法的寻优速度,由此带来了模型精度、测试效率和算法特性之间的矛盾。综上,如何解决高维参数空间优化搜索测试过程中的局部收敛和测试成本高的挑战,是目前亟需解决的难点问题。

本文提出一种包含采样-寻优-筛选的加速测试架构,设计了一种基于代理遗传优化(surrogate genetic optimization, SGO)的智能驾驶系统加速测试方法,主要内容如下:

(1)针对传统采样方法采样过程随机,重点不突出的缺陷,提出基于层次分析权值与优解区域特征的改进拉丁超立方采样方法。

(2)针对传统遗传算法优化过程局部收敛与搜索速度慢的弊端,提出基于重复度筛选的改进遗传算法。

(3)针对传统仿真测试过程全测试的成本与效率问题,提出基于代理模型筛选的仿真测试流程。

1 基于SGO的加速测试方法

基于SGO的智能驾驶系统加速测试方法主要包括参数采样模块、遗传寻优模块和代理筛选模块3个核心部分,架构如图1所示。

参数采样模块用于在场景参数空间中采样生成具体场景参数,以供遗传寻优模块进行种群生成和补充。遗传寻优模块通过采样数据和结果运算得出的适应度值进行场景要素的优化搜索。代理筛选模块以代理模型代替耗时的模拟仿真过程,可剔除大量非关键场景。3个模块相互耦合,共同完成高效而全面的加速测试工作。

1.1 参数采样模块

将待测智能驾驶系统看作黑箱,为了使遗传寻优模块更快找到具有潜在价值的测试场景,在测试开始时应采样到批量分散的场景参数。同时,为了提升对关键测试场景的持续探索能力,需要在迭代过程中进行分散场景的补充采样。

1.1.1 采样方法分析

为保证搜索到的测试场景参数足够分散,本文采用拉丁超立方采样(Latin hypercube sampling, LHS)生成初始采样点,LHS可以对每批次采样过程进行强制分层管理^[11],可提供均匀全面的样本数据。

拉丁超立方采样方法的基本过程如下:

- (1)对每一场景维度,按照累计密度函数分成 n 个相等分区(n 为场景要素维度)。
- (2)在单维度分区内随机选择数据点。
- (3)将各维度结果组合生成采样空间。
- (4)使用随机采样方法在采样空间中选取所需样本。

1.1.2 场景要素关键性层次分析

对于智能驾驶仿真测试而言,不同场景要素对测试结果具有不同影响程度,采样频次也应不一样。同时,伴随着优化过程的进行,优解区域的要素取值特征会被逐渐发掘,合理运用该特征可以进一步加速场景生成过程。因此,本文结合层次分析法,确定各场景要素的采样权值,生成基础采样区域,并在测试过程中结合要素取值特征更新采样区域,从而分配更为合理的采样数目。

层次分析法需要根据分析目标设计场景要素分类,并通过相互比较确立各要素分类的相对重要性权值。在后续的场景设计过程中,会根据要素分类的重要性权值确立拉丁超立方分区数量。

明确的待测系统是层次分析的前提,本文针对某款以毫米波雷达作为传感器的黑盒自动紧急制动(autonomous emergency braking, AEB)算法进行测试说明。如何迅速找到大量的关键危险场景是AEB系统测试的重点内容。

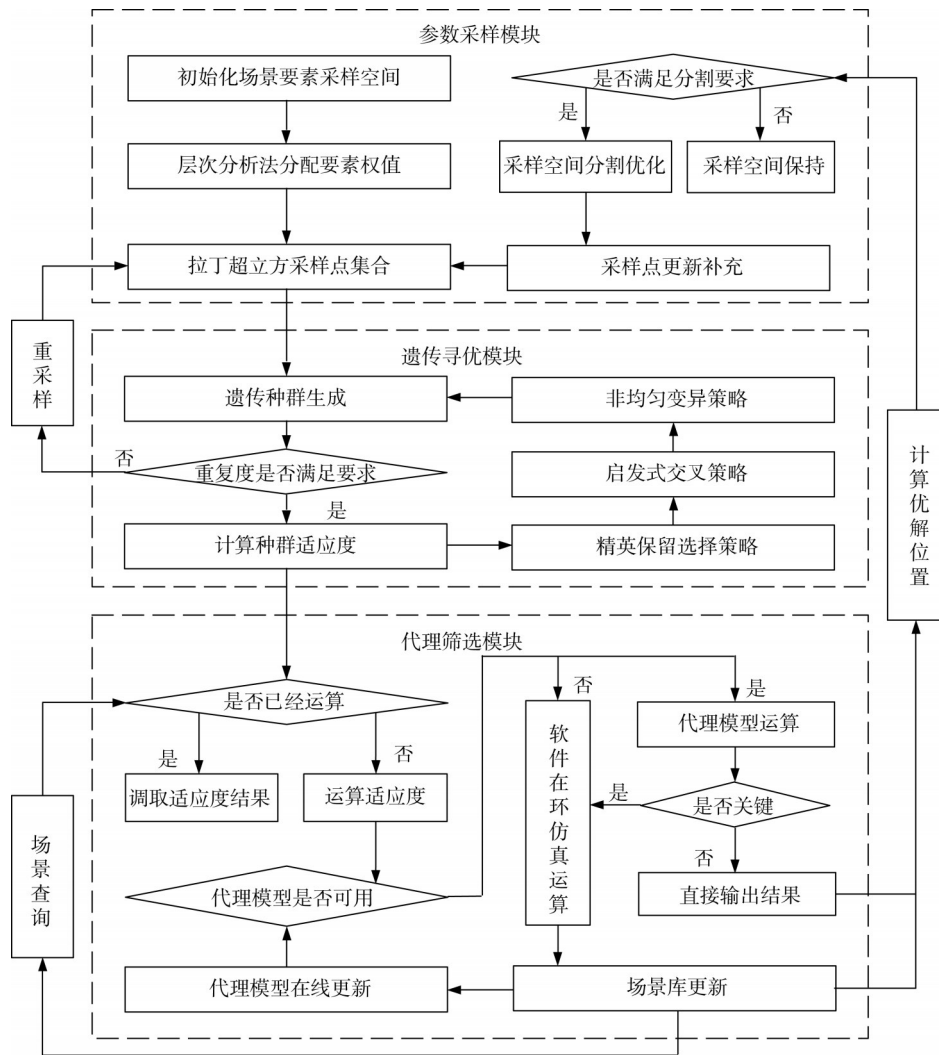


图1 算法整体架构

Fig.1 Overall architecture of algorithm

路面情况和运动过程是制动相关算法测试所关注的重点内容,相关研究表明降雨等天气对毫米波雷达感知效果有重要影响^[12]。所以,在AEB算法测试过程中,以对测试场景风险的影响为分析目标,可将相关场景要素分解为天气、路面和运动三方面内容,具体包含天气要素(W ,包括降雨强度等天气量化指标)、路面要素(P ,包括路面附着系数等)、距离要素(D ,包括各交通参与者的距离要素)、速度要素(V ,包括各交通参与者的速度)、加速度要素(A ,包括各交通参与者的加速度)5类。对于风险性层次中的要素,按照两两比较的方式确定相对重要性,分析结果如表1所示。

对其进行一致性检验,计算得出一致性指标值(consistency index, CI)为0.003 8,层次分析法5阶判断矩阵的随机一致性指标值(random index, RI)为1.12,一致性比例(consistency ratio, CR)计算如下:

表1 场景要素层次分析结果

Tab.1 Hierarchy analysis results of scenario elements

要素	特征向量	权值/%	最大特征根	一致性指标
W	0.303 7	6.07		
P	0.534 1	10.68		
D	0.899 3	17.99	5.015 3	0.003 8
V	1.631 5	32.63		
A	1.631 5	32.63		

$$CR = CI/RI = 0.003 4 \quad (1)$$

CR 小于限值0.1,所以层次分析结果满足一致性检验。

1.1.3 基于要素权值的采样区域生成

通过上述层次分析法,获得了场景要素分类的相对重要性权值。具体的场景要素采样过程即为在场景要素的取值空间中选取合适的采样点。拉丁超立方采样法在样本点生成前需要获取各场景维度的采样区间。

不同场景要素的测试步长不一致,由此引起了基准场景分区的差异,同时相对更重要的要素对应的采样频次应该更多,结合待测系统特性和表1中的层次分析法分析结果,对各种要素分类的采样分区设计如表2所示。基准分区与各要素取值特性有关,测试步长小、取值范围大的要素基准分区数设定较高。归一比例与层次分析权值有关,按最高值为1进行运算。实际分区数为基准分区数乘以归一比例并向上圆整。

表2 要素采样区间分配结果
Tab.2 Allocation results of element sampling interval

要素	W	P	D	V	A
基准分区	10	10	30	30	10
归一比例	0.19	0.33	0.55	1.00	1.00
实际分区	2	4	17	30	10

1.1.4 基于优解特征的采样区域更新

原始的拉丁超立方采样过程重点在于分散集群探索,分散的场景数据在测试前期有利于优化算法的全局寻优,但在测试后期,过于分散的数据会引入大量非关键场景要素区间,影响算法的局部寻优。为了更好地应用场景探索过程中的优解特征,与优化搜索算法结合,本文提出基于优解特征的采样区域更新思想,具体应用流程如图2所示。

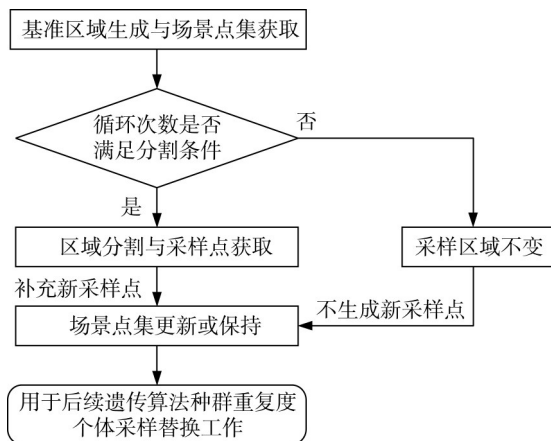


图2 采样区域和采样点集更新流程

Fig.2 Sampling area and point library update process

图2中,基准区域为各场景要素的基础取值范围,结合1.1.3中计算得出的实际分区,使用拉丁超立方采样获取均布的基础采样点并形成初始场景点集。本文中的场景点集主要用于后续遗传算法中种群个体的采样补充工作。如果后续遗传算法达到指定循环次数,则认为达到区域分割条件,对当前采样

区域进行分割并生成新采样点,与基准点集融合以补充更新场景点集。如果没有达到分割条件,则保持原始采样区域与场景点集不变。

如前文所述,如果满足分割条件,采用可变样本步长的手段对各场景要素区域进行分割优化,分割更新过程如图3所示。

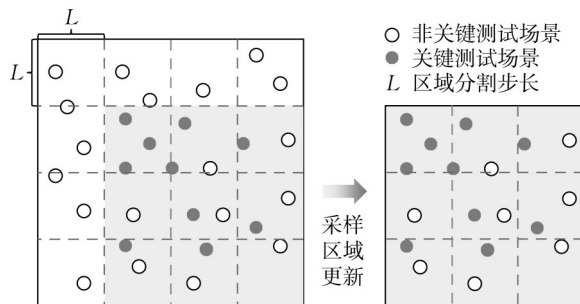


图3 基于优解特征的采样区域分割更新过程

Fig.3 Sampling region segmentation and update process based on optimal solution features

图3中关键测试场景即为需要搜索发掘的目标场景,非关键测试场景为搜索过程中的其他场景,区域分割步长为区域划分的基准长度。在满足分割条件时,通过分割步长对各场景要素进行区间划分,将当前场景空间内仅存在非关键测试场景的区域分割删除,由此更新生成新一轮场景区域。初始分割步长为各场景采样区间长度的1/4,并在后续每一轮缩减时将步长转变为当前步长的1/2。

1.2 遗传寻优模块

遗传算法是一种模仿生物进化中的基因变换过程的算法,符合“优胜劣汰、适者生存”准则,具有全局性搜索与启发式迭代能力。传统遗传算法的编码方式、适应度函数、选择、交叉和变异过程都具有一定局限性,难以高效应对初始参数空间巨大,目标解数量众多的智能驾驶系统加速测试问题,具体问题分析及改进思路如下。

1.2.1 编码方式设计

编码是算法寻优的关键,通过编码和解码,可以使用于仿真的表现型与可进化处理的基因型之间相互转换。经典遗传算法采用二进制编码,当优化搜索接近最优解时,由于基因位变化后表现型变化大且具有跳跃性,所以子代会偏离该最优解,难以稳定。因此,有学者提出使用格雷码代替二进制编码,提升遗传搜索的连续性^[13]。

此外,对于参数数量大的个体而言,需要使用较多位二进制数表示,虽会提高搜索精度,但会急剧增大参数空间,且提高的精度不能在模拟仿真上体现。

如图4a、4b所示,假定场景要素范围为 $[20, 60]\text{km}\cdot\text{h}^{-1}$,由仿真精度等确定的离散步长为 $4\text{km}\cdot\text{h}^{-1}$,共有11个参数取值,至少需要4位二进制数表示,从而表示了16位数,步长变为 $40/15 = 2.667$,但此步长不能在仿真过程中明显区分,还扩大了整体参数空间。如果采用多个二进制编码表示一个要素取值,以解决参数空间扩大问题,则会改变不同参数值的取值权重,影响参数优化过程。格雷码仍然采用二进制表示结果,所以也存在增大参数空间的问题。

实数编码直接使用个体的表现型进行编码,如图4c、4d所示,既保证了搜索过程的连续性,又不会增大参数空间,同时取消了解码模块,后续遗传算法的生成结果可直接用于仿真测试。因此,本文采用实数编码。

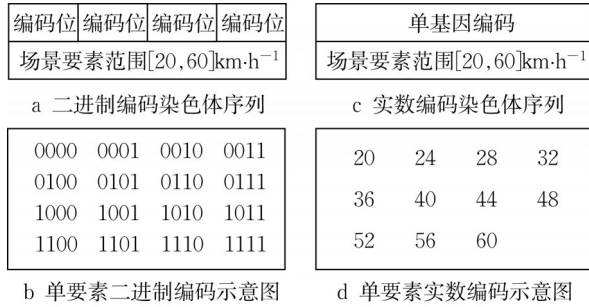


图4 场景要素编码方式对比

Fig.4 Comparison of scenario element encoding methods

1.2.2 目标函数设计

种群个体适应度是选择算子依据的主要指标,该值越大表明该个体在该环境下竞争力更强,也更有利于传递基因。经典遗传算法的目的为寻找全局最优解,尽量避免局部最优解生成,所以适应度值与目标函数有较强的线性对应关系。而对于智能驾驶加速测试问题,满足条件的优秀个体数量较多,更偏向于搜索目标解空间,所以需要适应度函数进行调整。

前文实例中以测试风险性层级为层次分析法分析目标,所以该实例的关键测试场景主要针对测试过程中的危险场景,本文以最大碰撞时间倒数(TTC_{\max}^{-1})为关键场景评价指标。 TTC_{\max}^{-1} 的计算公式如下:

$$\text{TTC}_{\max}^{-1} = \max((v_{\text{ego}} - v_{\text{front}})/s) \quad (2)$$

式中: v_{ego} 为主车车速; v_{front} 为前车车速; s 为主车与前车的初始相对位置; TTC_{\max}^{-1} 为计算出的最大碰撞时间倒数。

本文用于仿真测试的智能驾驶算法紧急制动碰撞时间阈值为 0.6s ,紧急制动表示该场景触发了算法的避险行为,所以将关键测试场景定义为 TTC_{\max}^{-1} 大于 1.6s^{-1} 的场景。为了保证区间内的寻优,而不是对某一个点进行寻优,本文设计如下适应度函数:

$$f_{\text{fit}} = \begin{cases} C + \frac{\text{TTC}_{\max}^{-1}}{100}, & \text{TTC}_{\max}^{-1} > 1.6\text{s}^{-1} \\ \text{TTC}_{\max}^{-1}, & \text{TTC}_{\max}^{-1} \leq 1.6\text{s}^{-1} \end{cases} \quad (3)$$

式中: C 为一个较大常数,为了引导整体搜索方向,加上 $\text{TTC}_{\max}^{-1}/100$ 是为了保证后续启发式交叉算子的应用效果。

1.2.3 选择算子设计

经典遗传算法父辈选择方式为轮盘赌选择法,即选择可能性与其适应度成正相关。此方法实现简单,但难以适配解空间搜索,导致生成的测试用例优秀个体较少。因此,需要对优秀个体进行保留,从而引导搜索方向,加速寻优过程。同时,由于优秀个体的适应度值较高,选择后的种群中可能存在大量重复个体,种群中个体分布如图5a所示,只存在少量个体类别,后续的启发式寻优易于趋向局部收敛,降低算法针对解空间的全局探索性。

因此,本文在轮盘赌选择的基础上,设计精英保留和重复个体筛选的选择策略。保留父辈中的最优个体,又通过种群筛选降低重复度,定义种群中个体重复度为当前个体在种群中的重复次数,将超过重复度阈值的个体在采样点集中进行采样替换。从而既保证了整体趋向于寻找更多的关键场景,又避免了过早收敛到局部最优解,改进算法的种群个体分布如图5b所示,相较于传统算法,显著提升了种群内的个体类别数量。

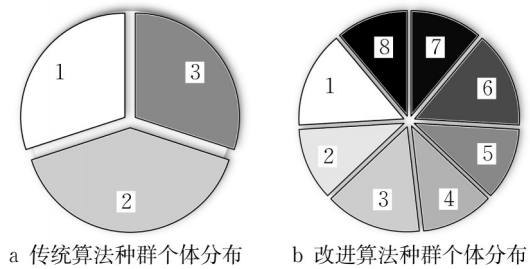


图5 种群内个体分布情况对比

Fig.5 Comparison of individual distribution within a population

1.2.4 交叉算子设计

经典的交叉算子主要应用于二进制编码,而实数编码采用直接离散交叉的方式,即对父辈某些变

量进行参数调换,对算法收敛没有贡献,仅生成了同等数量的新个体。如图6a所示,父代个体基因划分为两部分,随机组合后生成子代基因,不考虑父代自身的适应度差异。

为了在交叉层面突出加速效果,本文采用启发式交叉算子,即根据父代的适应度值进行子代个体的启发式变换,比较两父代的适应度值,保留优胜者,并将另一个体加上两者各参数差值的1.2倍,从而在交叉层面即体现进化过程,加速种群个体向优值收敛。如图6b所示,适应度值大的父代直接作为子代保留,适应度值小的父代在原来基因型的基础上加上其与适应度值大的基因型差值的1.2倍,学习了部分优秀个体特征。

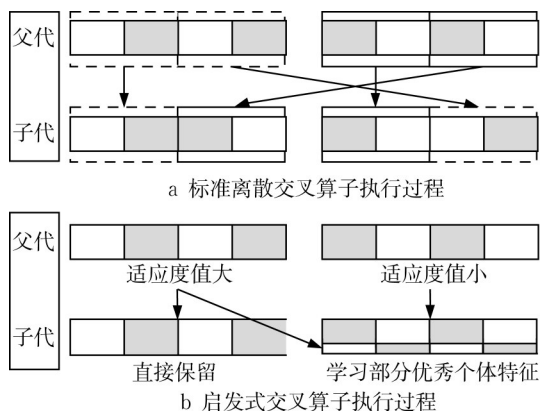


图6 交叉算子改进效果

Fig.6 Improvement effect of crossover operator

1.2.5 变异算子设计

经典的变异方式仅在一定概率下采用随机的方式变动参数值,同样对遗传算法收敛速度没有贡献。

本文使用重复度改进的非均匀变异方式,原理如下式所示:

$$x' = \begin{cases} x + (x_{\max} - x)(1 - r^{b(1-t/T)}), & r_m > 0.5 \\ x - (x - x_{\min})(1 - r^{b(1-t/T)}), & r_m \leq 0.5 \end{cases} \quad (4)$$

式中: x 为父代; x' 为子代; x_{\max} 和 x_{\min} 为各要素取值的最大值和最小值; r 和 r_m 为随机数; b 为非均匀系数; t 为当前个体在种群中的重复度; T 为整体重复度阈值。

变异效果与重复度相关,重复度越高,变动值越大,从而可以动态调整搜索步长,防止局部最优,促进全局优化收敛。

1.3 代理筛选模块

在智能驾驶仿真测试过程中,一次测试需要完成场景编辑、软件加载、模拟仿真以及结果运算等步

骤,需要消耗大量的时间。代理模型是指在分析和优化设计过程中可替代那些复杂而费时的模型的近似学习模型,以代理模型代替现有的仿真测试过程,可以显著缩短测试时间,降低测试成本。

本文在智能驾驶仿真测试的过程中融入代理运算机制,利用分层采样的改进遗传算法生成分散而全面的场景参数,有利于代理模型训练。利用代理模型运算消耗低,运算速度快等优势辅佐改进遗传算法,以期全面提升测试效率。

1.3.1 测试场景划分

在代理遗传优化算法执行过程中,存在多种不同类型的测试场景,可细分为:已仿真场景、代理预测场景、种群内场景,不同的场景分类有不同的用途。

(1)已仿真场景:表示经过模拟仿真软件仿真得出具体关键场景评价指标的场景,每次执行仿真测试后更新。用于计算遗传算法个体的仿真测试结果、避免重复执行仿真测试以及代理模型的训练与更新。根据仿真测试结果,已仿真场景可分为仿真关键场景和仿真非关键场景两部分,仿真关键场景即为本文重点搜索的目标场景。

(2)代理预测场景:表示经过代理模型预测出关键场景评价指标的场景,每次执行代理预测模型后更新。用于计算遗传个体的代理测试结果。考虑前期代理模型存在自身精度问题,代理模型的关键性阈值相较于软件在环仿真测试中的阈值更为宽泛,从而避免遗漏仿真测试关键边界处的场景,最大程度地平衡代理模型应用效果和场景测试结果的有效性。因此,根据代理预测结果,代理预测场景可分为代理关键场景和代理非关键场景两部分。代理关键场景即转入仿真测试阶段进行下一步验证,代理非关键场景视为仿真非关键场景,无需执行仿真测试。

(3)种群内场景:表示遗传算法种群内的场景,每次算法生成新一代场景后进行更新。主要用于种群内重复度计算和遗传算法的优化过程。种群内场景需通过代理模型或者仿真测试进行适应度计算。

1.3.2 代理模型构建

代理模型可以视作仿真结果的预测模型,现有的代理模型主要有神经网络、多项式插值、克里金模型等。由于智能驾驶仿真测试的每轮循环生成的场景数量有限,即用于代理模型训练的数据较少,所以不考虑训练过程需要较大数据量的神经网络模型。使用基础机器学习算法完成代理模型的设计。

随机森林模型以多个决策树集成学习为基础,

进行回归运算时,将所有决策树的预测结果加权平均作为模型输出,可以充分发挥各决策树的预测效果与泛化能力^[14]。因此,本文使用随机森林模型作为代理模型对测试结果进行预测。结合遗传算法的优化过程,设计随机森林基础结构如图7所示,集成50棵决策树,输入参数为各场景要素的取值,输出参数为场景的预测测试结果。

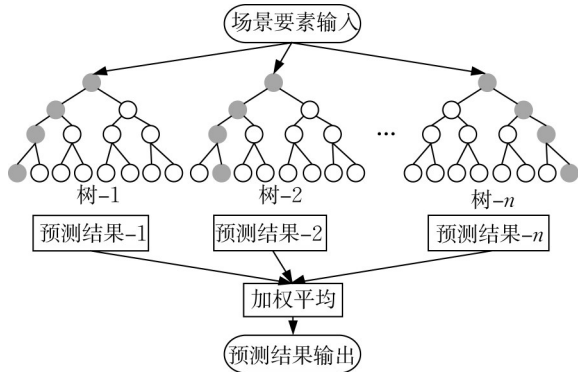


图7 随机森林代理模型结构

Fig.7 Structure of random forest surrogate model

1.3.3 代理模型训练与更新

针对随机森林模型在面对大量数据时训练耗时长,会丢失原有数据特征的缺陷,本文采用循环在线更新森林结构的训练方式,通过新的仿真测试数据补充森林结构。

代理模型的训练数据为已仿真场景和对应的风险测试结果,当已仿真场景数目大于100则执行初始化训练过程,划分训练集与测试集比例为7:3。训练完成后执行模型测试,并输出对应阶段的均方根

误差,作为后续代理模型应用依据。

在初始化训练后,后续执行仿真测试过程会补充新的已仿真场景,当新补充的仿真场景数据数量大于设定数值时,则进行代理模型的在线更新工作。在原有模型基础上,融入新的决策树,同样划分训练集与测试集比例为7:3,测试完毕同时输出整体均方根误差。

1.3.4 代理模型预测与应用

由于代理模型在开始时可能未达到足够精度,所以在新一轮待测场景生成后,需要根据该模型的测试精度来选择是否执行代理模型。具体过程如图8所示,如果代理模型未训练或测试均方根误差大于上限值 M ,则表示精度过低,不可用来预测场景输出,直接将批次场景转入仿真测试。如果代理模型精度小于 M ,则代理模型可用,考虑到代理模型仍然存在一定预测精度误差,直接将代理模型预测出的场景视为仿真关键场景会造成场景测试结果的可信度下降,所以本文根据均方根误差值设计场景划分判据,具体介绍如下:代理运算 TTC_{\max}^{-1} 的关键性界限为 $(1.6 - E/2)$, E 为代理模型此阶段测试的均方根误差。如果预测出的 TTC_{\max}^{-1} 大于 $(1.6 - E/2)$,则表示预测为代理关键场景,需要通过仿真测试计算出准确风险值,否则表示为代理非关键场景,直接将代理模型预测结果作为风险值输出。软件在环仿真测试得出的结果数据视为真值,将其保存到测试场景库,可用于避免重复执行仿真测试,并为代理模型训练提供数据。

整体仿真测试过程和代理运行过程统一使用

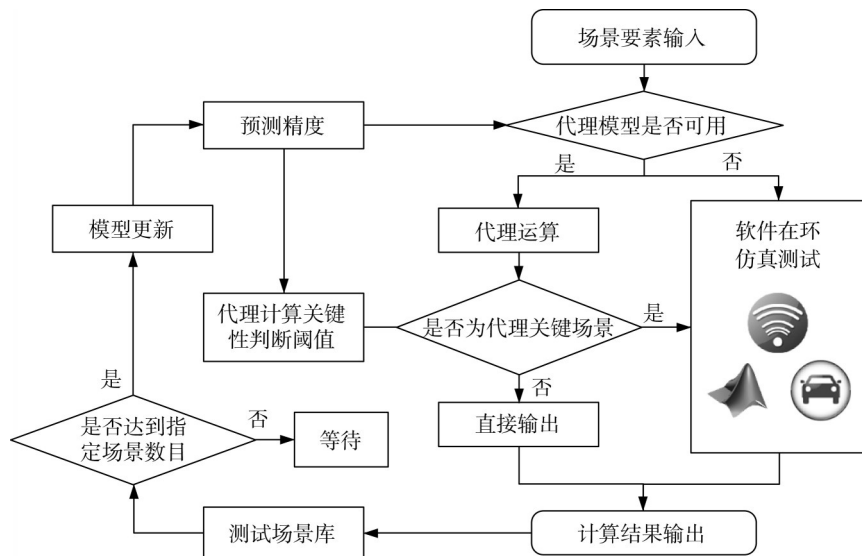


图8 代理模型应用流程

Fig.8 Application process of surrogate model

PreScan/Carsim/Simulink联合仿真平台,结合自动化测试代码与各部分算法代码实现自动化加速测试过程。

2 测试实验与结果分析

2.1 测试场景设计

在场景研究的不同阶段,对场景的抽象等级要求存在差距。在理论研究层面,场景应该采用抽象的,便于表述的逻辑语言来描述,以便于现实世界的语义化解释;而在仿真模拟层面,场景应该用详实、具体的参数来描述,以便于仿真软件的准确输入。据此,Menzel等^[15]将场景划分为功能场景、逻辑场景和具体场景3个等级。本节按照这种分级方式进行测试场景的设计。

2.1.1 功能场景设计

功能场景是最抽象级别的场景,通过语言来描述场景信息与驾驶任务。

实际驾驶场景往往是无限丰富且不可预测的,在跟车驾驶场景中,车辆往往在加速、减速和匀速等行驶工况中频繁切换以维持驾驶员的期望车速,前车突然减速是追尾的主要原因,也是AEB算法测试的主要场景^[16]。因此本文选取直线行驶下的前车减速跟驰场景作为功能场景,并对前车运动过程进行时序分解,将前车运动分解为先匀加速、再匀速、再匀减速3个阶段,同时引入降雨天气特征。

按照典型的场景要素分级结构,确定该场景的关键场景要素如下:主车初始车速、两车初始间距、前车初始车速、阶段1加速度、阶段2前车速度、阶段2匀速行驶距离、阶段3减速度、路面附着系数、场景降雨强度,通过这些要素可以完整表述测试场景。按照前文的要素层次分析结果,场景降雨强度属于天气要素,路面附着系数属于路面要素,两车初始间距和阶段2匀速行驶距离属于距离要素,主车和前车的初始车速属于速度要素,阶段1加速度和阶段3减速度属于加速度要素,采样分区数量按照表2查询。

为了简化仿真场景的构建过程,引入时间要素,结合运动学原理,将阶段2前车速度替换为阶段1持续时间,将阶段2匀速行驶距离替换为阶段2持续时间。阶段1持续时间和阶段2持续时间均为与时间相关的场景要素,时间可以通过位置、速度、加速度之间的运算得出,与三者均存在内部关联性,因此选取其中最大归一化比例1作为时间要素归一比例,

考虑到时间要素的离散步长,将其基础分区数量设置为10。场景示意图和前车速度图像如图9a、9b所示,图中, v_{ego} 和 v_{start} 表示主车与前车初始速度, L 为两车初始相对距离, t_{state1} 和 t_{state2} 表示阶段1和阶段2的持续时间, a_{state1} 和 a_{state3} 表示阶段1和阶段3的加速度,可以通过此类场景表述更多复杂的运动工况,满足真实世界中的驾驶特征。

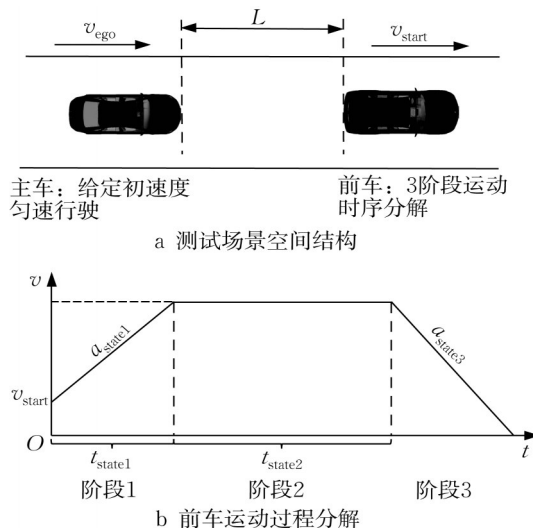


图9 前车变速跟驰场景示意图

Fig.9 Diagram of following scenario with variable speed front vehicle

2.1.2 逻辑场景设计

综合待测智能驾驶功能适用范围与相关测试规程^[17],确定逻辑场景参数空间如下:

主车初始车速 $v_{ego}=[20, 80]\text{km}\cdot\text{h}^{-1}$; 两车初始间距 $L=[10, 60]\text{m}$; 前车初始车速 $v_{start}=[20, 80]\text{km}\cdot\text{h}^{-1}$; 阶段1加速度 $a_{state1}=[1, 10]\text{m}\cdot\text{s}^{-2}$; 阶段1持续时间 $t_{state1}=[0, 5]\text{s}$; 阶段2持续时间 $t_{state2}=[0, 5]\text{s}$; 阶段3加速度 $a_{state3}=[-10, -1]\text{m}\cdot\text{s}^{-2}$; 路面附着系数 $\mu=[0.1, 0.9]$; 场景降雨强度 $I_{rain}=[0, 100]\text{mm}\cdot\text{h}^{-1}$ 。由此确定了一个由9个场景要素构成的庞大测试场景参数空间。

2.1.3 具体场景确定

具体场景的参数空间应为满足真实驾驶数据特征的分布形式,考虑到本文仅针对加速测试算法进行验证,所以采用离散均匀分布的形式。在场景元素范围与约束条件确定后,本节采用离散方式确定具体要素取值。本文距离离散步长为1m,速度离散步长为 $4\text{km}\cdot\text{h}^{-1}$,加速度离散步长为 $1\text{m}\cdot\text{s}^{-2}$,时间离散步长为0.5s,路面附着系数离散步长为0.05,降雨强度离散步长 $5\text{mm}\cdot\text{h}^{-1}$ 。

2.2 实验设计与执行

2.2.1 实验方案设置

设定种群数量50,迭代次数50,对功能场景执行SGO算法。以蒙特卡洛测试(Monte Carlo sampling, MC)、传统遗传算法(genetic algorithm, GA)为基本对照组。同时,为验证采样算法与代理模型的应用效果,加设MC+代理模型、GA+代理模型两个附加对照组。由于优化搜索算法会生成大量超出参数空间个体,为避免此类无意义个体浪费运算资源,对其进行标准化处理,即处于同一区间内的以最近区间端点值代替。

各实验方案分别为:①SGO算法(实验组),参数采样模块+遗传寻优模块+代理筛选模块。②MC算法(对照组),单一的蒙特卡洛采样测试。③MC算法+代理模型(对照组),在蒙特卡洛采样的基础上融合本文的代理筛选模块。④GA算法(对照组),基础的选择、交叉、变异算子,没有重复度筛选和分散采样功能。若算法连续两轮迭代未发现新的关键场景,则表示算法陷入局部最优,则重启算法进行重新运算。⑤GA算法+代理模型(对照组),在传统遗传算法的基础上融合本文的代理筛选模块。算法重启机制与④一致。各组别算法的结束条件均为达到设定的迭代次数。

2.2.2 实验结果与分析

为验证各算法的场景生成速度,将①~⑤5个组别每轮迭代生成的关键场景数据绘制如图10所示。

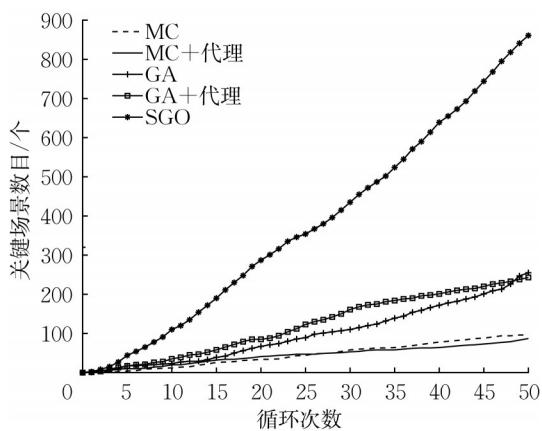


图10 关键测试场景生成过程示意图

Fig.10 Generation process of key test scenario

图10中横坐标表示整体的测试循环迭代次数(一次迭代理论上测试50次),纵坐标表示每轮迭代找到的关键场景数目。由图10中趋势可知,MC算法能够保持稳定探索过程,会不断地执行仿真测

试过程,但由于智能驾驶测试过程的“长尾效应”,算法探索到的关键场景数目极少。MC+代理中代理模型对MC算法整体测试过程的影响较小,整体场景生成效率变化趋势与单一MC算法类似。GA算法在探索前期具有较好的关键场景发掘能力,能比MC算法更快找到关键场景,但随着优化过程的执行,GA算法无法持续高效发掘关键场景,降低了整体寻优能力。GA+代理算法在测试中期小幅度提升了场景生成数量,但最终的整体关键场景数量与单一GA算法接近。SGO算法在测试全阶段均能表现出最好的探索能力和寻优能力,在每轮迭代中均可以通过仿真测试发掘新的关键测试场景。

为了分析SGO算法对于GA算法局部收敛性的影响,将包含遗传算法的①、④、⑤组别每轮迭代过程中种群个体的最高重复度数据绘制如图11所示。

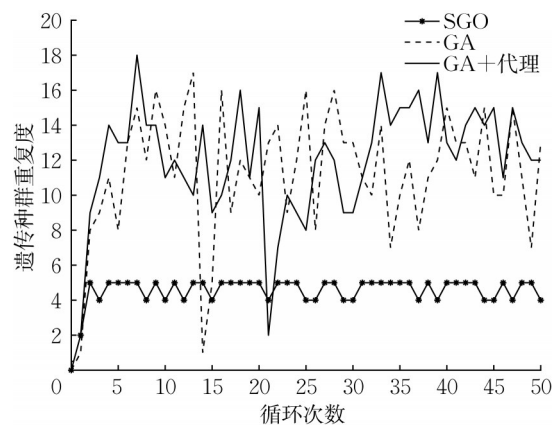


图11 遗传种群最高重复度示意图

Fig.11 Highest repeatability of genetic population

从图11可以看出,GA算法和GA+代理算法在探索过程中,种群中个体趋于重复,最高重复度会急剧增加,导致算法陷入局部收敛,探索性能变差,虽然在过程中可以通过算法重启实现重新探索,但由于没有利用前代种群的探索特征,整体场景生成效率相对较低。SGO算法通过分散采样和重复度筛选机制,可以有效降低种群内的重复个体数目,即改善传统遗传算法的局部收敛问题,提升场景生成效率。为了更准确地分析算法寻优效果,将各组别的测试数据记录如表3所示。表3中,结果运算次数表示累计进行测试运算的次数,单独仿真、代理运算为非关键、代理运算为关键并再次仿真都称为运算一次。关键场景数目表示累计找到的关键场景数目。关键场景比例表示整体测试过程中关键场景所占比例。

表3 关键场景生成数据对比

Tab.3 Comparison of key scenarios generation result

算法	MC	MC+代理	GA	GA+代理	SGO
结果运算次数	2 482	2 492	944	883	1 364
关键场景数目	96	87	255	243	861
关键场景比例/%	3.87	3.49	27.01	27.52	63.12

由表 3 可知,整体数据层面上,MC 算法和 MC+代理算法由于其随机性,采样场景重复性极低,整体结果运算次数高,但关键场景比例仅有 3.87% 和 3.49%,反映了测试过程的“长尾效应”。相较于 MC 算法,GA 算法和 GA+代理算法可以提升部分关键场景发掘数量,但由于种群中存在大量重复个体,尽管通过算法重启机制和多种遗传算子分别实现了 944 次和 883 次测试结果运算,但结果运算次数和关键场景比例均低于 SGO 算法。SGO 算法重点关注关键测试场景,结合重复度筛选机制可以降低种群重复度,从而提升结果运算次数,通过全面利用数据探索特征可以大幅度提升关键场景的测试比例,具有最佳的寻优效果。各组别每次迭代测试的时间消耗数据如图 12 所示。

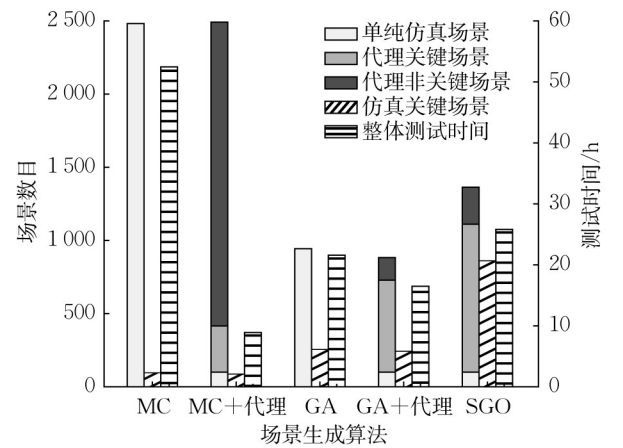


图 12 整体测试时间数据示意图
Fig.12 Diagram of overall testing time

为了更清晰地展现各算法的时间效率,定义探索效率为平均每小时找到的关键场景数目,通过该指标可以清晰展现各算法在时间消耗层面的加速测试效果,综合反映关键场景生成效率和整体测试时间。各算法探索效率数据如表 4 所示。

表 4 整体探索效率数据对比

Tab.4 Comparison of overall exploration efficiency

算法	MC	MC+代理	GA	GA+代理	SGO
探索效率	1.83	9.76	11.81	14.74	33.34

结合表 3、表 4 和图 12 分析可知,本文的代理模型可以一定程度上提升场景生成算法的探索效率,原先需要不断执行仿真测试的过程部分被代理模型代替,整体测试时间大幅度减少。从数据上看,MC 算法与 GA 算法的探索效率分别从 1.83 和 11.81 提升至 9.76 和 14.74。SGO 算法具有最佳的寻优能力,可以与代理模型进行充分结合,探索效率为 33.34,显著高于其他对照组。①、③、⑤组别的代理模型训练数据如表 5 所示。

表 5 代理模型训练数据对比

Tab.5 Comparison of surrogate model training

算法	MC+代理	GA+代理	SGO
有效训练量	400	700	1 100
训练次数	4	7	11
最高测试精度/%	87.08	89.34	91.26
代理关键数目	316	629	1 011
真实关键数目	83	236	853
关键场景查准率/%	26.27	43.43	84.37

表 5 中,代理模型的最高测试精度表示为自身训练过程中的测试阶段的最高精度,关键场景查准率表示真实仿真后的关键场景数目与代理关键场景数目的比值。

代理模型训练后,数据乱序补充时的整体测试精度受到训练次数和数据特征的影响。由于 MC 算法的随机探索特征和 GA 算法的局部收敛特征,算法探索到的关键场景较少,在限定迭代次数内,无法获得大量的仿真关键场景,因此代理模型训练次数较少。SGO 算法由于采样区域调整和重复度筛选机制,提升了探索能力和寻优能力,代理模型可以持续获得关键场景的仿真数据用于训练,测试精度高于其他对照组。

对比代理模型预测的关键场景和实际仿真后的真实关键场景数目可知,SGO 算法可以快速找到关键测试场景,同时补充大量关键场景的训练数据,所以代理模型的关键场景查准率显著高于 MC 算法和 GA 算法,在关键场景测试结果预测方面具有较高的应用价值。

3 结论

本文针对现有智能驾驶系统测试在高维场景领域面临的局部收敛性和测试成本高等问题,提出了一种代理遗传优化算法,对场景采样、遗传寻优、代理筛选模块进行融合设计,实现了测试效率与优化效果的协同提升。将该方法与蒙特卡洛采样、传统遗传算法、蒙特卡洛采样与代理模型结合、传统遗传算法与代理模型结合4种方法进行对比实验。结果表明,本文提出的代理遗传优化方法可以高效、迅速地探索关键具体测试场景,并提升代理模型对于关键场景的拟合精度,可以加速智能驾驶系统测试进程。同时,本文提出的种群重复度筛选思想和代理模型应用框架也可扩展用于其他种群优化算法。

作者贡献声明:

朱冰:命题提出与构思,仿真测试方案制定,论文修订。
 汤瑞:算法设计与优化,测试结果分析,论文撰写。
 赵健:仿真测试方案优化,论文修订与审核。
 张培兴:算法调试与仿真测试,结果分析,论文修订。
 李文旭:算法调试与仿真测试,论文修订。

参考文献:

- [1] 赵树廉,来飞,李志强,等.基于数字孪生技术的智能汽车测试方法研究[J].汽车工程,2023,45(1):42.
 ZHAO Shulian, LAI Fei, LI Keqiang, *et al.* Research on intelligent vehicle test method based on digital twin technology [J]. Automotive Engineering, 2023, 45(1): 42.
- [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.
 ZHU Bing, ZHANG Peixing, ZHAO Jian, *et al.* Review of scenario-based virtual validation methods for automated vehicles [J]. China Journal of Highway and Transport, 2019, 32(6): 1.
- [3] KALRA N, PADDOCK S M. Driving to safety: how many miles of driving would it take to demonstrate autonomous vehicle reliability? [J]. Transportation Research Part A: Policy & Practice, 2016, 94: 182.
- [4] 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.
 DENG Weiwen, LI Jiangkun, REN Bingtao, *et al.* A survey on automatic simulation scenario generation methods for autonomous driving [J]. China Journal of Highway and Transport, 2022, 35(1): 316.
- [5] ZHAO D, HUANG X, PENG H, *et al.* Accelerated evaluation of automated vehicles in car-following maneuvers [J]. IEEE Transactions on Intelligent Transportation Systems, 2018, 19(3): 733.
- [6] FENG S, SUN H, YAN X, *et al.* Dense reinforcement learning for safety validation of autonomous vehicles [J]. Nature, 2023, 615: 620.
- [7] 邢星宇,吴旭阳,刘力豪,等.基于目标优化的自动驾驶决策规划系统自动化测试方法[J].同济大学学报(自然科学版),2021,49(8):1162.
 XING Xingyu, WU Xuyang, LIU Lihao, *et al.* Automated testing method for autonomous driving decision planning system based on objective optimization [J]. Journal of Tongji University (Natural Science), 2021, 49(8): 1162.
- [8] KLÜCK F, ZIMMERMANN M, WOTAWA F, *et al.* Genetic algorithm-based test parameter optimization for ADAS system testing [C]// 2019 IEEE 19th International Conference on Software Quality, Reliability and Security (QRS). Sofia: IEEE, 2019: 418-425.
- [9] 张培兴,邱彬,朱冰,等.自动驾驶系统并行加速测试方法研究[J].汽车工程,2022,44(2):208.
 ZHANG Peixing, QIU Bin, ZHU Bing, *et al.* Study on parallel accelerated testing method for automated driving system [J]. Automotive Engineering, 2022, 44(2): 208.
- [10] FORRESTER A I J, SOBESTER A, KEANE A J. Engineering design via surrogate modelling: a practical guide [M]. Chichester: John Wiley and Sons, 2008.
- [11] VIANA F A C, VENTER G, BALABANOV V. An algorithm for fast optimal Latin hypercube design of experiments[J]. International Journal for Numerical Methods in Engineering, 2010, 82(2):135.
- [12] ZANG S, DING M, SMITH D, *et al.* The impact of adverse weather conditions on autonomous vehicles: how rain, snow, fog, and hail affect the performance of a self-driving car [J]. IEEE Vehicular Technology Magazine, 2019, 14(2): 103.
- [13] GAO F, HAN Z, ZHOU J, *et al.* Performance limit evaluation by evolution test with application to automatic parking system [J]. IEEE Transactions on Intelligent Vehicles, 2023, 8(4): 3096.
- [14] 周志华.机器学习[M].北京:清华大学出版社,2016.
 ZHOU Zhihua. Machine learning [M]. Beijing: Tsinghua University Press, 2016.
- [15] MENZEL T, BAGSCHIK G, MAURER M. Scenarios for development, test and validation of automated vehicles [C]// 2018 IEEE Intelligent Vehicles Symposium (IV). Changshu: IEEE, 2018: 1821-1827.
- [16] 段顺昌,白先旭,石琴,等.汽车自动紧急制动系统控制策略的预期功能安全设计[J].汽车工程,2022,44(9):1305.
 DUAN Shunchang, BAI Xianxu, SHI Qin, *et al.* Anticipated functional safety design of control strategy for automatic emergency braking system [J]. Automotive Engineering, 2022, 44(9): 1305.
- [17] 中华人民共和国工业和信息化部.乘用车自动紧急制动系统(AEBS)性能要求及试验方法:GB/T 39901—2021[S].北京:中国标准出版社,2021.
 Ministry of Industry and Information Technology of the People's Republic of China. Performance requirements and test methods for advanced emergency braking system (AEBS) for passenger car: GB/T 39901—2021[S]. Beijing: China Standard Press, 2021.