



(12) 发明专利申请

(10) 申请公布号 CN 118297143 A

(43) 申请公布日 2024. 07. 05

(21) 申请号 202410381743.1

(22) 申请日 2024.03.29

(71) 申请人 清华大学

地址 100084 北京市海淀区清华园

(72) 发明人 封硕 杨敬轩 张毅 白若瑄

何泓霖 李述 姚丹亚 裴欣

(74) 专利代理机构 北京安信方达知识产权代理

有限公司 11262

专利代理师 李丹 栗若木

(51) Int. Cl.

G06N 5/01 (2023.01)

G01M 17/007 (2006.01)

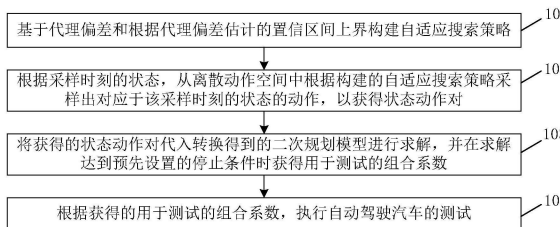
权利要求书3页 说明书12页 附图7页

(54) 发明名称

一种自适应测试策略调整方法及装置

(57) 摘要

本申请公开一种自适应测试策略调整方法及装置,本申请通过自适应搜索策略搜索到的状态动作对,平衡状态动作空间的探索与代理偏差并考虑自动驾驶环境中动作出现概率,提高了多个代理模型组合系数的调整效率,基于确定的组合系数执行自动驾驶汽车的测试,提高了自动驾驶汽车测试的鲁棒性以及测试评估的效率。



1. 一种自适应测试策略调整方法,其特征在于,包括:

基于代理偏差和根据代理偏差估计的置信区间上界构建自适应搜索策略,其中,代理偏差为自动驾驶汽车与多个代理模型之间安全性的差异,自适应搜索策略用于搜索平衡以下处理和参数:探索、代理偏差、以及自然驾驶环境中动作出现概率;

根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对;

将获得的状态动作对代入转换得到的二次规划模型进行求解,并在求解达到预先设置的停止条件时获得用于测试的组合系数,其中,二次规划模型为由多个代理模型组合系数的调整模型转换得到的关于多个代理模型组合系数的二次规划模型;

根据获得的用于测试的组合系数,执行自动驾驶汽车的测试。

2. 根据权利要求1所述的自适应测试策略调整方法,其特征在于,所述二次规划模型,由所述调整模型基于最优的重要性采样函数的组成结构确定。

3. 根据权利要求2所述的自适应测试策略调整方法,其特征在于,所述二次规划模型的表达式为:

$$\begin{aligned} \min_{\alpha} \quad & \frac{1}{2} \sum_{s \in \mathcal{S}, a \in \mathcal{A}} [Q^*(s, a) - Q_{\alpha}(s, a)]^2 \\ \text{s.t.} \quad & \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j = 1, \dots, J \end{aligned} \quad (1);$$

其中,  $\mathcal{S}$  表示离散状态空间,  $\mathcal{A}$  表示离散动作空间,  $Q^*(s, a)$  表示最优的兴趣事件发生概率,  $Q_{\alpha}(s, a) = \sum_{j=1}^J \alpha_j Q_j(s, a)$  表示所述代理模型的混合兴趣事件发生概率,  $Q_j(s, a)$  表示第  $j$  个所述代理模型的兴趣事件发生概率,  $J$  表示所述代理模型的总数,  $\alpha$  表示组合系数,  $a$  表示  $\mathcal{A}$  中的动作变量,  $s$  表示  $\mathcal{S}$  中的状态变量。

4. 根据权利要求3所述的自适应测试策略调整方法,其特征在于,所述自适应搜索策略的表达式为:

$$\eta(a | s) = \begin{cases} 1, & a = \arg \max_{a' \in \mathcal{A}} U(s, a) \phi(a | s), \\ 0, & \text{otherwise.} \end{cases} \quad (2);$$

其中,  $U(s, a) = g(Q \| Q_{\alpha})(s, a) + c \sqrt{\sum_{a' \in \mathcal{A}} N(s, a') / [1 + N(s, a)]}$  表示代理偏差置信区间上界,用于平衡探索与代理偏差,  $g(Q \| Q_{\alpha})$  表示代理偏差,

$$g(Q \| Q_{\alpha}) = \begin{cases} |Q - Q_{\alpha}| / Q_{\alpha}, & Q_{\alpha} > 0, \\ 0, & Q = Q_{\alpha} = 0, \\ +\infty, & Q > Q_{\alpha} = 0. \end{cases} \quad c \text{ 为探索与代理偏差的平衡系数,}$$

$\sqrt{\sum_{a' \in \mathcal{A}} N(s, a') / [1 + N(s, a)]}$  表示探索,  $N(s, a)$  表示状态动作对  $(s, a)$  的搜索次数,  $\phi(a | s)$  为背景车辆的自然驾驶策略;  $\arg \max_{a' \in \mathcal{A}} U(s, a) \phi(a | s)$  表示使得代理偏差置信区间上界与自然驾驶环境中动作出现概率的乘积取值最大的离散动作空间中的动作。

5. 根据权利要求4所述的自适应测试策略调整方法,其特征在于,所述根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对,包括:

首次搜索所述状态动作对时:均匀采样关键状态集合,获得初始状态 $s$ ,将初始状态 $s$ 作为采样时刻的状态通过自适应搜索策略从离散动作空间中采样获得对应于采样时刻的状态的动作 $a$ ,根据初始状态 $s$ 和采样获得的动作 $a$ 确定状态动作对 $(s, a)$ ;

第 $n$ 次搜索所述状态动作对时:通过在第 $n-1$ 次所述自适应搜索策略采样获得的动作 $a$ 观测第 $n$ 次搜索的状态 $s'$ ;将观测的状态 $s'$ 赋值为 $s$ ,通过第 $n$ 次所述自适应搜索策略采样获得对应于采样时刻的状态的动作 $a$ ,根据赋值的 $s$ 和第 $n$ 次搜索的动作 $a$ 确定状态动作对 $(s, a)$ , $n$ 大于或等于2。

6. 根据权利要求1所述的自适应测试策略调整方法,其特征在于,所述将获得的状态动作对代入转换得到的二次规划模型进行求解,包括:

通过凸优化求解器,对代入搜索到的所述状态动作对的所述二次规划模型进行求解。

7. 根据权利要求6所述的自适应测试策略调整方法,其特征在于,所述将搜索到的状态动作对代入转换获得的二次规划模型时,所述二次规划模型的表达式为:

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{(s,a) \in \mathcal{D}} [Q(s,a) - Q_{\alpha}(s,a)]^2 \\ \text{s.t.} & \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j=1, \dots, J \end{aligned} \quad (3);$$

其中, $\mathcal{D}$ 表示由搜索到的所述状态动作对构成的集合; $Q(s,a)$ 基于表格型密集强化学习方法得到; $Q_{\alpha}(s,a)$ 表示所述代理模型的混合兴趣事件发生概率。

8. 根据权利要求1所述的自适应测试策略调整方法,其特征在于,所述预设的求解停止条件,包括:

预先设定的平均滑动差分函数的取值小于预先设定的停止阈值;

其中,所述平均滑动差分函数为  $ASD(k) = \frac{1}{J} \sum_{j=1}^J \left| \sum_{k'=k-\Delta+1}^k [\alpha_j^{(k')} - \alpha_j^{(k'-\Delta)}] \right|$ ,  $k$ 表示测试次数,

$\Delta$ 为滑动步长, $\alpha_j^{(k)}$ 为第 $j$ 个代理模型的第 $k$ 次更新的组合系数。

9. 根据权利要求1所述的自适应测试策略调整方法,其特征在于,所述根据获得的用于测试的组合系数,执行自动驾驶汽车的测试,包括:

确定对应于所述用于测试的组合系数的混合的重要性采样函数;

根据确定的所述混合的重要性采样函数采样生成测试场景;

根据采样生成的所述测试场景对自动驾驶汽车进行测试。

10. 一种计算机存储介质,所述计算机存储介质中存储有计算机程序,所述计算机程序被处理器执行时实现如权利要求1至9中任一项所述的自适应测试策略调整方法。

11. 一种终端,包括:存储器和处理器,所述存储器中保存有计算机程序;其中,处理器被配置为执行存储器中的计算机程序;

所述计算机程序被所述处理器执行时实现如权利要求1至9中任一项所述的自适应测试策略调整方法。

12. 一种自适应测试策略调整装置,其特征在于,包括:构建单元、采样单元、求解单元和测试单元;其中,

构建单元,设置为:基于代理偏差和根据代理偏差估计的置信区间上界构建自适应搜索策略,其中,代理偏差为自动驾驶汽车与多个代理模型之间安全性的差异,自适应搜索策略用于搜索平衡以下处理和参数:探索、代理偏差、以及自然驾驶环境中动作出现概率;

采样单元,设置为:根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对;

求解单元,设置为:将获得的状态动作对代入转换得到的二次规划模型进行求解,并在求解达到预先设置的停止条件时获得用于测试的组合系数,其中,二次规划模型为由多个代理模型组合系数的调整模型转换得到的关于多个代理模型组合系数的二次规划模型;

测试单元,设置为:根据获得的用于测试的组合系数,执行自动驾驶汽车的测试。

## 一种自适应测试策略调整方法及装置

### 技术领域

[0001] 本申请涉及但不限于自动驾驶技术,其中涉及一种自适应测试策略调整方法及装置。

### 背景技术

[0002] 当前自动驾驶汽车面临严峻的安全问题,一些前沿的汽车车辆运营公司研发的自动驾驶汽车均遭遇了严重的交通事故,这些安全问题阻碍了自动驾驶汽车的规模化应用和商业化落地,因此,需对自动驾驶汽车进行测试评估。自动驾驶汽车的测试评估的基本流程包括:根据生成的一系列测试场景,对自动驾驶汽车进行测试;收集测试结果并对自动驾驶汽车的安全性进行评估;相关技术中的测试评估,通常可以得到事故率等测试指标的估计值。

[0003] 真实道路测试评估方法是相关技术中的测试评估方法,因为自然驾驶环境中的安全关键场景(如碰撞事故场景等)非常罕见,事故事件发生概率极低,存在真实道路测试的低效性问题。基于蒙特卡洛采样理论的真实道路测试评估方法,是自动驾驶汽车测试评估的一种方法,但是,由于真实道路测试的低效性问题,导致基于蒙特卡洛采样的测试方法需要非常多的测试次数才可以得到事故率的估计值,据测算一款自动驾驶汽车需要在自然驾驶环境中累积测试超过100亿公里,才能得到较高置信度的安全性能测试结果,而大量测试所消耗的时间成本和经济成本是不可接受的。重要性采样方法可以缓解自动驾驶汽车的事件事件发生概率极低的问题,其基本思想是提高那些容易发生事故的关键场景的采样概率,从而提高测试效率。但是,重要性采样方法面临维度灾难问题,无法应用于高维测试环境中。

[0004] 由于自动驾驶汽车系统的高度复杂性和所采用神经网络算法的黑箱性,利用单一代理模型(例如IDM、FVDM等模型)构造的重要性采样函数 $q$ 与多种类型的被测自动驾驶汽车对应的最优重要性采样函数 $q_1^*, q_2^*, \dots$ 之间存在很大差距,导致基于重要性采样方法的测试效率存在提升瓶颈,自适应测试评估方法则致力于解决这个问题;自适应测试评估方法主要包括如下三个阶段:第一阶段:在正式进行测评之前,通过对自动驾驶汽车进行少量测试得到测试结果,并利用这些测试结果优化重要性采样函数;第二阶段:使用优化后的重要性采样函数对自动驾驶汽车进行测试,得到测试结果;第三阶段:根据稀疏对抗采样方法得到测试指标的估计结果,直到测试指标满足预先设定的收敛条件。自适应测试策略调整方法是自适应测试评估方法的一种调整方法,自适应测试策略调整方法主要应用于自适应测试评估方法第一阶段。

[0005] 上述自适应测试评估方法只能用于简单低维的测试场景中,不能用于高维测试环境;如何获得可以应用于高维测试环境的测试评估方法,提高测试评估的效率,成为一个有待解决的问题。

## 发明内容

[0006] 以下是对本申请详细描述的主题的概述。本概述并非是为了限制权利要求的保护范围。

[0007] 本公开实施例提供一种自适应测试策略调整方法及装置,能够实现自动驾驶汽车的高维测试环境的测试评估,提高测试评估效率。

[0008] 本公开实施例提供了一种自适应测试策略调整方法,包括:

[0009] 基于代理偏差和根据代理偏差估计的置信区间上界构建自适应搜索策略,其中,代理偏差为自动驾驶汽车与多个代理模型之间安全性的差异,自适应搜索策略用于搜索平衡以下处理和参数:探索、代理偏差、以及自然驾驶环境中动作出现概率;

[0010] 根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对;

[0011] 将获得的状态动作对代入转换得到的二次规划模型进行求解,并在求解达到预先设置的停止条件时获得用于测试的组合系数,其中,二次规划模型为由多个代理模型组合系数的调整模型转换得到的关于多个代理模型组合系数的二次规划模型;

[0012] 根据获得的用于测试的组合系数,执行自动驾驶汽车的测试。

[0013] 另一方面,本公开实施例还提供一种计算机存储介质,所述计算机存储介质中存储有计算机程序,所述计算机程序被处理器执行时实现上述自适应测试策略调整方法。

[0014] 再一方面,本公开实施例还提供一种终端,包括:存储器和处理器,所述存储器中保存有计算机程序;其中,

[0015] 处理器被配置为执行存储器中的计算机程序;

[0016] 所述计算机程序被所述处理器执行时实现如上述自适应测试策略调整方法。

[0017] 与相关技术相比,本公开实施例可应用于自动驾驶汽车的高维测试环境的测试评估,可对多种类型被测自动驾驶汽车进行测试,通过自适应搜索策略搜索到的状态动作对,平衡探索、代理偏差、以及自然驾驶环境中动作出现概率,提高了组合系数的调整效率和自动驾驶汽车测试的鲁棒性,基于组合确定的组合系数执行自动驾驶汽车的测试,提高了测试评估效率。

[0018] 本申请的其它特征和优点将在随后的说明书中阐述,并且,部分地从说明书中变得显而易见,或者通过实施本申请而了解。本申请的其他优点可通过在说明书以及附图中所描述的方案来实现和获得。

## 附图说明

[0019] 附图用来提供对本申请技术方案的理解,并且构成说明书的一部分,与本申请的实施例一起用于解释本申请的技术方案,并不构成对本申请技术方案的限制。

[0020] 图1为本公开实施例自适应测试策略调整方法的流程图;

[0021] 图2为本公开实施例一种自适应测试策略调整装置的结构框图;

[0022] 图3为本公开实施例自动驾驶汽车超车场景的示意图;

[0023] 图4(a)为本公开实施例AV-I对应的自然驾驶环境与自然对抗驾驶环境事故率对比图;

[0024] 图4(b)为本公开实施例AV-II对应的模型自然驾驶环境与自然对抗驾驶环境事故

率对比图；

[0025] 图4(c)为本公开实施例AV-III对应的模型自然驾驶环境与自然对抗驾驶环境事故率对比图；

[0026] 图4(d)为本公开实施例AV-I对应的自然驾驶环境与自然对抗驾驶环境相对半宽对比图；

[0027] 图4(e)为本公开实施例AV-II对应的模型自然驾驶环境与自然对抗驾驶环境相对半宽对比图；

[0028] 图4(f)为本公开实施例AV-III对应的模型自然驾驶环境与自然对抗驾驶环境相对半宽对比图；

[0029] 图5(a)为本公开实施例AV-I对应的组合系数与测试次数关系的对比图；

[0030] 图5(b)为本公开实施例AV-II对应的组合系数与测试次数关系的对比图；

[0031] 图5(c)为本公开实施例AV-III对应的组合系数与测试次数关系的对比图；

[0032] 图5(d)为本公开实施例AV-I对应的平均滑动差异与测试次数关系的对比图；

[0033] 图5(e)为本公开实施例AV-II对应的平均滑动差异与测试次数关系的对比图；

[0034] 图5(f)为本公开实施例AV-III对应的平均滑动差异与测试次数关系的对比图；

[0035] 图6(a)为AV-I对应的本公开实施例方法与自然对抗驾驶环境事故率对比图；

[0036] 图6(b)为AV-II对应的本公开实施例方法与自然对抗驾驶环境事故率对比图；

[0037] 图6(c)为AV-III对应的本公开实施例方法与自然对抗驾驶环境事故率对比图；

[0038] 图6(d)为AV-I对应的本公开实施例方法与自然对抗驾驶环境相对半宽对比图；

[0039] 图6(e)为AV-II对应的本公开实施例方法与自然对抗驾驶环境相对半宽对比图；

[0040] 图6(f)为AV-III对应的本公开实施例方法与自然对抗驾驶环境相对半宽对比图；

[0041] 图6(g)为AV-I对应的本公开实施例方法与自然对抗驾驶环境所需测试次数对比图；

[0042] 图6(h)为AV-II对应的本公开实施例方法与自然对抗驾驶环境所需测试次数对比图；

[0043] 图6(i)为AV-III对应的本公开实施例方法与自然对抗驾驶环境所需测试次数对比图。

### 具体实施方式

[0044] 本申请描述了多个实施例,但是该描述是示例性的,而不是限制性的,并且对于本领域的普通技术人员来说显而易见的是,在本申请所描述的实施例包含的范围内可以有更多的实施例和实现方案。尽管在附图中示出了许多可能的特征组合,并在具体实施方式中进行了讨论,但是所公开的特征的许多其它组合方式也是可能的。除非特意加以限制的情况以外,任何实施例的任何特征或元件可以与任何其它实施例中的任何其他特征或元件结合使用,或可以替代任何其它实施例中的任何其他特征或元件。

[0045] 本申请包括并设想了与本领域普通技术人员已知的特征和元件的组合。本申请已经公开的实施例、特征和元件也可以与任何常规特征或元件组合,以形成由权利要求限定的独特的发明方案。任何实施例的任何特征或元件也可以与来自其它发明方案的特征或元件组合,以形成另一个由权利要求限定的独特的发明方案。因此,应当理解,在本申请中示

出和/或讨论的任何特征可以单独地或以任何适当的组合来实现。因此,除了根据所附权利要求及其等同替换所做的限制以外,实施例不受其它限制。此外,可以在所附权利要求的保护范围内进行各种修改和改变。

[0046] 此外,在描述具有代表性的实施例时,说明书可能已经将方法和/或过程呈现为特定的步骤序列。然而,在该方法或过程不依赖于本文所述步骤的特定顺序的程度上,该方法或过程不应限于所述的特定顺序的步骤。如本领域普通技术人员将理解的,其它的步骤顺序也是可能的。因此,说明书中阐述的步骤的特定顺序不应被解释为对权利要求的限制。此外,针对该方法和/或过程的权利要求不应限于按照所写顺序执行它们的步骤,本领域技术人员可以容易地理解,这些顺序可以变化,并且仍然保持在本申请实施例的精神和范围内。

[0047] 图1为本公开实施例自适应测试策略调整方法的流程图,如图1所示,包括:

[0048] 步骤101、基于代理偏差和根据代理偏差估计的置信区间上界构建自适应搜索策略,其中,代理偏差为自动驾驶汽车与多个代理模型之间安全性的差异,自适应搜索策略用于搜索平衡以下处理和参数:探索、代理偏差、以及自然驾驶环境中动作出现概率。

[0049] 本申请实施例中,探索为自动驾驶汽车的测试评估中的已有术语,是指以更大概率采样探索次数较少的动作,表达式可以为 $\sqrt{\sum_{a \in \mathcal{A}} N(s, a)} / [1 + N(s, a)]$ ,探索次数越少的动作,此表达式的值越大。

[0050] 步骤102、根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对;

[0051] 步骤103、将获得的状态动作对代入转换得到的二次规划模型进行求解,并在求解达到预先设置的停止条件时获得用于测试的组合系数,其中,二次规划模型为由多个代理模型组合系数的调整模型转换得到的关于多个代理模型组合系数的二次规划模型;

[0052] 步骤104、根据获得的用于测试的组合系数,执行自动驾驶汽车的测试。

[0053] 本公开实施例可应用于自动驾驶汽车的高维测试环境的测试评估,可对多种类型被测自动驾驶汽车进行测试,本申请通过构建的自适应搜索策略搜索到的状态动作对,平衡探索、代理偏差、以及自然驾驶环境中动作出现概率,提高了组合系数的调整效率和自动驾驶汽车测试的鲁棒性,基于组合确定的组合系数执行自动驾驶汽车的测试,提高了测试评估效率。

[0054] 在一种示例性实例中,本公开实施例二次规划模型(quadratic programming),由调整模型基于最优的重要性采样函数的组成结构确定。

[0055] 本公开实施例,多个代理模型组合系数的调整模型与二次规划模型均为本领域技术人员公知的技术术语,本公开实施例对此不作赘述。在一种示例性实例中,本公开实施例中多个代理模型组合系数的调整模型,可以参照相关技术,由自适应测试策略调整方法通过公式(4)建模获得,多个代理模型组合系数的调整模型为:

$$[0056] \quad \min_{q_\alpha \in \mathcal{Q}_J} \text{Var}_{q_\alpha} \left( \frac{\mathbb{I}_F(X)p(X)}{q_\alpha(X)} \right) \quad (4)$$

[0057] 公式(4)中, $\mathcal{Q}_J = \left\{ q_\alpha \in \mathcal{Q} : \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j=1, \dots, J \right\}$ , $\mathcal{Q}$ 为全部重要性采样函数的集合, $\alpha = [\alpha_1, \dots, \alpha_J]^T$ 为组合系数, $q_\alpha$ 为混合的重要性采样函数,混合的重要性采样函



数相应的混合重要性策略为  $\psi_a(a|s) = \sum_{j=1}^J \alpha_j \psi_j(a|s)$ ,  $\psi_j(a|s)$  为第  $j$  个重要性策略 (Importance policy) (重要性策略可以通过代理模型构建, 可以采用文献[1]中公开的方法获取),  $F$  表示兴趣事件,  $p$  表示测试场景  $X$  的自然概率分布,  $\mathbb{I}_F(X)$  为兴趣事件  $F$  的指示函数,  $\mathbb{I}_F(X) = \begin{cases} 1, & X \in F, \\ 0, & X \notin F. \end{cases}$   $a$  为动作变量 (可以包括驾驶环境中背景汽车的加速度等信息),  $s$

为状态变量 (可以包括驾驶环境中全部车辆的位置、速度等信息),  $J \geq 1$  为全部代理模型的数量, 重要性采样函数为未优化的重要性采样函数; 其中, 文献[1]为 Feng, S., Yan, X., Sun, H., Feng, Y. and Liu, H.X. 基于自然对抗驾驶环境的自动驾驶汽车智能性测试, 自然·通信 (Intelligent driving intelligence test for autonomous vehicles with naturalistic and adversarial environment. Nature Communications), 12, 748 (2021).

[0058] 在一种示例性实例中, 本公开实施例多个代理模型包括两个以上代理模型; 其中, 真实模型为待测实际自动驾驶模型, 在每个场景  $x$  有测试指标  $\mathbb{P}^*(A|x)$ , 本公开实施例中的代理模型是对真实自动驾驶模型的一个估计, 在每个场景  $x$  有测试指标  $\mathbb{P}(A|x)$ , 本公开实施例实际自动驾驶测试指标  $\mathbb{P}^*(A|x)$ , 使用  $\mathbb{P}(A|x)$  作为其估计。

[0059] 在一种示例性实例中, 本公开实施例二次规划模型的表达式为:

$$[0060] \quad \min_{\alpha} \frac{1}{2} \sum_{s \in \mathcal{S}, a \in \mathcal{A}} [Q^*(s, a) - Q_{\alpha}(s, a)]^2$$

$$[0061] \quad \text{s.t.} \quad \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j = 1, \dots, J \quad (1);$$

[0062] 其中,  $\mathcal{S}$  表示离散状态空间,  $\mathcal{A}$  表示离散动作空间,  $Q^*(s, a)$  表示最优的兴趣事件发生概率,  $Q_{\alpha}(s, a) = \sum_{j=1}^J \alpha_j Q_j(s, a)$  表示代理模型的混合兴趣事件发生概率,  $Q_j(s, a)$  表示第  $j$  个代理模型的兴趣事件发生概率,  $J$  表示代理模型的总数,  $\alpha$  表示组合系数,  $a$  表示  $\mathcal{A}$  中的动作变量,  $s$  表示  $\mathcal{S}$  中的状态变量。

[0063] 在一种示例性实例中, 本公开实施例自适应搜索策略通过以下处理搜索状态动作对中的动作: 代理偏差乘以自然驾驶环境中动作出现概率, 加上该自然驾驶环境中动作出现概率乘以平衡探索与代理偏差的常数、再乘以根号下已搜索的状态动作对的搜索次数的和、得到的和值除以当前状态动作对的访问次数获得中间变量, 这一中间变量在离散动作空间中取值最大的动作; 上述自适应搜索策略涉及到的自然驾驶环境中动作出现概率为本领域技术人员公知的定义; 平衡探索与代理偏差的常数可以由本领域技术人员分析设定, 并根据搜索状态动作对结合经验进行调整。

[0064] 在一种示例性实例中, 本公开实施例自适应搜索策略的表达式为:

$$[0065] \quad \eta(a|s) = \begin{cases} 1, & a = \arg \max_{a' \in \mathcal{A}} U(s, a) \phi(a|s), \\ 0, & \text{otherwise.} \end{cases} \quad (2);$$

[0066] 其中,  $U(s, a) = g(Q \| Q_{\alpha})(s, a) + c \sqrt{\sum_{a' \in \mathcal{A}} N(s, a') / [1 + N(s, a)]}$  表示代理偏差置信区

间上界,用于平衡探索与代理偏差, $g(Q \| Q_\alpha)$ 表示代理偏差,

$$g(Q \| Q_\alpha) = \begin{cases} |Q - Q_\alpha| / Q_\alpha, & Q_\alpha > 0, \\ 0, & Q = Q_\alpha = 0, \\ +\infty, & Q > Q_\alpha = 0. \end{cases} \quad c \text{为探索与代理偏差的平衡系数,}$$

$\sqrt{\sum_{a \in \mathcal{A}} N(s, a) / [1 + N(s, a)]}$ 表示探索, $N(s, a)$ 表示状态动作对 $(s, a)$ 的搜索次数, $\phi(a | s)$ 为背景车辆(背景车辆为自动驾驶测试过程中,本领域技术人员公知的相对于自动驾驶车辆的其他车辆)的自然驾驶策略; $\arg \max_{a \in \mathcal{A}} U(s, a) \phi(a | s)$ 表示使得代理偏差置信区间上界与自然驾驶环境中动作出现概率的乘积取值最大的离散动作空间中的动作。

[0067] 本公开实施例探索次数越少的动作表达式 $\sqrt{\sum_{a \in \mathcal{A}} N(s, a) / [1 + N(s, a)]}$ 的值越大, $c$ 是平衡系数, $c$ 越大则探索程度越高; $U(s, a)$ 是平衡探索与代理偏差,然后 $U(s, a) \phi(a | s)$ 添加了对自然驾驶环境中动作出现概率的考虑,最终实现探索、代理偏差、以及自然驾驶环境中动作出现概率的综合考虑。

[0068] 在一种示例性实例中,本公开实施例代理模型的混合兴趣事件发生概率和第 $j$ 个代理模型的兴趣事件发生概率,可以使用表格型密集强化学习方法获取,将代理模型的混合兴趣事件发生概率和第 $j$ 个代理模型的兴趣事件发生概率定义为未达到最优的兴趣事件的发生概率 $Q(s, a)$ ,则表格型密集强化学习方法,包括:初始化未达到最优的兴趣事件的发生概率 $Q(s, a)$ ,使得 $Q(s, a) = 0, \forall s \in \mathcal{S}, a \in \mathcal{A}$ ,其中, $\mathcal{S}$ 表示离散状态空间, $\mathcal{A}$ 表示离散动作空间;从关键状态集合中均匀采样初始状态,使得初始状态均为关键状态;根据学习率、时序差分误差及均匀采样后的关键状态集合的示性函数,对未达到最优的兴趣事件的发生概率 $Q(s, a)$ 进行更新学习;其中,根据学习率、时序差分误差及关键状态集合的示性函数,对未达到最优的兴趣事件的发生概率 $Q(s, a)$ 进行更新学习,包括基于以下表达式进行更新学习: $Q(s, a) \leftarrow Q(s, a) + \alpha \delta \mathbb{I}_{\mathcal{S}_c}(s)$ ;其中, $\alpha$ 为学习率, $\delta$ 为时序差分误差, $\mathcal{S}_c$ 表示关键状态集合, $\mathbb{I}_{\mathcal{S}_c}(s)$ 为关键状态集合的示性函数;当更新学习的时长 $t$ 小于预先设定的最大时长 $T$ 时,以相同的概率探索离散动作空间中的每个动作,以确定用于更新学习的动作,及搜索每一个动作后对应的观测到的状态及基于奖励函数获得的对应于该动作的奖励;其中,离散动作空间为马尔可夫决策过程中的参数。

[0069] 在一种示例性实例中,本公开实施例步骤103根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对,包括:

[0070] 首次搜索状态动作对时:均匀采样关键状态集合,获得初始状态 $s$ ,将均匀采样初始状态 $s$ 作为采样时刻的状态,通过自适应搜索策略从离散动作空间中采样获得对应于采样时刻的状态的动作 $a$ ,根据初始状态 $s$ 和采样获得的动作 $a$ 确定状态动作对 $(s, a)$ ;

[0071] 第 $n$ 次搜索状态动作对时:通过第 $n-1$ 次自适应搜索策略采样获得的动作 $a$ 观测第 $n$ 次搜索的状态 $s'$ ;将观测的状态 $s'$ 赋值为 $s$ ,通过第 $n$ 次自适应搜索策略采样获得动作 $a$ ,根据赋值的 $s$ 和第 $n$ 次搜索的动作 $a$ 确定状态动作对 $(s, a)$ , $n$ 大于或等于2。

[0072] 在一种示例性实例中,本公开实施例将获得的状态动作对代入转换得到的二次规划模型进行求解,包括:

[0073] 通过凸优化求解器,对代入搜索到的状态动作对的二次规划模型进行求解。

[0074] 在一种示例性实例中,本公开实施例将获得的状态动作对代入转换得到的二次规划模型进行求解时,二次规划模型的表达式为:

$$\begin{aligned}
 & \min_{\alpha} \frac{1}{2} \sum_{(s,a) \in \mathcal{D}} [Q(s,a) - Q_{\alpha}(s,a)]^2 \\
 & \text{s.t.} \quad \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j = 1, \dots, J
 \end{aligned} \quad (3);$$

[0076] 其中,  $\mathcal{D}$  表示由搜索到的状态动作对构成的集合;  $Q(s,a)$  基于表格型密集强化学习方法得到;  $Q_{\alpha}(s,a)$  表示代理模型的混合兴趣事件发生概率。

[0077] 在一种示例性实例中,本公开实施例预设的求解停止条件,包括:

[0078] 预先设定的平均滑动差分函数的取值小于预先设定的停止阈值;

[0079] 其中,平均滑动差分函数为  $ASD(k) = \frac{1}{J} \sum_{j=1}^J \left| \sum_{k'=k-\Delta+1}^k [\alpha_j^{(k')} - \alpha_j^{(k'-\Delta)}] \right|$ , 式中,  $k$  表示测试

次数,  $\Delta$  为滑动步长,  $\alpha_j^{(k)}$  为第  $j$  个代理模型的第  $k$  次更新的组合系数。

[0080] 在一种示例性实例中,本公开实施例停止阈值可以由本领域技术人员分析设定,例如,可以设置停止阈值为 0.02。

[0081] 在一种示例性实例中,本公开实施例根据获得的用于测试的组合系数,执行自动驾驶汽车的测试,包括:

[0082] 确定对应于该用于测试的组合系数的混合的重要性采样函数;

[0083] 根据确定的混合的重要性采样函数采样生成测试场景;

[0084] 根据采样生成的测试场景对自动驾驶汽车进行测试。

[0085] 本公开实施例还提供一种计算机存储介质,计算机存储介质中存储有计算机程序,计算机程序被处理器执行时实现上述自适应测试策略调整方法。

[0086] 本公开实施例还提供一种终端,包括:存储器和处理器,存储器中保存有计算机程序;

[0087] 其中,

[0088] 处理器被配置为执行存储器中的计算机程序;

[0089] 计算机程序被处理器执行时实现如上述自适应测试策略调整方法。

[0090] 图2为本公开实施例一种自适应测试策略调整装置的结构框图,如图2所示,包括:构建单元、采样单元、求解单元和测试单元;其中,

[0091] 构建单元,设置为:基于代理偏差和根据代理偏差估计的置信区间上界构建自适应搜索策略,其中,代理偏差为自动驾驶汽车与多个代理模型之间安全性的差异,自适应搜索策略用于搜索平衡以下处理和参数:探索、代理偏差、以及自然驾驶环境中动作出现概率;

[0092] 采样单元,设置为:根据采样时刻的状态,从离散动作空间中根据构建的自适应搜索策略采样出对应于该采样时刻的状态的动作,以获得状态动作对;

[0093] 求解单元, 设置为: 将获得的状态动作对代入转换得到的二次规划模型进行求解, 并在求解达到预先设置的停止条件时获得用于测试的组合系数, 其中, 二次规划模型为由多个代理模型组合系数的调整模型转换得到的关于多个代理模型组合系数的二次规划模型;

[0094] 测试单元, 设置为: 根据获得的用于测试的组合系数, 执行自动驾驶汽车的测试。

[0095] 在一种示例性实例中, 本公开实施例二次规划模型, 由调整模型基于最优的重要性采样函数的组成结构确定。

[0096] 在一种示例性实例中, 本公开实施例二次规划模型的表达式为:

$$[0097] \quad \min_{\alpha} \frac{1}{2} \sum_{s \in \mathcal{S}, a \in \mathcal{A}} [Q^*(s, a) - Q_{\alpha}(s, a)]^2$$

$$[0098] \quad \text{s.t.} \quad \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j = 1, \dots, J \quad (1);$$

[0099] 其中,  $\mathcal{S}$  表示离散状态空间,  $\mathcal{A}$  表示离散动作空间,  $Q^*(s, a)$  表示最优的兴趣事件发生概率,  $Q_{\alpha}(s, a) = \sum_{j=1}^J \alpha_j Q_j(s, a)$  表示代理模型的混合兴趣事件发生概率,  $Q_j(s, a)$  表示第  $j$  个代理模型的兴趣事件发生概率,  $J$  表示代理模型的总数,  $\alpha$  表示组合系数,  $a$  表示  $\mathcal{A}$  中的动作变量,  $s$  表示  $\mathcal{S}$  中的状态变量。

[0100] 在一种示例性实例中, 本公开实施例自适应搜索策略的表达式为:

$$[0101] \quad \eta(a|s) = \begin{cases} 1, & a = \arg \max_{a' \in \mathcal{A}} U(s, a) \phi(a|s), \\ 0, & \text{otherwise.} \end{cases} \quad (2);$$

[0102] 其中,  $U(s, a) = g(Q \| Q_{\alpha})(s, a) + c \sqrt{\sum_{a' \in \mathcal{A}} N(s, a') / [1 + N(s, a)]}$  表示代理偏差置信区间上

界, 用于平衡探索与代理偏差,  $g(Q \| Q_{\alpha})$  表示代理偏差,  $g(Q \| Q_{\alpha}) = \begin{cases} |Q - Q_{\alpha}| / Q_{\alpha}, & Q_{\alpha} > 0, \\ 0, & Q = Q_{\alpha} = 0, \\ +\infty, & Q > Q_{\alpha} = 0. \end{cases};$

为探索与代理偏差的平衡系数,  $\sqrt{\sum_{a' \in \mathcal{A}} N(s, a') / [1 + N(s, a)]}$  表示探索,  $N(s, a)$  表示状态动作对  $(s, a)$  的搜索次数,  $\phi(a|s)$  为背景车辆的自然驾驶策略;  $\arg \max_{a' \in \mathcal{A}} U(s, a) \phi(a|s)$  表示使得代理偏差置信区间上界与自然驾驶环境中动作出现概率的乘积取值最大的离散动作空间中的动作。

[0103] 在一种示例性实例中, 本公开实施例采样单元是设置为:

[0104] 首次搜索状态动作对时: 均匀采样关键状态集合, 获得初始状态  $s$ , 将初始状态  $s$  作为采样时刻的状态通过自适应搜索策略从离散动作空间中采样获得对应于采样时刻的状态的动作  $a$ , 根据初始状态  $s$  和采样获得的动作  $a$  确定状态动作对  $(s, a)$ ;

[0105] 第  $n$  次搜索状态动作对时: 通过第  $n-1$  次自适应搜索策略采样获得的动作  $a$  观测第  $n$  次搜索的状态  $s'$ ; 将观测的状态  $s'$  赋值为  $s$ , 通过第  $n$  次自适应搜索策略采样获得对应于采样时刻的状态的动作  $a$ , 根据赋值的  $s$  和第  $n$  次搜索的动作  $a$  确定状态动作对  $(s, a)$ ,  $n$  大于或等于 2。

[0106] 在一种示例性实例中, 本公开实施例求解单元是设置为:

[0107] 通过凸优化求解器,对代入搜索到的状态动作对的二次规划模型进行求解。

[0108] 在一种示例性实例中,本公开实施例求解单元将搜索到的状态动作对代入转换获得的二次规划模型时,二次规划模型的表达式为:

$$\begin{aligned} & \min_{\alpha} \frac{1}{2} \sum_{(s,a) \in \mathcal{D}} [Q(s,a) - Q_{\alpha}(s,a)]^2 \\ & \text{s.t.} \sum_{j=1}^J \alpha_j = 1, \alpha_j \geq 0, j=1, \dots, J \end{aligned} \quad (3);$$

[0110] 其中,  $\mathcal{D}$  表示由搜索到的状态动作对构成的集合;  $Q(s,a)$  基于表格型密集强化学习方法得到;  $Q_{\alpha}(s,a)$  表示代理模型的混合兴趣事件发生概率。

[0111] 在一种示例性实例中,本公开实施例预设的求解停止条件,包括:

[0112] 预先设定的平均滑动差分函数的取值小于预先设定的停止阈值;

[0113] 其中,平均滑动差分函数为  $ASD(k) = \frac{1}{J} \sum_{j=1}^J \left| \sum_{k'=k-\Delta+1}^k [\alpha_j^{(k')} - \alpha_j^{(k'-\Delta)}] \right|$ , 式中,  $k$  表示测试次数,  $\Delta$  为滑动步长,  $\alpha_j^{(k)}$  为第  $j$  个代理模型的第  $k$  次更新的组合系数。

[0114] 在一种示例性实例中,本公开实施例测试单元是设置为:

[0115] 确定对应于用于测试的组合系数的混合的重要性采样函数;

[0116] 根据确定的混合的重要性采样函数采样生成测试场景;

[0117] 根据采样生成的测试场景对自动驾驶汽车进行测试。

[0118] 以下通过应用示例对本公开实施例进行简要说明,应用示例仅用于陈述本公开实施例,并不用于限定本公开实施例的保护范围。

[0119] 应用示例

[0120] 针对自适应测试策略调整方法存在的只能应用于低维测试环境且测试效率低的问题,本公开实施例将自适应测试策略调整方法建模为多个代理模型组合系数的调整模型;求解上述代理模型组合系数的调整模型,针对多种类型的被测自动驾驶汽车得到调整的组合系数  $\alpha^*$ , 然后使用  $q_{\alpha^*}$  采样生成测试场景。本公开实施例提出基于密集强化学习方法的自适应测试策略调整方法,可以应用于高维测试环境中,并针对多种类型的被测自动驾驶汽车生成定制化的测试策略,提高了测试评估的效率;本公开实施例自适应测试策略调整方法,利用多个代理模型来提高对多种类型被测自动驾驶汽车进行测试的鲁棒性,并通过调整这些代理模型的组合系数来提高测试评估的效率;本公开实施例方法能够对每种类型的被测自动驾驶汽车生成定制化的测试场景,提高测试评估的鲁棒性和高效性。

[0121] 本公开实施例自适应测试策略调整方法,包括:

[0122] 步骤一、将预先确定的多个代理模型组合系数的调整模型,基于最优的重要性采样函数的组成结构简化为关于多个代理模型组合系数的二次规划模型;求解上述关于多个代理模型组合系数的二次规划模型,需要获得被测对象的兴趣事件发生概率  $Q^*(s,a)$ , 使用密集强化学习方法学习效率很低,需要大量的测试次数才能学到完整准确的最优的兴趣事件的发生概率  $Q^*(s,a)$ , 本公开实施例通过获取代理偏差较大的状态动作对  $(s,a)$ , 基于代理偏差较大的状态动作对  $(s,a)$  确定被测对象的兴趣事件发生概率的  $Q^*(s,a)$  值;

[0123] 步骤二、基于代理偏差和置信区间上界构建自适应搜索策略,通过自适应搜索策

略搜索得到状态动作对  $(s, a)$ ；

[0124] 步骤三、令  $\mathcal{D}$  表示本公开实施例搜索到的全部状态动作对  $(s, a)$  的集合, 则通过求解二次规划模型来自适应调整组合系数;

[0125] 本公开实施例通过凸优化求解器 (例如 CVXOPT) 对转换获得的关于多个代理模型组合系数的二次规划模型进行求解, 以更新组合系数  $\alpha$ , 直至满足预设的求解停止条件时, 输出调整的组合系数  $\alpha^*$ 。

[0126] 在一种示例性实例中, 本公开实施例对简化获得的关于多个代理模型组合系数的二次规划模型进行求解的停止条件为预先设定的平均滑动差分函数的取值小于预先设定的停止阈值 (例如 0.02)；

[0127] 步骤四、使用对应于组合系数  $\alpha^*$  的混合的重要性采样函数  $q_{\alpha^*}$  采样生成测试场景, 以对自动驾驶汽车进行测试;

[0128] 本公开实施例根据采样生成的测试场景对自动驾驶汽车进行测试, 获得事故率的评估结果; 本公开实施例得到事故率评估结果为:

$$[0129] \quad \hat{\mu}_{q_{\alpha^*}} = \frac{1}{n} \sum_{i=1}^n \frac{\mathbb{I}_F(X_i) p(X_i)}{q_{\alpha^*}(X_i)}, X_i \sim q_{\alpha^*} \quad (5);$$

[0130] 本公开实施例基于自适应密集强化学习的自适应测试策略调整方法, 可以针对多种类型的被测自动驾驶汽车得到调整的组合系数  $\alpha^*$ , 从而使用  $q_{\alpha^*}$  采样生成定制化的测试场景; 使用  $q_{\alpha^*}$  对自动驾驶汽车进行测试。

[0131] 本公开实施例提出基于自适应密集强化学习方法的自适应测试策略调整方法, 该方法可以应用于高维测试环境中, 并针对多种类型的被测自动驾驶汽车生成定制化的测试策略, 从而提高测试评估的效率, 减少测评所需测试成本。

[0132] 以下以自动驾驶测试的示例性参数对本公开实施例进行说明, 本公开实施例自适应测试策略调整方法, 包括:

[0133] 步骤201、获取驾驶场景的自然概率分布  $p$ , 第  $j$  个代理模型的兴趣事件发生概率  $Q_j, j=1, \dots, J$ , 最大测试时长  $T$ ;

[0134] 步骤202、初始化设定未优化的兴趣事件发生概率  $Q(s, a) = 0, N(s, a) = 0, \forall s \in \mathcal{S}, a \in \mathcal{A}, \alpha = [1/J, \dots, 1/J], i = 0, ASD = 1$ , 设定超参数  $c = 2, \Delta = 10$ ;

[0135] 步骤203、判断出平均滑动差分函数  $ASD$  的取值大于 0.02 时, 执行以下处理:

[0136] 步骤2031、令  $i \leftarrow i + 1$ ;

[0137] 步骤2032、从关键状态集合  $\mathcal{S}_c$  中均匀采样初始状态  $s$ ;

[0138] 步骤2033、令  $r \leftarrow 0, t \leftarrow 0$ ;

[0139] 步骤2034、当  $r = 0$  且  $t < T$  时, 执行以下处理:

[0140] 步骤20341、令  $t \leftarrow t + 1$ ;

[0141] 步骤20342、利用自适应搜索策略  $\eta$  采样动作  $a$ ;

[0142] 步骤20343、执行自适应搜索策略  $\eta$  采样获得的动作  $a$ , 并观测下一个状态  $s'$  与奖励  $r$  (若发生碰撞事故奖励为 1, 否则奖励为 0)；

[0143] 步骤20344、使用表格型密集强化学习方法更新学习  $Q(s, a)$ ;

[0144] 步骤20345、令  $s \leftarrow s'$ ; 本公开实施例执行步骤20345后, 通过赋值处理, 从新获得  $s$ ,

反馈步骤2034执行自适应搜索策略,以获得下一个状态动作对。

[0145] 步骤2035、将获得的状态动作对带入二次规划模型,使用CVXOPT求解二次规划模型来更新组合系数 $\alpha$ ;

[0146] 步骤2036、更新平均滑动差分ASD;

[0147] 步骤204、输出调整的组合系数 $\alpha^*$ 以及对应的混合的重要性采样函数 $q_{\alpha^*}$ ;

[0148] 步骤205、使用 $q_{\alpha^*}$ 采样生成测试场景对自动驾驶汽车进行测试,计算事故率的估计结果。

[0149] 在一种示例性实例中,本公开实施例可以根据估计结果,结合测试场景等设置不同的超参数数值,设置不同的超参数 $c$ 和 $\Delta$ 的数值。

[0150] 以自动驾驶汽车超车场景为例,如图3所示,将被测自动驾驶汽车记为AV,背景车辆记为BV,BV的前车记为LV。在AV超车BV与LV的过程中,若BV也换道超车LV,则BV可能与AV发生碰撞。

[0151] 超车场景的状态定义为 $s = [v_{BV}, R_1, \dot{R}_1, R_2, \dot{R}_2]^T$ ,其中, $v_{BV}$ 表示BV的速度, $R_1$ 表示LV与BV的相对距离, $\dot{R}_1$ 表示LV与BV的相对速度, $R_2$ 表示BV与AV的相对距离, $\dot{R}_2$ 表示BV与AV的相对速度。超车场景的动作定义为 $a = [a_{LV}, a_{BV}]^T$ ,其中 $a_{LV}$ 表示LV的加速度, $a_{BV}$ 表示BV的加速度。

[0152] 为了验证本公开实施例方法的有效性,选择3种不同类型的自动驾驶汽车模型进行测试:(1)文献[2](J.W.Ro,P.S.Roop,A.Malik和P.Ranjitkar,‘一种形式化建模和模拟人类跟车行为的形式化方法’,IEEE智能交通系统汇刊,第19卷,第2期,页639-648,2017(J.W.Ro,P.S.Roop,A.Malik,and P.Ranjitkar,“A formal approach for modeling and simulation of human car-following behavior,”IEEE Transactions on Intelligent Transportation Systems,vol.19,no.2,pp.639-648,2017))中的IDM模型,记为AV-I;(2)文献[3](J.Sangster,H.Rakha和J.Du,‘自然驾驶数据在驾驶员跟车行为建模中的应用’,交通研究记录,第2390卷,第1期,页20-33,2013(J.Sangster,H.Rakha,and J.Du,“Application of naturalistic driving data to modeling of driver car-following behavior,”Transportation research record,vol.2390,no.1,pp.20-33,2013))中的模型,记为AV-II;(3)使用文献[4](J.Schulman,F.Wolski,P.Dhariwal,A.Radford和O.Klimov,‘近端策略优化算法’,arXiv预印本arXiv:1707.06347,2017”(J.Schulman,F.Wolski,P.Dhariwal,A.Radford,and O.Klimov,“Proximal policy optimization algorithms,”arXiv preprint arXiv:1707.06347,2017))中的PPO算法训练得到的模型,记为AV-III。本公开实施例选取的3个代理模型如下:(1)与AV-I相同的IDM模型,记为SM-I;(2)文献[2]中的FVDM模型,最小加速度取 $-1\text{m/s}^2$ ,记为SM-II;(3)FVDM模型,最小加速度取 $-6\text{m/s}^2$ ,记为SM-III。

[0153] 本公开实施例使用文献[3]的方法生成自然驾驶环境(NDE,Naturalistic Driving Environment),并利用平均的组合系数(即 $\alpha = [1/3, 1/3, 1/3]$ )生成自然对抗驾驶环境(NADE,Naturalistic and Adversarial Driving Environment),这两个驾驶环境分别基于蒙特卡洛采样和稀疏对抗采样进行自动驾驶汽车事故率的评估,结果如图4(a)~

(f)所示,图4(a)~(f)中横坐标表示测试次数,图4(a)~(c)中纵坐标表示事故率,图4(d)~(f)中纵坐标表示相对半宽,图中曲线3-1对应于NDE的测试结果,曲线3-2对应于NADE的测试结果,曲线3-3表示事故率参考值(Ground truth),曲线3-4表示相对半宽等于0.3,图4(a)~(f)中相对半宽(RHW,Relative half-width)为事故率估计收敛性的判定标准,当相对半宽小于0.3时认为事故率的评估收敛;针对3种不同的被测自动驾驶汽车,NDE和NADE均可收敛到相同的事故率估计结果,且NADE所需测试次数更少。

[0154] 采用平均组合系数的NADE可以提高测试评估的鲁棒性,但是可能会损害测试效率,因为这样生成的NADE不针对任何被测自动驾驶汽车。采用本公开实施例自适应测试策略调整方法对3个代理模型的组合系数进行调整,得到结果如图5(a)~(f)所示,针对3种被测自动驾驶汽车,调整的组合系数分别为 $\alpha_{AV-I} = [0.95, 0.03, 0.02]$ (最优组合系数为 $\alpha_{AV-I}^* = [1, 0, 0]$ ), $\alpha_{AV-II} = [0.82, 0.16, 0.02]$ ,以及 $\alpha_{AV-III} = [0.65, 0.02, 0.33]$ ,所需的测试次数分别为 $3.8 \times 10^4$ , $4.9 \times 10^4$ ,以及 $3.5 \times 10^4$ 。

[0155] 利用调整的组合系数可以生成相应的定制化的测试策略,从而得到自适应测试环境(Adaptive Testing Environment,AdaTE,即本公开实施例方法);NADE与本公开实施例方法测试结果的对比如图6(a)~(i)所示,图中标号5-1对应的曲线和条状数据表示NADE对应的测试结果,标号5-2对应的曲线和条状数据表示AdaTE对应的测试结果,标号5-3的曲线表示事故率参考值(Ground truth),标号5-4表示相对半宽等于0.3的曲线;本公开实施例方法与NADE收敛到相同的事故率的估计值,且需要更少的测试次数;具体的,针对3种被测自动驾驶汽车,本公开实施例方法需要的测试次数分别减少了37.67%,21.64%,以及36.06%;由此可知,本公开实施例算法可以显著提高测试效率并增强测试鲁棒性。

[0156] 本领域普通技术人员可以理解,上文中所公开方法中的全部或某些步骤、系统、装置中的功能模块/单元可以被实施为软件、固件、硬件及其适当的组合。在硬件实施方式中,在以上描述中提及的功能模块/单元之间的划分不一定对应于物理组件的划分;例如,一个物理组件可以具有多个功能,或者一个功能或步骤可以由若干物理组件合作执行。某些组件或所有组件可以被实施为由处理器,如数字信号处理器或微处理器执行的软件,或者被实施为硬件,或者被实施为集成电路,如专用集成电路。这样的软件可以分布在计算机可读介质上,计算机可读介质可以包括计算机存储介质(或非暂时性介质)和通信介质(或暂时性介质)。如本领域普通技术人员公知的,术语计算机存储介质包括在用于存储信息(诸如计算机可读指令、数据结构、程序模块或其他数据)的任何方法或技术中实施的易失性和非易失性、可移除和不可移除介质。计算机存储介质包括但不限于RAM、ROM、EEPROM、闪存或其他存储器技术、CD-ROM、数字多功能盘(DVD)或其他光盘存储、磁盒、磁带、磁盘存储或其他磁存储装置、或者可以用于存储期望的信息并且可以被计算机访问的任何其他的介质。此外,本领域普通技术人员公知的是,通信介质通常包含计算机可读指令、数据结构、程序模块或者诸如载波或其他传输机制之类的调制数据信号中的其他数据,并且可包括任何信息递送介质。



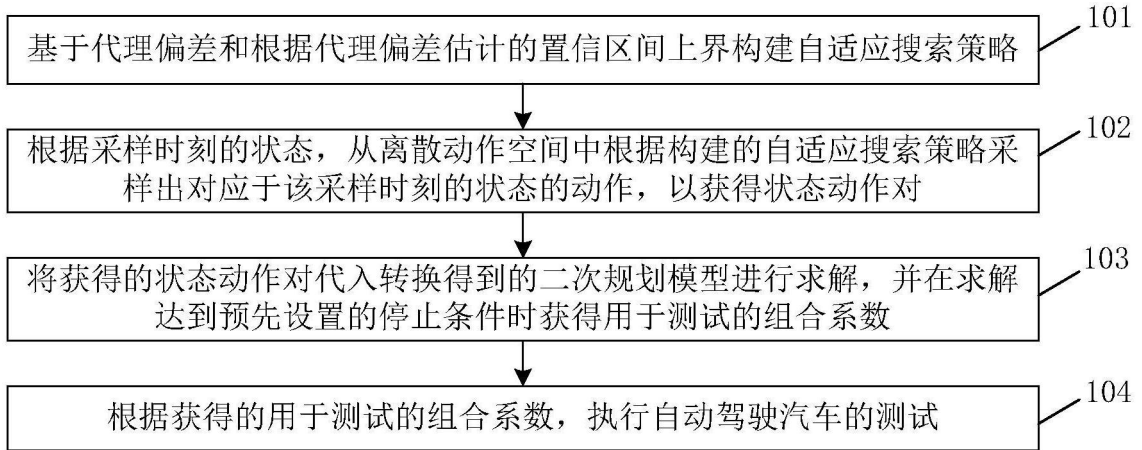


图1

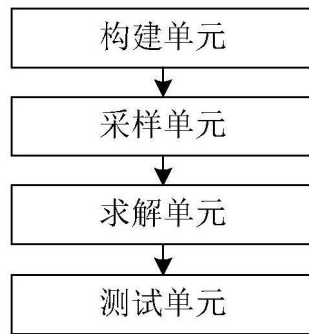


图2

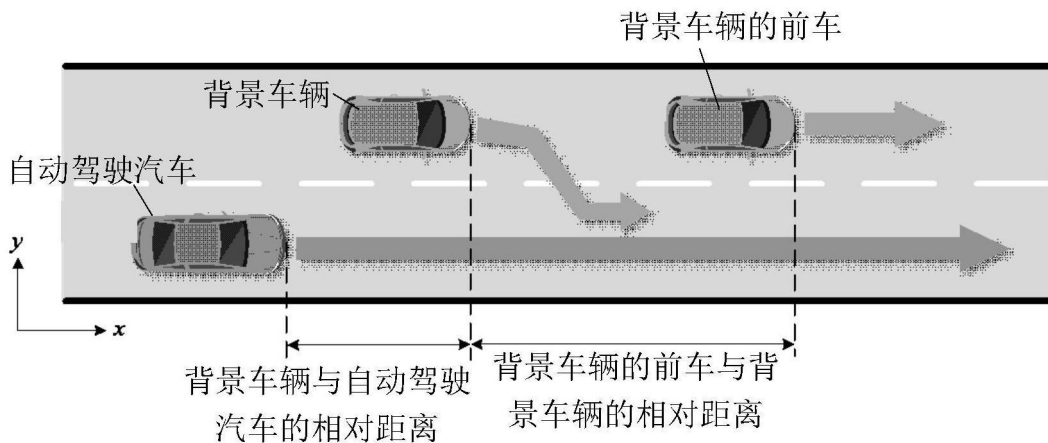


图3

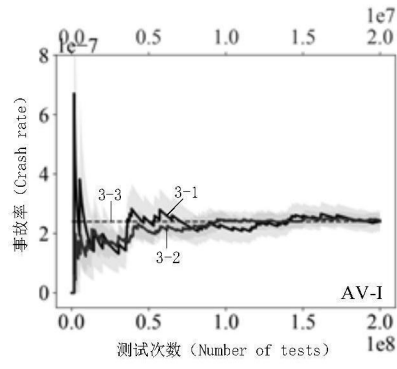


图4(a)

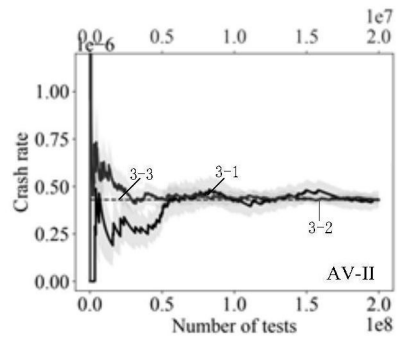


图4(b)

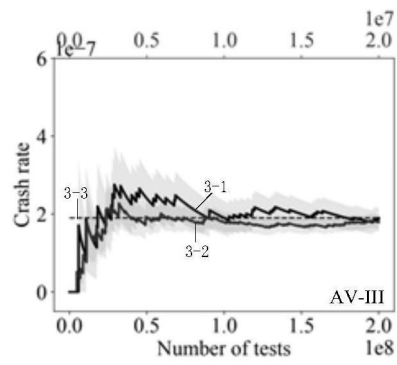


图4(c)

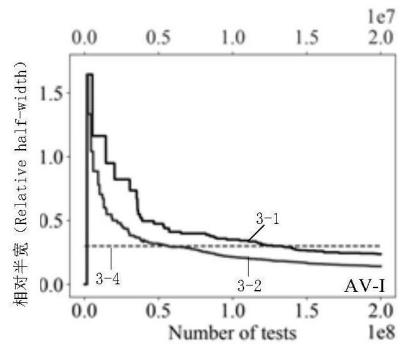


图4(d)

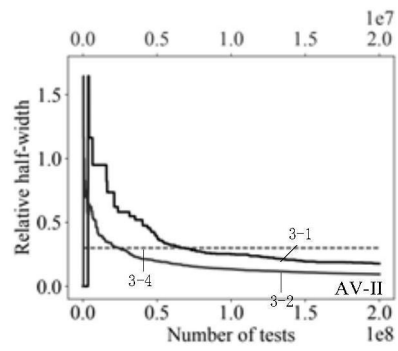


图4(e)

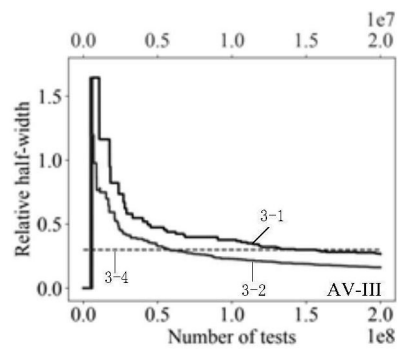


图4(f)

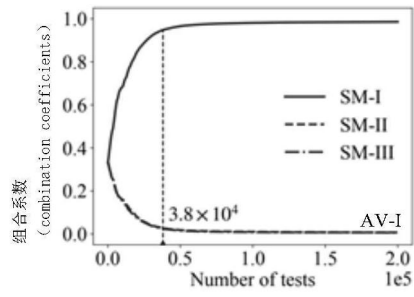


图5(a)

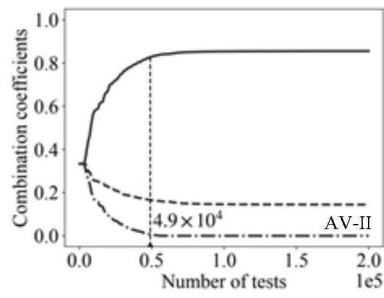


图5(b)

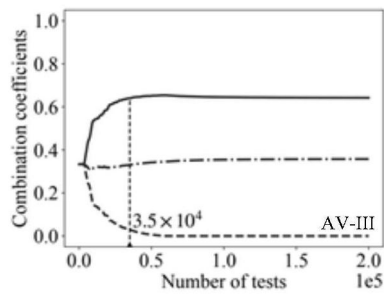


图5(c)

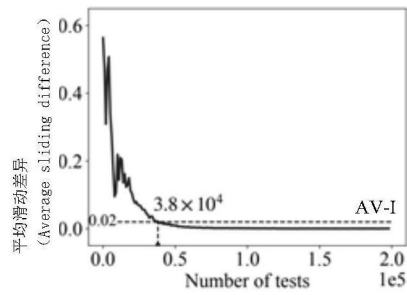


图5(d)

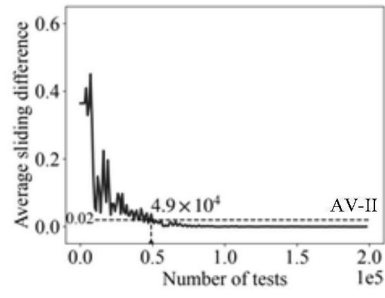


图5(e)

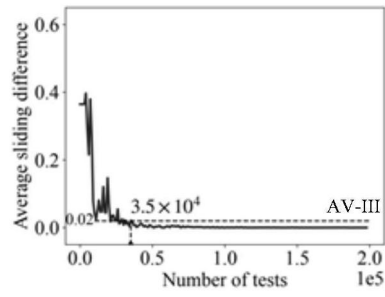


图5(f)

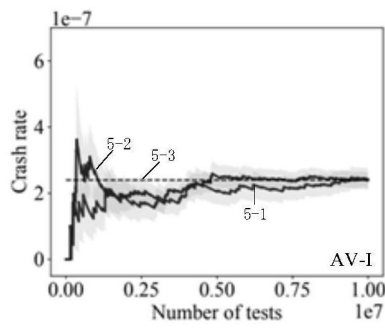


图6(a)

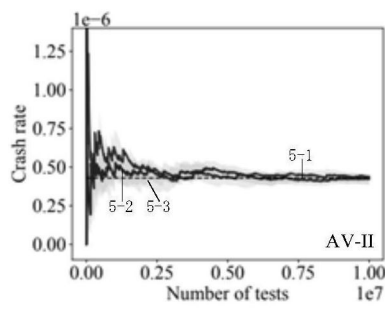


图6(b)

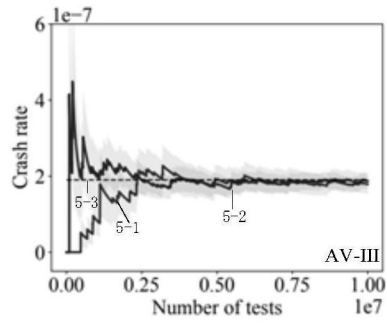


图6(c)

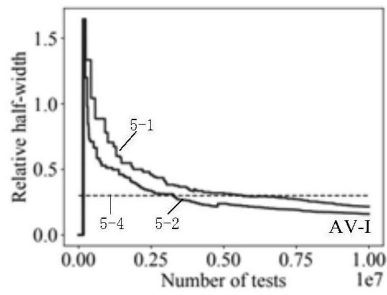


图6(d)

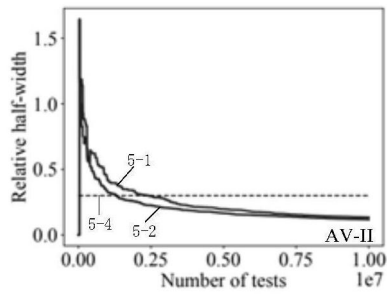


图6(e)

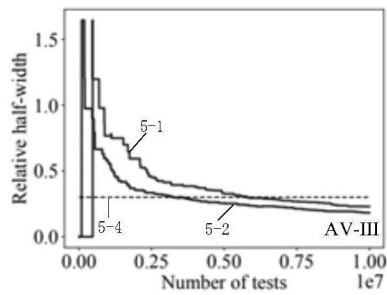


图6(f)

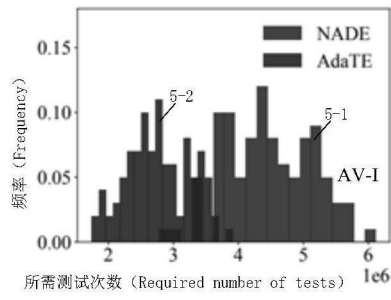


图6(g)

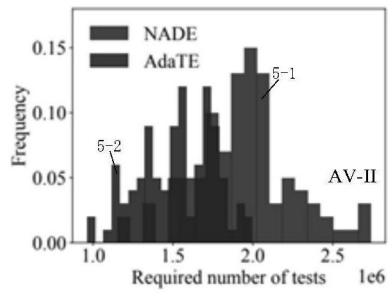


图6(h)

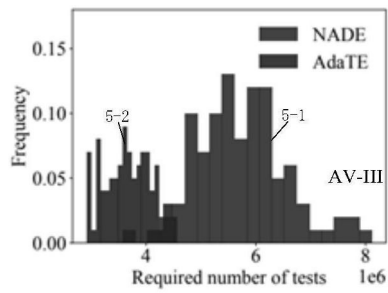


图6(i)