



(12) 发明专利

(10) 授权公告号 CN 112395777 B

(45) 授权公告日 2021.04.16

(21) 申请号 202110078648.0

审查员 夏守璐

(22) 申请日 2021.01.21

(65) 同一申请的已公布的文献号

申请公布号 CN 112395777 A

(43) 申请公布日 2021.02.23

(73) 专利权人 南栖仙策(南京)科技有限公司

地址 210038 江苏省南京市经济技术开发区红枫科技园C4

(72) 发明人 秦熔均 黄睿 毛华奇 彭万立

(74) 专利代理机构 南京乐羽知行专利代理事务所(普通合伙) 32326

代理人 李玉平

(51) Int. Cl.

G06F 30/20 (2020.01)

G06N 3/08 (2006.01)

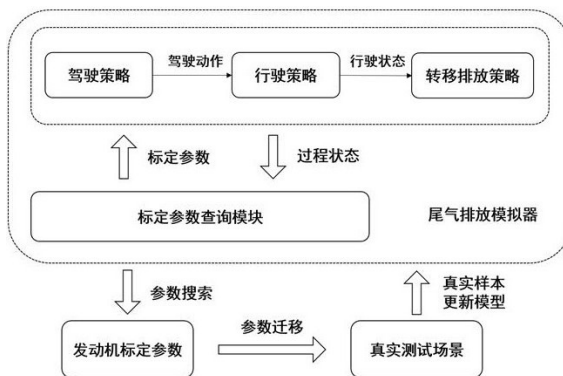
权利要求书2页 说明书6页 附图1页

(54) 发明名称

基于汽车尾气排放模拟环境的发动机标定参数寻优方法

(57) 摘要

本发明公开一种基于汽车尾气排放模拟环境的发动机标定参数寻优方法,以(1)基于机器学习的汽车排放测试模拟环境构建、(2)基于模拟环境的发动机标定最优参数的搜索,实现在模拟环境里低成本高效率寻优汽车尾气排放的发动机标定参数。本发明实现了在模拟环境中低成本高效率地寻优汽车发动机参数,达到了降本增效的技术效果。



1. 一种基于汽车尾气排放模拟环境的发动机标定参数寻优方法,其特征在于,包括汽车排放测试模拟环境构造、发动机标定最优参数搜索和参数迁移三个模块实现:

所述汽车排放测试模拟环境构造模块中,从真实的台架测试中获取的汽车排放过程时序数据,并对时序数据进行学习;学习过程中,基于驾驶员视角对时序过程数据作状态和动作数据分析,将状态数据和动作数据分别编码,记为S和a,然后根据数据的时序关系,将过程数据排列为对应的状态-动作序列,即 $(S_0, a_0, S_1, a_1, \dots, S_T)$ 序列;使用机器学习方法,通过此的状态-动作序列学习得到三个策略模型,即控制汽车行驶的驾驶策略模型,决定汽车行驶状态的行驶策略模型,汽车排放时序状态转移的转移排放策略模型;

将所述三个策略模型混合作为汽车排放策略模拟环境,该模拟环境输入标定参数状态值和汽车当前的状态数据,并输出动作数据执行完成后的下一个状态数据以及当前时刻的排放值;汽车排放策略模拟环境能够模拟司机整个驾驶过程,便于从中搜索出最优排放的发动机标定参数;

所述发动机标定最优参数搜索模块中,在训练好的模拟环境基础上,根据汽车排放测试需求设计一个奖励函数,根据在不同参数设置下,虚拟司机在与模拟环境的交互过程中获得的奖励函数,使用搜索算法寻找最优排放参数;

构建汽车排放测试模拟环境,分为模拟司机驾驶的驾驶策略网络;汽车行驶过程的行驶策略网络;汽车时序状态转移的转移排放策略模型;用数据驱动的方式得到可用于辅助真实场景决策的模拟环境,根据评价指标评估模拟环境,并在模拟环境中通过无梯度搜索算法去搜索最优的标定参数;包括如下步骤:

S101,获取原始实验数据;

S102,基于所述实验数据进行特征提取,得到用于模拟建模的特征数据集;

S103,获取已训练的汽车发动机相关状态变分自编码器模型;S103步骤,包括:

S103-1,提取所述用于模拟建模的特征集中的汽车发动机参数无关状态数据集;

S103-2,将上述数据集中的样本 y 映射到低维空间中的样本 $z \sim N(\mu, \sigma)$;

S103-3,将所述低维样本 z 映射回原始数据空间中的样本 y' ;

S103-4,最小化重构损失函数 $loss(y, y') = \|y - y'\|_2^2$,得到已训练的汽车发动机相关状态变分自编码器模型;

S104,固定所述已训练的汽车发动机相关状态变分编码器模型中的参数,将所述特征集中的汽车发动机参数相关状态数据输入变分编码器模型中得到转换后的状态向量;

S105,获取已训练的汽车尾气排放测试的模拟环境模型;

S106,获取汽车发动机的参数空间;

S107,固定已训练的模拟环境中的参数和汽车发动机的参数空间,通过优化不同参数组合在模拟环境中的排放指标,得到模拟环境上的最优发动机标定参数。

2. 根据权利要求1所述的基于汽车尾气排放模拟环境的发动机标定参数寻优方法,其特征在于,在实际应用情形下,所述参数迁移模块需要收集发动机标定最优参数搜索模块提供的参数,在真实汽车台架测试中产生的新的过程交互数据,用于对汽车排放策略模拟环境的进一步更新,并通过发动机标定最优参数搜索模块重新搜索参数。

3. 根据权利要求1所述的基于汽车尾气排放模拟环境的发动机标定参数寻优方法,其

特征在于,所述汽车排放测试模拟环境构造模块的构建过程主要分为四步:

首先,构建控制汽车行驶的驾驶策略模型;

其次,构建模拟司机按照驾驶策略模型驾驶后,表征发动机行驶状态的行驶策略模型,该模型用于模拟在虚拟司机驾驶操控下,发动机的状态模式;

然后,选择一个汽车排放时序状态转移的转移排放策略模型,该模型用于模拟司机驾驶行为在发动机的当前状态模式下,汽车排放相关状态量的转移和当前时刻的尾气排放值;

最后,将司机驾驶策略模型,发动机行驶策略模型以及汽车排放时序状态转移的转移排放策略模型组合得到一个混合策略模型,通过从给定数据关系中得到的标定参数查询模块,来进行发动机参数的时序迭代,混合策略模型和标定参数查询模块一起,作为模拟汽车在一段完整的测试时间里的整个尾气排放测试过程的模拟环境;所述标定参数查询模块是从当前时刻汽车发动机参数相关状态以及驾驶状态查询相应的瞬时发动机参数,是一个固定的函数模型。

4. 根据权利要求1所述的基于汽车尾气排放模拟环境的发动机标定参数寻优方法,其特征在于,所述构建汽车排放测试模拟环境中的混合策略模型时,采用神经网络作为混合策略模型各模块:

对于驾驶策略模型,模型的输入为当前一步的状态和动作,输出为下一步状态;将从真实环境中收集到的交互数据、即状态-动作序列 $(S_0, a_0, S_1, a_1, \dots, S_T)$ 进行切分,划分出多个 (S_t, a_t, S_{t+1}) 的元组,其中 S_t 为策略模型当前一步的状态, a_t 为当前一步的动作, S_{t+1} 为策略模型输出的下一步状态; (S_t, a_t) 作为驾驶策略模型的输入, S_{t+1} 为输出,使用驾驶策略模型学习该输入到输出的对应关系。

5. 根据权利要求1所述的基于汽车尾气排放模拟环境的发动机标定参数寻优方法,其特征在于,所述模拟环境使用对抗学习训练混合策略模型,该模型使用神经网络分别为混合策略的各模块建立模型;其中,判别器网络用于判别生成序列的置信度;通过混合策略模型,生成一批状态-动作序列来更新判别器;接下来使用更新后的判别器网络输出的置信度作为序列的期望奖励,用强化学习算法更新混合策略模型;将通过以上过程训练完成后的混合策略模型作为汽车排放测试的模拟环境使用。

6. 根据权利要求1所述的基于汽车尾气排放模拟环境的发动机标定参数寻优方法,其特征在于,所述策略迁移模块中,根据最优排放参数在真实环境台架测试上部署后的结果,对汽车排放测试模拟环境以及最优排放参数进行增量式的更新,具体的实施方式如下:

首先,将在汽车排放测试模拟环境中进行参数搜索得到的最优排放参数部署到真实的汽车台架测试中进行应用;在应用的过程中记录所有由该参数下汽车测试过程中交互生成的交互数据集;

其次,由新应用参数下产生的新交互数据集中与原数据集合并,并保存至一个新数据集中;

再次,对模拟环境和最优参数,按照模拟环境构建和参数搜索步骤定期使用重新整合的交互数据定期进行全量更新。

基于汽车尾气排放模拟环境的发动机标定参数寻优方法

技术领域

[0001] 本发明涉及一种基于模拟训练的发动机标定参数寻优方法,在基于机器学习技术构建的模拟环境,模拟汽车尾气排放过程,从而获得对发动机标定参数优化策略,属于智能制造领域。

背景技术

[0002] 汽车发动机标定过程通过调整电子控制单元内的各控制模块参数,来使得汽车驾驶过程中排放的尾气满足国家排放标准。当发动机、控制策略,外围器件确定之后,为了让汽车的排放满足国家标准,需要找出发动机在工作状态下的参数。一般的标定过程首先是通过台架实验使得发动机可以初步正常工作,然后得到最初的一组参数。接下来按照国家标准,例如第六阶段机动车污染物排放标准(国六标准)进行整车实验。通常,排放测试过程需要在固定的驾驶速度线条件下测试,这条速度线涵盖了启动、加速,减速等各种汽车行驶情况。标定过程需要对发动机参数进行调整,最终使得发动机在各个工况下都可以满足排放要求。

[0003] 发动机控制参数通常有上万个可调整变量,且由于发动机本身就是一个很复杂的系统,涉及各种物理模块,因此,目前市面上几乎没有一个可以自动标定参数的系统。目前为止,发动机的参数标定主要依赖于人工经验。而参数标定过程首先需要进行整车实验,整车实验需要司机进行驾驶,在符合速度线的前提下,收集、测量尾气数据,之后标定工作人员根据对排放数据的分析和标定经验对参数进行调整。

[0004] 这一过程主要面临两个问题:

[0005] (1) 费用高昂。整车实验需要配备一个完整的实验环境,搭建环境需要很高的费用。一次整车实验费用高达数万,而参数标定是一个迭代优化的序列决策过程,完成一个发动机的标定往往需要数百次的整车实验,成本高昂且难以避免。

[0006] (2) 高度依赖标定人员的经验。这些经验通常由标定人员之间交流以及实验中的试错获取,难以形成数字化的经验,新入行从业人员的学习成本较高。因此,现行参数标定的整个过程需要大量的人力和高昂的费用。

[0007] 另一方面,测试中的实验数据通常会保留下来,可以利用机器学习的方法,从数据中学习出智能的参数标定方法。

[0008] 机器学习中适用于序列决策过程的强化学习方法,其学过程要求智能体与环境不断交互试错,以帮助智能体在环境中逐渐学得最优控制策略(即最优决策策略),从而自动完成决策任务。然而,强化学习的过程中,智能体与环境进行大量交互试错的过程,在现实环境中往往难以实行的。例如,在实际的汽车测试过程中直接用强化学习,则需要使用大量不同的参数配置在真实的台架测试中进行试错,然后从这些参数条件下的汽车交互轨迹数据中,学习得到最优的参数,显然,这种方法会带来巨大的经济损失和时间成本,是无法接受的。因此,本发明提出了一种从历史汽车测试数据中构建模拟环境并在其中搜索最优发动机参数标准的方法。

发明内容

[0009] 发明目的:在汽车排放测试中,一般都是人工优化标定参数,需要不断地进行台架实验,耗费大量的人力和物力。为了降本增效,本发明提出了一种基于机器学习技术进行汽车尾气排放模拟测试的发动机参数寻优标定方法,首先从历史汽车标定实验数据中构建汽车排放测试模拟环境,然后通过算法在模拟环境中搜索最优标定参数,实现在模拟环境中低成本高效率地寻优汽车发动机参数。

[0010] 技术方案:一种基于汽车尾气排放模拟环境的发动机标定参数寻优方法,基于机器学习技术进行汽车尾气排放模拟测试的发动机参数寻优标定,通过在处理器上执行汽车排放测试模拟环境构造、发动机标定最优参数搜索和参数迁移三个模块实现:

[0011] 汽车排放测试模拟环境构造模块中,从真实的台架测试中获取的汽车排放过程时序数据,并对时序数据进行学习;学习过程中,基于驾驶员视角对时序过程数据作状态和动作数据分析,将状态数据和动作数据分别编码,记为 S 和 a ,然后根据数据的时序关系,将过程数据排列为对应的状态-动作序列,即 $(S_0, a_0, S_1, a_1, \dots, S_T)$ 序列。使用机器学习方法,通过此的状态-动作序列学习得到三个策略模型,即控制汽车行驶的驾驶策略模型,决定汽车行驶状态的行驶策略模型,汽车排放时序状态转移的转移排放策略模型。

[0012] 将三个策略模型混合作为汽车排放策略模拟环境,该模拟环境输入标定参数状态值和汽车当前的状态数据,并输出动作数据执行完成后的下一个状态数据以及当前时刻的排放值;汽车排放策略模拟环境能够模拟司机整个驾驶过程,便于从中搜索出最优排放的发动机标定参数。

[0013] 在实际应用情形下,参数迁移模块需要收集发动机标定最优参数搜索模块提供的参数,在真实汽车台架测试中产生的新的过程交互数据,用于对汽车排放策略模拟环境的进一步更新,并通过发动机标定最优参数搜索模块重新搜索参数。

[0014] 汽车排放测试模拟环境构造模块的构建过程主要分为四步:

[0015] 首先,构建控制汽车行驶的驾驶策略模型;

[0016] 其次,构建模拟司机按照驾驶策略模型驾驶后,表征发动机行驶状态的行驶策略模型,该模型用于模拟在虚拟司机驾驶操控下,发动机的状态模式;

[0017] 然后,选择一个汽车排放时序状态转移的转移排放策略模型,该模型用于模拟司机驾驶行为在发动机的当前状态模式下,汽车排放相关状态量的转移(即下一时刻状态量)和当前时刻的尾气排放值;

[0018] 最后,将司机驾驶策略模型,发动机行驶策略模型以及汽车排放时序状态转移的转移排放策略模型组合得到一个混合策略模型,通过从给定数据关系中得到的标定参数查询模块(从当前时刻汽车发动机参数相关状态以及驾驶状态查询相应的瞬时发动机参数,是一个固定的函数模型),来进行发动机参数的时序迭代,混合策略模型和标定参数查询模块一起,作为模拟汽车在一段完整的测试时间里的整个尾气排放测试过程的模拟环境。

[0019] 构建汽车排放测试模拟环境中的混合策略模型时,采用神经网络作为混合策略模型各模块:

[0020] 以驾驶策略模型举例,模型的输入为当前一步的状态和动作,输出为下一步状态;将从真实环境中收集到的交互数据、即状态-动作序列 $(S_0, a_0, S_1, a_1, \dots, S_T)$ 进行切分,划分出多个 (S_t, a_t, S_{t+1}) 的元组,其中 S_t 为策略模型当前一步的状态, a_t 为当前一步的动

作, S_{t+1} 为策略模型输出的下一步状态; (S_t, a_t) 作为驾驶策略模型的输入, S_{t+1} 为输出, 使用驾驶策略模型学习该输入到输出的对应关系。

[0021] 模拟环境使用对抗学习训练混合策略模型, 该模型使用神经网络分别为混合策略的各模块建立模型。其中, 判别器网络用于判别生成序列的置信度。通过混合策略模型, 生成一批状态-动作序列来更新判别器。接下来使用更新后的判别器网络输出的置信度作为序列的期望奖励, 用强化学习算法更新混合策略模型。将通过以上过程训练完成后的混合策略模型作为汽车排放测试的模拟环境使用。

[0022] 发动机标定最优参数搜索模块中, 在所述训练好的模拟环境基础上, 根据汽车排放测试需求设计一个奖励函数, 根据在不同参数设置下, 虚拟司机在与模拟环境的交互过程中获得的奖励函数, 使用搜索算法寻找最优排放参数。

[0023] 策略迁移模块中, 根据最优排放参数在真实环境台架测试上部署后的结果, 对汽车排放测试模拟环境以及最优排放参数进行增量式的更新, 具体的实施方式如下:

[0024] 首先, 将在汽车排放测试模拟环境中进行参数搜索得到的最优排放参数部署到真实的汽车台架测试中进行应用; 在应用的过程中记录所有由该参数下汽车测试过程中交互生成的交互数据集;

[0025] 其次, 由新应用参数下产生的新交互数据集中与原数据集合并, 并保存至一个新数据集中;

[0026] 再次, 对模拟环境和最优参数, 按照模拟环境构建和参数搜索步骤定期使用重新整合的交互数据定期进行全量更新。

附图说明

[0027] 图1是本发明方法的训练流程图。

具体实施方式

[0028] 下面结合具体实施例, 进一步阐明本发明, 应理解这些实施例仅用于说明本发明而并不用于限制本发明的范围, 在阅读了本发明之后, 本领域技术人员对本发明的各种等价形式的修改均落于本申请所附权利要求所限定的范围。

[0029] 基于汽车尾气排放模拟环境的发动机标定参数寻优方法, 主要包括构建汽车排放测试模拟环境、评估模拟环境准确性、在模拟环境中搜索最优标定参数, 其中构建汽车排放测试模拟环境, 分为模拟司机驾驶的驾驶策略网络; 汽车行驶过程的行驶策略网络; 汽车时序状态转移的转移排放策略模型等; 用数据驱动的方式得到可用于辅助真实场景决策的模拟环境, 根据我们设计的评价指标评估模拟环境, 并在模拟环境中通过无梯度搜索算法去搜索最优的标定参数。包括如下步骤:

[0030] S101, 获取原始实验数据; S101步骤, 具体包括:

[0031] S101-1, 获取标定实验的参数设定, 简记为 P;

[0032] S101-2, 获取基于上述 P 的测试过程数据;

[0033] S101-3, 获取基于上述 P 的测试结果数据。

[0034] S102, 基于所述实验数据进行特征提取, 得到用于模拟建模的特征数据集合; S102步骤, 包括:

- [0035] 步骤 S102-1,获取标定实验排放测试时间内的瞬时结果数据;
- [0036] 步骤 S102-2,获取标定实验排放测试时间内的瞬时过程数据;
- [0037] 步骤 S102-3,如果瞬时过程数据和结果数据的时间间隔粒度不同,则通过插值的方式按细粒度时间间隔对齐;
- [0038] 步骤 S102-4,根据时间对齐后的瞬时结果数据和瞬时过程数据,拼接成排放测试时间内的过程-结果数据;
- [0039] 步骤 S102-5,将过程-结果数据按时间 t 进一步划分:汽车发动机参数相关状态 $S_{t,1}$ 、汽车发动机参数相关状态 $S_{t,2}$ 、汽车驾驶动作 $A_{drive\ t}$ 、汽车附加动作 $A_{aux\ t}$ 、汽车排放 O_t 、汽车下一时刻参数无关状态 $S_{t+1,1}$ 、汽车下一时刻参数相关状态 $S_{t+1,2}$,时间先后顺序,得到状态-动作序列 $(S_{0,1}, S_{0,2}, A_{drive0}, A_{aux0}, O_0, S_{1,1}, S_{1,2}, A_{drive1}, A_{aux1}, O_1, \dots, S_{T,1}, S_{T,2})$ 。
- [0040] S103,获取已训练的汽车发动机相关状态变分自编码器模型;S103步骤,包括:
- [0041] S103-1,提取所述用于模拟建模的特征集合中的汽车发动机参数无关状态数据集;
- [0042] S103-2,将上述数据集中的样本 y 映射到更低维空间中的样本 $z \sim N(\mu, \sigma)$;
- [0043] S103-3,将所述低维样本 z 映射回原始数据空间中的样本 y' ;
- [0044] S103-4,最小化重构损失函数 $loss(y, y') = \|y - y'\|_2^2$,得到已训练的汽车发动机相关状态变分自编码器模型。
- [0045] S104,固定所述已训练的汽车发动机相关状态变分编码器模型中的参数,将所述强化学习特征集中的汽车发动机参数相关状态数据输入变分编码器模型中得到转换后的状态向量;
- [0046] S105,获取已训练的汽车尾气排放测试的模拟环境模型;S105步骤,包括:
- [0047] 步骤 S105-1,构建驾驶策略模型,该模型用于模拟车辆驾驶测试过程中的驾驶员控制行为,例如驾驶车辆按照国标标准速度曲线行驶的驾控过程。从真实的台架测试中获取汽车排放过程时序数据,并对时序数据进行学习;学习过程中,基于驾驶员视角对时序过程数据作状态和动作数据分析,将状态数据和动作数据分别编码,记为 S 和 a ,然后根据数据的时序关系,将过程数据排列为对应的状态-动作序列,即 $(S_0, a_0, S_1, a_1, \dots, S_T)$ 序列,将该状态-动作序列切分出多个 (S_t, a_t, S_{t+1}) 的元组,其中 S_t 为策略模型当前一步的状态, a_t 为当前一步的动作, S_{t+1} 为策略模型输出的下一步状态;其中 (S_t, a_t) 作为驾驶策略模型的输入, S_{t+1} 为输出;驾驶策略模型需要学习该输入到输出的对应关系;使用监督学习方法对这个对应关系进行学习,具体方案如下:将 (S_t, a_t) 作为样本, S_{t+1} 作为预测目标,从划分后的数据集中训练出模拟环境;最后,训练后的模拟环境根据输入的状态-动作对 (S_t, a_t) ,预测下一个状态 S_{t+1} ,供混合策略模型中的驾驶策略模型使用;(该过程在S105-2,S105-3中类似)
- [0048] 步骤 S105-2,构建行驶策略模型,该模型用于模拟车辆驾驶测试过程中驾驶控制行为对发动机状态的影响模式;
- [0049] 步骤 S105-3,构建汽车排放时序状态转移的转移排放策略模型,该模型用于模拟汽车的状态转移和气体排放的影响方式;
- [0050] 步骤 S105-4,获取发动机参数查询模块,该模块用于从当前时刻汽车发动机参数

相关状态以及驾驶状态查询相应的瞬时发动机参数,是一个固定的函数模型;

[0051] 步骤 S105-5,通过策略选择函数将所述驾驶控制策略模型、驾驶附加策略模型和驾驶状态转移与排放策略模型以及发动机参数查询模块组合成一个混合策略模型,作为汽车尾气排放测试的场景模拟环境。具体方案如下:

[0052] (1)用若干神经网络分别为混合策略各模块建立模型;

[0053] (2)建立一个神经网络作为判别器,该判别器分别接收两组序列数据,即真实实验中采样的序列数据和生成的序列数据。判别器可以理解成一个二分类器,对于真实的序列数据输出1,对于生成的数据输出0,使用判别器可以得到某条生成序列 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 的置信度;

[0054] (3)通过混合策略模型的交互训练,生成一批 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列,用生成的 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列和真实汽车排放测试数据的 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列更新判别器,将汽车排放测试的真实数据的 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列数据集记为 D ,训练过程中的生成的 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列数据集记为 D' ,更新目标如下:

$$[0055] \quad \min_f E_{\tau \sim D} [\ln f(\tau)] + E_{\tau' \sim D'} [\ln(1 - f(\tau'))]$$

[0056] 其中 \ln 为自然对数函数, f 为判别器,为输出为0~1之间实数(置信度),越接近1,表示和真实历史数据相似度越高, $f(\tau)$, $f(\tau')$ 分别表示判别器在 D 中单条 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列和 D' 中单条 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列上输出的置信度。具体实施时,可令判别器 f 计算出 $f(S_t, a_t, S_{t+1})$ 作为状态-动作-状态对 (S_t, a_t, S_{t+1}) 三元组的置信度,并按照: $f(\tau) = \frac{1}{T} \sum_{i=0}^{T-1} f(S_t, a_t, S_{t+1})$ 的方式计算判别器对单条序列输出的置信度;

[0057] (4)用更新后的判别器给生成的 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列打分,即输出置信度,作为序列的期望奖励,用强化学习算法更新混合策略模型;

[0058] (5)重复(3)(4)两步,直到达到最大循环次数;

[0059] (6)将通过以上过程训练完成后的混合策略模型作为汽车排放测试模拟环境中的汽车排放策略模拟环境使用。

[0060] S106,获取汽车发动机的参数空间;S106步骤,包括:

[0061] 步骤 S106-1,设置参数的类型以及取值范围。若该参数为离散值,确定可能的取值集合;若该参数为连续值,确定参数的下界和上界;

[0062] 步骤 S106-2,设置参数的约束条件,包括参数之间应满足的数值关系和实际条件下的限制关系。

[0063] S107,固定所述已训练的模拟环境中的参数和汽车发动机的参数空间,通过优化不同参数组合在所述模拟环境中的排放指标,得到模拟环境上的最优发动机标定参数。

[0064] S107步骤,包括:

[0065] 步骤 S107-1,如S105,模型建立一个搜索参数的模拟环境;

[0066] 在一具体实施中,S107-1中的模拟环境构建如下:

[0067] (1)通过神经网络搭建驾驶策略模型、驾驶附加策略模型,状态转移与排放策略模型。驾驶策略模型旨在模拟司机的驾驶行为,即司机如何根据特定的速度曲线对汽车进行控制,如踩油门,刹车等行为。驾驶附加策略模型可以通过模拟出的驾驶行为预测出发动机的状态信息,发动机的状态信息与车辆的排放息息相关,因此驾驶附加策略模型得到的输出可以作为状态转移与排放策略模型的输入,从而完成模拟环境的搭建;

[0068] (2)实验过程中收集得到的数据与发动机参数存在一元或二元的函数关系,通过该函数关系,使用线性插值的方式建立发动机参数的查询模块。

[0069] (3)对各模型整体进行对抗学习训练。训练过程中,用神经网络建立判别器,用于判别某条生成序列 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 的置信度,取值为0到1之间的实数,越接近1则越像真实数据,越接近0则越像生成数据。在训练过程中使用强化学习算法(如PPO)去更新各模块,生成序列过程如下:

[0070] 首先是模拟环境的输入数据,模拟环境的所有状态变量中采样获得初始值 s_0 ,和该初始值下通过发动机参数查询模块得到的过程量 p_0 ,一起记为输入数据 S_0 ;

[0071] 其次,将该输入数据输入驾驶控制策略模型网络中,采样得到的虚拟司机驾驶动作 A_{drive0} ,然后级联上 S_0 之后,一起输入到驾驶附加策略模型网络中,得到车辆行驶过程中的状态 A_{aux0} ,再级联 A_{drive0}, S_0 后,一起输入到状态转移与排放策略模型网络中,得到当前时刻的排放 O_0 和下一时刻的状态 S_1 。再根据当前状态 S_1 ,重复上述过程得到一条轨迹。重复M次此过程得到M条轨迹,每条轨迹都可表述为 $(S_0, a_0, S_1, a_1, \dots, S_n)$ 序列。

[0072] 按所述S105-5训练模拟环境各模块,最后通过评价指标选择出最佳的模拟环境,具体的评价指标可参考如下:记真实环境中的尾气排放为 $O=(O_1, O_2, \dots, O_T)$,模拟环境得到的排放为 $O'=(O'_1, O'_2, \dots, O'_T)$,使用两者的均方误差作为评价指标,即

$$MSE = \left\| O - O' \right\|_2^2。$$

[0073] 步骤 S107-2,如所述S106,建立参数的搜索空间;

[0074] 步骤 S107-3,在参数的搜索空间中采样多组参数,在不同的参数的下,虚拟司机在模拟环境进行排放测试的步骤具体如下:

[0075] 步骤 S107-3-1,根据汽车排放测试的需求(如国标VI),设计一个奖励函数,用于评估不同参数的好坏;

[0076] 步骤 S107-3-2,使用无梯度优化算法或演化算法,结合模拟环境给出的多组参数的奖励函数信息,去除奖励函数值较差(小)的参数组合,不断减小待搜索的参数空间,直到算法遍历参数空间,得到最优参数为止,最后得到的参数作为汽车排放测试模拟环境中的最佳排放参数。

[0077] 在一具体实施中,所述 S107-3-1 中的奖励函数的设计可以参考如下:

[0078] 在时长为 T 秒的过程 $t=1, 2, \dots, T$,汽车排放测试实验每秒生成 N 中气体 $O_n^t, n=1, 2, \dots, N$,每种气体的权重为 W_n ,奖励函数可以设置为在模拟环境中生成的所有排放气体的加权值的相反数: $-\sum_{t=1}^T \sum_{n=1}^N W_n O_n^t$,生成排放气体量越少,越可能满足国标,其奖励值越大。反之,则给较小的奖励值。

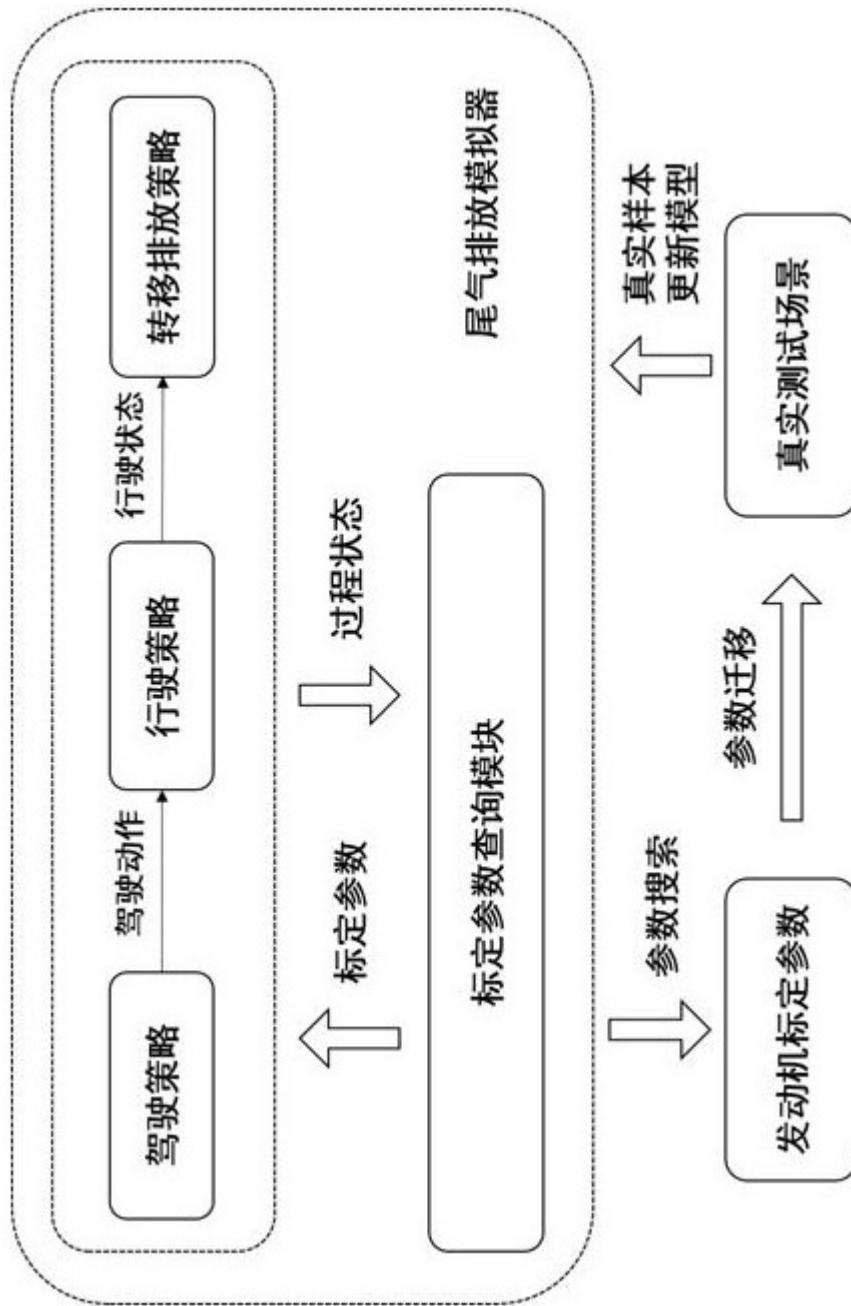


图1