



(12) 发明专利申请

(10) 申请公布号 CN 115713860 A

(43) 申请公布日 2023. 02. 24

(21) 申请号 202211472405.6

(22) 申请日 2022.11.23

(71) 申请人 浙江省交通运输科学研究院

地址 311305 浙江省杭州市西湖区留下街
道大龙驹坞705号

(72) 发明人 金波 娄刃 何亚强 汪成立
张俊烨 杨松 张冶芳

(74) 专利代理机构 浙江千克知识产权代理有限公司 33246

专利代理师 黎双华

(51) Int.Cl.

G08G 1/01 (2006.01)

G08G 1/08 (2020.01)

G06F 30/27 (2006.01)

G06N 20/00 (2019.01)

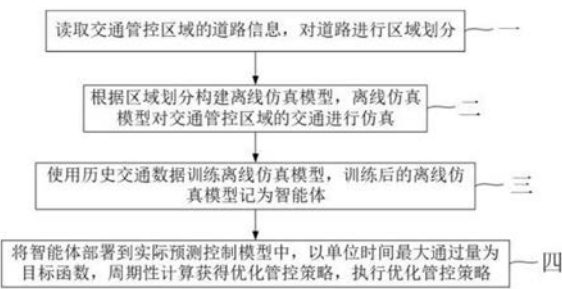
权利要求书3页 说明书8页 附图5页

(54) 发明名称

一种基于强化学习的高速公路交通管控方法

(57) 摘要

本发明涉及交通控制技术领域,具体涉及一种基于强化学习的高速公路交通管控方法,包括以下步骤:步骤一:读取交通管控区域的道路信息,对道路进行区域划分;步骤二:根据区域划分构建离线仿真模型,所述离线仿真模型对交通管控区域的交通进行仿真;步骤三:训练所述离线仿真模型,训练后的离线仿真模型记为智能体;步骤四:将所述智能体部署到实际预测控制模型中,以单位时间最大通过量为目标函数,周期性计算获得优化管控策略,执行所述优化管控策略。本发明的有益技术效果包括:利用强化学习训练智能体,克服模型预测控制对于交通流预测模型过度依赖的缺点,提高了预测的准确度,从而实现有效管控策略的准确解算,获得更为优化的管控策略。



1. 一种基于强化学习的高速公路交通管控方法,其特征在于,包括以下步骤:

步骤一:读取交通管控区域的道路信息,对道路进行区域划分;

步骤二:根据区域划分构建离线仿真模型,所述离线仿真模型对交通管控区域的交通进行仿真;

步骤三:使用历史交通数据训练所述离线仿真模型,训练后的离线仿真模型记为智能体;

步骤四:将所述智能体部署到实际预测控制模型中,以单位时间最大通过量为目标函数,周期性计算获得优化管控策略,执行所述优化管控策略。

2. 根据权利要求1所述的一种基于强化学习的高速公路交通管控方法,其特征在于,

步骤一中,进行区域划分的方法包括:

将交通管控区域的道路划分为以下区域类型:可变限速区域K0、加速区域J、合流区域H和匝道区域Z0;

将可变限速区域K0的上游区域划分为L-1个主路检测区域,分别记为 K_1, K_2, \dots, K_{L-1} ;

将匝道区域Z0的上游区域划分为L-1个匝道检测区域,分别记为 Z_1, Z_2, \dots, Z_{L-1} 。

3. 根据权利要求2所述的一种基于强化学习的高速公路交通管控方法,其特征在于,

步骤一中,构建离线仿真模型的方法包括:

设置离线仿真模型的状态量,所述状态量包括:可变限速区域K0及其上游区域的交通流密度 $\{q_{K_0}, q_{K_1}, \dots, q_{K_{L-1}}\}$,合流区域交通流密度 q_H 、匝道区域及其上游区域的交通流密度 $\{q_{Z_0}, q_{Z_1}, \dots, q_{Z_{L-1}}\}$,状态量记为 $\{q_{K_0}, \dots, q_{K_{L-1}}, q_H, q_{Z_0}, \dots, q_{Z_{L-1}}\}$;

设置离线仿真模型的动作,设置控制步长L,将所述动作设置为可变限速区域K0内在L步长内的限速值序列 $\{V_0, V_1, \dots, V_{L-1}\}$ 和匝道区域Z0在L步长内的汇入率序列 $\{P_0, P_1, \dots, P_{L-1}\}$;

设置离线仿真模型的奖励,在控制步长L内的总通行时间TTT作为奖励,

$$TTT = T \sum_{l=0}^{L-1} \left[T \sum_{t=0}^{i-1} (q_{in}(t) - q_{out}(t)) \right]$$

其中,T表示控制步长的间隔时长, $q_{in}(t)$ 表示t时刻控制区域内的流入交通流密度, $q_{out}(t)$ 表示t时刻控制区域内的流出交通流密度。

4. 根据权利要求3所述的一种基于强化学习的高速公路交通管控方法,其特征在于,

步骤一中,构建离线仿真模型的方法还包括构建仿真环境,

构建仿真环境的方法包括:

根据交通管控区域的道路信息建立道路模型;

将流入交通流数据及流出交通流数据导入到道路模型,设置车辆在道路区域的运动模型,构成交通流仿真模型;

所述交通流仿真模型作为离线仿真模型。

5. 根据权利要求4所述的一种基于强化学习的高速公路交通管控方法,其特征在于,

设置车辆在道路区域的运动模型的方法包括:

设置车辆属性,所述车辆属性包括位置和速度,所述位置和速度由流入交通流数据确

定；

设置车辆变速控制函数，所述车辆变速控制函数以预设周期改变车辆的速度，所述车辆变速控制函数的输入为限速、与前车的距离、车辆加速度和车辆当前周期的速度，所述限速由车辆位置所处的区域类型确定。

6. 根据权利要求3至5任一项所述的一种基于强化学习的高速公路交通管控方法，其特征在于，

步骤三中，训练所述离线仿真模型的方法包括：

选定交通管控模式，所述交通管控模式包括可变限速控制模式和匝道限流模式；

根据选定的交通管控模式训练所述离线仿真模型；

步骤四中，尝试生成的优化管控策略与选定的交通管控模式相符，使得单位时间内取得最大通过量的优化管控策略进行执行。

7. 根据权利要求6所述的一种基于强化学习的高速公路交通管控方法，其特征在于，

选定交通管控模式为可变限速控制模式时，训练所述离线仿真模型的方法包括：

步骤A1) 获取当前状态 $s_t = \{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$ ；

步骤A2) 生成并执行一个动作 $a_t, a_t = \{V_0, V_1, \dots, V_{L-1}\}$ ；

步骤A3) 通过离线仿真模型演算，得到下一状态 s_{t+1} ，并获取强化信号 r_{t+1} ；

步骤A4) 更新Q值，依据强化信号 r_{t+1} 按照下式计算Q值：

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \gamma \cdot \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1})$$

其中，其中 s_t 表示时刻表t的状态， a_t 表示时刻表t的动作， r_t 表示时刻表t的奖励， γ 为Q学习的预设参数；

步骤A5) 若Q值收敛，则离线仿真模型训练结束；反之，则返回步骤A1)。

8. 根据权利要求3至5任一项所述的一种基于强化学习的高速公路交通管控方法，其特征在于，

选定交通管控模式为匝道限流模式时，训练所述离线仿真模型的方法包括：

步骤B1) 获取当前状态 $s_t = \{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$ ；

步骤B2) 生成并执行一个动作 $a_t, a_t = \{P_0, P_1, \dots, P_{L-1}\}$ ；

步骤B3) 通过离线仿真模型演算，得到下一状态 s_{t+1} ，并获取强化信号 r_{t+1} ；

步骤B4) 更新Q值，依据强化信号 r_{t+1} 按照下式计算Q值：

$$Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \gamma \cdot \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1})$$

其中，其中 s_t 表示时刻表t的状态， a_t 表示时刻表t的动作， r_t 表示时刻表t的奖励， γ 为Q学习的预设参数；

步骤B5) 若Q值收敛，则离线仿真模型训练结束；反之，则返回步骤A1)。

9. 根据权利要求3至5任一项所述的一种基于强化学习的高速公路交通管控方法，其特征在于，

步骤四中，将所述智能体部署到实际预测控制模型中的方法包括：

获取目标交通管控区域的道路信息；

根据交通管控区域的道路信息建立道路模型，替换智能体中的道路模型；

将当前流入交通流数据及流出交通流数据导入到智能体，完成智能体的部署。

10. 根据权利要求3至5任一项所述的一种基于强化学习的高速公路交通管控方法，其

特征在于，

步骤四中，周期性计算获得优化管控策略的方法包括：

步骤C1) 获取目标交通管控区域的前状态；

步骤C2) 生成管控策略，将所述管控策略与当前状态导入智能体，获得当前管控策略下的单位时间通过量；

步骤C3) 不断更改管控策略，直到找到单位时间通过量最大的管控策略或者达到最大更改次数；

步骤C4) 使得单位时间通过量最大的管控策略作为优化管控策略，记为 $\{u_0^*, u_1^*, \dots, u_{L-1}^*\}$ ，执行所述优化管控策略。

一种基于强化学习的高速公路交通管控方法

技术领域

[0001] 本发明涉及交通控制技术领域,具体涉及一种基于强化学习的高速公路交通管控方法。

背景技术

[0002] 由于土地资源紧张、环境保护等因素制约,高速公路建设步伐变缓,同时汽车保有量不断上升,部分路段交通流量趋于饱和,高速公路在供给能力和效率方面越来越难以满足日益增长的服务需求,亟需通过信息化和智能化手段来提升高速公路运行效率与安全。交通管控作为提升高速公路运行效率与安全的有效手段,其主要通过可变限速和匝道限流等措施来优化道路资源的时空分配,从而提高道路运行效率,进而减缓交通拥堵和提升服务质量。可变限速控制方法通过调节瓶颈区上游区域的限速,将进入瓶颈区的车辆数控制在一定范围内,从而提高瓶颈区运行效率。匝道限流控制方法通过限制匝道汇入的车流量,从而缓解瓶颈区拥堵。

[0003] 目前在高速公路交通管控过程中,交通管控策略大多是管理人员依据经验来制定的。由于仅基于经验,现有交通管控策略的实施效果并不理想,难以符合实际需求。同时,人工方法也无法适应动态变化的交通需求与流量。因此,一种自动化和智能化的交通管控方法显得至关重要。

[0004] 现有技术公开了一种模型预测控制(Model Predictive Control,MPC)技术。作为自动化交通管控的重要方法之一,能够基于交通流预测模型,实现对交通运行态势的预测,并通过求解全局最优或局部最优化问题来确定最佳策略。然而这种方法受限于交通流预测模型的精度,而实际交通运行条件时变复杂,交通流的预测精准低,控制效果无法保证。因此有必要研究新的能够更为有效、更为准确的对交通流进行管控的技术。

发明内容

[0005] 本发明所要解决的技术问题:目前缺乏对交通流进行准确管控方案的技术问题。提出了一种基于强化学习的高速公路交通管控方法,能够有效提高高速公路的通行能力。

[0006] 解决上述技术问题,本发明采用如下技术方案:一种基于强化学习的高速公路交通管控方法,包括以下步骤:

[0007] 步骤一:读取交通管控区域的道路信息,对道路进行区域划分;

[0008] 步骤二:根据区域划分构建离线仿真模型,所述离线仿真模型对交通管控区域的交通进行仿真;

[0009] 步骤三:使用历史交通数据训练所述离线仿真模型,训练后的离线仿真模型记为智能体;

[0010] 步骤四:将所述智能体部署到实际预测控制模型中,以单位时间最大通过量为目标函数,周期性计算获得优化管控策略,执行所述优化管控策略。

[0011] 作为优选,步骤一中,进行区域划分的方法包括:

[0012] 将交通管控区域的道路划分为以下区域类型：可变限速区域K0、加速区域J、合流区域H和匝道区域Z0；

[0013] 将可变限速区域K0的上游区域划分为L-1个主路检测区域，分别记为 K_1, K_2, \dots, K_{L-1} ；

[0014] 将匝道区域Z0的上游区域划分为L-1个匝道检测区域，分别记为 Z_1, Z_2, \dots, Z_{L-1} 。

[0015] 作为优选，步骤一中，构建离线仿真模型的方法包括：

[0016] 设置离线仿真模型的状态量，所述状态量包括：可变限速区域K0及其上游区域的交通流密度 $\{qK_0, qK_1, \dots, qK_{L-1}\}$ ，合流区域交通流密度qH、匝道区域及其上游区域的交通流密度 $\{qZ_0, qZ_1, \dots, qZ_{L-1}\}$ ，状态量记为 $\{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$ ；

[0017] 设置离线仿真模型的动作，设置控制步长L，将所述动作设置为可变限速区域K0内在L步长内的限速值序列 $\{V_0, V_1, \dots, V_{L-1}\}$ 和匝道区域Z0在L步长内的汇入率序列 $\{P_0, P_1, \dots, P_{L-1}\}$ ；

[0018] 设置离线仿真模型的奖励，在控制步长L内的总通行时间TTT作为奖励，

$$[0019] \quad TTT = T \sum_{l=0}^{L-1} \left[T \sum_{t=0}^{i-1} (q_{in}(t) - q_{out}(t)) \right]$$

[0020] 其中，T表示控制步长的间隔时长， $q_{in}(t)$ 表示t时刻控制区域内的流入交通流密度， $q_{out}(t)$ 表示t时刻控制区域内的流出交通流密度。

[0021] 作为优选，步骤一中，构建离线仿真模型的方法还包括构建仿真环境，

[0022] 构建仿真环境的方法包括：

[0023] 根据交通管控区域的道路信息建立道路模型；

[0024] 将流入交通流数据及流出交通流数据导入到道路模型，设置车辆在道路区域的运动模型，构成交通流仿真模型；

[0025] 所述交通流仿真模型作为离线仿真模型。

[0026] 作为优选，设置车辆在道路区域的运动模型的方法包括：

[0027] 设置车辆属性，所述车辆属性包括位置和速度，所述位置和速度由流入交通流数据确定；

[0028] 设置车辆变速控制函数，所述车辆变速控制函数以预设周期改变车辆的速度，所述车辆变速控制函数的输入为限速、与前车的距离、车辆加速度和车辆当前周期的速度，所述限速由车辆位置所处的区域类型确定。

[0029] 分别设置车辆在可变限速区域K0、加速区域J、合流区域H和匝道区域Z0的车辆变速控制函数。

[0030] 作为优选，步骤三中，训练所述离线仿真模型的方法包括：

[0031] 选定交通管控模式，所述交通管控模式包括可变限速控制模式和匝道限流模式；

[0032] 根据选定的交通管控模式训练所述离线仿真模型；

[0033] 步骤四中，尝试生成的优化管控策略与选定的交通管控模式相符，使得单位时间内取得最大通过量的优化管控策略进行执行。

[0034] 作为优选，选定交通管控模式为可变限速控制模式时，训练所述离线仿真模型的方法包括：

- [0035] 步骤A1) 获取当前状态 $s_t = \{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$;
- [0036] 步骤A2) 生成并执行一个动作 $a_t, a_t = \{V_0, V_1, \dots, V_{L-1}\}$;
- [0037] 步骤A3) 通过离线仿真模型演算, 得到下一状态 s_{t+1} , 并获取强化信号 r_{t+1} ;
- [0038] 步骤A4) 更新Q值, 依据强化信号 r_{t+1} 按照下式计算Q值:
- [0039] $Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \gamma \cdot \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1})$
- [0040] 其中, 其中 s_t 表示时刻表t的状态, a_t 表示时刻表t的动作, r_t 表示时刻表t的奖励, γ 为Q学习的预设参数;
- [0041] 步骤A5) 若Q值收敛, 则离线仿真模型训练结束; 反之, 则返回步骤A1)。
- [0042] 作为优选, 选定交通管控模式为匝道限流模式时, 训练所述离线仿真模型的方法包括:
- [0043] 步骤B1) 获取当前状态 $s_t = \{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$;
- [0044] 步骤B2) 生成并执行一个动作 $a_t, a_t = \{P_0, P_1, \dots, P_{L-1}\}$;
- [0045] 步骤B3) 通过离线仿真模型演算, 得到下一状态 s_{t+1} , 并获取强化信号 r_{t+1} ;
- [0046] 步骤B4) 更新Q值, 依据强化信号 r_{t+1} 按照下式计算Q值:
- [0047] $Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \gamma \cdot \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1})$
- [0048] 其中, 其中 s_t 表示时刻表t的状态, a_t 表示时刻表t的动作, r_t 表示时刻表t的奖励, γ 为Q学习的预设参数;
- [0049] 步骤B5) 若Q值收敛, 则离线仿真模型训练结束; 反之, 则返回步骤A1)。
- [0050] 作为优选, 步骤四中, 将所述智能体部署到实际预测控制模型中的方法包括:
- [0051] 获取目标交通管控区域的道路信息;
- [0052] 根据交通管控区域的道路信息建立道路模型, 替换智能体中的道路模型;
- [0053] 将当前流入交通流数据及流出交通流数据导入到智能体, 完成智能体的部署。
- [0054] 作为优选, 步骤四中, 周期性计算获得优化管控策略的方法包括:
- [0055] 步骤C1) 获取目标交通管控区域的前状态;
- [0056] 步骤C2) 生成管控策略, 将所述管控策略与当前状态导入智能体, 获得当前管控策略下的单位时间通过量;
- [0057] 步骤C3) 不断更改管控策略, 直到找到单位时间通过量最大的管控策略或者达到最大更改次数;
- [0058] 步骤C4) 使得单位时间通过量最大的管控策略作为优化管控策略, 记为 $\{u_0^*, u_1^*, \dots, u_{L-1}^*\}$, 执行所述优化管控策略。
- [0059] 本发明的有益技术效果包括: 1) 利用强化学习训练智能体, 克服模型预测控制对于交通流预测模型过度依赖的缺点, 提高了预测的准确度, 从而实现有效管控策略的准确解算, 获得更为优化的管控策略; 2) 利用预测控制模型滚动的优化管控策略, 根据管控区域的交通状态计算最优管控策略并不断更新, 实现全时域的优化控制; 3) 与传统管控策略优化方法不同, 本发明管控策略的优化不仅仅根据可变限速和匝道区域的交通流数据, 还根据了可变限速和匝道区域上游的交通流数据, 取得了更优化的交通管控效果, 更加有效的降低高速公路瓶颈区的车辆总通行时间, 从而显著提升交通运行效率。
- [0060] 本发明的其他特点和优点将会下面的具体实施方式、附图中详细的揭露。

附图说明

- [0061] 下面结合附图对本发明做进一步的说明：
- [0062] 图1为本发明实施例高速公路交通管控方法流程示意图。
- [0063] 图2为本发明实施例进行区域划分方法流程示意图。
- [0064] 图3为本发明实施例高速公路区域划分示意图。
- [0065] 图4为本发明实施例构建仿真环境方法流程示意图。
- [0066] 图5为本发明实施例训练离线仿真模型方法流程示意图。
- [0067] 图6为本发明实施例可变限速控制模式训练方法流程示意图。
- [0068] 图7为本发明实施例匝道限流模式训练方法流程示意图。
- [0069] 图8为本发明实施例智能体部署到实际预测控制模型方法流程示意图。
- [0070] 图9为本发明实施例周期性计算获得优化管控策略方法流程示意图。
- [0071] 其中：100、划分区域，101、新增划分区域，102、传统划分区域。

具体实施方式

[0072] 下面结合本发明实施例的附图对本发明实施例的技术方案进行解释和说明，但下述实施例仅为本发明的优选实施例，并非全部。基于实施方式中的实施例，本领域技术人员在没有做出创造性劳动的前提下所获得其他实施例，都属于本发明的保护范围。

[0073] 在下文描述中，出现诸如术语“内”、“外”、“上”、“下”、“左”、“右”等指示方位或者位置关系仅是为了方便描述实施例和简化描述，而不是指示或暗示所指的装置或者元件必须具有特定的方位、以特定的方位构造和操作，因此不能理解为对本发明的限制。

[0074] 在介绍本实施例方案前，对本实施例的应用背景做相关介绍。

[0075] 正如背景技术所指出，由于土地资源紧张、环境保护等因素制约，高速公路建设步伐变缓，同时汽车保有量不断上升，部分路段交通流量趋于饱和，高速公路在供给能力和效率方面越来越难以满足日益增长的服务需求，亟需通过信息化和智能化手段来提升高速公路运行效率与安全。借助信息化和智能化的手段，能够更加科学和高效的管控车辆在高速公路上的通行，从而提高高速公路的通行效率。但目前缺乏有效实现高效率生成交通管控策略的方案。

[0076] 车辆在高速公路上通行，车辆的速度不仅和车流量密度有关，还与车辆所处的高速公路的区域有关。因此，欲实现生成交通管控策略，需要建立尽可能贴近实际目标交通管控区域车辆的通行规律的车辆运动的模拟。或者直接对道路进行建模，或者对交通流进行建模仿真。由于车辆的多样性和司机技术的不同，车辆并没有统一的运动模型。因此需要借助深度学习的手段，通过对实际的历史数据的学习模拟，完成对车辆运动模型的建立，进而构建出车流量模型，进而建立出相应的道路模型，即本实施例所称的智能体。

[0077] 一种基于强化学习的高速公路交通管控方法，请参阅附图1，包括以下步骤：

[0078] 步骤一：读取交通管控区域的道路信息，对道路进行区域划分；

[0079] 步骤二：根据区域划分构建离线仿真模型，离线仿真模型对交通管控区域的交通进行仿真；

[0080] 步骤三：使用历史交通数据训练离线仿真模型，训练后的离线仿真模型记为智能体；

[0081] 步骤四：将智能体部署到实际预测控制模型中，以单位时间最大通过量为目标函数，周期性计算获得优化管控策略，执行优化管控策略。

[0082] 对道路进行区域划分后，不仅能够对车辆进行速度控制模式的限定，同时也为道路的交通管控提供了对象。本实施例中，管控策略即是针对相应的道路区域进行的。不同的道路区域能够进行不同的限速及限制通行，从而形成管控策略。在不同的道路区域，车辆的速度控制和通行规则不同。对应的离线仿真模型将提供这些道路区域的仿真。

[0083] 请参阅附图2，步骤一中，进行区域划分的方法包括：

[0084] 步骤S101) 将交通管控区域的道路划分为以下区域类型：可变限速区域K0、加速区域J、合流区域H和匝道区域Z0；

[0085] 步骤S102) 将可变限速区域K0的上游区域划分为L-1个主路检测区域，分别记为 K_1, K_2, \dots, K_{L-1} ；

[0086] 步骤S103) 将匝道区域Z0的上游区域划分为L-1个匝道检测区域，分别记为 Z_1, Z_2, \dots, Z_{L-1} 。

[0087] 可变限速区域K0及其上游区域能够划分不同的限速值，加速区域L的通信规则由公知常识所确定，在合流区域H内，若车辆之间的间距超过规定值，则车辆能够不减速、不等待的交替通行，但若车辆之间的间距达不到规定值，则车辆需要减速或者停车等待，并交替通行。匝道区域Z0具有匝道限速，且匝道区域Z0能够被限流，即被关闭。

[0088] 请参阅附图3，为本实施例将高速公路划分的四个区域类型。相对于传统划分区域102仅划分可变限速区域K0、加速区域J、合流区域H和匝道区域Z0，本实施例新增划分区域101包括可变限速区域K0的上游区域和匝道区域Z0的上游区域。通过改进区域划分方式，细化了高速公路管控策略的控制，提高了管控策略的准确度，能够更有效的提高高速公路的通行效率。

[0089] 步骤一中，构建离线仿真模型的方法包括：

[0090] 设置离线仿真模型的状态量，状态量包括：可变限速区域K0及其上游区域的交通流密度 $\{q_{K_0}, q_{K_1}, \dots, q_{K_{L-1}}\}$ ，合流区域交通流密度 q_H 、匝道区域及其上游区域的交通流密度 $\{q_{Z_0}, q_{Z_1}, \dots, q_{Z_{L-1}}\}$ ，状态量记为 $\{q_{K_0}, \dots, q_{K_{L-1}}, q_H, q_{Z_0}, \dots, q_{Z_{L-1}}\}$ ；

[0091] 设置离线仿真模型的动作，设置控制步长L，将动作设置为可变限速区域K0内在L步长内的限速值序列 $\{V_0, V_1, \dots, V_{L-1}\}$ 和匝道区域Z0在L步长内的汇入率序列 $\{P_0, P_1, \dots, P_{L-1}\}$ ；

[0092] 设置离线仿真模型的奖励，在控制步长L内的总通行时间TTT作为奖励，

$$[0093] \quad TTT = T \sum_{l=0}^{L-1} \left[T \sum_{t=0}^{i-1} (q_{in}(t) - q_{out}(t)) \right]$$

[0094] 其中，T表示控制步长的间隔时长， $q_{in}(t)$ 表示t时刻控制区域内的流入交通流密度， $q_{out}(t)$ 表示t时刻控制区域内的流出交通流密度。

[0095] 步骤一中，构建离线仿真模型的方法还包括构建仿真环境，

[0096] 请参阅附图4，构建仿真环境的方法包括：

[0097] 步骤S201) 根据交通管控区域的道路信息建立道路模型；

[0098] 步骤S202) 将流入交通流数据及流出交通流数据导入到道路模型，设置车辆在道

路区域的运动模型,构成交通流仿真模型;

[0099] 步骤S203) 交通流仿真模型作为离线仿真模型。

[0100] 现有技术公开了多种车辆的运动模型的仿真方式,通过建立离线仿真模型能够建立起高速公路的通行仿真,与历史数据进行对比,能够优化车辆的运动模型,使车辆的运动模型与实际高速公路上车辆的运动尽可能相符。但不同司机驾驶的不同类型或者不同型号的车辆,车辆的运动规律不会全部相同。只是建立的统一的车辆的运动模型,能够整体上反映高速公路上车辆的运行规律,从而实现交通流的仿真。

[0101] 另一方面,本实施例提供了具体的设置车辆在道路区域的运动模型的方法,包括:设置车辆属性,车辆属性包括位置和速度,位置和速度由流入交通流数据确定;

[0102] 设置车辆变速控制函数,车辆变速控制函数以预设周期改变车辆的速度,车辆变速控制函数的输入为限速、与前车的距离、车辆加速度和车辆当前周期的速度,限速由车辆位置所处的区域类型确定。通过设置车辆的位置和速度属性,结合车辆变速控制函数,能够通过控制车辆的速度,不断仿真车辆的位置变化,实现车辆从交通管控区域的入口位置,移动到交通管控区域出口位置的过程中的运行规律。

[0103] 更进一步的,分别设置车辆在可变限速区域K0、加速区域J、合流区域H和匝道区域Z0的车辆变速控制函数。由于不同的划分区域的通行规则不同,因此有可能引起车辆变速控制函数的不准确,为进一步提高仿真的准确度,分别设置车辆在可变限速区域K0、加速区域J、合流区域H和匝道区域Z0的车辆变速控制函数。

[0104] 请参阅附图5,步骤三中,训练离线仿真模型的方法包括:

[0105] 步骤S301) 选定交通管控模式,交通管控模式包括可变限速控制模式和匝道限流模式;

[0106] 步骤S302) 根据选定的交通管控模式训练离线仿真模型;

[0107] 步骤S303) 步骤四中,尝试生成的优化管控策略与选定的交通管控模式相符,使得单位时间内取得最大通过量的优化管控策略进行执行。

[0108] 可变限速控制模式和匝道限流模式均能够起到改善高速公路通行效率的效果。具体选择哪种方式由实际的管控条件确定。

[0109] 请参阅附图6,当选定交通管控模式为可变限速控制模式时,训练离线仿真模型的方法包括:

[0110] 步骤A1) 获取当前状态 $s_t = \{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$;

[0111] 步骤A2) 生成并执行一个动作 $a_t, a_t = \{V_0, V_1, \dots, V_{L-1}\}$;

[0112] 步骤A3) 通过离线仿真模型演算,得到下一状态 s_{t+1} ,并获取强化信号 r_{t+1} ;

[0113] 步骤A4) 更新Q值,依据强化信号 r_{t+1} 按照下式计算Q值:

[0114] $Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \gamma \cdot \max_{a_{t+1}} Q_t(s_{t+1}, a_{t+1})$

[0115] 其中,其中 s_t 表示时刻表t的状态, a_t 表示时刻表t的动作, r_t 表示时刻表t的奖励, γ 为Q学习的预设参数;

[0116] 步骤A5) 若Q值收敛,则离线仿真模型训练结束;反之,则返回步骤A1)。

[0117] 动作 a_t 决定了可变限速区域K0及其上游区域的限速值,分别记为 $V_0, V_1, \dots, V_{L-10}$ 。当离线仿真模型的仿真结果使Q值收敛时,表明找到了最佳的动作 a_t ,即为最佳的管控策略。

[0118] 请参阅附图7,当选定交通管控模式为匝道限流模式时,训练离线仿真模型的方法包括:

[0119] 步骤B1) 获取当前状态 $s_t = \{qK_0, \dots, qK_{L-1}, qH, qZ_0, \dots, qZ_{L-1}\}$;

[0120] 步骤B2) 生成并执行一个动作 $a_t, a_t = \{P_0, P_1, \dots, P_{L-1}\}$;

[0121] 步骤B3) 通过离线仿真模型演算,得到下一状态 s_{t+1} ,并获取强化信号 r_{t+1} ;

[0122] 步骤B4) 更新Q值,依据强化信号 r_{t+1} 按照下式计算Q值:

[0123] $Q_{t+1}(s_t, a_t) \leftarrow Q_t(s_t, a_t) + \gamma \cdot \max_{a'} Q_t(s_{t+1}, a')$

[0124] 其中,其中 s_t 表示时刻表t的状态, a_t 表示时刻表t的动作, r_t 表示时刻表t的奖励, γ 为Q学习的预设参数;

[0125] 步骤B5) 若Q值收敛,则离线仿真模型训练结束;反之,则返回步骤A1)。

[0126] 动作 a_t 决定了匝道区域Z0及其上游区域的汇入状态,分别记为 P_0, P_1, \dots, P_{L-1} 。 P_0, P_1, \dots, P_{L-1} 的取值均在0和1之间取值。取值为0表示对应的匝道被限流关闭,取值为1,表示对应的匝道能够正常通行。离线仿真模型的仿真结果使Q值收敛时,表明找到了最佳的动作 a_t ,即为最佳的管控策略。

[0127] 请参阅附图8,步骤四中,将智能体部署到实际预测控制模型中的方法包括:

[0128] 步骤S401) 获取目标交通管控区域的道路信息;

[0129] 步骤S402) 根据交通管控区域的道路信息建立道路模型,替换智能体中的道路模型;

[0130] 步骤S403) 将当前流入交通流数据及流出交通流数据导入到智能体,完成智能体的部署。

[0131] 智能体包括了道路模型和交通流仿真模型,交通流仿真模型又包含了车辆的运动模型。对于不同的交通管控区域,车辆的运动模型是相同的。将智能体的道路模型替换为目标交通管控区域的道路模型,即可完成智能体的部署。

[0132] 步骤四中,周期性计算获得优化管控策略的方法包括:

[0133] 步骤C1) 获取目标交通管控区域的前状态;

[0134] 步骤C2) 生成管控策略,将管控策略与当前状态导入智能体,获得当前管控策略下的单位时间通过量;

[0135] 步骤C3) 不断更改管控策略,直到找到单位时间通过量最大的管控策略或者达到最大更改次数;

[0136] 步骤C4) 使得单位时间通过量最大的管控策略作为优化管控策略,记为 $\{u_0^*, u_1^*, \dots, u_{L-1}^*\}$,执行优化管控策略。

[0137] 通过周期性的不断优化计算,持续的根据实际情况的变化,获得最优的管控策略,有效提升高速公路的通行效率。当选定交通管控模式为可变限速控制模式时, $\{u_0^*, u_1^*, \dots, u_{L-1}^*\}$ 表示取得单位时间最大通过量的 $\{V_0, V_1, \dots, V_{L-1}\}$,当选定交通管控模式为匝道限流模式时, $\{u_0^*, u_1^*, \dots, u_{L-1}^*\}$ 表示取得单位时间最大通过量的 $\{P_0, P_1, \dots, P_{L-1}\}$ 。

[0138] 本实施例的有益技术效果包括:1) 利用强化学习训练智能体,克服模型预测控制对于交通流预测模型过度依赖的缺点,提高了预测的准确度,从而实现有效管控策略的准确解算,获得更为优化的管控策略;2) 利用预测控制模型滚动的优化管控策略,根据管控区

域的交通状态计算最优管控策略并不断更新,实现全时域的优化控制;3) 与传统管控策略优化方法不同,本发明管控策略的优化不仅仅根据可变限速和匝道区域的交通流数据,还根据了可变限速和匝道区域上游的交通流数据,取得了更优化的交通管控效果,更加有效的降低高速公路瓶颈区的车辆总通行时间,从而显著提升交通运行效率。

[0139] 以上所述,仅为本发明的具体实施方式,但本发明的保护范围并不局限于此,熟悉该本领域的技术人员应该明白本发明包括但不限于附图和上面具体实施方式中描述的内容。任何不偏离本发明的功能和结构原理的修改都将包括在权利要求书的范围中。

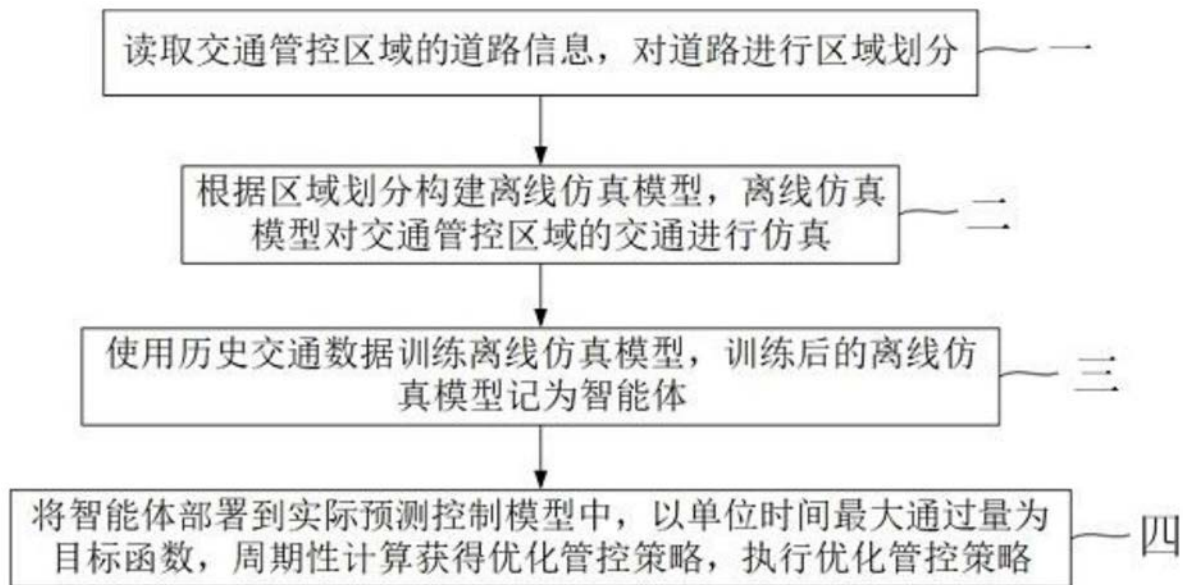


图1



图2

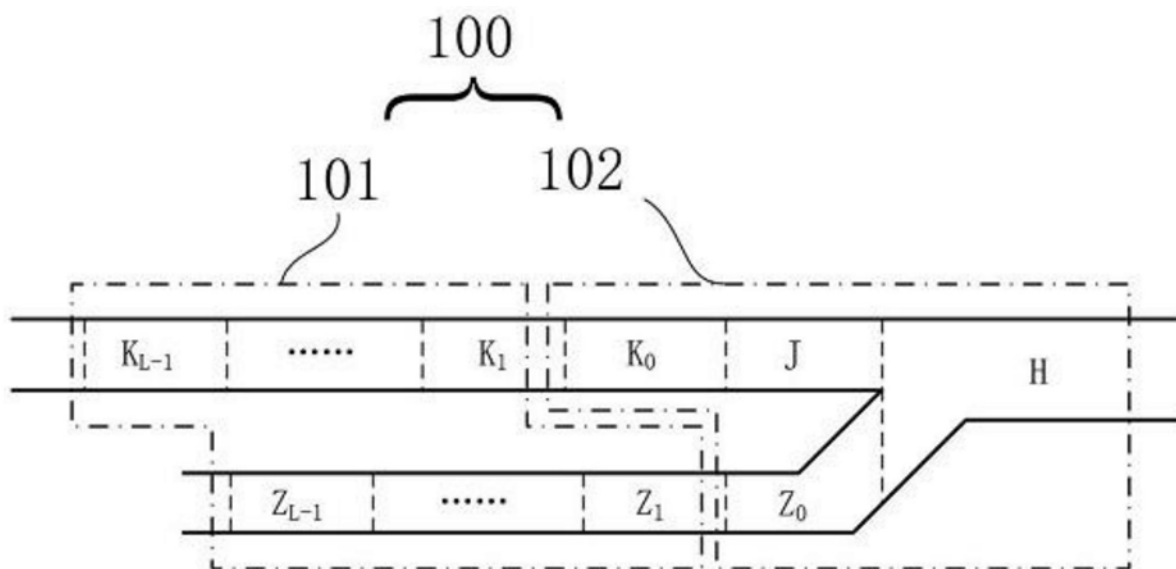


图3

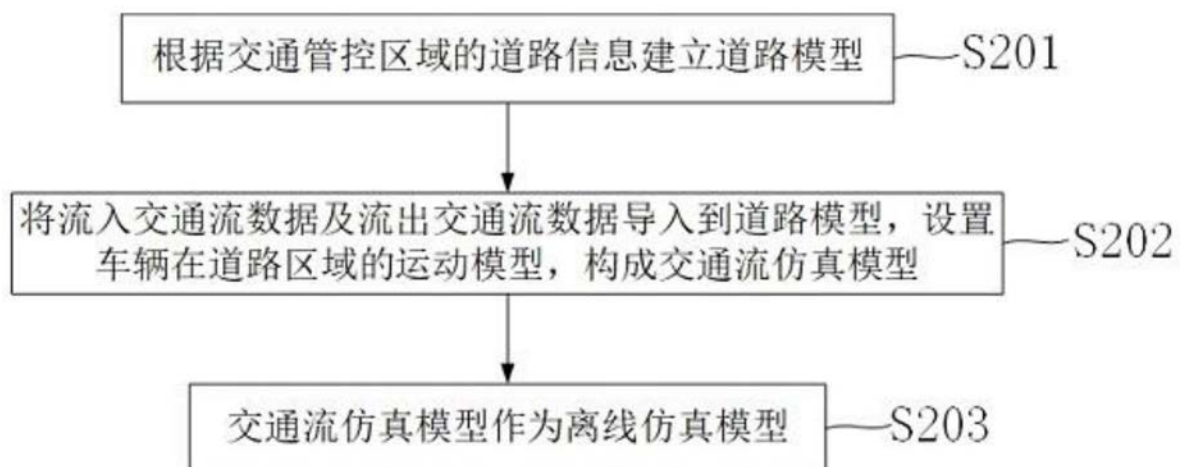


图4

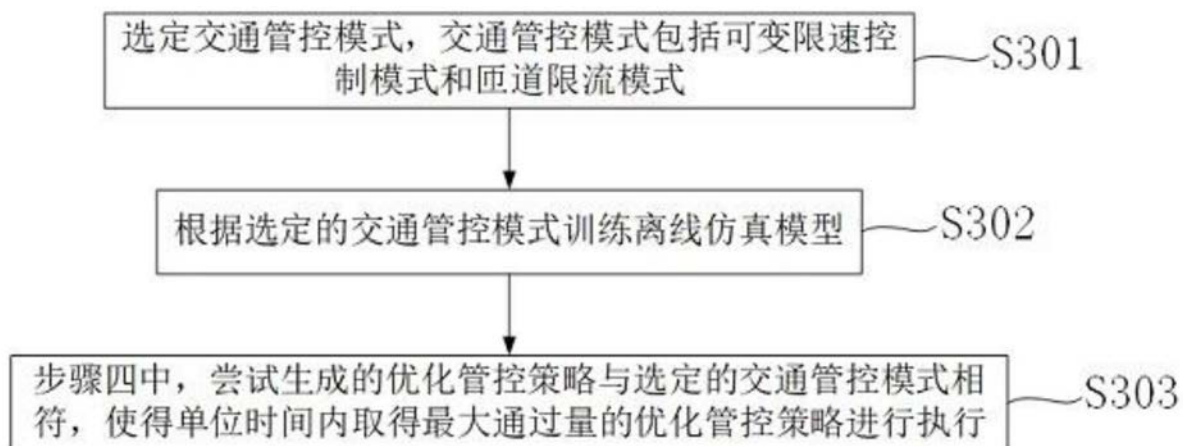


图5

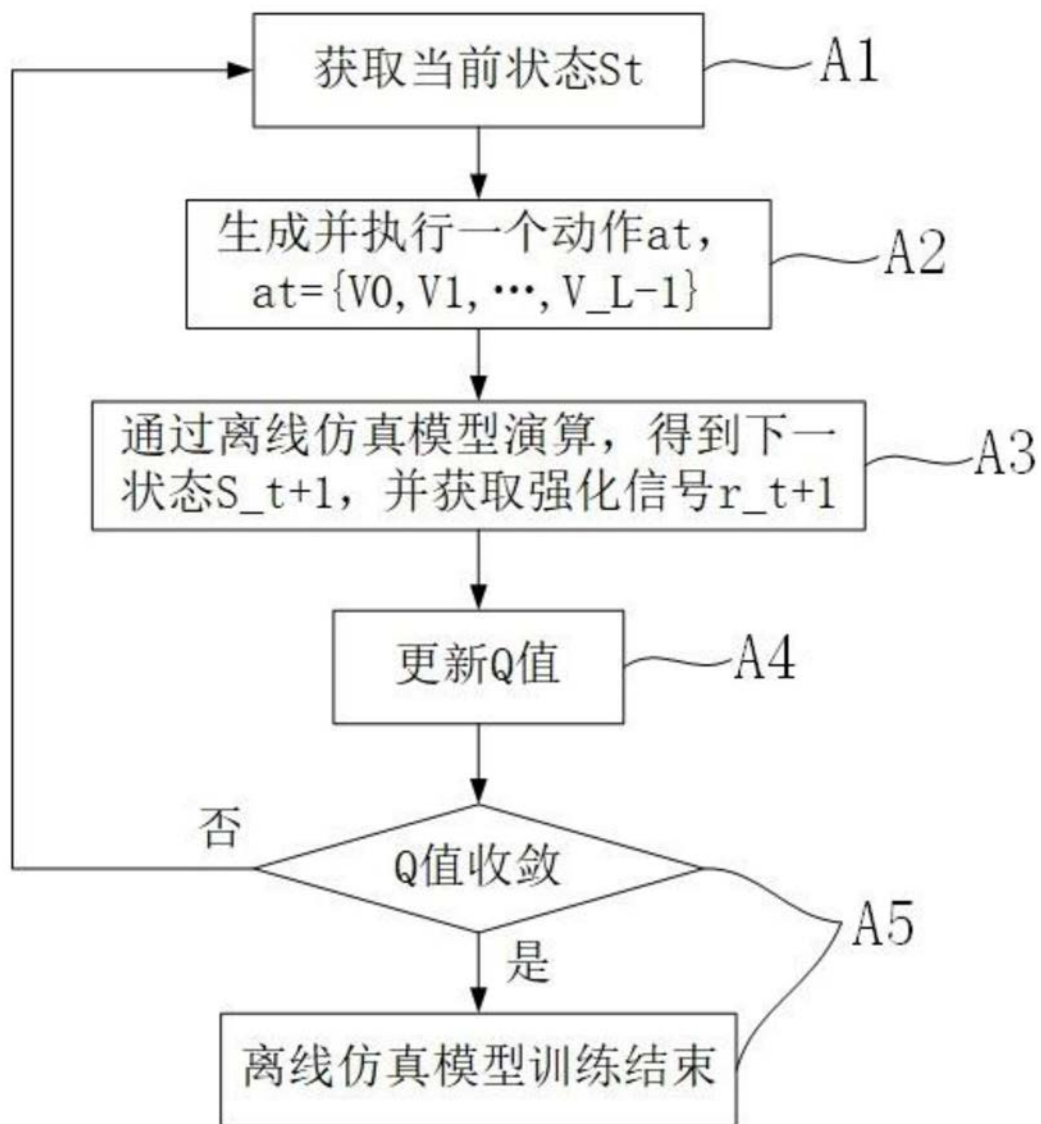


图6

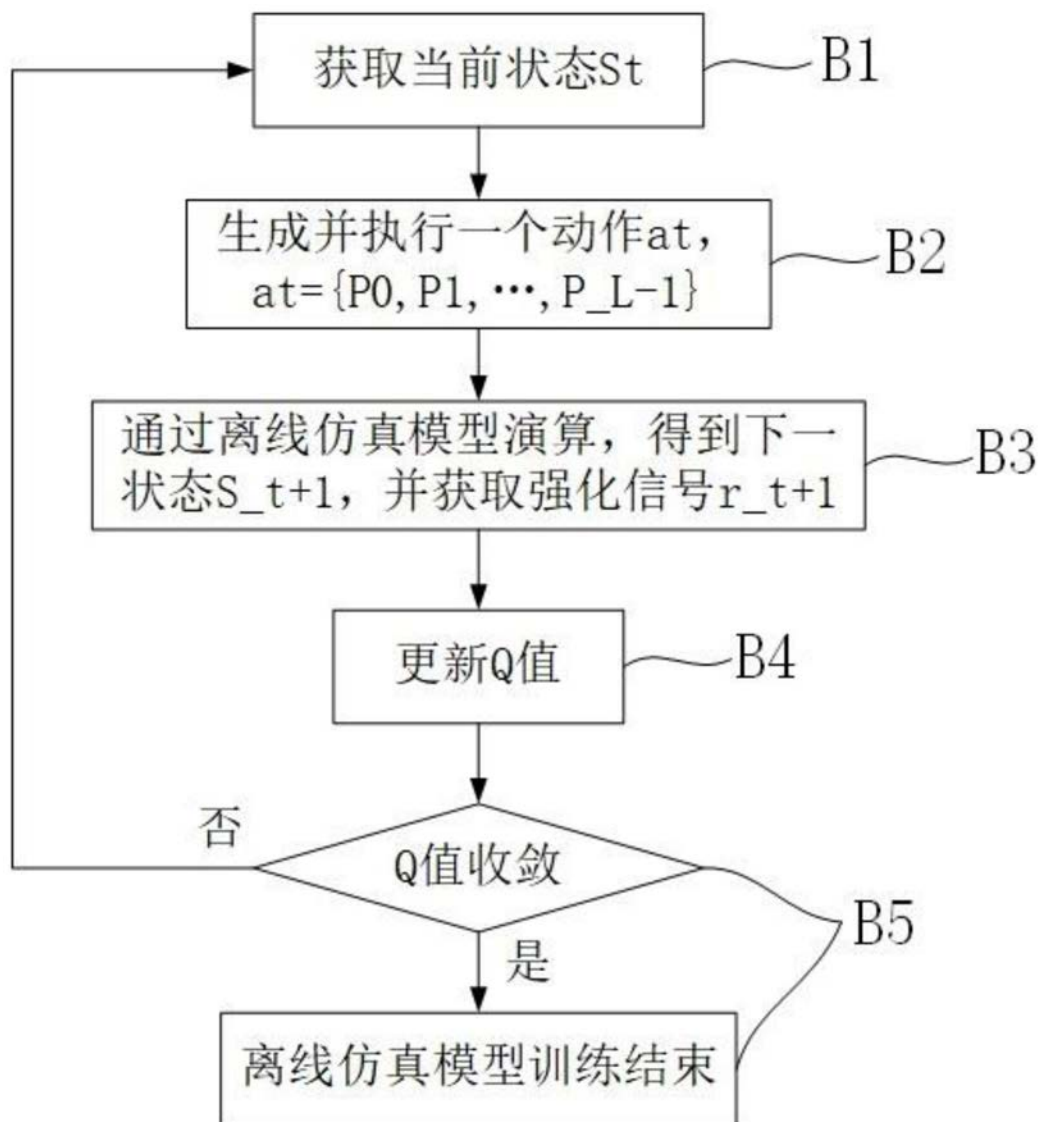


图7

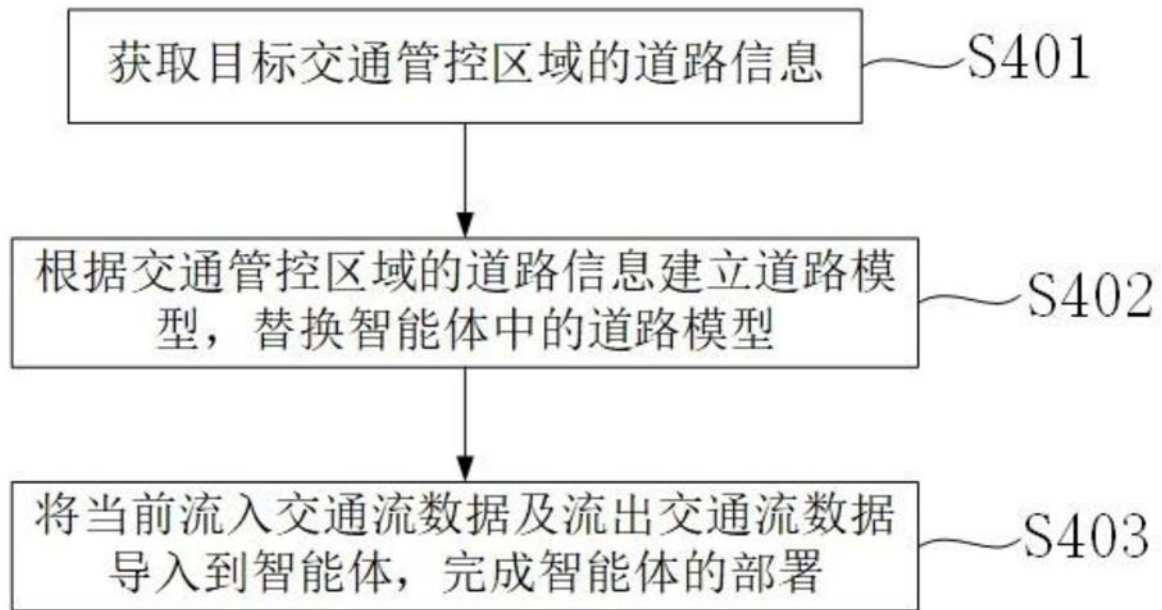


图8

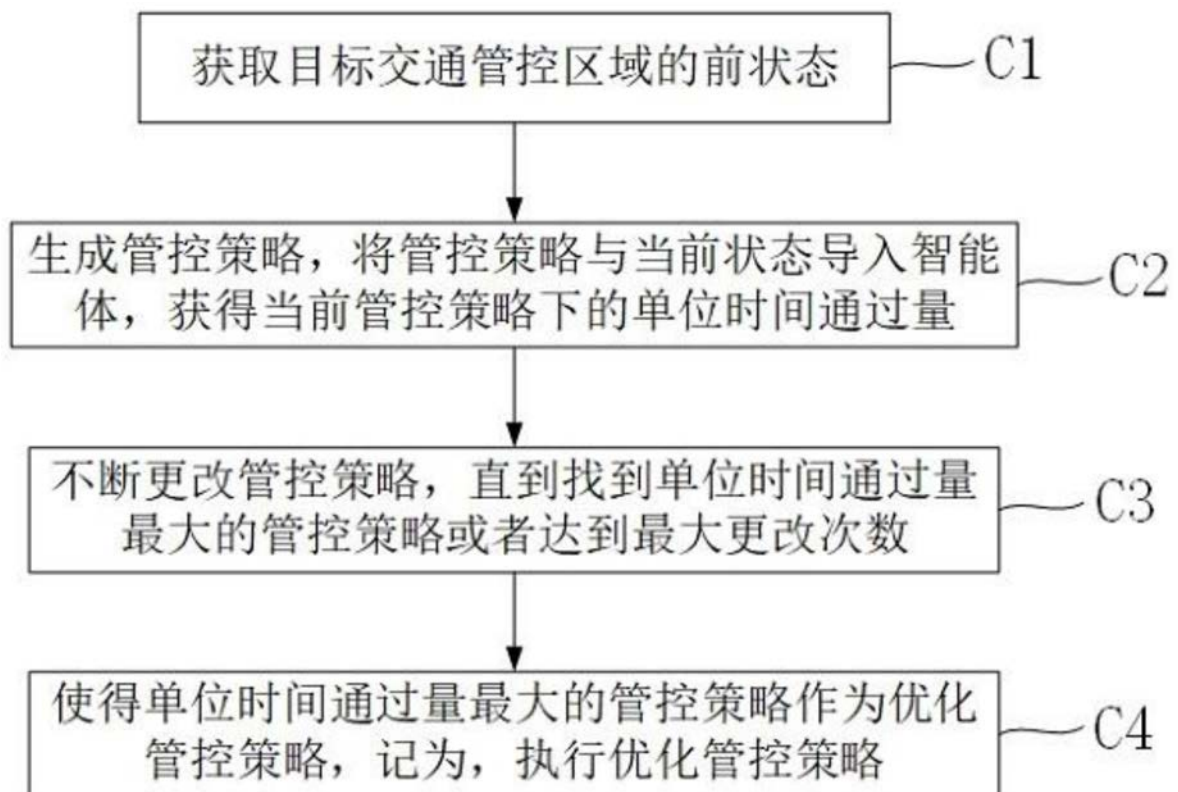


图9