



(12) 发明专利申请

(10) 申请公布号 CN 112292719 A

(43) 申请公布日 2021. 01. 29

(21) 申请号 201980042562.3

(74) 专利代理机构 中国专利代理(香港)有限公司 72001

(22) 申请日 2019.05.22

代理人 臧永杰 刘春元

(30) 优先权数据

102018210280.5 2018.06.25 DE

(51) Int.Cl.

G08G 1/16 (2006.01)

(85) PCT国际申请进入国家阶段日

B60W 30/095 (2006.01)

2020.12.24

G06K 9/00 (2006.01)

(86) PCT国际申请的申请数据

PCT/EP2019/063232 2019.05.22

(87) PCT国际申请的公布数据

WO2020/001867 DE 2020.01.02

(71) 申请人 罗伯特·博世有限公司

地址 德国斯图加特

(72) 发明人 S·J·厄特萨米

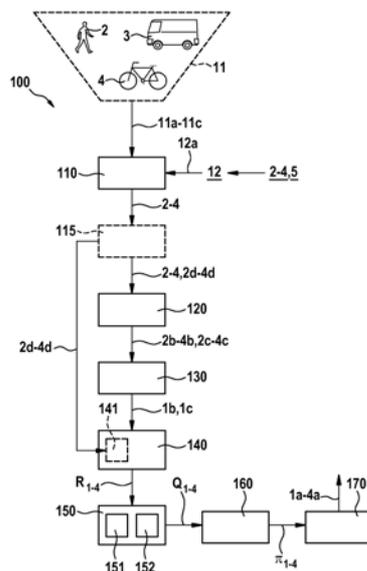
权利要求书2页 说明书9页 附图4页

(54) 发明名称

使自我车辆的轨迹与运动的外来对象适配

(57) 摘要

一种用于预报自我车辆(1)的环境(11)中的外来对象(2-4)的轨迹(2a-4a)以及用于为自我车辆(1)确定与之适配的自身未来轨迹(1a)的方法(100、200),所述方法具有以下步骤:标识(110)外来对象(2-4);确定(120)所述外来对象(2-4)中的每一个的运动通向哪个近距离目的地(2b-4b)并且该运动根据哪些基本规则(2c-4c)进行;确定(130)自我车辆(1)的运动通向哪个近距离目的地(1b)以及该运动根据哪些基本规则(1c)进行;为自我车辆(1)以及为外来对象(2-4)分别设置(140)质量函数RI-4;为自我车辆(1)以及为外来对象(2-4)分别设置(150)质量度量 Q_{1-4} ;确定(160)自我车辆和外来对象(2-4)的使质量度量 Q_{1-4} 最大化的最佳运动策略 π_{1-4} ;从最佳运动策略 π_{1-4} 中确定(170)所寻求的轨迹(1a-4a)。用于控制自我车辆1的方法(100)。所属的计算机程序。



1. 一种用于预报自我车辆(1)的环境(11)中的外来对象(2-4)的轨迹(2a-4a)以及用于为所述自我车辆(1)确定与之适配的自身未来轨迹(1a)的方法(100),所述方法具有以下步骤:

·从所述环境(11)的物理观测的时间序列(11a-11c)中和/或从经由所述车辆(1)的无线接口(12)从所述外部对象(2-4)本身和/或从基础设施(5)接收的信息(12a)中,标识(110)所述外来对象(2-4);

·确定(120)所述外来对象(2-4)中的每一个的运动通向哪个近距离目的地(2b-4b)以及该运动根据哪些基本规则(2c-4c)进行;

·确定(130)所述自我车辆(1)的运动通向哪个近距离目的地(1b)以及该运动根据哪些基本规则(1c)进行;

·为所述自我车辆(1)以及为所述外来对象(2-4)分别设置(140)质量函数 R_{1-4} ,所述质量函数给由所述自我车辆(1)和所述外来对象(2-4)的当前状态构成的总状况 x 和可能的紧接着的运动行动 a_{1-4} 分配针对在当前总状况 x 中该行动 a_{1-4} 对于分别观察的交通参与者(1-4)如何好的度量;

·为所述自我车辆(1)以及为所述外来对象(2-4)分别设置(150)质量度量 Q_{1-4} ,所述质量度量除了值 $R_{1-4}(x, a_{1-4})$ 之外还将状态变化 x' 的概率 $P(x')$ 的分布的期望值 $E(P(x'))$ 分配给所述总状况 x 和所述可能的紧接着的运动行动 a_{1-4} ,其中其余交通参与者(2-4;1、3、4;1-3)以所述概率对紧接着的运动行动 a_{1-4} 作出反应;

·确定(160)所述自我车辆和所述外来对象(2-4)的那些使所述质量度量 Q_{1-4} 最大化的最佳运动策略 π_{1-4} ;

·从所述最佳运动策略 π_{1-4} 中确定(170)所寻求的轨迹(1a-4a)。

2. 根据权利要求1所述的方法(100),其中选择(151)其关于所述运动策略 π_{1-4} 的最佳值由贝尔曼最佳值给出的质量度量 Q_{1-4} 。

3. 根据权利要求1至2中任一项所述的方法(100),其中选择(152)波尔兹曼-吉布斯分布作为状态变化 x' 的概率 $P(x')$ 的分布。

4. 一种用于预报在自我车辆(1)的环境(11)中的外来对象(2-4)的轨迹(2a-4a)并且用于为所述自我车辆(1)确定与之适配的自身未来轨迹(1a)的方法(200),所述方法具有以下步骤:

·从所述环境(11)的物理观测的时间序列(11a-11c)中和/或从经由所述车辆(1)的无线接口(12)从所述外部对象(2-4)本身和/或从基础设施(5)接收的信息(12a)中,标识(210)所述外来对象(2-4);

·确定(220)所述外来对象(2-4)中的每一个的运动通向哪个近距离目的地(2b-4b)以及该运动根据哪些基本规则(2c-4c)进行;

·确定(230)所述自我车辆(1)的运动通向哪个近距离目的地(1b)以及该运动根据哪些基本规则(1c)进行;

·为所述自我车辆(1)以及为所述外来对象(2-4)分别设置(240)特征函数 F_{1-4} 为使得将 F_{1-4} 应用于仍自由的参数的组 θ_{1-4} 提供质量函数 R_{1-4} ,其中所述质量函数 R_{1-4} 给由所述自我车辆(1)和所述外来对象(2-4)的当前状态构成的总状况 x 以及可能的紧接着的运动行动 a_{1-4}

分配针对在当前总状况 x 中行动 a_{1-4} 对于分别观察的交通参与者(1-4)如何好的度量;

·将所述自我车辆(1)和所述外来对象(2-4)的运动策略 π_{1-4} 确定(250)为那些在所述总状况 x 中导致所述自我车辆(1)和所述外来对象(2-4)的运动行动 a_{1-4} 的最大因果熵 $H(a_{1-4}||x)$ 的策略;

·从所述运动策略 π_{1-4} 中确定(260)所寻求的轨迹(1a-4a)。

5. 根据权利要求4所述的方法(200),其中关于所述运动策略 π_{1-4} 的因果熵 $H(a_{1-4}||x)$ 的最大值在以下边界条件下被确定(251),即不仅对于所述自主车辆(1)而且对于所述外来对象(2-4)而言,各自特征函数 F_{1-4} 在所有可能的总状况 x 和所有可能的紧接着的运动行动 a_{1-4} 上的期望值与在迄今为止的轨迹中根据经验观测的特征函数 F_{1-4} 的平均值相等。

6. 根据权利要求1至5中任一项所述的方法(100、200),其中所述最佳运动策略 π_{1-4} 在以下边界条件下被确定(160、250),即所述最佳运动策略在相同的来历 H^i 情况下彼此无关。

7. 根据权利要求1至6中任一项所述的方法(100、200),其中所述最佳运动策略 π_{1-4} 在以下边界条件下被确定(160、250),即所述最佳运动策略分别统计地围绕使各自质量函数 R_{1-4} 最大化的策略分布。

8. 根据权利要求1至7中任一项所述的方法(100、200),其中所述外来对象(2-4)分别在其类型(2d-4d)方面被分类(115、215),并且其中各自质量函数 R_{2-4} 或各自特征函数 F_{2-4} 根据这种类型(2d-4d)被选择(141、241)。

9. 一种用于在自我车辆(1)的环境中在具有运动的外来对象(2-4)的交通状况中控制所述自我车辆(1)的方法(300),所述方法具有以下步骤:

·利用根据权利要求1至8中任一项所述的方法(100、200)来确定(310)所述自我车辆(1)的与所述外来对象(2-4)的行为适配的轨迹(1a);

·适配的轨迹(1a)被传送(320)给所述自我车辆(1)的运动规划器(13);

·通过所述运动规划器(13)确定(330)用于所述自我车辆(1)的驱动系统(14)、转向系统(15)和/或制动系统(16)的操控程序(13a),其中所述操控程序(13a)被构造用于在系统极限的范围内使所述车辆(1)的实际行为尽可能好地与所述轨迹(1a)一致;

·根据所述操控程序(13a)操控(340)所述驱动系统(14)、转向系统(15)和/或制动系统(16)。

10. 一种包含机器可读指令的计算机程序,当在计算机上和/或在控制设备上执行所述机器可读指令时,所述机器可读指令促使所述计算机和/或所述控制设备执行根据权利要求1至9中任一项所述的方法(100、200、300)。

使自我车辆的轨迹与运动的外来对象适配

技术领域

[0001] 本发明涉及尤其是在具有人控制的外来对象的混合交通中用于至少部分自动化的方法的轨迹规划。

背景技术

[0002] 在街道交通中至少部分自动化地运动的车辆不将突然地排挤由人控制的车辆,并且也将不在单独的路线上与人控制的交通隔离。相反地,这些车辆将必须在具有人控制的外来对象的混合交通中安全地运动,其中这些外来对象也例如包括作为较弱的交通参与者的行人。在人控制的外来对象情况下,始终如下存在不安全性:这些外来对象接下来执行哪种运动行动。因此,用于至少部分地自动化地驾驶的控制系统的指示至少部分地从对迄今为止的行为的观测中来推断外来对象的未来行为。

[0003] WO 2017/197 170 A1公开一种用于运动的自主单元的控制单元,所述运动的自主单元可以是机器人或也可以是车辆。控制单元首先确定基本轨迹,利用所述基本轨迹来跟踪自主单元的首要目的地、例如行驶目的地。然后,通过安全模块改变基本轨迹,使得避免与人或其他人控制的单元碰撞。为此,对各自的人控制的运动进行预报。

发明内容

[0004] 在本发明的范围内,开发了一种用于预报在自我车辆的环境中的外来对象的轨迹以及用于为自我车辆确定与之适配的自身未来轨迹的方法。

[0005] 自我车辆是应该对其轨迹产生影响以便避免与外来对象碰撞的车辆。外来对象尤其是可以是人或由人控制的车辆、诸如常规的机动车或自行车。但是,也考虑不可控制的或仅受限制地可控制的外来对象、诸如停放在坡上之后滚落的车辆或从其牵引车挣脱的拖车。

[0006] 首先标识外来对象。为此,可以考虑环境的物理观测的时间序列,诸如摄像机图像系列或基于事件的传感器输出的事件系列。可替代地或也以组合的方式,可以考虑经由车辆的无线接口接收的信息。这些信息可以由外部对象本身例如经由车辆对车辆(V2V)接口来传送。但是,信息也可以由基础设施例如经由车辆对基础设施(V2I)接口传送。

[0007] 就此而论,标识意味着至少检测在自我车辆的环境中哪些外来对象可以彼此无关地运动。就此而论,检测外来对象详细地是什么也是有利的,但是并非是强制性需要的。

[0008] 确定外来对象中的每一个的运动通向哪个近距离目的地,并且该运动根据哪些基本规则进行。如何详细地执行该确定取决于哪些信息是可用的。因此,例如可以单独地从轨迹的时间变化过程外推出特定的近距离目的地比其他的更有可能。使用越多的附加信息,近距离目的地的预报(Vorhersage)就变得越准确。如果例如识别出作为外来对象的车辆已经设置了方向指示灯,则以高概率规划拐弯过程。但是作为外来对象的车辆也可以例如直接通过V2V通信宣布其当前的近距离或甚至远距离目的地。

[0009] 外来对象的运动进行所根据的基本规则尤其是可以包括街道交通秩序的规则,并

且还取决于外来对象的类型。因此,例如车辆必须使用行车道和两条行车道中的右行车道。而行人例如被保持在人行道上行走,并且如果存在如用于横穿行车道的斑马线或交通信号灯之类的人行横道,则也使用这些人行横道。

[0010] 此外确定自我车辆的运动通向哪个近距离目的地,并且该运动根据哪些基本规则进行。基本规则在这里尤其是可以再次包括街道交通秩序的规则,并且不必在所有状况下都相同。因此,当车辆拖曳拖车或利用防滑链行驶时,例如允许的最高速度分别被限制。因此,基本规则的确定例如也可以包括对自我车辆的配置的分析。

[0011] 不仅为自我车辆而且为外来对象分别设置质量函数 R_{1-4} ,所述质量函数给由自我车辆和外来对象的当前状态构成的总状况 x 以及可能的紧接着的运动行动 a_{1-4} 分配针对在当前总状况 x 中行动 a_{1-4} 对于分别所观察的交通参与者如何好的度量。质量函数 R_{1-4} 尤其是可以包括例如针对在状况 x 中运动行动 a_{1-4} 在何种程度上力争到达各自近距离目的地并且力争遵守规则的度量。在此,从1达到直至4的数字索引不应被理解为关于可处置的外来对象的数量是限制性的,而是仅仅是解释性的,以便能够根据示例来阐述该方法。通常,还可以谈及质量函数 R_i 和紧接着的运动行动 a_i 。

[0012] 术语“状态”通常包括以下参量,即利用所述参量可以表征自我车辆或外来对象对交通状况的贡献。状态尤其是可以包括位置或也可以包括时间导数、即例如速度和加速度。

[0013] 不仅为自我车辆而且为外来对象分别建立质量度量 Q_{1-4} ,所述质量度量除了值 $R_{1-4}(x, a_{1-4})$ 之外也将状态变化 x' 的概率 $P(x')$ 的分布的期望值 $E(P(x'))$ 分配给总状况 x 和可能的紧接着的运动行动 a_{1-4} ,其中其余交通参与者以所述概率对该紧接着的运动行动 a_{1-4} 作出反应。例如,质量度量 Q_{1-4} 可以是由质量函数的值 $R_{1-4}(x, a_{1-4})$ 和期望值 $E(P(x'))$ 组成的加权和。

[0014] 确定自我车辆和外来对象的那些使质量度量 Q_{1-4} 最大化的最佳运动策略 π_{1-4} 。从最佳运动策略 π_{1-4} 中确定自我车辆和外来对象的所寻求的轨迹。

[0015] 在此,运动策略的术语通常包括将数值 $\pi_{1-4}(x, a_{1-4})$ 分配给总状况 x 和紧接着的运动行动 a_{1-4} 的每个函数 π_{1-4} 。因此,与使该术语与确定性规则相关联的常见语言惯用法相比,该术语是广义的。确定性规则例如可以说明,在存在特定的总状况 x 时,应该由自我车辆执行或由外来对象完成刚好一个紧接着的运动行动 a_{1-4} 。

[0016] 已经认识到,尤其是外来对象的行为不总是遵循确定性规则。如果外来对象例如由人控制,则该控制虽然是智能的,但是不必然引向刚好对于跟踪各自的近距离目的地是最佳的该运动行动。即使人类驾驶员原则上决定正确的驾驶机动车动作,这也适用。因此例如,从其上为此未明确标记车行道的道路向左拐弯可以围绕理想的行驶路线发散(streuen)。车辆当在红色交通信号灯之前多次制动时虽然每次将大致在停车线处停止,但是速度的时间变化过程可能是不同的。驾驶员可以例如从开始有时较强而有时较弱地踩踏制动踏板,并且稍后下意识地再调节制动压力,以便最终在正确的位置处停止。对此的更深原因在于,驾驶任务总体上太复杂而不被完全有意识地被执行。为了完全能够以所要求的速度胜任多重任务执行(Multitasking),学习型驾驶员必须首先将特定的过程“自动化”到潜意识中。

[0017] 甚至行人的正确行为也不是完全确定性的。当行人例如横穿行车道时,该行人不总是刚好以与行驶方向成直角地做这一点,而是以与之有随机偏差地做这一点。

[0018] 只有当进行控制的人错误地决定时,外来对象的行为才不再遵循确定性规则。从

而例如设置右方向指示灯不是对以下的保证：驾驶员实际上将从优先行驶路段向右转入下一道路并且相对于来自该道路的另一车辆放弃其优先行驶。相反地，也可能发生以下情况，即在驾驶员断定该驾驶员搞错了并且道路稍后才必须向右拐之后，该驾驶员笔直地继续行驶。人类驾驶员例如也不能对隐藏在其镜子的盲区中的对象作出反应。行人也总是再次有意识地越过使用安全人行横道的义务或在红色交通信号灯处的等候义务。

[0019] 此外，甚至自我车辆的行为以一定方式是概率性的。如果例如将特定的制动压力给予到制动设备的制动缸用以停车，则自我车辆的由此引起的减速可能例如根据行车道的状态以及制动液的温度和水含量而变化。

[0020] 通过现在将其余交通参与者的状态变化概括为概率分布 $P(x')$ 的方式以及通过所有交通参与者的运动策略 π_{1-4} 同样可以是概率性的方式，从而可以如此改良自我车辆对总状况 x 的反应，使得所述反应实际上以较高的概率是符合交通规则，并且尤其是避免碰撞。以一定的方式，因此在技术上模仿每个人类驾驶员必须在驾驶学校中学习的预见性驾驶，以便用于至少部分自动化地驾驶的系统可以至少与人类驾驶员一样好地完成驾驶任务。

[0021] 在一种特别有利的构型中，选择质量度量 Q_{1-4} ，其关于运动策略 π_{1-4} 的最佳值由贝尔曼最佳值给出。以一定的方式，这是质量度量 Q_{1-4} 的相互耦合和递归定义的组合。

[0022] 例如，在无限的时间范围情况下，质量度量 Q_i 在完成优化的最终状态 Q^* 下可以具有形式：

$$Q_i^*(\mathbf{x}, a_i) = R_i(x, a_i) + \gamma \cdot E_{\pi^*(-i)}(V_i^*(\mathbf{x}') | \mathbf{x}, a_i) \quad (1)$$

其中 $\pi^*(-i)$ 是其他交通参与者的最佳运动策略，所述其他交通参与者的索引是不同于 i 的索引，以及

$$V_i^*(\mathbf{x}') = \underset{a'}{\operatorname{softmax}} Q_i^*(\mathbf{x}', a').$$

[0023] 期望值 E 在概率性状态过渡以及其他交通参与者的策略上进行，所述其他交通参与者的索引是不同于 i 的索引。所述期望值通过

$$E = \sum_{\mathbf{a}_{-i}, \mathbf{x}} P(\mathbf{x}' | \mathbf{x}, \mathbf{a}) \cdot \pi^*(\mathbf{a}_{-i} | \mathbf{x}) \cdot V_i^*(\mathbf{x})$$

给出。

[0024] 在另一特别有利的构型中，在最佳运动策略 π_{1-4} 在相同的来历 H^t 情况下彼此无关的前提条件下确定所述最佳运动策略 π_{1-4} ：

$$\pi^*(a_{-i}(t) | H^t) = \prod_{j \in -i} \pi_j^*(a_j(t) | H^t) \quad (2).$$

[0025] 如果现在在此外选择波尔兹曼-吉布斯(Boltzmann-Gibbs)分布作为状态变化 x' 的概率 $P(x')$ 的分布，则交通参与者分别根据最大熵原理选择其运动策略：

$$\pi_i^*(a_j(t) | H^t) \propto \exp(Q_i^*(\mathbf{x}(t), a_j(t))) \quad (3).$$

[0026] 方程式(1)至(3)构成成一组 M 个耦合的方程式，其中 M 是所观察的交通参与者的数

量。所述方程可以被联合为

$$Q_i^* = T_i(Q_{-i}^*, Q_i^*), i \in [M] \quad (4).$$

[0027] 在这里 T_i 是方程式(1)的右侧。方程式(4)具有刚好一个最佳解 Q_i^* ,其可以利用以下算法获得:

```

算法      1 - MMCE-I
输入      Ht, {Rj}
1: 初始化   :  $Q_i^0$  und  $Q_{-i}^0$ 
2:  $s=0$ 
3: while Konvergenz do
4:   for  $j \in -i$  do
5:      $Q_j^{s+1} \leftarrow T_i(Q_{-i}^{s+1}, Q_i^s)$ 
6:   end for
7:    $Q_i^{s+1} \leftarrow T_i(Q_{-i}^{s+1}, Q_i^s)$ 
8:    $s \leftarrow s+1$ 
9: end while
输出       $Q^s$ 

```

如果时间范围是有限的,则问题轻微不同。第 i 个交通参与者的质量函数 Q 在完成优化的状态下在时间步 $\tau \in [t, t+T]$ 时具有形式:

$$Q_i^\tau(\mathbf{x}, a_i) := R_i(\mathbf{x}, a_i) + E_{\pi^\tau(-i)}(V_i^{\tau+1}(\mathbf{x}') | \mathbf{x}, a_i) \quad (5),$$

$$V_i^{\tau+1}(\mathbf{x}') = \operatorname{softmax}_{a'} Q_i^{\tau+1}(\mathbf{x}', a')$$

其具有边界条件:

$$V_i^{t+T}(\mathbf{x}) = R_{i,F}(\mathbf{x})$$

是在时间范围结束时处于最终优化状态下质量函数 R_i 的值。类似于有限时间范围的情况,期望值再次取决于其余交通参与者的策略,所述策略又有利地是波尔兹曼分布的:

$$\pi_{-i}^\tau(a_j | \mathbf{x}) \propto \prod_{j \in -i} \exp(Q_j^\tau(\mathbf{x}, a_j)).$$

[0028] 因此,方程式(5)可以被写为:

$$Q_i^\tau = U_i(Q_{-i}^\tau, V_i^{\tau+1}), i \in [M], \quad (6)$$

其中, U_i 是方程式(5)的右侧。例如,可以利用以下算法获得最佳解:

```

    算法    2 - MMCE-F
    输入    {Rj, Rj,F}, T
    1: for j=1, ..., M do
    2:     Vjt+1 ← Rj,F
    3: end for
    4: 初始化 : Q0 ← [Q10, ..., QM0]
    5: for κ=T-1, ..., 0 do
    6:     s=0
    7:     while Konvergenz do
    8:         for i=1, ..., M do
    9:             Qis+1 ← Ui(Qis, Vit+κ+1)
    10:        end for
    11:        s ← s+1
    12:    end while
    13:    Q0 ← Qs
    14:    for j=1, ..., M do
    15:        Vjt+κ ← softmaxaj Uj(Qj0, Vjt+κ+1)
    16:    end for
    17: end for

```

输出 {V^t, ..., V^{t+T}}

在本发明的范围中，开发了用于预报自我车辆的环境中的外来对象的轨迹以及用于为自我车辆确定与之适配的自身未来轨迹的另一方法。该方法首先如先前描述的方法一样开始，即标识外来对象，并且不仅为自我车辆而且为外来对象确定运动的基本规则和近距离目的地。

[0029] 与先前描述的方法不同地，不仅为自我车辆而且为外来对象分别设置特征函数 F_{1-4} 为使得将 F_{1-4} 应用于仍自由的参数的组 θ_{1-4} 提供质量函数 R_{1-4} ，其中该质量函数 R_{1-4} 给由自我车辆和外来对象的当前状态构成的总状况 x 以及可能的紧接着的运动行动 a_{1-4} 分配针对在当前总状况 x 中行动 a_{1-4} 对于分别所观察的交通参与者如何好的度量。质量函数 R_{1-4} 尤其是可以包括例如针对运动行动 a_{1-4} 在状况 x 中在何种程度上力争达到各自的近距离目的地并且力争遵守规则的度量。

[0030] 特征函数 F_{1-4} 可以例如体现各自交通参与者的特性和目的地，例如行人移向的目的地地点也或者所述行人的行进速度。在车辆情况下，除了目的地地点外，例如行驶应该安全地、流畅地和舒适地进行的要求可以进入特征函数 F_{1-4} 中。因此，特征函数 F_{1-4} 尤其是可以例如由涉及不同目的地的多个部分组成，其中这些目的地也可以是互相相反的。于是，参数的组 θ_{1-4} 可以例如体现权重，不同的目的地和要求以所述权重包含在最终质量函数 R_{1-4} 中。参数的组 θ_{1-4} 尤其是可以例如作为参数的向量存在并且例如包含系数，由在特征函数 F_{1-4} 中包含的不同目的地组成的线性组合以所述系统进入质量函数 R_{1-4} 中。

[0031] 自我车辆和外来对象的运动策略 π_{1-4} 被确定为那些在总状况 x 中导致自我车辆和

外来对象的运动行动 a_{1-4} 的最大因果熵 $H(a_{1-4}||\mathbf{x})$ 的策略。从运动策略 π_{1-4} 中确定所寻求的轨迹。

[0032] 最终获得的结果具有与根据先前描述的方法获得的结果相同的优点。特别是该方法的优点是，与对于直接确定质量函数 R_{1-4} 相比，对于确定参数组 θ_{1-4} 需要关于各自交通参与者的更加少的信息。但是在参数组 θ_{1-4} 的另一侧上可以考虑每个附加信息，不管来自何种源。在优化时根据运动策略 π_{1-4} 确定自由的参数 θ_{1-4} 。

[0033] 因果熵 $H(a_{1-4}||\mathbf{x})$ 可以被写为

$$H(a_{1-4} || \mathbf{x}) = -E_{\mathbf{a}, \mathbf{x}} \left[\sum_{t \leq T} \log \pi^t(\mathbf{a}(t) | H^t) \right].$$

[0034] 有利地，在以下边界条件下关于运动策略 π_{1-4} 确定因果熵 $H(a_{1-4}||\mathbf{x})$ 的最大值，即不仅对于自我车辆而且对于外来对象而言，各自特征函数 F_{1-4} 在所有可能的总状况 \mathbf{x} 和所有可能的紧接着的运动行动 a_{1-4} 上的期望值与在迄今为止的轨迹中根据经验观测的特征函数 F_{1-4} 的平均值相等。该平均值尤其是可以在迄今根据经验观测的所有状况 \mathbf{x} 和运动 a_{1-4} 上被构成：

$$\begin{aligned} E_{a_i, \mathbf{x}} [F_i(\mathbf{x}, a_i)] &= \tilde{E}_{a_i, \mathbf{x}} [F_i(\mathbf{x}, a_i)] \forall i, \\ \pi_i^t(a_i(t) | H^t) &\geq 0 \forall i, a_i(t), H^t, \\ \sum_{a_i(t)} \pi_i^t(a_i(t) | H^t) &= 1 \forall i, H^t. \end{aligned} \quad (7)$$

[0035] 结合以下其他边界条件，即在相同的来历 H^t 情况下最佳运动策略 π_{1-4} 彼此无关，并且所述最佳运动策略分别以统计方式围绕使各自质量函数 R_{1-4} 最大化的策略分布，在使用

$$E_{a_i, \mathbf{x}} [F_i(\mathbf{x}, a_i)] = \sum_{t \leq T} E [F_i(\mathbf{x}(t), a_i(t))]$$

的情况下可以说明方程式(7)的递归解：

$$\begin{aligned} \pi_i^\tau &= \frac{1}{Z_i(\tau)} \exp(W_i^\tau(H^\tau, a_i(\tau))), \\ W_i^\tau(H^\tau, a_i(\tau)) &= \theta_i^T F_i(x(\tau), a_i(\tau)) + E_{\pi^\tau(-i)} [\log Z_i(\tau + 1)], \\ \log Z_i(\tau) &= \operatorname{softmax}_{a'} W_i^\tau(H^\tau, a'), \\ \log Z_i(T) &= \operatorname{softmax}_a \theta_i^T F_i(\mathbf{x}(T), a). \end{aligned}$$

[0036] 在这里是

$$\pi_{-i}^t = \prod_{j \in -i} \pi_j^t.$$

[0037] 边界条件对于所有交通参与者是 $Z_i(T+1)=1$ 。

[0038] 递归解类似于根据方程式 (5) 的完成优化的质量度量 Q 。 $W_i^t(H^t, a_i(\tau))$ 起质量度量 Q_i 的作用, 并且质量函数 R_i 作为线性组合由特征函数 F_i 组成。

[0039] 因此, 最终从自我车辆的角度来看, 可以促进“逆向强化学习 (inverse reinforcement learning)”, 即在知道自我车辆的质量函数 R_i 的情况下, 可以单独地从其余交通参与者的观测中推断其质量函数 R_{2-4} 。这例如可以利用以下算法发生:

算法3: 用于自主车辆的MMCE-IRL

```

1: for j ∈ -i do
2:    $\tilde{F}_j^{(0)} \leftarrow \sum_{(x, a_j) \in H^t} F_j(x, a_j)$ 
3:    $R_j^{(0)} \leftarrow F_j^T \theta_j$ 
4: end for
5:  $\pi \leftarrow \text{MMCE}(H^t, \{R_k^{(0)}: k \in -i\}, R_i)$ 
6: for j ∈ -i do
7:    $\bar{F}_j \leftarrow E_\pi[F_j]$ 
8:    $\theta_j \leftarrow \text{Projektion}_{\|\theta_j\| \leq B} (\theta_j - \alpha(\tilde{F}_j(t) - \bar{F}_j))$ 
9: end for

```

在另一特别有利的构型中, 外来对象分别在其类型方面被分类, 并且各自质量函数 R_{2-4} 或各自特征函数 F_{2-4} 根据该类型被选择。以这种方式, 外来对象的最终轨迹以及因此还有自我车辆的与之适配的轨迹的确定可以更快速地收敛 (konvergieren) 并且获得更精确的结果。如先前所阐述的, 运动的基本规则尤其是可以取决于对象的类型。可以根据物理观测和/或根据经由无线接口接收的信息进行分类。

[0040] 如先前所阐述的, 自我车辆的与运动的外来对象的存在适配的轨迹的确定不是目的本身, 而是具有以下目的: 特别是对于具有由人控制的外来对象的混合式交通而言改善至少部分自动化驾驶的车辆的适宜性。因此, 本发明还涉及一种用于在自我车辆的环境中在具有运动的外来对象的交通状况下控制自我车辆的方法。

[0041] 在该方法情况下, 利用先前描述的方法之一确定自我车辆的与外来对象的行为适配的轨迹。适配的轨迹被传送给自我车辆的运动规划器。通过运动规划器确定用于自主车辆的驱动系统、转向系统和/或制动系统的操控程序, 其中所述操控程序被构造用于在系统极限的范围内使车辆的实际行为尽可能好地与所确定的轨迹一致。

[0042] 根据操控程序操控驱动系统、转向系统和/或制动系统。

[0043] 因为由于借助于CAN总线的内部联网典型地存在从车辆中的每个位置对利用传感器系统检测的或经由无线接口获取的关于车辆环境中的外来对象的信息的访问, 所以可以在自我车辆的任意已经存在的控制设备中实现该方法。同样, 可以经由CAN总线从车辆中的每个位置操控运动规划器。该方法可以例如以软件的形式来实现, 所述软件可以作为用于这种控制设备的更新或升级来出售, 并且就此而言是自身的产品。因此, 本发明还涉及一种

具有机器可读指令的计算机程序,当在计算机上和/或在控制设备上执行所述计算机可读指令时,所述计算机可读指令促使计算机和/或控制设备执行由本发明提供的方法。同样,本发明还涉及具有计算机程序的机器可读数据载体或下载产品。

附图说明

[0044] 下面根据附图与对本发明的优选实施例的描述一起更详细地描绘对本发明进行改善的其他措施。

实施例

[0045] 图1示出方法100的实施例;

图2示出方法200的实施例;

图3是示出方法300的实施例;

图4示出具有自我车辆1和三个人控制的外来对象2-4的示例性交通场景。

具体实施方式

[0046] 图1示出方法100的实施例。在步骤110中,与经由无线接口12接收的信息12a一起在图1中未绘入的自我车辆1的环境11的物理观测的时间序列11a-11c进行处理。这些信息12a来自车辆环境11本身中的外来对象2-4和/或来自基础设施5。在步骤110中,标识外来对象2-4,即断定存在三个外来对象2-4,所述外来对象2-4以不同的方式运动。

[0047] 在可选的步骤115中,根据类型2d-4d对外来对象2-4进行分类。在步骤120中,预测由外来对象2-4分别力图的近距离目的地2b-4b,并且确定基本规则2c-4c,其中外来对象2-4的运动根据所述基本规则进行。与此类似地,在步骤130中确定自我车辆1的运动通向哪个近距离目的地1b,并且该运动根据哪些基本规则1c进行。

[0048] 在步骤140中,基于存在的信息为自我车辆1以及为外来对象2-4设置各自的质量函数 R_{1-4} ,其中当外来对象2-4的各自类型2d-4d在可选的步骤115中已被确定时,根据可选的子步骤141可以考虑所述类型。

[0049] 在步骤150中,将质量函数 R_{1-4} 扩展成质量度量 Q_{1-4} ,所述质量度量 Q_{1-4} 附加地还包含状态变化 x' 的概率 $P(x')$ 的分布的期望值 $E(P(x'))$,并且就此而言质量度量 Q_{1-4} 也彼此耦合。在此,根据子步骤151,选择质量度量 Q_{1-4} ,所述质量度量关于运动策略 π_{1-4} 的最佳值由贝尔曼最佳值(Bellman-Optimum)给出。根据子步骤152,选择波尔兹曼-吉布斯分布作为状态变化 x' 的概率 $P(x')$ 的分布。

[0050] 在步骤160中,确定自我车辆和外来对象2-4的那些使质量度量 Q_{1-4} 最大化的运动策略 π_{1-4} 。由此,最后在步骤170中,从中确定外来对象2-4的所寻求的轨迹2a-4a以及自我车辆1的与之适配的额定轨迹1a。

[0051] 图2示出方法200的实施例。步骤210、215、220和230与方法100的步骤110、115、120和130相同。

[0052] 在方法200的步骤240中,与方法100的步骤140不同地,不确定完整的质量函数 R_{1-4} ,而是确定特征函数 F_{1-4} ,所述特征函数利用仍自由的参数的组 θ_{1-4} 被参数化并且只有与这些参数 θ_{1-4} 相结合才构成完整的质量函数 R_{1-4} 。只要在步骤215中确定了外来对象2-4的类

型2d-4d,就可以在可选的子步骤241中考虑所述类型用以选择各自特征函数 F_{2-4} 。

[0053] 在步骤250中,将自我车辆和外来对象的运动策略 π_{1-4} 确定为那些使最大因果熵最大化的策略。同时,还确定特征函数 F_{1-4} 的参数 θ_{1-4} 。在此,根据子步骤251,预先给定边界条件,所述边界条件使得能够以递归方式确定运动策略 π_{1-4} 。

[0054] 在步骤260中,类似于方法100的步骤170,从运动策略 π_{1-4} 中确定外来对象2-4的所寻求的轨迹2a-4a以及自我车辆1的与之适配的额定轨迹1a。

[0055] 图3示出方法300的实施例。在步骤310中,利用方法100或200为自我车辆1确定与自我车辆1的环境11中的外来对象2-4的行为适配的额定轨迹1a。在步骤320中,将该适配的轨迹1a传送给自我车辆1的运动规划器13。在步骤330中,通过运动规划器13确定用于自我车辆1的驱动系统24、转向系统15和/或制动系统16的操控程序13a。

[0056] 就此而论重要的是,轨迹的术语通常涉及组合式空间坐标和时间坐标中的轨道(Bahn)。这意味着不仅可以通过改变运动方向而且可以通过改变速度、诸如刹车、等待并且稍后重新开动来改变轨迹。

[0057] 在步骤340中,根据操控程序13a来操控驱动系统14、转向系统15或制动系统16。

[0058] 图4示出复杂的交通场景,其中可以有利地使用所描述的方法100、200、300。自我车辆1在道路50的右通车车道上笔直地朝近距目的地1b的方向行驶。

[0059] 第一外来对象2是另一车辆,该另一车辆的方向指示灯2e表明其驾驶员打算拐向通往车辆2的近距目的地2b的横街51。第二外来对象3是另一车辆,该另一车辆从自我车辆1的角度看在道路50的相反行车道上处于笔直地朝其近距目的地3b的方向的路上。第三外来对象4是行人,所述行人在道路50的从其看来相对的侧上向近距目的地4b行进。

[0060] 在图4中所示的状况下,行人4必须使用道路50上的人行横道52,这同时责成车辆3的驾驶员等待。因此,原则上,车辆2的驾驶员可以直接加速并且如打算的那样向左拐,这对于其快速到达近距目的地2b会是最佳的。与此相应地,自我车辆1在其车道中具有至少直至人行横道52的自由行驶。在车辆2的驾驶员将做对其而言最佳事情的简化假设下的控制方法因此将会使自我车辆1加速。但是,如果车辆2的驾驶员现在如下错误地估计该状况,即该驾驶员必须首先让迎面车流中的车辆3驶过(这当在人行横道52上无行人4的情况下确实也会是正确的),则自我车辆从后面撞在车辆2上。根据本发明的方法使得能够考虑这种不安全性。从而例如可以在如此程度上限制用于继续行驶的速度,使得对于车辆2实际上停住的情况,仍然可以利用全制动来防止相撞。

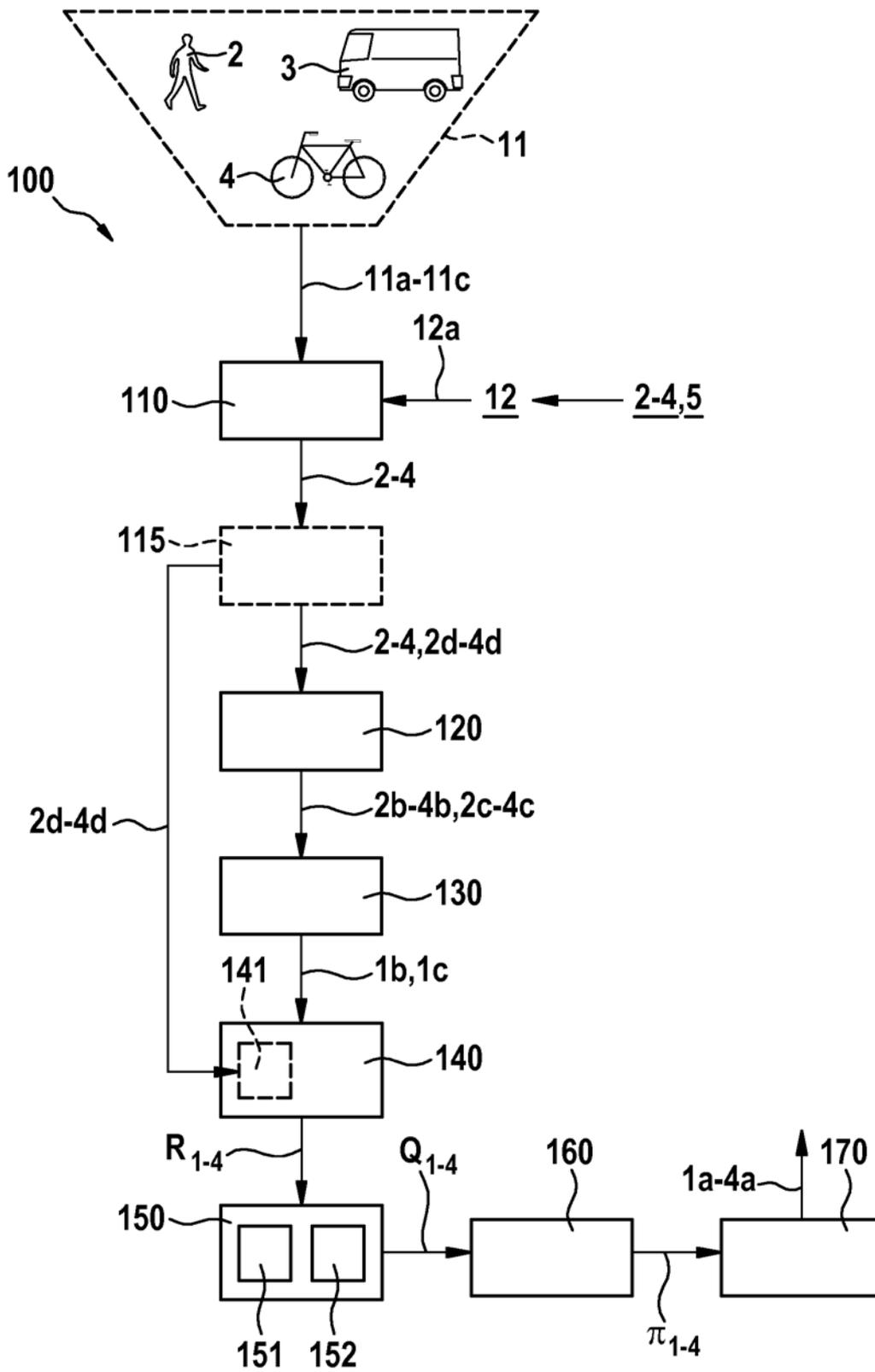


图 1

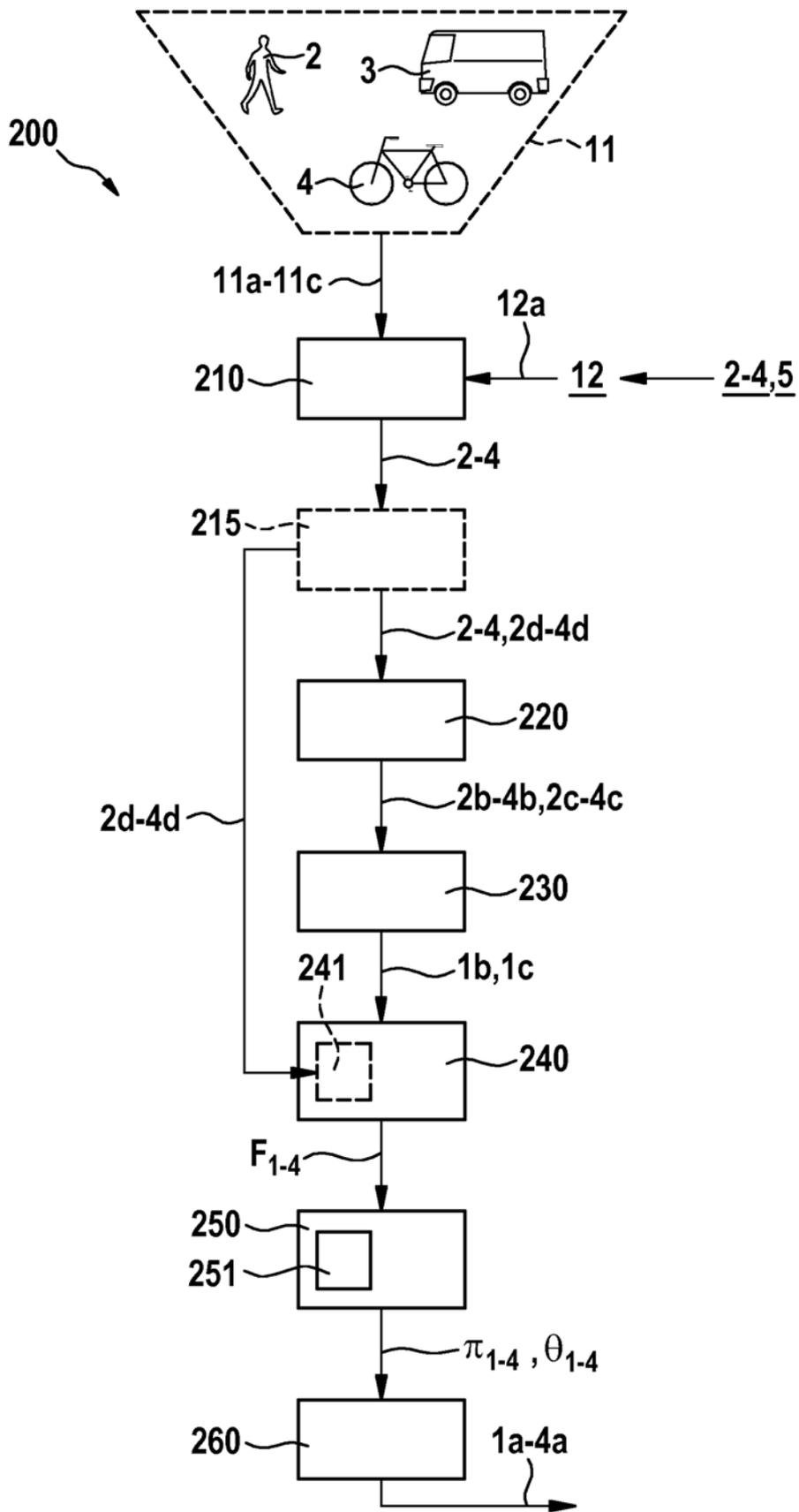


图 2

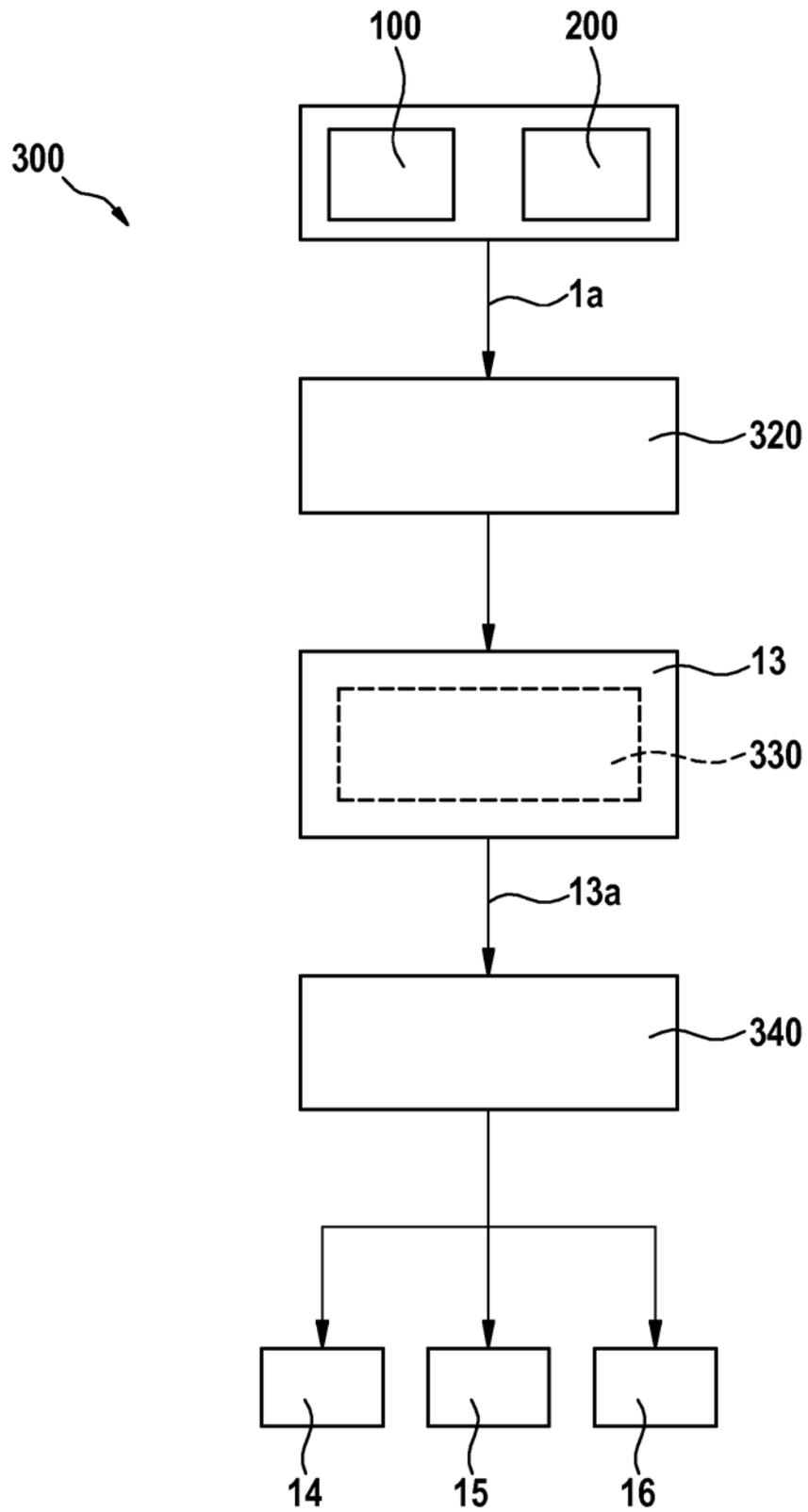


图 3

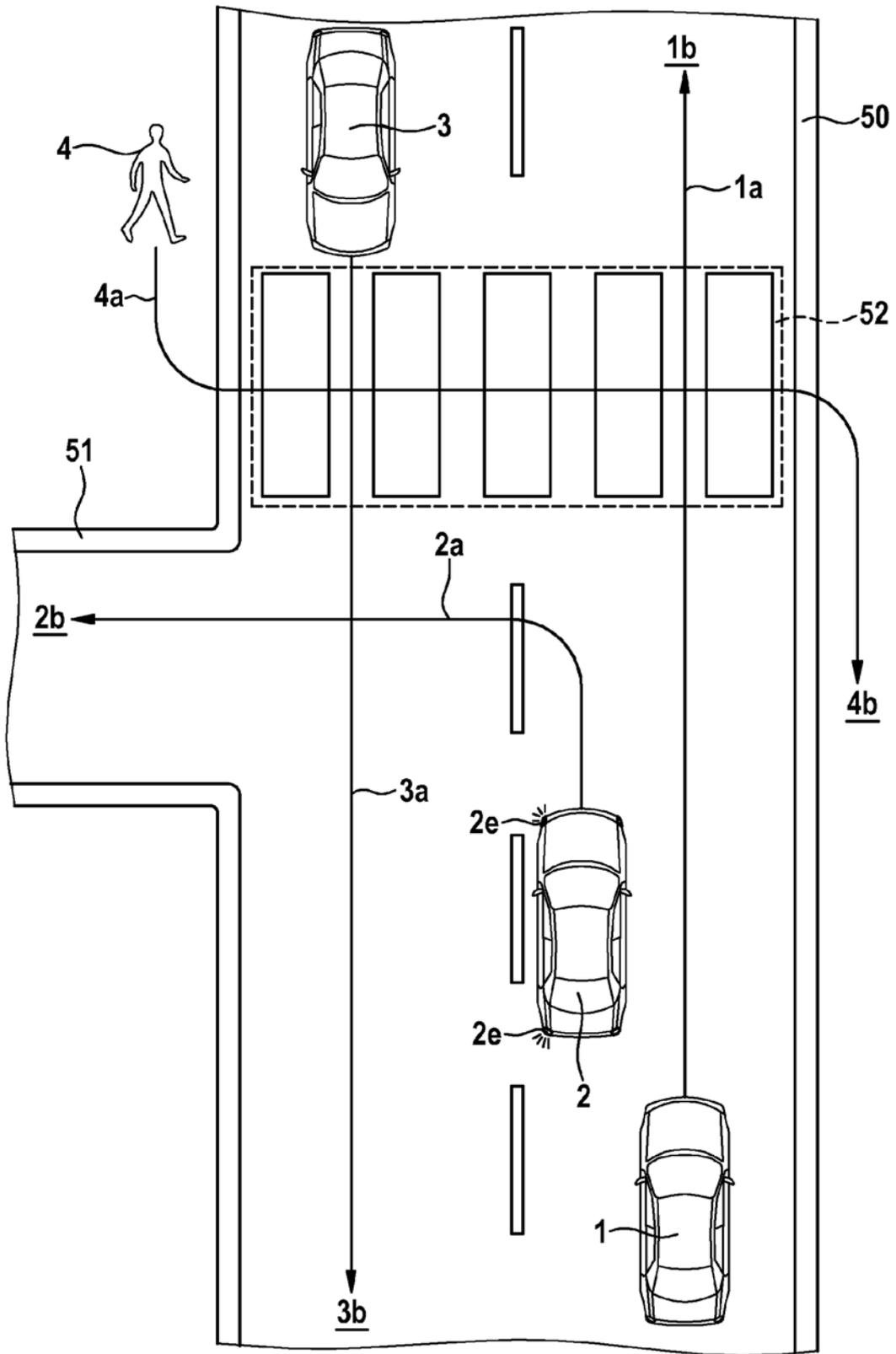


图 4