



(12) 发明专利申请

(10) 申请公布号 CN 117141450 A

(43) 申请公布日 2023. 12. 01

(21) 申请号 202310626296.7

(22) 申请日 2023.05.30

(30) 优先权数据

A50386/2022 2022.05.31 AT

(71) 申请人 AVL李斯特有限公司

地址 奥地利格拉茨

(72) 发明人 A·穆萨

(74) 专利代理机构 北京同达信恒知识产权代理

有限公司 11291

专利代理师 黄志华 何月华

(51) Int. Cl.

B60W 20/00 (2016.01)

B60W 20/13 (2016.01)

B60W 30/182 (2020.01)

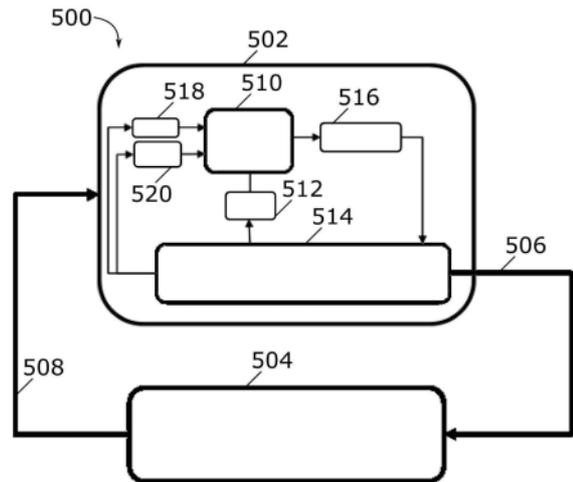
权利要求书1页 说明书10页 附图5页

(54) 发明名称

用于控制混合动力车的驱动系的装置和方法

(57) 摘要

本发明涉及一种用于在行驶路线上控制混合动力车的驱动系(100,504)以达到控制目标的控制装置(110,502),包括:基于规则的选择器(514),能够由基于规则的选择器(514)依据车辆状态来配置的模式使能器(512),以及与该模式使能器(512)处于信号技术连接中的智能体(280,402,510),其中,该智能体(280,402,510)具有基于增强学习的配置,所述配置基于在混合动力车运行期间或模拟期间所采集的数据,并且该智能体(280,402,510)设立用于依据该模式使能器(512)的配置从至少一个不可变的工作模式和/或至少一个可变的工作模式中选择一个工作模式以实行该驱动系(100,504)的调设。



1. 一种用于在行驶路线上控制混合动力车的驱动系(100,504)以达到控制目标的控制装置(110,502),包括:

- 基于规则的选择器(514),
- 能够由该基于规则的选择器(514)依据车辆状态来配置的模式使能器(512),以及
- 与该模式使能器(512)处于信号技术连接中的智能体(280,402,510),

其中,该智能体(280,402,510)具有基于增强学习的配置,所述配置基于在该混合动力车运行期间或模拟期间所采集的数据,并且

其中,该智能体(280,402,510)设立用于依据该模式使能器(512)的配置从不可变的工作模式和/或至少一个可变的工作模式中选择一个工作模式以实行该驱动系(100,504)的调设。

2. 根据权利要求1所述的控制装置(110,502),其中,该不可变的工作模式是附加增压模式或能量回收模式或停止模式。

3. 根据权利要求1或2所述的控制装置(110,502),其中,该至少一个可变工作模式包括燃烧模式和/或电驱动模式和/或充电优化模式。

4. 根据权利要求1或2所述的控制装置(110,502),其中,该智能体(280,402,510)还设立用于基于该模式使能器(512)的配置从紧急运行模式中选择一个工作模式。

5. 根据权利要求1或2所述的控制装置(110,502),其中,该驱动系(100,504)具有P2架构,在该P2架构中,电机(120)安置在传动装置输入端。

6. 根据权利要求1或2所述的控制装置(110,502),其中,该智能体(280,402,510)包括人工智能的网络结构。

7. 根据权利要求6所述的控制装置(110,502),其中,该智能体(280,402,510)包括是深度Q网络的网络结构。

8. 根据权利要求1或2所述的控制装置(110,502),其中,该智能体(280,402,510)包括应用区域和独立于该应用区域的训练区域。

9. 根据权利要求1或2所述的控制装置(110,502),其中,该控制装置(110,502)设立用于按照周期性前后相继的时步分别重新选择该工作模式。

10. 一种用于在行驶路线上控制混合动力车的驱动系(100,504)以达到控制目标的方法,包括以下步骤:

- 依据车辆状态使用基于规则的选择器来配置模式使能器;
- 将该模式使能器的配置传输给智能体(280,402,510),其中,该智能体(280,402,510)具有基于增强学习的配置,该配置基于在该混合动力车的操作期间或模拟期间所采集的数据;并且

-基于该模式使能器(512)的配置从不变的工作模式和/或至少一个可变的工作模式中使用该智能体(280,402,510)选择一个工作模式以实行该驱动系(100,504)的调设。

11. 根据权利要求10所述的方法,还包括以下步骤:使用在该混合动力车工作或模拟期间所采集的关于增强学习的数据来设定该智能体(280,402,510)。

用于控制混合动力车的驱动系的装置和方法

技术领域

[0001] 本发明涉及一种用于在行驶路线上控制混合动力车的驱动系以达成控制目标的控制装置和一种用于在行驶路线上控制混合动力车的驱动系的方法。该控制装置基于智能功能架构,其将基于规则的选择过程与用增强学习作出的选择过程相互组合。

背景技术

[0002] 用于控制混合动力车的驱动系的控制装置的主要任务在于找到用于驱动系组成部件、即内燃机、电动机、电池、变速器、离合器等工作点,其符合期望的控制目标且同时满足关于驱动系部件的安全性、故障诊断和保护的一定设定条件。控制目标是参数(例如像油耗、系统的总效率、排放或驱动系部件的使用寿命的)优化目标。

[0003] 用于控制混合动力车的驱动系的控制装置扮演协调角色,做法是其观察司机要求、行车状况和组成部件状态,确定驱动系组成部件的工作点,并且将相应要求转送至驱动系组成部件。

[0004] 一般,人们能将混合动力车的工作范围分为多个彼此间具有短暂动态过渡阶段的准静态混合模式。混合模式的已知例子是纯电行驶、能量回收、附加增压和充电模式。这些工作方式的特点是组成部件工作点的特定设定条件。例如在电驱动模式下的纯电行驶中,牵引仅由一台或多台电机负担,而内燃机与牵引无关。在附加增压模式下,内燃机和电机都以正扭矩来促成牵引。在充电模式下,发动机扭矩超过期望的牵引力矩,并且这个或这些电机施加负扭矩,由此电池被充电。其中几个模式通过预定规则来确定。例如在纯电行驶下电机扭矩等于期望牵引力矩。内燃机在纯电行驶下被停止并且与驱动装置分离,混合驱动系的其它组成部件受到优化。

[0005] 很常见地,混合工作方式在混合动力车驱动系控制中未被明确考虑,并且其优化直接针对在发动机和电动机之间分配扭矩来执行,而没有参照工作方式。但“将工作范围明确划分为工作模式”有以下将解释的几个优点。

[0006] 在这种做法中,优化任务可被分为两个层级的部分任务:在工作模式内的工作优化和工作模式的选择。待达成的最佳条件通常不仅涉及单独的控制目标,也涉及大多可能相互矛盾的各不同控制目标的加权组合。控制目标例如可以是或包括:系统效率最大化,排放和NVH(噪声-振动-舒适性)最小化,驱动系的单独或全部组成部件的耐用性提高和行驶舒适性提高。目标的权重可以在各不同行驶状况下改变。鉴于此,显然工作模式的评估和选择也并非微不足道的任务,尤其当工作范围被分为多个工作模式、预测信息可供使用和/或应该考虑附加优化目标时。

[0007] 此外,用于控制驱动系的控制装置经受下述调整,其受到社会、政治和技术的发展的影响。一方面,降低真实行驶中的能耗和排放的必要性意味着车辆工作的线上(即在行驶期间内)优化。另一方面,近年来出现许多新的车辆功能,例如像自主行驶、预瞻性功能、队列管理或车辆个体化。这两种趋势也显著影响到混合动力车的控制:新的附加因素和复杂的目标组合明显影响优化判断。

[0008] 优化问题因以下理由而无法再通过传统的利用线下校准规则的基于规则的控制来解决:校准规则的做法一般基于校准参数的线下优化,其针对少量几个行驶周期进行。真实行驶中的优化必须覆盖数量明显更大的情况,其大部分无法通过仅短暂的实验室条件下的试运行被覆盖。

[0009] 所需校准参数的数量及其相互影响可能性随着混合驱动系组件的数量和优化目标数量的增加而明显提升,故通常在基于规则的控制中所存在的控制选择透明度和合理性优点已经基本上消失。

[0010] 另外,校准用支出和成本也随着校准参数数量增加而显著提高。

[0011] 其中一个可能对克服该问题有帮助的极有潜力的技术是人工智能(KI)、尤其是增强学习(Reinforcement Learning,RL)。基于RL算法的混合控制相关主题近几年来被研发人员视为用于其中几个已有问题的潜在解决方案。在关于该主题的大多数公开文献中,仅对完全由RL替代的能量管理优化部分予以考虑。但此做法有如下缺点。

[0012] 首先难以实现RL算法以用于实时的车辆线上控制,因为其应用是计算成本极其高的。此外,系统限制条件、安全方面、故障诊断装置和组成部件保护主题在混合动力车驱动系用控制中占主导地位并且未被RL控制装置纳入考虑,这将RL做法限制到原型车和演示车辆。

发明内容

[0013] 本发明所基于的任务是克服上述缺点。

[0014] 该任务通过根据权利要求1的装置和根据权利要求9的方法来完成。

[0015] 根据第一方面,本发明提供一种用于在行驶路线上控制混合动力车的驱动系(100,504)以达到控制目标的控制装置,其包括:基于规则的选择器,由基于规则的选择器依据车辆状态可配置的模式使能器,以及与该模式使能器处于信号技术连接中的智能体,其中,该智能体具有基于增强学习的配置,所述配置基于在运行期间或在混合动力车模拟期间所采集的数据,并且智能体设立用于依据该模式使能器的配置从不可变的工作模式和/或至少一个可变的工作模式中选择工作模式以实行该驱动系的调设。

[0016] 模式使能器可根据预先选择的工作模式被配置。根据该预先选择,可以从多个原则上可用的不变和可变的工作模式中例如仅选择一个不变的工作模式或多个可变的工作模式或“一个不变的工作模式和一个可变的工作模式”。

[0017] 该控制装置用于在当前所行驶的行驶路线上控制混合动力车的驱动系。作为本发明意义上的混合动力车是指具有使用两个不同能量源的驱动系的车辆。混合动力车的驱动系尤其可以是插电混合动力车的驱动系并且包括内燃机和电机。或者,混合动力车也可以是具有如下驱动系的车辆,其利用燃料电池和蓄电池作为能量源。此外,混合动力车也可以是具有如下驱动系的车辆,其以燃料电池和内燃机为能量源。控制装置原则上用于在混合动力车行驶期间的控制。行驶路线可以是预先编程的行驶路线和/或由驱动系组成部件预测在车辆前方的行驶路线。尤其也可以仅将一个路段视为行驶路线。车辆状态可以除了车辆工作参数外还包括关于行驶路线的信息。不变的工作模式无法由控制装置改变并因此以始终相同的方式对相同输入作出反应。可变的工作模式通过增强学习可由智能体调整,因此原则上取决于在混合动力车工作期间或在混合动力车模拟期间所采集的数据。该数据例

如可以包含关于环境的环境数据、关于车辆的车辆数据、描述车辆工作的工作数据或还有关于驾驶车辆的人员的状态的司机数据。

[0018] 当不变的工作模式是附加增压模式或能量回收模式或停止模式时获得优点。

[0019] 附加增压模式、能量回收模式和停止模式原则上可以在任何选择过程中可供使用,但在这里,由基于规则的选择器分别仅将其中一个不变的工作模式和或许一个或多个可变的工作模式转交给智能体以供选择。当所需扭矩在一定转速下高于内燃机最大扭矩时启用附加增压模式。内燃机最大扭矩被定义为PHEV系统内的组成部件限制,其无法通过HCU判断被失效。当所需扭矩为负时,总是选择能量回收模式。利用P2-PHEV中的校准参数,已经确定再生用SoC范围为94%到21%。停止/静止被视为固定模式,因为它在所需扭矩和速度两者都等于零时被启动。

[0020] 当至少一个可自由选择的工作模式包括燃烧模式和/或电驱动模式和/或充电优化模式时将获得其它优点。

[0021] 优选可以规定智能体还设立用于基于模式使能器(512)的配置从紧急运行模式中选择一个工作模式。

[0022] 当电池充电量近似达到最小阈值时,为了保护电池使用寿命而选择紧急运行模式。最小阈值可以为20%充电量。内燃机在紧急运行模式中一般不在最佳负荷点工作,而是在如下负荷点,此时存在油耗最小值、尤其是油耗全域最小值。为了找到油耗全域最小值,智能体应该在可变的工作模式可用的时段中找到最佳模式选择。

[0023] 当紧急运行模式包含至少一个怠速模式和/或最小耗量模式时获得其它优点。

[0024] 尤其可以规定驱动系具有P2架构,此时电机安置在变速器输入端。

[0025] 在也称为P2插电混合架构的P2架构中,电机安置在变速器的输入端,这允许内燃机通过离合器与电机分离。由此减小牵引力矩损失并由此通过选择并控制工作模式来实现更大的增效。

[0026] 当驱动系具有P2架构、P3架构或P4架构时也将获得优点。

[0027] 在P3架构下电机安置在变速器输出端。布置在变速器后的电机直接作用于差速器。P3架构在能量回收时提供很高的效率,因为其与发动机和变速器的全部机械和液压损失无关。P4架构意味着设有电动的次级轴例如后轴。P4架构中的电驱动轴的功率可简单扩缩。P4架构的空间需求几乎与全轮驱动装置的分动变速器一样高。本发明在采用P2、P3和P4架构中任一个时提供特殊优点。

[0028] 尤其优选地可以规定,智能体包括人工智能的、尤其是深度Q网络的网络结构。

[0029] 人工智能的且尤其是深度Q网络的网络结构是人工神经元网,其基于多个相互连接的被称为人工神经元且模仿生物大脑神经元的单元和/或节点的集合。每处连接可以如生物大脑中的突触那样传输信号至其它神经元。人工神经元接收信号,处理它并向与之相连的神经元发出信号。在一个连接处的“信号”是实数,且每个神经元的输出通过其输入之和的非线性函数来计算。该连接被称为边缘。神经元和边缘一般具有在学习过程进展中得到调整的权重。

[0030] 当智能体包括应用区域和独立于应用区域的训练区域时将获得其它优点。

[0031] 应用区域设立用于在车辆运行期间实行驱动系的配置和/或设定。训练区域用于借助于增强学习来调整智能体配置。

- [0032] 更优选规定了该控制装置设立用于在前后接续的时步中分别重新选择工作模式。
- [0033] 时步尤其也可以周期性前后接续。
- [0034] 根据第二方面,本发明提供用于在行驶路线上控制混合动力车的驱动系以达成控制目标的方法,包括以下步骤:依据车辆状态用基于规则的选择器配置模式使能器;将模式使能器的配置传输给智能体,其中,该智能体具有基于增强学习的配置,其基于在混合动力车的操作期间或模拟期间所采集的数据;并且基于模式使能器的配置从不变的工作模式和/或至少一个可变的工作模式中用智能体选择工作模式以实行驱动系的调设。
- [0035] 在根据本发明第二方面的方法中优选规定,该方法还包括以下步骤:用在混合动力车工作或模拟期间所采集的关于增强学习的数据设定智能体。
- [0036] 优选可以规定该方法由根据权利要求1至8之一的控制装置执行。

附图说明

- [0037] 以下将结合如图所示的实施例来详述本发明,在此,图示意性示出:
- [0038] 图1示出具有P2插电混合配置的混合动力车驱动系;
- [0039] 图2示出混合驱动系的单纯基于规则的控制装置的示例性结构;
- [0040] 图3示出根据本发明的一个特殊实施方式的混合驱动系控制装置的示例性结构;
- [0041] 图4示出传统的RL架构;
- [0042] 图5示出混合RL架构;
- [0043] 图6示出基于增强学习的智能体判断系统的结构;
- [0044] 图7示出基于增强学习的具有模式使能器的智能体的判断系统的结构。

具体实施方式

- [0045] 图1示出具有也被称为P2配置的P2插电混合配置的混合动力车的驱动系100。驱动系100包括控制装置110、内燃机112、变速器114和由电池116经由变换器118供电的电机120。内燃机112和电机120可接入到变速器114中。不同于驱动系的其它配置,在P2配置中在内燃机112与电机120之间设有离合器122,借此能将内燃机112与电机120和变速器114脱开。在此配置中的分离允许高效的纯电驱动而没有由内燃机112造成的牵引力矩损失。
- [0046] 内燃机112、变速器114、电池116、变换器118、电机120和离合器是驱动系100的组成部件,其分别具有一个或多个工作点。除了这些组成部件,驱动系100可以具有出于概览考虑未被明确示出的其它组成部件。组成部件的工作点原则上取决于各自所选的工作模式。
- [0047] 例如在燃烧模式中内燃机112在工作点以正常转速工作。电机120就像电池116一样被关断。离合器122被闭合。当所需扭矩小于可由内燃机112提供的最大扭矩时,能够选择燃烧模式。
- [0048] 而在电驱动模式中关断内燃机112。电机120在驱动工作点工作,电池116放电,而离合器122打开。当所需扭矩小于可由电机120提供的最大扭矩时,能够选择该工作模式。
- [0049] 在充电优化模式中,所需扭矩小于有效负荷点。与此相应,内燃机112的扭矩被提升到更有效的较高水平,且附加扭矩被用于通过电机120的发电机模式来给电池116充电。
- [0050] 控制装置110设立用于接收关于驱动系100的各个组成部件的状态的信息并发送

控制信号至驱动系100的各个组成部件。在图1中出于概览考虑仅示出控制装置110与内燃机112的这种连接。但原则上规定,控制装置110接收关于驱动系100的多个或全部组成部件的状态的信息并且将控制信号发送至驱动系100的多个或全部组成部件。

[0051] 控制装置110的主要任务在于找到用于混合驱动系100的组成部件的工作点,其符合期望的优化目标。

[0052] 控制装置110观察车辆司机的要求、行驶状态和组成部件的状态,确定驱动系组成部件的工作点并且发送相应要求至驱动系组成部件。

[0053] 通常,人们可以将混合动力车的由组成部件的工作点所限定的工作范围分为彼此之间具有短暂动态过渡阶段的若干准静态工作模式。其中几个工作模式通过预定的规则来确定。例如在纯电行驶时,电机120的扭矩等于期望的牵引力矩。内燃机112被停止且与变速器114分离。其它工作模式可以关于其工作点而受到优化并且被称为可变工作模式。

[0054] 如果工作空间被分为若干工作模式,则一个优化任务可以被分为两个部分任务。第一部分任务是优化在各工作模式内的工作。第二部分任务是选择工作模式。待达到的最佳条件在此尤其可以是各不同目标的加权组合。目标在此可以彼此矛盾。示例性优化目标可以是:系统效率最大化,排放和NVH最小化,组成部件耐用性提高和舒适性提高。目标的权重可以在各不同的行驶状况下改变。鉴于此,混合模式的评估和选择也并非微不足道的任务,尤其就将工作范围分解为多个工作模式、提供预测信息和/或将附加优化维度纳入考虑而言。例如变速器挡的选择由控制装置负责。

[0055] 图2示出混合驱动系100的控制装置的示例性结构200,其中未采用增强学习,而是利用仅基于规则的判断路径来代替。

[0056] 在观察者子系统210中,观察装置212、214确定混合驱动系100的不同部件的数据。观察者子系统210包括用于对传感器和控制单元的为控制混合驱动系100所需的信息进行处理的软件部分,并且在其输出端予以提供。

[0057] 第一观察装置212确定并限定混合驱动系100的可供使用的扭矩和功率能力并且在其输出端提供所述值。混合驱动系100、内燃机112和电机120的扭矩极限和功率极限的限定依据车辆部件和驱动系部件的当前状态来发生。

[0058] 第二观察装置214的任务是检测并确定混合驱动系100的驱动系统和组成部件的发动机停止许可和起动要求。关于发动机停止许可和起动要求的前提条件包括在何种条件下能够或应该起动和停止混合驱动系100的部分、尤其是电机和内燃机。

[0059] 在扭矩确定单元220中确定扭矩需求。为此,用油门踏板确定单元222通过油门踏板位置来确定司机所需求的扭矩。在刹车踏板确定单元224中查明刹车踏板位置。在驱动力矩确定单元226中,关于油门踏板位置和刹车踏板位置的信息被汇总并且确定所需的扭矩要求。

[0060] 在驱动力矩确定单元226中所确定的扭矩需求和在观察装置212、214中所确定的瞬间提供的扭矩和混合驱动系100的启停条件被汇总在工作模式计算单元230中且被用来计算各不同工作模式的效率和/或其它评估标准。各不同工作模式包括电工作模式232(电动模式)、能量回收工作模式234、增压工作模式236和其它工作模式238。

[0061] 由工作模式计算单元230所评估的不同工作模式的选择与混合驱动系100的几个主要组成部件的设定以及可在各自工作模式中所选择的关于所需扭矩 T_{req} 的配属关系一起

在表1中被列出。

工作模式	所需扭矩	内燃机转速			电机		电池		离合器	
		优化	标准	无	马达	发电机	充电	放电	闭	开
燃烧模式	$T_{req} < T_{BRMmax}$	-	×	-	-	-	-	-	×	-
附加增压	$T_{req} >>$	-	×	-	×	-	-	×	×	-
备用增压	$T_{req} > T_{BM}/T_{EM}$	×	-	-	×	-	-	×	×	-
[0062] 充电优化	$T_{req} < T_{BRM}$	-	×	-	-	×	×	-	×	-
最小负载	$T_{req} < T_{BRM}$	-	×	-	-	×	-	-	×	-
怠速模式	$T_{req} = 0$	×	-	-	-	×	×	-	-	×
电动模式	$T_{req} < T_{EM}$	-	-	×	×	-	-	×	-	×
能量回收	$T_{req} < 0$	-	×		-	×	×	-	×	
静止模式	$T_{req} = 0$	-	-	×	-	-	-	×	-	×

[0063] 表1:在不同工作模式下的驱动系组成部件的设定

[0064] 所计算的评估标准被共同传输给工作模式评估单元240。在工作模式评估单元240中,评定单元242完成各不同工作模式的评估,输出评估结果至选择单元244。选择单元244才就在当前时刻选择哪个工作模式而作出选择并将其传输给控制单元250。

[0065] 在控制单元250中动态控制该驱动系。该控制由限制单元252、序列单元254和协调单元256执行并依据所选工作模式发生。在限制单元252中,扭矩和其它限制条件被检查和设定。限制区域包含其来自选择单元244、序列单元254和协调单元256的输入。序列单元254实现电机120和内燃机112的启动序列和停止序列。在此要确定电机120和/或内燃机112是否和何时被启动或停止。序列单元254从选择单元244和协调单元256获得其输入。协调单元256执行序列单元254和限制单元252的动态协调。

[0066] 由限制单元252检查的且协调的值尤其与扭矩过滤相关地才被相转送至混合驱动系100的驱动系控制器260。

[0067] 图3示出根据本发明的一个特殊实施方式的混合驱动系100的控制单元110的结构300。

[0068] 结构300相比于结构200有多处被改变。改变尤其涉及附加使用基于增强学习(Reinforcement Learning, RL)的系统,明确来说是环境数据收集器270和通过增强学习可配置的智能体280的附加存在的系统。

[0069] 环境数据收集器270包括准备装置272和奖励计算装置274。处理装置272设立用于将由观察者子系统210采集并转发给环境数据收集装置270的信息处理成智能体280所需的形式。该形式通常被称为观察(observation)。奖励计算装置274设立用于计算由所用策略的特征限定的奖励和/或惩罚。奖励代表待优化的控制目的,其可能包含油耗和/或能耗、排放、司机舒适性评估、NVH、车辆行驶特性等。如果未达所述目标,则也按这些标准计算惩罚。计算的奖励和/或惩罚被传输给智能体280。

[0070] 智能体280可被分为应用区域282和训练区域283。

[0071] 应用区域282包括基于增强学习的策略器284和模式使能器285。策略器284提供智能体280关于环境和在工作中所执行的策略已获得的当前知识。模式使能器285是过滤器，其从可用工作模式范围排除了不可选的工作模式。在本发明的一个特殊实施方式中，模式使能器285是长度对应于可用工作模式数量的向量。模式使能器285的向量除了状态外还被用作策略器284的输入，以在每个时步中对所选工作模式作出判断。

[0072] 训练区域包括经验存储器286和学习算法288。经验存储器286收集并存储在车辆与环境互动时的智能体280的经验。经验存储器286设立用于存储关于环境、驱动系100状态、由控制装置110所选的工作模式和奖励的信息。

[0073] 学习算法288负责收集一定数量的经验实例，从中根据预定的学习算法学习并用新的“更富经验”的版本更新RL策略。

[0074] 所提出的架构的一个优点是智能体280具有两个具有不同功能性的独立区域，即，一个应用区域282和一个训练区域283。应用区域参与预训练策略的执行，以便对行为作出判断。这种做法通常被称为线下RL。而训练区域283依据收集的经验线上训练当前策略，且也被称为线上RL。这意味着训练功能性能完全与工作功能性分开地在单独的计算单元或在云中执行。

[0075] “将系统限制条件、安全性、故障诊断和保护纳入算法中”通过行动屏蔽技术来获得。由系统功能所规定的边界条件被定义为“模式使能器”向量，其由智能体280在每个时步中选择合适行动时予以考虑。通过这种方式，由智能体280只选择如下工作模式，其从系统角度看是允许的，由此避免冲突。由智能体280选择的工作模式被回传给选择单元244，其以与由传统选择过程选定的工作模式相同的方式处理该信息。

[0076] 传统的RL架构400在图4中被示出，在此，智能体402根据其策略自由选择行动404并且环境406以系统状态408和奖励410作出反应。依据系统状态408和奖励410，智能体402随后选择新行动404以扩大奖励410。行动例如可以是工作状态的选择。

[0077] 图5示出混合RL架构500，其基于行动屏蔽概念。

[0078] 控制装置502在信号技术上与混合动力车(未被明确示出)的混合驱动系504相连。控制信号506从控制装置502被发送至混合驱动系504以实行混合驱动系504的设定。混合驱动系504配备有多个传感器(未被明确示出)并且将其状态508回传给控制装置502。

[0079] 状态508也像在传统的RL架构400中那样包括系统状态518和奖励520。所述数据在控制装置内部从基于规则的选择器514被传输至智能体510。

[0080] 基于规则的选择器514设立用于依据车辆状态从多个不变的工作模式和/或多个可变的工作模式选择一个或多个工作模式。该选择用于实行混合驱动系504的调设。基于规则的选择器514还设立用于按照在工作期间内的预设的行驶条件和/或环境条件根据所做选择来配置模式使能器512。模式使能器512是逻辑向量，其值通常被选择为逻辑0或逻辑1并且借此允许启用智能体的各个可变工作模式。模式使能器512的逻辑向量由基于规则的选择器514构建且被传输给智能体510。

[0081] 智能体510是基于增强学习的选择器。在工作中，智能体510基于系统状态518获得输入变量并且通过增强学习所基于的神经网络选择优先工作模式。优先工作模式由控制单元516来处理并被传输给基于规则的选择器514。

[0082] 智能体510所选的优先工作模式是可变的工作模式或不变的工作模式并且可以包

括用于驱动系部件的工作点。智能体510所做的选择基于“该智能体510所基于的策略”，其又基于增强学习且通过奖励520而可变。该控制的基础是首先由基于规则的选择器514对可用的工作模式做判断，智能体510从可用的工作模式中根据其策略选择最终被认定为最佳的工作状态。

[0083] 与图6和图7相关地解释该机制的一个具体实现方式。

[0084] 图6示出智能体510的基于增强学习的判断系统的结构600。结构600是人工智能、在此情况下是深度Q网络的网络结构。

[0085] 在输入层610，输入变量630作为输入用于网络。第一输入变量620可以是电池充电量。第二输入变量622可以是当前存在的扭矩。第三输入变量624可以是当前存在的车速。第四输入变量626可以是当前所选的变速器挡。第五输入变量628可以是电机工作状态(通或断)。除了这些输入变量630外，还能存在其它输入变量630。也可以选择不同于所述输入变量622~628的输入变量。

[0086] 除了输入层610外，结构600还包括多个中间层612~616和一个输出层618。在输出层618，针对不同的工作状态的Q值632被输出。Q值632呈现由智能体510预期的用于在一定状态中选择的工作状态的奖励520。智能体510的目的是优化奖励520。

[0087] 图7示出智能体510的基于增强学习的判断系统的本发明结构700。结构700基本上对应于图6的结构600并且通过模式使能器向量710和策略掩码712被扩展。在结构700中，输出层618将由人工神经网络预测的Q值632乘以策略掩码712，其是与模式使能器向量710长度相同的向量。策略掩码712提供大的负值给通过模式使能器710因零值被停用的工作模式。在第二输出层718中，仅如下的Q值632是高的，其已由模式使能器710启用，即，具有逻辑1。由选择器514通过模式使能器710停用的工作模式自动包含高负值并且能够与此相应地不被选择为优先工作模式，而是在由智能体完成选择的情况下被忽视。通过这种方式，智能体510遵守系统限制，而没有违反它。

[0088] 以上实施方式只以例子形式描述本发明。

[0089] 附图标记列表

[0090] 100驱动系

[0091] 110控制装置

[0092] 112内燃机

[0093] 114变速器

[0094] 116电池

[0095] 118变换器

[0096] 120电机

[0097] 122离合器

[0098] 200结构

[0099] 210观察者子系统

[0100] 212观察装置

[0101] 214观察装置

[0102] 220扭矩确定单元

[0103] 222油门踏板确定单元

- [0104] 224刹车踏板确定单元
- [0105] 226驱动力矩确定单元
- [0106] 230工作模式计算单元
- [0107] 232电工作模式
- [0108] 234能量回收工作模式
- [0109] 236增压工作模式
- [0110] 238其它工作模式
- [0111] 240工作模式评估单元
- [0112] 242评定单元
- [0113] 244选择单元
- [0114] 250控制单元
- [0115] 252限制单元
- [0116] 254序列单元
- [0117] 256协调单元
- [0118] 260驱动系控制器
- [0119] 270环境数据收集装置
- [0120] 272处理装置
- [0121] 274奖励计算装置
- [0122] 280智能体
- [0123] 282应用区域
- [0124] 283训练区域
- [0125] 284策略器
- [0126] 286经验区域
- [0127] 288学习算法
- [0128] 300结构
- [0129] 400 RL架构
- [0130] 402智能体
- [0131] 404行动
- [0132] 406环境
- [0133] 408系统状态
- [0134] 410奖励
- [0135] 500混合RL架构
- [0136] 502控制装置
- [0137] 504驱动系
- [0138] 506控制信号
- [0139] 508状态
- [0140] 510智能体
- [0141] 512模式使能器
- [0142] 514选择器

- [0143] 516控制单元
- [0144] 518系统状态
- [0145] 520奖励
- [0146] 600结构
- [0147] 610输入层
- [0148] 612第一中间层
- [0149] 614第二中间层
- [0150] 616第n中间层
- [0151] 618输出层
- [0152] 620第一输入变量
- [0153] 622第二输入变量
- [0154] 624第三输入变量
- [0155] 626第四输入变量
- [0156] 628第五输入变量
- [0157] 630输入变量
- [0158] 632 Q值
- [0159] 700结构
- [0160] 710模式使能器向量
- [0161] 712策略掩码
- [0162] 718第二输出层

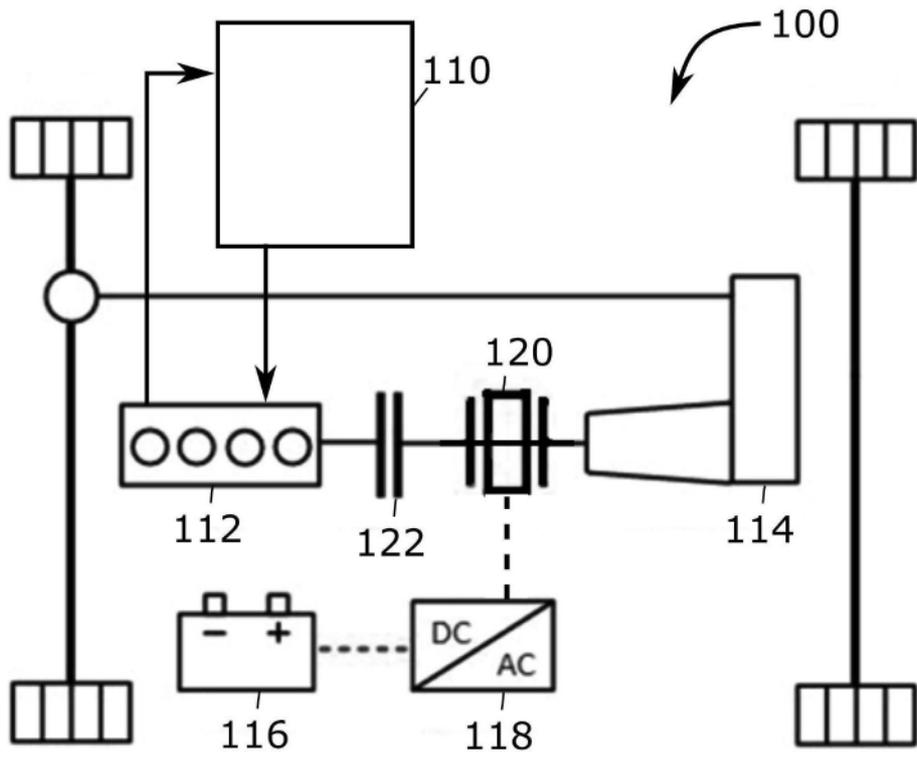


图1

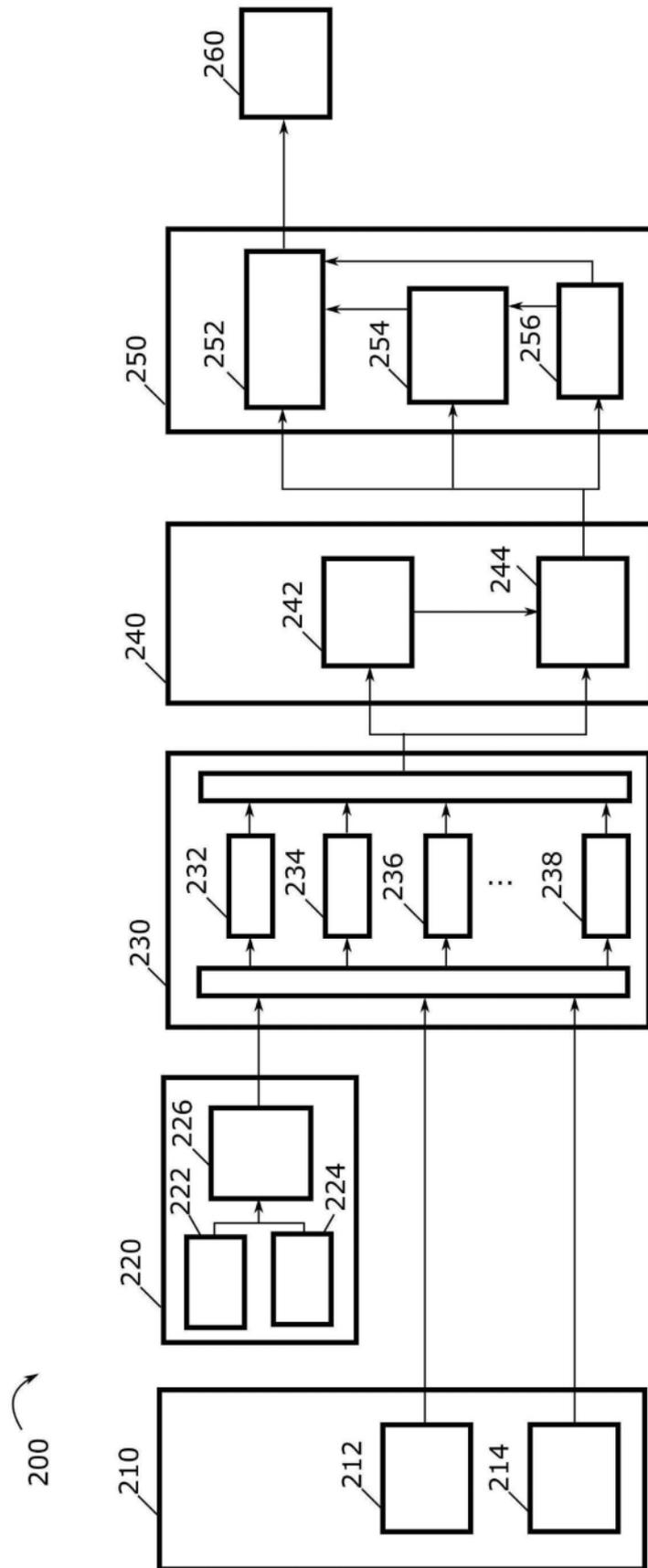


图2

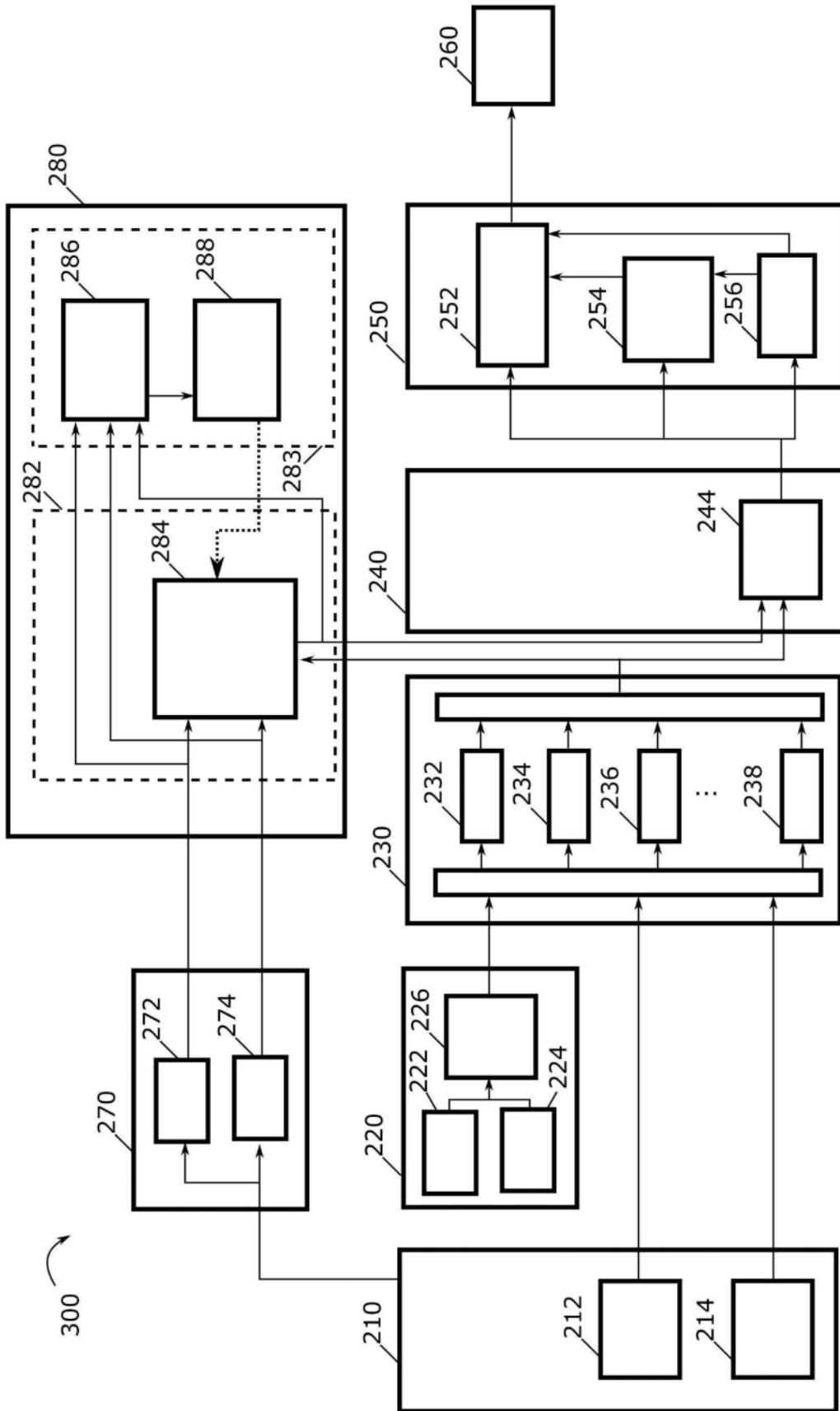


图3

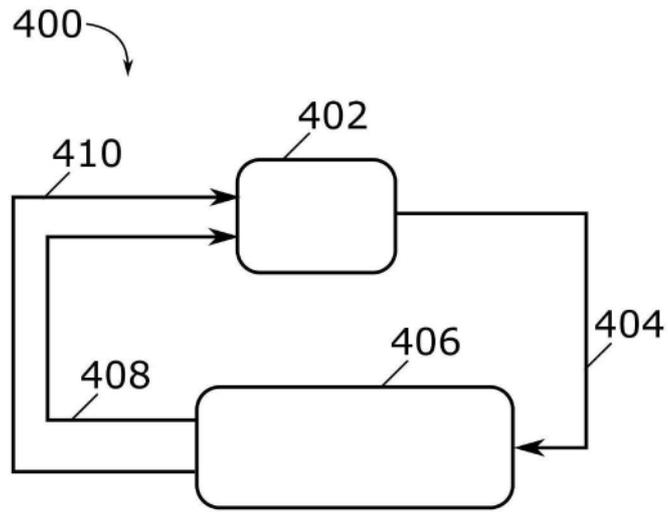


图4

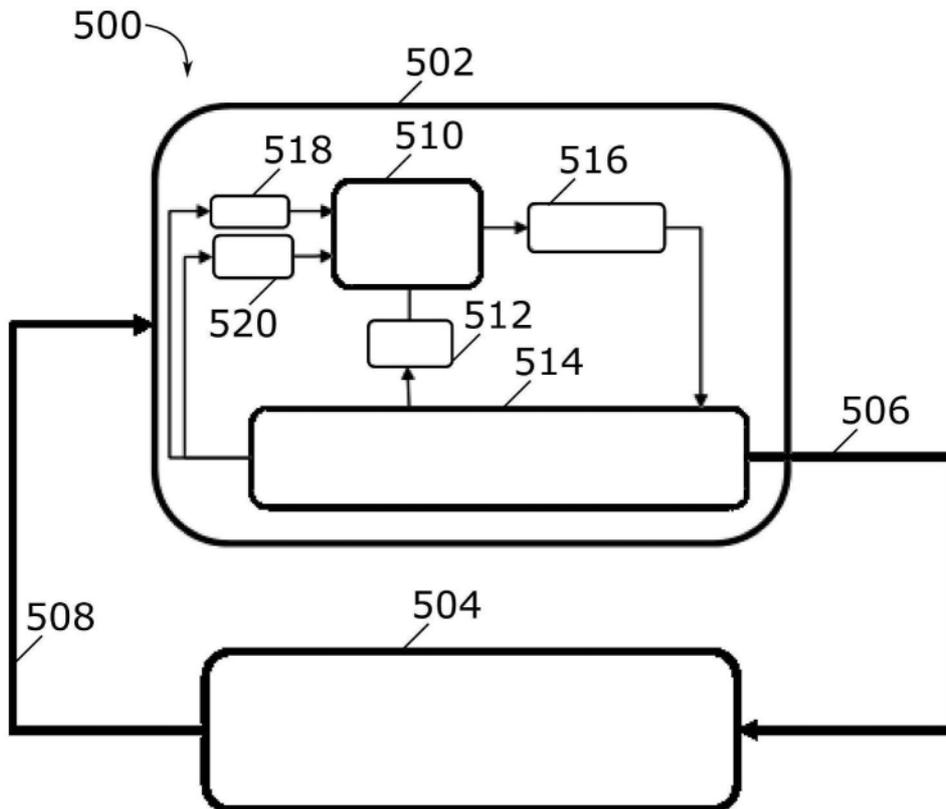


图5

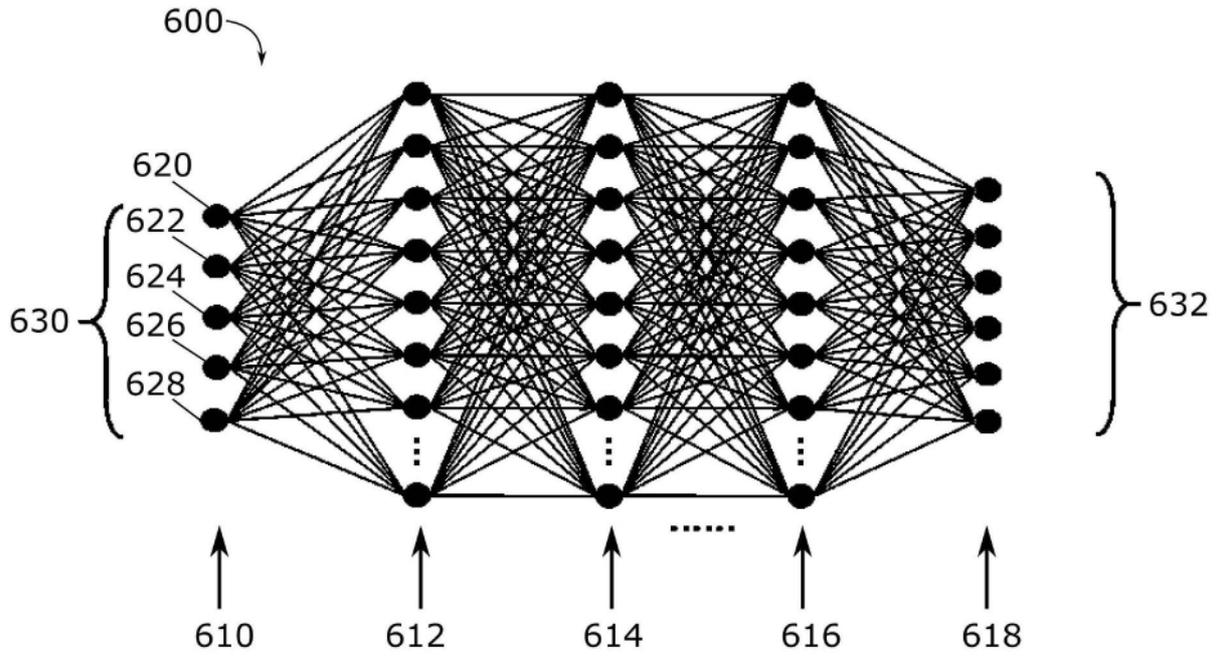


图6

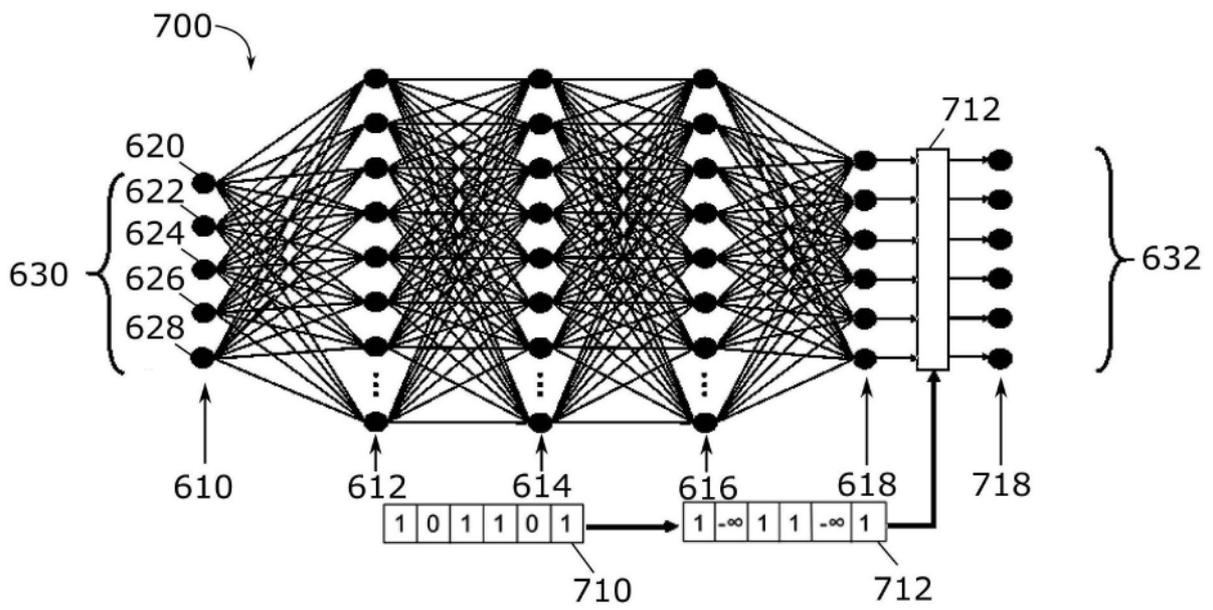


图7