



(12) 发明专利

(10) 授权公告号 CN 109955115 B

(45) 授权公告日 2021.04.23

(21) 申请号 201811603191.5

B25J 9/16 (2006.01)

(22) 申请日 2018.12.26

(56) 对比文件

(65) 同一申请的已公布的文献号
申请公布号 CN 109955115 A

CN 105729228 A, 2016.07.06

CN 106736825 A, 2017.05.31

CN 106557070 A, 2017.04.05

(43) 申请公布日 2019.07.02

CN 104750026 A, 2015.07.01

(30) 优先权数据
2017-249339 2017.12.26 JP

JP 2016124054 A, 2016.07.11

Nagdeve, L, V. Chaturvedi, and J.

(73) 专利权人 发那科株式会社
地址 日本山梨县

Vimal. Parametric optimization of abrasive water jet machining using Taguchi methodology.《Int'l J. Research in Engineering and Applied Sciences》.2012,第2卷

(72) 发明人 樽井贺彦

审查员 杨捷斐

(74) 专利代理机构 北京银龙知识产权代理有限公司 11243
代理人 范胜杰 文志

(51) Int. Cl.

B23Q 11/10 (2006.01)

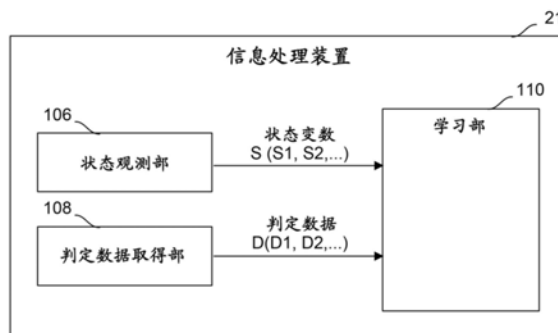
权利要求书2页 说明书9页 附图8页

(54) 发明名称

切屑去除装置以及信息处理装置

(57) 摘要

本发明提供一种切屑去除装置以及信息处理装置。该信息处理装置决定为了去除切屑而喷出物体的切屑去除装置的喷出条件,该信息处理装置观测表示切屑的去除效率的数据,来作为表示环境的当前状态的状态变量,另外,取得表示喷出条件的标签数据,然后将这些观测到的状态变量与取得的标签数据关联起来进行学习。



1. 一种信息处理装置,其决定为了去除切屑而喷出物体的切屑去除装置的喷出条件,其特征在于,

所述信息处理装置具有:

状态观测部,其观测表示进行了预定的切屑去除作业时的所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;

判定数据取得部,其取得判定数据,该判定数据表示在观测到表示所述去除效率的数据后进行了调整的所述喷出条件;以及

学习部,其使用所述状态变量和所述判定数据,将切屑的去除效率和所述喷出条件关联起来进行学习。

2. 一种信息处理装置,其决定为了去除切屑而喷出物体的切屑去除装置的喷出条件,其特征在于,

所述信息处理装置具有:

状态观测部,其观测表示在预定的喷出条件下进行了切屑去除作业时的所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;

判定数据取得部,其取得判定数据,该判定数据表示所述切屑的去除效率的恰当与否判定结果;以及

学习部,其使用所述状态变量和所述判定数据,学习所述喷出条件的调整与所述切屑的去除效率的相关性,

所述学习部具有:

回报计算部,其求出与所述恰当与否判定结果相关联的回报;以及

价值函数更新部,其使用所述回报对函数进行更新,该函数表示与所述切屑的去除效率相对的所述喷出条件的价值,

所述学习部通过学习在观测到所述切屑的去除效率时进行的所述喷出条件的调整行为的价值,来学习所述喷出条件的调整与所述切屑的去除效率的相关性。

3. 根据权利要求1或2所述的信息处理装置,其特征在于,

所述学习部通过多层构造来运算所述状态变量与所述判定数据。

4. 根据权利要求1或2所述的信息处理装置,其特征在于,

所述信息处理装置还具有决策部,该决策部基于所述学习部的学习结果,输出基于所述喷出条件的指令值。

5. 根据权利要求1或2所述的信息处理装置,其特征在于,

所述学习部使用针对多个所述切屑去除装置分别得到的所述状态变量以及所述判定数据来学习所述喷出条件。

6. 根据权利要求1或2所述的信息处理装置,其特征在于,

所述信息处理装置存在于云服务器。

7. 一种切屑去除装置,其为了去除切屑而喷出物体,其特征在于,

所述切屑去除装置包含决定喷出条件的信息处理装置,

所述信息处理装置具有:

状态观测部,其观测表示进行了预定的切屑去除作业时的所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;

判定数据取得部,其取得判定数据,该判定数据表示在观测到表示所述去除效率的数据后进行了调整的所述喷出条件;以及

学习部,其使用所述状态变量和所述判定数据,将切屑的去除效率和所述喷出条件关联起来进行学习。

8. 一种切屑去除装置,其为了去除切屑而喷出物体,其特征在于,

所述切屑去除装置包含决定喷出条件的信息处理装置,

所述信息处理装置具有:

状态观测部,其观测表示在预定的喷出条件下进行了切屑去除作业时的所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;

判定数据取得部,其取得判定数据,该判定数据表示所述切屑的去除效率的恰当与否判定结果;以及

学习部,其使用所述状态变量和所述判定数据,学习所述喷出条件的调整与所述切屑的去除效率的相关性,

所述学习部具有:

回报计算部,其求出与所述恰当与否判定结果相关联的回报;以及

价值函数更新部,其使用所述回报对函数进行更新,该函数表示与所述切屑的去除效率相对的所述喷出条件的价值,

所述学习部通过学习在观测到所述切屑的去除效率时进行的所述喷出条件的调整行为的价值,来学习所述喷出条件的调整与所述切屑的去除效率的相关性。

切屑去除装置以及信息处理装置

技术领域

[0001] 本发明涉及一种切屑去除装置以及信息处理装置,特别是涉及一种能够计算出最佳的切削液喷出条件的切屑去除装置以及信息处理装置。

背景技术

[0002] 近年来,工厂的自动化不断推进,通过无人运转来生产大量的产品。在机械加工中也不断推进自动化,通过使用机器人等来进行工件的交换,从而实现长时间的无人运转。

[0003] 在上述的自动化的机械加工系统中,为了防止加工不良,需要从加工区域自动地去除由于加工而产生的切屑的单元。切屑堆积的位置各种各样,因此要求正确地确定该堆积位置并高效地去除。

[0004] 例如在日本特开2016-120589号公报中记载有如下技术,即,在机器人等能够比较自由地移动的装置中设有视觉相机等切屑堆积状况检测单元和切削液喷嘴等切屑去除单元,通过向切屑堆积状况检测单元检测出的切屑堆积位置喷出切削液来去除切屑。

[0005] 但是,在日本特开2016-120589号公报记载的技术中,未使切削液的喷出条件(典型地为流量等)最佳化。因此,由于切屑的堆积状况等,存在以下问题:喷出超过需求的较多的切削液,对泵、软管等机械要素施加负担,或者消耗电力或去除切屑所需的时间。

发明内容

[0006] 因此,本发明的目的在于提供一种能够计算最佳的切削液喷出条件的切屑去除装置以及信息处理装置。

[0007] 本发明的一实施方式的信息处理装置决定为了去除切屑而喷出物体的切屑去除装置的喷出条件,其具有:状态观测部,其观测表示所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;标签数据取得部,其取得表示所述喷出条件的标签数据;以及学习部,其将所述状态变量与所述标签数据关联起来进行学习。

[0008] 本发明的另一实施方式的信息处理装置决定为了去除切屑而喷出物体的切屑去除装置的喷出条件,其具有:状态观测部,其观测表示所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;判定数据取得部,其取得判定数据,该判定数据表示所述切屑的去除效率的恰当与否判定结果;以及学习部,其使用所述状态变量和所述判定数据,将所述喷出条件与所述切屑的去除效率关联起来进行学习。

[0009] 所述学习部能够具备:回报计算部,其求出与所述恰当与否判定结果相关联的回报;价值函数更新部,其使用所述回报对表示所述喷出条件相对于所述切屑的去除效率的价值的函数进行更新。

[0010] 所述学习部能够通过多层构造运算所述状态变量和所述判定数据。

[0011] 所述信息处理装置还能够具备决策部,该决策部基于所述学习部的学习结果,输出基于所述喷出条件的指令值。

[0012] 所述学习部能够使用针对多个所述切屑去除装置分别得到的所述状态变量以及

所述判定数据来学习所述喷出条件。

[0013] 所述信息处理装置能够存在于云服务器。

[0014] 本发明的一实施方式的切屑去除装置为了去除切屑而喷出物体,其包含决定喷出条件的信息处理装置。并且,所述信息处理装置具有:状态观测部,其观测表示所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;标签数据取得部,其取得表示所述喷出条件的标签数据;以及学习部,其将所述状态变量与所述标签数据关联起来进行学习。

[0015] 本发明的另一实施方式的切屑去除装置为了去除切屑而喷出物体,其包含决定喷出条件的信息处理装置。并且,所述信息处理装置具有:状态观测部,其观测表示所述切屑的去除效率的数据,来作为表示环境的当前状态的状态变量;判定数据取得部,其取得判定数据,该判定数据表示所述切屑的去除效率的恰当与否判定结果;以及学习部,其使用所述状态变量和所述判定数据,将所述喷出条件与所述切屑的去除效率关联起来进行学习。

[0016] 通过本发明,能够提供一种能够计算最佳的切削液喷出条件的切屑去除装置以及信息处理装置。

附图说明

[0017] 图1表示切屑去除装置的一个例子。

[0018] 图2是信息处理装置的概要硬件结构图。

[0019] 图3是一实施方式的信息处理装置的概要功能框图。

[0020] 图4是表示信息处理装置的一个方式的概要功能框图。

[0021] 图5是表示图4的信息处理装置中的学习部执行的机器学习方法的一个方式的概要流程图。

[0022] 图6是另一实施方式的信息处理装置的概要功能框图。

[0023] 图7A说明神经元。

[0024] 图7B说明神经网络。

[0025] 图8是另一实施方式的信息处理装置的概要功能框图。

[0026] 图9是表示组装了切屑去除装置的切屑去除系统的一个方式的概要功能框图。

[0027] 图10是表示组装了切屑去除装置的切屑去除系统的一个方式的概要功能框图。

具体实施方式

[0028] 图1表示使用了机器人的机床的切屑去除装置的一个例子。

[0029] 切屑去除装置由以下部分构成:机床1;设置在该机床1外的作为移动装置的机器人2;安装在该机器人2的用于喷出切削液等的冷却枪4;安装在机器人2的视觉相机等视觉传感器3;以及搭载在机器人2的信息处理装置21。

[0030] 机床1在底座5上具备将作为加工空间的机械内部与外部切断的防溅罩7,防止切削液、由于工件的加工而产生的切屑从机械内部向外部飞散。在防溅罩7设有开口部9a、9b,在这些开口部9a、9b分别安装有可自由开闭的开闭门8a、8b。在机械内部配置了用于放置工件(未图示)的工作台6以及更换工具的工具更换装置10。机床1通过控制装置(未图示)来控制,进行工件的加工。

[0031] 机器人2是具有通过机器人2的控制装置20进行控制的多关节的机械臂的机器人，在机械臂的前端安装有冷却枪4。机器人2的控制装置20具有信息处理装置21。另外，冷却枪4也可以通过安装在机器人的机械臂前端的抓持进行抓持。

[0032] 机器人2在机床1不进行工件加工的期间，使机械臂的前端从开口部9b进入机械内部，通过安装在机器人2的视觉传感器3等检测装置来确认切屑的堆积状况，从安装在机械臂前端的冷却枪4对于堆积的切屑喷射用于工件加工的切削液即冷却液，清扫在机械内部堆积的切屑，将切屑排出到机外。另外，切屑去除不仅是喷射切屑液的方式，也可以是将切屑与空气，切削液等流体一起吸引的方式。

[0033] 信息处理装置21进行以下处理：对由视觉传感器3等检测装置检测出的切屑的堆积状况的信息进行处理，判定需要去除切屑的部位。另外，信息处理装置21既可以内置在机器人的控制装置20中，也可以相对于机器人的控制装置20独立地构成。另外，信息处理装置21进行的与切屑的堆积状况有关的信息处理在上述的日本特开2016-120589号公报中进行了具体的记载，是公知的技术，因此在这里省略详细说明。

[0034] 图2是表示图1所示的机器人控制装置20中的信息处理装置21的主要部分的概要硬件结构图。

[0035] 该信息处理装置21所具有的CPU11是整体地控制信息处理装置21的处理器。CPU11经由总线19读出存储在ROM12或者非易失性存储器14中的程序，按照程序来控制整个信息处理装置21。在RAM13中暂时存储计算数据、显示数据等。另外，ROM12或者非易失性存储器14中存储的程序、各种数据可以在执行时或使用时在RAM13中展开。

[0036] 输入输出装置60例如包含显示器、键盘、通信接口等，经由接口18接受来自CPU11的信息并进行显示，并且，将通过键盘、无线/有线通信从外部输入的各种数据向CPU11传递。

[0037] 图3是第1实施方式的信息处理装置21的概要功能框图。

[0038] 该信息处理装置21包含软件(学习算法等)以及硬件(处理器(CPU11)等)，该软件以及硬件用于通过所谓的机器学习来自己学习与切屑的去除效率(去除率、去除所需时间或去除所需能量等)相对的喷出条件(切削液等的流量或泵输出等)。该信息处理装置21要学习的东西相当于表示切屑的去除效率与喷出条件之间的相关性的模型构造。

[0039] 在图3中如功能块所示，信息处理装置21具有：状态观测部106，其观测包含数据S1的表示环境的当前状态的状态变量S，其中，上述数据S1表示切屑的去除效率(去除率、去除所需时间、或者去除所需能量等)；判定数据取得部108，其取得包含表示喷出条件(切削液等的流量或泵输出等)的数据D1的判定数据D；以及学习部110，其使用状态变量S和判定数据D，将切屑的去除效率和喷出条件关联起来进行学习。

[0040] 状态观测部106例如既可以构成为处理器(CPU11)的一个功能，也可以构成为用于使处理器(CPU11)发挥功能的存储在ROM12中的软件。

[0041] 关于状态观测部106观测的状态变量S中的切屑的去除效率S1，例如在使用日本特开2016-120589号公报的装置来进行切屑清扫时，可通过对清扫前的切屑量与清扫后残余的切屑量之间的比(去除率)、为了达成预定的去除率所需的清扫时间(去除所需时间)、为了达成预定的去除率所需的电能等能量(去除所需能量)等进行记录来取得上述去除效率S1。

[0042] 另外,除了切屑的去除效率S1以外,状态观测部106还可以在状态变量S中包含切屑的堆积状况S2、冷却枪4的喷嘴的位置以及朝向S3、切屑自身的形状S4等。关于切屑的堆积状况S2,能够作为由视觉传感器3得到的切屑的分布面积、由重量传感器(未图示)得到的切屑的重量来取得。例如,将加工空间即机械内部预先分成有限的区域编号,针对每个区域编号,基于分布面积或重量将分布状况用级别1、级别2……等多个级别来表现。关于喷嘴的位置以及朝向S3,能够通过日本特开2016-120589号公报的装置,作为坐标以及向量来取得。关于切屑自身的形状S4,能够通过视觉传感器3来取得。例如,能够针对切屑的每个大小或每种形状类型定义唯一的识别编号,输入该定义的识别编号来作为切屑自身的形状S4。

[0043] 判定数据取得部108例如既可以构成为处理器(CPU11)的一个功能,也可以构成为用于使处理器(CPU11)发挥功能的存储在ROM12的软件。

[0044] 关于判定数据取得部108取得的判定数据D中的喷出条件(切削液等的流量或泵输出等)D1,例如可通过记录在使用日本特开2016-120589号公报的装置进行切屑清扫时喷出的切削液等的流量、泵输出等来取得。

[0045] 另外,除了喷出条件D1以外,判定数据取得部108还可以在判定数据D中含有表示喷嘴的剖面面积或剖面形状的数据D2等。

[0046] 学习部110例如既可以构成为处理器(CPU11)的一个功能,也可以构成为用于使处理器(CPU11)发挥功能的存储在ROM12的软件。学习部110按照被总称为机器学习的任意的学习算法,来学习切屑的去除效率与喷出条件之间的关联性。学习部110能够基于包含上述的状态变量S和判定数据D的数据集合来反复执行学习。

[0047] 通过反复进行上述的学习循环,学习部110能够自动地识别出暗示切屑的去除效率与喷出条件之间的相关性的特征。在开始学习算法时,切屑的去除效率与喷出条件之间的相关性实质上是未知的,但是,随着学习的推进,学习部110逐渐地识别特征来解释相关性。如果将切屑的去除效率与喷出条件之间的相关性解释到能够某种程度信赖的水准,则学习部110反复输出的学习结果能够用于对当前状态(即切屑的去除效率)选择行为(即决策),该行为是指如何决定喷出条件。即,伴随着学习算法的推进,学习部110能够使切屑的去除效率与如何决定喷出条件这样的行为之间的相关性逐渐接近最佳解。

[0048] 如上所述,在信息处理装置21中,学习部110使用状态观测部106观测到的状态变量S和判定数据取得部108取得的判定数据D,按照机器学习算法,学习切屑的去除效率与喷出条件之间的关联性。状态变量S由切屑的去除效率这样的难以受到干扰影响的数据构成,另外,通过取得喷出条件来唯一地求出判定数据D。因此,根据信息处理装置21,通过使用学习部110的学习结果,不依赖于运算或估算,能够自动且正确地求出与切屑的去除效率相对应的喷出条件。

[0049] 并且,若不依赖于运算或估算而能够自动地求出喷出条件,则仅通过指定切屑的去除效率就能够迅速地确定与之相适合的喷出条件,因此能够高效地决定喷出条件。

[0050] 在具有上述结构的信息处理装置21中,学习部110所执行的学习算法没有特别限定,能够采用作为机器学习的公知的学习算法。图4是图3所示的信息处理装置21的一个方式,表示具有学习部110的结构,作为学习算法的一个例子该学习部110执行强化学习。

[0051] 强化学习是如下的方法:以试错的方式重复进行观测学习对象所在的环境的当前状态(即,输入),并且在当前状态下执行预定的行为(即,输出),对于该行为赋予某种回报

这样的循环,学习使回报的总和最大化的行为(在图4的信息处理装置21中能够使切屑的去除效率最大化的喷出条件)来作为最佳解。

[0052] 在图4所示的信息处理装置21中,学习部110具有:回报计算部112,其求出基于状态变量S决定了喷出条件的情况下的与切屑的去除效率的恰当与否判定结果相关联的回报R;以及价值函数更新部114,其使用回报R来更新表示喷出条件的价值的函数Q。在学习部110中,通过由价值函数更新部114反复进行函数Q的更新,来学习与切屑的去除效率相对应的喷出条件。

[0053] 对学习部110执行的强化学习的算法的一个例子进行说明。该例子的算法作为Q学习(Q-learning)而被知晓,是将行为主体的状态s以及在该状态s下行为主体可选择的行为a作为独立变量,对用于表示在状态s下选择了行为a时的行为a的价值的函数Q(s,a)进行学习的方法。在状态s下选择价值函数Q最高的行为a是最佳解。在状态s与行为a之间的相关性是未知的状态下开始进行Q学习,通过重复进行在任意的状态s下选择各种行为a的试错来反复更新价值函数Q使其接近最佳解。在此,作为在状态s下选择了行为a的结果,环境(即状态s)发生了变化时,得到与该变化相对应的回报(即,行为a的加权)r,以选择得到更高回报r的行为a的方式来诱导学习,由此能够在比较短的时间内使价值函数Q接近最佳解。

[0054] 关于价值函数Q的更新式,通常能够如下述式(1)那样来表示。在式(1)中, s_{t+1} 以及 a_t 分别是时刻t的状态以及行为,通过行为 a_t 状态变为 s_{t+1} 。 r_{t+1} 是通过状态从 s_t 变为 s_{t+1} 而得到的回报。 $\max Q$ 的项表示在时刻 t_{+1} 进行了成为最大价值Q(在时刻t考虑)的行为a时的Q。 α 以及 γ 分别是学习系数以及折扣率,在 $0 < \alpha \leq 1$ 、 $0 < \gamma \leq 1$ 的范围内任意地设定。

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha (r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t))$$

[0055]

... .. (1)

[0056] 在学习部110执行Q学习的情况下,状态观测部106观测到的状态变量S以及判定数据取得部108取得的判定数据D相当于更新式的状态s,应该如何决定喷出条件的行为相当于更新式的行为a,回报计算部112求出的回报R相当于更新式的回报r。由此,价值函数更新部114通过使用了回报R的Q学习来反复更新表示喷出条件的价值的函数Q。

[0057] 关于回报计算部112求出的回报R,例如在决定了喷出条件后基于决定的喷出条件进行了切屑去除时,在将切屑的去除效率判定为“适合”的情况(例如,去除率变高的情况、去除所需时间、去除所需能量变低的情况等)下设为正的回报R,在将切屑的去除效率判定为“否”的情况(例如,去除率降低的情况、去除所需时间,去除所需能量变高的情况等)下,设为负的回报R。正负的回报R的绝对值既可以彼此相同也可以不同。另外,作为判定的条件,可以组合判定数据D中包含的多个值来进行判定。

[0058] 价值函数更新部114能够持有将状态变量S、判定数据D、回报R与通过函数Q表示的行为价值(例如数值)关联起来进行了整理的行为价值表。在该情况下,价值函数更新部114更新函数Q的行为与价值函数更新部114更新行为价值表的行为为相同含义。在开始Q学习时,环境的当前状态与喷出条件之间的相关性是未知的,因此在行为价值表中,通过与任意决定的行为价值的值(函数Q)关联起来的方式来准备各种的状态变量S、判定数据D以及回报R。另外,若知晓判定数据D,则回报计算部112能够直接计算出与其对应的回报R,并将计算出的值R写入行为价值表。

[0059] 当使用对于行为赋予的回报R来推进Q学习时,向选择可得到更高回报R的行为的方向诱导学习,并且根据作为在当前状态下执行所选择的行为的结果而发生了变化的环境状态(即,状态变量S以及判定数据D),改写与当前状态下进行的行为有关的行为价值的值(函数Q)来更新行为价值表。通过反复进行该更新,以越是适当的行为值变得越大的方式改写在行为价值表中显示的行为价值的值(函数Q)。如此,未知的环境的当前状态(切屑的去除效率)和与之相对的行为(喷出条件)之间的相关性逐渐地明确。即,通过行为价值表的更新,切屑的去除效率与喷出条件之间的关系逐渐接近最佳解。

[0060] 参照图5,进一步说明图4的信息处理装置21中的学习部110所执行的上述Q学习的流程(即,机器学习方法的一个方式)。

[0061] 首先,在步骤SA01中,价值函数更新部114一边参照该时间点的行为价值表,一边任意地选择喷出条件来作为在状态观测部106观测到的状态变量S所示的当前状态下进行的行为。然后,价值函数更新部114在步骤SA02中,取入状态观测部106观测到的当前状态的状态变量S,在步骤SA03中,取入判定数据取得部108取得的当前状态的判定数据D。然后,价值函数更新部114在步骤SA04中,基于判定数据D判断喷出条件是否恰当,在恰当的情况下,在步骤SA05中,将回报计算部112求出的正的回报R用于函数Q的更新式,接着,在步骤SA06中,使用当前状态下的状态变量S以及判定数据D、回报R、行为价值的值(更新后的函数Q)来更新行为价值表。另一方面,在步骤SA04中,在判断为喷出条件不恰当的情况下,在步骤SA07中,将回报计算部112求出的负的回报R用于函数Q的更新式,接着,在步骤SA06中,使用当前状态下的状态变量S以及判定数据D、回报R、行为价值的值(更新后的函数Q)来更新行为价值表。学习部110通过反复进行步骤SA01~SA07的处理来反复更新行为价值表,从而推进喷出条件的学习。另外,针对判定数据D中包含的各个数据来执行步骤SA04到步骤SA07的求出回报R的处理以及价值函数的更新处理。

[0062] 图6是与图3所示的信息处理装置不同的另一方式的信息处理装置21,其具有学习部110,作为学习算法的另一个例子,该学习部110执行有监督学习。

[0063] 有监督学习与在输入和输出的关系未知的状态下开始学习的上述强化学习不同,是如下方法:预先大量地赋予输入和与该输入对应的输出为已知的数据组(称为训练数据),从这些训练数据中识别出暗示输入与输出之间的相关性的特征,由此来学习相关性模型,该相关性模型用于推定与新的输入相对的所需的输出(在图6的信息处理装置21中为切屑的去除效率与喷出条件之间的对应关系)。

[0064] 在图6所示的信息处理装置21中,学习部110具有:误差计算部116,其计算相关性模型M与从预先准备的训练数据T识别的相关性特征之间的误差E,上述相关性模型M用于从状态变量S以及判定数据D导出与切屑的去除效率相对应的喷出条件;以及模型更新部118,其更新相关性模型M使得缩小该误差E。学习部110通过由模型更新部118反复更新相关性模型M,来学习切屑的去除效率与喷出条件之间的对应关系。

[0065] 相关性模型M的初始值例如简单地(例如用一次函数)表现了切屑的去除效率与喷出条件之间的相关性,在开始有监督学习之前向学习部110赋予该相关性模型M的初始值。训练数据T例如由在过去的切屑的去除作业中取得的实际成绩值(切屑的去除效率与喷出条件为已知的数据组)构成,在开始有监督学习之前向学习部110赋予训练数据T。误差计算部116从向学习部110赋予的大量的训练数据T中识别出暗示切屑的去除效率与喷出条件之

间的相关性的相关性特征,并求出该相关性特征与当前状态下的状态变量S以及判定数据D所对应的相关性模型M之间的误差E。模型更新部118例如按照预先决定的更新规则,向该误差E变小的方向更新相关性模型M。

[0066] 在下次的学习循环中,误差计算部116使用通过按照更新后的相关性模型M来尝试去除切屑从而发生了变化的状态变量S以及判定数据D,关于与这些发生了变化的状态变量S以及判定数据D对应的相关性模型M求出误差E,模型更新部118再次更新相关性模型M。如上所述,未知环境的当前状态(切屑的去除效率)和与其相对的行为(喷出条件的决定)之间的相关性逐渐变得明确。即,通过相关性模型M的更新,使切屑的去除效率与喷出条件之间的关系逐渐接近最佳解。

[0067] 另外,在信息处理装置21中,还能够在学习的初始阶段,由学习部110执行有监督学习,在学习进行了某种程度的阶段,将通过有监督学习得到的切屑的去除效率所对应的喷出条件作为初始值,由学习部110执行强化学习。在该情况下,强化学习中的初始值具有某种程度的可靠性,因此,能够比较迅速地到达最佳解。

[0068] 在进行上述强化学习、有监督学习时,例如,能够使用神经网络来代替Q学习。图7A示意地表示神经元的模型。图7B示意地表示将图7A所示的神经元组合而成的三层的神经网络的模型。神经网络例如能够通过模拟神经元模型的运算装置、存储装置等来构成。

[0069] 图7A所示的神经元输出与多个输入x(在这里作为一个例子为输入 $x_1 \sim x_3$)相对的结果y。对各输入 $x_1 \sim x_3$ 乘以与该输入x对应的权重w($w_1 \sim w_3$)。由此,神经元输出由下式(2)表现的结果y。另外,在式(2)中,输入x、结果y以及权重w全部为向量。另外, θ 为偏置, f_k 为激活函数。

$$[0070] \quad y = f_k \left(\sum_{i=1}^n x_i w_i - \theta \right) \quad \dots \dots (2)$$

[0071] 图7B所示的三层的神经网络从左侧输入多个输入x(在这里作为一个例子为输入 $x_1 \sim$ 输入 x_3),从右侧输出结果y(在这里作为一个例子为结果 $y_1 \sim$ 结果 y_3)。在图示的例子中,对输入 x_1 、 x_2 、 x_3 分别乘以对应的权重(全体由 w_1 表示),向三个神经元N11、N12、N13都输入各个输入 x_1 、 x_2 、 x_3 。

[0072] 在图7B中,神经元N11~N13的各个输出全体由 z_1 表示。能够将 z_1 看做是提取了输入向量的特征量而得到的特征向量。在图示的例子中,对特征向量 z_1 分别乘以对应的权重(全体由 w_2 表示),将各个特征向量 z_1 均输入给两个神经元N21、N22。特征向量 z_1 表示权重 w_1 与权重 w_2 之间的特征。

[0073] 在图7B中,神经元N21~N22的各个输出全体由 z_2 表示。能够将 z_2 看做是提取了特征向量 z_1 的特征量而得到的特征向量。在图示的例子中,对特征向量 z_2 分别乘以对应的权重(全体由 w_3 表示),将各个特征向量 z_2 均输入给三个神经元N31、N32、N33。特征向量 z_2 表示权重 w_2 与权重 w_3 之间的特征。最后,神经元N31~N33分别输出结果 $y_1 \sim y_3$ 。

[0074] 另外,也可以采用使用了三层以上的层次的神经网络的所谓的深度学习的方法。

[0075] 在信息处理装置21中,能够将状态变量S和判定数据D作为输入x,由学习部110按照上述的神经网络进行多层构造的运算,由此能够输出喷出条件(结果y)。另外,在信息处理装置21中,还能够将神经网络用作强化学习中的价值函数,将状态变量S和行为a作为输入x,由学习部110按照上述的神经网络进行多层构造的运算,由此能够输出某个状态下的

该行为的价值(结果 y)。另外,在神经网络的动作模式中存学习模式和价值预测模式,例如能够在学习模式下使用学习数据组来学习权重 w ,在价值预测模式下使用学习到的权重 w 来进行行为的价值判定。另外,在价值预测模式下还能够进行检测、分类、推论等。

[0076] 能够将上述的信息处理装置21的结构记载为由处理器(CPU11)执行的机器学习方法(或软件)。该机器学习方法是学习喷出条件的机器学习方法,由计算机的CPU执行如下步骤:观测切屑的去除效率 S_1 来作为表示进行机器人控制的环境的当前状态的状态变量 S ;取得判定数据 D ,该判定数据 D 表示按照调整后的喷出条件的切屑的去除效率的恰当与否判定结果;使用状态变量 S 和判定数据 D ,将切屑的去除效率与喷出条件关联起来进行学习。

[0077] 图8表示第2实施方式的信息处理装置21。

[0078] 信息处理装置21从状态数据取得部130取得状态观测部106观测的状态变量 S 的切屑的去除效率 S_1 等来作为状态数据 S_0 。状态数据取得部130能够从机床1的各部、机器人2具有的各种传感器、作业者进行的适当数据的输入等取得该状态数据 S_0 。

[0079] 信息处理装置21除了包含用于通过机器学习自己学习喷出条件的软件(学习算法等)以及硬件(处理器(CPU11)等)以外,还包含用于将基于学习结果而求出的喷出条件作为针对机器人2的指令而输出的软件(运算算法等)以及硬件(处理器(CPU11)等)。

[0080] 决策部122例如既可以构成为处理器(CPU11)的一个功能,也可以构成为用于使处理器(CPU11)发挥功能的在ROM12中存储的软件。决策部122基于学习部110学习的结果生成与切屑的去除效率相对的指令值 C ,该指令值 C 包含用于决定喷出条件的指令,并将生成的指令值 C 输出。在决策部122将指令值 C 已输出给机器人2的情况下,与之相对应地环境的状态发生变化。

[0081] 状态观测部106在下一个学习周期观测在决策部122输出了针对环境的指令值 C 后发生了变化的状态变量 S 。学习部110使用发生了变化的状态变量 S ,例如更新价值函数 Q (即行为价值表),由此来学习喷出条件(强化学习的情况)。

[0082] 决策部122将用于指示基于学习结果求出的喷出条件的指令值 C 向机器人2输出。通过反复进行该学习周期,信息处理装置21推进喷出条件的学习,逐渐提高其可靠性。

[0083] 具有上述结构的信息处理装置21实现与第1实施方式的信息处理装置21(图3)相同的效果。特别是图8所示的第2实施方式的信息处理装置21能够通过决策部122的输出而使环境的状态变化。另一方面,在第1实施方式的信息处理装置21中,能够向外部装置请求相当于决策部的功能,以用于在环境中反映学习部110的学习结果。

[0084] 图9表示具有多个切屑去除装置160的切屑去除系统170。

[0085] 切屑去除系统170具有:至少进行同一作业的多个切屑去除装置160、160';以及将这些切屑去除装置160、160'相互连接的有线/无线的网络172,多个切屑去除装置160中的至少一个切屑去除装置160构成为具有上述的信息处理装置21的切屑去除装置160。另外,切屑去除系统170能够包含不具有信息处理装置21的切屑去除装置160'。切屑去除装置160、160'具有相同目的的作业所需的机构。

[0086] 在具有上述结构的切屑去除系统170中,多个切屑去除装置160、160'中的具有信息处理装置21的切屑去除装置160使用学习部110的学习结果,不依赖于运算、估算,能够自动且正确地求出与期望的切屑的去除效率相对应的喷出条件。另外,能够构成为至少一个切屑去除装置160的信息处理装置21基于针对其他的多个切屑去除装置160、160'分别得到

的状态变量S以及判定数据D,学习全部的切屑去除装置160、160'共用的喷出条件,由全部的切屑去除装置160、160'共享该学习结果。因此,通过切屑去除系统170,能够将更多样的数据集合(包含状态变量S以及判定数据D)作为输入,从而提高喷出条件的学习速度和可靠性。

[0087] 图10表示具有切屑去除装置160'的另一实施方式的切屑去除系统170'。

[0088] 该切屑去除系统170'具有:信息处理装置21、具有同一机械结构的多个切屑去除装置160'、将这些切屑去除装置160'和信息处理装置21相互连接的有线/无线网络172。

[0089] 在具有上述结构的切屑去除系统170'中,信息处理装置21基于针对多个切屑去除装置160'分别得到的状态变量S以及判定数据D,学习全部的切屑去除装置160'共用的与切屑的去除效率相对的喷出条件,使用该学习结果,不依赖于运算、估算,能够自动且正确地求出与切屑的去除效率相对的喷出条件。

[0090] 切屑去除系统170'能够具有信息处理装置21存在于在网络172中准备的云服务器等的结构。通过该结构,与多个切屑去除装置160'分别存在的场所、时期无关,能够在需要时将所需数量的切屑去除装置160'与信息处理装置21连接。

[0091] 从事于图9的切屑去除系统170、图10的切屑去除系统170'的作业者能够在信息处理装置21开始学习后的适当时期,判定信息处理装置21的对于喷出条件的学习达成度(喷出量的可靠性)是否达到要求级别。

[0092] 以上对本发明的实施方式进行了说明,但本发明不仅限于实施方式的例子,通过施加适当的变更,能够以各种方式进行实施。

[0093] 例如,信息处理装置21执行的学习算法、运算算法、控制算法等不限于上述的算法,能够采用各种各样的算法。

[0094] 另外,例如,可以使信息处理装置21能够输入、输出从哪个区域开始以何种顺序进行扫除这样的信息来作为状态变量S。例如,将机床内的加工空间分成有限的区域(区间)。针对每个区间,向学习器输入上述的状态变量S1、S2、……、标签以及表示扫除顺序的序数(若全部为9个区间,则是区间1至区间9中的任意一个)。由此,例如能够学习用于实现所期望的切屑去除效率的扫除顺序。

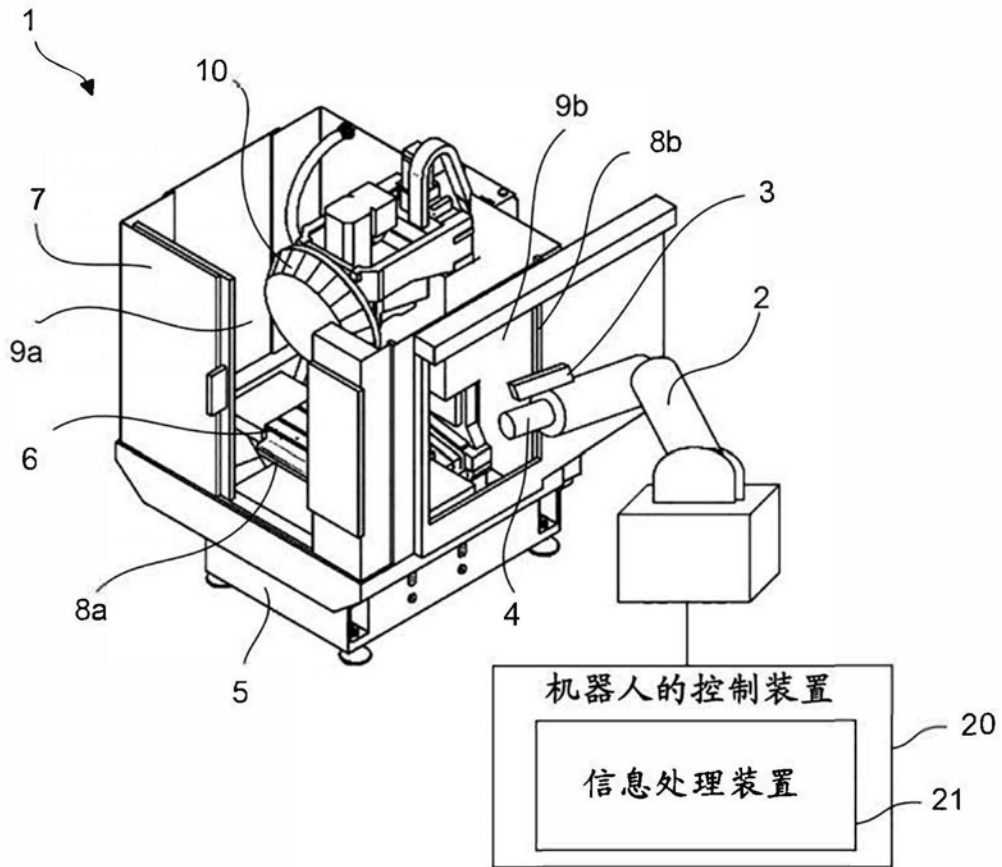


图1

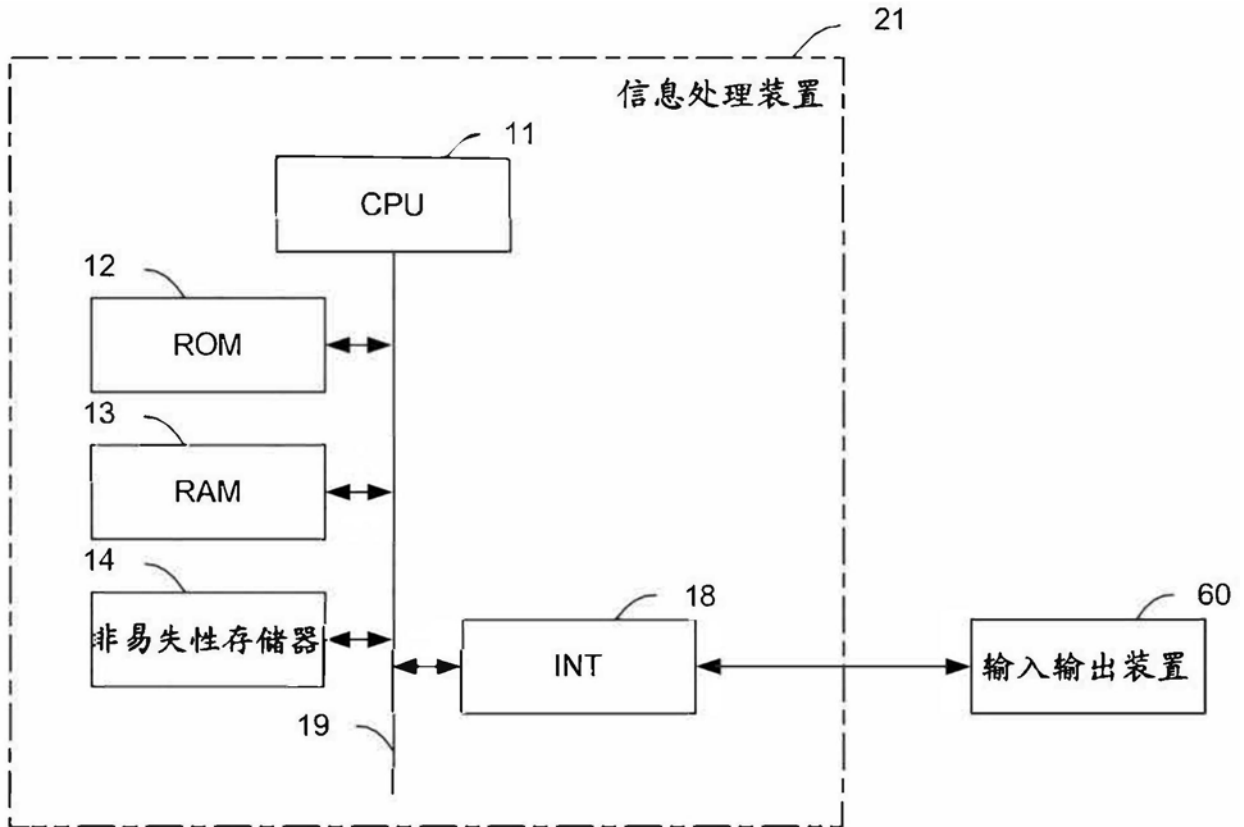


图2

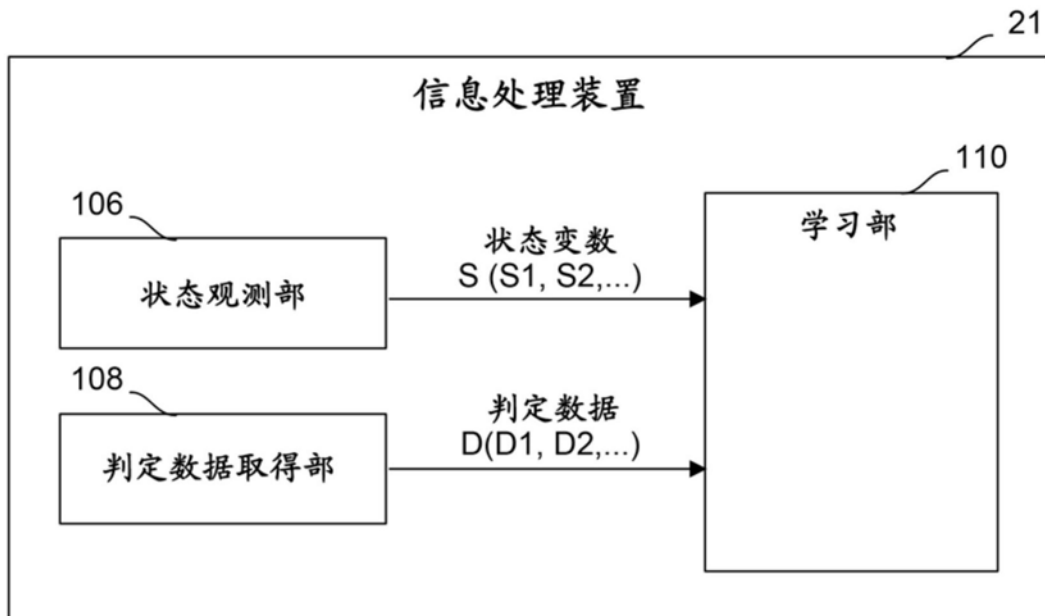


图3

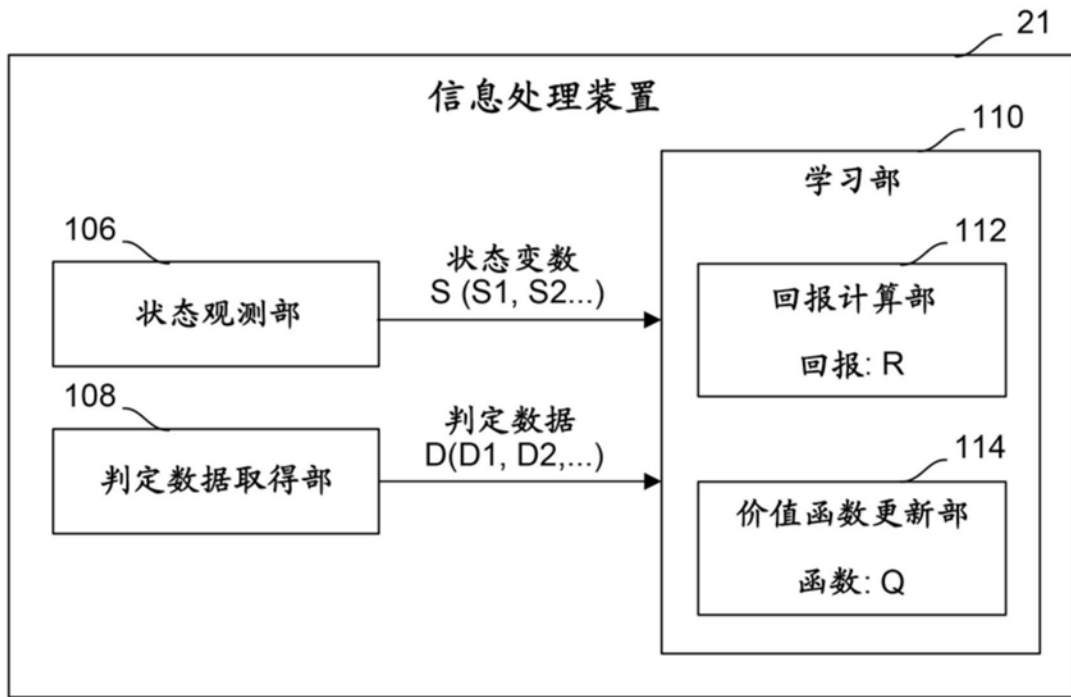


图4

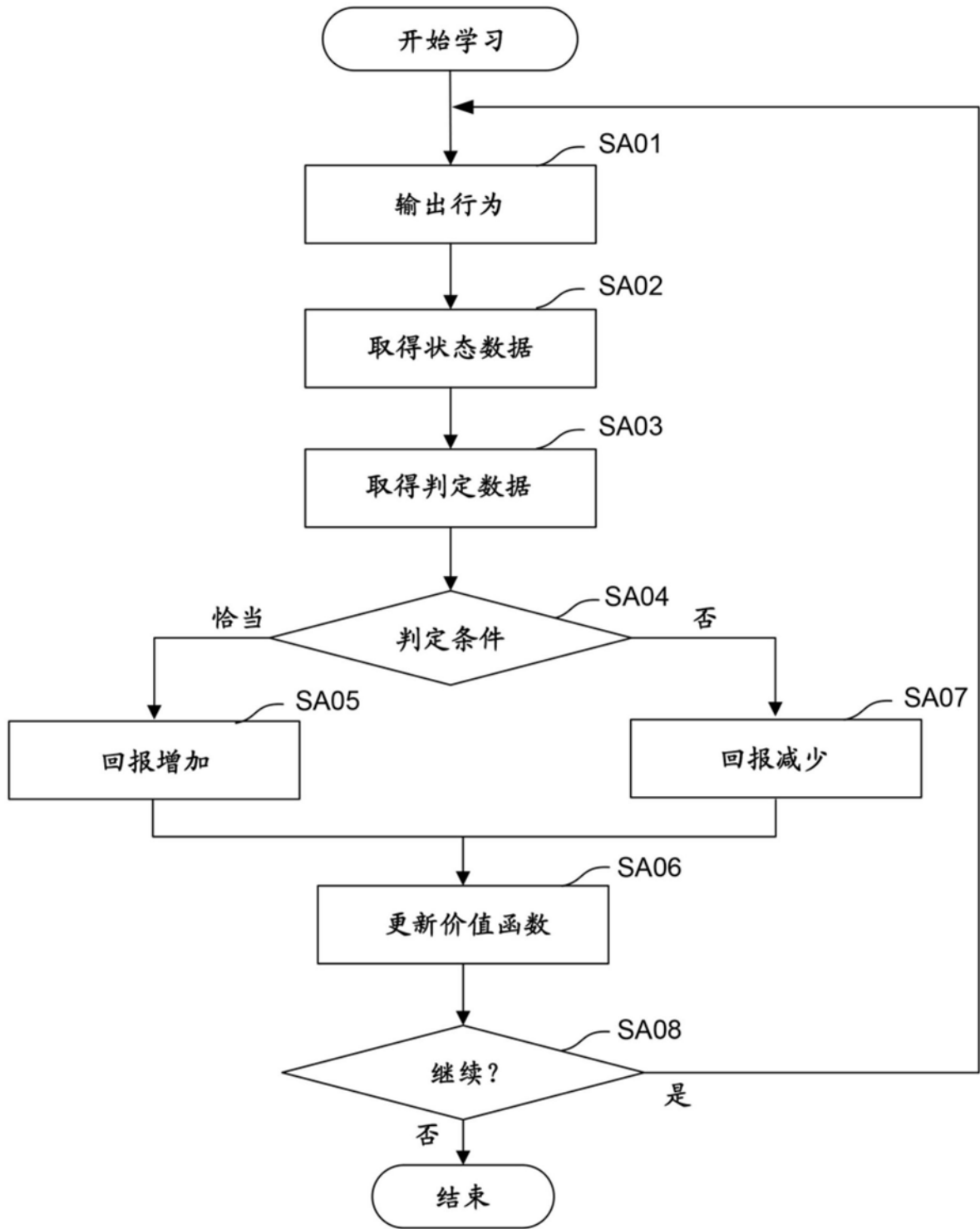


图5

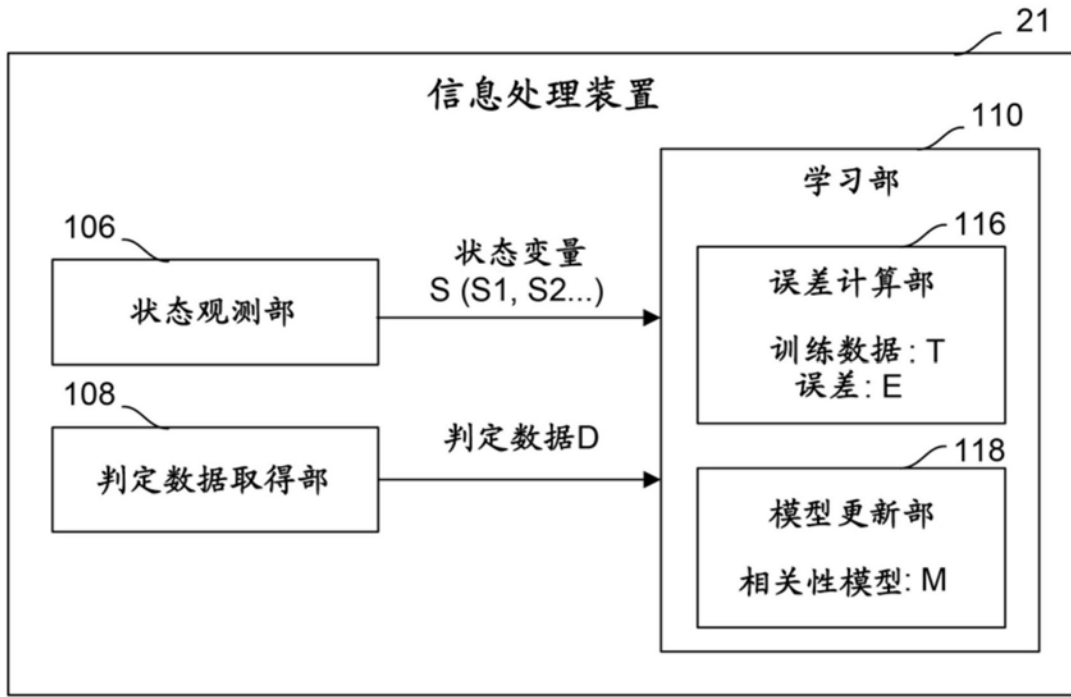


图6

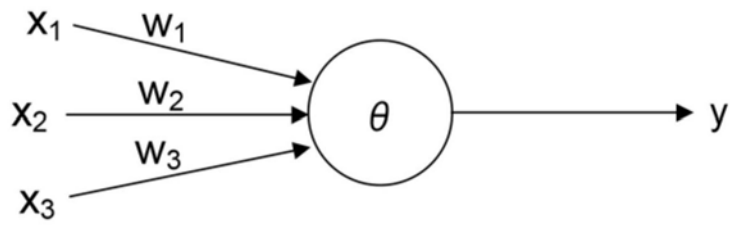


图7A

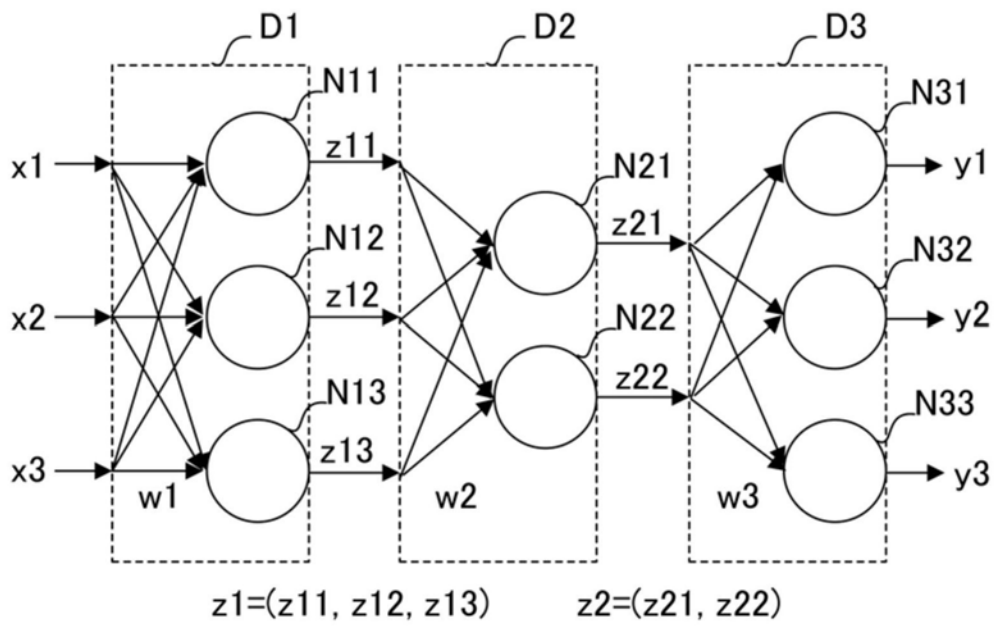


图7B

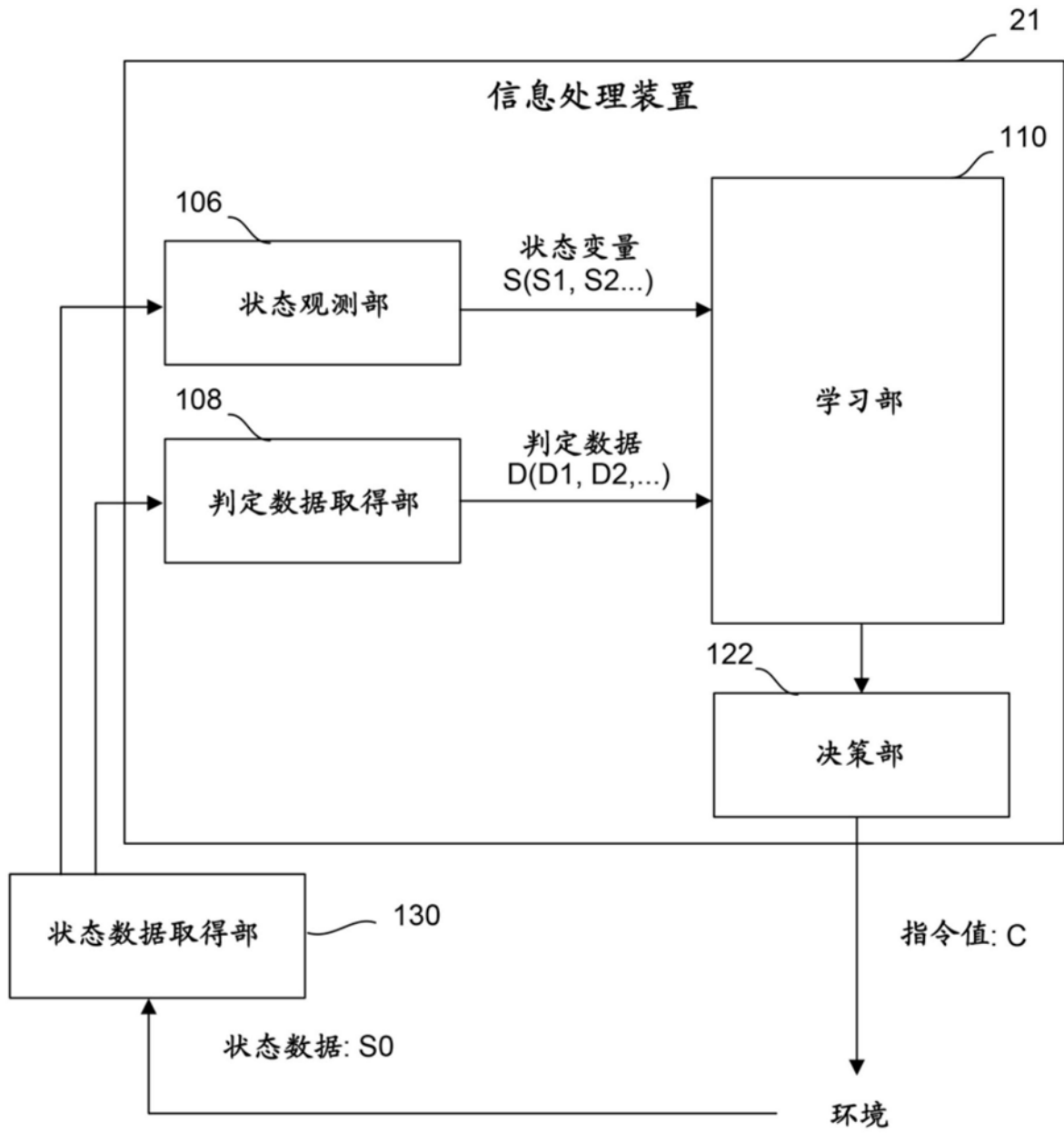


图8

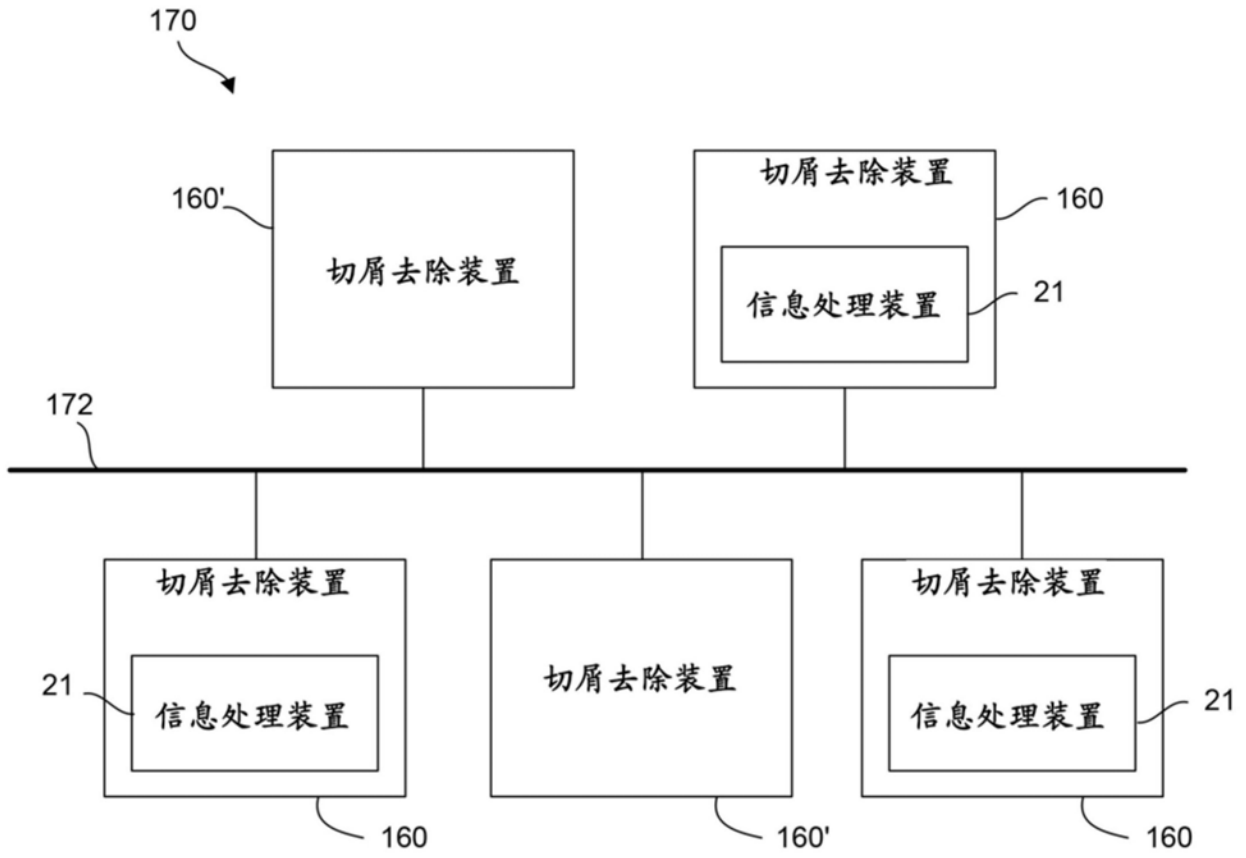


图9

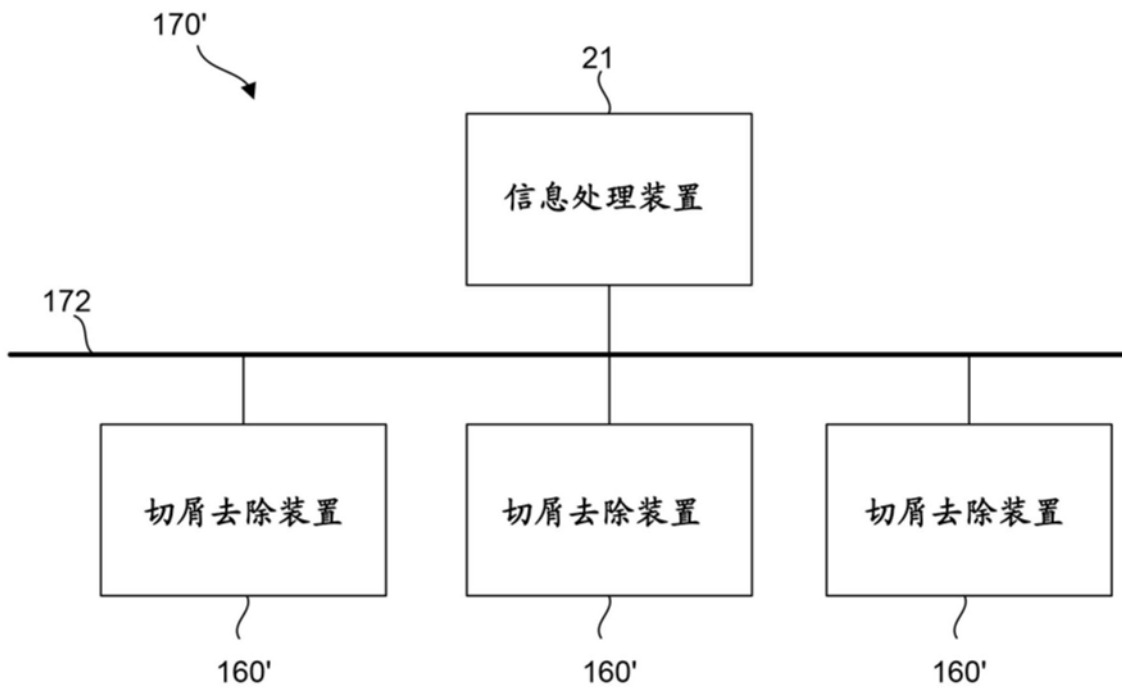


图10