(19)国家知识产权局

(12)发明专利申请

(10)申请公布号 CN 116783561 A
(43)申请公布日 2023.09.19

(21)申请号 202280009199.7
(22)申请日 2022.01.07
(30)优先权数据
2021-004401 2021.01.14 JP
(85)PCT国际申请进入国家阶段日
2023.07.06
(86)PCT国际申请的申请数据
PCT/JP2022/000336 2022.01.07
(87)PCT国际申请的公布数据
WO2022/153936 JA 2022.07.21
(71)申请人 发那科株式会社
地址 日本山梨县
(72)发明人 堀川祐太郎 安田将司
(74)专利代理机构 北京银龙知识产权代理有限公司 11243
专利代理师 许静 范胜杰
(51)Int.Cl.
G05B 19/4093(2006.01)

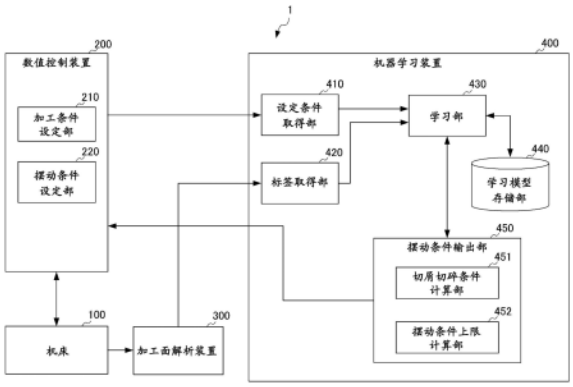
权利要求书2页 说明书19页 附图12页

(54)发明名称

机器学习装置

(57)摘要

本发明提供一种机器学习装置，其可以计算能够产生工具的空摆且能够实现良好的加工完成工件的评价数据的摆动条件。一种机器学习装置，其学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件，具有：设定条件取得部，其取得用于所述摆动加工的设定条件；标签取得部，其取得所述机床的加工完成工件的评价数据作为标签；学习部，其将所述设定条件与所述标签的组作为训练数据来进行监督学习，所述学习部具有：学习模型，其学习使所述加工完成工件的评价数据最佳化的摆动条件。



CN 116783561 A

1.一种机器学习装置,其学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件,其特征在于,

所述机器学习装置具有:

设定条件取得部,其取得用于所述摆动加工的设定条件;

标签取得部,其取得所述机床的加工完成工件的评价数据作为标签;以及

学习部,其将所述设定条件与所述标签的组作为训练数据来进行监督学习,

所述学习部具有:学习模型,其学习使所述加工完成工件的评价数据最佳化的摆动条件。

2.根据权利要求1所述的机器学习装置,其特征在于,

所述机器学习装置还具有:摆动条件输出部,其根据所述学习模型,输出使所述加工完成工件的评价数据最佳化的最佳摆动条件。

3.根据权利要求2所述的机器学习装置,其特征在于,

所述摆动条件输出部具有:切屑切碎条件计算部,其计算所述工件的切屑能够切碎的切屑切碎摆动条件,

所述摆动条件输出部根据所述学习模型,输出满足所述切屑切碎摆动条件的所述最佳摆动条件。

4.根据权利要求2或3所述的机器学习装置,其特征在于,

所述摆动条件输出部具有:摆动条件上限计算部,其计算不超过预先设定的上限值的上限摆动条件,

所述摆动条件输出部根据所述学习模型,输出满足所述上限摆动条件的所述最佳摆动条件。

5.根据权利要求1～4中任一项所述的机器学习装置,其特征在于,

所述加工完成工件的所述评价数据包含所述加工完成工件的面粗糙度、圆度或尺寸精度的数据。

6.一种机器学习装置,其学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件,其特征在于,

所述机器学习装置具有:

设定条件取得部,其取得用于所述摆动加工的设定条件作为状态信息;

判定信息取得部,其取得所述机床的加工完成工件的评价数据作为判定信息;

行为信息输出部,其输出表示针对当前的状态应当如何变更所述摆动条件的行为信息;

回报计算部,其根据所述判定信息,计算强化学习中回报的值;

价值函数更新部,其根据所述状态信息、所述行为信息以及所述回报来更新价值函数,该价值函数用于确定所述机床的摆动条件的价值;以及

摆动条件输出部,其根据所述价值函数,输出使所述加工完成工件的评价数据最佳化的最佳摆动条件。

7.根据权利要求6所述的机器学习装置,其特征在于,

在所述加工完成工件的所述评价数据小于规定阈值时,所述回报计算部将所述回报设为正值,在所述评价数据超过所述规定阈值时,将所述回报设为负值。

8.根据权利要求6或7所述的机器学习装置,其特征在于,

所述摆动条件输出部具有:切屑切碎条件计算部,其计算所述工件的切屑能够切碎的切屑切碎摆动条件,

所述摆动条件输出部输出满足所述切屑切碎摆动条件的所述最佳摆动条件。

9.根据权利要求6～8中任一项所述的机器学习装置,其特征在于,

所述摆动条件输出部具有:摆动条件上限计算部,其计算不超过预先设定的上限值的上限摆动条件,

所述摆动条件输出部输出满足所述上限摆动条件的所述最佳摆动条件。

10.根据权利要求6～9中任一项所述的机器学习装置,其特征在于,

所述加工完成工件的所述评价数据包含所述加工完成工件的面粗糙度、圆度或尺寸精度的数据。

11.根据权利要求10所述的机器学习装置,其特征在于,

在所述评价数据是所述加工完成工件的面粗糙度数据时,所述回报计算部根据理论面粗糙度的近似式来决定所述规定阈值。

12.根据权利要求11所述的机器学习装置,其特征在于,

所述面粗糙度数据包含算术平均粗糙度、最大高度、最大峰高、最大谷深、平均高度、最大截面高度以及负载长度比中的至少一个。

13.根据权利要求1～12中任一项所述的机器学习装置,其特征在于,

所述设定条件具有:

加工条件,其包含所述机床中的所述工具的进给速度、主轴转速、坐标值、所述工具的刀尖、所述工具的种类以及所述工件的材料中的至少一个;以及

所述摆动条件,其包含所述机床的摆动频率和摆动振幅中的至少一个。

14.根据权利要求1～13中任一项所述的机器学习装置,其特征在于,

所述机器学习装置由多个数值控制装置共享。

15.根据权利要求1～13中任一项所述的机器学习装置,其特征在于,

所述机器学习装置设置在云服务器上。

机器学习装置

技术领域

[0001] 本发明涉及一种机器学习装置。

背景技术

[0002] 以往,已知有如下的摆动切削(例如,参照专利文献1):一边将工件和切削工具沿加工方向进给一边进行切削,使工件和切削工具同时相对地往复振动。在摆动切削加工中,为了切碎切屑,需要设定频率以及振幅等摆动条件,以使工具相对于工件以一定周期空摆(air cut)。

[0003] 然后,操作员通过确认所设定的摆动条件的波形,来确认是否产生了上次路径与当前路径的公差,即是否产生了切屑切断。

[0004] 现有技术文献

[0005] 专利文献

[0006] 专利文献1:日本特开2019-28597号公报

发明内容

[0007] 发明要解决的课题

[0008] 例如,通过摆动切削进行加工的加工完成工件的面粗糙度、圆度、尺寸精度等那样的、与加工完成工件的精加工情况相关的评价数据,与不应用摆动切削时相比恶化的情况较多。这是因为摆动切削的振动和切削量的变化等导致的,但评价数据恶化的机理非常复杂。因此,要求计算能够产生工具的空摆且能够实现良好的加工完成工件的评价数据的摆动条件。

[0009] 用于解决课题的手段

[0010] 本公开的一方式的机器学习装置,其学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件,其中,所述机器学习装置具有:设定条件取得部,其取得用于所述摆动加工的设定条件;标签取得部,其取得所述机床的加工完成工件的评价数据作为标签;以及学习部,其将所述设定条件与所述标签的组作为训练数据来进行监督学习,所述学习部具有:学习模型,其学习使所述加工完成工件的评价数据最佳化的摆动条件。

[0011] 本公开的一方式的机器学习装置,其学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件,其中,所述机器学习装置具有:设定条件取得部,其取得用于所述摆动加工的设定条件作为状态信息;判定信息取得部,其取得所述机床的加工完成工件的评价数据作为判定信息;行为信息输出部,其输出表示针对当前的状态应当如何变更所述摆动条件的行为信息;回报计算部,其根据所述判定信息,计算强化学习中回报的值;价值函数更新部,其根据所述状态信息、所述行为信息以及所述回报来更新价值函数,该价值函数用于确定所述机床的摆动条件的价值;以及摆动条件输出部,其根据所述价值函数,输出使所述加工完成工件的评价数据最佳化的最佳摆动条件。

[0012] 发明效果

[0013] 根据本发明,可以计算能够产生工具的空摆且能够实现良好的加工完成工件的评价数据的摆动条件。

附图说明

[0014] 图1是表示第一实施方式的控制系统的概要的图。

[0015] 图2是表示第一实施方式的机器学习装置进行的学习模型的构筑处理流程的流程图。

[0016] 图3是表示第一实施方式的机器学习装置进行的摆动条件的输出处理流程的流程图。

[0017] 图4是表示第二实施方式的控制系统的概要的图。

[0018] 图5是表示第二实施方式的机器学习装置进行的价值函数的更新处理流程的流程图。

[0019] 图6是表示第二实施方式的机器学习装置进行的输出摆动条件的处理流程的流程图。

[0020] 图7是表示第三实施方式的控制系统的概要的图。

[0021] 图8是表示第三实施方式的机器学习装置进行的学习模型的构筑处理流程的流程图。

[0022] 图9是表示第三实施方式的机器学习装置进行的摆动条件的输出处理流程的流程图。

[0023] 图10是表示第四实施方式的控制系统的概要的图。

[0024] 图11是表示第四实施方式的机器学习装置进行的价值函数的更新处理流程的流程图。

[0025] 图12是表示第四实施方式的机器学习装置进行的输出摆动条件的处理流程的流程图。

具体实施方式

[0026] [第一实施方式]

[0027] 以下,对本发明的实施方式的一例进行说明。图1是表示第一实施方式的控制系统1的概要的图。控制系统1的目的在于:通过使用机器学习计算摆动条件,来计算在机床100中产生工具的空摆且能够实现良好的加工完成工件的评价数据的摆动条件。特别是,第一实施方式的控制系统1的目的在于:通过使用机器学习计算摆动条件,来计算在机床100中产生工具的空摆且能够实现良好的面粗糙度数据的摆动条件。

[0028] 控制系统1具有:机床100、数值控制装置200、加工面解析装置300以及机器学习装置400。在此,机床100和数值控制装置200为一对一的组,以能够通信的方式连接。此外,机床100和数值控制装置200可以经由连接接口直接连接,或者也可以经由LAN(Local Area Network)等网络连接。

[0029] 另外,数值控制装置200、加工面解析装置300以及机器学习装置400分别经由连接接口直接连接,或者分别经由网络连接,能够相互进行通信。此外,网络例如是在工厂内构筑的LAN、因特网、公共电话网、或者它们的组合。网络中的具体通信方式、是有线连接和无

线连接中的哪一个等不做特别限定。

[0030] 机床100按照数值控制装置200的控制，使用摆动切削来加工工件。机床100具有工具、主轴、进给轴等用于进行摆动切削加工的一般结构。

[0031] 数值控制装置200是用于控制机床100的装置。数值控制装置200具有加工条件设定部210和摆动条件设定部220。

[0032] 加工条件设定部210设定设定条件，该设定条件包含用于在机床100中进行摆动加工的加工条件以及摆动条件。加工条件设定部210将所设定的设定条件输出给机器学习装置400。在此，设定条件具有：加工条件，其包含机床100中工具的进给速度、主轴转速、坐标值、工具的刀尖、工具的种类以及工件的材料中的至少一个；摆动条件，其包含机床100的摆动频率和摆动振幅中的至少一个。

[0033] 摆动条件设定部220在数值控制装置200中设定从机器学习装置400输出的摆动条件，并将摆动条件输出给机床100。

[0034] 加工面解析装置300测定以及计算在机床100中加工后的加工完成工件的面粗糙度数据，并将计算出的面粗糙度数据输出给机器学习装置400。面粗糙度数据例如包含算术平均粗糙度、最大高度、最大峰高、最大谷深、平均高度、最大截面高度以及负载长度比(load length ratio)中的至少一个。

[0035] 此外，在摆动切削加工中，工具周期性振动，因此，面粗糙度在工件(例如，圆筒形状)的各种部位发生变化。因此，在进行面粗糙度数据的测定时，需要在工件的圆筒侧面多次测定线粗糙度。另外，面粗糙度数据的计算可以使用在圆筒侧面测定多个部位的线粗糙度并计算平均面粗糙度的方法(算术平均粗糙度Ra)、以及在圆筒侧面测定多个部位的线粗糙度并计算成为最大值的线粗糙度的方法(最大高度Rz)等方法。

[0036] 机器学习装置400使用从数值控制装置200受理的设定条件以及从加工面解析装置300受理的面粗糙度数据进行机器学习。然后，机器学习装置400通过进行机器学习，来构筑用于输出摆动条件的学习模型。机器学习装置400具有：设定条件取得部410、标签取得部420、学习部430、学习模型存储部440以及摆动条件输出部450。

[0037] 设定条件取得部410从数值控制装置200取得包含摆动加工用的加工条件以及摆动条件的设定条件。

[0038] 标签取得部420从加工面解析装置300取得机床100的加工完成工件的面粗糙度数据作为标签。在此，标签是指在机器学习中应该与输入对应的正确的输出。

[0039] 这样，作为学习对象的用于摆动加工的设定条件和面粗糙度数据的标签成为组，输入到学习部430。该设定条件和标签的组相当于机器学习中的训练数据。在此，优选准备多个用于进行机器学习的训练数据。

[0040] 学习部430根据输入的训练数据进行机器学习，由此，构筑学习模型。即，学习部430学习使加工完成工件的面粗糙度最佳化的摆动条件。将学习部430构筑的学习模型输出给学习模型存储部440。

[0041] 学习模型存储部440存储由学习部430学习到的学习模型。学习模型存储部440存储的学习模型由摆动条件输出部450利用。此外，在构筑了学习模型后，在新取得了训练数据时，学习部430可以通过对学习模型存储部440存储的学习模型进一步进行监督学习，来更新曾经构筑的学习模型。

[0042] 另外，机器学习装置400也可以与其他机器学习装置之间共享学习模型存储部440存储的学习模型。如果在多个机器学习装置中共享学习模型，则能够在各机器学习装置中分散地进一步进行监督学习，因此，能够提高监督学习的效率和精度。

[0043] 摆动条件输出部450根据学习模型存储部440存储的学习模型，输出使加工完成工件的面粗糙度最佳化的最佳摆动条件。

[0044] 另外，摆动条件输出部450具有切屑切碎条件计算部451和摆动条件上限计算部452。

[0045] 切屑切碎条件计算部451计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。并且，摆动条件输出部450根据学习模型，输出满足切屑切碎摆动条件的最佳摆动条件。在此，切屑切碎摆动条件的计算可以预先将能够切碎切屑的摆动条件的组合保存为表数据，或者也可以在根据摆动条件计算的加工路径中，计算在上次路径与本次路径之间产生重叠这样的各摆动条件的关系式。

[0046] 摆动条件上限计算部452计算不超过预先设定的上限值的上限摆动条件。然后，摆动条件输出部450根据学习模型，输出满足上限摆动条件的最佳摆动条件。在此，预先设定的上限值例如表示摆动加工中的频率或振幅的上限值、或者摆动加工中的速度或加速度的上限值。判别是否满足上限摆动条件例如可以通过将各种上限值保存为参数或者通过加工程序进行指定、根据摆动条件(频率以及振幅)计算速度以及加速度来实现。

[0047] 接下来，参照图2的流程图对机器学习装置400的学习模型的构筑处理进行说明。

[0048] 图2是表示第一实施方式的机器学习装置400进行的学习模型的构筑处理流程的流程图。

[0049] 在步骤S1中，设定条件取得部410从数值控制装置200取得包含摆动加工用的加工条件以及摆动条件的设定条件。

[0050] 在步骤S2中，标签取得部420从加工面解析装置300取得机床100的加工完成工件的面粗糙度数据作为标签。

[0051] 在步骤S3中，学习部430在被输入以设定条件和标签为组的训练数据时，根据输入的训练数据来执行机器学习。学习部430使用输入的训练数据来执行机器学习。学习部430例如通过回归分析、神经网络、最小二乘法、阶梯法等进行监督学习。

[0052] 监督学习可以通过在线学习进行，也可以通过批量学习进行，还可以通过小批量学习进行。在线学习是从声音数据提取特征量，每当被输入标签而制作训练数据时，立即进行监督学习的学习方法。另外，批量学习是指如下学习方法：从声音数据提取特征量，在反复被输入标签而制作训练数据的期间，收集与反复对应的多个训练数据，使用收集到的全部训练数据来进行监督学习。

[0053] 并且，小批量学习是指在线学习和批量学习的中间的、每当积攒某种程度训练数据时进行监督学习的学习方法。

[0054] 在步骤S4中，学习部430判定是否结束监督学习。在此，使监督学习结束的条件可以任意决定。例如，学习部430可以在神经网络的输出与标签的误差值为规定值以下时，结束监督学习。另外，学习部430也可以在反复进行了预先确定的次数的监督学习时，结束监督学习。在结束监督学习时(是)，处理转移到步骤S5。另一方面，在未结束监督学习时(否)，处理转移到步骤S3。

[0055] 在步骤S5中，学习部430将通过到此时间点为止的监督学习而构筑的学习模型输出并存储在学习模型存储部440中。另外，在取得了新的训练数据时，机器学习装置400可以对学习模型进行进一步的机器学习。之后，机器学习装置400结束本处理。

[0056] 接下来，参照图3的流程图对机器学习装置400的摆动条件的输出处理进行说明。

[0057] 图3是表示第一实施方式的机器学习装置400进行的摆动条件的输出处理流程的流程图。

[0058] 在步骤S11中，设定条件取得部410从数值控制装置200取得包含摆动加工用的加工条件以及摆动条件的设定条件。

[0059] 在步骤S12中，切屑切碎条件计算部451计算通过机床100的摆动加工能够切碎工件切屑的切屑切碎摆动条件。

[0060] 在步骤S13中，摆动条件上限计算部452计算不超过预先设定的上限值的上限摆动条件。

[0061] 在步骤S14中，摆动条件输出部450根据在步骤S11中取得的设定条件和学习模型存储部440存储的学习模型，输出使加工完成工件的面粗糙度最佳化的最佳摆动条件。

[0062] 在步骤S15中，摆动条件输出部450判定在步骤S13中输出的最佳摆动条件是否满足在步骤S12中计算出的切屑切碎摆动条件。在最佳摆动条件满足切屑切碎摆动条件时(是)，处理转移到步骤S16。另一方面，在最佳摆动条件不满足切屑切碎摆动条件时(否)，处理再次转移到步骤S14。

[0063] 在步骤S16中，摆动条件输出部450判定在步骤S13中输出的最佳摆动条件是否满足在步骤S13中计算出的上限摆动条件。在最佳摆动条件满足上限摆动条件时(是)，处理转移到步骤S16。另一方面，在最佳摆动条件不满足上限摆动条件时(否)，处理再次转移到步骤S14。

[0064] 在步骤S17中，摆动条件输出部450向数值控制装置200输出最佳摆动条件，之后结束处理。

[0065] 如以上说明那样，第一实施方式的机器学习装置400具有：设定条件取得部410，其取得用于摆动加工的设定条件；标签取得部420，其取得机床100的加工完成工件的面粗糙度数据作为标签；学习部430，其将设定条件与标签的组作为训练数据来进行监督学习，学习部430具有：学习模型，其学习使加工完成工件的面粗糙度最佳化的摆动条件。

[0066] 由此，与难以设定考虑了面粗糙度的摆动条件的现有技术相比，机器学习装置400能够学习考虑了面粗糙度而最佳化的最佳摆动条件。

[0067] 另外，机器学习装置400还具有：摆动条件输出部450，其根据学习模型输出使加工完成工件的面粗糙度最佳化的最佳摆动条件。由此，与难以设定考虑了面粗糙度的摆动条件的现有技术相比，机器学习装置400能够输出考虑了面粗糙度而最佳化的最佳摆动条件。另外，机器学习装置400能够使摆动条件的设定自动化，因此，能够减轻作业者的负担。

[0068] 另外，摆动条件输出部450具有：切屑切碎条件计算部451，其计算工件的切屑能够切碎的切屑切碎摆动条件，摆动条件输出部450根据学习模型，输出满足切屑切碎摆动条件的最佳摆动条件。由此，机器学习装置400能够设定能切碎切屑且考虑了面粗糙度而最佳化的最佳摆动条件。

[0069] 另外，摆动条件输出部450具有：摆动条件上限计算部452，其计算不超过预先设定

的上限值的上限摆动条件，摆动条件输出部450根据学习模型输出满足上限摆动条件的最佳摆动条件。由此，机器学习装置400能够设定能切碎切屑、且不超过摆动条件的上限、且考虑了面粗糙度而最佳化的最佳摆动条件。

[0070] [第二实施方式]

[0071] 接着，对第二实施方式的控制系统10进行说明。此外，在进行第二实施方式的说明时，对与第一实施方式相同的结构要件标注相同符号，省略或简化其说明。第二实施方式的控制系统10与第一实施方式的主要不同点在于，使用强化学习来代替监督学习，其他结构具有与第一实施方式一样的结构。

[0072] 图4是表示第二实施方式的控制系统10的概要的图。控制系统10的目的在于，通过使用机器学习计算摆动条件，计算能够在机床100中产生工具的空摆且能够实现良好的面粗糙度数据的摆动条件。

[0073] 控制系统1具有：机床100、数值控制装置200、加工面解析装置300以及机器学习装置500。在此，机床100和数值控制装置200为一对一的组，以能够通信的方式连接。此外，机床100和数值控制装置200可以经由连接接口直接连接，或者也可以经由LAN(Local Area Network)等网络连接。

[0074] 另外，数值控制装置200、加工面解析装置300以及机器学习装置500分别经由连接接口直接连接，或者分别经由网络连接，能够相互进行通信。此外，网络例如是在工厂内构筑的LAN、因特网、公共电话网、或者它们的组合。网络中的具体通信方式、是有线连接和无线连接中的哪一个等不做特别限定。

[0075] 机床100、数值控制装置200以及加工面解析装置300如上述那样具有与第一实施方式一样的结构。

[0076] 机器学习装置500是进行强化学习的装置。在进行机器学习装置500中包含的各功能块的说明之前，首先，对强化学习的基本结构进行说明。智能体(相当于本实施方式中的机器学习装置500)观测环境的状态，选择某个行为，环境根据该行为而变化。随着环境的变化，给予某种回报，智能体学习更好的行为的选择(决策)。

[0077] 相对于监督学习表示完全的正确解，强化学习中的回报大多是基于环境的部分变化的片段值。

[0078] 因此，智能体学习选择行为，使得将来的回报合计为最大。

[0079] 这样，在强化学习中，通过学习行为，根据行为对环境给予的相互作用来学习适当的行为，即学习用于使将来得到的回报为最大的学习方法。这表示在本实施方式中，例如能够获得选择用于缩短交接时间T，并且缩短开放时间T2的行为信息这样的、影响未来的行为。

[0080] 在此，作为强化学习可以使用任意的学习方法，在以下说明中，以在某种环境的状态s下，使用Q学习(Q-learning)的情况为例进行说明，所述Q学习是学习选择行为a的价值Q(s、a)的方法。

[0081] Q学习以在某种状态s时从能够取得的行为a中将价值Q(s、a)最高的行为a选择为最佳行为为目的。

[0082] 但是，在最初开始Q学习的时间点，对于状态s与行为a的组合来说，完全不知晓价值Q(s、a)的正确值。因此，智能体在某种状态s下选择各种行为a，针对当时的行为a，根据给

予的回报，选择更好的行为，由此，继续学习正确的价值Q(s、a)。

[0083] 另外，想要使将来获得的回报合计最大化，因此，目标是最终成为Q(s、a)＝E[Σ(γ^t)r_t]。在此，E[]表示期待值，t表示时刻、γ表示后述的称为折扣率的参数，r_t表示时刻t的回报，Σ是时刻t的合计。该式中的期待值是按最佳行为状态变化时的期望值。但是在Q学习的过程中，由于不知道最佳行为，因此通过进行各种行为，一边搜索一边进行强化学习。这样的价值Q(s、a)的更新式例如可以通过下式(1)来表示。

[0084] [数学式1]

[0085] $$Q(s_{t+1},a_{t+1}) \leftarrow Q(s_t,a_t)+\alpha(r_{t+1}+\gamma \max_a Q(s_{t+1},a)-Q(s_t,a_t)) \quad \cdots(1)$$

[0086] 在上述的式(1)中，S_t表示时刻t的环境状态，a_t表示时刻t的行为。通过行为a_t，状态变化为s_{t+1}。r_{t+1}表示通过该状态的变化而得的回报。另外，带有max的项是：在状态s_{t+1}下，将γ乘以选择出当时知道的Q值最高的行为a时的Q值而得的。在此，γ是0＜γ≤1的参数，称为折扣率。此外，α是学习系数，设α的范围为0＜α≤1。

[0087] 上述的式(1)表示如下方法：根据尝试a_t的结果而反馈回来的回报r_{t+1}，更新状态s_t下的行为a_t的价值Q(s_t、a_t)。该更新式表示了：若行为a_t导致的下一状态s_{t+1}下的最佳行为的价值$\max_a$ Q(s_{t+1}、a)比状态s_t下的行为a_t的价值Q(s_t、a_t)大，则增大价值Q(s_t、a_t)，反之如果小，则减小价值Q(s_t、a_t)。即，使某种状态下的某种行为的价值接近该行为导致的下一状态下的最佳行为的价值。其中，该差因折扣率γ和回报r_{t+1}的存在形式而变化，但基本上是某种状态下的最佳行为的价值传播至其前一个状态下的行为价值的结构。

[0088] 在此，Q学习存在如下方法：制作针对全部状态行为对(s、a)的价值Q(s、a)的表格，来进行学习。但是，有时为了求出全部状态行为对的价值Q(s、a)的值状态数会过多，使得Q学习收敛需要较多的时间。

[0089] 因此，可以利用公知的称为DQN(Deep Q-Network)的技术。具体而言，可以使用适当的神经网络来构成价值函数Q，调整神经网络的参数，由此，通过适当的神经网络来近似价值函数Q来计算价值Q(s、a)的值。通过利用DQN，能够缩短Q学习收敛所需的时间。此外，关于DQN，例如在以下的非专利文献中有详细的记载。

[0090] 〈非专利文献〉

[0091] “Human-level control through deep reinforcement learning”，Volodymyr Mnih1著[online]，[令和2年12月15日检索]，因特网〈URL:http://files.davidqiu.com/research/nature14236.pdf〉

[0092] 机器学习装置500进行上述说明的Q学习。具体而言，机器学习装置500学习如下的价值函数Q：将在数值控制装置200中设定的设定条件(加工条件、摆动条件等)设为状态s，将与该状态s有关的数值控制装置200中的摆动条件变更选择为行为a。

[0093] 机器学习装置500观测在数值控制装置200中设定的设定条件(加工条件、摆动条件等)等状态s，决定行为a。机器学习装置500每当进行行为a时返回回报。机器学习装置500试错性地搜索最佳的行为a，使得将来的回报合计为最大。由此，机器学习装置500能够针对在数值控制装置200中设定的设定条件(加工条件、摆动条件等)等即状态s，选择最佳的行为a。

[0094] 即，根据由机器学习装置500学习到的价值函数Q，针对某种状态s的设定条件(加

工条件、摆动条件等)，选择应用的行为a中的价值函数Q的值最大那样的行为a，由此，能够选择面粗糙度数据最小(最佳)那样的行为a。

[0095] 为了进行以上的强化学习，机器学习装置500具有：设定条件取得部510、判定信息取得部520、行为信息输出部530、学习部540、价值函数存储部550以及摆动条件输出部560。

[0096] 设定条件取得部510从数值控制装置200取得在数值控制装置200中设定的设定条件(加工条件、摆动条件等)作为状态信息(状态s)。该状态s相当于Q学习中的环境的状态s。

[0097] 具体而言，第二实施方式中的状态s表示在数值控制装置200中设定的设定条件(加工条件、摆动条件等)。设定条件具有包含机床100的主轴转速、进给速度等的加工条件和包含机床100的摆动振幅、摆动频率等的摆动条件。

[0098] 判定信息取得部520取得用于计算进行Q学习的回报的判定信息。具体而言，判定信息取得部520取得机床100的加工完成工件的面粗糙度数据作为用于计算进行Q学习的回报的判定信息。

[0099] 行为信息输出部530将学习部540生成的行为信息(行为a)发送给数值控制装置200。如上所述，数值控制装置200根据该行为a来变更当前的状态s，即当前设定的设定条件，由此，转变为下一个状态s’(即在机床100中执行了变更后的摆动条件时的状态)。

[0100] 学习部540学习在某种环境的状态s下选择某个行为a时的价值Q(s,a)。具体而言，学习部540具有：回报计算部541、价值函数更新部542以及行为信息生成部543。

[0101] 回报计算部541根据判定信息计算在某种状态s下选择了行为a时的回报。在此，回报计算部541在加工完成工件的面粗糙度数据小于规定阈值时，将回报设为正值，在面粗糙度数据超过规定阈值时，将回报设为负值。

[0102] 另外，回报计算部541根据理论面粗糙度的近似式来决定规定阈值。具体而言，回报计算部541将对理论面粗糙度的近似式乘以校正系数a而得的值决定为规定阈值。在此，规定阈值表示为$a \cdot F^2/8R$，F表示进给量[mm]，R表示工具半径。

[0103] 价值函数更新部542根据状态s、行为a、将行为a应用于状态s时的状态s’以及如上述那样计算出的回报值来进行Q学习，由此，更新价值函数存储部550存储的价值函数Q。

[0104] 价值函数Q的更新可以通过在线学习进行，也可以通过批量学习进行，还可以通过小批量学习进行。在线学习是指如下学习方法：通过将某个行为a应用于当前的状态s，每当状态s转变为新的状态s’时，立即进行价值函数Q的更新。

[0105] 另外，批量学习是指如下学习方法：反复进行将某个行为a应用于当前的状态s，状态s转变为新的状态s’，由此，收集学习用的数据，使用收集到的全部学习用数据来进行价值函数Q的更新。并且，小批量学习是指在线学习和批量学习的中间的、每当积攒某种程度的学习用数据时进行价值函数Q的更新的学习方法。

[0106] 行为信息生成部543在Q学习的过程中生成行为a，将生成的行为a输出到行为信息输出部530。具体而言，行为信息生成部543针对当前的状态s，选择Q学习的过程中的行为a。第二实施方式中的行为a包含针对当前的状态s应该如何变更摆动条件。

[0107] 并且，行为信息生成部543也可以采取如下策略：通过在当前的推定的行为a的价值中选择价值Q(s,a)最高的行为a’的贪婪算法、以某个小的概率ε随机地选择行为a’，除此以外选择价值Q(s,a)最高的行为a’的ε贪婪算法这样的公知的方法，来选择行为a’。

[0108] 价值函数存储部550是存储价值函数Q的存储装置。存储在价值函数存储部550中

的价值函数Q由价值函数更新部542更新。

[0109] 另外，机器学习装置500根据通过价值函数更新部542进行Q学习而更新的价值函数Q，生成作为用于使机床100进行价值Q(s，a)为最大的摆动加工的摆动条件的行为a(以下，称为“最佳化行为信息”)。

[0110] 摆动条件输出部560取得价值函数存储部550存储的价值函数Q。如上所述，该价值函数Q通过价值函数更新部542进行Q学习而进行了更新。然后，摆动条件输出部560根据价值函数Q，生成使加工完成工件的面粗糙度最佳化的最佳摆动条件作为最佳化行为信息，将生成的最佳摆动条件(最佳化行为信息)输出给数值控制装置200。

[0111] 另外，摆动条件输出部560具有切屑切碎条件计算部561和摆动条件上限计算部562。

[0112] 切屑切碎条件计算部561计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。然后，摆动条件输出部560输出满足切屑切碎摆动条件的最佳摆动条件。

[0113] 摆动条件上限计算部562计算不超过预先设定的上限值的上限摆动条件。然后，摆动条件输出部560输出满足上限摆动条件的最佳摆动条件。

[0114] 数值控制装置200根据该最佳摆动条件(最佳化行为信息)来修正当前设定的摆动条件，生成动作指令，由此，机床100能够以加工完成工件的面粗糙度成为最小(最佳)的方式进行动作。

[0115] 接着，参照图5的流程图对机器学习装置500的价值函数的更新处理进行说明。图5是表示第二实施方式的机器学习装置500进行的价值函数的更新处理流程的流程图。

[0116] 首先，在步骤S21中，设定条件取得部510从数值控制装置200取得作为状态信息的设定条件。取得的设定条件输出到价值函数更新部542和行为信息生成部543。如上所述，该设定条件(状态信息)是相当于Q学习中的环境的状态s的信息。

[0117] 在步骤S22中，行为信息生成部543生成作为新的行为信息的摆动条件，将生成的新的行为信息(行为a)经由行为信息输出部530输出给数值控制装置200。接收到行为信息的数值控制装置200通过根据接收到的行为信息变更了与当前的状态s有关的摆动条件后的状态s’，来驱动机床100，进行工件的摆动加工。如上所述，该行为信息相当于Q学习中的行为a。

[0118] 在步骤S23中，判定信息取得部520取得机床100的加工完成工件的面粗糙度数据作为用于计算进行Q学习的回报的判定信息。

[0119] 在步骤S24中，回报计算部531根据输入的判定信息(加工完成工件的面粗糙度)计算回报。回报计算部541判定加工完成工件的面粗糙度数据是否小于规定阈值。在面粗糙度数据小于规定阈值时(是)，处理转移到步骤S25。另一方面，在面粗糙度数据超过规定阈值时(否)，处理转移到步骤S26。

[0120] 在步骤S25中，回报计算部541将回报的值计算为正值。

[0121] 在步骤S26中，回报计算部541将回报的值计算为负值。

[0122] 在步骤S27中，价值函数更新部542根据上述计算出的回报的值，更新价值函数存储部550存储的价值函数Q。然后，学习部640再次返回到步骤S21，通过重复上述的处理，价值函数Q收敛为适当的值。此外，学习部640也可以将上述的处理重复了规定次数、重复了规定时间作为条件而结束。以上，对机器学习装置500的动作进行了说明，但例如计算回报的

值的处理是一例，不限于此。

[0123] 图6是表示第二实施方式的机器学习装置500进行的输出摆动条件的处理流程的流程图。

[0124] 在步骤S31中，摆动条件输出部560取得价值函数存储部550存储的价值函数Q。如上所述，该价值函数Q通过价值函数更新部542进行Q学习而进行了更新。

[0125] 在步骤S32中，摆动条件输出部560根据价值函数Q，例如在当前设定的状态s下，从能够取得的行为a中选择价值Q(s,a)最高的行为a作为最佳行为，由此，生成最佳摆动条件(最佳化行为信息)。

[0126] 在步骤S33中，切屑切碎条件计算部561计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。

[0127] 在步骤S34中，摆动条件上限计算部562计算不超过预先设定的上限值的上限摆动条件。

[0128] 在步骤S35中，摆动条件输出部560判定在步骤S32中生成的最佳摆动条件是否满足在步骤S33中计算出的切屑切碎摆动条件。在最佳摆动条件满足切屑切碎摆动条件时(是)，处理转移到步骤S35。另一方面，在最佳摆动条件不满足切屑切碎摆动条件时(否)，处理再次转移到步骤S31。

[0129] 在步骤S36中，摆动条件输出部560判定在步骤S32中输出的最佳摆动条件是否满足在步骤S34中计算出的上限摆动条件。在最佳摆动条件满足上限摆动条件时(是)，处理转移到步骤S37。另一方面，在最佳摆动条件不满足上限摆动条件时(否)，处理再次转移到步骤S31。

[0130] 在步骤S32中，摆动条件输出部560对数值控制装置200输出生成的最佳摆动条件(最佳化行为信息)。

[0131] 根据以上，数值控制装置200根据最佳摆动条件修正当前设定的状态s(即，当前设定的摆动条件)，生成动作指令。然后，数值控制装置200将生成的动作指令输送给机床100，由此，机床100能够以加工完成工件的面粗糙度数据成为最佳(最小)的方式进行动作。

[0132] 如以上说明的那样，第二实施方式的机器学习装置500具有：设定条件取得部510，其取得用于摆动加工的设定条件作为状态信息；判定信息取得部520，其取得机床100的加工完成工件的面粗糙度数据作为判定信息；行为信息输出部530，其输出表示针对当前的状态应当如何变更所述摆动条件的行为信息；回报计算部541，其根据判定信息，计算强化学习中的回报的值；价值函数更新部542，其根据状态信息、行为信息以及回报来更新价值函数，该价值函数用于确定机床100的摆动条件的价值；摆动条件输出部560，其根据价值函数，输出使加工完成工件的面粗糙度最佳化的最佳摆动条件。

[0133] 由此，与难以设定考虑了面粗糙度的摆动条件的现有技术相比，机器学习装置500能够输出考虑了面粗糙度而最佳化的最佳摆动条件。另外，机器学习装置500能够使摆动条件的设定自动化，因此，能够减轻作业者的负担。

[0134] 另外，回报计算部541在加工完成工件的面粗糙度数据小于规定阈值时，将回报设为正值，在面粗糙度数据超过规定阈值时，将回报设为负值。由此，机器学习装置500能够决定考虑了面粗糙度数据的回报值。

[0135] 另外，回报计算部541根据理论面粗糙度的近似式来决定规定阈值。由此，机器学

习装置500能够决定考虑了理论面粗糙度的回报值。

[0136] 另外，摆动条件输出部560具有：切屑切碎条件计算部561，其计算工件的切屑能够切碎的切屑切碎摆动条件，摆动条件输出部560输出满足切屑切碎摆动条件的最佳摆动条件。由此，机器学习装置500能够设定进行切屑切碎且考虑了面粗糙度而最佳化的最佳摆动条件。

[0137] 摆动条件输出部560具有：摆动条件上限计算部562，其计算不超过预先设定的上限值的上限摆动条件，摆动条件输出部560输出满足上限摆动条件的最佳摆动条件。由此，机器学习装置500能够设定进行切屑切碎且不超过摆动条件的上限且考虑了面粗糙度而最佳化的最佳摆动条件。

[0138] [第三实施方式]

[0139] 接着，对第三实施方式的控制系统20进行说明。此外，在进行第三实施方式的说明时，对与第一实施方式相同的结构要件标注相同符号，省略或简化其说明。第三实施方式的控制系统10与第一实施方式的主要不同点在于，使用加工完成工件的圆度代替加工完成工件的面粗糙度数据，其他结构具有与第一实施方式一样的结构。

[0140] 图7是表示第三实施方式的控制系统20的概要的图。控制系统20的目的在于，通过使用机器学习计算摆动条件，计算在机床100中产生工具的空摆且能够实现良好的圆度数据的摆动条件。

[0141] 控制系统20具有：机床100、数值控制装置200、加工面解析装置600以及机器学习装置700。在此，机床100和数值控制装置200为一对一的组，以能够通信的方式连接。此外，机床100和数值控制装置200可以经由连接接口直接连接，或者也可以经由LAN(Local Area Network)等网络连接。

[0142] 另外，数值控制装置200、加工面解析装置600以及机器学习装置700分别经由连接接口直接连接，或者分别经由网络连接，能够相互进行通信。此外，网络例如是在工厂内构筑的LAN、因特网、公共电话网、或者它们的组合。网络中的具体通信方式、是有线连接和无线连接中的哪一个等不做特别限定。

[0143] 如上所述，机床100和数值控制装置200具有与第一实施方式一样的结构。

[0144] 加工面解析装置600测定以及计算在机床100中加工后的加工完成工件的圆度数据，将计算出的面粗糙度数据输出给机器学习装置700。

[0145] 机器学习装置700使用从数值控制装置200受理到的设定条件以及从加工面解析装置600受理到的圆度数据进行机器学习。然后，机器学习装置700通过进行机器学习，构筑用于输出摆动条件的学习模型。机器学习装置700具有：设定条件取得部710、标签取得部720、学习部730、学习模型存储部740以及摆动条件输出部750。

[0146] 设定条件取得部710从数值控制装置200取得包含摆动加工用的加工条件以及摆动条件的设定条件。

[0147] 标签取得部720从加工面解析装置600取得机床100的加工完成工件的圆度数据作为标签。在此，标签是指在机器学习中应与输入对应的正确的输出。

[0148] 这样，作为学习对象的用于摆动加工的设定条件和圆度数据的标签成为组，输入到学习部430。该设定条件和标签的组相当于机器学习中的训练数据。在此，优选准备多个用于进行机器学习的训练数据。

[0149] 学习部730根据输入的训练数据进行机器学习，由此，构筑学习模型。即，学习部730学习使加工完成工件的圆度最佳化的摆动条件。学习部730构筑的学习模型输出到学习模型存储部740。

[0150] 学习模型存储部740存储由学习部730学习到的学习模型。学习模型存储部740存储的学习模型由摆动条件输出部750使用。此外，在构筑了学习模型之后，在新取得了训练数据时，学习部730也可以通过对学习模型存储部740存储的学习模型进一步进行监督学习，来更新曾经构筑的学习模型。

[0151] 另外，机器学习装置700可以与其他机器学习装置之间共享学习模型存储部740存储的学习模型。如果在多个机器学习装置中共享学习模型，则能够在各机器学习装置中分散地进一步进行监督学习，因此，能够提高监督学习的效率和精度。

[0152] 摆动条件输出部750根据学习模型存储部740存储的学习模型，输出使加工完成工件的圆度最佳化的最佳摆动条件。

[0153] 另外，摆动条件输出部750具有切屑切碎条件计算部751和摆动条件上限计算部752。

[0154] 切屑切碎条件计算部751计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。然后，摆动条件输出部750根据学习模型，输出满足切屑切碎摆动条件的最佳摆动条件。

[0155] 摆动条件上限计算部752计算不超过预先设定的上限值的上限摆动条件。然后，摆动条件输出部750根据学习模型，输出满足上限摆动条件的最佳摆动条件。

[0156] 接下来，参照图8的流程图对机器学习装置700的学习模型的构筑处理进行说明。

[0157] 图8是表示第三实施方式的机器学习装置700进行的学习模型的构筑处理流程的流程图。

[0158] 在步骤S41中，设定条件取得部710从数值控制装置200取得包含摆动加工用的加工条件以及摆动条件的设定条件。

[0159] 在步骤S42中，标签取得部720从加工面解析装置600取得机床100的加工完成工件的圆度数据作为标签。

[0160] 在步骤S43中，学习部730在被输入以设定条件和标签为组的训练数据时，根据输入的训练数据来执行机器学习。学习部730使用输入的训练数据来执行机器学习。学习部730例如通过回归分析、神经网络、最小二乘法、阶梯法等进行监督学习。

[0161] 在步骤S4中，学习部730判定是否结束监督学习。在此，使监督学习结束的条件可以任意决定。例如，学习部730可以在神经网络的输出与标签的误差值为规定值以下时，结束监督学习。另外，学习部730也可以在反复进行了预先决定的次数的监督学习时，结束监督学习。在结束监督学习时(是)，处理转移到步骤S45。另一方面，在未结束监督学习时(否)，处理转移到步骤S43。

[0162] 在步骤S45中，学习部730将通过到该时间点为止的监督学习而构筑的学习模型，输出并存储在学习模型存储部740中。另外，在取得了新的训练数据时，机器学习装置700能够对学习模型进行进一步的机器学习。之后，机器学习装置700结束本处理。

[0163] 接下来，参照图9的流程图对机器学习装置700的摆动条件的输出处理进行说明。

[0164] 图9是表示第三实施方式的机器学习装置700进行的摆动条件的输出处理流程的

流程图。

[0165] 在步骤S51中，设定条件取得部710从数值控制装置200取得包含摆动加工用的加工条件以及摆动条件的设定条件。

[0166] 在步骤S52中，切屑切碎条件计算部751计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。

[0167] 在步骤S53中，摆动条件上限计算部752计算不超过预先设定的上限值的上限摆动条件。

[0168] 在步骤S54中，摆动条件输出部750根据在步骤S51中取得的设定条件和学习模型存储部740存储的学习模型，输出使加工完成工件的圆度最佳化的最佳摆动条件。

[0169] 在步骤S55中，摆动条件输出部750判定在步骤S53中输出的最佳摆动条件是否满足在步骤S52中计算出的切屑切碎摆动条件。在最佳摆动条件满足切屑切碎摆动条件时(是)，处理转移到步骤S56。另一方面，在最佳摆动条件不满足切屑切碎摆动条件时(否)，处理再次转移到步骤S54。

[0170] 在步骤S56中，摆动条件输出部750判定在步骤S53中输出的最佳摆动条件是否满足在步骤S53中计算出的上限摆动条件。在最佳摆动条件满足上限摆动条件时(是)，处理转移到步骤S56。另一方面，在最佳摆动条件不满足上限摆动条件时(否)，处理再次转移到步骤S54。

[0171] 在步骤S57中，摆动条件输出部750向数值控制装置200输出最佳摆动条件，之后结束处理。

[0172] 如以上说明的那样，第三实施方式的机器学习装置700具有：设定条件取得部710，其取得用于摆动加工的设定条件；标签取得部720，其取得机床100的加工完成工件的圆度数据作为标签；学习部730，其将设定条件与标签的组作为训练数据来进行监督学习，学习部730具有：学习模型，其学习使加工完成工件的圆度最佳化的摆动条件。

[0173] 由此，与难以设定考虑了圆度的摆动条件的现有技术相比，机器学习装置700能够学习考虑圆度而最优化的最佳摆动条件。

[0174] 另外，机器学习装置700还具有：摆动条件输出部750，其根据学习模型，输出使加工完成工件的圆度最佳化的最佳摆动条件。由此，与难以设定考虑了圆度的摆动条件的现有技术相比，机器学习装置700能够输出考虑圆度而最佳化的最佳摆动条件。另外，机器学习装置700能够使摆动条件的设定自动化，因此，能够减轻作业者的负担。

[0175] 另外，摆动条件输出部750具有：切屑切碎条件计算部751，其计算工件的切屑能够切碎的切屑切碎摆动条件，摆动条件输出部750根据学习模型，输出满足切屑切碎摆动条件的最佳摆动条件。由此，机器学习装置700能够设定进行切屑切碎且考虑了面粗糙度而最佳化的最佳摆动条件。

[0176] 另外，摆动条件输出部750具有：摆动条件上限计算部752，其计算不超过预先设定的上限值的上限摆动条件，摆动条件输出部750根据学习模型，输出满足上限摆动条件的最佳摆动条件。由此，机器学习装置700能够设定进行切屑切碎且不超过摆动条件的上限且考虑了面粗糙度而最佳化的最佳摆动条件。

[0177] [第四实施方式]

[0178] 接着，对第四实施方式的控制系统30进行说明。此外，在进行第四实施方式的说明

时，对与第三实施方式相同的结构要件标注相同符号，省略或简化其说明。第四实施方式的控制系统30与第三实施方式的主要不同点在于，使用强化学习代替监督学习，其他结构具有与第三实施方式一样的结构。

[0179] 图10是表示第四实施方式的控制系统30的概要的图。控制系统30的目的在于，通过使用机器学习计算摆动条件，计算在机床100中产生工具的空摆且能够实现良好的圆度数据的摆动条件。

[0180] 控制系统30具有：机床100、数值控制装置200、加工面解析装置600以及机器学习装置800。在此，机床100和数值控制装置200为一对一的组，以能够通信的方式连接。此外，机床100和数值控制装置200可以经由连接接口直接连接，或者也可以经由LAN(Local Area Network)等网络连接。

[0181] 另外，数值控制装置200、加工面解析装置600以及机器学习装置800分别经由连接接口直接连接，或者分别经由网络连接，能够相互进行通信。此外，网络例如是在工厂内构筑的LAN、因特网、公共电话网、或者它们的组合。网络中的具体通信方式、是有线连接和无线连接中的哪一个等不做特别限定。

[0182] 机床100、数值控制装置200以及加工面解析装置600如上述那样具有与第三实施方式一样的结构。

[0183] 机器学习装置800是进行强化学习的装置。此外，与强化学习相关的处理与第二实施方式的机器学习装置500一样，因此，省略说明。为了进行强化学习，机器学习装置800具有：设定条件取得部810、判定信息取得部820、行为信息输出部830、学习部840、价值函数存储部850以及摆动条件输出部860。

[0184] 设定条件取得部810从数值控制装置200取得在数值控制装置200中设定的设定条件(加工条件、摆动条件等)作为状态信息(状态s)。该状态s相当于Q学习中的环境的状态s。

[0185] 具体而言，第四实施方式中的状态s表示在数值控制装置200中设定的设定条件(加工条件、摆动条件等)。设定条件具有：包含机床100的主轴转速、进给速度等的加工条件；包含机床100的摆动振幅、摆动频率等的摆动条件。

[0186] 判定信息取得部820取得用于计算进行Q学习的回报的判定信息。具体而言，判定信息取得部820取得机床100的加工完成工件的圆度数据，作为用于计算进行Q学习的回报的判定信息。

[0187] 行为信息输出部830将学习部840生成的行为信息(行为a)发送给数值控制装置200。如上所述，数值控制装置200根据该行为a来变更当前的状态s，即当前设定的设定条件，由此，转变为下一个状态s’(即在机床100中执行了变更后的摆动条件时的状态)。

[0188] 学习部840学习在某种环境的状态s下选择某个行为a时的价值Q(s,a)。具体而言，学习部840具有：回报计算部841、价值函数更新部842以及行为信息生成部843。

[0189] 回报计算部841根据判定信息计算在某种状态s下选择了行为a时的回报。在此，回报计算部841在加工完成工件的圆度数据小于规定阈值时，将回报设为正值，在圆度数据超过了规定阈值时，将回报设为负值。

[0190] 价值函数更新部842根据状态s、行为a、将行为a应用于状态s时的状态s’以及如上述那样计算出的回报的值来进行Q学习，由此，更新价值函数存储部850存储的价值函数Q。

[0191] 价值函数Q的更新可以通过在线学习进行，也可以通过批量学习进行，还可以通过

小批量学习进行。

[0192] 行为信息生成部843在Q学习的过程中生成行为a，将生成的行为a输出到行为信息输出部830。具体而言，行为信息生成部843针对当前的状态s，选择Q学习的过程中的行为a。第四实施方式中的行为a包含：针对当前的状态s应该如何变更摆动条件。

[0193] 并且，行为信息生成部843也可以采取如下策略：通过在当前的推定的行为a的价值中选择价值Q(s，a)最高的行为a'的贪婪算法、以某个小的概率ε随机地选择行为a'，除此以外选择价值Q(s，a)最高的行为a'的ε贪婪算法这样的公知的方法，来选择行为a'。

[0194] 价值函数存储部850是存储价值函数Q的存储装置。存储在价值函数存储部850中的价值函数Q由价值函数更新部842更新。

[0195] 另外，机器学习装置800根据通过价值函数更新部842进行Q学习而更新的价值函数Q，生成作为用于使机床100进行价值Q(s，a)为最大的摆动加工的摆动条件的行为a(以下，称为“最佳化行为信息”)。

[0196] 摆动条件输出部860取得价值函数存储部850存储的价值函数Q。如上所述，该价值函数Q通过价值函数更新部842进行Q学习而进行了更新。然后，摆动条件输出部860根据价值函数Q，生成使加工完成工件的面粗糙度最佳化的最佳摆动条件作为最佳化行为信息，将生成的最佳摆动条件(最佳化行为信息)输出给数值控制装置200。

[0197] 另外，摆动条件输出部860具有切屑切碎条件计算部861和摆动条件上限计算部862。

[0198] 切屑切碎条件计算部861计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。然后，摆动条件输出部860输出满足切屑切碎摆动条件的最佳摆动条件。

[0199] 摆动条件上限计算部862计算不超过预先设定的上限值的上限摆动条件。然后，摆动条件输出部860输出满足上限摆动条件的最佳摆动条件。

[0200] 数值控制装置200根据该最佳摆动条件(最佳化行为信息)修正当前设定的摆动条件，生成动作指令，由此，机床100能够以加工完成工件的圆度为最佳的方式进行动作。

[0201] 接着，参照图11的流程图对机器学习装置800的价值函数的更新处理进行说明。图11是表示第四实施方式的机器学习装置800进行的价值函数的更新处理流程的流程图。

[0202] 首先，在步骤S61中，设定条件取得部810从数值控制装置200取得作为状态信息的设定条件。取得的设定条件输出到价值函数更新部842和行为信息生成部843。如上所述，该设定条件(状态信息)是相当于Q学习中的环境的状态s的信息。

[0203] 在步骤S62中，行为信息生成部843生成作为新的行为信息的摆动条件，将生成的新的行为信息(行为a)经由行为信息输出部830输出给数值控制装置200。接收到行为信息的数值控制装置200通过根据接收到的行为信息变更了与当前的状态s有关的摆动条件的状态s'来驱动机床100，进行工件的摆动加工。如上所述，该行为信息相当于Q学习中的行为a。

[0204] 在步骤S63中，判定信息取得部820取得机床100的加工完成工件的圆度数据，作为用于计算进行Q学习的回报的判定信息。

[0205] 在步骤S64中，回报计算部831根据输入的判定信息(加工完成工件的圆度数据)计算回报。回报计算部841判定加工完成工件的圆度数据是否小于规定阈值。在圆度数据小于规定阈值时(是)，处理转移到步骤S65。另一方面，在圆度数据超过规定阈值时(否)，处理转

移到步骤S66。

[0206] 在步骤S65中，回报计算部841将回报的值计算为正值。

[0207] 在步骤S66中，回报计算部841将回报的值计算为负值。

[0208] 在步骤S67中，价值函数更新部842根据上述计算出的回报值，更新价值函数存储部850存储的价值函数Q。然后，学习部840再次返回到步骤S61，通过重复上述的处理，价值函数Q收敛为适当的值。此外，学习部840也可以将上述的处理重复了规定次数、重复了规定时间作为条件而结束。以上，对机器学习装置800的动作进行了说明，但例如计算回报的值的处理是一例，不限于此。

[0209] 图12是表示第四实施方式的机器学习装置800进行的输出摆动条件的处理流程的流程图。

[0210] 在步骤S71中，摆动条件输出部860取得价值函数存储部850存储的价值函数Q。如上所述，该价值函数Q通过价值函数更新部842进行Q学习而进行了更新。

[0211] 在步骤S72中，摆动条件输出部860根据价值函数Q，例如在当前设定的状态s下，从能够取得的行为a中选择价值Q(s,a)最高的行为a作为最佳行为，由此，生成最佳摆动条件(最佳化行为信息)。

[0212] 在步骤S73中，切屑切碎条件计算部861计算通过机床100的摆动加工，工件的切屑能够切碎的切屑切碎摆动条件。

[0213] 在步骤S74中，摆动条件上限计算部862计算不超过预先设定的上限值的上限摆动条件。

[0214] 在步骤S75中，摆动条件输出部860判定在步骤S72中生成的最佳摆动条件是否满足在步骤S73中计算出的切屑切碎摆动条件。在最佳摆动条件满足切屑切碎摆动条件时(是)，处理转移到步骤S75。另一方面，在最佳摆动条件不满足切屑切碎摆动条件时(否)，处理再次转移到步骤S71。

[0215] 在步骤S76中，摆动条件输出部860判定在步骤S72中输出的最佳摆动条件是否满足在步骤S74中计算出的上限摆动条件。在最佳摆动条件满足上限摆动条件时(是)，处理转移到步骤S77。另一方面，在最佳摆动条件不满足上限摆动条件时(否)，处理再次转移到步骤S71。

[0216] 在步骤S72中，摆动条件输出部860对数值控制装置200输出生成的最佳摆动条件(最佳化行为信息)。

[0217] 根据以上，数值控制装置200根据最佳摆动条件修正当前设定的状态s(即，当前设定的摆动条件)，生成动作指令。然后，数值控制装置200将生成的动作指令输送给机床100，由此，机床100能够以加工完成工件的圆度数据最佳的方式进行动作。

[0218] 如以上说明的那样，第四实施方式的机器学习装置800具有：设定条件取得部810，其取得用于摆动加工的设定条件作为状态信息；判定信息取得部820，其取得机床100的加工完成工件的圆度数据作为判定信息；行为信息输出部830，其输出表示针对当前的状态应当如何变更所述摆动条件的行为信息；回报计算部841，其根据判定信息，计算强化学习中的回报的值；价值函数更新部842，其根据状态信息、行为信息以及回报来更新价值函数，该价值函数用于确定机床100的摆动条件的价值；摆动条件输出部860，其根据价值函数，输出使加工完成工件的圆度最佳化的最佳摆动条件。

[0219] 由此，与难以设定考虑了面粗糙度的摆动条件的现有技术相比，机器学习装置800能够输出考虑圆度而最佳化的最佳摆动条件。另外，机器学习装置800能够使摆动条件的设定自动化，因此，能够减轻作业者的负担。

[0220] 另外，回报计算部841在加工完成工件的圆度数据小于规定阈值时，将回报设为正值，在圆度数据超过规定阈值时，将回报设为负值。由此，机器学习装置800能够决定考虑了面粗糙度数据的回报的值。

[0221] 另外，摆动条件输出部860具有：切屑切碎条件计算部861，其计算工件的切屑能够切碎的切屑切碎摆动条件，摆动条件输出部860输出满足切屑切碎摆动条件的最佳摆动条件。由此，机器学习装置800能够设定进行切屑切碎且考虑了面粗糙度而最佳化的最佳摆动条件。

[0222] 摆动条件输出部860具有：摆动条件上限计算部862，其计算不超过预先设定的上限值的上限摆动条件，摆动条件输出部860输出满足上限摆动条件的最佳摆动条件。由此，机器学习装置800能够设定进行切屑切碎且不超过摆动条件的上限且考虑了面粗糙度而最佳化的最佳摆动条件。

[0223] [其他实施方式]

[0224] 在上述的实施方式中，机器学习装置400、500、700、800使用了加工完成工件的面粗糙度数据以及圆度数据，但其他实施方式的机器学习装置也可以使用其他数据作为加工完成工件的评价数据。例如，在其他实施方式的机器学习装置中，加工完成工件的评价数据也可以包含面粗糙度、圆度或尺寸精度。在此，尺寸精度表示是否按照由加工程序指令的形状那样的尺寸进行了加工。

[0225] 即，在使用监督学习时，其他实施方式的机器学习装置学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件，该机器学习装置具有：设定条件取得部，其取得用于摆动加工的设定条件；标签取得部，其取得机床的加工完成工件的评价数据作为标签；学习部，其将设定条件与标签的组作为训练数据来进行监督学习，学习部具有：学习模型，其学习使加工完成工件的评价数据最佳化的摆动条件。另外，机器学习装置还可以具有：摆动条件输出部，其根据学习模型，输出使加工完成工件的评价数据最佳化的最佳摆动条件。

[0226] 由此，与难以设定考虑了加工完成工件的评价数据的摆动条件的现有技术相比，机器学习装置能够输出考虑加工完成工件的评价数据而最佳化的最佳摆动条件。另外，机器学习装置能够使摆动条件的设定自动化，因此，能够减轻作业者的负担。

[0227] 另外，在使用强化学习时，其他实施方式的机器学习装置学习一边使工具与工件相对地摆动一边进行摆动加工的机床的摆动条件，该机器学习装置具有：设定条件取得部，其取得用于摆动加工的设定条件作为状态信息；判定信息取得部，其取得机床的加工完成工件的评价数据作为判定信息；行为信息输出部，其输出表示针对当前的状态应当如何变更摆动条件的行为信息；回报计算部，其根据判定信息，计算强化学习中回报的值；价值函数更新部，其根据状态信息、行为信息以及回报来更新价值函数，该价值函数用于确定机床的摆动条件的价值；摆动条件输出部，其根据价值函数，输出使加工完成工件的评价数据最佳化的最佳摆动条件。另外，回报计算部在加工完成工件的评价数据小于规定阈值时，将回报设为正值，在评价数据超过规定阈值时，将回报设为负值。

[0228] 由此，与难以设定考虑了加工完成工件的评价数据的摆动条件的现有技术相比，机器学习装置能够输出考虑加工完成工件的评价数据而最佳化的最佳摆动条件。另外，机器学习装置能够使摆动条件的设定自动化，因此，能够减轻作业者的负担。

[0229] 另外，第一实施方式至第四实施方式的设定条件包含：加工条件，其包含机床100中的工具的进给速度、主轴转速、坐标值、工具的刀尖、工具的种类以及工件的材料中的至少一个；摆动条件，其包含机床100的摆动频率和摆动振幅中的至少一个。由此，机器学习装置400、500、700、800能够取得包含适当的加工条件以及摆动条件的设定条件。

[0230] 第一实施方式和第二实施方式的面粗糙度数据包含算术平均粗糙度、最大高度、最大峰高、最大谷深、平均高度、最大截面高度和负载长度比中的至少一个。由此，机器学习装置400、500能够通过适当的方法取得面粗糙度数据。

[0231] 另外，机器学习装置400、500、700、800可以由多个数值控制装置200共享。由此，机器学习装置400、500、700、800能够通过多个数值控制装置200共享以及更新学习模型、价值函数等。

[0232] 另外，机器学习装置400、500、700、800可以设置在云服务器上。由此，机器学习装置400、500能够在云服务器上共享以及更新学习模型、价值函数等。

[0233] 上述的数值控制装置200以及机器学习装置400、500、700、800具有CPU(Central Processing Unit)等运算处理装置。另外，数值控制装置200以及机器学习装置400、500、700、800还具有：储存了应用软件、OS(Operating System)等各种控制用程序的HDD(Hard Disk Drive)等辅助存储装置、用于储存运算处理装置执行程序时暂时需要的数据的RAM(Random Access Memory)这样的主存储装置。

[0234] 然后，关于数值控制装置200以及机器学习装置400、500、700、800，运算处理装置从辅助存储装置读入应用软件、OS，一边在主存储装置中展开读入的应用软件、OS，一边进行基于这些应用软件、OS的运算处理。另外，数值控制装置200以及机器学习装置400、500、700、800根据该运算结果来控制各装置具有的各种硬件。由此，实现本实施方式的功能块。即，上述的实施方式能够通过硬件与软件协作来实现。

[0235] 作为具体例，数值控制装置200能够通过在一般的机床100的控制装置中组装用于实现实施方式的应用软件来实现。另外，机器学习装置400、500、700、800能够通过在一般的个人计算机中组装用于实现本实施方式的应用软件来实现。

[0236] 其中，关于机器学习装置400、500、700、800，由于伴随机器学习的运算量增多，因此例如利用在个人计算机搭载GPU(Graphics Processing Units)，称为GPGPU(General-Purpose computing on Graphics Processing Units)的技术，在将GPU用于伴随机器学习的运算处理时可以进行高速处理。并且，为了进行更高速的处理，机器学习装置400、500可以使用多台搭载了这样的GPU的计算机来构筑计算机集群，通过该计算机集群所包含的多个计算机来进行并行处理。

[0237] 以上，对本发明的实施方式进行了说明，上述控制系统1、10、20、30可以通过硬件、软件或它们的组合来实现。另外，由上述的控制系统1、10、20、30进行的控制方法也可以通过硬件、软件或它们的组合来实现。在此，通过软件实现是指通过计算机读入并执行程序来实现。

[0238] 程序能够使用各种类型的非暂时性的计算机可读介质(non-transitory

computer readable medium)来储存,并提供给计算机。非暂时性的计算机可读介质包含各种类型的有实体的记录介质(tangible storage medium)。非暂时性的计算机可读介质的例子包含磁记录介质(例如,硬盘驱动器)、磁光记录介质(例如,磁光盘)、CD-ROM(Read Only Memory)、CD-R、CD-R/W和半导体存储器(例如,掩模ROM、PROM(Programmable ROM)、EPROM(Erasable PROM)、闪存ROM和RAM(random access memory))。

[0239] 另外,上述的各实施方式是本发明的优选的实施方式,但本发明的范围并不仅限定于上述各实施方式。在不脱离本发明精神的范围内能够以实施了各种变更的方式进行实施。

[0240] 符号说明

[0241] 1 控制系统

[0242] 10 控制系统

[0243] 100 机床

[0244] 200 数值控制装置

[0245] 300 加工面解析装置

[0246] 400 机器学习装置

[0247] 410 设定条件取得部

[0248] 420 标签取得部

[0249] 430 学习部

[0250] 440 学习模型存储部

[0251] 450 摆动条件输出部

[0252] 500 机器学习装置

[0253] 510 设定条件取得部

[0254] 520 判定信息取得部

[0255] 530 行为信息输出部

[0256] 540 学习部

[0257] 550 价值函数存储部

[0258] 560摆动条件输出部。

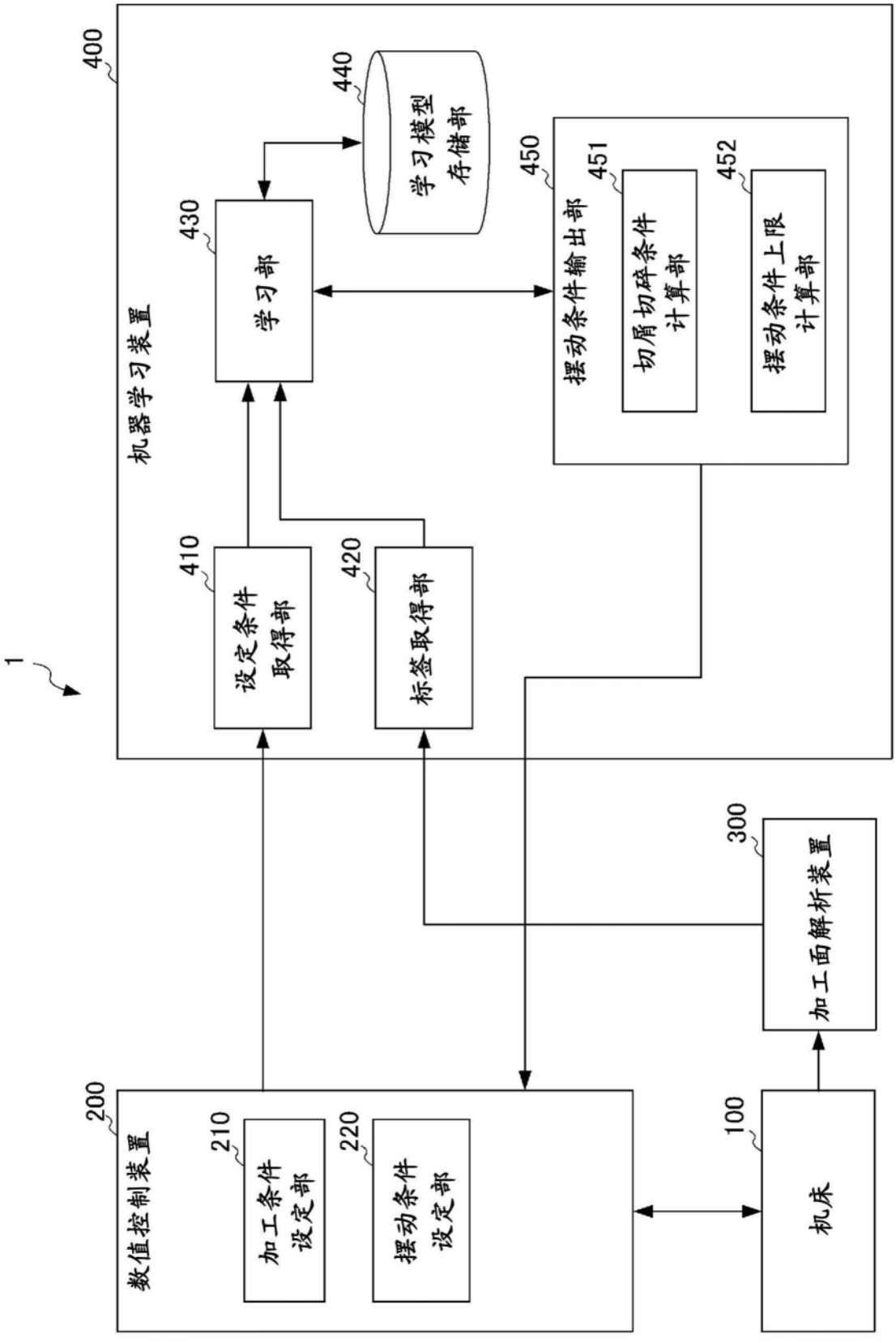

1
400
机器学习装置
440
学习模型存储部
430
学习部
450
摆动条件输出部
451
切屑切碎条件计算部
452
摆动条件上限计算部
410
设定条件取得部
420
标签取得部
300
加工面解析装置
200
数值控制装置
210
加工条件设定部
220
摆动条件设定部
100
机床

图1

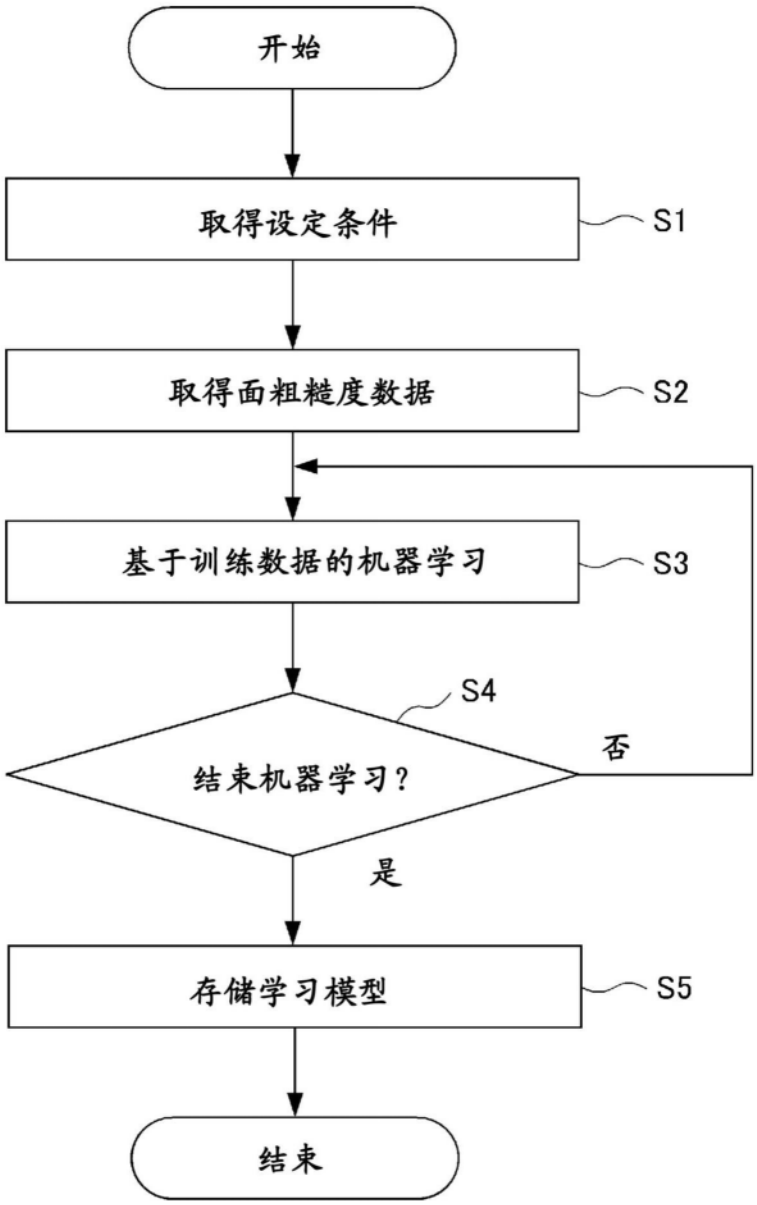

开始
取得设定条件
S1
取得面粗糙度数据
S2
基于训练数据的机器学习
S3
S4
结束机器学习？
否
是
存储学习模型
S5
结束

图2

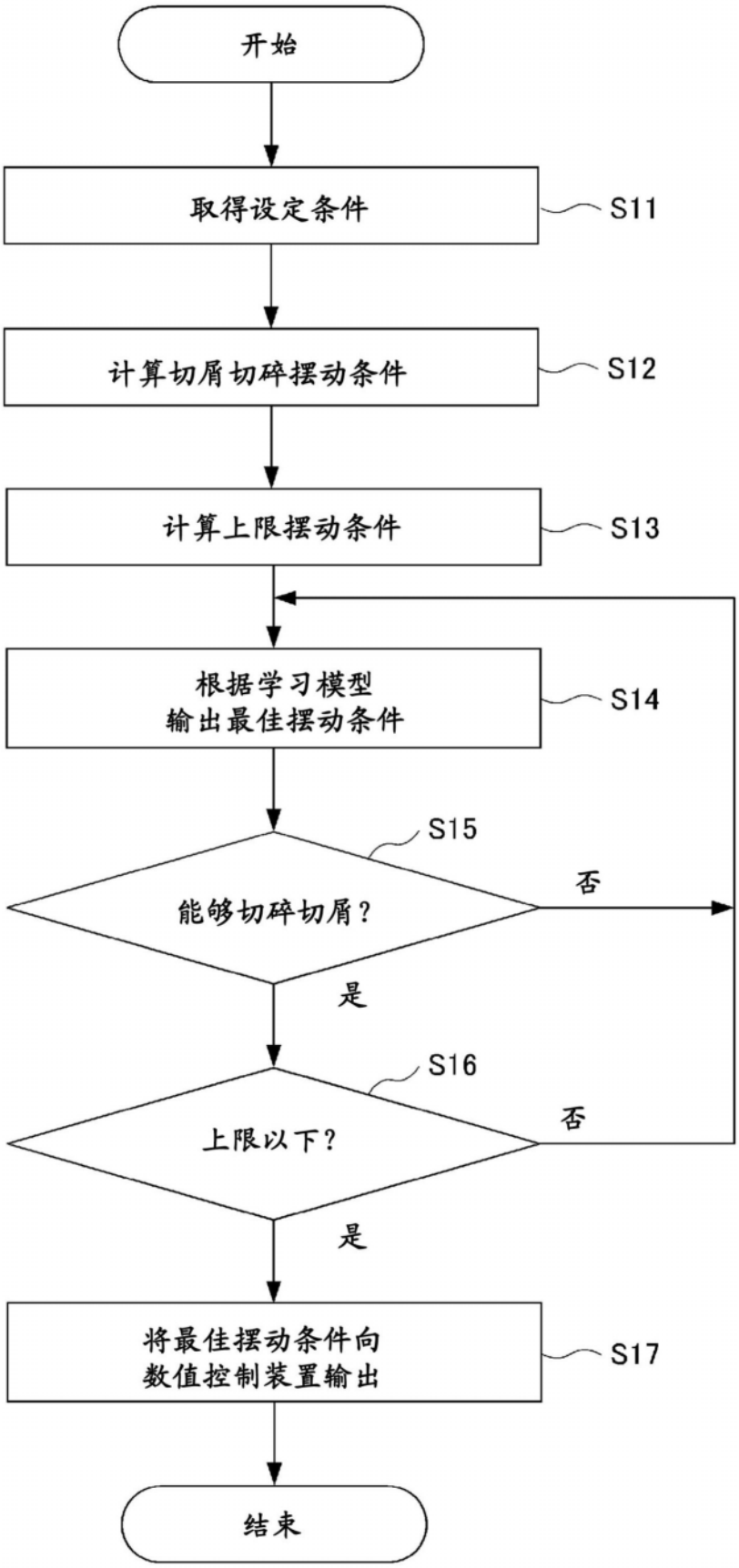

开始
取得设定条件
S11
计算切屑切碎摆动条件
S12
计算上限摆动条件
S13
根据学习模型
输出最佳摆动条件
S14
S15
能够切碎切屑？
否
是
S16
上限以下？
否
是
将最佳摆动条件向
数值控制装置输出
S17
结束

图3

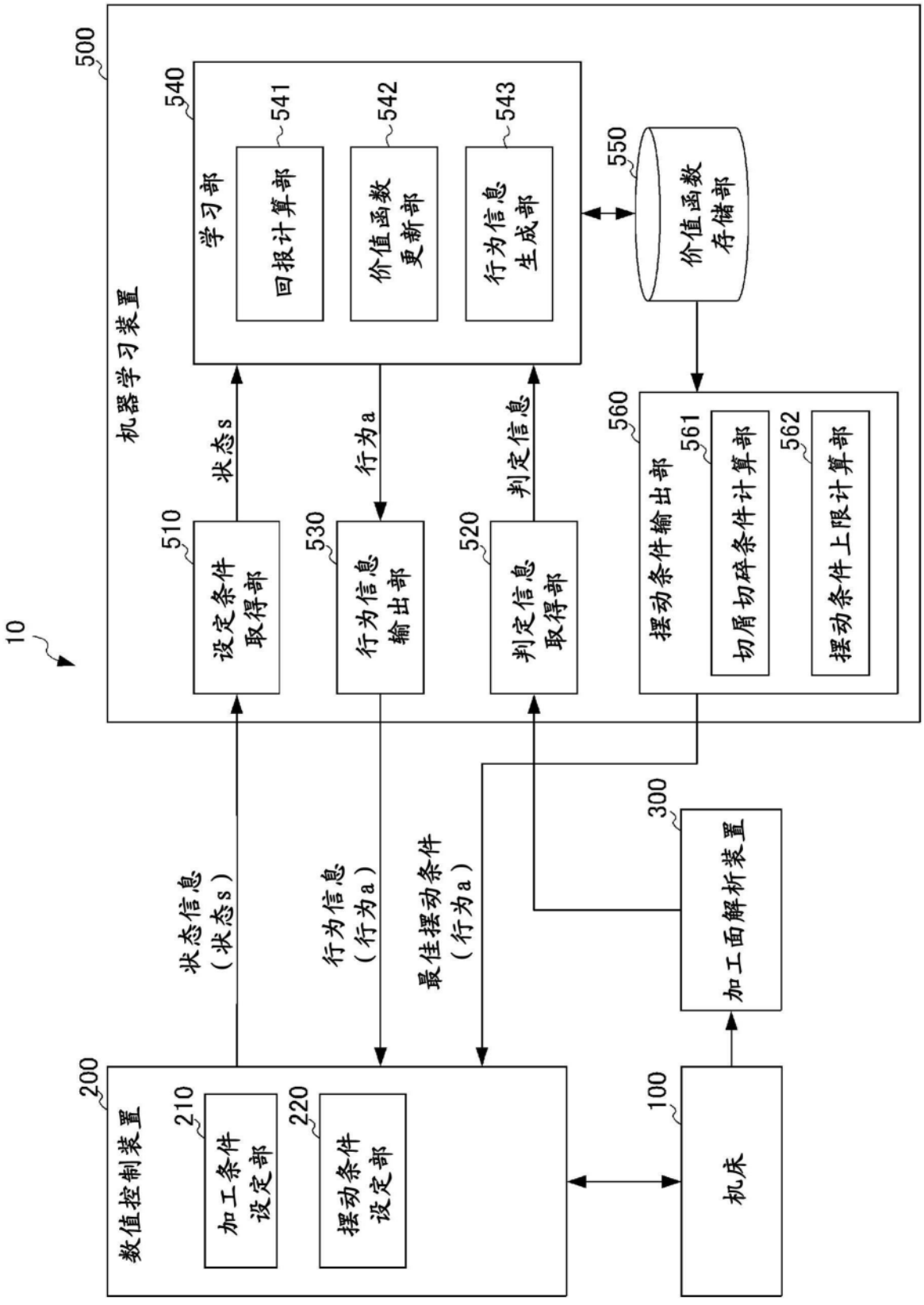

10
500
机器学习装置
540
学习部
541
回报计算部
542
价值函数更新部
543
行为信息生成部
550
价值函数存储部
560
摆动条件输出部
561
切屑切碎条件计算部
562
摆动条件上限计算部
510
设定条件取得部
530
行为信息输出部
520
判定信息取得部
状态s
行为a
判定信息
状态信息（状态s）
行为信息（行为a）
最佳摆动条件（行为a）
300
加工面解析装置
200
数值控制装置
210
加工条件设定部
220
摆动条件设定部
100
机床

图4

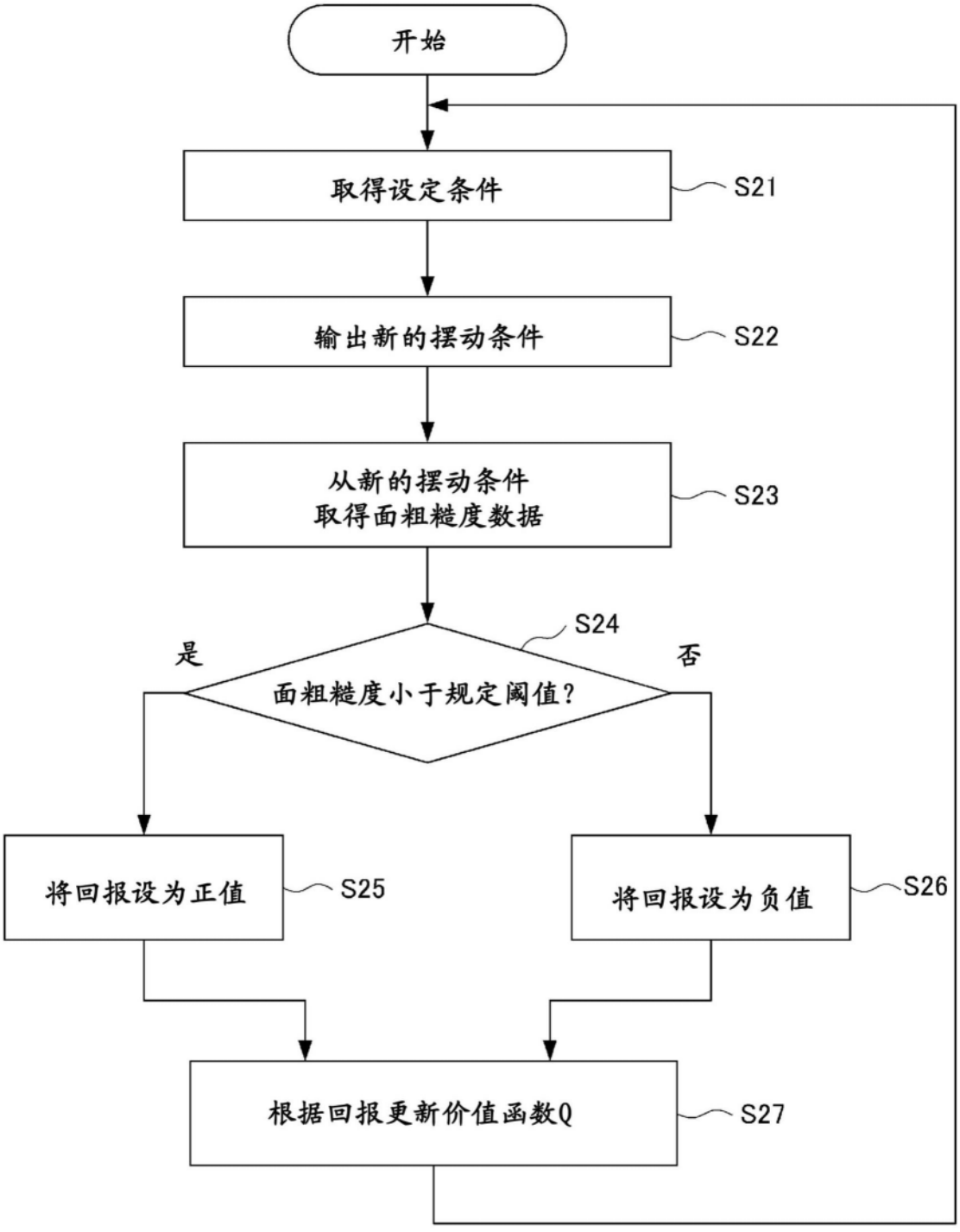

开始
取得设定条件
S21
输出新的摆动条件
S22
从新的摆动条件
取得面粗糙度数据
S23
S24
是
否
面粗糙度小于规定阈值?
将回报设为正值
S25
将回报设为负值
S26
根据回报更新价值函数Q
S27

图5

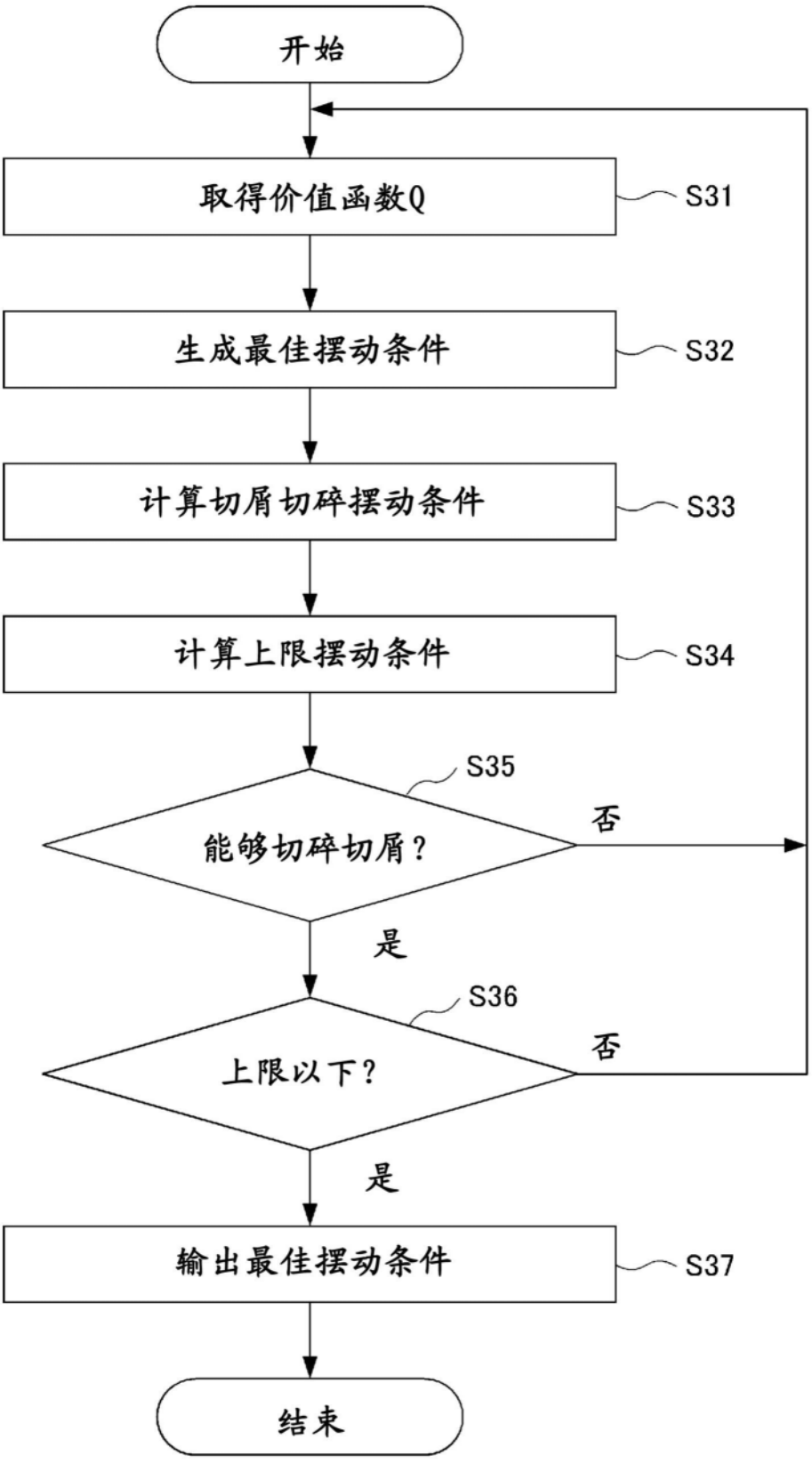

开始
取得价值函数Q
S31
生成最佳摆动条件
S32
计算切屑切碎摆动条件
S33
计算上限摆动条件
S34
S35
能够切碎切屑？
否
是
S36
上限以下？
否
是
输出最佳摆动条件
S37
结束

图6

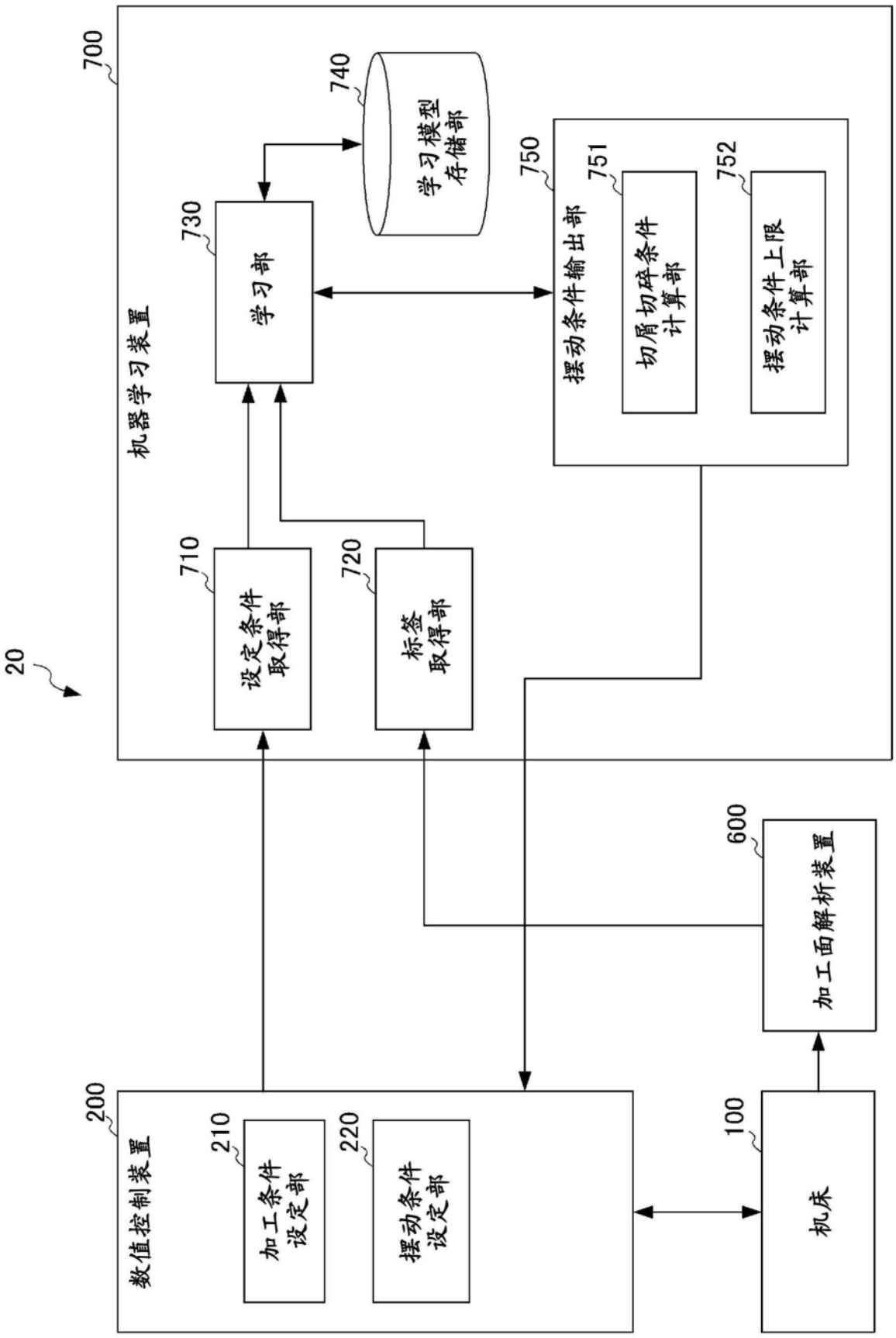

20
700
机器学习装置
740
学习模型存储部
730
学习部
750
摆动条件输出部
751
切屑切碎条件计算部
752
摆动条件上限计算部
710
设定条件取得部
720
标签取得部
600
加工面解析装置
200
数值控制装置
210
加工条件设定部
220
摆动条件设定部
100
机床

图7

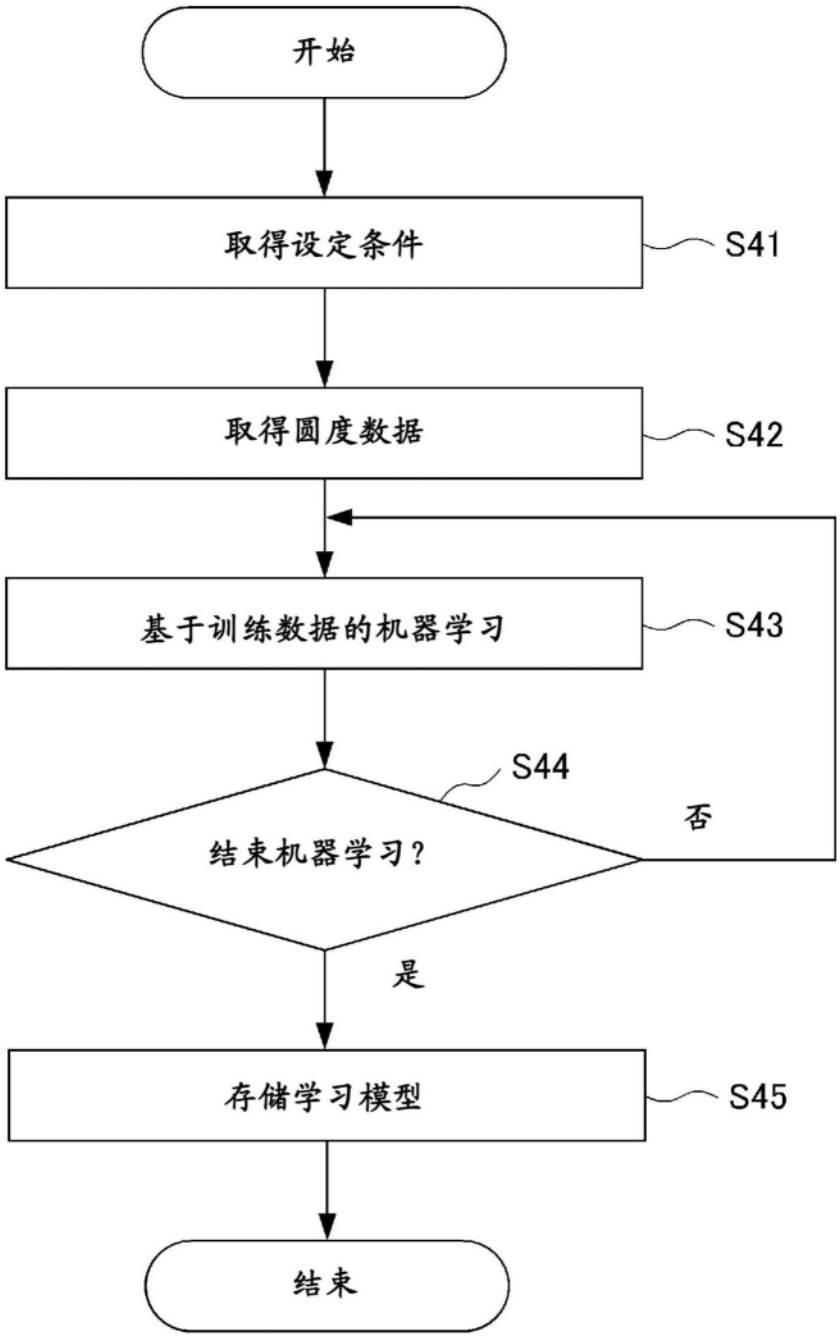

开始
取得设定条件
S41
取得圆度数据
S42
基于训练数据的机器学习
S43
S44
结束机器学习？
否
是
存储学习模型
S45
结束

图8

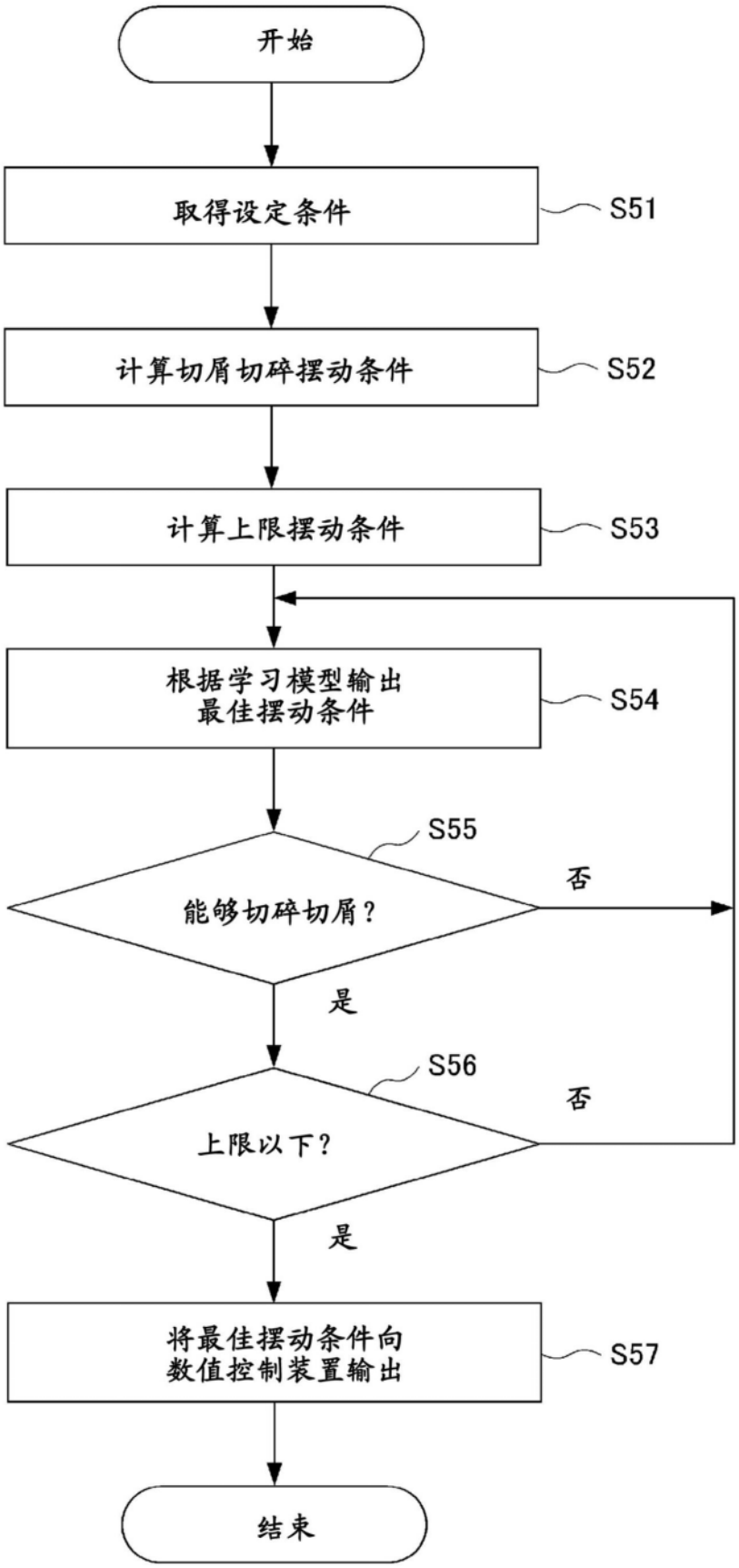

开始
取得设定条件
S51
计算切屑切碎摆动条件
S52
计算上限摆动条件
S53
根据学习模型输出
最佳摆动条件
S54
S55
能够切碎切屑？
否
是
S56
上限以下？
否
是
将最佳摆动条件向
数值控制装置输出
S57
结束

图9

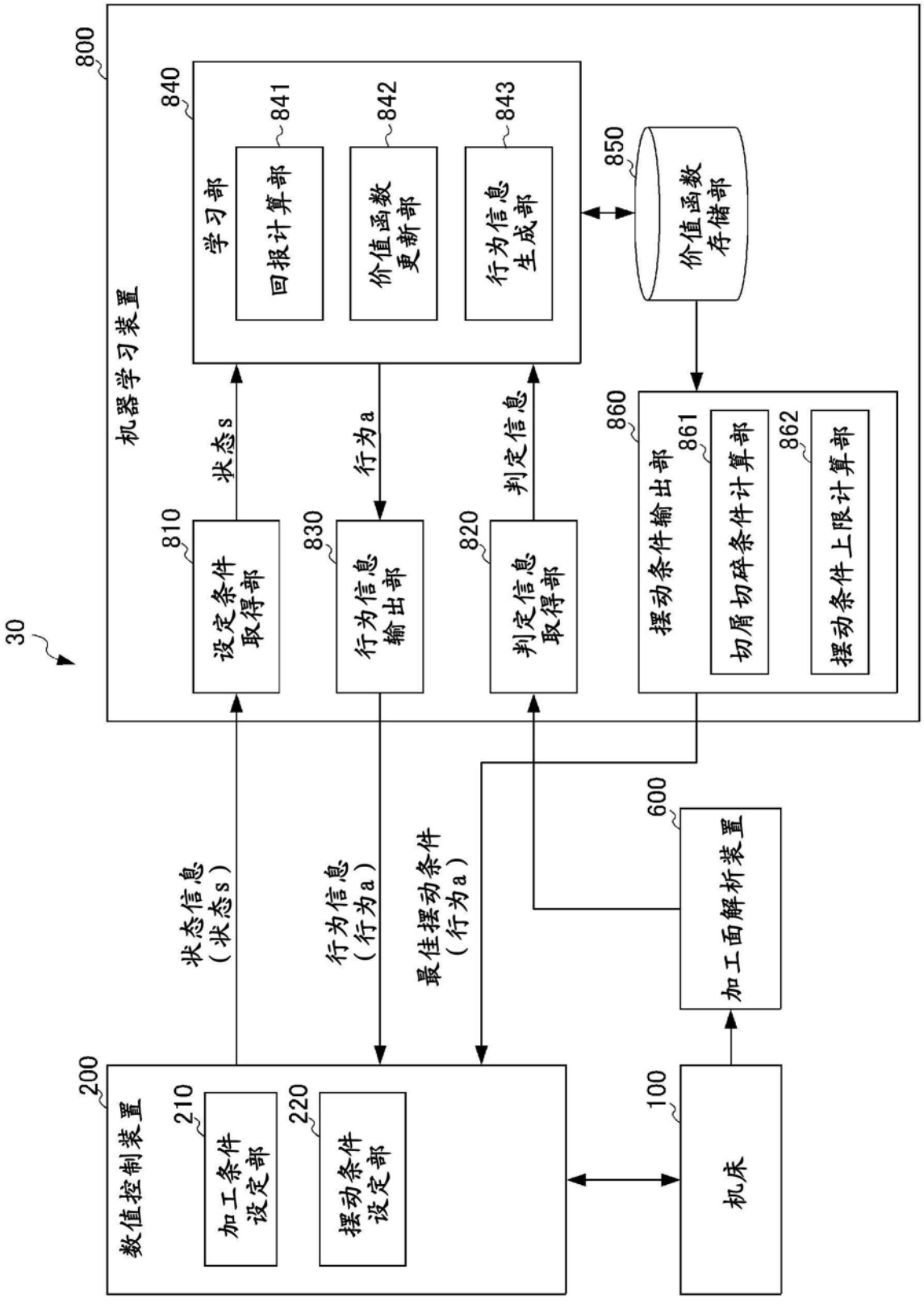

30
800
机器学习装置
840
学习部
841
回报计算部
842
价值函数更新部
843
行为信息生成部
850
价值函数存储部
860
摆动条件输出部
861
切屑切碎条件计算部
862
摆动条件上限计算部
810
设定条件取得部
830
行为信息输出部
820
判定信息取得部
状态s
行为a
判定信息
600
加工面解析装置
状态信息（状态s）
行为信息（行为a）
最佳摆动条件（行为a）
200
数值控制装置
210
加工条件设定部
220
摆动条件设定部
100
机床

图10

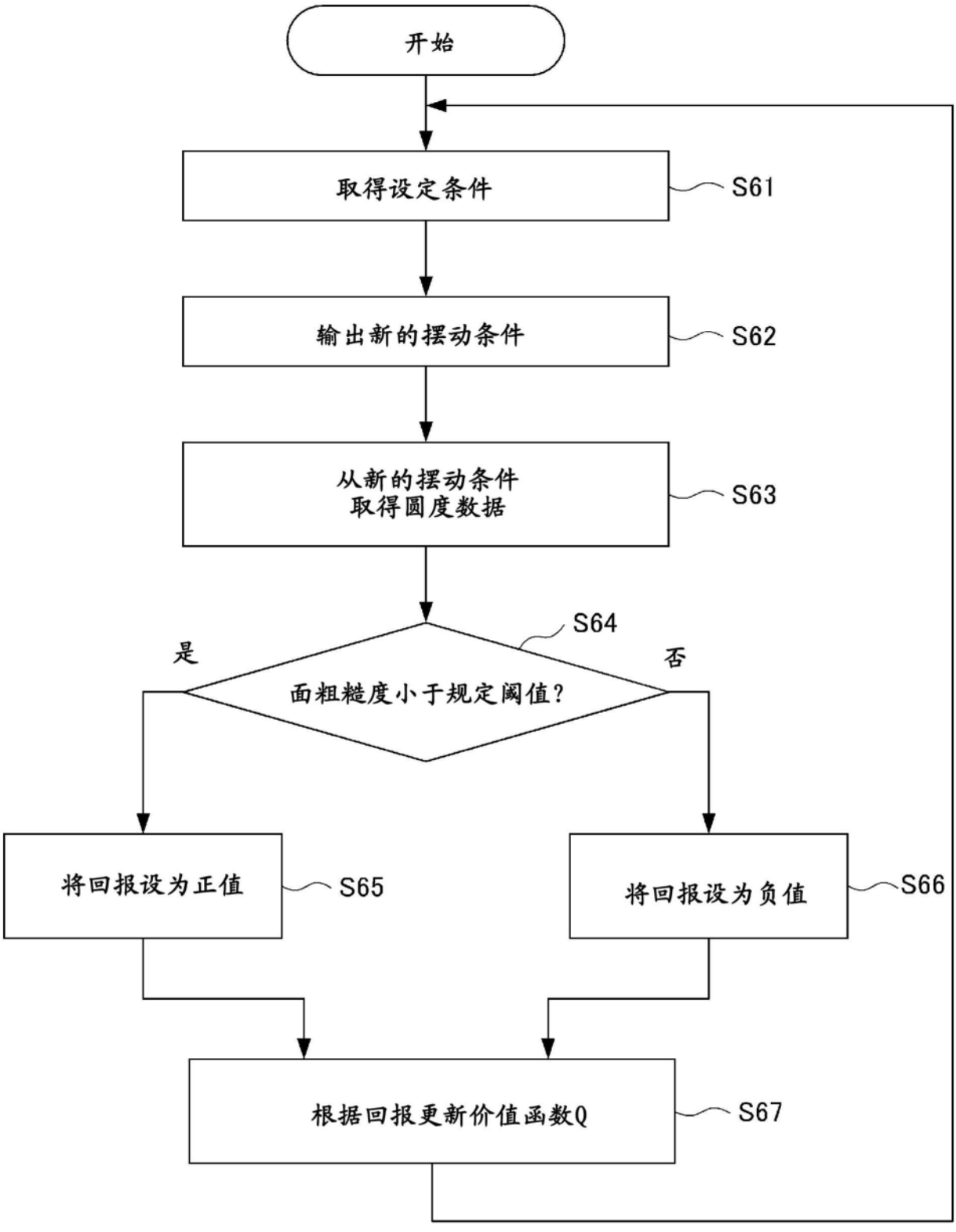

开始
取得设定条件
S61
输出新的摆动条件
S62
从新的摆动条件
取得圆度数据
S63
S64
是
否
面粗糙度小于规定阈值?
将回报设为正值
S65
将回报设为负值
S66
根据回报更新价值函数Q
S67

图11

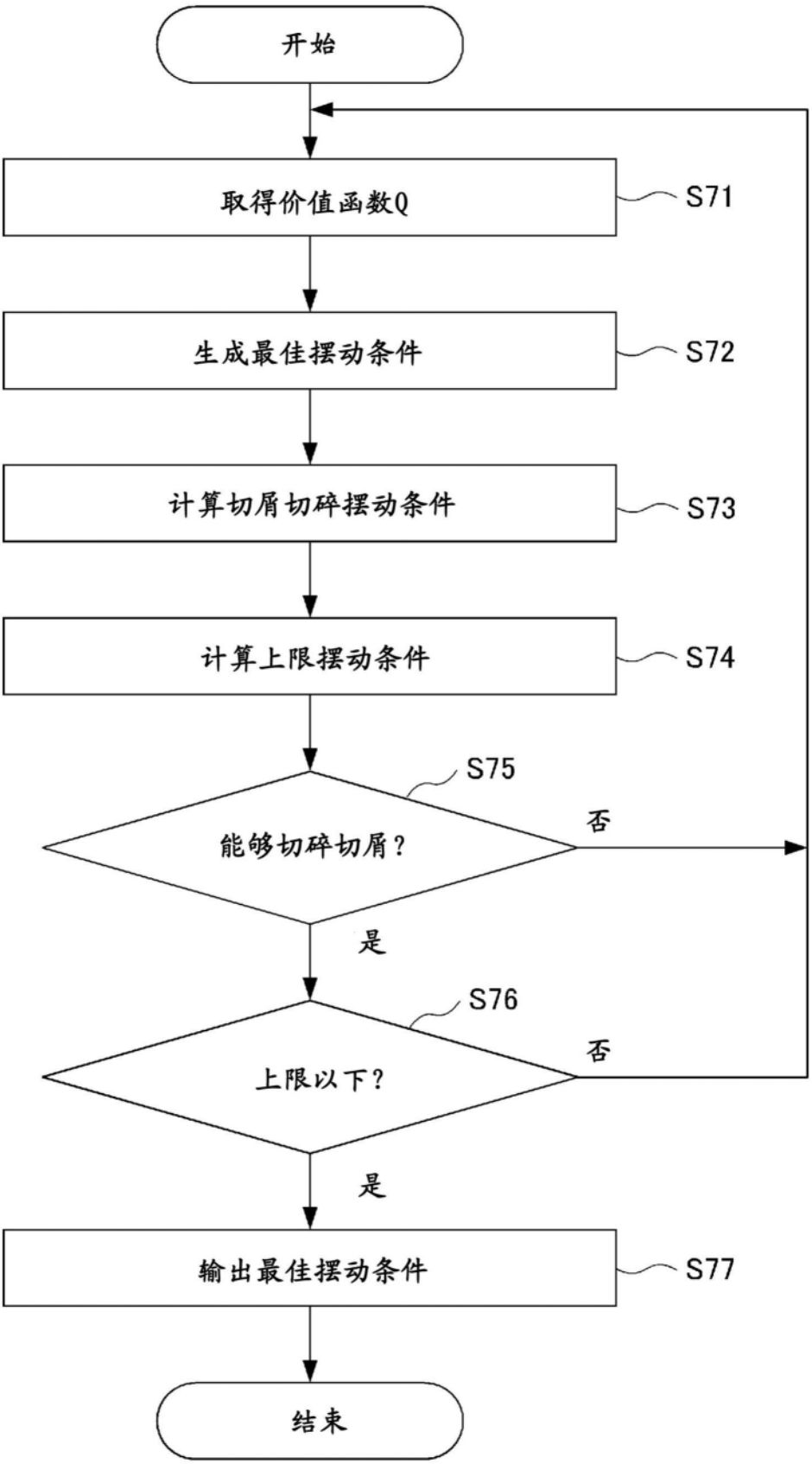

开始
取得价值函数Q
S71
生成最佳摆动条件
S72
计算切屑切碎摆动条件
S73
计算上限摆动条件
S74
S75
能够切碎切屑？
否
是
S76
上限以下？
否
是
输出最佳摆动条件
S77
结束

图12