

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年7月21日(21.07.2022)



(10) 国際公開番号

WO 2022/153936 A1

- (51) 国際特許分類:
G05B 19/4093 (2006.01) G06N 20/00 (2019.01)
G05B 19/4155 (2006.01)
- (21) 国際出願番号: PCT/JP2022/000336
- (22) 国際出願日: 2022年1月7日(07.01.2022)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2021-004401 2021年1月14日(14.01.2021) JP
- (71) 出願人: ファナック株式会社 (FANUC CORPORATION) [JP/JP]; 〒4010597 山梨県南都留郡忍野村忍草字古馬場 3 5 8 0 番地 Yamanashi (JP).
- (72) 発明者: 堀川 祐太郎 (HORIKAWA Yuutarou); 〒4010597 山梨県南都留郡忍野村忍草字古馬場

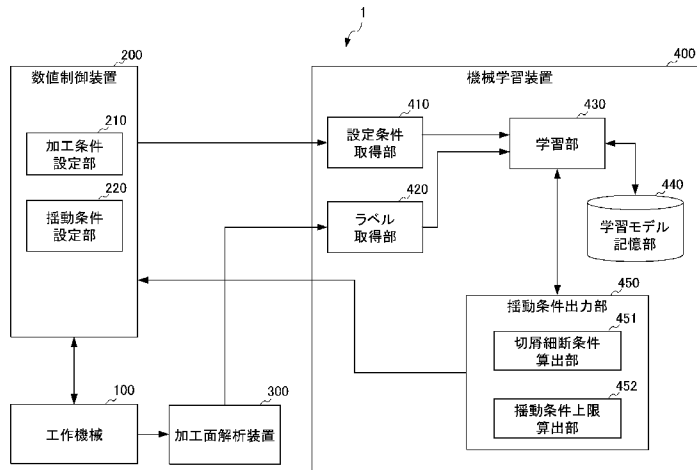
3 5 8 0 番地 ファナック株式会社内 Yamanashi (JP). 安田 将司(YASUDA Masashi); 〒4010597 山梨県南都留郡忍野村忍草字古馬場 3 5 8 0 番地 ファナック株式会社内 Yamanashi (JP).

(74) 代理人: 正林 真之, 外(SHOBAYASHI Masayuki et al.); 〒1000005 東京都千代田区丸の内 1 - 7 - 1 2 サピアタワー Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: MACHINE LEARNING DEVICE

(54) 発明の名称: 機械学習装置



100 Machine tool
200 Numerical control device
210 Machining conditions setting unit
220 Orbiting conditions setting unit
300 Machined surface analyzing device
400 Machine learning device
410 Set conditions acquisition unit
420 Label acquisition unit
430 Learning unit
440 Learning model storage unit
450 Orbiting conditions output unit
451 Chipping/shredding conditions calculation unit
452 Orbiting conditions upper limit calculation unit

(57) Abstract: Provided is a machine learning device capable of causing idling of a tool, and of calculating orbiting conditions that realize favorable evaluation data of finished workpieces. The machine learning device that learns orbiting conditions of a machine tool that performs orbiting machining in which a tool and a workpiece are orbited relative to each other is provided with a set conditions acquisition unit that acquires set conditions for the orbiting machining, a label acquisition unit that acquires evaluation data of finished workpieces by the machine tool as a label, and a learning unit that performs supervised learning using a set of the set conditions and the label as training data, the learning unit being provided with a learning model for training orbiting conditions that optimize the evaluation data of the finished workpieces.

ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告 (条約第21条(3))

(57) 要約: 工具の空振りが生じ、かつ良好な加工済みワークの評価データを実現できる揺動条件を算出することができる機械学習装置を提供すること。工具とワークと相対的に揺動させながら揺動加工する工作機械の揺動条件を学習する機械学習装置は、前記揺動加工のための設定条件を取得する設定条件取得部と、前記工作機械による加工済みワークの評価データを、ラベルとして取得するラベル取得部と、前記設定条件と前記ラベルとの組みを教師データとして教師あり学習を行う学習部と、を備え、前記学習部は、前記加工済みワークの評価データを最適化する揺動条件を学習する学習モデルを備える。

明 細 書

発明の名称：機械学習装置

技術分野

[0001] 本発明は、機械学習装置に関する。

背景技術

[0002] 従来、ワークと切削工具を加工方向に送りながら切削を行い、同時にワークと切削工具とを相対的に往復振動させる揺動切削が知られている（例えば、特許文献1参照）。揺動切削加工では、切屑を細断するために、工具がワークに対して一定周期で空振り（エアカット）するように周波数及び振幅等の揺動条件を設定する必要がある。

[0003] そして、オペレーターは、設定された揺動条件の波形を確認することによって、前回のパスと現在のパスとの公差が発生しているか、つまり、切屑切断が発生しているかを確認する。

先行技術文献

特許文献

[0004] 特許文献1：特開2019-28597号公報

発明の概要

発明が解決しようとする課題

[0005] 例えば、揺動切削により加工される加工済みワークの面粗さ、真円度、寸法精度等のような、加工済みワークの仕上がりに関する評価データは、揺動切削を適用しないときと比べて悪化することが多い。これは、揺動切削の振動及び切削量の変化等が原因となっているが、評価データが悪化することのメカニズムは非常に複雑である。したがって、工具の空振りが生じ、かつ良好な加工済みワークの評価データを実現できる揺動条件を算出することが求められている。

課題を解決するための手段

[0006] 本開示の一態様に係る機械学習装置は、工具とワークと相対的に揺動させ

ながら揺動加工する工作機械の揺動条件を学習する機械学習装置であって、前記揺動加工のための設定条件を取得する設定条件取得部と、前記工作機械による加工済みワークの評価データを、ラベルとして取得するラベル取得部と、前記設定条件と前記ラベルとの組みを教師データとして教師あり学習を行う学習部と、を備え、前記学習部は、前記加工済みワークの評価データを最適化する揺動条件を学習する学習モデルを備える。

[0007] 本開示の一態様に係る機械学習装置は、工具とワークと相対的に揺動させながら揺動加工する工作機械の揺動条件を学習する機械学習装置であって、前記揺動加工のための設定条件を、状態情報として取得する設定条件取得部と、前記工作機械による加工済みワークの評価データを、判定情報として取得する判定情報取得部と、現在の状態に対して前記揺動条件をどのように変更すべきかを示す行動情報を出力する行動情報出力部と、前記判定情報に基づいて、強化学習における報酬の値を算出する報酬算出部と、前記状態情報、前記行動情報及び前記報酬に基づいて、前記工作機械の揺動条件の価値を定める価値関数を更新する価値関数更新部と、前記価値関数に基づいて、前記加工済みワークの評価データを最適化する最適揺動条件を出力する揺動条件出力部と、を備える。

発明の効果

[0008] 本発明によれば、工具の空振りが生じ、かつ良好な加工済みワークの評価データを実現できる揺動条件を算出することができる。

図面の簡単な説明

[0009] [図1]第1実施形態に係る制御システムの概要を示す図である。

[図2]第1実施形態に係る機械学習装置による学習モデルの構築処理の流れを示すフローチャートである。

[図3]第1実施形態に係る機械学習装置による揺動条件の出力処理の流れを示すフローチャートである。

[図4]第2実施形態に係る制御システムの概要を示す図である。

[図5]第2実施形態に係る機械学習装置による価値関数の更新処理の流れを示す

すフローチャートである。

[図6]第2実施形態に係る機械学習装置による揺動条件を出力する処理の流れを示すフローチャートである。

[図7]第3実施形態に係る制御システムの概要を示す図である。

[図8]第3実施形態に係る機械学習装置による学習モデルの構築処理の流れを示すフローチャートである。

[図9]第3実施形態に係る機械学習装置による揺動条件の出力処理の流れを示すフローチャートである。

[図10]第4実施形態に係る制御システムの概要を示す図である。

[図11]第4実施形態に係る機械学習装置による価値関数の更新処理の流れを示すフローチャートである。

[図12]第4実施形態に係る機械学習装置による揺動条件を出力する処理の流れを示すフローチャートである。

発明を実施するための形態

[0010] [第1実施形態]

以下、本発明の実施形態の一例について説明する。図1は、第1実施形態に係る制御システム1の概要を示す図である。制御システム1は、機械学習を用いて揺動条件を算出することによって、工作機械100において工具の空振りが生じ、かつ良好な加工済みワークの評価データを実現できる揺動条件を算出することを目的とする。特に、第1実施形態に係る制御システム1は、機械学習を用いて揺動条件を算出することによって、工作機械100において工具の空振りが生じ、かつ良好な面粗さデータを実現できる揺動条件を算出することを目的とする。

[0011] 制御システム1は、工作機械100と、数値制御装置200と、加工面解析装置300と、機械学習装置400と、を備える。ここで、工作機械100と数値制御装置200とは、1対1の組とされて、通信可能に接続されている。なお、工作機械100と数値制御装置200とは、接続インタフェースを介して直接接続されてもよく、又は、LAN (Local Area

Network)等のネットワークを介して接続されてもよい。

[0012] また、数値制御装置200、加工面解析装置300及び機械学習装置400は、それぞれ接続インタフェースを介して直接に接続、又は、それぞれネットワークを介して接続されており、相互に通信を行うことが可能である。なお、ネットワークは、例えば、工場内に構築されたLANや、インターネット、公衆電話網、或いは、これらの組み合わせである。ネットワークにおける具体的な通信方式や、有線接続及び無線接続のいずれであるか等については、特に限定されない。

[0013] 工作機械100は、数値制御装置200の制御に従い、揺動切削を用いてワークを加工する。工作機械100は、工具、主軸、送り軸等の揺動切削加工を行うための一般的な構成を有する。

[0014] 数値制御装置200は、工作機械100を制御するための装置である。数値制御装置200は、加工条件設定部210と、揺動条件設定部220と、を備える。

[0015] 加工条件設定部210は、工作機械100において揺動加工を行うための加工条件及び揺動条件を含む設定条件を設定する。加工条件設定部210は、設定した設定条件を機械学習装置400へ出力する。ここで、設定条件は、工作機械100における工具の送り速度、主軸回転速度、座標値、工具の刃先、工具の種類及びワークの材料の少なくとも1つを含む加工条件と、工作機械100の揺動周波数及び揺動振幅の少なくとも1つを含む揺動条件と、を有する。

[0016] 揺動条件設定部220は、機械学習装置400から出力された揺動条件を数値制御装置200において設定し、揺動条件を工作機械100へ出力する。

[0017] 加工面解析装置300は、工作機械100において加工された加工済みワークの面粗さデータを測定及び算出し、算出した面粗さデータを機械学習装置400へ出力する。面粗さデータは、例えば、算術平均粗さ、最大高さ、最大山高さ、最大谷深さ、平均高さ、最大断面高さ及び負荷長さ率の少なく

とも1つを含む。

- [0018] なお、揺動切削加工では、周期的に工具が振動するため、ワーク（例えば、円筒形状）の様々な箇所では表面粗さが変わる。そこで、面粗さデータの測定時には、ワークの円筒側面で複数回、線粗さを測定する必要がある。また、面粗さデータの算出は、円筒側面で複数箇所の線粗さを測定し、平均の面粗さを算出する方法（算術平均粗さ R_a ）、及び円筒側面で複数箇所の線粗さを測定し、最大の値となる線粗さを算出する方法（最大高さ R_z ）等のような方法を用いることができる。
- [0019] 機械学習装置400は、数値制御装置200から受け付けた設定条件及び加工面解析装置300から受け付けた面粗さデータを用いて機械学習を行う。そして、機械学習装置400は、機械学習を行うことにより、揺動条件を出力するための学習モデルを構築する。機械学習装置400は、設定条件取得部410と、ラベル取得部420と、学習部430と、学習モデル記憶部440と、揺動条件出力部450と、を備える。
- [0020] 設定条件取得部410は、揺動加工のための加工条件及び揺動条件を含む設定条件を数値制御装置200から取得する。
- [0021] ラベル取得部420は、工作機械100による加工済みワークの面粗さデータを、加工面解析装置300から、ラベルとして取得する。ここで、ラベルとは、機械学習において入力に対応すべき正しい出力である。
- [0022] このようにして、学習対象とした揺動加工のための設定条件と、面粗さデータのラベルとが組となり、学習部430に入力される。この設定条件とラベルの組は、機械学習における教師データに相当する。ここで、機械学習を行うための教師データは、多数用意されることが望ましい。
- [0023] 学習部430は、入力された教師データに基づいて機械学習を行うことにより、学習モデルを構築する。すなわち、学習部430は、加工済みワークの面粗さを最適化する揺動条件を学習する。学習部430が構築した学習モデルは、学習モデル記憶部440に対して出力される。
- [0024] 学習モデル記憶部440は、学習部430によって学習された学習モデル

を記憶する。学習モデル記憶部440が記憶した学習モデルは、揺動条件出力部450により利用される。なお、学習モデルを構築した後に、新たに教師データを取得した場合には、学習部430は、学習モデル記憶部440が記憶した学習モデルに対して更に教師あり学習を行うことにより、一度構築した学習モデルを更新するようにしてもよい。

[0025] また、機械学習装置400は、学習モデル記憶部440が記憶した学習モデルを、他の機械学習装置との間で共有するようにしてもよい。学習モデルを複数の機械学習装置で共有するようにすれば、各機械学習装置にて分散して、更に教師あり学習を行うことが可能になるので、教師あり学習の効率及び精度を向上させることができる。

[0026] 揺動条件出力部450は、学習モデル記憶部440が記憶した学習モデルに基づいて、加工済みワークの面粗さを最適化する最適揺動条件を出力する。

また、揺動条件出力部450は、切屑細断条件算出部451と、揺動条件上限算出部452と、を有する。

[0027] 切屑細断条件算出部451は、工作機械100の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。そして、揺動条件出力部450は、学習モデルに基づいて、切屑細断揺動条件を満たす最適揺動条件を出力する。ここで、切屑細断揺動条件の算出は、予め切屑細断可能な揺動条件の組み合わせをテーブルデータとして保持してもよく、又は揺動条件に基づいて計算される加工経路において、前回のパスと今回のパスとの間で重なりが発生するような各揺動条件の関係式を算出してもよい。

[0028] 揺動条件上限算出部452は、予め設定された上限値を超えない上限揺動条件を算出する。そして、揺動条件出力部450は、学習モデルに基づいて、上限揺動条件を満たす最適揺動条件を出力する。ここで、予め設定された上限値は、例えば、揺動加工における周波数又は振幅の上限値、若しくは揺動加工における速度又は加速度の上限値を示す。上限揺動条件を満たすかを判別することは、例えば、各種上限値をパラメータに保持する又は加工プロ

グラムで指定することや、速度及び加速度を揺動条件（周波数及び振幅）から算出することによって実現できる。

[0029] 次に、図2のフローチャートを参照して機械学習装置400による学習モデルの構築処理について説明をする。

図2は、第1実施形態に係る機械学習装置400による学習モデルの構築処理の流れを示すフローチャートである。

[0030] ステップS1において、設定条件取得部410は、揺動加工のための加工条件及び揺動条件を含む設定条件を数値制御装置200から取得する。

[0031] ステップS2において、ラベル取得部420は、工作機械100による加工済みワークの面粗さデータを、加工面解析装置300から、ラベルとして取得する。

[0032] ステップS3において、学習部430は、設定条件とラベルとを組にした教師データが入力されると、入力された教師データに基づいて機械学習を実行する。学習部430は、入力された教師データを用いて機械学習を実行する。学習部430は、例えば、回帰分析、ニューラルネットワーク、最小二乗法、ステップワイズ法等により、教師あり学習を行う。

[0033] 教師あり学習は、オンライン学習で行ってもよく、バッチ学習で行ってもよく、ミニバッチ学習で行ってもよい。オンライン学習とは、音声データから特徴量を抽出し、ラベルが入力されて教師データが作成される都度、即座に教師あり学習を行うという学習方法である。また、バッチ学習とは、音声データから特徴量を抽出し、ラベルが入力されて、教師データが作成されることが繰り返される間に、繰り返しに応じた複数の教師データを収集し、収集した全ての教師データを用いて、教師あり学習を行うという学習方法である。

[0034] 更に、ミニバッチ学習とは、オンライン学習と、バッチ学習の中間的な、ある程度教師データが溜まるたびに教師あり学習を行うという学習方法である。

[0035] ステップS4において、学習部430は、教師あり学習を終了するか否か

を判定する。ここで、教師あり学習を終了させる条件は、任意に定めることができる。例えば、学習部430は、ニューラルネットワークの出力とラベルとの誤差の値が所定値以下となった場合に、教師あり学習を終了させるようにしてもよい。また、学習部430は、予め定めておいた回数だけ教師あり学習を繰り返した場合に、教師あり学習を終了させるようにしてもよい。教師あり学習を終了する場合（YES）、処理は、ステップS5へ移る。一方、教師あり学習を終了しない場合（NO）、処理は、ステップS3へ移る。

[0036] ステップS5において、学習部430は、その時点までの教師あり学習により構築した学習モデルを、学習モデル記憶部440に出力して記憶させる。また、新たな教師データを取得した場合に、機械学習装置400は、学習モデルに対して更なる機械学習を行うことができる。その後、機械学習装置400は、本処理を終了する。

[0037] 次に、図3のフローチャートを参照して機械学習装置400による揺動条件の出力処理について説明をする。

図3は、第1実施形態に係る機械学習装置400による揺動条件の出力処理の流れを示すフローチャートである。

[0038] ステップS11において、設定条件取得部410は、揺動加工のための加工条件及び揺動条件を含む設定条件を数値制御装置200から取得する。

[0039] ステップS12において、切屑細断条件算出部451は、工作機械100の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。

ステップS13において、揺動条件上限算出部452は、予め設定された上限値を超えない上限揺動条件を算出する。

[0040] ステップS14において、揺動条件出力部450は、ステップS11において取得された設定条件及び学習モデル記憶部440が記憶した学習モデルに基づいて、加工済みワークの面粗さを最適化する最適揺動条件を出力する。

- [0041] ステップS 1 5において、揺動条件出力部4 5 0は、ステップS 1 3において出力された最適揺動条件が、ステップS 1 2において算出された切屑細断揺動条件を満たすか否かを判定する。最適揺動条件が切屑細断揺動条件を満たす場合（YES）、処理は、ステップS 1 6へ移る。一方、最適揺動条件が切屑細断揺動条件を満たさない場合（NO）、処理は、再度ステップS 1 4へ移る。
- [0042] ステップS 1 6において、揺動条件出力部4 5 0は、ステップS 1 3において出力された最適揺動条件が、ステップS 1 3において算出された上限揺動条件を満たすか否かを判定する。最適揺動条件が上限揺動条件を満たす場合（YES）、処理は、ステップS 1 6へ移る。一方、最適揺動条件が上限揺動条件を満たさない場合（NO）、処理は、再度ステップS 1 4へ移る。
- [0043] ステップS 1 7において、揺動条件出力部4 5 0は、最適揺動条件を数値制御装置2 0 0へ出力し、その後、処理を終了する。
- [0044] 以上説明したように、第1実施形態に係る機械学習装置4 0 0は、揺動加工のための設定条件を取得する設定条件取得部4 1 0と、工作機械1 0 0による加工済みワークの面粗さデータを、ラベルとして取得するラベル取得部4 2 0と、設定条件とラベルとの組みを教師データとして教師あり学習を行う学習部4 3 0と、を備え、学習部4 3 0は、加工済みワークの面粗さを最適化する揺動条件を学習する学習モデルを備える。
- [0045] これにより、機械学習装置4 0 0は、面粗さを考慮した揺動条件の設定が難しかった従来技術と比較して、面粗さを考慮して最適化された最適揺動条件を学習することができる。
- [0046] また、機械学習装置4 0 0は、学習モデルに基づいて、加工済みワークの面粗さを最適化する最適揺動条件を出力する揺動条件出力部4 5 0を更に備える。これにより、機械学習装置4 0 0は、面粗さを考慮した揺動条件の設定が難しかった従来技術と比較して、面粗さを考慮して最適化された最適揺動条件を出力することができる。また、機械学習装置4 0 0は、揺動条件の設定を自動化できるため、作業者の負担を低減することができる。

[0047] また、揺動条件出力部450は、ワークの切屑が細断可能となる切屑細断揺動条件を算出する切屑細断条件算出部451を備え、揺動条件出力部450は、学習モデルに基づいて、切屑細断揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置400は、切屑細断がされ、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0048] また、揺動条件出力部450は、予め設定された上限値を超えない上限揺動条件を算出する揺動条件上限算出部452を備え、揺動条件出力部450は、学習モデルに基づいて、上限揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置400は、切屑細断がされ、かつ揺動条件の上限を超えず、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0049] [第2実施形態]

次に、第2実施形態に係る制御システム10について説明する。なお、第2実施形態の説明にあたっては、第1実施形態と同一の構成要件については同一符号を付し、その説明を省略若しくは簡略化する。第2実施形態に係る制御システム10は、教師あり学習に代えて、強化学習を用いる点が第1実施形態とは主として異なり、他の構成は、第1実施形態と同様の構成を備える。

[0050] 図4は、第2実施形態に係る制御システム10の概要を示す図である。制御システム10は、機械学習を用いて揺動条件を算出することによって、工作機械100において工具の空振りが生じ、かつ良好な面粗さデータを実現できる揺動条件を算出することを目的とする。

[0051] 制御システム1は、工作機械100と、数値制御装置200と、加工面解析装置300と、機械学習装置500と、を備える。ここで、工作機械100と数値制御装置200とは、1対1の組とされて、通信可能に接続されている。なお、工作機械100と数値制御装置200とは、接続インタフェースを介して直接接続されてもよく、又は、LAN (Local Area Network) 等のネットワークを介して接続されてもよい。

[0052] また、数値制御装置200、加工面解析装置300及び機械学習装置500は、それぞれ接続インタフェースを介して直接に接続、又は、それぞれネットワークを介して接続されており、相互に通信を行うことが可能である。なお、ネットワークは、例えば、工場内に構築されたLANや、インターネット、公衆電話網、或いは、これらの組み合わせである。ネットワークにおける具体的な通信方式や、有線接続及び無線接続のいずれであるか等については、特に限定されない。

[0053] 工作機械100、数値制御装置200及び加工面解析装置300は、上述したように第1実施形態と同様の構成を有する。

[0054] 機械学習装置500は、強化学習を行う装置である。機械学習装置500に含まれる各機能ブロックの説明に先立って、まずは、強化学習の基本的な仕組みについて説明する。エージェント（本実施形態における機械学習装置500に相当）は、環境の状態を観測し、ある行動を選択し、当該行動に基づいて環境が変化する。環境の変化に伴って、何らかの報酬が与えられ、エージェントはより良い行動の選択（意思決定）を学習する。

教師あり学習が、完全な正解を示すのに対して、強化学習における報酬は、環境の一部の変化に基づく断片的な値であることが多い。

このため、エージェントは、将来にわたっての報酬の合計を最大にするよう行動を選択するように学習する。

[0055] このように、強化学習では、行動を学習することにより、環境に行動が与える相互作用を踏まえて適切な行動を学習、すなわち将来的に得られる報酬を最大にするための学習する方法を学ぶ。これは、本実施形態において、例えば、受渡時間 T を短縮し、更に、開放時間 T_2 を短縮するための行動情報を選択するという、未来に影響をおよぼすような行動を獲得できることを表している。

[0056] ここで、強化学習としては、任意の学習方法を用いることができるが、以下の説明では、或る環境の状態 s の下で、行動 a を選択する価値 $Q(s, a)$ を学習する方法である Q 学習（ Q -learning）を用いる場合を例

にとって説明をする。

Q学習では、或る状態 s のとき、取り得る行動 a のなかから、価値 $Q(s, a)$ の最も高い行動 a を最適な行動として選択することを目的とする。

[0057] しかしながら、Q学習を最初に開始する時点では、状態 s と行動 a との組み合わせについて、価値 $Q(s, a)$ の正しい値は全く分かっていない。そこで、エージェントは、或る状態 s の下で様々な行動 a を選択し、その時の行動 a に対して、与えられる報酬に基づいて、より良い行動の選択をすることにより、正しい価値 $Q(s, a)$ を学習していく。

[0058] また、将来にわたって得られる報酬の合計を最大化したいので、最終的に価値 $Q(s, a) = E[\sum(\gamma^t) r_t]$ となるようにすることを目指す。ここで $E[\]$ は期待値を表し、 t は時刻、 γ は後述する割引率と呼ばれるパラメータ、 r_t は時刻 t における報酬、 $\sum$ は時刻 t による合計である。この式における期待値は、最適な行動にしたがって状態変化した場合の期待値である。しかしQ学習の過程において、最適な行動が何であるのかは不明であるので、様々な行動を行うことにより、探索しながら強化学習をする。このような価値 $Q(s, a)$ の更新式は、例えば、次の式(1)により表すことができる。

[0059] [数1]

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

[0060] 上記の式(1)において、 s_t は、時刻 t における環境の状態を表し、 a_t は、時刻 t における行動を表す。行動 a_t により、状態は s_{t+1} に変化する。 r_{t+1} は、その状態の変化により得られる報酬を表している。また、 $\max$ の付いた項は、状態 s_{t+1} の下で、その時に分かっている最もQ値の高い行動 a を選択した場合のQ値に γ を乗じたものになる。ここで、 γ は、 $0 < \gamma \leq 1$ のパラメータで、割引率と呼ばれる。また、 α は、学習係数で、 $0 < \alpha \leq 1$ の範囲とする。

[0061] 上述した式(1)は、試行 a_t の結果、返ってきた報酬 r_{t+1} を元に、状態

s_t における行動 a_t の価値 $Q(s_t, a_t)$ を更新する方法を表している。

この更新式は、状態 s_t における行動 a_t の価値 $Q(s_t, a_t)$ よりも、行動 a_t による次の状態 s_{t+1} における最良の行動の価値 $\max_a Q(s_{t+1}, a)$ の方が大きければ、価値 $Q(s_t, a_t)$ を大きくし、逆に小さければ、価値 $Q(s_t, a_t)$ を小さくすることを示している。つまり、或る状態における或る行動の価値を、それによる次の状態における最良の行動の価値に近づける。ただし、その差は、割引率 γ と報酬 r_{t+1} のあり方により変わってくるが、基本的には、ある状態における最良の行動の価値が、それに至る一つ前の状態における行動の価値に伝播していく仕組みになっている。

[0062] ここで、Q学習では、すべての状態行動ペア (s, a) についての価値 $Q(s, a)$ のテーブルを作成して、学習を行う方法がある。しかし、すべての状態行動ペアの価値 $Q(s, a)$ の値を求めるには状態数が多すぎて、Q学習が収束するのに多くの時間を要してしまう場合がある。

[0063] そこで、公知のDQN (Deep Q-Network) と呼ばれる技術を利用するようにしてもよい。具体的には、価値関数 Q を適当なニューラルネットワークを用いて構成し、ニューラルネットワークのパラメータを調整することにより、価値関数 Q を適当なニューラルネットワークで近似することにより価値 $Q(s, a)$ の値を算出するようにしてもよい。DQNを利用することにより、Q学習が収束するのに要する時間を短くすることが可能となる。なお、DQNについては、例えば、以下の非特許文献に詳細な記載がある。

[0064] <非特許文献>

「Human-level control through deep reinforcement learning」、Volodymyr Mnih 1 著 [online]、[令和2年12月15日検索]、インターネット<URL: <http://files.davidqiu.com/research/nature14236.pdf>>

[0065] 機械学習装置500は、上記において説明をしたQ学習を行う。具体的に

は、機械学習装置 500 は、数値制御装置 200 において設定された設定条件（加工条件、揺動条件等）を状態 s とし、当該状態 s に係る数値制御装置 200 における揺動条件の変更を行動 a として、選択する価値関数 Q を学習する。

[0066] 機械学習装置 500 は、数値制御装置 200 において設定された設定条件（加工条件、揺動条件等）等の状態 s を観測して、行動 a を決定する。機械学習装置 500 は、行動 a をするたびに報酬が返ってくる。機械学習装置 500 は、将来にわたっての報酬の合計が最大になるように、最適な行動 a を試行錯誤的に探索する。そうすることで、機械学習装置 500 は、数値制御装置 200 において設定された設定条件（加工条件、揺動条件等）等である状態 s に対して、最適な行動 a を選択することが可能となる。

[0067] すなわち、機械学習装置 500 により学習された価値関数 Q に基づいて、或る状態 s に係る設定条件（加工条件、揺動条件等）に対して、適用される行動 a のうち、価値関数 Q の値が最大となるような行動 a を選択することで、面粗さデータが最小（最適）になるような行動 a を選択することが可能となる。

[0068] 以上の強化学習を行うために、機械学習装置 500 は、設定条件取得部 510 と、判定情報取得部 520 と、行動情報出力部 530 と、学習部 540 と、価値関数記憶部 550 と、揺動条件出力部 560 と、を備える。

[0069] 設定条件取得部 510 は、数値制御装置 200 において設定された設定条件（加工条件、揺動条件等）を、状態情報（状態 s ）として数値制御装置 200 から取得する。この状態 s は、 Q 学習における、環境の状態 s に相当する。

[0070] 具体的には、第 2 実施形態における状態 s は、数値制御装置 200 において設定された設定条件（加工条件、揺動条件等）を示す。設定条件は、工作機械 100 の主軸回転速度、送り速度等を含む加工条件と、工作機械 100 の揺動振幅、揺動周波数等を含む揺動条件と、を有する。

[0071] 判定情報取得部 520 は、 Q 学習を行うための報酬を算出するための判定

情報を取得する。具体的には、判定情報取得部 520 は、工作機械 100 による加工済みワークの面粗さデータを、Q 学習を行うための報酬を算出するための判定情報として取得する。

[0072] 行動情報出力部 530 は、学習部 540 が生成した行動情報（行動 a）を、数値制御装置 200 に対して送信する。数値制御装置 200 は、上述したように、この行動 a に基づいて、現在の状態 s、すなわち現在設定されている設定条件を変更することで、次の状態 s'（すなわち変更された揺動条件を工作機械 100 において実行した場合の状態）に遷移する。

[0073] 学習部 540 は、或る環境の状態 s の下で、ある行動 a を選択する場合の価値 $Q(s, a)$ を学習する。具体的には、学習部 540 は、報酬算出部 541 と、価値関数更新部 542 と、行動情報生成部 543 と、を備える。

[0074] 報酬算出部 541 は、或る状態 s の下で、行動 a を選択した場合の報酬を、判定情報に基づいて算出する。ここで、報酬算出部 541 は、加工済みワークの面粗さデータが所定の閾値未満のときに、報酬を正の値とし、面粗さデータが所定の閾値を超えたときに、報酬を負の値とする。

[0075] また、報酬算出部 541 は、理論面粗さの近似式に基づいて所定の閾値を決定する。具体的には、報酬算出部 541 は、理論面粗さの近似式に補正係数 a を掛けた値を所定の閾値として決定する。ここで、所定の閾値は、 $a \cdot F^2 / 8R$ と表され、F は送り量 [mm]、R は工具半径を表す。

[0076] 価値関数更新部 542 は、状態 s と、行動 a と、行動 a を状態 s に適用した場合の状態 s' と、上記のようにして算出された報酬の値とに基づいて Q 学習を行うことにより、価値関数記憶部 550 が記憶する価値関数 Q を更新する。

[0077] 価値関数 Q の更新は、オンライン学習で行ってもよく、バッチ学習で行ってもよく、ミニバッチ学習で行ってもよい。オンライン学習とは、或る行動 a を現在の状態 s に適用することにより、状態 s が新たな状態 s' に遷移する都度、即座に価値関数 Q の更新を行うという学習方法である。

[0078] また、バッチ学習とは、或る行動 a を現在の状態 s に適用することにより

、状態 s が新たな状態 s' に遷移することを繰り返すことにより、学習用のデータを収集し、収集したすべての学習用データを用いて、価値関数 Q の更新を行うという学習方法である。更に、ミニバッチ学習とは、オンライン学習と、バッチ学習の中間的な、ある程度学習用データが溜まるたびに価値関数 Q の更新を行うという学習方法である。

[0079] 行動情報生成部 543 は、 Q 学習の過程において、行動 a を生成して、生成した行動 a を行動情報出力部 530 に対して出力する。具体的には、行動情報生成部 543 は、現在の状態 s に対して、 Q 学習の過程における行動 a を選択する。第 2 実施形態における行動 a は、現在の状態 s に対して揺動条件をどのように変更すべきかを含む。

[0080] 更に、行動情報生成部 543 は、現在の推定される行動 a の価値の中で、最も価値 $Q(s, a)$ の高い行動 a' を選択するグリーディ法や、ある小さな確率 ε でランダムに行動 a' を選択し、それ以外では最も価値 $Q(s, a)$ の高い行動 a' を選択する ε グリーディ法といった公知の方法により、行動 a' を選択する方策を取るようにしてもよい。

[0081] 価値関数記憶部 550 は、価値関数 Q を記憶する記憶装置である。価値関数記憶部 550 に記憶された価値関数 Q は、価値関数更新部 542 により更新される。

[0082] また、機械学習装置 500 は、価値関数更新部 542 が Q 学習を行うことにより更新した価値関数 Q に基づいて、価値 $Q(s, a)$ が最大となる揺動加工を工作機械 100 に行わせるための揺動条件としての行動 a （以下、「最適化行動情報」と呼ぶ。）を生成する。

[0083] 揺動条件出力部 560 は、価値関数記憶部 550 が記憶している価値関数 Q を取得する。この価値関数 Q は、上述したように、価値関数更新部 542 が Q 学習を行うことにより更新したものである。そして、揺動条件出力部 560 は、価値関数 Q に基づいて、最適化行動情報として、加工済みワークの面粗さを最適化する最適揺動条件を生成し、生成した最適揺動条件（最適化行動情報）を数値制御装置 200 に対して出力する。

- [0084] また、揺動条件出力部560は、切屑細断条件算出部561と、揺動条件上限算出部562と、を有する。
- [0085] 切屑細断条件算出部561は、工作機械100の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。そして、揺動条件出力部560は、切屑細断揺動条件を満たす最適揺動条件を出力する。
- [0086] 揺動条件上限算出部562は、予め設定された上限値を超えない上限揺動条件を算出する。そして、揺動条件出力部560は、上限揺動条件を満たす最適揺動条件を出力する。
- [0087] 数値制御装置200が、この最適揺動条件（最適化行動情報）に基づいて現在設定されている揺動条件を修正して、動作指令を生成することにより、工作機械100は、加工済みワークの面粗さが最小（最適）になるように動作することができる。
- [0088] 次に、図5のフローチャートを参照して機械学習装置500による価値関数の更新処理について説明をする。図5は、第2実施形態に係る機械学習装置500による価値関数の更新処理の流れを示すフローチャートである。
- [0089] まず、ステップS21において、設定条件取得部510は、数値制御装置200から状態情報としての設定条件を取得する。取得した設定条件は、価値関数更新部542及び行動情報生成部543に対して出力される。上述したように、この設定条件（状態情報）は、Q学習における環境の状態sに相当する情報である。
- [0090] ステップS22において、行動情報生成部543は、新たな行動情報としての揺動条件を生成し、生成した新たな行動情報（行動a）を、行動情報出力部530を介して数値制御装置200に対して出力する。行動情報を受信した数値制御装置200は、受信した行動情報に基づいて現在の状態sに係る揺動条件を変更した状態s'により、工作機械100を駆動させて、ワークの揺動加工を行う。上述したように、この行動情報は、Q学習における行動aに相当するものである。
- [0091] ステップS23において、判定情報取得部520は、工作機械100によ

る加工済みワークの面粗さデータを、Q学習を行うための報酬を算出するための判定情報として取得する。

[0092] ステップS 2 4において、報酬算出部5 3 1は、入力された判定情報（加工済みワークの面粗さ）に基づいて報酬を算出する。報酬算出部5 4 1は、加工済みワークの面粗さデータが所定の閾値未満であるか否かを判定する。面粗さデータが所定の閾値未満である場合（Y E S）、処理は、ステップS 2 5へ移る。一方、面粗さデータが所定の閾値を超える場合（N O）、処理は、ステップS 2 6へ移る。

[0093] ステップS 2 5において、報酬算出部5 4 1は、報酬の値を正の値として算出する。

ステップS 2 6において、報酬算出部5 4 1は、報酬の値を負の値として算出する。

[0094] ステップS 2 7において、価値関数更新部5 4 2は、上述にて算出された報酬の値に基づいて、価値関数記憶部5 5 0が記憶している価値関数Qを更新する。そして、学習部6 4 0は、再度ステップS 2 1に戻り、上述した処理を繰り返すことにより、価値関数Qは、適切な値に収束していく。なお、学習部6 4 0は、上述した処理を、所定回数繰り返したことや、所定時間繰り返したことを条件として終了するようにしてもよい。以上、機械学習装置5 0 0の動作について説明したが、例えば、報酬の値を算出する処理は、一例であって、これに限定されない。

[0095] 図6は、第2実施形態に係る機械学習装置5 0 0による揺動条件を出力する処理の流れを示すフローチャートである。

ステップS 3 1において、揺動条件出力部5 6 0は、価値関数記憶部5 5 0が記憶している価値関数Qを取得する。この価値関数Qは、上述したように価値関数更新部5 4 2がQ学習を行うことにより更新したものである。

[0096] ステップS 3 2において、揺動条件出力部5 6 0は、価値関数Qに基づいて、例えば現在設定されている状態sにおいて、取り得る行動aのなかから価値Q（s， a）の最も高い行動aを最適な行動として選択することで最適

揺動条件（最適化行動情報）を生成する。

- [0097] ステップS 3 3において、切屑細断条件算出部5 6 1は、工作機械1 0 0の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。
- [0098] ステップS 3 4において、揺動条件上限算出部5 6 2は、予め設定された上限値を超えない上限揺動条件を算出する。
- [0099] ステップS 3 5において、揺動条件出力部5 6 0は、ステップS 3 2において生成された最適揺動条件が、ステップS 3 3において算出された切屑細断揺動条件を満たすか否かを判定する。最適揺動条件が切屑細断揺動条件を満たす場合（Y E S）、処理は、ステップS 3 5へ移る。一方、最適揺動条件が切屑細断揺動条件を満たさない場合（N O）、処理は、再度ステップS 3 1へ移る。
- [0100] ステップS 3 6において、揺動条件出力部5 6 0は、ステップS 3 2において出力された最適揺動条件が、ステップS 3 4において算出された上限揺動条件を満たすか否かを判定する。最適揺動条件が上限揺動条件を満たす場合（Y E S）、処理は、ステップS 3 7へ移る。一方、最適揺動条件が上限揺動条件を満たさない場合（N O）、処理は、再度ステップS 3 1へ移る。
- [0101] ステップS 3 2において、揺動条件出力部5 6 0は、生成した最適揺動条件（最適化行動情報）を数値制御装置2 0 0に対して出力する。
- [0102] 以上により、数値制御装置2 0 0は、最適揺動条件に基づいて現在設定されている状態s（すなわち、現在設定されている揺動条件）を修正して、動作指令を生成する。そして、数値制御装置2 0 0は、生成した動作指令を工作機械1 0 0に送ることにより、工作機械1 0 0は、加工済みワークの面粗さデータが最適（最小）になるように動作することができる。
- [0103] 以上説明したように、第2実施形態に係る機械学習装置5 0 0は、揺動加工のための設定条件を、状態情報として取得する設定条件取得部5 1 0と、工作機械1 0 0による加工済みワークの面粗さデータを、判定情報として取得する判定情報取得部5 2 0と、現在の状態に対して前記揺動条件をどのよ

うに変更すべきかを示す行動情報を出力する行動情報出力部530と、判定情報に基づいて、強化学習における報酬の値を算出する報酬算出部541と、状態情報、行動情報及び報酬に基づいて、工作機械100の揺動条件の価値を定める価値関数を更新する価値関数更新部542と、価値関数に基づいて、加工済みワークの面粗さを最適化する最適揺動条件を出力する揺動条件出力部560と、を備える。

[0104] これにより、機械学習装置500は、面粗さを考慮した揺動条件の設定が難しかった従来技術と比較して、面粗さを考慮して最適化された最適揺動条件を出力することができる。また、機械学習装置500は、揺動条件の設定を自動化できるため、作業者の負担を低減することができる。

[0105] また、報酬算出部541は、加工済みワークの面粗さデータが所定の閾値未満のときに、報酬を正の値とし、面粗さデータが所定の閾値を超えたときに、報酬を負の値とする。これにより、機械学習装置500は、面粗さデータを考慮した報酬の値を決定することができる。

[0106] また、報酬算出部541は、理論面粗さの近似式に基づいて所定の閾値を決定する。これにより、機械学習装置500は、理論面粗さを考慮した報酬の値を決定することができる。

[0107] また、揺動条件出力部560は、ワークの切屑が細断可能となる切屑細断揺動条件を算出する切屑細断条件算出部561を備え、揺動条件出力部560は、切屑細断揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置500は、切屑細断がされ、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0108] 揺動条件出力部560は、予め設定された上限値を超えない上限揺動条件を算出する揺動条件上限算出部562を備え、揺動条件出力部560は、上限揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置500は、切屑細断がされ、かつ揺動条件の上限を超えず、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0109] [第3実施形態]

次に、第3実施形態に係る制御システム20について説明する。なお、第3実施形態の説明にあたっては、第1実施形態と同一の構成要件については同一符号を付し、その説明を省略若しくは簡略化する。第3実施形態に係る制御システム10は、加工済みワークの面粗さデータに代えて、加工済みワークの真円度を用いる点が第1実施形態とは主として異なり、他の構成は、第1実施形態と同様の構成を備える。

[0110] 図7は、第3実施形態に係る制御システム20の概要を示す図である。制御システム20は、機械学習を用いて揺動条件を算出することによって、工作機械100において工具の空振りが生じ、かつ良好な真円度データを実現できる揺動条件を算出することを目的とする。

[0111] 制御システム20は、工作機械100と、数値制御装置200と、加工面解析装置600と、機械学習装置700と、を備える。ここで、工作機械100と数値制御装置200とは、1対1の組とされて、通信可能に接続されている。なお、工作機械100と数値制御装置200とは、接続インタフェースを介して直接接続されてもよく、又は、LAN (Local Area Network) 等のネットワークを介して接続されてもよい。

[0112] また、数値制御装置200、加工面解析装置600及び機械学習装置700は、それぞれ接続インタフェースを介して直接に接続、又は、それぞれネットワークを介して接続されており、相互に通信を行うことが可能である。なお、ネットワークは、例えば、工場内に構築されたLANや、インターネット、公衆電話網、或いは、これらの組み合わせである。ネットワークにおける具体的な通信方式や、有線接続及び無線接続のいずれであるか等については、特に限定されない。

[0113] 工作機械100及び数値制御装置200は、上述したように第1実施形態と同様の構成を有する。

加工面解析装置600は、工作機械100において加工された加工済みワークの真円度データを測定及び算出し、算出した面粗さデータを機械学習装置700へ出力する。

- [0114] 機械学習装置 700 は、数値制御装置 200 から受け付けた設定条件及び加工面解析装置 600 から受け付けた真円度データを用いて機械学習を行う。そして、機械学習装置 700 は、機械学習を行うことにより、揺動条件を出力するための学習モデルを構築する。機械学習装置 700 は、設定条件取得部 710 と、ラベル取得部 720 と、学習部 730 と、学習モデル記憶部 740 と、揺動条件出力部 750 と、を備える。
- [0115] 設定条件取得部 710 は、揺動加工のための加工条件及び揺動条件を含む設定条件を数値制御装置 200 から取得する。
- [0116] ラベル取得部 720 は、工作機械 100 による加工済みワークの真円度データを、加工面解析装置 600 から、ラベルとして取得する。ここで、ラベルとは、機械学習において入力に対応すべき正しい出力である。
- [0117] このようにして、学習対象とした揺動加工のための設定条件と、真円度データのラベルとが組となり、学習部 430 に入力される。この設定条件とラベルの組は、機械学習における教師データに相当する。ここで、機械学習を行うための教師データは、多数用意されることが望ましい。
- [0118] 学習部 730 は、入力された教師データに基づいて機械学習を行うことにより、学習モデルを構築する。すなわち、学習部 730 は、加工済みワークの真円度を最適化する揺動条件を学習する。学習部 730 が構築した学習モデルは、学習モデル記憶部 740 に対して出力される。
- [0119] 学習モデル記憶部 740 は、学習部 730 によって学習された学習モデルを記憶する。学習モデル記憶部 740 が記憶した学習モデルは、揺動条件出力部 750 により利用される。なお、学習モデルを構築した後に、新たに教師データを取得した場合には、学習部 730 は、学習モデル記憶部 740 が記憶した学習モデルに対して更に教師あり学習を行うことにより、一度構築した学習モデルを更新するようにしてもよい。
- [0120] また、機械学習装置 700 は、学習モデル記憶部 740 が記憶した学習モデルを、他の機械学習装置との間で共有するようにしてもよい。学習モデルを複数の機械学習装置で共有するようになれば、各機械学習装置にて分散し

て、更に教師あり学習を行うことが可能になるので、教師あり学習の効率及び精度を向上させることができる。

[0121] 揺動条件出力部 750 は、学習モデル記憶部 740 が記憶した学習モデルに基づいて、加工済みワークの真円度を最適化する最適揺動条件を出力する。

また、揺動条件出力部 750 は、切屑細断条件算出部 751 と、揺動条件上限算出部 752 と、を有する。

[0122] 切屑細断条件算出部 751 は、工作機械 100 の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。そして、揺動条件出力部 750 は、学習モデルに基づいて、切屑細断揺動条件を満たす最適揺動条件を出力する。

[0123] 揺動条件上限算出部 752 は、予め設定された上限値を超えない上限揺動条件を算出する。そして、揺動条件出力部 750 は、学習モデルに基づいて、上限揺動条件を満たす最適揺動条件を出力する。

[0124] 次に、図 8 のフローチャートを参照して機械学習装置 700 による学習モデルの構築処理について説明をする。

図 8 は、第 3 実施形態に係る機械学習装置 700 による学習モデルの構築処理の流れを示すフローチャートである。

[0125] ステップ S41 において、設定条件取得部 710 は、揺動加工のための加工条件及び揺動条件を含む設定条件を数値制御装置 200 から取得する。

[0126] ステップ S42 において、ラベル取得部 720 は、工作機械 100 による加工済みワークの真円度データを、加工面解析装置 600 から、ラベルとして取得する。

[0127] ステップ S43 において、学習部 730 は、設定条件とラベルとを組にした教師データが入力されると、入力された教師データに基づいて機械学習を実行する。学習部 730 は、入力された教師データを用いて機械学習を実行する。学習部 730 は、例えば、回帰分析、ニューラルネットワーク、最小二乗法、ステップワイズ法等により、教師あり学習を行う。

[0128] ステップS4において、学習部730は、教師あり学習を終了するか否かを判定する。ここで、教師あり学習を終了させる条件は、任意に定めることができる。例えば、学習部730は、ニューラルネットワークの出力とラベルとの誤差の値が所定値以下となった場合に、教師あり学習を終了させるようにしてもよい。また、学習部730は、予め定めておいた回数だけ教師あり学習を繰り返した場合に、教師あり学習を終了させるようにしてもよい。教師あり学習を終了する場合（YES）、処理は、ステップS45へ移る。一方、教師あり学習を終了しない場合（NO）、処理は、ステップS43へ移る。

[0129] ステップS45において、学習部730は、その時点までの教師あり学習により構築した学習モデルを、学習モデル記憶部740に出力して記憶させる。また、新たな教師データを取得した場合に、機械学習装置700は、学習モデルに対して更なる機械学習を行うことができる。その後、機械学習装置700は、本処理を終了する。

[0130] 次に、図9のフローチャートを参照して機械学習装置700による揺動条件の出力処理について説明をする。

図9は、第3実施形態に係る機械学習装置700による揺動条件の出力処理の流れを示すフローチャートである。

[0131] ステップS51において、設定条件取得部710は、揺動加工のための加工条件及び揺動条件を含む設定条件を数値制御装置200から取得する。

[0132] ステップS52において、切屑細断条件算出部751は、工作機械100の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。

ステップS53において、揺動条件上限算出部752は、予め設定された上限値を超えない上限揺動条件を算出する。

[0133] ステップS54において、揺動条件出力部750は、ステップS51において取得された設定条件及び学習モデル記憶部740が記憶した学習モデルに基づいて、加工済みワークの真円度を最適化する最適揺動条件を出力する

。

[0134] ステップS 5 5において、揺動条件出力部7 5 0は、ステップS 5 3において出力された最適揺動条件が、ステップS 5 2において算出された切屑細断揺動条件を満たすか否かを判定する。最適揺動条件が切屑細断揺動条件を満たす場合（Y E S）、処理は、ステップS 5 6へ移る。一方、最適揺動条件が切屑細断揺動条件を満たさない場合（N O）、処理は、再度ステップS 5 4へ移る。

[0135] ステップS 5 6において、揺動条件出力部7 5 0は、ステップS 5 3において出力された最適揺動条件が、ステップS 5 3において算出された上限揺動条件を満たすか否かを判定する。最適揺動条件が上限揺動条件を満たす場合（Y E S）、処理は、ステップS 5 6へ移る。一方、最適揺動条件が上限揺動条件を満たさない場合（N O）、処理は、再度ステップS 5 4へ移る。

[0136] ステップS 5 7において、揺動条件出力部7 5 0は、最適揺動条件を数値制御装置2 0 0へ出力し、その後、処理を終了する。

[0137] 以上説明したように、第3実施形態に係る機械学習装置7 0 0は、揺動加工のための設定条件を取得する設定条件取得部7 1 0と、工作機械1 0 0による加工済みワークの真円度データを、ラベルとして取得するラベル取得部7 2 0と、設定条件とラベルとの組みを教師データとして教師あり学習を行う学習部7 3 0と、を備え、学習部7 3 0は、加工済みワークの真円度を最適化する揺動条件を学習する学習モデルを備える。

[0138] これにより、機械学習装置7 0 0は、真円度を考慮した揺動条件の設定が難しかった従来技術と比較して、真円度を考慮して最適化された最適揺動条件を学習することができる。

[0139] また、機械学習装置7 0 0は、学習モデルに基づいて、加工済みワークの真円度を最適化する最適揺動条件を出力する揺動条件出力部7 5 0を更に備える。これにより、機械学習装置7 0 0は、真円度を考慮した揺動条件の設定が難しかった従来技術と比較して、真円度を考慮して最適化された最適揺動条件を出力することができる。また、機械学習装置7 0 0は、揺動条件の

設定を自動化できるため、作業者の負担を低減することができる。

[0140] また、揺動条件出力部 750 は、ワークの切屑が細断可能となる切屑細断揺動条件を算出する切屑細断条件算出部 751 を備え、揺動条件出力部 750 は、学習モデルに基づいて、切屑細断揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置 700 は、切屑細断がされ、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0141] また、揺動条件出力部 750 は、予め設定された上限値を超えない上限揺動条件を算出する揺動条件上限算出部 752 を備え、揺動条件出力部 750 は、学習モデルに基づいて、上限揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置 700 は、切屑細断がされ、かつ揺動条件の上限を超えず、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0142] [第 4 実施形態]

次に、第 4 実施形態に係る制御システム 30 について説明する。なお、第 4 実施形態の説明にあたっては、第 3 実施形態と同一の構成要件については同一符号を付し、その説明を省略若しくは簡略化する。第 4 実施形態に係る制御システム 30 は、教師あり学習に代えて、強化学習を用いる点が第 3 実施形態とは主として異なり、他の構成は、第 3 実施形態と同様の構成を備える。

[0143] 図 10 は、第 4 実施形態に係る制御システム 30 の概要を示す図である。制御システム 30 は、機械学習を用いて揺動条件を算出することによって、工作機械 100 において工具の空振りが生じ、かつ良好な真円度データを実現できる揺動条件を算出することを目的とする。

[0144] 制御システム 30 は、工作機械 100 と、数値制御装置 200 と、加工面解析装置 600 と、機械学習装置 800 と、を備える。ここで、工作機械 100 と数値制御装置 200 とは、1 対 1 の組とされて、通信可能に接続されている。なお、工作機械 100 と数値制御装置 200 とは、接続インタフェースを介して直接接続されてもよく、又は、LAN (Local Area

Network)等のネットワークを介して接続されてもよい。

[0145] また、数値制御装置200、加工面解析装置600及び機械学習装置800は、それぞれ接続インタフェースを介して直接に接続、又は、それぞれネットワークを介して接続されており、相互に通信を行うことが可能である。なお、ネットワークは、例えば、工場内に構築されたLANや、インターネット、公衆電話網、或いは、これらの組み合わせである。ネットワークにおける具体的な通信方式や、有線接続及び無線接続のいずれであるか等については、特に限定されない。

[0146] 工作機械100、数値制御装置200及び加工面解析装置600は、上述したように第3実施形態と同様の構成を有する。

[0147] 機械学習装置800は、強化学習を行う装置である。なお、強化学習に関する処理は、第2実施形態に係る機械学習装置500と同様であるため、説明を省略する。強化学習を行うために、機械学習装置800は、設定条件取得部810と、判定情報取得部820と、行動情報出力部830と、学習部840と、価値関数記憶部850と、揺動条件出力部860と、を備える。

[0148] 設定条件取得部810は、数値制御装置200において設定された設定条件（加工条件、揺動条件等）を、状態情報（状態s）として数値制御装置200から取得する。この状態sは、Q学習における、環境の状態sに相当する。

[0149] 具体的には、第4実施形態における状態sは、数値制御装置200において設定された設定条件（加工条件、揺動条件等）を示す。設定条件は、工作機械100の主軸回転速度、送り速度等を含む加工条件と、工作機械100の揺動振幅、揺動周波数等を含む揺動条件と、を有する。

[0150] 判定情報取得部820は、Q学習を行うための報酬を算出するための判定情報を取得する。具体的には、判定情報取得部820は、工作機械100による加工済みワークの真円度データを、Q学習を行うための報酬を算出するための判定情報として取得する。

[0151] 行動情報出力部830は、学習部840が生成した行動情報（行動a）を

、数値制御装置 200 に対して送信する。数値制御装置 200 は、上述したように、この行動 a に基づいて、現在の状態 s 、すなわち現在設定されている設定条件を変更することで、次の状態 s' （すなわち変更された揺動条件を工作機械 100 において実行した場合の状態）に遷移する。

[0152] 学習部 840 は、或る環境の状態 s の下で、ある行動 a を選択する場合の価値 $Q(s, a)$ を学習する。具体的には、学習部 840 は、報酬算出部 841 と、価値関数更新部 842 と、行動情報生成部 843 と、を備える。

[0153] 報酬算出部 841 は、或る状態 s の下で、行動 a を選択した場合の報酬を、判定情報に基づいて算出する。ここで、報酬算出部 841 は、加工済みワークの真円度データが所定の閾値未満のときに、報酬を正の値とし、真円度データが所定の閾値を超えたときに、報酬を負の値とする。

[0154] 価値関数更新部 842 は、状態 s と、行動 a と、行動 a を状態 s に適用した場合の状態 s' と、上記のようにして算出された報酬の値とに基づいて Q 学習を行うことにより、価値関数記憶部 850 が記憶する価値関数 Q を更新する。

[0155] 価値関数 Q の更新は、オンライン学習で行ってもよく、バッチ学習で行ってもよく、ミニバッチ学習で行ってもよい。

[0156] 行動情報生成部 843 は、 Q 学習の過程において、行動 a を生成して、生成した行動 a を行動情報出力部 830 に対して出力する。具体的には、行動情報生成部 843 は、現在の状態 s に対して、 Q 学習の過程における行動 a を選択する。第 4 実施形態における行動 a は、現在の状態 s に対して揺動条件をどのように変更すべきかを含む。

[0157] 更に、行動情報生成部 843 は、現在の推定される行動 a の価値の中で、最も価値 $Q(s, a)$ の高い行動 a' を選択するグリーディ法や、ある小さな確率 ε でランダムに行動 a' を選択し、それ以外では最も価値 $Q(s, a)$ の高い行動 a' を選択する ε グリーディ法といった公知の方法により、行動 a' を選択する方策を取るようにしてもよい。

[0158] 価値関数記憶部 850 は、価値関数 Q を記憶する記憶装置である。価値関

数記憶部 850 に記憶された価値関数 Q は、価値関数更新部 842 により更新される。

[0159] また、機械学習装置 800 は、価値関数更新部 842 が Q 学習を行うことにより更新した価値関数 Q に基づいて、価値 $Q(s, a)$ が最大となる揺動加工を工作機械 100 に行わせるための揺動条件としての行動 a （以下、「最適化行動情報」と呼ぶ。）を生成する。

[0160] 揺動条件出力部 860 は、価値関数記憶部 850 が記憶している価値関数 Q を取得する。この価値関数 Q は、上述したように、価値関数更新部 842 が Q 学習を行うことにより更新したものである。そして、揺動条件出力部 860 は、価値関数 Q に基づいて、最適化行動情報として、加工済みワークの面粗さを最適化する最適揺動条件を生成し、生成した最適揺動条件（最適化行動情報）を数値制御装置 200 に対して出力する。

[0161] また、揺動条件出力部 860 は、切屑細断条件算出部 861 と、揺動条件上限算出部 862 と、を有する。

[0162] 切屑細断条件算出部 861 は、工作機械 100 の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。そして、揺動条件出力部 860 は、切屑細断揺動条件を満たす最適揺動条件を出力する。

[0163] 揺動条件上限算出部 862 は、予め設定された上限値を超えない上限揺動条件を算出する。そして、揺動条件出力部 860 は、上限揺動条件を満たす最適揺動条件を出力する。

[0164] 数値制御装置 200 が、この最適揺動条件（最適化行動情報）に基づいて現在設定されている揺動条件を修正して、動作指令を生成することにより、工作機械 100 は、加工済みワークの真円度が最適になるように動作することができる。

[0165] 次に、図 11 のフローチャートを参照して機械学習装置 800 による価値関数の更新処理について説明をする。図 11 は、第 4 実施形態に係る機械学習装置 800 による価値関数の更新処理の流れを示すフローチャートである。

[0166] まず、ステップS 6 1において、設定条件取得部8 1 0は、数値制御装置2 0 0から状態情報としての設定条件を取得する。取得した設定条件は、価値関数更新部8 4 2及び行動情報生成部8 4 3に対して出力される。上述したように、この設定条件（状態情報）は、Q学習における環境の状態sに相当する情報である。

[0167] ステップS 6 2において、行動情報生成部8 4 3は、新たな行動情報としての揺動条件を生成し、生成した新たな行動情報（行動a）を、行動情報出力部8 3 0を介して数値制御装置2 0 0に対して出力する。行動情報を受信した数値制御装置2 0 0は、受信した行動情報に基づいて現在の状態sに係る揺動条件を変更した状態s'により、工作機械1 0 0を駆動させて、ワークの揺動加工を行う。上述したように、この行動情報は、Q学習における行動aに相当するものである。

[0168] ステップS 6 3において、判定情報取得部8 2 0は、工作機械1 0 0による加工済みワークの真円度データを、Q学習を行うための報酬を算出するための判定情報として取得する。

[0169] ステップS 6 4において、報酬算出部8 3 1は、入力された判定情報（加工済みワークの真円度データ）に基づいて報酬を算出する。報酬算出部8 4 1は、加工済みワークの真円度データが所定の閾値未満であるか否かを判定する。真円度データが所定の閾値未満である場合（Y E S）、処理は、ステップS 6 5へ移る。一方、真円度データが所定の閾値を超える場合（N O）、処理は、ステップS 6 6へ移る。

[0170] ステップS 6 5において、報酬算出部8 4 1は、報酬の値を正の値として算出する。

ステップS 6 6において、報酬算出部8 4 1は、報酬の値を負の値として算出する。

[0171] ステップS 6 7において、価値関数更新部8 4 2は、上述にて算出された報酬の値に基づいて、価値関数記憶部8 5 0が記憶している価値関数Qを更新する。そして、学習部8 4 0は、再度ステップS 6 1に戻り、上述した処

理を繰り返すことにより、価値関数 Q は、適切な値に収束していく。なお、学習部840は、上述した処理を、所定回数繰り返したことや、所定時間繰り返したことを条件として終了するようにしてもよい。以上、機械学習装置800の動作について説明したが、例えば、報酬の値を算出する処理は、一例であって、これに限定されない。

[0172] 図12は、第4実施形態に係る機械学習装置800による揺動条件を出力する処理の流れを示すフローチャートである。

ステップS71において、揺動条件出力部860は、価値関数記憶部850が記憶している価値関数 Q を取得する。この価値関数 Q は、上述したように価値関数更新部842が Q 学習を行うことにより更新したものである。

[0173] ステップS72において、揺動条件出力部860は、価値関数 Q に基づいて、例えば現在設定されている状態 s において、取り得る行動 a のなかから価値 $Q(s, a)$ の最も高い行動 a を最適な行動として選択することで最適揺動条件（最適化行動情報）を生成する。

[0174] ステップS73において、切屑細断条件算出部861は、工作機械100の揺動加工によってワークの切屑が細断可能となる切屑細断揺動条件を算出する。

[0175] ステップS74において、揺動条件上限算出部862は、予め設定された上限値を超えない上限揺動条件を算出する。

[0176] ステップS75において、揺動条件出力部860は、ステップS72において生成された最適揺動条件が、ステップS73において算出された切屑細断揺動条件を満たすか否かを判定する。最適揺動条件が切屑細断揺動条件を満たす場合（YES）、処理は、ステップS75へ移る。一方、最適揺動条件が切屑細断揺動条件を満たさない場合（NO）、処理は、再度ステップS71へ移る。

[0177] ステップS76において、揺動条件出力部860は、ステップS72において出力された最適揺動条件が、ステップS74において算出された上限揺動条件を満たすか否かを判定する。最適揺動条件が上限揺動条件を満たす場

合（ＹＥＳ）、処理は、ステップＳ７７へ移る。一方、最適揺動条件が上限揺動条件を満たさない場合（ＮＯ）、処理は、再度ステップＳ７１へ移る。

[0178] ステップＳ７２において、揺動条件出力部８６０は、生成した最適揺動条件（最適化行動情報）を数値制御装置２００に対して出力する。

[0179] 以上により、数値制御装置２００は、最適揺動条件に基づいて現在設定されている状態ｓ（すなわち、現在設定されている揺動条件）を修正して、動作指令を生成する。そして、数値制御装置２００は、生成した動作指令を工作機械１００に送ることにより、工作機械１００は、加工済みワークの真円度データが最適になるように動作することができる。

[0180] 以上説明したように、第４実施形態に係る機械学習装置８００は、揺動加工のための設定条件を、状態情報として取得する設定条件取得部８１０と、工作機械１００による加工済みワークの真円度データを、判定情報として取得する判定情報取得部８２０と、現在の状態に対して前記揺動条件をどのように変更すべきかを示す行動情報を出力する行動情報出力部８３０と、判定情報に基づいて、強化学習における報酬の値を算出する報酬算出部８４１と、状態情報、行動情報及び報酬に基づいて、工作機械１００の揺動条件の価値を定める価値関数を更新する価値関数更新部８４２と、価値関数に基づいて、加工済みワークの真円度を最適化する最適揺動条件を出力する揺動条件出力部８６０と、を備える。

[0181] これにより、機械学習装置８００は、面粗さを考慮した揺動条件の設定が難しかった従来技術と比較して、真円度を考慮して最適化された最適揺動条件を出力することができる。また、機械学習装置８００は、揺動条件の設定を自動化できるため、作業者の負担を低減することができる。

[0182] また、報酬算出部８４１は、加工済みワークの真円度データが所定の閾値未満のときに、報酬を正の値とし、真円度データが所定の閾値を超えたときに、報酬を負の値とする。これにより、機械学習装置８００は、面粗さデータを考慮した報酬の値を決定することができる。

[0183] また、揺動条件出力部８６０は、ワークの切屑が細断可能となる切屑細断

揺動条件を算出する切屑細断条件算出部 861 を備え、揺動条件出力部 860 は、切屑細断揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置 800 は、切屑細断がされ、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0184] 揺動条件出力部 860 は、予め設定された上限値を超えない上限揺動条件を算出する揺動条件上限算出部 862 を備え、揺動条件出力部 860 は、上限揺動条件を満たす最適揺動条件を出力する。これにより、機械学習装置 800 は、切屑細断がされ、かつ揺動条件の上限を超えず、かつ面粗さを考慮して最適化された最適揺動条件を設定することができる。

[0185] [他の実施形態]

上述した実施形態では、機械学習装置 400, 500, 700, 800 は、加工済みワークの面粗さデータ及び真円度データを用いたが、他の実施形態に係る機械学習装置は、加工済みワークの評価データとして、他のデータを用いてもよい。例えば、他の実施形態に係る機械学習装置において、加工済みワークの評価データは、面粗さ、真円度又は寸法精度を含んでもよい。ここで、寸法精度は、加工プログラムによって指令された形状通りの寸法で加工されているかを示す。

[0186] すなわち、教師あり学習を用いる場合、他の実施形態に係る機械学習装置は、工具とワークと相対的に揺動させながら揺動加工する工作機械の揺動条件を学習する機械学習装置であって、揺動加工のための設定条件を取得する設定条件取得部と、工作機械による加工済みワークの評価データを、ラベルとして取得するラベル取得部と、設定条件とラベルとの組みを教師データとして教師あり学習を行う学習部と、を備え、学習部は、加工済みワークの評価データを最適化する揺動条件を学習する学習モデルを備える。また、機械学習装置は、学習モデルに基づいて、加工済みワークの評価データを最適化する最適揺動条件を出力する揺動条件出力部を更に備えてもよい。

[0187] これにより、機械学習装置は、加工済みワークの評価データを考慮した揺動条件の設定が難しかった従来技術と比較して、加工済みワークの評価デー

タを考慮して最適化された最適揺動条件を出力することができる。また、機械学習装置は、揺動条件の設定を自動化できるため、作業者の負担を低減することができる。

[0188] また、強化学習を用いる場合、他の実施形態に係る機械学習装置は、工具とワークと相対的に揺動させながら揺動加工する工作機械の揺動条件を学習する機械学習装置であって、揺動加工のための設定条件を、状態情報として取得する設定条件取得部と、工作機械による加工済みワークの評価データを、判定情報として取得する判定情報取得部と、現在の状態に対して揺動条件をどのように変更すべきかを示す行動情報を出力する行動情報出力部と、判定情報に基づいて、強化学習における報酬の値を算出する報酬算出部と、状態情報、行動情報及び報酬に基づいて、工作機械の揺動条件の価値を定める価値関数を更新する価値関数更新部と、価値関数に基づいて、加工済みワークの評価データを最適化する最適揺動条件を出力する揺動条件出力部と、を備える。また、報酬算出部は、加工済みワークの評価データが所定の閾値未満のときに、報酬を正の値とし、評価データが所定の閾値を超えたときに、報酬を負の値とする。

[0189] これにより、機械学習装置は、加工済みワークの評価データを考慮した揺動条件の設定が難しかった従来技術と比較して、加工済みワークの評価データを考慮して最適化された最適揺動条件を出力することができる。また、機械学習装置は、揺動条件の設定を自動化できるため、作業者の負担を低減することができる。

[0190] また、第1実施形態から第4実施形態に係る設定条件は、工作機械100における工具の送り速度、主軸回転速度、座標値、工具の刃先、工具の種類及びワークの材料の少なくとも1つを含む加工条件と、工作機械100の揺動周波数及び揺動振幅の少なくとも1つを含む揺動条件と、を含む。これにより、機械学習装置400、500、700、800は、適切な加工条件及び揺動条件を含む設定条件を取得することができる。

[0191] 第1実施形態及び第2実施形態に係る面粗さデータは、算術平均粗さ、最

大高さ、最大山高さ、最大谷深さ、平均高さ、最大断面高さ及び負荷長さ率の少なくとも1つを含む。これにより、機械学習装置400、500は、適切な手法で面粗さデータを取得することができる。

[0192] また、機械学習装置400、500、700、800は、複数の数値制御装置200によって共有されてもよい。これにより、機械学習装置400、500、700、800は、学習モデルや、価値関数等を複数の数値制御装置200によって共有及び更新することができる。

[0193] また、機械学習装置400、500、700、800は、クラウドサーバ上に設けられてもよい。これにより、機械学習装置400、500は、学習モデルや、価値関数等をクラウドサーバ上で共有及び更新することができる。

[0194] 上述した数値制御装置200及び機械学習装置400、500、700、800は、CPU (Central Processing Unit) 等の演算処理装置を備える。また、数値制御装置200及び機械学習装置400、500、700、800は、アプリケーションソフトウェアやOS (Operating System) 等の各種の制御用プログラムを格納したHDD (Hard Disk Drive) 等の補助記憶装置や、演算処理装置がプログラムを実行する上で一時的に必要とされるデータを格納するためのRAM (Random Access Memory) といった主記憶装置も備える。

[0195] そして、数値制御装置200及び機械学習装置400、500、700、800は、演算処理装置が補助記憶装置からアプリケーションソフトウェアやOSを読み込み、読み込んだアプリケーションソフトウェアやOSを主記憶装置に展開させながら、これらのアプリケーションソフトウェアやOSに基づいた演算処理を行う。また、数値制御装置200及び機械学習装置400、500、700、800は、この演算結果に基づいて、各装置が備える各種のハードウェアを制御する。これにより、本実施形態の機能ブロックは実現される。つまり、上述した実施形態は、ハードウェアとソフトウェアが

協働することにより実現することができる。

[0196] 具体例として、数値制御装置200は、一般的な工作機械100の制御装置に実施形態を実現するためのアプリケーションソフトウェアを組み込むことにより実現できる。また、機械学習装置400、500、700、800は、一般的なパーソナルコンピュータに、本実施形態を実現するためのアプリケーションソフトウェアを組み込むことにより実現できる。

[0197] ただし、機械学習装置400、500、700、800については、機械学習に伴う演算量が多いため、例えば、パーソナルコンピュータにGPU (Graphics Processing Units) を搭載し、GPGPU (General-Purpose computing on Graphics Processing Units) と呼ばれる技術により、GPUを機械学習に伴う演算処理に利用するようにすると、高速処理できるようになるのでよい。更には、より高速な処理を行うために、機械学習装置400、500は、このようなGPUを搭載したコンピュータを複数台用いてコンピュータ・クラスターを構築し、このコンピュータ・クラスターに含まれる複数のコンピュータにて並列処理を行うようにしてもよい。

[0198] 以上、本発明の実施形態について説明したが、上記の制御システム1、10、20、30は、ハードウェア、ソフトウェア又はこれらの組み合わせにより実現することができる。また、上記の制御システム1、10、20、30により行なわれる制御方法も、ハードウェア、ソフトウェア又はこれらの組み合わせにより実現することができる。ここで、ソフトウェアによって実現されるとは、コンピュータがプログラムを読み込んで実行することにより実現されることを意味する。

[0199] プログラムは、様々なタイプの非一時的なコンピュータ可読媒体 (non-transitory computer readable medium) を用いて格納され、コンピュータに供給することができる。非一時的なコンピュータ可読媒体は、様々なタイプの実体のある記録媒体 (tangible storage medium) を含む。非一時的なコンピュー

タ可読媒体の例は、磁気記録媒体（例えば、ハードディスクドライブ）、光磁気記録媒体（例えば、光磁気ディスク）、CD-ROM（Read Only Memory）、CD-R、CD-R/W、半導体メモリ（例えば、マスクROM、PROM（Programmable ROM）、EPROM（Erasable PROM）、フラッシュROM、RAM（random access memory））を含む。

[0200] また、上述した各実施形態は、本発明の好適な実施形態ではあるが、上記各実施形態のみに本発明の範囲を限定するものではない。本発明の要旨を逸脱しない範囲において種々の変更を施した形態での実施が可能である。

符号の説明

- [0201] 1 制御システム
- 10 制御システム
- 100 工作機械
- 200 数値制御装置
- 300 加工面解析装置
- 400 機械学習装置
- 410 設定条件取得部
- 420 ラベル取得部
- 430 学習部
- 440 学習モデル記憶部
- 450 揺動条件出力部
- 500 機械学習装置
- 510 設定条件取得部
- 520 判定情報取得部
- 530 行動情報出力部
- 540 学習部
- 550 価値関数記憶部
- 560 揺動条件出力部

請求の範囲

- [請求項1] 工具とワークと相対的に揺動させながら揺動加工する工作機械の揺動条件を学習する機械学習装置であって、
- 前記揺動加工のための設定条件を取得する設定条件取得部と、
- 前記工作機械による加工済みワークの評価データを、ラベルとして取得するラベル取得部と、
- 前記設定条件と前記ラベルとの組みを教師データとして教師あり学習を行う学習部と、を備え、
- 前記学習部は、前記加工済みワークの評価データを最適化する揺動条件を学習する学習モデルを備える、
- 機械学習装置。
- [請求項2] 前記学習モデルに基づいて、前記加工済みワークの評価データを最適化する最適揺動条件を出力する揺動条件出力部を更に備える、請求項1に記載の機械学習装置。
- [請求項3] 前記揺動条件出力部は、前記ワークの切屑が細断可能となる切屑細断揺動条件を算出する切屑細断条件算出部を備え、
- 前記揺動条件出力部は、前記学習モデルに基づいて、前記切屑細断揺動条件を満たす前記最適揺動条件を出力する、請求項2に記載の機械学習装置。
- [請求項4] 前記揺動条件出力部は、予め設定された上限値を超えない上限揺動条件を算出する揺動条件上限算出部を備え、
- 前記揺動条件出力部は、前記学習モデルに基づいて、前記上限揺動条件を満たす前記最適揺動条件を出力する、請求項2又は3に記載の機械学習装置。
- [請求項5] 前記加工済みワークの前記評価データは、前記加工済みワークの面粗さ、真円度又は寸法精度のデータを含む、請求項1から4のいずれか一項に記載の機械学習装置。
- [請求項6] 工具とワークと相対的に揺動させながら揺動加工する工作機械の揺

動条件を学習する機械学習装置であって、

前記揺動加工のための設定条件を、状態情報として取得する設定条件取得部と、

前記工作機械による加工済みワークの評価データを、判定情報として取得する判定情報取得部と、

現在の状態に対して前記揺動条件をどのように変更すべきかを示す行動情報を出力する行動情報出力部と、

前記判定情報に基づいて、強化学習における報酬の値を算出する報酬算出部と、

前記状態情報、前記行動情報及び前記報酬に基づいて、前記工作機械の揺動条件の価値を定める価値関数を更新する価値関数更新部と、

前記価値関数に基づいて、前記加工済みワークの評価データを最適化する最適揺動条件を出力する揺動条件出力部と、
を備える機械学習装置。

[請求項7] 前記報酬算出部は、前記加工済みワークの前記評価データが所定の閾値未満のときに、前記報酬を正の値とし、前記評価データが前記所定の閾値を超えたときに、前記報酬を負の値とする、請求項6に記載の機械学習装置。

[請求項8] 前記揺動条件出力部は、前記ワークの切屑が細断可能となる切屑細断揺動条件を算出する切屑細断条件算出部を備え、

前記揺動条件出力部は、前記切屑細断揺動条件を満たす前記最適揺動条件を出力する、請求項6又は7に記載の機械学習装置。

[請求項9] 前記揺動条件出力部は、予め設定された上限値を超えない上限揺動条件を算出する揺動条件上限算出部を備え、

前記揺動条件出力部は、前記上限揺動条件を満たす前記最適揺動条件を出力する、請求項6から8のいずれか一項に記載の機械学習装置。
。

[請求項10] 前記加工済みワークの前記評価データは、前記加工済みワークの面

粗さ、真円度又は寸法精度のデータを含む、請求項 6 から 9 のいずれか一項に記載の機械学習装置。

[請求項11] 前記評価データが前記加工済みワークの面粗さデータである場合、前記報酬算出部は、理論面粗さの近似式に基づいて前記所定の閾値を決定する、請求項 10 に記載の機械学習装置。

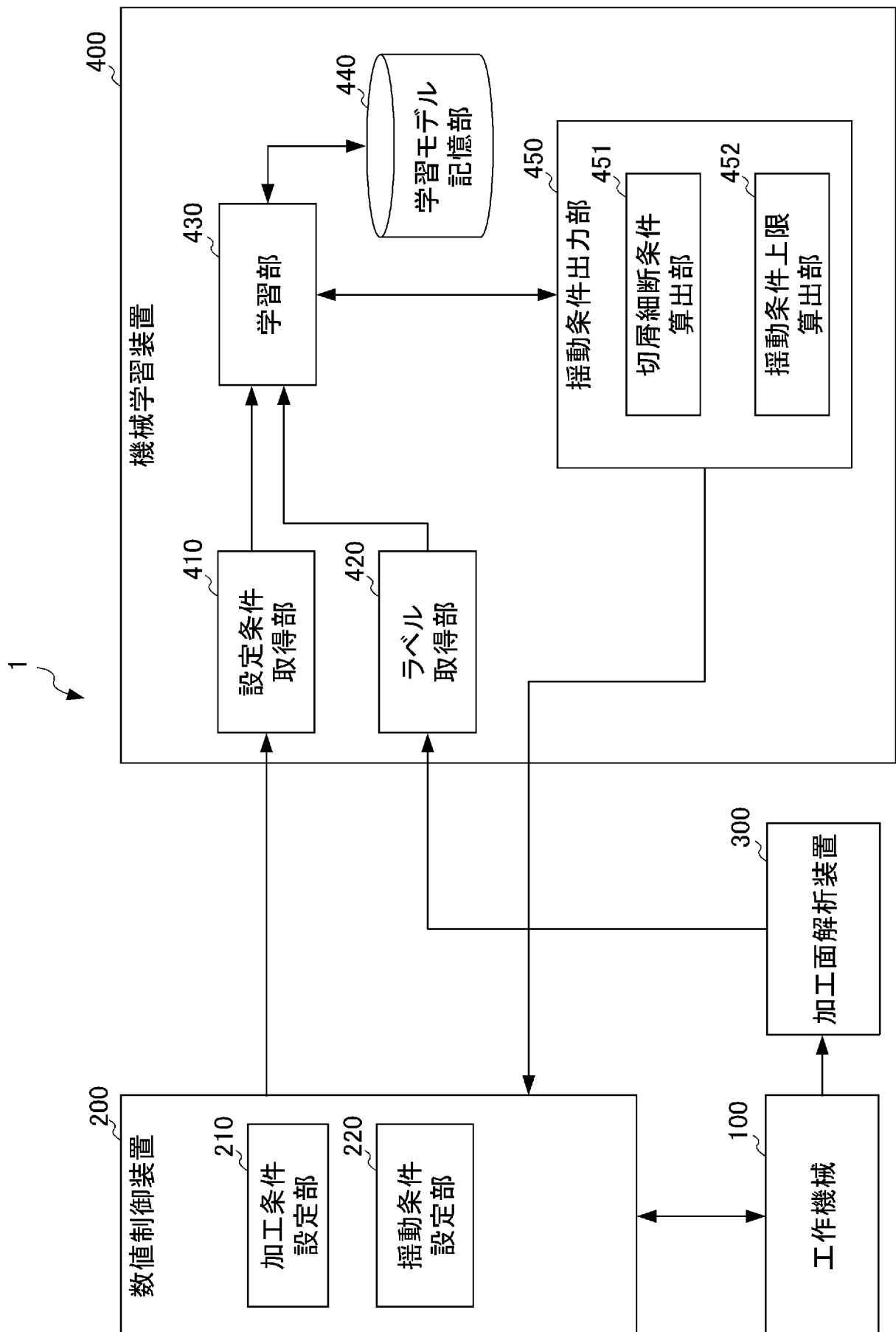
[請求項12] 前記面粗さデータは、算術平均粗さ、最大高さ、最大山高さ、最大谷深さ、平均高さ、最大断面高さ及び負荷長さ率の少なくとも 1 つを含む請求項 11 に記載の機械学習装置。

[請求項13] 前記設定条件は、
前記工作機械における前記工具の送り速度、主軸回転速度、座標値、前記工具の刃先、前記工具の種類及び前記ワークの材料の少なくとも 1 つを含む加工条件と、
前記工作機械の揺動周波数及び揺動振幅の少なくとも 1 つを含む前記揺動条件と、
を有する、請求項 1 から 12 のいずれか一項に記載の機械学習装置。

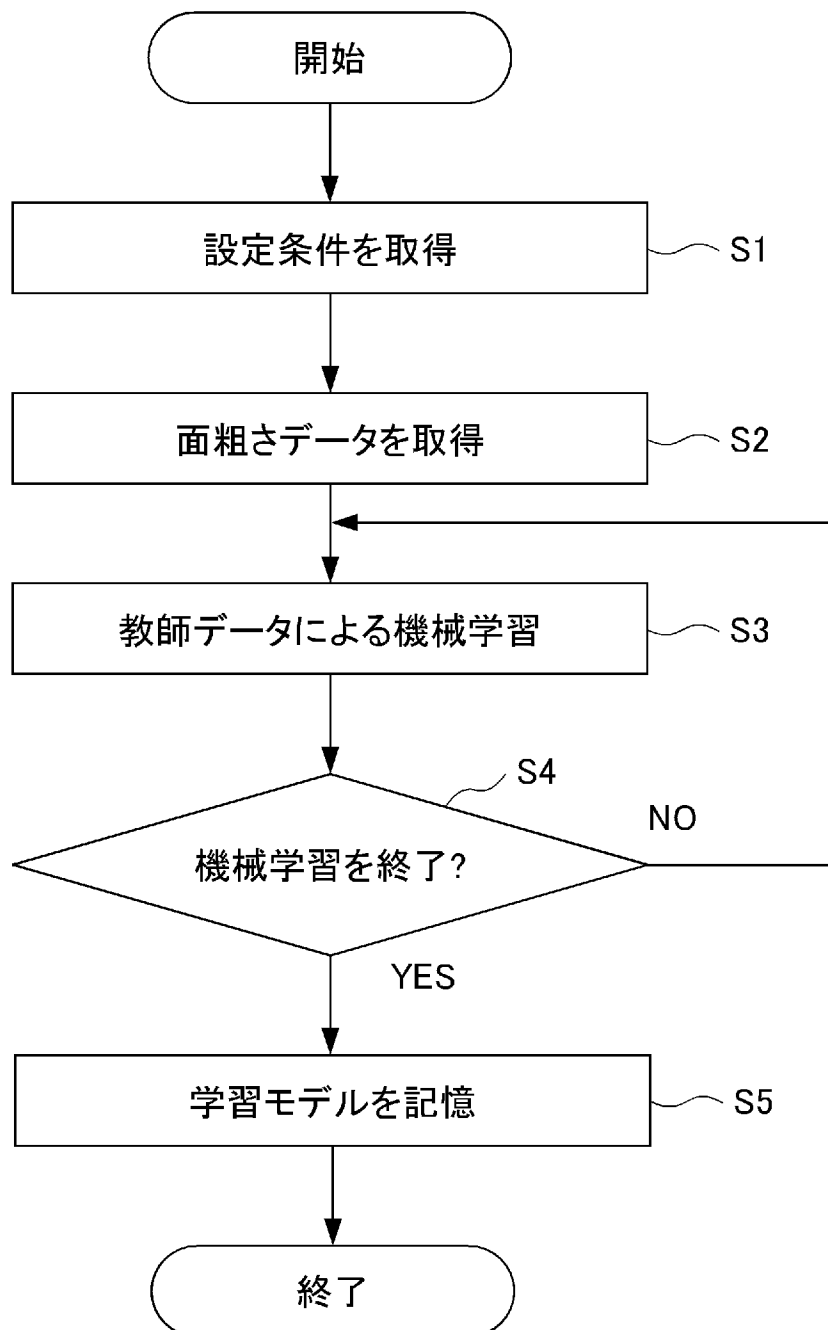
[請求項14] 前記機械学習装置は、複数の数値制御装置によって共有される、請求項 1 から 13 のいずれか一項に記載の機械学習装置。

[請求項15] 前記機械学習装置は、クラウドサーバ上に設けられる、請求項 1 から 13 のいずれか一項に記載の機械学習装置。

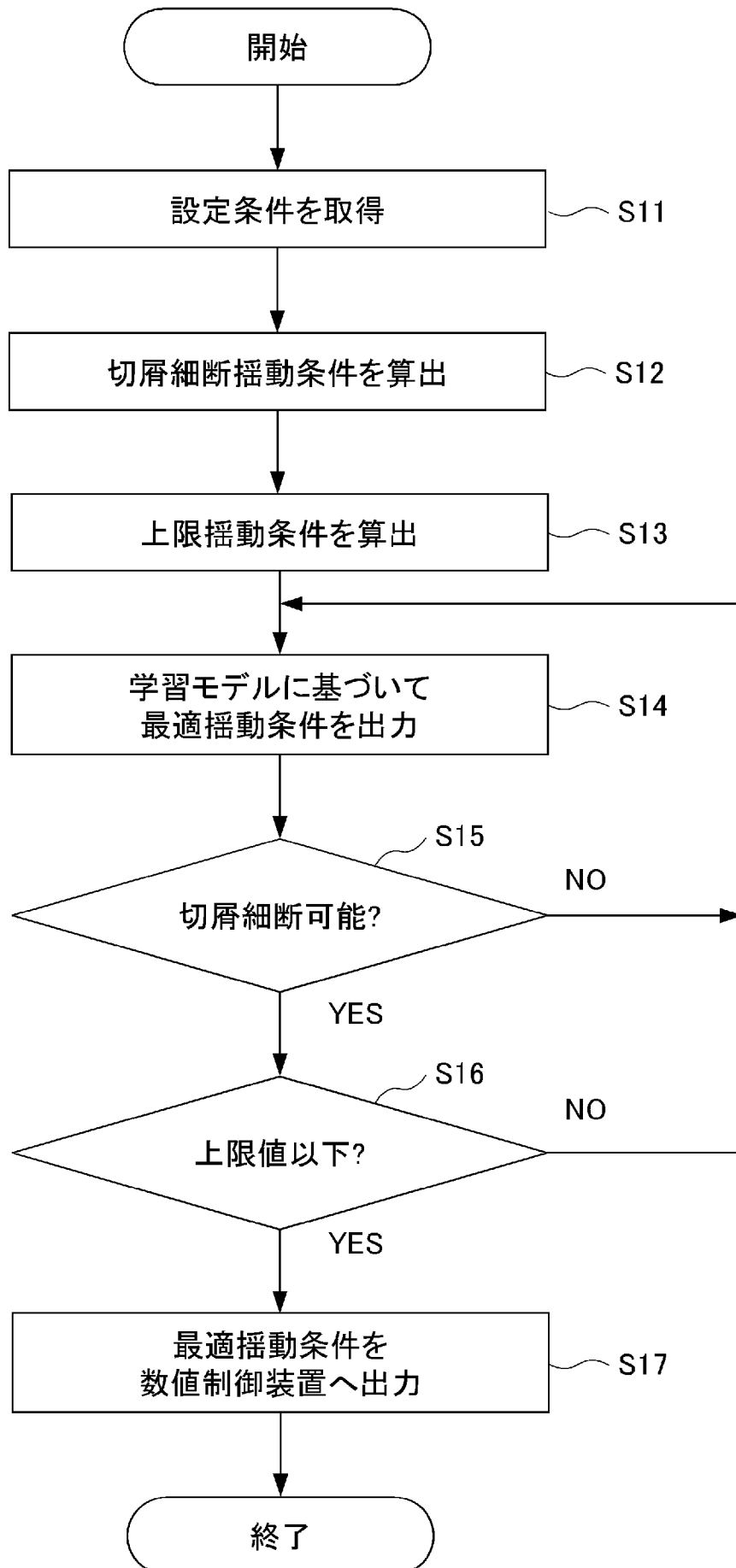
[図1]



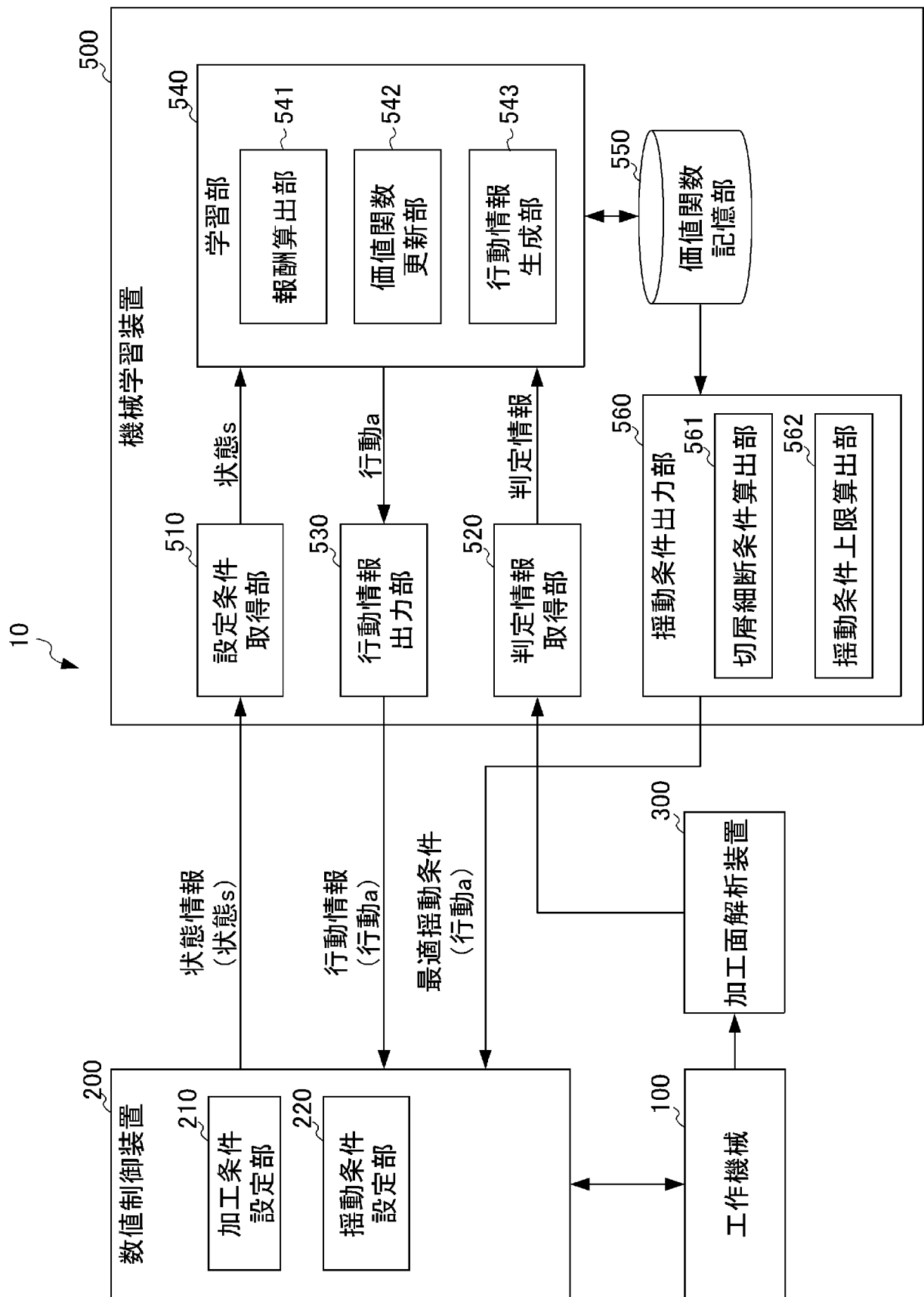
[図2]



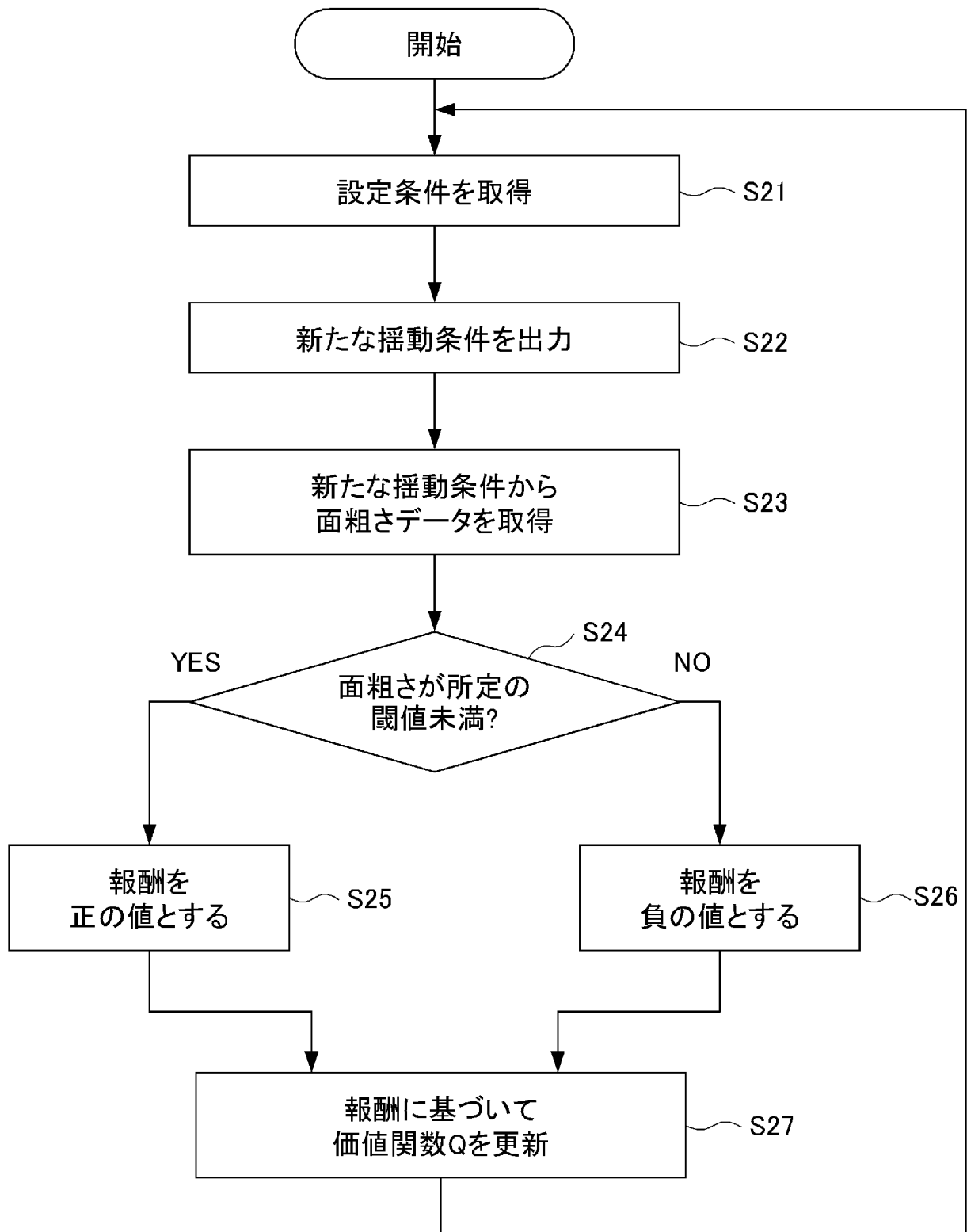
[図3]



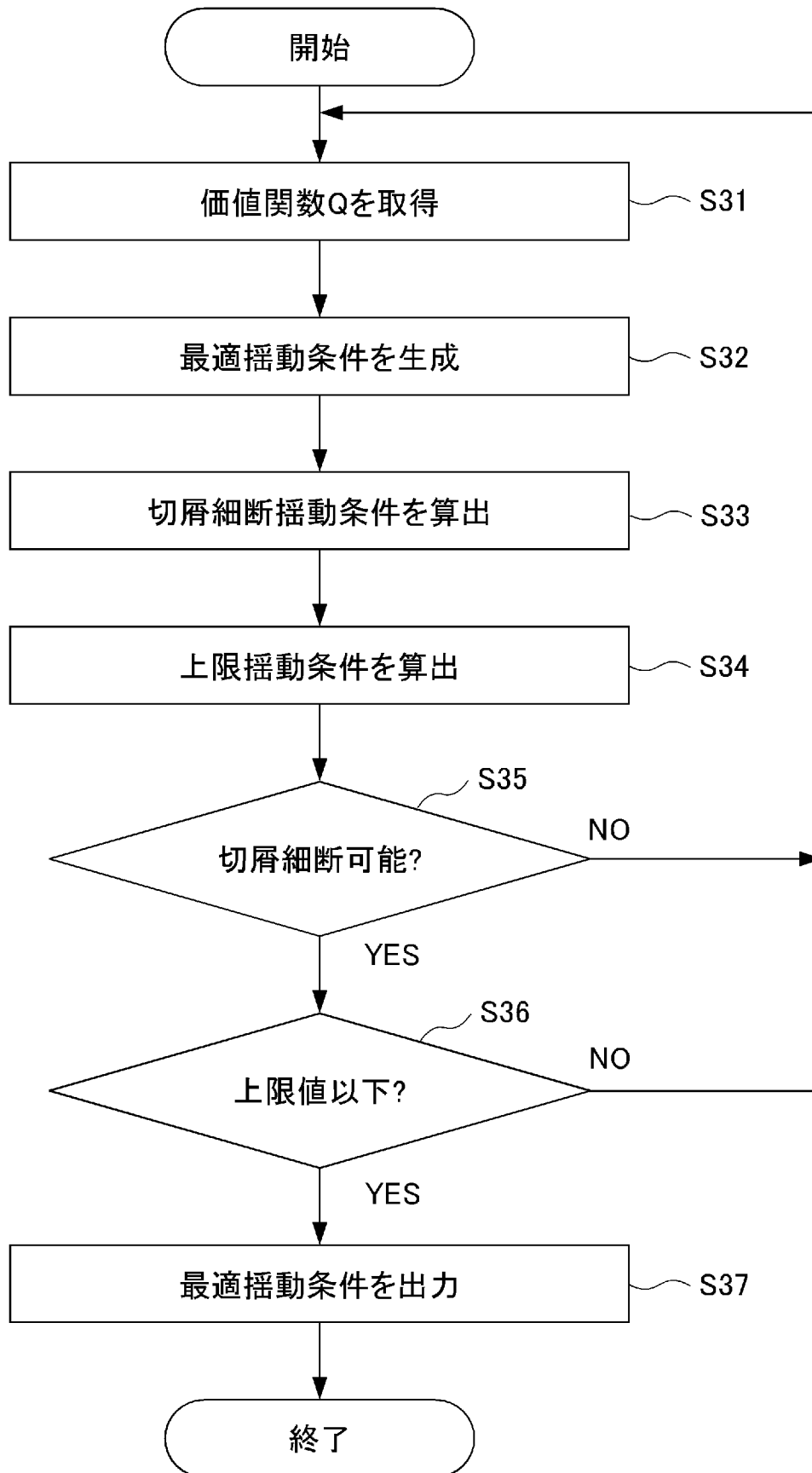
[図4]



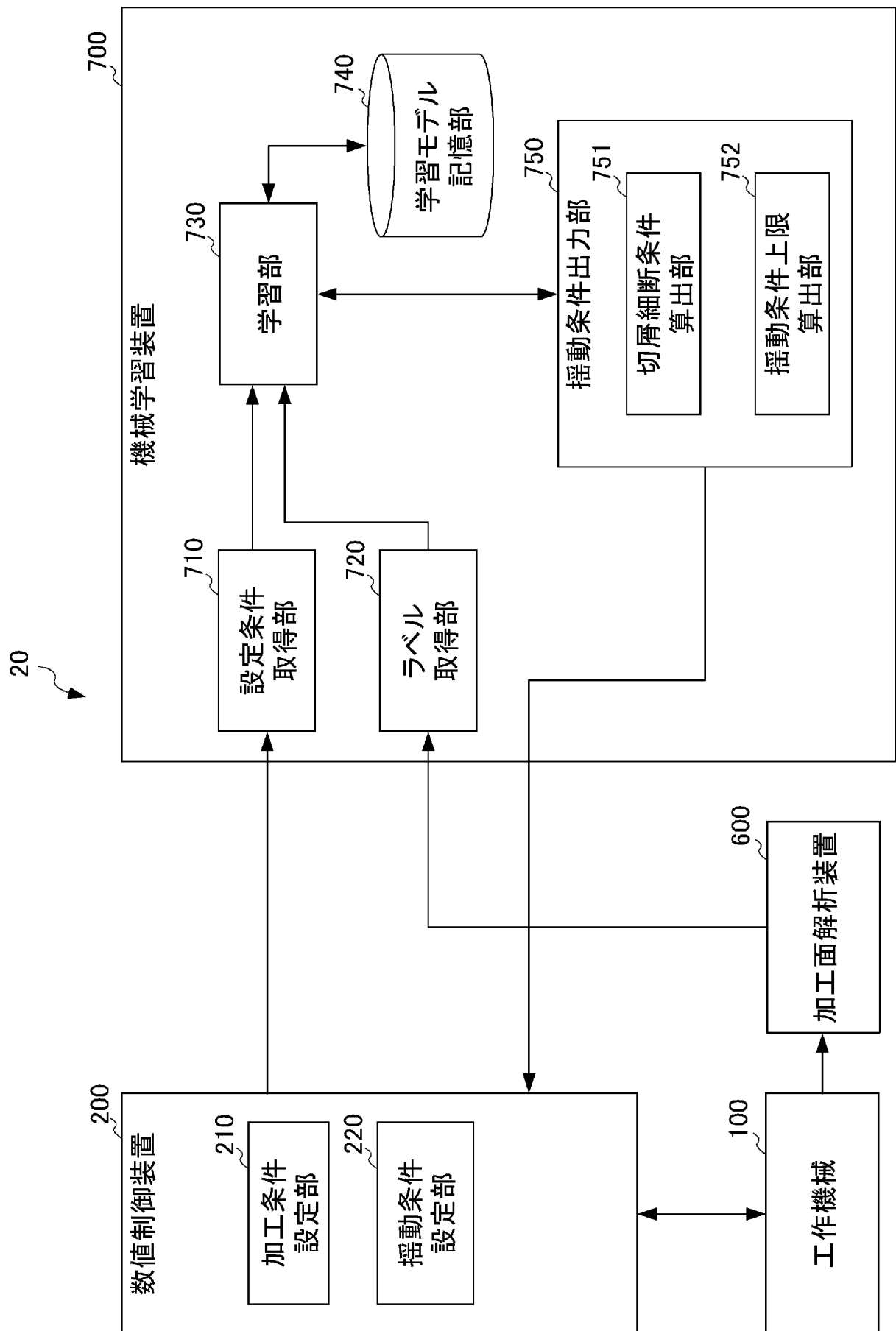
[図5]



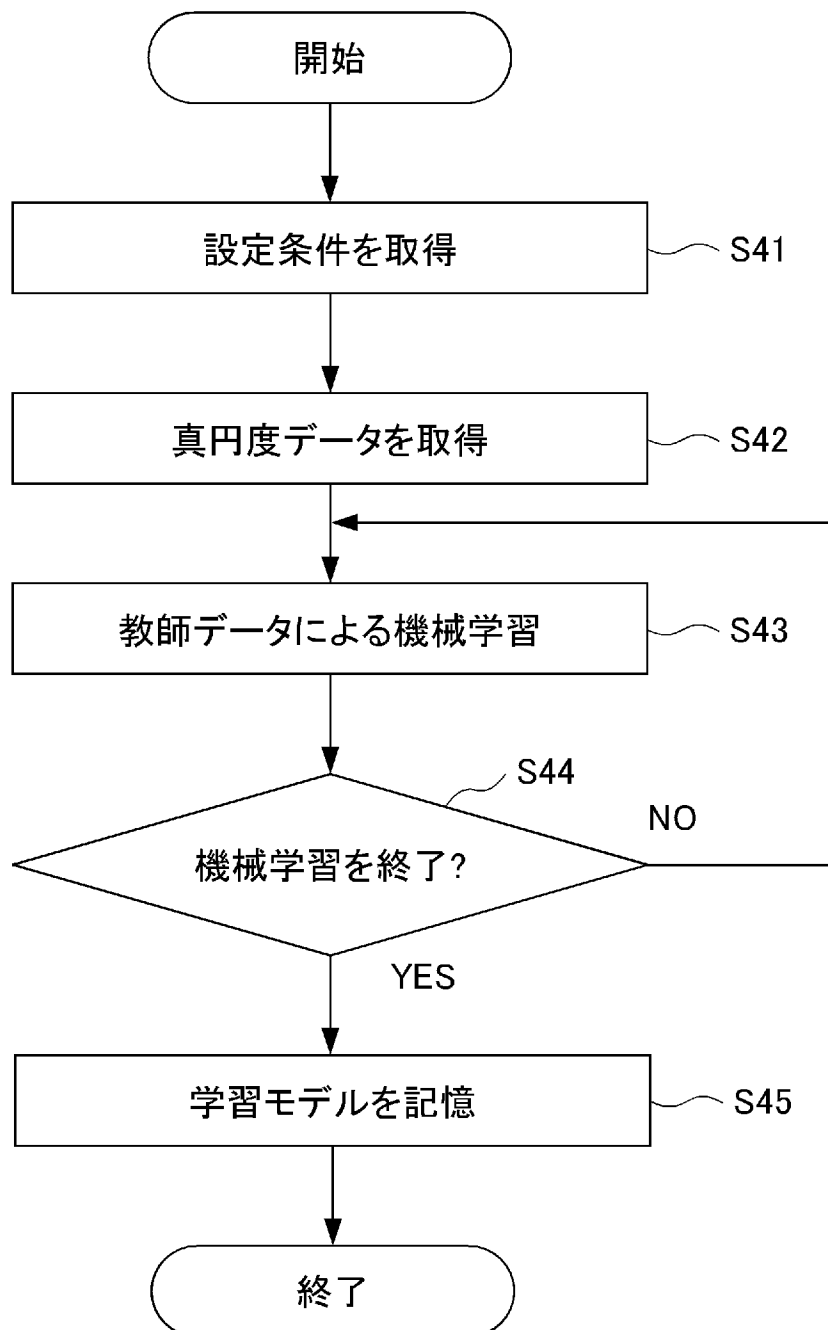
[図6]



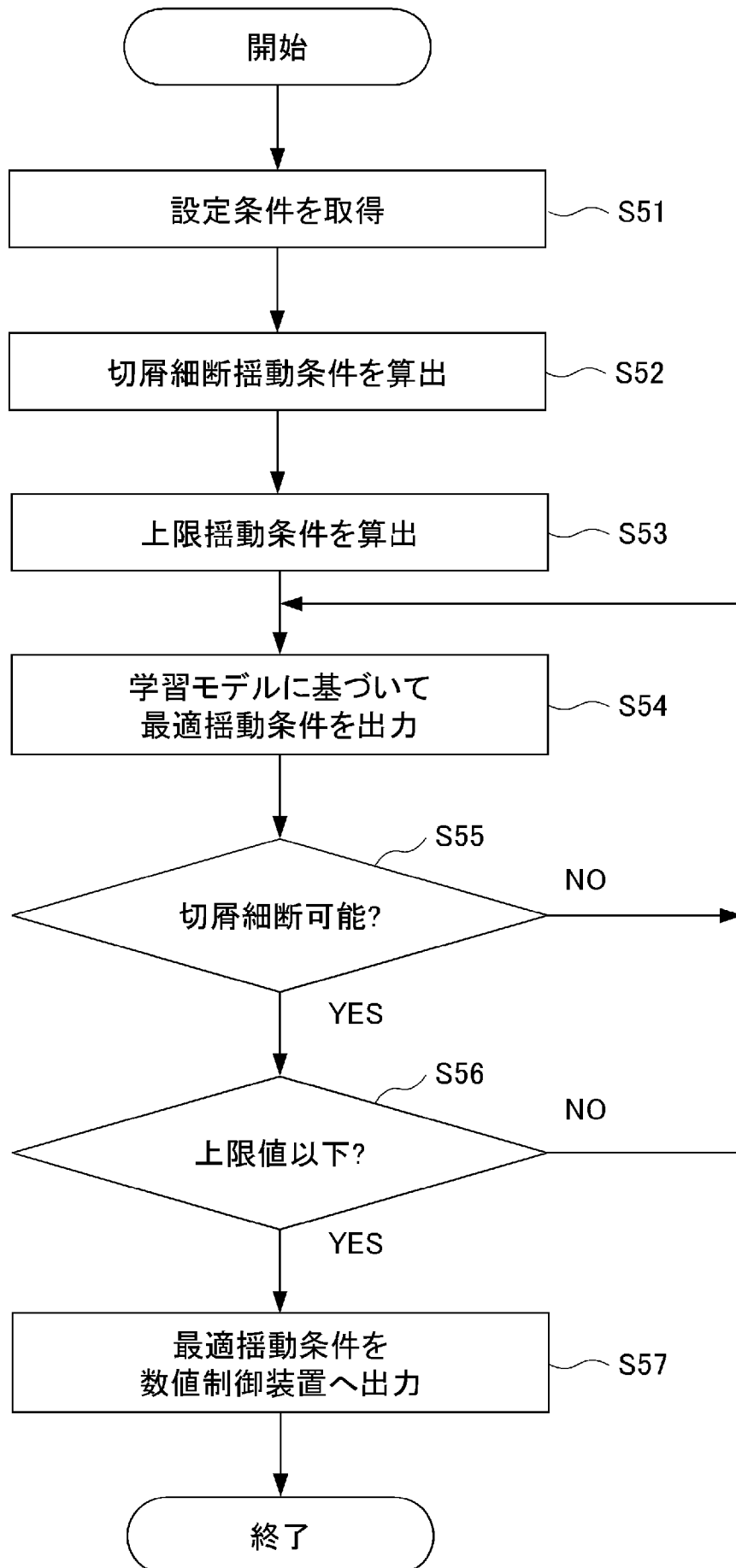
[図7]



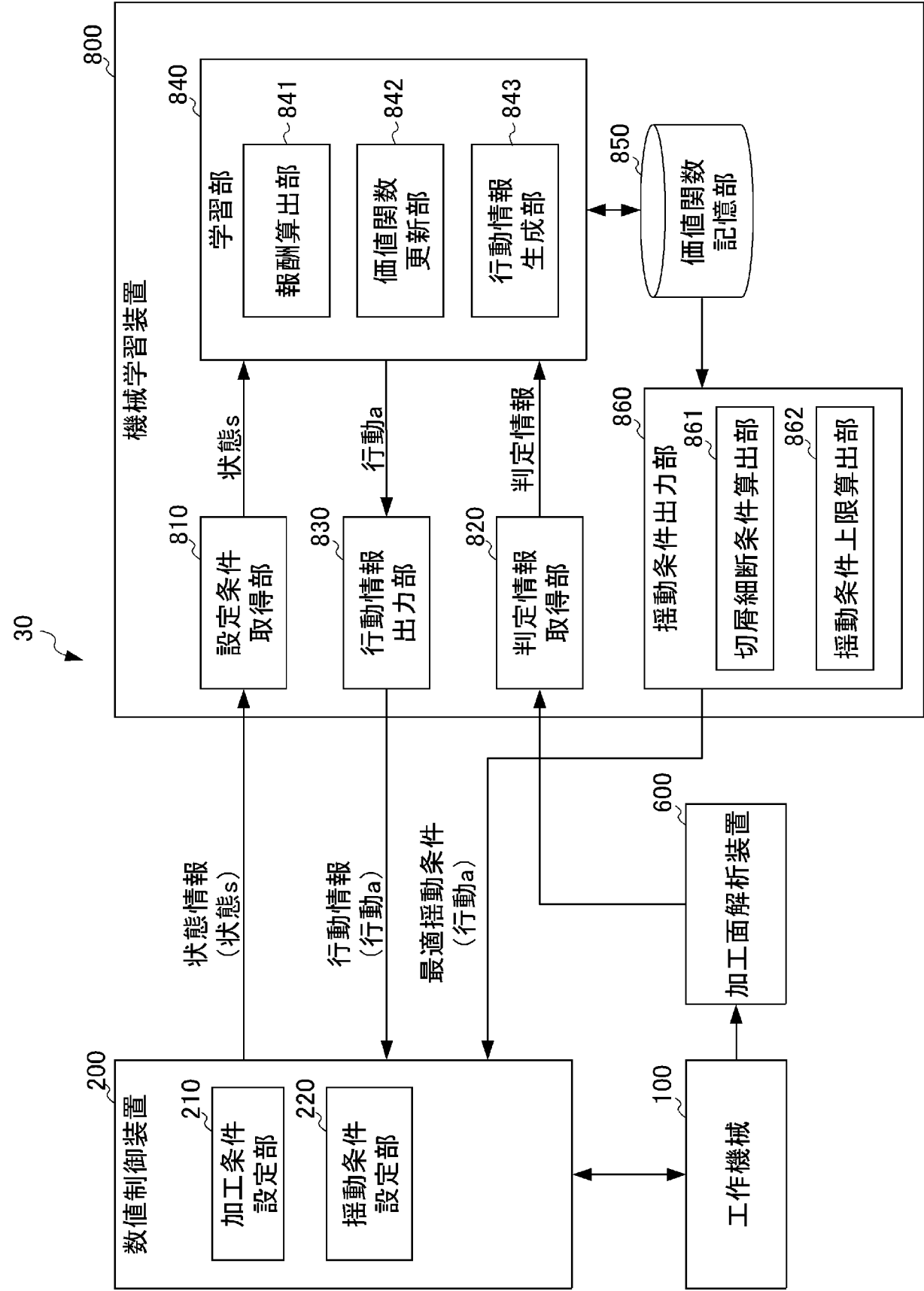
[図8]



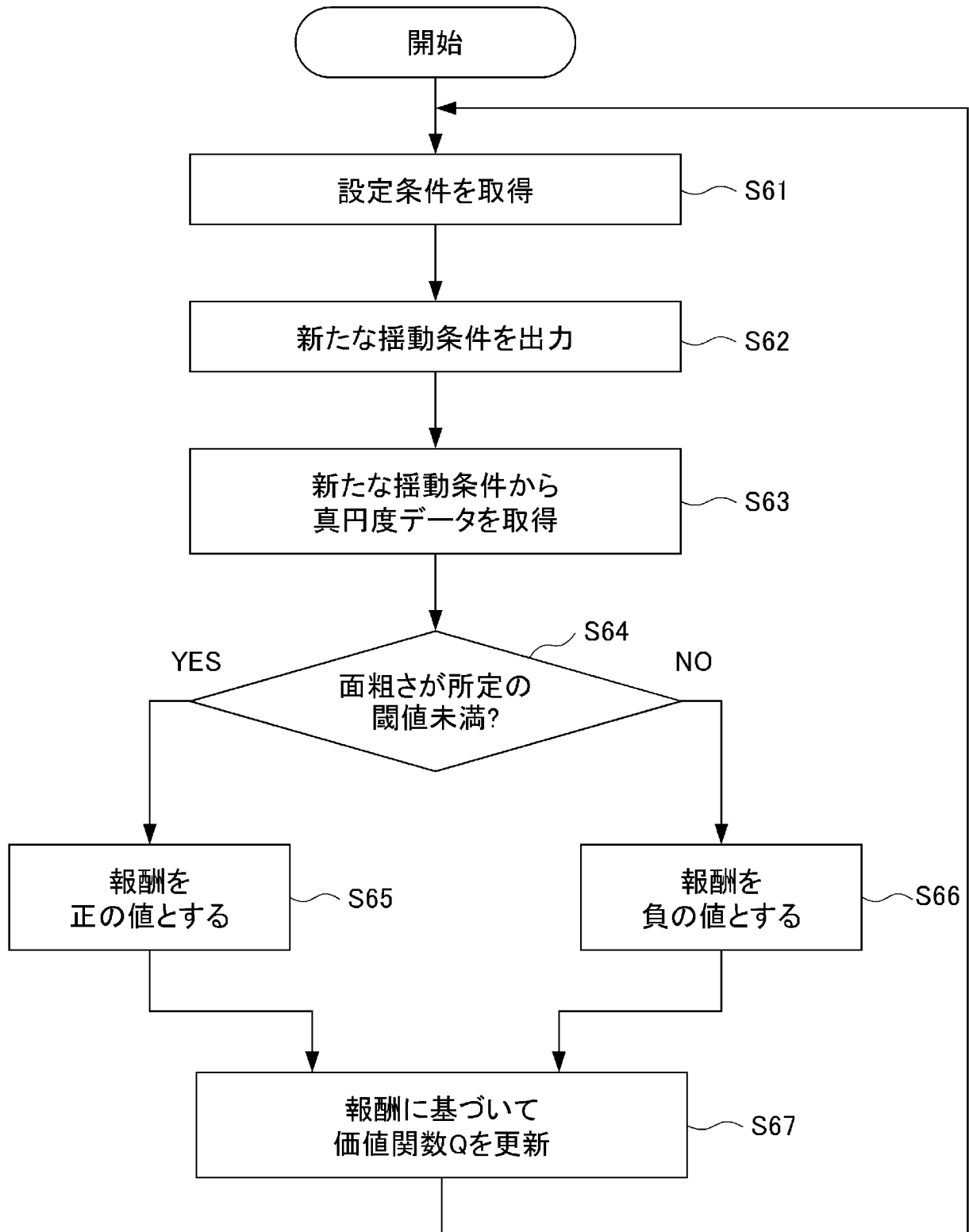
[図9]



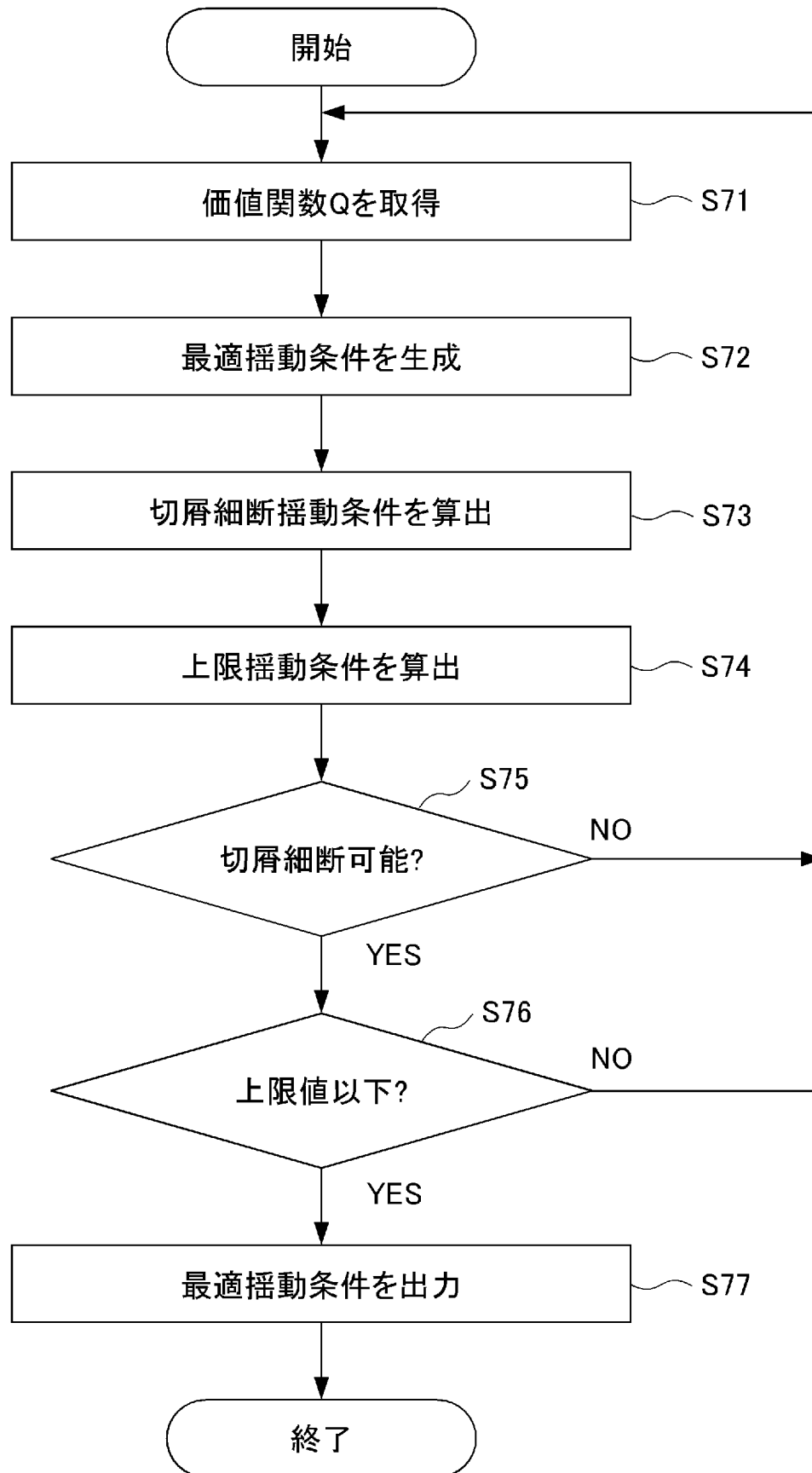
[図10]



[図11]



[図12]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/000336

A. CLASSIFICATION OF SUBJECT MATTER**G05B 19/4093**(2006.01)i; **G05B 19/4155**(2006.01)i; **G06N 20/00**(2019.01)i

FI: G06N20/00 130; G05B19/4155 V; G05B19/4093 M

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G05B19/4093; G05B19/4155; G06N20/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2022

Registered utility model specifications of Japan 1996-2022

Published registered utility model applications of Japan 1994-2022

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 6775720 B1 (MITSUBISHI ELECTRIC CORPORATION) 28 October 2020 (2020-10-28) paragraphs [0012]-[0081], fig. 1-4	1-15
A	JP 2018-181217 A (FANUC LTD) 15 November 2018 (2018-11-15) paragraphs [0019]-[0068], fig. 1-9	1-15
A	JP 2020-157425 A (JTEKT CORP) 01 October 2020 (2020-10-01) fig. 7-12	1-15

☐ Further documents are listed in the continuation of Box C.☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

25 February 2022

Date of mailing of the international search report

08 March 2022

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915
Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2022/000336

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	6775720	B1	28 October 2020	(Family: none)	
JP	2018-181217	A	15 November 2018	US 2018/0307211 A1 paragraphs [0029]-[0080], fig. 1-9	
				DE 102018108778 A1	
				CN 108723889 A	
JP	2020-157425	A	01 October 2020	US 2020/0306928 A1 fig. 7-12	
				CN 111745467 A	

A. 発明の属する分野の分類（国際特許分類（IPC）） G05B 19/4093(2006.01)i; G05B 19/4155(2006.01)i; G06N 20/00(2019.01)i FI: G06N20/00 130; G05B19/4155 V; G05B19/4093 M														
B. 調査を行った分野														
調査を行った最小限資料（国際特許分類（IPC）） G05B19/4093; G05B19/4155; G06N20/00														
最小限資料以外の資料で調査を行った分野に含まれるもの														
<table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2022年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2022年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2022年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2022年	日本国実用新案登録公報	1996 - 2022年	日本国登録実用新案公報	1994 - 2022年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2022年													
日本国実用新案登録公報	1996 - 2022年													
日本国登録実用新案公報	1994 - 2022年													
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）														
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
A	JP 6775720 B1（三菱電機株式会社）28.10.2020（2020 - 10 - 28） 段落[0012]-[0081], 図1-4	1-15												
A	JP 2018-181217 A（ファナック株式会社）15.11.2018（2018 - 11 - 15） 段落[0019]-[0068], 図1-9	1-15												
A	JP 2020-157425 A（株式会社ジェイテクト）01.10.2020（2020 - 10 - 01） 図7-12	1-15												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 25.02.2022	国際調査報告の発送日 08.03.2022													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 樋口 幸太郎 3U 5561 電話番号 03-3581-1101 内線 3364													

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2022/000336

引用文献			公表日	パテントファミリー文献	公表日
JP	6775720	B1	28.10.2020	(ファミリーなし)	
JP	2018-181217	A	15.11.2018	US 2018/0307211 A1	
				段落[0029]-[0080], 図1-9	
				DE 102018108778 A1	
				CN 108723889 A	
JP	2020-157425	A	01.10.2020	US 2020/0306928 A1	
				図7-12	
				CN 111745467 A	