

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6499710号
(P6499710)

(45) 発行日 平成31年4月10日(2019.4.10)

(24) 登録日 平成31年3月22日(2019.3.22)

(51) Int.Cl.		F I			
G05B 19/416	(2006.01)	G05B	19/416		K
G05B 19/4155	(2006.01)	G05B	19/4155		V
G06N 20/00	(2019.01)	G06N	99/00	150	
G06N 3/08	(2006.01)	G06N	3/08		

請求項の数 11 (全 17 頁)

(21) 出願番号	特願2017-84039 (P2017-84039)	(73) 特許権者	390008235
(22) 出願日	平成29年4月20日 (2017.4.20)		ファナック株式会社
(65) 公開番号	特開2018-181217 (P2018-181217A)		山梨県南都留郡忍野村忍草字古馬場358
(43) 公開日	平成30年11月15日 (2018.11.15)		〇番地
審査請求日	平成30年5月15日 (2018.5.15)	(74) 代理人	110001151
早期審査対象出願			あいわ特許業務法人
		(72) 発明者	相澤 誠彰
			山梨県南都留郡忍野村忍草字古馬場358
			〇番地 ファナック株式会社内
		(72) 発明者	萱沼 友一
			山梨県南都留郡忍野村忍草字古馬場358
			〇番地 ファナック株式会社内
		審査官	白井 卓巳
			最終頁に続く

(54) 【発明の名称】 加減速制御装置

(57) 【特許請求の範囲】

【請求項1】

ワークの加工を行う工作機械を制御する加減速制御装置であって、
前記工作機械の各軸の速度のN階時間微分要素（Nは自然数）を学習する機械学習装置を備え、

前記機械学習装置は、

前記各軸の速度のN階時間微分要素を示す第1の状態データを、環境の現在状態を表す状態変数として観測する状態観測部と、

加工済みワークの面品位及び加工時間の適否判定結果を示す判定データを取得する判定データ取得部と、

前記状態変数と前記判定データとを用いて、前記各軸の速度のN階時間微分要素を、前記加工済みワークの面品位及び加工時間と関連付けて学習する学習部と、

を備える加減速制御装置。

【請求項2】

前記状態観測部は、前記状態変数として、加工タイプを識別する第2の状態データを更に観測し、

前記学習部は、前記第1の状態データ及び前記第2の状態データを、前記加工済みワークの面品位及び加工時間と関連付けて学習する、

請求項1に記載の加減速制御装置。

【請求項3】

前記加工済みワークの面品位に関する前記判定データは、前記ワークの表面粗さ S_a 、表面の最大高さ S_v 、表面性状のアスペクト比 S_{tr} 、尖り度 S_{ku} 、偏り度 S_{sk} 、界面の展開面積比 S_{dr} 、光反射率、画像の特徴のうち、少なくとも1つを利用して得られた値を含む、

請求項1に記載の加減速制御装置。

【請求項4】

前記学習部は、

前記適否判定結果に関連する報酬を求める報酬計算部と、

前記報酬を用いて、前記各軸の速度のN階時間微分要素の価値を表す関数を更新する価値関数更新部とを備える、

10

請求項1に記載の加減速制御装置。

【請求項5】

前記学習部は、前記状態変数と前記判定データとを多層構造で演算する、

請求項1に記載の加減速制御装置。

【請求項6】

前記学習部による学習結果に基づいて、前記各軸の速度のN階時間微分要素に基づく指令値を出力する意思決定部を更に備える、

請求項1に記載の加減速制御装置。

【請求項7】

前記学習部は、複数の前記工作機械から得られた前記状態変数及び前記判定データを用いて、該工作機械における前記各軸の速度のN階時間微分要素を学習する、

20

請求項1に記載の加減速制御装置。

【請求項8】

前記機械学習装置は、クラウドサーバに存在する、

請求項1に記載の加減速制御装置。

【請求項9】

ワークの加工を行う工作機械の各軸の速度のN階時間微分要素を学習する機械学習装置であって、

前記各軸の速度のN階時間微分要素を示す第1の状態データを、環境の現在状態を表す状態変数として観測する状態観測部と、

30

加工済みワークの面品位及び加工時間の適否判定結果を示す判定データを取得する判定データ取得部と、

前記状態変数と前記判定データとを用いて、前記各軸の速度のN階時間微分要素を、前記加工済みワークの面品位及び加工時間と関連付けて学習する学習部と、

を備える機械学習装置。

【請求項10】

ワークの加工を行う工作機械を制御する加減速制御装置であって、

前記工作機械の各軸の速度のN階時間微分要素(Nは自然数)と、前記加工をしたワークの面品位及び加工時間と、の相関性を示すモデルと、

前記モデルに基づいて、前記各軸の速度のN階時間微分要素に基づく指令値を出力する意思決定部と、

40

を備える加減速制御装置。

【請求項11】

ワークの加工を行う工作機械の各軸の速度のN階時間微分要素を学習する機械学習方法であって、

前記各軸の速度のN階時間微分要素示す第1の状態データを、環境の現在状態を表す状態変数として観測するステップと、

加工済みワークの面品位及び加工時間の適否判定結果を示す判定データを取得するステップと、

前記状態変数と前記判定データとを用いて、前記各軸の速度のN階時間微分要素を、前

50

記加工済みワークの面品位及び加工時間と関連付けて学習するステップと、
を含む機械学習方法。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、加減速制御装置に関し、特に工作機械の各軸の制御における最適な加減速を生成する技術に関する。

【背景技術】

【0002】

従来、加工プログラムを作成し、該加工プログラムに基づいて工作機械を制御して、部品や金型などの加工が行われている。加工を行う際の加工速度は加工プログラム内で軸の移動速度として指令するが、これは工具と加工物の相対移動（工具移動）の最大速度であり、実際に工作機械に対しては、加工開始時やコーナー部、曲線部分などにおいて、各軸の加減速時定数に従い軸の移動速度を変動させた移動データが出力されていた。また、加工物には許容誤差及び目標加工時間があるが、これは工作機械の操作者が加工物の加工面品位を確認しながら、加減速時定数を変更したり、プログラム内で指令する速度を変更したりするなどの方法で調整していた。なお、本明細書においては、指令位置と実際の加工物の形状誤差を意味する加工精度、加工物の形状変化が滑らかである（加工面の傷や筋目が目立たず、光が均一に反射する）度合いを加工面品位とする。

【0003】

このような加工における加減速時定数や速度などのパラメータ調整に関連する従来技術として特許文献1がある。特許文献1には、工作機械の振動データを用いて機械学習を行うことにより、加工面品位と加工時間とのバランスの取れた最適な速度分布を求める方法が記載されている。これにより、工作機械を、加工面の不良が発生しない適切な加工速度で加工を実施するよう制御することが可能となる。

【先行技術文献】

【特許文献】

【0004】

【特許文献1】特願2015-189542

【発明の概要】

【発明が解決しようとする課題】

【0005】

図8に示すように、特許文献1記載の技術は、工作機械における機械振動データに基づいて加工速度を調整するというものである。しかしながら、この技術においては、最終的に評価されるべき加工ワークの加工面品位に関する情報は利用されていない。このため、特許文献1記載の技術で報酬判定対象とされる振動データのしきい値は、加工面品位が所望の状態となるような値となるようあらかじめ算出されていなければならない。このように、特許文献1記載技術では、テスト加工を行うなどして、加工面品位を考慮したしきい値をあらかじめ決定する手間が必要であった。

【0006】

本発明はこのような問題を解決することが可能であり、工作機械の各軸の制御における最適な加減速を生成することが可能な加減速制御装置を提供することである。

【課題を解決するための手段】

【0007】

本発明の一実施の形態における加減速制御装置は、ワークの加工を行う工作機械を制御する加減速制御装置であって、前記工作機械の各軸の速度のN階時間微分要素（Nは自然数）を学習する機械学習装置を備え、前記機械学習装置は、前記各軸の速度のN階時間微分要素を示す第1の状態データを、環境の現在状態を表す状態変数として観測する状態観測部と、加工済みワークの面品位及び加工時間の適否判定結果を示す判定データを取得する判定データ取得部と、前記状態変数と前記判定データとを用いて、前記各軸の速度のN

10

20

30

40

50

階時間微分要素を、前記加工済みワークの面品位及び加工時間と関連付けて学習する学習部と、を備える。

【0008】

本発明の一実施の形態における加減速制御装置は、前記状態観測部は、前記状態変数として、加工タイプを識別する第2の状態データを更に観測し、前記学習部は、前記第1の状態データ及び前記第2の状態データを、前記加工済みワークの面品位及び加工時間と関連付けて学習する。

【0009】

本発明の一実施の形態における加減速制御装置は、前記加工済みワークの面品位に関する前記判定データは、前記ワークの表面粗さ S_a 、表面の最大高さ S_v 、表面性状のアスペクト比 S_{tr} 、尖り度 S_{ku} 、偏り度 S_{sk} 、界面の展開面積比 S_{dr} 、光反射率、画像の特徴のうち、少なくとも1つを利用して得られた値を含む。

10

【0010】

本発明の一実施の形態における加減速制御装置は、前記学習部は、前記適否判定結果に関連する報酬を求める報酬計算部と、前記報酬を用いて、前記各軸の速度のN階時間微分要素の価値を表す関数を更新する価値関数更新部とを備える。

【0011】

本発明の一実施の形態における加減速制御装置は、前記学習部は、前記状態変数と前記判定データとを多層構造で演算する。

【0012】

本発明の一実施の形態における加減速制御装置は、前記学習部による学習結果に基づいて、前記各軸の速度のN階時間微分要素に基づく指令値を出力する意思決定部を更に備える。

20

【0013】

本発明の一実施の形態における加減速制御装置は、前記学習部は、複数の前記工作機械から得られた前記状態変数及び前記判定データを用いて、該工作機械における前記各軸の速度のN階時間微分要素を学習する。

【0014】

本発明の一実施の形態における加減速制御装置は、前記機械学習装置は、クラウドサーバに存在する。

30

【0015】

本発明の一実施の形態における機械学習装置は、ワークの加工を行う工作機械の各軸の速度のN階時間微分要素を学習する機械学習装置であって、前記各軸の速度のN階時間微分要素を示す第1の状態データを、環境の現在状態を表す状態変数として観測する状態観測部と、加工済みワークの面品位及び加工時間の適否判定結果を示す判定データを取得する判定データ取得部と、前記状態変数と前記判定データとを用いて、前記各軸の速度のN階時間微分要素を、前記加工済みワークの面品位及び加工時間と関連付けて学習する学習部と、を備える。

【0016】

本発明の一実施の形態における機械学習方法は、ワークの加工を行う工作機械の各軸の速度のN階時間微分要素を学習する機械学習方法であって、前記各軸の速度のN階時間微分要素を示す第1の状態データを、環境の現在状態を表す状態変数として観測するステップと、加工済みワークの面品位及び加工時間の適否判定結果を示す判定データを取得するステップと、前記状態変数と前記判定データとを用いて、前記各軸の速度のN階時間微分要素を、前記加工済みワークの面品位及び加工時間と関連付けて学習するステップと、を含む。

40

【発明の効果】

【0017】

本発明により、工作機械の各軸の制御における最適な加減速を生成することが可能な加減速制御装置を提供することができる。

50

【図面の簡単な説明】

【0018】

【図1】第1の実施形態による加減速制御装置の概略的な機能ブロック図である。

【図2】制御装置の一形態を示す概略的な機能ブロック図である。

【図3A】機械学習方法の一形態を示す概略的なフローチャートである。

【図3B】機械学習方法の一形態を示す概略的なフローチャートである。

【図4A】ニューロンを説明する図である。

【図4B】ニューラルネットワークを説明する図である。

【図5】第2の実施形態による加減速制御装置の概略的な機能ブロック図である。

【図6】加減速制御システムの一形態を示す概略的な機能ブロック図である。

10

【図7】加減速制御システムの他の形態を示す概略的な機能ブロック図である。

【図8】従来技術による加減速制御を説明する図である。

【図9】本発明による加減速制御を説明する図である。

【発明を実施するための形態】

【0019】

本発明は、図9に示すように、加工ワーク加工面の情報に基づいて最適な加減速制御を行うため、機械学習の手法を適用することを特徴とする。これにより、要求される加工面質に応じて速度分布（加速度や加加速度）を決定することができるので、しきい値の事前設定などの必要がない。また、加工前の段階で加工面質をコントロールできる、すなわち仕上げ加工の際にどこでどのような加減速をするかを自動的に調整できるので、研磨の工程を省略又は短縮できる。以下、本発明の実施形態を図面と共に説明する。

20

【0020】

本発明では、加工プログラムに基づくワークの加工における工作機械の各軸の加減速度調整を目的とした移動量の決定に関する機械学習を行うことで、ワークの加工における工作機械の各軸の速度分布が最適なものになるように決定する。ここで各軸の速度分布とは速度のN階時間微分要素（Nは任意の自然数）のことをいい、例えば各軸の加速度や加加速度である。各軸の速度分布（速度のN階時間微分要素）の決定においては、より速い工具移動と、加工精度や加工面品位がより良くなることを目的とし、より短時間で加工精度や加工面品位を維持したワークの加工を実現する。

【0021】

30

図1は、第1の実施形態による加減速制御装置10の概略的な機能ブロック図である。加減速制御装置10は、例えば、工作機械の各軸の加減速を制御する機能を有する数値制御装置（CNC）として実装することができる。加減速制御装置10は、加工精度や加工面品位がより良く、かつ加工時間がより短くなる各軸の速度分布（速度のN階時間微分要素）を、いわゆる機械学習により自ら学習するためのソフトウェア（学習アルゴリズム等）及びハードウェア（コンピュータのCPU等）を含む機械学習装置20を備える。加減速制御装置10が備える機械学習装置20が学習する各軸の速度分布（速度のN階時間微分要素）は、加工ワークの加工精度、面品位及び加工時間に係る状態を示す情報と、当該加工精度、面品位及び加工時間における速度分布（速度のN階時間微分要素）との、相関性を表すモデル構造に相当する。

40

【0022】

図1に機能ブロックで示すように、加減速制御装置10が備える機械学習装置20は、各軸の速度分布（速度のN階時間微分要素）を表す状態変数Sを観測する状態観測部22と、設定された状態変数Sの元での加工ワークの面品位を表す判定データD1、及び加工時間を表す判定データD2を含む判定データDを取得する判定データ取得部24と、状態変数Sと判定データDとを用いて、最適な各軸の速度分布（速度のN階時間微分要素）を学習する学習部26とを備える。

【0023】

状態観測部22は、例えばコンピュータのCPUの一機能として構成できる。或いは状態観測部22は、例えばコンピュータのCPUを機能させるためのソフトウェアとして構

50

成できる。状態観測部 22 が観測する状態変数 S は、例えば、加減速制御装置 10 が加工プログラムを解析することで得られる各軸の速度分布（速度の N 階時間微分要素） S_1 や、加減速制御装置 10 が実測した各軸の速度分布（速度の N 階時間微分要素） S_1 などを含む。

【0024】

判定データ取得部 24 は、例えばコンピュータの CPU の一機能として構成できる。或いは判定データ取得部 24 は、例えばコンピュータの CPU を機能させるためのソフトウェアとして構成できる。判定データ取得部 24 が取得する判定データ D_1 は、加工面の検査結果を示す数値データであって、例えば検査装置（図示しない）又は検査装置に付設されるセンサから得られるデータや、該データを利用乃至変換して得られるデータを用いることができる。検査装置には、例えば加工面解析装置（典型的にはレーザ顕微鏡）、加工面画像撮影装置、光反射率測定装置などがある。検査装置により計測可能な、面品位を表すデータとしては、例えば表面粗さ S_a 、表面の最大高さ S_v 、表面性状のアスペクト比 S_{tr} 、尖り度 S_{ku} 、偏り度 S_{sk} 、界面の展開面積比 S_{dr} 、加工ワークの光反射率、加工面の画像の特徴などがある。あるいは、判定データ D_1 は、熟練の作業者による面品位の評価結果を記録したファイルの入力、面品位の評価結果のキーボード等のインターフェイスからの直接入力により得られるデータや、該データを利用乃至変換して得られるデータを用いても良い。判定データ取得部 24 が取得する判定データ D_2 は、例えば加減速制御装置 10 が実測した加工時間に関するデータや、該データを利用乃至変換して得られるデータを用いることができる。

【0025】

このように、加減速制御装置 10 が備える機械学習装置 20 における学習は、工作機械における所定の速度分布（速度の N 階時間微分要素）での加工、加工ワークの面品位の測定、及び加工時間の測定を 1 サイクルとして実施される。

【0026】

学習部 26 は、例えばコンピュータの CPU の一機能として構成できる。或いは学習部 26 は、例えばコンピュータの CPU を機能させるためのソフトウェアとして構成できる。学習部 26 は、機械学習と総称される任意の学習アルゴリズムに従い、最適な各軸の速度分布（速度の N 階時間微分要素）を学習する。学習部 26 は、前述した状態変数 S と判定データ D とを含むデータ集合に基づく学習を反復実行することができる。学習サイクルの反復中、状態変数 S は、前回までの学習サイクルで得た各軸の速度分布（速度の N 階時間微分要素）とし、また判定データ D は、当該決定した各軸の速度分布（速度の N 階時間微分要素）に基づく加工ワークの面品位及び加工時間の評価結果とする。

【0027】

このような学習サイクルを繰り返すことにより、学習部 26 は、各軸の速度分布（速度の N 階時間微分要素）と加工ワークの面品位及び加工時間との相関性を暗示する特徴を自動的に識別することができる。各軸の速度分布（速度の N 階時間微分要素）と加工ワークの面品位及び加工時間との相関性は実質的に未知であるが、学習部 26 は、学習を進めるに従い徐々に特徴を識別して相関性を解釈する。各軸の速度分布（速度の N 階時間微分要素）と加工ワークの面品位及び加工時間との相関性が、ある程度信頼できる水準まで解釈されると、学習部 26 が反復出力する学習結果は、現在状態（つまり各軸の速度分布（速度の N 階時間微分要素））に対してどのような加工ワークの面品位及び加工時間が導かれるべきであるかという行動の選択（つまり意思決定）を行うために使用できるものとなる。つまり学習部 26 は、学習アルゴリズムの進行に伴い、各軸の速度分布（速度の N 階時間微分要素）と、当該状態から導かれる加工ワークの面品位及び加工時間という行動との、相関性を最適解に徐々に近づけることができる。

【0028】

上記したように、加減速制御装置 10 が備える機械学習装置 20 は、状態観測部 22 が観測した状態変数 S と判定データ取得部 24 が取得した判定データ D とを用いて、学習部 26 が機械学習アルゴリズムに従い、各軸の速度分布（速度の N 階時間微分要素）を学習

10

20

30

40

50

するものである。状態変数 S は、各軸の速度分布（速度の N 階時間微分要素）といった、外乱の影響を受け難いデータで構成され、また判定データ D は、検査装置による検査結果や加工時間の実測値を取得したりすることにより一義的に求められる。したがって、加減速制御装置 10 が備える機械学習装置 20 によれば、学習部 26 の学習結果を用いることで、各軸の速度分布（速度の N 階時間微分要素）に応じた、加工ワークの面品位及び加工時間を、演算や目算によらずに自動的に、しかも正確に求めることができるようになる。

【0029】

加工ワークの面品位及び加工時間を、演算や目算によらずに自動的に求めることができれば、加工を開始する前に各軸の速度分布（速度の N 階時間微分要素）の設定値を把握するだけで、加工ワークの面品位及び加工時間を迅速に推定することができる。

10

【0030】

加減速制御装置 10 が備える機械学習装置 20 の一変形例として、状態観測部 22 は、状態変数 S として、加工経路の形状などを表す加工タイプ S_2 を更に観測することができる。加工タイプ S_2 は例えば、加工経路の形状（直線部、コーナー部、 R 加工部、同心円状の加工部などを特定するための識別データ）を含むことができる。加工タイプ S_2 はさらに、直線部以外についてはコーナー等の大きさを示すデータ（角度、半径など）を含んでいても良い。角度や半径は、予め複数の段階を定義しておき、加工経路に含まれる角度や半径がどの段階に当たるかを示す識別データを含めるようにしても良い。各軸の速度分布（速度の N 階時間微分要素）を変化させる必要が生ずるのは一般に直線部以外のコーナー部などであるが、コーナー部などの形状によって、最適な各軸の速度分布（速度の N 階時間微分要素）は異なる可能性がある。加工タイプ S_2 を観測するならば、学習部 26 は、加工ワークの面品位及び加工時間を、各軸の速度分布（速度の N 階時間微分要素） S_1 及び加工タイプ S_2 の双方と関連付けて学習することができる。つまり、面品位及び加工時間と、各軸の速度分布（速度の N 階時間微分要素）との相関性を示すモデルを、加工タイプ S_2 ごとに独立に構築することが可能である。したがって、コーナー部などの形状に応じた、最適な各軸の速度分布（速度の N 階時間微分要素）を学習できる。

20

【0031】

加減速制御装置 10 が備える機械学習装置 20 の他の変形例として、学習部 26 は、同一の構成を有する複数の工作機械のそれぞれについて得られた状態変数 S 及び判定データ D を用いて、それら工作機械における各軸の速度分布（速度の N 階時間微分要素）に対応する加工ワークの面品位及び加工時間を学習することができる。この構成によれば、一定時間で得られる状態変数 S と判定データ D とを含むデータ集合の量を増加できるので、より多様なデータ集合を入力として、各軸の速度分布（速度の N 階時間微分要素）に対応する加工ワークの面品位及び加工時間の学習の速度や信頼性を向上させることができる。

30

【0032】

上記構成を有する機械学習装置 20 では、学習部 26 が実行する学習アルゴリズムは特に限定されず、機械学習として公知の学習アルゴリズムを採用できる。図 2 は、図 1 に示す加減速制御装置 10 の一形態であって、学習アルゴリズムの一例として強化学習を実行する学習部 26 を備えた構成を示す。強化学習は、学習対象が存在する環境の現在状態（つまり入力）を観測するとともに現在状態で所定の行動（つまり出力）を実行し、その行動に対し何らかの報酬を与えるというサイクルを試行錯誤的に反復して、報酬の総計が最大化されるような方策（本願の機械学習装置では加工ワークの面品位がより良く、かつ加工時間より短くなるような各軸の速度分布（速度の N 階時間微分要素））を最適解として学習する手法である。

40

【0033】

図 2 に示す加減速制御装置 10 が備える機械学習装置 20 において、学習部 26 は、ある状態変数 S に基づいて行われた加工の結果（加工ワークの面品位及び加工時間を示す判定データ D ）に関連する報酬 R を求める報酬計算部 28 と、報酬 R を用いて、各軸の速度分布（速度の N 階時間微分要素）の値を表す関数 Q を更新する価値関数更新部 30 とを備える。学習部 26 は、価値関数更新部 30 が関数 Q の更新を繰り返すことによって、加

50

工ワークの面品位がより良く、かつ加工時間より短くなるような各軸の速度分布（速度のN階時間微分要素）を学習する。

【0034】

学習部26が実行する強化学習のアルゴリズムの一例を説明する。この例によるアルゴリズムは、Q学習（Q-learning）として知られるものであって、行動主体の状態sと、その状態sで行動主体が選択し得る行動aとを独立変数として、状態sで行動aを選択した場合の行動の価値を表す関数 $Q(s, a)$ を学習する手法である。状態sで価値関数Qが最も高くなる行動aを選択することが最適解となる。状態sと行動aとの相関性が未知の状態sでQ学習を開始し、任意の状態sで種々の行動aを選択する試行錯誤を繰り返すことで、価値関数Qを反復して更新し、最適解に近付ける。ここで、状態sで行動aを選択した結果として環境（つまり状態s）が変化したときに、その変化に応じた報酬（つまり行動aの重み付け）rが得られるように構成し、より高い報酬rが得られる行動aを選択するように学習を誘導することで、価値関数Qを比較的短時間で最適解に近付けることができる。

10

【0035】

価値関数Qの更新式は、一般に下記の数1式のように表すことができる。数1式において、 s_t 及び a_t はそれぞれ時刻tにおける状態及び行動であり、行動 a_t により状態は s_{t+1} に変化する。 r_{t+1} は、状態が s_t から s_{t+1} に変化したことで得られる報酬である。 $\max_a Q$ の項は、時刻t+1で最大の価値Qになる（と時刻tで考えられている）行動aを行ったときのQを意味する。及び α はそれぞれ学習係数及び割引率であり、 $0 < \alpha < 1$ 、 $0 < \gamma < 1$ で任意設定される。

20

【0036】

【数1】

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

【0037】

学習部26がQ学習を実行する場合、状態観測部22が観測した状態変数S及び判定データ取得部24が取得した判定データDは、更新式の状態sに該当し、現在状態（加工ワークの面品位及び加工時間）に対して各軸の速度分布（速度のN階時間微分要素）をどのように変更するべきかという行動は、更新式の行動aに該当し、報酬計算部28が求める報酬Rは、更新式の報酬rに該当する。よって価値関数更新部30は、各軸の速度分布（速度のN階時間微分要素）の価値を表す関数Qを、報酬Rを用いたQ学習により繰り返し更新する。

30

【0038】

報酬計算部28が求める報酬Rは、例えば、各軸の速度分布（速度のN階時間微分要素）を決定した後該速度分布に基づいてワークの加工を実施したときに、加工ワークの面品位及び加工時間が「適」と判定される場合（例えば、ワークの面品位を表す指標が所定のしきい値以上である場合、加工時間が所定のしきい値以下である場合など）に正（プラス）の報酬Rとし、加工ワークの面品位及び加工時間が「否」と判定される場合（例えば、ワークの面品位を表す指標が所定のしきい値未満である場合、加工時間が所定のしきい値を超える場合など）に負（マイナス）の報酬Rとすることができる。正負の報酬Rの絶対値は、互いに同一であってもよいし異なってもよい。また、判定の条件として、判定データDに含まれる複数の値を組み合わせる判定するようにしても良い。ここで、しきい値は任意に設定することとしても良く、判定データDの累積的な平均値などの統計値を用いることとしても良い。あるいは、学習の初期段階は、加工ワークの面品位や加工時間の許容度が比較的大きくなるようにしきい値を設定し、学習が進行するにつれて許容度を縮小するようしきい値を変更していく構成とすることもできる。

40

【0039】

また、加工ワークの面品位及び加工時間の適否判定結果を、「適」及び「否」の二通り

50

だけでなく複数段階に設定することができる。例として、加工時間の想定最大値が T_{max} の場合、加工時間 T が、 $0 < T < T_{max} / 5$ のときは報酬 $R = 5$ を与え、 $T_{max} / 5 < T < T_{max} / 2$ のときは報酬 $R = 2$ を与え、 $T_{max} / 2 < T < T_{max}$ のときは報酬 $R = 1$ を与えるような構成とすることができる。あるいは、判定データ D の累積的な平均値からの偏差といった統計値を用いて、複数段階を定義することとしても良い。さらに、学習の初期段階は T_{max} を比較的大きく設定し、学習が進行するにつれて T_{max} を縮小する構成とすることもできる。同様に、ワークの面品位を表す指標についても複数段階で評価することが可能である。

【0040】

価値関数更新部 30 は、状態変数 S と判定データ D と報酬 R とを、関数 Q で表される行動価値（例えば数値）と関連付けて整理した行動価値テーブルを持つことができる。この場合、価値関数更新部 30 が関数 Q を更新するという行為は、価値関数更新部 30 が行動価値テーブルを更新するという行為と同義である。 Q 学習の開始時には環境の現在状態と各軸の速度分布（速度の N 階時間微分要素）との相関性は未知であるから、行動価値テーブルにおいては、種々の状態変数 S と判定データ D と報酬 R とが、無作為に定めた行動価値の値（関数 Q ）と関連付けた形態で用意されている。なお報酬計算部 28 は、判定データ D が分かればこれ対応する報酬 R を直ちに算出でき、算出した値 R が行動価値テーブルに書き込まれる。

10

【0041】

加工ワークの面品位及び加工時間の適否判定結果に応じた報酬 R を用いて Q 学習を進めると、より高い報酬 R が得られる行動を選択する方向へ学習が誘導され、選択した行動を現在状態で実行した結果として変化する環境の状態（つまり状態変数 S 及び判定データ D ）に応じて、現在状態で行う行動についての行動価値の値（関数 Q ）が書き替えられて行動価値テーブルが更新される。この更新を繰り返すことにより、行動価値テーブルに表示される行動価値の値（関数 Q ）は、適正な行動ほど大きな値となるように書き換えられる。このようにして、未知であった環境の現在状態（加工ワークの面品位及び加工時間）とそれに対する行動（各軸の速度分布（速度の N 階時間微分要素））との相関性が徐々に明らかになる。つまり行動価値テーブルの更新により、加工ワークの面品位及び加工時間と、各軸の速度分布（速度の N 階時間微分要素）との関係が最適解に徐々に近づけられる。

20

【0042】

図 3 A 及び図 3 B を参照して、学習部 26 が実行する上記した Q 学習のフロー（つまり機械学習方法の一形態）をさらに説明する。まずステップ $SA01$ で、価値関数更新部 30 は、その時点での行動価値テーブルを参照しながら、状態観測部 22 が観測した状態変数 S が示す現在状態で行う行動として各軸の速度分布（速度の N 階時間微分要素）を無作為に選択する。次に価値関数更新部 30 は、ステップ $SA02$ で、状態観測部 22 が観測している現在状態の状態変数 S を取り込み、ステップ $SA03$ で、判定データ取得部 24 が取得している現在状態の判定データ $D1$ 及び $D2$ を取り込む。次に価値関数更新部 30 は、ステップ $SA04$ で、判定データ $D1$ に基づき、各軸の速度分布（速度の N 階時間微分要素）が適当であったか否かを判断し、適当であった場合、ステップ $SA05$ で、報酬計算部 28 が求めた正の報酬 $R1$ を関数 Q の更新式に適用する。一方、ステップ $SA04$ で、各軸の速度分布（速度の N 階時間微分要素）が適当でなかったと判断した場合、ステップ $SA06$ で、報酬計算部 28 が求めた負の報酬 $R1$ を関数 Q の更新式に適用する。

30

40

【0043】

同様に、価値関数更新部 30 は、ステップ $SA07$ で、判定データ $D2$ に基づき、各軸の速度分布（速度の N 階時間微分要素）が適当であったか否かを判断し、適当であった場合、ステップ $SA08$ で、報酬計算部 28 が求めた正の報酬 $R2$ を関数 Q の更新式に適用する。一方、ステップ $SA07$ で、各軸の速度分布（速度の N 階時間微分要素）が適当でなかったと判断した場合、ステップ $SA09$ で、報酬計算部 28 が求めた負の報酬 $R2$ を関数 Q の更新式に適用する。

【0044】

50

次いでステップ S A 1 0 で、現在状態における状態変数 S 及び判定データ D 1 及び D 2 と報酬 R 1 及び R 2 と行動価値の値（更新後の関数 Q）とを用いて行動価値テーブルを更新する。学習部 2 6 は、ステップ S A 0 1 ~ S A 1 0 を繰り返すことで行動価値テーブルを反復して更新し、各軸の速度分布（速度の N 階時間微分要素）の学習を進行させる。

【 0 0 4 5 】

前述した強化学習を進める際に、例えば Q 学習の代わりに、ニューラルネットワークを用いることができる。図 4 A は、ニューロンのモデルを模式的に示す。図 4 B は、図 4 A に示すニューロンを組み合わせ構成した三層のニューラルネットワークのモデルを模式的に示す。ニューラルネットワークは、例えば、ニューロンのモデルを模した演算装置や記憶装置等によって構成できる。

10

【 0 0 4 6 】

図 4 A に示すニューロンは、複数の入力 x （ここでは一例として、入力 $x_1 \sim$ 入力 x_3 ）に対する結果 y を出力するものである。各入力 $x_1 \sim x_3$ には、この入力 x に対応する重み w （ $w_1 \sim w_3$ ）が掛けられる。これにより、ニューロンは、次の数 2 式により表現される出力 y を出力する。なお、数 2 式において、入力 x 、出力 y 及び重み w は、すべてベクトルである。また、 θ はバイアスであり、 f_k は活性化関数である。

【 0 0 4 7 】

【数 2】

$$y = f_k \left(\sum_{i=1}^n x_i w_i - \theta \right)$$

20

【 0 0 4 8 】

図 4 B に示す三層のニューラルネットワークは、左側から複数の入力 x （ここでは一例として、入力 $x_1 \sim$ 入力 x_3 ）が入力され、右側から結果 y （ここでは一例として、結果 $y_1 \sim$ 結果 y_3 ）が出力される。図示の例では、入力 x_1 、 x_2 、 x_3 のそれぞれに対応の重み（総称して w_1 で表す）が乗算されて、個々の入力 x_1 、 x_2 、 x_3 がいずれも 3 つのニューロン N 1 1、N 1 2、N 1 3 に入力されている。

【 0 0 4 9 】

図 4 B では、ニューロン N 1 1 ~ N 1 3 の各々の出力を、総称して z_1 で表す。 z_1 は、入力ベクトルの特徴量を抽出した特徴ベクトルと見なすことができる。図示の例では、特徴ベクトル z_1 のそれぞれに対応の重み（総称して w_2 で表す）が乗算されて、個々の特徴ベクトル z_1 がいずれも 2 つのニューロン N 2 1、N 2 2 に入力されている。特徴ベクトル z_1 は、重み w_1 と重み w_2 との間の特徴を表す。

30

【 0 0 5 0 】

図 4 B では、ニューロン N 2 1 ~ N 2 2 の各々の出力を、総称して z_2 で表す。 z_2 は、特徴ベクトル z_1 の特徴量を抽出した特徴ベクトルと見なすことができる。図示の例では、特徴ベクトル z_2 のそれぞれに対応の重み（総称して w_3 で表す）が乗算されて、個々の特徴ベクトル z_2 がいずれも 3 つのニューロン N 3 1、N 3 2、N 3 3 に入力されている。特徴ベクトル z_2 は、重み w_2 と重み w_3 との間の特徴を表す。最後にニューロン N 3 1 ~ N 3 3 は、それぞれ結果 $y_1 \sim y_3$ を出力する。

40

【 0 0 5 1 】

加減速制御装置 1 0 が備える機械学習装置 2 0 においては、状態変数 S と判定データ D とを入力 x として、学習部 2 6 が上記したニューラルネットワークに従う多層構造の演算を行うことで、各軸の速度分布（速度の N 階時間微分要素）（結果 y ）を出力することができる。なおニューラルネットワークの動作モードには、学習モードと価値予測モードとがあり、例えば学習モードで学習データセットを用いて重み w を学習し、学習した重み w を用いて価値予測モードで行動の価値判断を行うことができる。なお価値予測モードでは、検出、分類、推論等を行うこともできる。

【 0 0 5 2 】

上記した加減速制御装置 1 0 の構成は、コンピュータの CPU が実行する機械学習方法

50

(或いはソフトウェア)として記述できる。この機械学習方法は、各軸の速度分布(速度のN階時間微分要素)を学習する機械学習方法であって、コンピュータのCPUが、各軸の速度分布(速度のN階時間微分要素)を示す状態データS1と、必要に応じ加工タイプを示す状態データS2を、現在状態を表す状態変数Sとして観測するステップと、加工ワークの面品位及び加工時間の適否判定結果を示す判定データDを取得するステップと、状態変数Sと判定データDとを用いて、状態変数Sと判定データDとを関連付けて学習するステップとを有する。

【0053】

図5は、第2の実施形態による加減速制御装置40を示す。加減速制御装置40は、機械学習装置50と、状態観測部22が観測する状態変数Sの各軸の速度分布(速度のN階時間微分要素)S1と、必要に応じて加工タイプを示す状態データS2と、を状態データS0として取得する状態データ取得部42とを備える。

10

【0054】

加減速制御装置40が有する機械学習装置50は、各軸の速度分布(速度のN階時間微分要素)を機械学習により自ら学習するためのソフトウェア(学習アルゴリズム等)及びハードウェア(コンピュータのCPU等)に加えて、学習した各軸の速度分布(速度のN階時間微分要素)を工作機械(図示せず)への指令として出力するためのソフトウェア(演算アルゴリズム等)及びハードウェア(コンピュータのCPU等)を含むものである。加減速制御装置40が含む機械学習装置50は、1つの共通のCPUが、学習アルゴリズム、演算アルゴリズム等の全てのソフトウェアを実行する構成を有することもできる。

20

【0055】

意思決定部52は、例えばコンピュータのCPUの一機能として構成できる。或いは意思決定部52は、例えばコンピュータのCPUを機能させるためのソフトウェアとして構成できる。意思決定部52は、学習部26が学習した各軸の速度分布(速度のN階時間微分要素)に基づいて加工を行う工作機械への指令値Cを生成し、生成した指令値Cとして出力する。意思決定部52が学習した各軸の速度分布(速度のN階時間微分要素)に基づく指令値Cを工作機械に対して出力した場合、これに応じて、環境の状態(各軸の速度分布(速度のN階時間微分要素)S1)が変化する。

【0056】

状態観測部22は、意思決定部52による環境への各軸の速度分布(速度のN階時間微分要素)の表示、または出力後に変化した各軸の速度分布(速度のN階時間微分要素)S1を含む状態変数Sを次の学習サイクルにおいて観測する。学習部26は、変化した状態変数Sを用いて、例えば価値関数Q(すなわち行動価値テーブル)を更新することで、各軸の速度分布(速度のN階時間微分要素)を学習する。意思決定部52は、学習した各軸の速度分布(速度のN階時間微分要素)の下で状態変数Sに応じて各軸の速度分布(速度のN階時間微分要素)に基づく指令値Cを工作機械へと出力する。このサイクルを繰り返すことにより、機械学習装置50は各軸の速度分布(速度のN階時間微分要素)の学習を進め、自身が決定する各軸の速度分布(速度のN階時間微分要素)の信頼性を徐々に向上させる。

30

【0057】

上記構成を有する加減速制御装置40が備える機械学習装置50は、前述した機械学習装置20と同等の効果を奏する。特に機械学習装置50は、意思決定部52の出力によって環境の状態を変化させることができる。他方、機械学習装置20では、学習部26の学習結果を環境に反映させるための意思決定部に相当する機能を、外部装置(例えば工作機械の制御装置)に求めることができる。

40

【0058】

図6は、工作機械60を備えた一実施形態による加減速制御システム70を示す。加減速制御システム70は、同一の機械構成を有する複数の工作機械60、60'と、それら工作機械60、60'を互いに接続するネットワーク72とを備え、複数の工作機械60、60'のうち少なくとも1つが、上記した加減速制御装置40を備える工作機械60と

50

して構成される。また加減速制御システム70は、加減速制御装置40を備えない工作機械60'を含むことができる。工作機械60、60'は、ワークを加工するために必要とされる一般的な構成を有する。

【0059】

上記構成を有する加減速制御システム70は、複数の工作機械60、60'のうちで加減速制御装置40を備える工作機械60が、学習部26の学習結果を用いて、加工ワークの面品位及び加工時間に応じた各軸の速度分布(速度のN階時間微分要素)を、演算や目算によらずに自動的に、しかも正確に求めることができる。また、少なくとも1つの工作機械60の加減速制御装置40が、他の複数の工作機械60、60'のそれぞれについて得られた状態変数S及び判定データDに基づき、複数の工作機械60、60'の各軸の速度分布(速度のN階時間微分要素)を学習し、その学習結果を1以上の工作機械60、60'が利用するように構成できる。すなわち、複数の工作機械60、60'から取得した状態変数S及び判定データDに基づく学習結果を、1つの工作機械60の制御に適用することができ、又は複数の工作機械60、60'の制御に適用することもできる。したがって加減速制御システム70によれば、より多様なデータ集合(状態変数S及び判定データDを含む)を入力として、各軸の速度分布(速度のN階時間微分要素)の学習の速度や信頼性を向上させることができる。

10

【0060】

図7は、工作機械60'を備えた他の実施形態による加減速制御システム70'を示す。加減速制御システム70'は、機械学習装置50(又は20)と、同一の機械構成を有する複数の工作機械60'と、それら工作機械60'と機械学習装置50(又は20)とを互いに接続するネットワーク72とを備える。

20

【0061】

上記構成を有する加減速制御システム70'は、機械学習装置50(又は20)が、複数の工作機械60'のそれぞれについて得られた状態変数S及び判定データDに基づき、全ての工作機械60'に共通する各軸の速度分布(速度のN階時間微分要素)を学習し、その学習結果を用いて、加工ワークの面品位及び加工時間に応じた各軸の速度分布(速度のN階時間微分要素)を、演算や目算によらずに自動的に、しかも正確に求めることができる。

【0062】

加減速制御システム70'は、機械学習装置50(又は20)が、ネットワーク72に用意されたクラウドサーバやセルコントローラ等に存在する構成を有することができる。この構成によれば、複数の工作機械60'のそれぞれが存在する場所や時期に関わらず、必要なときに必要な数の工作機械60'を機械学習装置50(又は20)に接続することができる。

30

【0063】

加減速制御システム70、70'に従事する作業者は、機械学習装置50(又は20)による学習開始後の適当な時期に、機械学習装置50(又は20)による各軸の速度分布(速度のN階時間微分要素)に対応する加工ワークの面品位及び加工時間の学習の到達度(すなわち各軸の速度分布(速度のN階時間微分要素)と加工ワークの面品位及び加工時間との相関関係の信頼性)が要求レベルに達したか否かの判断を実行することができる。

40

【0064】

以上、本発明の実施の形態について説明したが、本発明は上述した実施の形態の例のみに限定されることなく、適宜の変更を加えることにより様々な態様で実施することができる。

【0065】

例えば、機械学習装置20、50が実行する学習アルゴリズム、機械学習装置50が実行する演算アルゴリズム、加減速制御装置10、40が実行する制御アルゴリズム等は、上述したものに限定されず、様々なアルゴリズムを採用できる。

【0066】

50

また、加減速制御装置 10, 40 の機械学習装置 20, 50 は、オンラインで機械学習することとしても良く、加減速制御装置 10, 40 の制御によるワーク加工時に状態データ S や判定データ D をログデータとして記録しておき、記録したログデータを収集して、収集したログデータから取得した状態データ S や判定データ D に基づいて機械学習装置 20, 50 が機械学習するようにしても良い。

【0067】

また、上述の実施の形態では、加減速制御装置 10, 40 の機械学習装置 20, 50 は、加工ワークの面品位及び加工時間の両方を判定データ D として用いる例を示したが、本発明はこれに限定されず、例えば加工ワークの加工精度、面品位及び加工時間のうち少なくともいずれか 1 つを判定データ D として用いることとしても良い。

10

【0068】

あるいは、加工ワークの面品位 D1、加工時間 D2 に加え、加工精度 D3 を判定データ D として用いることとしても良い。この場合、図 3B のフローチャートにおいては、価値観数更新処理 (SA10) の前に加工精度 D3 を用いて判定処理を実施することができる。すなわち、価値関数更新部 30 は、判定データ D3 に基づき、各軸の速度分布 (速度の N 階時間微分要素) が適当であったか否かを判断し、適当であった場合、報酬計算部 28 が正の報酬 R3 を求めて関数 Q の更新式に適用する。一方、各軸の速度分布 (速度の N 階時間微分要素) が適当でなかったと判断した場合、報酬計算部 28 が負の報酬 R3 を求めて関数 Q の更新式に適用する。そしてステップ SA10 で、現在状態における状態変数 S 及び判定データ D1、D2、D3 と報酬 R1 R2、R3 と行動価値の値 (更新後の関数 Q) とを用いて行動価値テーブルを更新する。

20

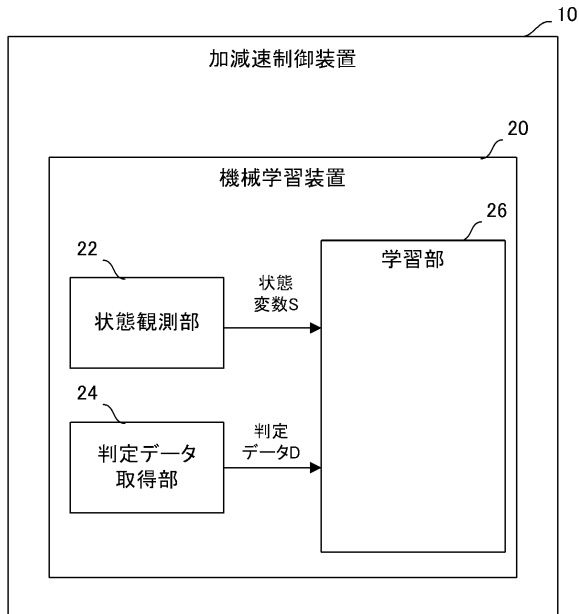
【符号の説明】

【0069】

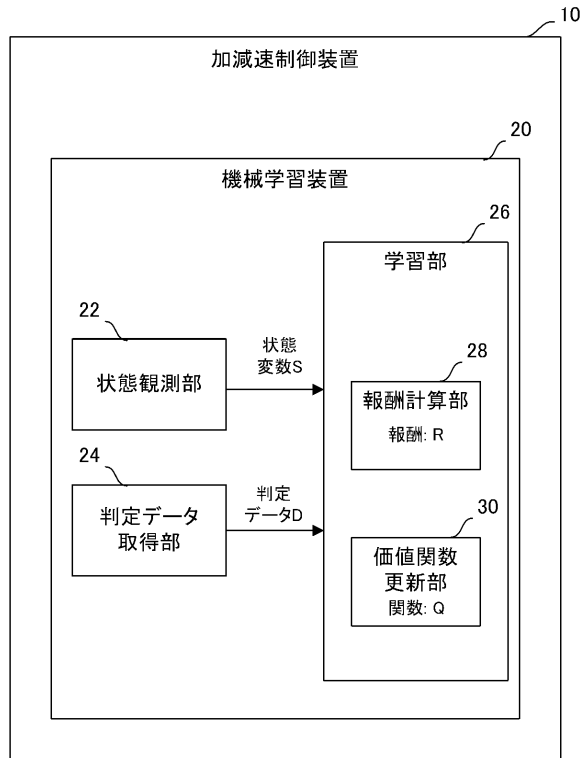
- 10, 40 加減速制御装置
- 20, 50 機械学習装置
- 22 状態観測部
- 24 判定データ取得部
- 26 学習部
- 28 報酬計算部
- 30 価値関数更新部
- 42 状態データ取得部
- 52 意思決定部
- 60, 60' 工作機械
- 70, 70' 加減速制御システム
- 72 ネットワーク

30

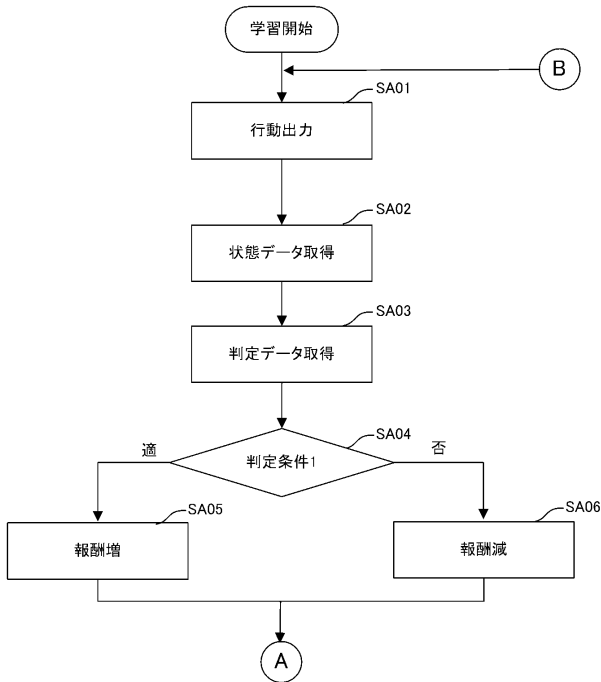
【図1】



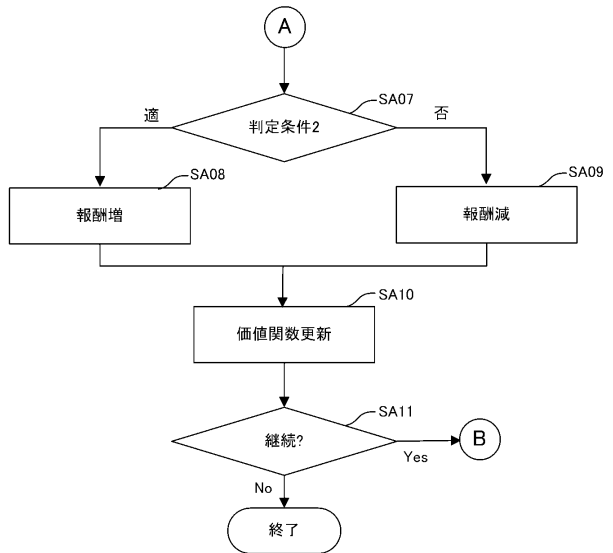
【図2】



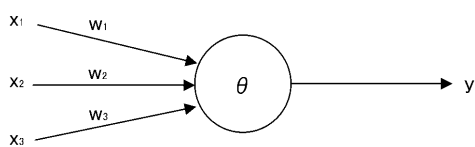
【図3A】



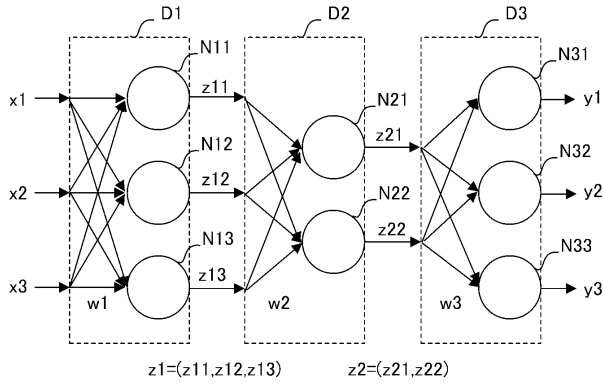
【図3B】



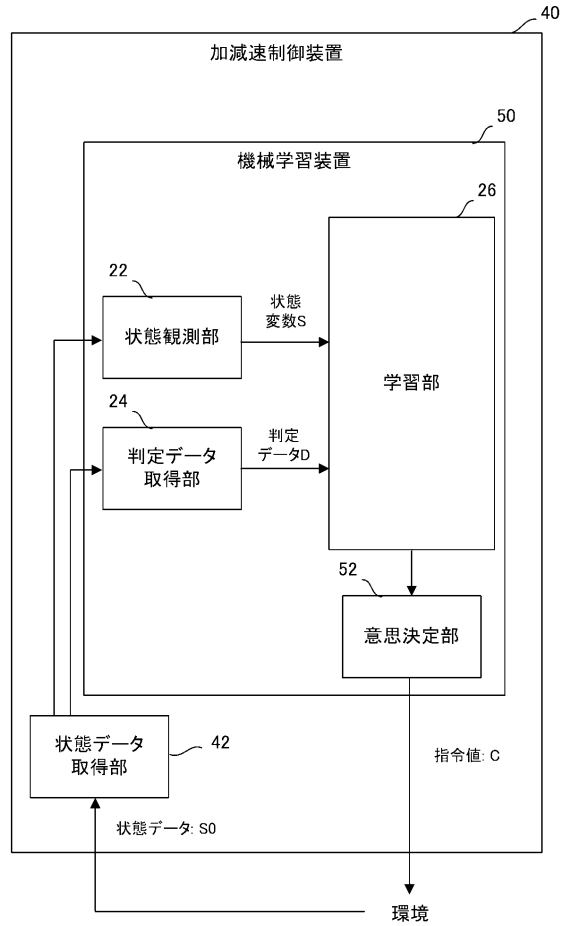
【図4A】



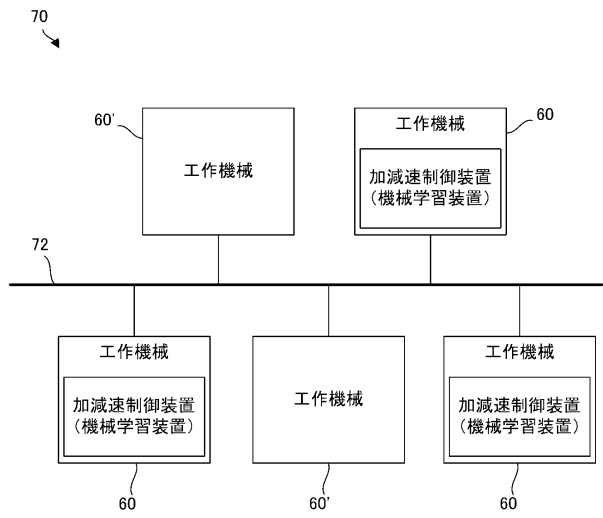
【図4B】



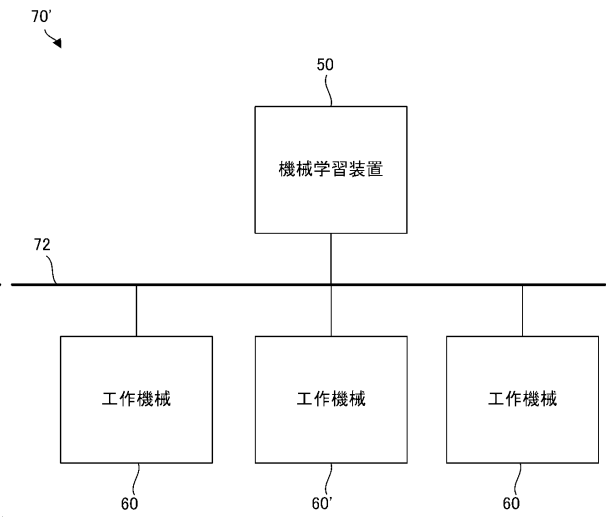
【図5】



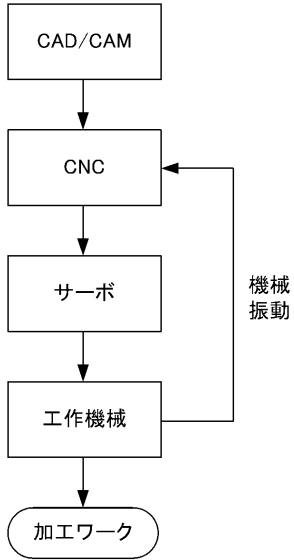
【図6】



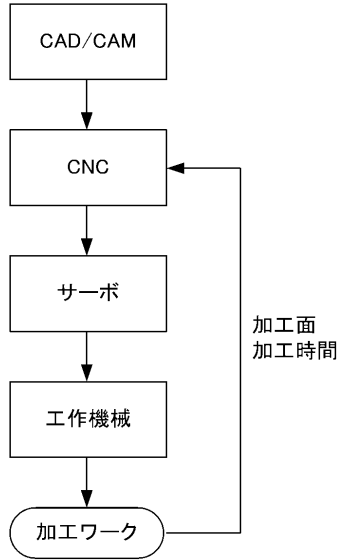
【図7】



【図 8】



【図 9】



フロントページの続き

- (56)参考文献 特開2017-068325(JP,A)
特開2017-030067(JP,A)
特開2017-045406(JP,A)
米国特許出願公開第2015/0331402(US,A1)
特開2003-005813(JP,A)
特開平04-135205(JP,A)
特開2017-062695(JP,A)
中国特許出願公開第106156079(CN,A)

(58)調査した分野(Int.Cl., DB名)

G05B 19/18 - 19/416
G06N 3/08 - 20/00