(19)日本国特許庁(JP)　(12)公開特許公報(A)　(11)特許出願公開番号
特開2017-185577
(P2017-185577A)
(43)公開日　平成29年10月12日(2017.10.12)

(51)Int.Cl.		F I			テーマコード（参考）
B25J 9/22	(2006.01)	B25J	9/22	A	3C269
G06N 99/00	(2010.01)	G06N	99/00	150	3C707
G05B 19/4155	(2006.01)	G05B	19/4155	V	
G05B 19/4069	(2006.01)	G05B	19/4069		

審査請求 有　請求項の数 15　ＯＬ　（全 15 頁）

(21)出願番号　特願2016-75476(P2016-75476)
(22)出願日　平成28年4月4日(2016.4.4)

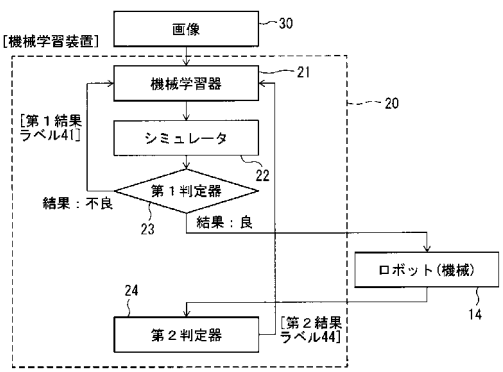
(71)出願人　390008235
ファナック株式会社
山梨県南都留郡忍野村忍草字古馬場３５８０番地
(74)代理人　100099759
弁理士　青木　篤
(74)代理人　100092624
弁理士　鶴田　準一
(74)代理人　100114018
弁理士　南山　知広
(74)代理人　100165191
弁理士　河合　章
(74)代理人　100151459
弁理士　中村　健一

最終頁に続く

(54)【発明の名称】シミュレーション結果を利用して学習を行う機械学習装置，機械システム，製造システムおよび機械学習方法

(57)【要約】
【課題】機械学習の安全性を保証することができる機械学習装置，機械システム，製造システムおよび機械学習方法の提供を図る。
【解決手段】機械学習により機械１４に対する制御指令を学習する機械学習装置２０であって、前記機械学習を行って前記制御指令を出力する機械学習器２１と、前記制御指令に基づいて、前記機械の作業動作のシミュレーションを実行するシミュレータ２２と、前記シミュレータによる前記シミュレーションの実行結果に基づいて、前記制御指令を判定する第１判定器２３と、を備える。
【選択図】図２

図2



【特許請求の範囲】
【請求項１】
機械学習により機械に対する制御指令を学習する機械学習装置であって、
前記機械学習を行って前記制御指令を出力する機械学習器と、
前記制御指令に基づいて、前記機械の作業動作のシミュレーションを実行するシミュレータと、
前記シミュレータによる前記シミュレーションの実行結果に基づいて、前記制御指令を判定する第１判定器と、を備える、
ことを特徴とする機械学習装置。
【請求項２】
前記第１判定器は、
前記シミュレーションの実行結果に問題がなければ、前記制御指令を良好と判定して、前記機械に対する前記制御指令の入力を実行し、
前記シミュレーションの実行結果に問題があれば、前記制御指令を不良と判定して、前記機械に対する前記制御指令の入力を停止する、
ことを特徴とする請求項１に記載の機械学習装置。
【請求項３】
前記第１判定器による判定結果が不良の場合、
前記機械に対する前記制御指令の入力を停止すると共に、
前記第１判定器により得られた第１結果ラベルを、訓練データとして前記機械学習器に対して与えることで学習を行う、
ことを特徴とする請求項２に記載の機械学習装置。
【請求項４】
前記第１判定器による判定結果が不良の場合、
前記第１結果ラベルと共に、前記シミュレータから得られる計算結果データの一部または全てを含む第１状態量を、前記機械学習器に対して与えることで学習を行う、
ことを特徴とする請求項３に記載の機械学習装置。
【請求項５】
さらに、
前記制御指令による前記機械の作業結果を判定する第２判定器を備え、
前記第１判定器による判定結果が良好の場合、
前記機械に対する前記制御指令の入力を実行すると共に、
前記第２判定器により得られる第２結果ラベルを、訓練データとして前記機械学習器に与えることで学習を行う、
ことを特徴とする請求項２乃至請求項４のいずれか１項に記載の機械学習装置。
【請求項６】
前記第１判定器による判定結果が良好の場合、
前記第２結果ラベルと共に、前記機械または周囲環境の状態を検出するセンサの出力データ，前記機械を制御する制御ソフトウェアの内部データ，および，前記センサの出力データまたは前記制御ソフトウェアの内部データに基づいて得られる計算データの少なくとも１つを含む第２状態量を、前記機械学習器に対して与えることで学習を行う、
ことを特徴とする請求項５に記載の機械学習装置。
【請求項７】
前記機械学習器の状態を、定期的または予め指定した条件成立時に第１状態として保存し、
前記第１判定器による判定結果が良好になる頻度が低下した場合、前記機械学習器の状態を、前記第１状態に戻す、
ことを特徴とする請求項２乃至請求項６のいずれか１項に記載の機械学習装置。
【請求項８】
前記第１判定器による判定結果が良好および不良を含む３つ以上の状態を取るとき、前

記状態に基づいて前記制御指令に含まれる前記機械の指令速度を変化させる、
ことを特徴とする請求項２乃至請求項７のいずれか１項に記載の機械学習装置。
【請求項９】
前記機械学習器は、
前記機械の作業状態に基づいて報酬を計算する報酬計算部と、
前記制御指令の価値を定める価値関数を有し、前記報酬に基づいて前記価値関数を更新する価値関数更新部と、を含む、
ことを特徴とする請求項２乃至請求項８のいずれか１項に記載の機械学習装置。
【請求項１０】
前記機械学習装置は、ニューラルネットワークを備える、
ことを特徴とする請求項１乃至請求項９のいずれか１項に記載の機械学習装置。
【請求項１１】
請求項１乃至請求項１０のいずれか１項に記載の機械学習装置、および、前記機械を備える機械システムであって、
前記機械を制御する制御装置と、
前記機械学習器，前記シミュレータおよび前記第１判定器を含む前記機械学習装置を有するコンピュータ装置と、を備え、
前記コンピュータ装置と、１台または複数台の前記機械がネットワークを介して接続される、
ことを特徴とする機械システム。
【請求項１２】
前記機械学習器の学習を、
前記機械による実際の生産作業中にも継続して行い、あるいは、
事前に実施して、前記機械による実際の生産作業中には行わない、
ことを特徴とする請求項１１に記載の機械システム。
【請求項１３】
請求項１１または請求項１２に記載の機械システムを複数備える製造システムであって、
前記機械学習装置は、それぞれの前記機械システムに設けられ、
複数の前記機械システムに設けられた複数の前記機械学習装置は、通信媒体を介して相互にデータを共有または交換するようになっている、
ことを特徴とする製造システム。
【請求項１４】
前記機械学習装置は、クラウドサーバ上に存在する、
ことを特徴とする請求項１３に記載の製造システム。
【請求項１５】
機械学習により機械に対する制御指令を学習する機械学習方法であって、
前記機械学習を行って前記制御指令を出力し、
前記制御指令に基づいて、前記機械の作業動作のシミュレーションを実行し、
前記シミュレーションの実行結果に基づいて、前記制御指令を判定し、
前記シミュレーションの実行結果に問題がなければ、前記制御指令を良好と判定して、前記機械に対する前記制御指令の入力を実行し、
前記シミュレーションの実行結果に問題があれば、前記制御指令を不良と判定して、前記機械に対する前記制御指令の入力を停止する、
ことを特徴とする機械学習方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本発明は、シミュレーション結果を利用して学習を行う機械学習装置，機械システム，製造システムおよび機械学習方法に関する。

【背景技術】
【０００２】
近年、例えば、ロボットなどの機械に対して機械学習を適用することが注目され、実用化のための研究・開発が行われている。例えば、機械の制御指令を学習するシステムにおいて、その機械を制御する制御指令が問題を有するものでないことを保証する必要がある。
【０００３】
ところで、従来、位置誤差をニューラルネットワークにより推定し、動作プログラムのデータを補正して動作する組立ロボットが開示されている(例えば、特許文献１参照)。
【０００４】
また、従来、仮想的なシステムモデルに対して、作業指令候補を繰り返し入力して学習することで、干渉しない最適な作業動作を獲得する装置も開示されている(例えば、特許文献２参照)。
【先行技術文献】
【特許文献】
【０００５】
【特許文献１】特開平０９－３１９４２０号公報
【特許文献２】特許第５１３７５３７号公報
【発明の概要】
【発明が解決しようとする課題】
【０００６】
上述したように、機械に対して機械学習を適用したものとしては、様々な提案がなされているが、例えば、特許文献１において、得られた学習結果によっては、機械の制御指令が機械や周囲環境に対して干渉などを引き起こす虞がある。すなわち、機械学習では、学習結果の安全性は保証されないため、例えば、干渉などを防ぐことは困難であり、生産ラインに対して機械学習を導入する場合の障害になっている。
【０００７】
また、例えば、特許文献２では、システムモデルの評価結果を用いて学習を行うため、システムモデルが正確である必要があり、このシステムモデルの精度を改善するためのモデル再構成部を設けることで、最適化を可能にしている。しかしながら、この手法では、最適化結果の精度がシステムモデルの精度に依存するため、システムモデルの改善に長時間を要することになり、現実的ではない。
【０００８】
本発明の目的は、上述した従来技術の課題に鑑み、機械学習の安全性を保証することができる機械学習装置，機械システム，製造システムおよび機械学習方法の提供にある。
【課題を解決するための手段】
【０００９】
本発明に係る第１実施形態によれば、機械学習により機械に対する制御指令を学習する機械学習装置であって、前記機械学習を行って前記制御指令を出力する機械学習器と、前記制御指令に基づいて、前記機械の作業動作のシミュレーションを実行するシミュレータと、前記シミュレータによる前記シミュレーションの実行結果に基づいて、前記制御指令を判定する第１判定器と、を備える機械学習装置が提供される。
【００１０】
前記第１判定器は、前記シミュレーションの実行結果に問題がなければ、前記制御指令を良好と判定して、前記機械に対する前記制御指令の入力を実行し、前記シミュレーションの実行結果に問題があれば、前記制御指令を不良と判定して、前記機械に対する前記制御指令の入力を停止するのが好ましい。前記第１判定器による判定結果が不良の場合、前記機械に対する前記制御指令の入力を停止すると共に、前記第１判定器により得られた第１結果ラベルを、訓練データとして前記機械学習器に対して与えることで学習を行うことができる。前記第１判定器による判定結果が不良の場合、前記第１結果ラベルと共に、前

記シミュレータから得られる計算結果データの一部または全てを含む第１状態量を、前記機械学習器に対して与えることで学習を行うこともできる。

【００１１】

前記機械学習装置は、さらに、前記制御指令による前記機械の作業結果を判定する第２判定器を備え、前記第１判定器による判定結果が良好の場合、前記機械に対する前記制御指令の入力を実行すると共に、前記第２判定器により得られる第２結果ラベルを、訓練データとして前記機械学習器に与えることで学習を行うことができる。前記第１判定器による判定結果が良好の場合、前記第２結果ラベルと共に、前記機械または周囲環境の状態を検出するセンサの出力データ，前記機械を制御する制御ソフトウェアの内部データ，および，前記センサの出力データまたは前記制御ソフトウェアの内部データに基づいて得られる計算データの少なくとも１つを含む第２状態量を、前記機械学習器に対して与えることで学習を行うこともできる。

【００１２】

前記機械学習器の状態を、定期的または予め指定した条件成立時に第１状態として保存し、前記第１判定器による判定結果が良好になる頻度が低下した場合、前記機械学習器の状態を、前記第１状態に戻すことができる。前記第１判定器による判定結果が良好および不良を含む３つ以上の状態を取るとき、前記状態に基づいて前記制御指令に含まれる前記機械の指令速度を変化させることもできる。前記機械学習器は、前記機械の作業状態に基づいて報酬を計算する報酬計算部と、前記制御指令の価値を定める価値関数を有し、前記報酬に基づいて前記価値関数を更新する価値関数更新部と、を含むのが好ましい。前記機械学習装置は、ニューラルネットワークを備えてもよい。

【００１３】

本発明に係る第２実施形態によれば、上述した第１実施形態の機械学習装置、および、前記機械を備える機械システムであって、前記機械を制御する制御装置と、前記機械学習器，前記シミュレータおよび前記第１判定器を含む前記機械学習装置を有するコンピュータ装置と、を備え、前記コンピュータ装置と、１台または複数台の前記機械がネットワークを介して接続される機械システムが提供される。前記機械学習器の学習を、前記機械による実際の生産作業中にも継続して行い、あるいは、事前に実施して、前記機械による実際の生産作業中には行わないように構成してもよい。

【００１４】

本発明に係る第３実施形態によれば、上述した第２実施形態の機械システムを複数備える製造システムであって、前記機械学習装置は、それぞれの前記機械システムに設けられ、複数の前記機械システムに設けられた複数の前記機械学習装置は、通信媒体を介して相互にデータを共有または交換するようになっている製造システムが提供される。前記機械学習装置は、クラウドサーバ上に存在してもよい。

【００１５】

本発明に係る第４実施形態によれば、機械学習により機械に対する制御指令を学習する機械学習方法であって、前記機械学習を行って前記制御指令を出力し、前記制御指令に基づいて、前記機械の作業動作のシミュレーションを実行し、前記シミュレーションの実行結果に基づいて、前記制御指令を判定し、前記シミュレーションの実行結果に問題がなければ、前記制御指令を良好と判定して、前記機械に対する前記制御指令の入力を実行し、前記シミュレーションの実行結果に問題があれば、前記制御指令を不良と判定して、前記機械に対する前記制御指令の入力を停止する機械学習方法が提供される。

【発明の効果】

【００１６】

本発明に係る機械学習装置，機械システム，製造システムおよび機械学習方法によれば、機械学習の安全性を保証することができるという効果を奏する。

【図面の簡単な説明】

【００１７】

【図１】図１は、本発明が適用される機械システムの一例を模式的に示す図である。

【図２】図２は、本発明に係る機械学習装置の第１実施例を示すブロック図である。
【図３】図３は、図２に示す機械学習装置における処理の一例を説明するための図である。
【図４】図４は、図３に示す全体画像から部分画像を切り出す処理の一例を説明するための図である。
【図５】図５は、図２に示す機械学習装置に適用したニューラルネットワークの入出力の構成例を説明するための図である。
【図６】図６は、本発明に係る機械学習装置の第２実施例を示すブロック図である。
【図７】図７は、一実施形態の機械学習装置における機械学習器の一例を説明するための図である。
【発明を実施するための形態】
【００１８】
まず、本発明に係るシミュレーション結果を利用して学習を行う機械学習装置，機械システム，製造システムおよび機械学習方法の実施形態を詳述する前に、本発明が適用される機械システムの一例、および、その問題点を、図１を参照して説明する。
【００１９】
図１は、本発明が適用される機械システムの一例を模式的に示す図であり、エンドエフェクタ１３を取り付けたロボット１４が、ワーク１２を籠１１から取り出す作業を行う様子を示すものである。図１に示されるように、機械システムの一例において、籠１１の上方には、三次元計測器１５が設けられ、この三次元計測器１５によりワーク１２の深度画像を撮影する。さらに、三次元計測器１５により撮影した深度画像(画像３０)に基づいて、ロボット１４がワーク１２を取り出す動作を学習する。ここで、ロボット１４は、例えば、６軸多関節型ロボットであり、ロボット１４およびエンドエフェクタ１３のそれぞれの駆動軸は、制御装置(機械学習装置)からの制御指令(指令)により制御される。
【００２０】
ところで、機械学習によりロボット１４に対する指令を学習する場合、機械学習装置(２０)により獲得された指令によっては、干渉が発生する虞がある。例えば、ロボット１４に対する指令が不適切(不良)な場合、具体的に、図１において、エンドエフェクタ１３が指令Ａにより軌跡１６のように移動(作業)する場合、エンドエフェクタ１３(ロボット１４)と籠１１が干渉することになる。
【００２１】
すなわち、ロボット１４に対する指令は、適切(良好)であることが求められ、具体的に、図１において、ロボット１４は、指令Ｂによるエンドエフェクタ１３の軌跡１７のように、エンドエフェクタ１３と籠１１が干渉することなく作業を行うことが要求されるが、例えば、上述した指令Ａがロボット１４に入力されると干渉が発生し、籠１１，エンドエフェクタ１３，ロボット１４およびエンドエフェクタ１３、あるいは、周辺装置などの破損を招く虞がある。特に、実際の生産ラインにおいては、生産停止につながるため、大きな問題になる。さらに、籠１１内のワーク１２の状態によっては、他のワーク１２との相互作用などに起因して、ロボット１４がワーク１２をうまく取り出せないことも生じ得る。この場合にも、作業のタクトタイムが遅れるという問題が発生する。
【００２２】
以下、本発明に係るシミュレーション結果を利用して学習を行う機械学習装置，機械システム，製造システムおよび機械学習方法の実施形態を、添付図面を参照して詳述する。なお、本明細書において、機械として、制御装置により制御される産業用ロボットを例として説明するが、本発明の適用は、例えば、制御装置により制御される工作機械，鍛圧機械，射出成形機，あるいは，サービス用ロボットなどであってもよいのはもちろんである。
【００２３】
図２は、本発明に係る機械学習装置の第１実施例を示すブロック図であり、例えば、図１を参照して説明した機械システム(ロボットシステム)に適用される機械学習装置を示す

ものである。機械学習装置２０は、三次元計測器１５からの画像情報(画像)３０を受け取って機械学習を行い、ロボット(機械)１４に制御指令(指令)を出力して作業動作(作業)を制御する。

【００２４】

図２に示されるように、機械学習装置２０は、機械学習を行って制御指令を出力する機械学習器２１と、制御指令に基づいてロボット１４の作業のシミュレーションを実行するシミュレータ２２と、シミュレータ２２によるシミュレーションの実行結果に基づいて制御指令を判定する第１判定器２３と、制御指令によるロボット１４の作業結果を判定する第２判定器２４を備える。

【００２５】

第１判定器２３は、シミュレータ２２によるシミュレーションの実行結果に問題がなければ、機械学習器２１から出力される制御指令を良好と判定してロボット１４に入力する。そして、ロボット１４は、第１判定器２３による判定結果が良好とされた制御指令に基づいて作業を行う。また、第１判定器２３は、シミュレータ２２によるシミュレーションの実行結果に問題があれば、制御指令を不良と判定してロボット１４に対する制御指令の入力を停止する。すなわち、第１判定器２３による判定結果が不良とされた制御指令はロボット１４に入力されず、ロボット１４は、シミュレータ２２によるシミュレーションの実行結果で問題があるとされた制御指令に基づいて作業を行うことから回避される。

【００２６】

ここで、第１判定器２３による判定結果が不良の場合、ロボット１４に対する制御指令の入力を停止するだけでなく、例えば、第１判定器２３により得られた第１結果ラベル４１を、訓練データとして機械学習器２１に対して与えることでさらなる学習を行うようにすることができる。

【００２７】

このように、第１実施例の機械学習装置２０によれば、実際にロボット１４を動作させる前に、シミュレータ２２により制御指令を実行し、例えば、周辺装置とロボット１４の距離や、ワーク１２の動きを事前にシミュレーションし、干渉の発生やワーク１２の取り出し失敗の可能性を第１判定器２３によって判定する。そして、シミュレータ２２によるシミュレーションの実行結果に基づいて、問題ないと評価(判定)された場合にのみ実機(ロボット１４)を動作させることで、実機の破損や作業の失敗などを防止することができる。すなわち、第１実施例によれば、機械学習の安全性を保証する、すなわち、機械の破損や作業の失敗あるいは周辺装置の破損などといった問題の発生を防止すことが可能になる。

【００２８】

さらに、第１実施例の機械学習装置２０において、第２判定器２４は、制御指令によるロボット１４の作業結果を判定する。そして、第１判定器２３による判定結果が良好の場合、ロボット１４に対する制御指令の入力を実行すると共に、第２判定器２４により得られる第２結果ラベル４４を、訓練データとして機械学習器２１に与えることで学習を行うようになっている。

【００２９】

すなわち、制御指令は、シミュレータ２２によって評価され、シミュレータの評価結果が第１判定器２３によって判定され、判定結果が良好であればロボット１４に対する制御指令の入力を実行して、ロボット１４に実際の動作を行わせる。そして、ロボット１４の動作結果(作業結果)は、第２判定器２４により評価(判定)され、第２判定器２４は、第２結果ラベル４４を出力する。そして、機械学習器２１は、この第２結果ラベル４４，上述した第１判定器２３からの第１結果ラベル４１および画像３０による訓練データに基づいて、制御指令の学習(機械学習)を行う。

【００３０】

なお、第１判定器２３による判定結果が不良の場合、ロボット１４に対する制御指令の入力は行われず、第１判定器２３から機械学習器２１に対して、第１結果ラベル４１が出

力される。そして、機械学習器２１は、この第１結果ラベル４１および画像３０による訓練データに基づいて、制御指令の学習を行う。

【００３１】

図３は、図２に示す機械学習装置における処理の一例を説明するための図であり、図４は、図３に示す全体画像から部分画像を切り出す処理の一例を説明するための図である。次に、図３および図４を参照して、例えば、図１に示す三次元計測器１５によって取得された全体画像３１から、画像の一部を切り出して部分画像３２を作成する処理の一例を説明する。

【００３２】

図４に示されるように、部分画像３２は、全体画像３１の中から任意の部分を切り出すことで作成することができるが、通常は一定間隔ごとにオフセットしながら、全領域をカバーするまで切り出し処理を繰り返して行う。具体的に、例えば、横３２０ピクセル，縦２４０ピクセルの全体画像３１に対して、横８０ピクセル，縦が８０ピクセルの部分画像(３２)を、縦横ともに４０ピクセルごとオフセットしながら部分画像を作成すると、｛(３２０－８０)／４０＋１｝×｛(２４０－８０)／４０＋１｝＝７×５＝３５枚の部分画像が得られることになる。

【００３３】

このようにして作成した部分画像３２を、例えば、ニューラルネットワークを用いた機械学習器２５に入力し、機械学習器２５の出力として、取り出し成功確率２６を得る。これを全ての部分画像に対して繰り返し行い、その結果の中から、最高取り出し成功確率を持つ部分画像(３２)を選択し、その部分画像３２を切り出した位置情報(最高取り出し成功確率位置の情報)２７を取得する。

【００３４】

なお、ニューラルネットワーク(２５)は、ニューロンのモデルを模した演算装置およびメモリなどで構成され、汎用の計算機やプロセッサを用いることもできるが、例えば、ＧＰＧＰＵ(General-Purpose computing on Graphics Processing Units)や大規模ＰＣクラスターなどを適用すると、より高速処理が可能になる。

【００３５】

そして、最高取り出し成功確率位置２７の位置情報をシミュレータ２２に入力し、ロボット１４の動作の評価(作業動作のシミュレーション)を行う。例えば、ワーク１２の取り出し位置から、籠１１の直上の位置を経由して、搬出位置へワーク１２を搬送するような動作を、予めシミュレータ２２に設定しておくことで、取り出し動作の軌跡を生成することができる。

【００３６】

ここで、簡略化されたシミュレータ２２としては、例えば、取り出し位置からの取り出し動作を生成し、ロボット先端位置(エンドエフェクタ１３)の時系列データを出力するものであってもよい。この時系列データは、第１判定器２３に入力され、ロボット先端位置が事前にユーザーが指定した範囲を逸脱しないかどうかをチェックすることにより、干渉の有無をチェックすることができ、動作結果の良否の判定、すなわち、制御指令が良好か不良かの判定を行うことができる。

【００３７】

また、より高度なシミュレータ２２としては、さらに、ロボット１４，ハンド(エンドエフェクタ１３)，籠１１，ワーク１２および周辺環境などの形状モデルや、重量などの物性値まで含めてモデリングし、例えば、重力の影響やワーク相互の干渉、ロボットに使用しているモータの負荷なども計算対象とする、いわゆる物理シミュレーションを行うシミュレータとすることもできる。

【００３８】

この場合、ロボット先端位置の範囲をチェックするだけではなく、ワークとワークが相互に干渉することによる取り出しの失敗、あるいは、ロボットが過負荷になってエラー停止するなどについても、シミュレート(評価)することが可能になる。また、シミュレータ

２２から得られる評価結果は、干渉によるロボット１４やワーク１２へのダメージ量、あるいは、干渉からの余裕度など多岐にわたるため、第１判定器２３による良否判定の基準も、それらの中から必要なものを取捨選択して組み合わせたものとすることで、より一層高度な判定を行うことが可能になる。

【００３９】

なお、前述したように、第１判定器２３により不良と判定された場合、実際のロボット１４(実機)には制御指令を入力せず、第１判定器２３からの第１結果ラベル４１および使用した部分画像４３による訓練データを、機械学習器(ニューラルネットワーク)２５に入力することで、制御指令の学習を行う。これにより、例えば、不良な(問題のある)制御指令を実機に入力することを避けることができ、また、機械学習器２５が不良な制御指令を出力しないように学習させることができる。

【００４０】

図５は、図２に示す機械学習装置に適用したニューラルネットワークの入出力の構成例を説明するための図である。図５に示されるように、ニューラルネットワーク２５は、入力として部分画像３２を受け取り、取り出し成功確率２６を出力する。ここで、学習する場合には、例えば、使用した部分画像４３と第１結果ラベル４１のセットを訓練データとし、教師あり学習によって制御指令を学習する。

【００４１】

すなわち、使用した部分画像４３から得られた取り出し成功確率２６と、シミュレータ２２から得られた第１結果ラベル４１との間の誤差を、誤差計算部２８により計算し、この誤差を逆伝播させることで、ニューラルネットワーク２５の学習を行うことができる。具体的に、取り出し成功確率２６が、失敗を表す『０』から成功を表す『１』までの範囲の数値を取るようにした場合、第１結果ラベル４１は、例えば、取り出し成功／失敗の結果を受けて『０』または『１』の値を取るようにすれば、誤差は、両者の差を取ることで計算することが可能である。

【００４２】

第１判定器２３により良好(良)と判定された場合、ロボット１４(実機)に制御信号を入力しても問題なく動作(作業)すると保証されたことを意味する。これにより、実機(ロボット１４)に制御指令を入力し、実際に動作を行わせる。そして、その結果をカメラなどのセンサにより観測し、実際に結果が良になったかどうかを、第２判定器２４によって判定する。

【００４３】

第２判定器２４からは、結果の良否(良好／不良)が第２結果ラベル４４として、ニューラルネットワーク２５に出力される。すなわち、ニューラルネットワーク２５は、第２結果ラベル４４および使用した部分画像４３で構成された訓練データを受け取って学習を行う。この学習の方法は、前述したシミュレータ２２から得られた第１結果ラベル４１を用いた場合の学習と同様である。

【００４４】

また、第１判定器２３および第２判定器２４は良否判定を行っているが、それぞれの判定器が出力する結果は、必ずしも二値(『０』または『１』)である必要はない。例えば、閾値を２段階設け、『良』，『可』，『不可』の３つに分類する判定を行うこともできる。これは、例えば、ロボット１４が籠１１などの障害物の付近を通過する場合、シミュレータ２２の精度によっては、干渉の可能性が疑われる状況が生じるためである。

【００４５】

例えば、干渉の可能性が疑われる場合、第１判定器２３の判定を『良(良好)』および『不可(不良)』の二値の判定ではなく、さらに、『可(中間)』を含めた三値、あるいは、それ以上の多値とし、制御指令をより細かく制御することができる。すなわち、第１判定器２３による判定結果が、良好および不良を含む３つ以上の状態(多値)を取るとき、その状態に基づいて、例えば、制御指令に含まれるロボット１４の指令速度を変化させる、具体的に、干渉の可能性が高い場合には、指令速度を下げてロボット１４を動作させるといっ

た制御を行うことができる。

【００４６】

これにより、例えば、ロボット１４による干渉が万一生じた場合でも、ロボット１４や周辺装置の損傷などを小さく抑えることが可能になる。さらに、閾値を３段階，４段階と増やし、指令速度を段階ごとに変化させて干渉のリスクに対して、より細かく対応することもできる。

【００４７】

ここで、図２を参照して説明した機械学習装置２０は、ロボット１４の制御装置内に設けることもできるが、例えば、ロボット１４を、機械学習を行わない作業にも使用することを考えると、コストの面から過剰な装備になってしまう。そこで、機械学習を行うコンピュータ装置(機械学習装置２０)を、ロボット１４の制御装置とは別に設けてもよい。このようなハードウェア構成にすることで、ロボット１４の制御装置側には、制御指令を受けるためのインターフェイスを準備するだけでよくなり、ロボット１４の制御装置のコストを抑えることができる。また、機械学習装置２０が設けられたコンピュータ装置のみを入れ替えることで、機械学習器２１やシミュレータ２２の能力を容易に向上させることもできるといった利点もある。

【００４８】

さらに、機械学習装置２０が設けられたコンピュータ装置とロボット１４の制御装置を、ネットワークを介して接続するため、コンピュータ装置を複数台のロボット１４(制御装置)と接続することが可能になり、学習を行うロボット１４を適宜切り替えて接続することにより、１台のコンピュータ装置で多数(全て)のロボット１４の学習を行うことも可能になる。すなわち、例えば、機械学習装置２０およびロボット１４を備えるロボットシステム(機械システム)が、ロボット１４を制御する制御装置と、機械学習器２１，シミュレータ２２および第１判定器２３を含む機械学習装置２０を有するコンピュータ装置と、を備え、そのコンピュータ装置と、１台または複数台のロボット１４を、ネットワークを介して接続することができる。

【００４９】

また、実際に生産を行う機械システムにおいては、生産を継続しながら学習によって動作を改善したい場合がある。特に、機械学習は、学習に多くの訓練データを必要とするため、学習を行う数(データ数)を増やして学習効果を上げるには、実際に生産を行うデータも使いながら学習するのが好ましい。

【００５０】

この場合、問題になるのは、学習によって動作が常に改善することは保証されず、逆に動作の成功率が低下する場合があることである。この問題も、干渉の問題と並んで、実際の生産ラインに学習機能を適用することの障害の一つになる虞がある。そこで、動作の成功率が低下するような場合には、実際に生産を行う機械システムへの適用時には、学習を行わないようにすることもできる。すなわち、機械学習器２１による学習を、例えば、ロボット１４による実際の生産作業中にも継続して行い、あるいは、事前に実施して、ロボット１４による実際の生産作業中には行わないようにすることができる。

【００５１】

また、機械学習器２１の状態を、実際の生産ライン投入段階など予め指定したタイミングで保存し、または、定期的に保存し、第１判定器２３による判定結果が不良になる確率が上がった場合には、機械学習器２１を以前の状態に戻して機械学習器２１の性能を回復し、実際の生産効率を保証しながら、学習による更なる効率改善を可能にすることができる。すなわち、機械学習器２１の状態を、定期的または予め指定した条件成立時に第１状態として保存し、第１判定器２３による判定結果が良好になる頻度が低下した場合、機械学習器２１の状態を、第１状態に戻すように構成することもできる。なお、上述した第１実施例では、「教師あり学習」を適用した例を示したが、「強化学習」などの他の学習方法を適用することもできる。

【００５２】

図６は、本発明に係る機械学習装置の第２実施例を示すブロック図であり、強化学習を適用した機械学習装置の例を示すものである。なお、図６は、取り出し位置が既に決定しているものとして、ロボット１４の取り出し動作の軌跡を学習する例を示している。図６に示す第２実施例では、機械学習器５５によりロボット１４の動作軌跡が生成され、シミュレータ５２に入力される。シミュレータ５２では、ロボット１４の動作が評価され、第１判定器５３によってその動作の良否(良好／不良)が判定される。

【００５３】

第１判定器５３による判定結果が良好の場合、制御指令がロボット１４に対する制御指令の入力を実行し、ロボット１４が動作する。ロボット１４の動作結果は、第２判定器５４で判定され、その判定結果は、第２結果ラベル６４として機械学習器５５に送られると共に、ロボット１４または周囲環境の状態を表すデータ、ロボット１４を制御する制御ソフトウェアの内部データ、および、それらの両者の一方または両方に基づいて得られる計算データの少なくとも１つを含む第２状態量６５が機械学習器５５に送られる。すなわち、第１判定器５３による判定結果が良好の場合、第２結果ラベル６４と共に、ロボット１４または周囲環境の状態を検出するセンサの出力データ，ロボット１４を制御する制御ソフトウェアの内部データ，および，センサの出力データまたは制御ソフトウェアの内部データに基づいて得られる計算データの少なくとも１つを含む第２状態量６５を、機械学習器５５に対して与えることで学習を行うようになっている。

【００５４】

また、第１判定器５３による判定結果が不良の場合、制御指令はロボット１４に入力されず、第１判定器５３による判定によって得られた第１結果ラベル６１と、シミュレータ５２の内部データで構成される第１状態量６２が機械学習器５５に送られる。すなわち、第１判定器５３による判定結果が不良の場合、第１結果ラベル６１と共に、シミュレータ５２から得られる計算結果データの一部または全てを含む第１状態量６２を、機械学習器５５に対して与えることで学習を行うようになっている。

【００５５】

このとき、機械学習器５５で生成される動作軌跡は、例えば、一定時間分、または、一定距離分の短いステップで生成され、その都度シミュレーションと判定が行われ、それを繰り返すことで、最終的に全体の動作軌跡が生成される。

【００５６】

ここで、第１結果ラベル６１および第１状態量６２、または、第２結果ラベル６４および第２状態量６５を受け取った機械学習器５５は、例えば、「強化学習」を行う。具体的に、「Ｑ学習」を適用して学習を行う場合について説明する。

【００５７】

Ｑ学習は、ある環境状態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。つまり、ある状態ｓのとき、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ればよい。しかし、最初は、状態ｓと行動ａとの組合せについて、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(行動主体)は、ある状態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、報酬が与えられる。それにより、エージェントは、より良い行動の選択、すなわち、正しい価値Ｑ(ｓ，ａ)を学習していく。

【００５８】

ここで、第２実施例の機械学習装置において、Ｑ学習を適用して学習を行う場合、状態量ｓは、第１判定器５３から出力された第１結果ラベル６１およびシミュレータ５２から出力された第１状態量６２、または、第２判定器５４から出力された第２結果ラベル６４およびロボット１４や周辺機器などから出力された第２状態量６５で構成される。

【００５９】

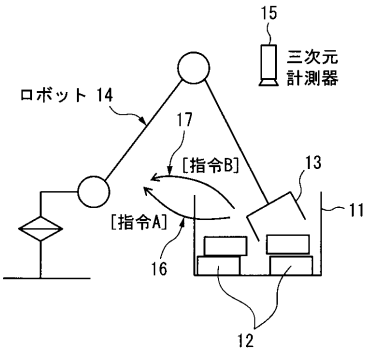
第１結果ラベル６１および第２結果ラベル６４には、例えば、干渉の有無，タクトタイム，最終位置までの距離、ワーク１２の把持の成功／失敗などの結果が含まれる。また、行動ａは、ロボット１４に対する制御指令であり、例えば、ロボット１４の先端位置(エンドエフェクタ１３)の移動方向および移動速度などとすることができる。

【００６０】
図７は、一実施形態の機械学習装置における機械学習器の一例を説明するための図であり、「強化学習(Ｑ学習)」を適用した場合の一例を示すものである。図７に示されるように、機械学習器５５は、報酬計算部７１および価値関数更新部７２を備える。報酬計算部７１は、例えば、第１結果ラベル６１および第１状態量６２、または、第２結果ラベル６４および第２状態量６５に基づいて、報酬、例えば、スコアを計算する。
【００６１】
ここで、報酬は、例えば、動作の成功に対しては高くなるように設定され、動作の失敗に対しては低くなるように設定される。また、報酬は、例えば、動作のタクトタイムに基づいて計算してもよい。さらに、報酬に関して、例えば、ハンド部(エンドエフェクタ１３)による把持の成功や、ハンド部による運搬の成功などといった各段階に応じて報酬の計算を行うようにしてもよい。また、例えば、干渉が発生したと判断した場合、報酬が低くなるように設定すると共に、その時点で動作を中断して初期位置からやり直すといった処理を行うことで、例えば、問題のある状態を発生または継続させずに、学習を進めることができる。
【００６２】
次に、価値関数更新部７２は、例えば、動作の価値を定める価値関数Ｑ(ｓ，ａ)を有し、上述した報酬計算部７１による報酬に応じて価値関数を更新する。そして、制御指令(動作軌跡)は、更新された価値関数に基づいて生成され、シミュレータ５２に出力されて上述した処理を行い、ロボット１４(機械)を制御するために使用される。これにより、機械学習の安全性を保証(機械の破損や作業の失敗、あるいは、周辺装置の破損などといった問題の発生を防止)した制御指令により機械を制御することができる。
【００６３】
なお、本発明に係る機械学習装置に対しては、上述した「教師あり学習」や「強化学習(Ｑ学習)」だけでなく、例えば、知られている様々な機械学習の手法(例えば、様々な「教師あり学習」，「教師なし学習」および「強化学習」)、あるいは、これらの手法を実現するうえで、特徴量そのものの抽出を学習する「深層学習(ディープラーニング：Deep Learning)」などの様々な技術を幅広く適用することができる。さらに、本発明に係る機械学習装置の適用は、上述した産業用ロボットなどに限定されるものではなく、例えば、制御装置により制御される工作機械，鍛圧機械，射出成形機，あるいは，サービス用ロボットなどであってもよいのはいうまでもない。
【００６４】
以上において、機械学習装置および機械(ロボット)を備える機械システムを複数備える製造システムを考えた場合、例えば、機械学習装置をそれぞれの機械システムに設け、複数の機械システムに設けられた複数の機械学習装置を、通信媒体を介して相互にデータを共有または交換するように構成してもよい。また、機械学習装置は、例えば、インターネットなどの通信媒体を介して、クラウドサーバ上に存在するように構成することも可能である。
【００６５】
以上、実施形態を説明したが、ここに記載した全ての例や条件は、発明および技術に適用する発明の概念の理解を助ける目的で記載されたものであり、特に記載された例や条件は発明の範囲を制限することを意図するものではない。また、明細書のそのような記載は、発明の利点および欠点を示すものでもない。発明の実施形態を詳細に記載したが、各種の変更、置き換え、変形が発明の精神および範囲を逸脱することなく行えることが理解されるべきである。
【符号の説明】
【００６６】
１１　籠
１２　ワーク
１３　エンドエフェクタ

１４　　ロボット(機械)
１５　　三次元計測器
２０　　機械学習装置
２１，２５，５５　　機械学習器(ニューラルネットワーク)
２２，５２　　シミュレータ
２３，５３　　第１判定器
２４，５４　　第２判定器
７１　　報酬計算部
７２　　価値関数更新部

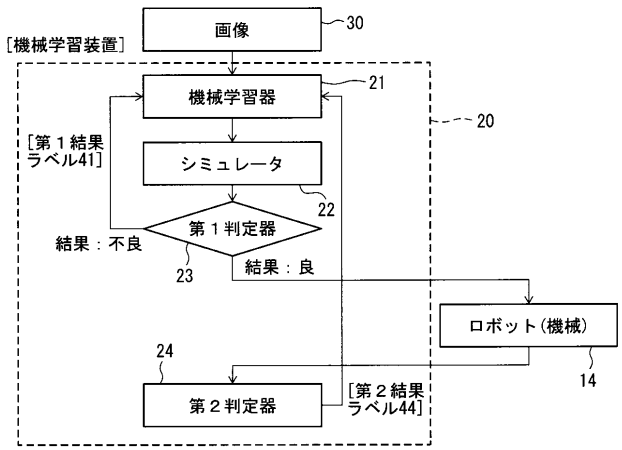
【図１】

図1



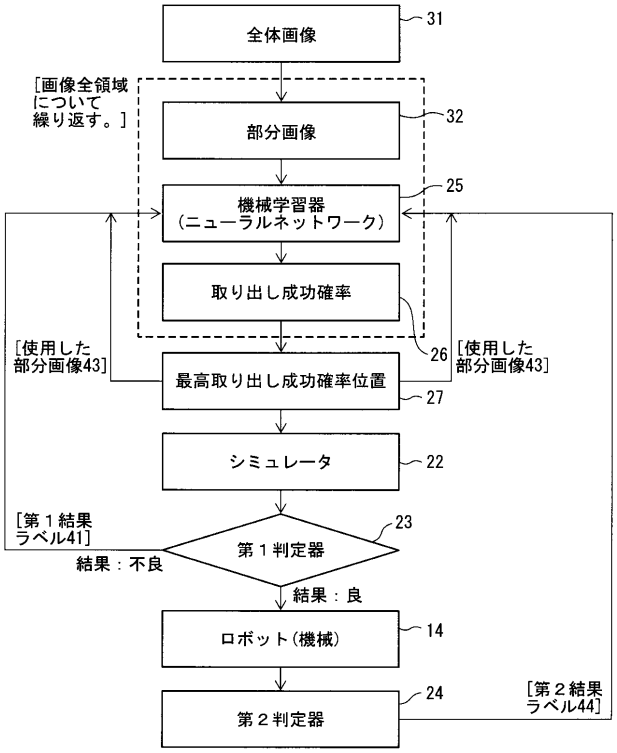
【図２】

図2



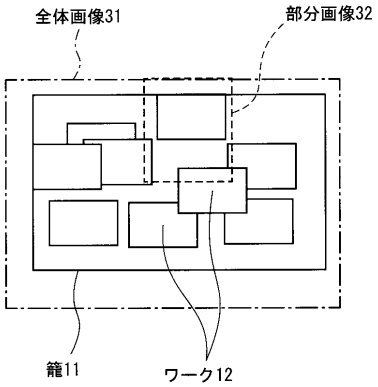
【図３】

図3



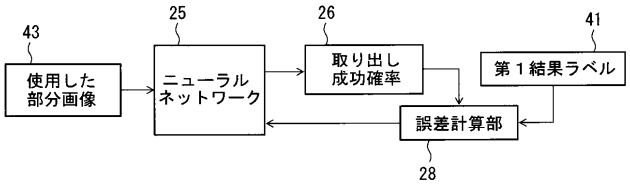
【図４】

図4



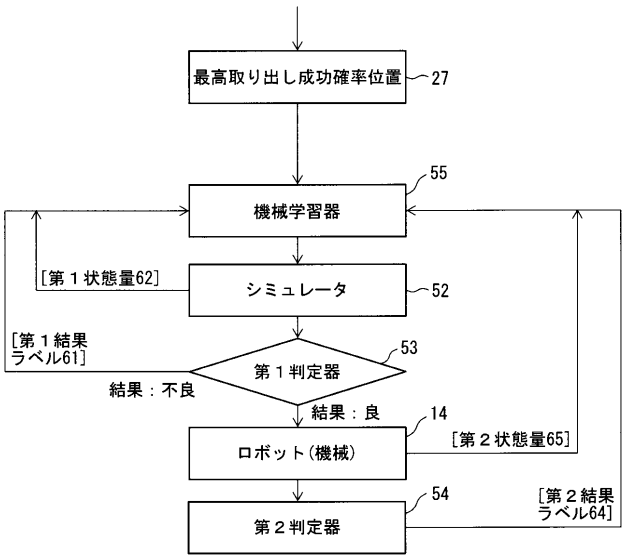
【図５】

図5



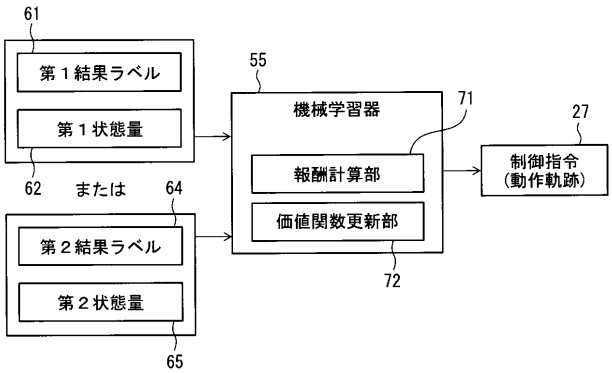
【図６】

図6



【図７】

図7



フロントページの続き

(72)発明者　中川　浩
山梨県南都留郡忍野村忍草字古馬場３５８０番地　ファナック株式会社内
(72)発明者　梶山　貴史
山梨県南都留郡忍野村忍草字古馬場３５８０番地　ファナック株式会社内

Ｆターム(参考) 3C269 AB33 BB14 MN41 MN44
3C707 AS04 KT03 KT06 LS08 LS15 LS20 LT06 LU05 LW15