

(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1月20日(20.01.2022)



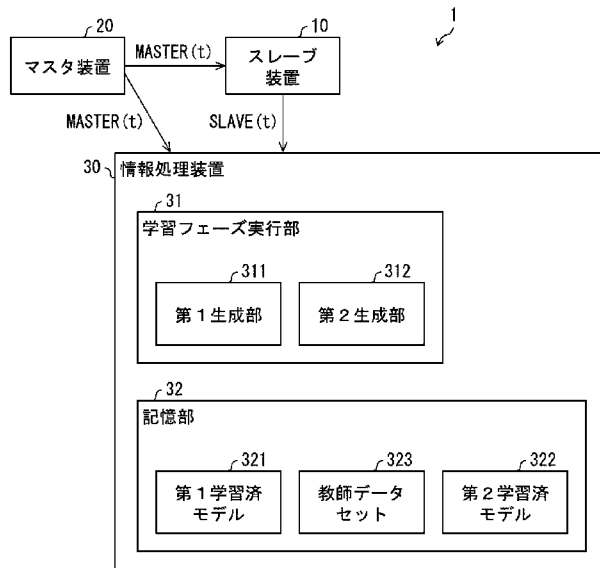
(10) 国際公開番号

WO 2022/014485 A1

- (51) 国際特許分類:
B25J 3/00 (2006.01) B25J 13/00 (2006.01)
- (21) 国際出願番号: PCT/JP2021/025932
- (22) 国際出願日: 2021年7月9日(09.07.2021)
- (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 2020-120740 2020年7月14日(14.07.2020) JP
特願 2020-217788 2020年12月25日(25.12.2020) JP
- (71) 出願人: 国立大学法人筑波大学(UNIVERSITY OF TSUKUBA) [JP/JP]; 〒3058577 茨城県つくば市天王台一丁目1番1 Ibaraki (JP).
- (72) 発明者: 境野 翔(SAKAINO Sho); 〒3058577 茨城県つくば市天王台一丁目1番1 国立大学法人筑波大学内 Ibaraki (JP).
- (74) 代理人: 棚井 澄雄, 外 (TANAI Sumio et al.); 〒1006620 東京都千代田区丸の内一丁目9番2号 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,

(54) Title: INFORMATION PROCESSING DEVICE, METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、方法、及びプログラム



- 10 Slave device
20 Master device
30 Information processing device
31 Training phase execution unit
32 Storage unit
311 First generation unit
312 Second generation unit
321 First trained model
322 Second trained model
323 Training data set

(57) Abstract: A first inference unit (331) uses a first trained model (321) to infer command information at a first unit time. A second inference unit (332) uses a second trained model (322) to infer command information at a second unit time that is shorter than the first unit time from information that corresponds to command information derived by means of the first trained model (321). An operation control unit (333) uses the command information inferred by the second inference unit (332) to operate a control target device (10B).

(57) 要約: 第1推定部(331)は、第1学習済モデル(321)を用いて、第1の単位時間毎に指令情報を推定する。第2推定部(332)は、第2学習済モデル(322)を用いて、第1学習済モデル(321)により導出される指令情報に応じた情報から、第1の単位時間よりも短い第2の単位時間毎に指令情報を推定する。動作制御部(333)は、第2推定部(332)が推定した指令情報を用いて制御対象装置(10B)を動作させる。

[続葉有]

QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
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))

明 細 書

発明の名称： 情報処理装置、方法、及びプログラム

技術分野

[0001] 本発明は、情報処理装置、方法、及びプログラムに関する。

背景技術

[0002] ロボットを制御する技術が知られている。例えば、特許文献1には、マスタ・スレーブシステムによって遠隔制御を行うロボットが記載されている。また、マスタ・スレーブシステムを用いたロボットの遠隔操作を機械学習させて学習済モデルを生成し、この学習済モデルを用いてロボットを動作させる技術が提案されている。

先行技術文献

特許文献

[0003] 特許文献1：WO2015/041046号公報

発明の概要

発明が解決しようとする課題

[0004] 機械学習させた学習済モデルを用いてロボットを動作させる場合、動作条件（動作速度、等）や動作環境（操作対象物の形状や位置、柔軟性などの物理特性、障害物位置、等）が変化した場合に、ロボットが適切に動作しない場合があった。これら、動作条件や動作環境の変化に対応しロボットの動作予測を安定させるためには、膨大な教師データが必要になるとともに、計算時間も膨大になってしまうという問題があった。

[0005] 本発明の一態様は、ロボット等、制御対象装置の動作制御を行う際に、動作予測の安定した制御対象装置を実現することを目的とする。

課題を解決するための手段

[0006] 本発明の態様1に係る情報処理装置は、指令情報に従い動作する制御対象装置が出力する、当該制御対象装置の状態を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第1学習済モデルを用いて、第1の

単位時間毎に前記指令情報を推定する第1推定部と、前記制御対象装置が出力した前記状態を表す情報と、当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報とから、前記制御対象装置に供給される前記指令情報を導出する第2学習済モデルを用いて、前記指令情報を、前記第1の単位時間よりも短い第2の単位時間毎に推定する第2推定部と、前記第2推定部が推定した前記指令情報を用いて前記制御対象装置を動作させる動作制御部と、を備えている。

[0007] 上記の構成によれば、情報処理装置は、状態を表す情報から指令情報を導出する第1学習済モデルと、第1学習済モデルの出力に応じた情報から指令情報を導出する第2学習済モデルとを用いて、制御対象装置を動作させる。これらの複数の学習済モデルが制御対象装置の動作制御において階層的に用いられることにより、膨大な教師データを必要とすることなく、制御対象装置の動作予測を安定させることができる。また、ここでは2階層の学習済モデルについて記したが、さらに階層を増やしてn階層のモデルを用いても、制御対象装置の動作目的に応じ各階層モデルを用いる、用いないの選択をしても良い。

[0008] 本発明の態様2に係る情報処理装置は、前記態様1において、前記指令情報は、前記指令情報を出力する装置から発行される情報であり、前記第2学習済モデルに入力される情報は、前記第1学習済モデルが導出した指令情報、及び、当該指令情報と前記状態を表す情報を含んでもよい。

[0009] 上記の構成によれば、第2学習済モデルの入力データは、前記第1の単位時間で推定される前記第1学習済モデルの出力と前記第2の単位時間で出力される前記制御対象装置が出力した前記状態の組となる。前記第2単位時間と比較して前記第1単位時間を長く設定しているので、前記第1学習済モデルの出力の変動が少ないものとして前記第2学習済モデルは前記制御対象装置の動作予測することができるようになり、予測を安定させることができる。

[0010] 本発明の態様3に係る情報処理装置は、前記態様1又は2において、前記

第2学習済モデルに入力される情報は、前記第1学習済モデルが導出した指令情報と前記状態を表す情報との組を低相関化（相関が小さくなるように変換）したデータであってもよい。例えば、二つのデータからそれらの和と差のデータを生成することで低相関化できる。

[0011] 上記の構成によれば、第2学習済モデルの入力データ間の相関が小さくなる。これにより、制御対象装置の動作予測を安定させることができる。データ間の相関が低ければ低いほど、制御対象装置の動作予測を安定させることができる。

[0012] 本発明の態様4に係る情報処理装置は、前記態様1～3において、前記指令情報を出力する装置、及び前記制御対象装置はそれぞれ、ロボットアームを備え、前記指令情報及び前記状態を表す情報はそれぞれ、前記ロボットアームの動作を示す情報を含んでもよい。

[0013] 上記の構成によれば、情報処理装置は、ロボットアームの動作を示す情報を含む指令情報を導出する階層的な学習済モデルを用いて、制御対象装置の動作制御を行う。これにより、ロボットアームを備える制御対象装置の動作予測を安定させることができる。

[0014] 本発明の態様5に係る情報処理装置は、前記態様1～4において、前記指令情報を出力する装置は、ユーザ操作に応じて動作するマスタ装置であり、前記制御対象装置は、前記マスタ装置から出力される前記指令情報に従い動作するスレーブ装置であり、前記第1学習済モデル及び前記第2学習済モデルは、前記マスタ装置及び前記スレーブ装置を備えるマスタ・スレーブシステムにおいて収集される前記指令情報及び前記状態を表す情報を用いて学習されてもよい。

[0015] 上記の構成によれば、情報処理装置は、マスタ・スレーブシステムにおいて収集される指令情報及び状態を表す情報を用いて生成された、階層的な複数の学習済モデルを用いて制御対象装置を動作させる。これにより、制御対象装置の動作予測を安定させることができる。

[0016] 本発明の態様6に係る情報処理装置は、指令情報に従い動作する制御対象

装置が出力した、当該制御対象装置の状態を表す情報と、前記制御対象装置に供給された前記指令情報との組を用いて、前記状態を表す情報と前記指令情報との相関関係を学習させた第1学習済モデルを生成する第1生成部と、前記制御対象装置が出力した前記状態を表す情報と当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報との関係を表す情報を含む入力情報と、前記制御対象装置に供給された前記指令情報との組を用いて、前記入力情報と前記指令情報との相関関係を学習させた第2学習済モデルを生成する第2生成部と、を備えている。

[0017] 上記の構成によれば、情報処理装置は、状態を表す情報から指令情報を導出する第1学習済モデルと、第1学習済モデルの出力に応じた情報から指令情報を導出する第2学習済モデルとを生成する。これらの複数の推定する単位時間を持つ学習済モデルが制御対象装置の動作制御において階層的に用いられることにより、第1生成部と第2生成部を独立して動作させることができる。すなわち、第1学習済モデルと第2学習済モデルを独立して得ることができるため、膨大な教師データを必要とすることなく、学習や制御対象装置の動作予測を安定させることができる。

[0018] 本発明の態様7に係る情報処理装置は、前記態様6において、前記制御対象装置が出力した前記状態を表す情報を第1の単位時間毎に前記第1学習済モデルに入力することにより前記指令情報を推定する第1推定部と、前記制御対象装置が出力した前記状態を表す情報と当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報との関係を表す情報を含む入力データを、前記第1の単位時間よりも短い第2の単位時間毎に前記第2学習済モデルに入力することにより前記指令情報を推定する第2推定部と、前記第2推定部が推定した前記指令情報を用いて前記制御対象装置を動作させる動作制御部と、を更に備えてもよい。

[0019] 上記の構成によれば、情報処理装置は、状態を表す情報から指令情報を導出する第1学習済モデルと、第1学習済モデルの出力に応じた情報から指令情報を導出する第2学習済モデルとを用いて、制御対象装置を動作させる。

これらの複数の学習済モデルが制御対象装置の動作制御において階層的に用いられることにより、膨大な教師データを必要とすることなく、制御対象装置の動作予測を安定させることができる。

[0020] 本発明の態様 8 に係る情報処理装置は、前記態様 6 又は 7 において、前記指令情報を出力する装置は、ユーザ操作に応じて動作するマスタ装置であり、前記制御対象装置は、前記マスタ装置から出力される前記指令情報に従い動作するスレーブ装置であり、前記第 1 生成部は、前記マスタ装置及び前記スレーブ装置を備えるマスタ・スレーブシステムにおいて収集される前記指令情報及び前記状態を表す情報を用いて、前記第 1 学習済モデルを生成し、前記第 2 生成部は、前記マスタ・スレーブシステムにおいて収集される前記指令情報及び前記状態を表す情報を用いて、前記第 2 学習済モデルを生成してもよい。

[0021] 上記の構成によれば、情報処理装置は、マスタ・スレーブシステムにおいて収集される指令情報及び状態を表す情報を用いて、階層的な複数の学習済モデルを生成する。マスタ装置はユーザ操作に応じ動作することから、スレーブ装置の制御の遅延や環境との動的なインタラクションをマスタ装置を通して人間が体感しながら制御する情報を前記指令情報として収集することができる。これらの複数の学習済モデルが制御対象装置の動作制御において用いられることにより、制御対象装置の動作予測を安定させることができる。

[0022] 本発明の態様 9 に係る情報処理装置は、前記態様 1 ～ 8 において、前記第 1 学習済モデルは、前記状態を表す情報と、前記制御対象装置の動作を規定する、動作速度を含む情報とを含む入力データから、前記制御対象装置が当該状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報を導出する学習済モデルであってもよい。

[0023] 上記の構成によれば、情報処理装置は、制御対象装置の動作を規定する、動作速度を含む情報を含む教師データを用いて学習された第 1 学習済モデルを用いて、指令情報を推定する。これにより、動作速度が可変である制御対象装置の動作予測を安定させることができる。

- [0024] 本発明の態様１０に係る情報処理装置は、前記態様９において、前記動作速度を含む情報は、前記制御対象装置の動作周期を表す周波数又は動作実行時間を示す情報を含んでもよい。
- [0025] 上記の構成によれば、情報処理装置は、制御対象装置の動作周期を示す周波数を含む教師データを用いて学習された第１学習済モデルを用いて、指令情報を推定する。これにより、周期的な動作を行う制御対象装置の動作予測を安定させることができる。
- [0026] 本発明の態様１１に係る情報処理装置は、前記態様９又は１０において、前記第１学習済モデルは、前記入力データから、前記制御対象装置が前記状態を表す情報を出力してから予め定められた単位時間が経過した後に前記制御対象装置に供給される前記指令情報を導出する学習済モデルであってもよい。
- [0027] 上記の構成によれば、情報処理装置は、制御対象装置が状態を表す情報を出力してから単位時間が経過した後に制御対象装置に供給される指令情報を導出する第１学習済モデルを用いて、指令情報を推定する。これにより、動作速度が可変である制御対象装置の動作予測を安定させることができる。
- [0028] 本発明の態様１２に係る情報処理装置は、前記態様６又は７において、前記第１学習済モデルは、前記状態を表す情報と、前記制御対象装置に指示する動作を規定する、動作速度を含む情報とを含む入力データから、前記制御対象装置が当該状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報を出力データとして導出する学習済モデルであり、前記第１生成部は、前記制御対象装置の動作速度を変えながら前記制御対象装置を動作させることにより収集される前記指令情報、前記状態を表す情報、及び前記動作速度を含む情報を用いて、前記第１学習済モデルを生成してもよい。
- [0029] 上記の構成によれば、制御対象装置の動作速度を変えながら制御対象装置を動作させることにより収集される教師データを用いて、第１学習済モデルが生成される。情報処理装置がこの学習済モデルを用いて制御対象装置の動作制御を行うことにより、動作速度が可変である制御対象装置の動作予測を

安定させることができる。

[0030] 本発明の態様 1 3 に係る情報処理装置は、前記態様 1 ～ 1 2 において、前記第 1 学習済モデルは、前記指令情報から、前記制御対象装置に当該指令情報が供給された後に前記制御対象装置が出力する第 2 の状態を表す情報を導出する第 3 学習済モデルにより導出される第 2 の状態を表す情報を入力情報として学習させた学習済モデルであってもよい。

[0031] 上記の構成によれば、第 3 学習済モデルが導出する状態を表す情報が、第 1 学習済モデルの入力情報として学習が行われ、第 1 学習済モデルが生成される。情報処理装置は、この第 1 学習済モデルを用いて制御対象装置の動作制御を行うことにより、制御対象装置の動作予測を安定させることができる。

[0032] 本発明の態様 1 4 に係る方法は、指令情報に従い動作する制御対象装置が出力する、当該制御対象装置の状態を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第 1 学習済モデルを用いて、前記指令情報を、第 1 の単位時間毎に推定する第 1 ステップと、前記制御対象装置が出力した前記状態を表す情報と、当該状態を表す情報を前記第 1 学習済モデルに入力することにより導出される前記指令情報との関係を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第 2 学習済モデルを用いて、前記指令情報を、前記第 1 の単位時間よりも短い第 2 の単位時間毎に推定する第 2 ステップと、前記第 2 ステップにおいて推定した前記指令情報を用いて前記制御対象装置を動作させる第 3 ステップと、を情報処理装置が実行する方法である。ここでは第 3 ステップまでを記したが、ステップ数は有限整数である n ステップであっても良い。

[0033] 上記の構成によれば、情報処理装置は、状態を表す情報から指令情報を導出する第 1 学習済モデルと、第 1 学習済モデルの出力に応じた情報から指令情報を導出する第 2 学習済モデルとを用いて、制御対象装置を動作させる。これらの複数の学習済モデルが制御対象装置の動作制御において階層的に用いられることにより、膨大な教師データを必要とすることなく、制御対象装

置の動作予測を安定させることができる。

[0034] 本発明の態様 1 5 に係る情報処理装置は、前記態様 1 ～ 5 において、前記第 2 推定部は、 n 個（ n は 2 以上の自然数）の前記第 2 学習済モデルを用いて前記指令情報を推定し、 n 個の前記第 2 学習済モデルは階層的に用いられ、各層の前記第 2 学習済モデルから導出される指令情報に応じた関係を表す情報が、当該第 2 学習済モデルの下層の第 2 学習済モデルの入力情報となってもよい。

[0035] 上記の構成によれば、複数の学習済モデルが階層的に用いられることにより、制御対象装置の動作予測を安定させることができる。

[0036] 本発明の態様 1 6 に係る情報処理装置は、前記態様 1 ～ 5 において、前記制御対象装置は、対象物に対する作業を行う装置であり、前記第 1 の単位時間及び前記第 2 の単位時間の一方又は両方を、前記対象物の硬軟の程度に応じて変更する変更制御部、を更に備えてもよい。

[0037] 上記の構成によれば、対象物の硬軟の程度に応じて推定周期を変更することができる。これにより、制御対象装置の動作予測を安定させることができる。

[0038] 本発明の態様 1 7 に係る情報処理装置は、前記態様 1 ～ 1 2 において、前記第 1 学習済モデルは、前記状態を表す情報から、所定の時間が経過した後の前記制御対象装置の第 2 の状態を表す情報、及び、前記指令情報を導出する学習済モデルであって、当該第 1 学習済モデルが導出した前記第 2 の状態を表す情報を入力情報として学習させた学習済モデルであってもよい。

[0039] 上記の構成によれば、第 1 学習済モデルが導出する状態を表す情報を、第 1 学習済モデルの入力情報として学習が行われ、第 1 学習済モデルが生成される。情報処理装置は、この第 1 学習済モデルを用いて制御対象装置の動作制御を行うことにより、制御対象装置の動作予測を安定させることができる。

[0040] 本発明の態様 1 8 に係る情報処理装置は、前記態様 1 ～ 1 2 において、前記第 1 学習済モデルは、前記指令情報及び前記状態を表す情報から、所定の

時間が経過した後の前記制御対象装置の第2の状態を表す情報、及び、前記指令情報を導出する学習済モデルであって、当該第1学習済モデルが導出した前記指令情報及び前記第2の状態を表す情報を、入力情報として学習させた学習済モデルであってもよい。

[0041] 上記の構成によれば、第1学習済モデルが導出する状態を表す情報を、第1学習済モデルの入力情報として学習が行われ、第1学習済モデルが生成される。情報処理装置は、この第1学習済モデルを用いて制御対象装置の動作制御を行うことにより、制御対象装置の動作予測を安定させることができる。

[0042] 本発明の態様19に係る情報処理装置は、前記態様15において、前記n個の前記第2学習済モデルは、それぞれ動作周波数が異なってもよい。

[0043] 上記の構成によれば、複数の学習済モデルを階層的に用いることにより、制御対象装置が行う複数の動作のうちそれぞれの動作に好適な動作周波数で学習させることができる。

[0044] 本発明の態様20に係る情報処理装置は、前記態様13～18において、前記入力情報とは、定量的に表される値であってもよい。

[0045] 上記の構成によれば、前記入力情報とは定量的に表される値であるため、情報処理装置は、実際の動作速度に応じて、指令速度を修正することができる。すなわち、情報処理装置は自律的に学習することができるため、容易に学習済モデルを学習させることができる。

[0046] 本発明の態様21に係る情報処理装置は、前記態様1～5において、前記制御対象装置は、把持対象物を把持又は接触対象物に接触し、前記第1学習済モデルと、前記第2学習済モデルとには、前記把持対象物を把持又は前記接触対象物に接触する動作が記憶され、前記把持対象物を把持又は前記接触対象物に接触する動作は、前記把持対象物又は前記接触対象物の硬軟の程度に応じて前記第1学習済モデル又は前記第2学習済モデルのうちいずれか学習済モデルに学習されていてもよい。

[0047] 上記の構成によれば、同一の動作であっても、対象物の硬軟に応じて、学

習される学習済みモデルが異なるため、より好適なサンプリングタイムで学習済みモデルに学習させることができる。したがって、上記の構成によれば、学習に要する時間を低減させ、学習済みモデルの記憶に要するメモリ容量を低減させることができる。

[0048] 本発明の態様 2 2 に係る情報処理装置は、前記態様 2 3 において、前記第 1 学習済みモデルは、前記把持対象物が柔らかい物体である場合の動作を学習し、前記第 2 学習済みモデルは、前記把持対象物が硬い物体である場合の動作を学習し、前記第 1 学習済みモデルと前記第 2 学習済みモデルとは、動作周波数が異なってもよい。

[0049] 上記の構成によれば、対象物が柔らかい物体である場合には、動作周波数が低い方が好適であるため、低い動作周波数の学習済みモデルに学習され、対象物が硬い物体である場合には、動作周波数が高い方が好適であるため、高い動作周波数の学習済みモデルに学習される。したがって、上記の構成によれば、より好適なサンプリングタイムで学習済みモデルに学習させることができる。したがって、上記の構成によれば、学習に要する時間を低減させ、学習済みモデルの記憶に要するメモリ容量を低減させることができる。

[0050] 本発明の態様 2 3 に係るプログラムは、上記態様 1 ～ 1 3、1 5 ～ 2 2 のいずれかの情報処理装置としてコンピュータを機能させるためのプログラムであって、前記各部としてコンピュータを機能させるためのプログラムである。

[0051] 本発明の態様 2 4 に係る情報処理装置は、前記態様 1 ～ 5 及び 7 において、前記第 1 推定部は、複数の時点における前記指令情報を推定し、前記第 2 推定部は、推定された複数の時点における前記指令情報ごとに、前記指令情報を推定してもよい。

上記の構成によれば、上位層の学習モデルと下位層の学習モデルをそれぞれ別個独立に学習させることができる。また、上記の構成によれば、少ない学習により、より正確に状態を推定することができる。

[0052] 本発明の範疇には、態様 1 9 に係るプログラム及びおよびそれを記録した

コンピュータ読み取り可能な記録媒体も含まれる。

[0053] 本発明の各態様に係る情報処理装置は、コンピュータによって実現してもよく、この場合には、コンピュータを前記情報処理装置が備える各部（ソフトウェア要素）として動作させることにより前記情報処理装置をコンピュータにて実現させる情報処理装置の制御プログラム、およびそれを記録したコンピュータ読み取り可能な記録媒体も、本発明の範疇に入る。

発明の効果

[0054] 本発明の一態様によれば、学習済モデルを用いて制御対象装置の動作制御を行う際の動作予測を安定させることができる。

図面の簡単な説明

[0055] [図1]本発明の実施形態1に係るシステムの構成を概略的に示す図である。
[図2]第1学習済モデルの一例を模式的に示した図である。
[図3]第2学習済モデルの一例を模式的に示した図である。
[図4]情報処理装置が行う処理の流れを例示するフローチャートである。
[図5]教師データを収集するためのシステムを模式的に示す図である。
[図6]本発明の実施形態2に係るシステムの構成を概略的に示す図である。
[図7]情報処理装置が行う処理の流れを例示するフローチャートである。
[図8]階層化された学習済モデルを模式的に示す図である。
[図9]本発明の実施形態3に係るシステムの構成を概略的に示す図である。
[図10]本発明の実施形態4に係るシステムの構成を概略的に示す図である。
[図11]従来手法により制御されたロボットアームの状態値の時間的な変化を表すグラフである。
[図12]従来手法により制御されたロボットアームの状態値の時間的な変化を表すグラフである。
[図13]ロボットアームの状態値の時間的な変化を表すグラフである。
[図14]本発明の実施形態5に係るシステムの構成を概略的に示す図である。
[図15]本発明の実施形態6に係るシステムの構成を概略的に示す図である。
[図16]第1学習済モデルと第3学習済モデルとの関係を例示する図である。

[図17]第1学習済モデルと第3学習済モデルとの関係を例示する図である。

[図18]本発明の実施形態7に係るシステムの構成を概略的に示す図である。

[図19]本発明の実施形態8に係るシステムの構成を概略的に示す図である。

[図20]本発明の実施形態9に係るシステムの構成を概略的に示す図である。

[図21]第1学習済モデルの内容を例示する図である。

[図22]第1学習済モデルの内容を例示する図である。

[図23]第1学習済モデルの内容を例示する図である。

[図24]本発明の実施形態9の変形例に係るシステムの構成を概略的に示す図である。

[図25]本発明の実施形態10に係るシステムの構成を概略的に示す図である。

[図26]本発明の実施形態10に係るシステムの機能構成を説明する図である。

[図27]本発明の実施形態10に係る階層化された動作生成器をより詳細に説明するための模式図である。

[図28]本発明の実施形態10に係るスレーブ装置が筆記具により文字を書く場合の一例を模式的に示す図である。

[図29]本発明の実施形態10に係るスレーブ装置により紙面に書かれた文字の一例を模式的に示す図である。

[図30]本発明の実施形態11に係るシステムを用いた場合における理想値と応答値について説明するための図である。

[図31]従来技術に係る状態の予測について説明する図である。

[図32]本発明に係るずれの蓄積の問題点について説明する図である。

[図33]本発明の実施形態6に係る状態の予測について説明する図である。

[図34]本発明の実施形態9に係る状態の予測について説明する図である。

[図35]本発明の実施形態12に係る階層化された学習済モデルをより詳細に説明するための模式図である。

[図36]本発明の実施形態13に係る状態の予測について説明する図である。

発明を実施するための形態

[0056] 〔実施形態１〕

以下、本発明の一実施形態について説明する。

[0057] 図１は、本発明の実施形態１に係る情報処理装置３０を含むシステム１の構成を概略的に示す図である。システム１は、制御対象装置を動作させるために用いられる学習済モデルを生成するシステムである。システム１は、スレーブ装置１０、マスタ装置２０、及び情報処理装置３０を備える。

[0058] スレーブ装置１０及びマスタ装置２０は、例えばロボットアームを備えるロボットである。スレーブ装置１０は、時刻 t において供給される指令情報 $MASTER(t)$ に従い動作する装置である。スレーブ装置１０は、本発明に係る制御対象装置の一例である。スレーブ装置１０は、時刻 t において、スレーブ装置１０の状態を表す状態情報 $SLAVE(t)$ を出力する。状態情報 $SLAVE(t)$ は例えば、スレーブ装置１０が備えるロボットアームの時刻 t における動作を示す情報を含む。ロボットアームの動作を示す情報は例えば、ロボットアームの位置、関節の角度などの姿勢、関節の角速度、トルク、周囲状況とロボットアームとの関係を表す画像を示す情報を含む。なお、状態情報 $SLAVE(t)$ に含まれる情報はこれらに限定されるものではなく、他の情報が含まれていてもよい。

[0059] マスタ装置２０は、ユーザ操作に応じて動作する装置である。マスタ装置２０は、指令情報を出力する装置の一例である。マスタ装置２０はスレーブ装置１０に指令情報 $MASTER(t)$ を供給する。指令情報 $MASTER(t)$ は、指令情報を出力する装置から発光される情報である。本実施形態において、指令情報 $MASTER(t)$ は、マスタ装置２０に対するユーザ操作に応じて変化したマスタ装置２０の状態を表す情報である。指令情報 $MASTER(t)$ は例えば、マスタ装置２０が備えるロボットアームの時刻 t における動作を示す情報を含む。ロボットアームの動作を示す情報は例えば、ロボットアームの位置、関節の角度などの姿勢、関節の角速度、トルク、周辺状況とロボットアームとの関係を示す画像を表す情報を含む。なお、

指令情報MASTER (t) に含まれる情報はこれらに限定されるものではなく、他の情報が含まれていてもよい。

[0060] スレーブ装置10及びマスタ装置20が行う動作は、人が行うほぼすべての動作、例えば、筆記具を用いた文字又は画像の描画動作、掃除道具を用いた掃除動作、調理動作、物品の把持動作、塗装動作、研磨動作、切削動作、穴あけ動作、又は組み付け動作等、及び、人とロボットが協調して行う動作、例えば重量物の運搬や移動などである。

[0061] 情報処理装置30は、マスタ装置20から出力される指令情報MASTER (t)、及びスレーブ装置10から出力される状態情報SLAVE (t)を用いて、制御対象装置を動作させるための学習済モデルを生成する装置である。情報処理装置30は例えば、サーバ装置、又はパーソナルコンピュータである。

[0062] 情報処理装置30は、学習フェーズ実行部31、及び、記憶部32を備える。学習フェーズ実行部31は、第1生成部311、及び第2生成部312を備える。記憶部32は、第1学習済モデル321、第2学習済モデル322、及び教師データセット323を記憶する。

[0063] 第1生成部311は、入力情報と出力情報との組を用いて、入力情報と出力情報との相関関係を学習させた第1学習済モデル321を生成する。本実施形態において、第1学習済モデル321に入力される入力情報は、スレーブ装置10が時刻tに出力した状態情報SLAVE (t)である。出力情報は、時刻tから予め定められた単位時間T1が経過した後の時刻(t + T1)において、スレーブ装置10に供給される指令情報MASTER (t + T1)である。なお、単位時間T1は、ゼロや負の値であってもよい。

[0064] 第1学習済モデル321は、状態情報SLAVE (t)から指令情報MASTER (t + T1)を生成可能な任意の学習モデルであり得る。第1学習済モデル321は例えば、CNN (Convolutional Neural Network)、RNN (Recurrent Neural Network)、LSTM (Long Short-Term Memory)、DNN (Deep Neural Network)、又はこれらの組み合わせで実現可能である。

。

[0065] 図2は、第1学習済モデル321の一例を模式的に示した図である。第1学習済モデル321には、上述のように、状態情報SLAVE(t)が入力データとして入力される。第1学習済モデル321は、例えば、畳み込み層と、プーリング層と、結合層とから成る。畳み込み層において、入力データはフィルタリングによる情報の畳み込みがなされる。畳み込みを経たデータは、プーリング層においてプーリング処理が施される。これにより、データ中の特徴の位置変化に対するモデルの認識能力が向上する。プーリング処理を経たデータは、結合層で処理されることによって、第1学習済モデル321の出力データ、すなわち、時刻(t+T1)においてスレーブ装置10に供給される指令情報MASTER(t+T1)の形式に変換されて出力される。単位時間T1は、例えばLSTMのサンプリング周期であり、例えば400msecである。

[0066] マスタ装置20及びスレーブ装置10がロボットアームを備える場合、状態情報SLAVE(t)は、例えば、ロボットアームの時刻tにおける関節の角度($\theta_{s1}(t)$, $\theta_{s2}(t)$, $\theta_{s3}(t)$)、関節の角速度($\theta_{s1}'(t)$, $\theta_{s2}'(t)$, $\theta_{s3}'(t)$)、及びトルク($\tau_{s1}(t)$, $\tau_{s2}(t)$, $\tau_{s3}(t)$)の9つのパラメータを含む。指令情報MASTER(t)も、ロボットアームの時刻tにおける関節の角度($\theta_{m1}(t)$, $\theta_{m2}(t)$, $\theta_{m3}(t)$)、関節の角速度($\theta_{m1}'(t)$, $\theta_{m2}'(t)$, $\theta_{m3}'(t)$)、及びトルク($\tau_{m1}(t)$, $\tau_{m2}(t)$, $\tau_{m3}(t)$)の9つのパラメータを含む。

[0067] 第2生成部312は、入力情報と出力情報との組を用いて、入力情報と出力情報との相関関係を学習させた第2学習済モデル322を生成する。第2学習済モデル322の入力情報は、スレーブ装置10が出力した状態情報SLAVE(t)と第1学習済モデル321に入力することにより導出される指令情報MASTER(t+T1)や、これらの情報を低相関化したデータであってもよい。例えば、図3に示すようにスレーブ装置10が出力した状

態情報SLAVE (t)と状態情報SLAVE (t)を第1学習済モデル321に入力することにより導出される指令情報MASTER ($t+T1$)との関係を表す情報DATA ($t, t+T1$)を含む形態としてもよい。また、第2学習済モデル322の出力情報は、第2の単位時間 $T2$ が経過した後に、スレーブ装置10に供給された指令情報MASTER ($t+T2$)である。すなわち、第2学習済モデル322の出力情報は、スレーブ装置10が上記第1学習済モデル321の入力情報である状態情報SLAVE (t)を出力してから第1の単位時間 $T1$ よりも短い第2の単位時間 $T2$ が経過した後にスレーブ装置10に供給された指令情報MASTER ($t+T2$)である。本実施形態において、単位時間 $T2$ は、単位時間 $T1$ よりも短くてもよい。

[0068] 第2学習済モデル322は、状態情報SLAVE (t)から指令情報MASTER (t)を生成可能な任意の学習モデルであり得る。第2学習済モデル322は例えば、CNN、RNN、LSTM、DNN、又はこれらの組み合わせで実現可能である。単位時間 $T2$ は、例えばLSTMのサンプリング周期であり、例えば20 msecである。単位時間 $T2$ は、ゼロであってもよい。

[0069] 情報DATA ($t, t+T1$)は、状態情報SLAVE (t)と指令情報MASTER ($t+T1$)との関係を表す情報である。本実施形態において、情報DATA ($t, t+T1$)は、第1学習済モデル321が出力する指令情報MASTER ($t+T1$)と相関の小さいデータである。相関の小さいデータとは、指令情報MASTER ($t+T1$)と相関関係を有しないデータ、又は、相関関係が小さいデータである。情報DATA ($t, t+T1$)は例えば、状態情報SLAVE (t)と指令情報MASTER ($t+T1$)とから導出される情報である。

[0070] 情報DATA ($t, t+T1$)は例えば、状態情報SLAVE (t)と指令情報MASTER ($t+T1$)との差分、又は比を表す情報である。情報DATA ($t, t+T1$)は、一例として、以下の(1)式で表される。状

態情報 S L A V E (t) と指令情報 M A S T E R (t + T 1) との差分、又は比を用いることにより、第 2 学習済モデル 3 2 2 の入力と出力との相関が低減する。

$$[0071] \quad DATA(t, t+T1) = SLAVE(t) - MASTER(t+T1) \dots (1)$$

図 3 は、第 2 学習済モデル 3 2 2 の一例を模式的に示した図である。図 3 の例において、第 2 学習済モデル 3 2 2 に入力される入力情報は、指令情報 M A S T E R (t + T 1) 、及び、情報 D A T A (t , t + T 1) を含む。すなわち、図 3 の例において、第 2 学習済モデル 3 2 2 には、指令情報 (t + T 1) 、及び、情報 D A T A (t , t + T 1) が入力データとして入力される。

[0072] 第 2 学習済モデル 3 2 2 は、例えば、畳み込み層と、プーリング層と、結合層とから成る。畳み込み層において、入力データはフィルタリングによる情報の畳み込みがなされる。畳み込みを経たデータは、プーリング層においてプーリング処理が施される。これにより、データ中の特徴の位置変化に対するモデルの認識能力が向上する。プーリング処理を経たデータは、結合層で処理されることによって、第 2 学習済モデル 3 2 2 の出力データ、すなわち、時刻 (t + T 2) においてスレーブ装置 1 0 に供給された指令情報 M A S T E R (t + T 2) の形式に変換されて出力される。

[0073] 図 4 は、情報処理装置 3 0 が行う学習済モデルの生成処理の流れを示すフローチャートである。なお、一部のステップは並行して、又は順序を替えて実行されてもよい。ステップ S 1 0 0 において、第 1 生成部 3 1 1 は、教師データセット 3 2 3 に含まれる教師データを用いて、第 1 学習済モデル 3 2 1 を生成する。第 1 学習済モデル 3 2 1 の学習で用いられる教師データは、上述したように、状態情報 S L A V E (t) と、指令情報 (t + T 1) との組である。

[0074] ステップ S 1 0 1 において、第 2 生成部 3 1 2 は、教師データセット 3 2 3 に含まれる教師データを用いて、第 2 学習済モデル 3 2 2 を生成する。第

2 学習済モデル 3 2 2 の学習で用いられる教師データは、上述したように、指令情報 ($t + T 1$) 及び情報 DATA ($t, t + T 1$) と、指令情報 ($t + T 2$) との組である。

[0075] 本実施形態により生成された第 1 学習済モデル 3 2 1 及び第 2 学習済モデル 3 2 2 は、制御対象装置を動作させる際に、階層的に用いられる。より具体的には、第 1 学習済モデル 3 2 1 の出力に応じた情報が、第 2 学習済モデル 3 2 2 の入力情報として用いられ、第 2 学習済モデル 3 2 2 の出力である指令情報 ($t + T 2$) が、制御対象装置の動作制御に用いられる。制御対象装置の動作制御については、後述の実施形態 2 で説明する。

[0076] 次いで、本実施形態の構成の具体例について、図面を参照しつつ説明する。

[0077] (実施例)

図 5 は、本実施形態に係る教師データを収集するためのシステムを模式的に示す図である。図 5 において、システム 2 は、マスタ装置 2 0 及びスレーブ装置 1 0 を備え、マスタ装置 2 0 の動作に応じてスレーブ装置 1 0 が動作するマスタ・スレーブシステムである。システム 2 において収集される指令情報 MASTER (t) 及び状態情報 SLAVE (t) を教師データとして、第 1 学習済モデル 3 2 1 及び第 2 学習済モデル 3 2 2 が学習される。

[0078] システム 2 は、消しゴムを用いて線消しを行う動作における指令情報 MASTER (t) 及び状態情報 SLAVE (t) を収集するためのシステムである。マスタ装置 2 0 は、ベース部 2 1 1 と、ベース部 2 1 1 の上に置かれた回転部 2 1 2 と、回転部 2 1 2 に連結された第 1 のアーム部 2 1 3 と、第 1 のアーム部 2 1 3 に連結された第 2 のアーム部 2 1 4 と、を含む。回転部 2 1 2 は、ベース部 2 1 1 に対して Z 軸周りに回転可能になっている。第 1 のアーム部 2 1 3 は、第 1 の端部が回転部 2 1 2 に連結されており、第 1 の端部を支点として Y 軸周りに揺動可能になっている。その結果、第 1 のアーム部 2 1 3 の第 2 の端部は Z 軸方向に揺動可能になっている。第 2 のアーム部 2 1 4 は、第 1 の端部が第 1 のアーム部 2 1 3 に連結されており、第 1 の

端部を支点としてY軸周りに揺動可能になっている。その結果、第2のアーム部214の第2の端部はX軸方向に揺動可能になっている。これによって、操作者は、第2のアーム部214の第2の端部を把持してX、Y、Z軸方向に移動操作することができる。

[0079] スレーブ装置10は、マスタ装置20から出力される指令情報MASTER(t)により動作する。スレーブ装置10は、ベース部111と、ベース部111の上に置かれた回転部112と、回転部112に連結された第1のアーム部113と、第1のアーム部113に連結された第2のアーム部114と、を含む。回転部112は、ベース部111に対してZ軸周りに回転可能になっている。第1のアーム部113は、第1の端部が回転部112に連結されており、第1の端部を支点としてY軸周りに揺動可能になっている。その結果、第1のアーム部113の第2の端部はZ軸方向に揺動可能になっている。第2のアーム部114は、第1の端部が第1のアーム部113に連結されており、第1の端部を支点としてY軸周りに揺動可能になっている。その結果、第2のアーム部114の第2の端部はX軸方向に揺動可能になっている。第2のアーム部114の第2の端部には、消しゴムEが取り付けられている。消しゴムEは、第2のアーム部214の第2の端部に対する操作者の移動操作に応じて移動する。

[0080] スレーブ装置10は、マスタ装置20から供給された指令情報MASTER(t)の表す状態（例えば、ロボットアームの関節の角度、角速度、及びトルク）と、自装置の状態との差分がゼロとなるよう、自装置の状態を変化させる制御を行う。

[0081] 第2のアーム部114の下には、線が描かれた紙面Pが置かれており、紙面PのZ軸方向の高さは調整可能になっている。例えば、紙面Pの高さを20mmに調整した状態において、操作者は、まず、第2のアーム部214の第2の端部を把持してZ軸方向に下ろす移動操作を行う。第2のアーム部114の第2の端部に取り付けられた消しゴムEが紙面Pに接触すると、操作者は、消しゴムEが紙面Pに接触した状態のまま第2のアーム部214の第

2の端部をX軸およびY軸の方向に移動操作する。これにより、操作者は、紙面Pに描かれている線を消しゴムEによって消す操作を行う。

[0082] 情報処理装置30は、上記の一連の操作における時系列の指令情報MASTER(t)及び状態情報SLAVE(t)を収集し、記憶部32に保存する。情報処理装置30は、収集した指令情報MASTER(t)及び状態情報SLAVE(t)を用いて、上述した第1学習済モデル321及び第2学習済モデル322の生成処理を実行する。

[0083] なお、上記のような消しゴムEによる線消しの操作を、紙面Pの高さを50mmおよび80mmに調整したそれぞれの状態においても同様に行ってもよい。

[0084] 〔変形例〕

上述の実施形態では、マスタ装置20がスレーブ装置10に指令情報MASTER(t)を供給する場合の構成について説明した。スレーブ装置10に指令情報MASTER(t)を供給する装置は、マスタ装置20に限られず、他の装置であってもよい。

[0085] 上述の実施形態では、上位層の学習済モデル（第1学習済モデル321）と、下位層の学習済モデル（第2学習済モデル322）とを含む2階層の学習済モデルを生成する場合を説明した。学習済モデルの階層は2に限らず、3以上であってもよい。階層が3以上である場合も、上述の実施形態と同様に、上位層の学習済モデルの出力に応じた情報が、下位層の学習済モデルの入力として用いられる。

[0086] すなわち、第2推定部332は、n個（nは2以上の自然数）の第2学習済モデル322を用いて前記指令情報を推定し、n個の第2学習済モデル322は階層的に用いられ、各層の第2学習済モデル322から導出される指令情報MASTERに応じた情報DATAが、第2学習済モデル322の下層の第2学習済モデル322の入力情報となる。

[0087] このとき、上位層の学習済モデルのサンプリング周期（推定周期）は、下位層の学習済モデルのサンプリング周期（推定周期）よりも長いことが好ま

しい。

[0088] 例えば、第1学習済モデル321と、2つの第2学習済モデル322a、322bとを含む3階層の学習済モデルが用いられてもよい。この場合、第1学習済モデル321の入力は、状態情報SLAVE(t)であり、出力は指令情報MASTER(t+T1)である。第1学習済モデル321の出力である指令情報MASTER(t+T1)と、情報DATA(t, t+T1)とが、第2学習済モデル322aの入力となる。第2学習済モデル322aの出力は、指令情報MASTER(t+T2a)である。

[0089] また、第2学習済モデル322aの出力である指令情報MASTER(t+T2a)と、情報DATA(t, t+T2a)とが、第2学習済モデル322aの下位層の第2学習済モデル322bの入力となる。第2学習済モデル322bの出力は、指令情報MASTER(t+T2b)である。最下層の第2学習済モデル322bの出力である指令情報MASTER(t+T2b)が、スレーブ装置10の動作制御に用いられる。

[0090] 上述の実施形態では、1台のスレーブ装置10と1台のマスタ装置20とを含むシステム1を例示したが、システム1に含まれるスレーブ装置10の台数及びマスタ装置20の台数は、2以上であってもよい。この場合、情報処理装置30は、複数のスレーブ装置10及び複数のマスタ装置20から状態情報SLAVE(t)及び指令情報MASTER(t)を収集し、収集した情報を教師データとして学習済モデルを生成する。

[0091] また、上述の実施形態では、スレーブ装置10と情報処理装置30とが別体である場合の構成を説明したが、スレーブ装置10と情報処理装置30とが一体の装置として構成されていてもよい。

[0092] [実施形態2]

本発明の他の実施形態について、以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0093] 図6は、本発明の実施形態2に係る情報処理装置30Bを含むシステム1

Bの構成を概略的に示す図である。システム1Bは、学習させた複数階層の学習済モデルを用いて制御対象装置10Bを動作させるシステムである。システム1Bは、制御対象装置10B及び情報処理装置30Bを備える。制御対象装置10Bは、指令情報MASTER(t)に従い動作する装置である。制御対象装置10Bは例えば、ロボットアームを備える装置であり、例えば、実施形態1に係るスレーブ装置10である。情報処理装置30Bは制御対象装置10Bに指令情報MASTER(t)を供給することにより、制御対象装置10Bを動作させる。

[0094] 情報処理装置30Bは、推定フェーズ実行部33、及び記憶部32を備える。推定フェーズ実行部33は、第1推定部331、第2推定部332、及び動作制御部333を備える。第1推定部331は、第1学習済モデル321を用いて、状態情報SLAVE(t)から指令情報MASTER(t+T1)を、第1の単位時間 τ_1 毎に推定する。第2推定部332は、第2学習済モデル322を用いて、情報DATA(t, t+T1)を含む入力データから指令情報MASTER(t+T2)を、第2の単位時間 τ_2 毎に推定する。動作制御部333は、第2推定部332が推定した指令情報MASTER(t+T2)を用いて、制御対象装置10Bを動作させる。第2の単位時間 τ_2 は、第1の単位時間 τ_1 よりも短い。第1の単位時間 τ_1 は、例えば400msである。第2の単位時間 τ_2 は、例えば20msである。

[0095] 図7は、情報処理装置30Bが行う制御対象装置10Bの動作制御処理の流れを示すフローチャートである。なお、一部のステップは並行して、又は順序を替えて実行されてもよい。

[0096] ステップS200において、第1推定部331は、制御対象装置10Bから状態情報SLAVE(t)を取得する。ステップS201（第1ステップの一例）において、第1推定部331は、取得した状態情報SLAVE(t)と第1学習済モデル321とを用いて、指令情報MASTER(t+T1)を、単位時間 τ_1 毎に推定する。すなわち、第1推定部331は、状態情報SLAVE(t)を第1学習済モデル321に入力することにより導出さ

れる指令情報MASTER ($t + T1$) を推定する。

[0097] ステップS202において、第1推定部331は、第1学習済モデル321に入力した状態情報SLAVE (t) と、ステップS200で推定した指令情報MASTER ($t + T1$) とを用いて、第2学習済モデル322の入力情報を算出する。例えば、指令情報MASTER ($t + T1$) と情報DATA ($t, t + T1$) の組である。情報DATA ($t, t + T1$) は、例えば上述の(1)式により算出される。

[0098] ステップS203 (第2ステップの一例) において、第2推定部332は、状態情報SLAVE (t) と指令情報MASTER ($t + T1$) との関係を表す情報DATA ($t, t + T1$) から、第2学習済モデル322を用いて、指令情報MASTER ($t, t + T2$) を単位時間 τ_2 毎に推定する。より具体的には、例えば、第2推定部332は、指令情報MASTER ($t + T1$) と情報DATA ($t, t + T1$) とを入力データとして第2学習済モデル322に入力することにより導出される指令情報MASTER ($t + T2$) を推定する。なお、本実施形態では、単位時間 τ_2 は単位時間 τ_1 よりも短いため、第2学習済モデル322に入力データが入力される頻度は、第1学習済モデル321から指令情報MASTER ($t + T1$) が出力される頻度よりも高い。そのため、第1学習済モデル321から出力されたひとつの指令情報MASTER ($t + T1$) が、複数回にわたって第2学習済モデル322に入力される場合がある。

[0099] すなわち、本実施形態では、複数の学習済モデル (第1学習済モデル321及び第2学習済モデル322) が階層化され、上位の学習済モデルの出力に応じた情報が下位層の学習済モデルの入力とされる。

[0100] 図8は、階層化された学習済モデルを模式的に示す図である。図8の例では、第1学習済モデル321、及び第2学習済モデル322が、階層的に用いられる。第1学習済モデル321は上位層の学習済モデルであり、例えばサンプリング周期が400msecのLSTMである。第2学習済モデル322は下位層の学習済モデルであり、例えばサンプリング周期が20msec

cのLSTMである。すなわち、本実施形態において、上位層の学習済モデルのサンプリング周期は、下位層の学習済モデルのサンプリング周期よりも長い。

[0101] 図8の例では、上位層である第1学習済モデル321の出力指令情報MASTER ($t+T1$) に応じた情報DATA ($t, t+T1$) が、下位層の第2学習済モデル322の入力とされる。

[0102] 図8において、第1学習済モデル321の入力情報である状態情報SLAVE (t) は、例えば、ロボットアームの時刻 t における関節の角度 ($\theta_{s1}(t), \theta_{s2}(t), \theta_{s3}(t)$)、関節の角速度 ($\theta_{s1}'(t), \theta_{s2}'(t), \theta_{s3}'(t)$)、及びトルク ($\tau_{s1}(t), \tau_{s2}(t), \tau_{s3}(t)$) の9つのパラメータを含む。また、第1学習済モデル321の出力情報である指令情報MASTER ($t+T1$) は、例えば、ロボットアームの時刻 t における関節の角度 ($\theta_{m1}(t+T1), \theta_{m2}(t+T1), \theta_{m3}(t+T1)$)、関節の角速度 ($\theta_{m1}'(t+T1), \theta_{m2}'(t+T1), \theta_{m3}'(t+T1)$)、及びトルク ($\tau_{m1}(t+T1), \tau_{m2}(t+T1), \tau_{m3}(t+T1)$) の9つのパラメータを含む。

[0103] 一方、図8において、第2学習済モデル322の入力情報は、例えば、9つのパラメータを含む指令情報MASTER ($t+T1$) と、9つのパラメータを含む情報DATA ($t, t+T1$) との計18個のパラメータを含む。

[0104] 図7の説明に戻る。ステップS204（第3ステップの一例）において、動作制御部333は、時刻 ($t+T2$) において、ステップS202で推定された指令情報MASTER ($t+T2$) を用いて制御対象装置10Bを動作させる。上述したように、指令情報MASTER ($t+T2$) は例えば、ロボットアームの時刻 ($t+T2$) における関節の角度、関節の角速度、及びトルクを含む。制御対象装置10Bは、時刻 ($t+T2$) において情報処理装置30Bから供給される指令情報MASTER ($t+T2$) に従い、ロ

ロボットアームを動作させる。例えば、制御対象装置 10B は、情報処理装置 30B から供給された指令情報 MASTER (t) の表す状態（例えば、ロボットアームの関節の角度、角速度、及びトルク）と、自装置の状態との差分がゼロとなるよう、自装置の状態を変化させる制御を行う。

[0105] ステップ S205 において、動作制御部 333 は、動作制御を終了するかを判定する。この判定は例えば、操作者により動作制御を終了する旨の操作が行われたか否かを判定することにより行われてもよい。動作制御を終了する場合（ステップ S205；YES）、動作制御部 333 は処理を終了する。一方、動作制御を終了しない場合（ステップ S205；NO）、動作制御部 333 はステップ S200 の処理に戻り、状態情報 SLAVE (t) を用いた指令情報 MASTER (t + T2) の推定処理、及び制御対象装置 10B の動作制御処理を継続する。

[0106] ところで、制御対象装置 10B に供給する指令情報 MASTER (t) をひとつの学習済モデルで推定しようとする、動作予測を安定させるためにはその学習済モデルの階層を深くする必要がある。そのため、膨大な量の教師データが必要となるとともに、計算時間も膨大になってしまう。

[0107] それに対し本実施形態では、情報処理装置 30B は、状態情報 SLAVE (t) から指令情報 MASTER (t + T1) を導出する第 1 学習済モデル 321 と、第 1 学習済モデルの出力に応じた情報から指令情報 (t + T2) を導出する第 2 学習済モデル 322 とを用いて、制御対象装置 10B を動作させる。すなわち、本実施形態では、複数の学習済モデルが階層化され、上位層の学習済モデル（第 1 学習済モデル 321）の出力が、下位層の学習済モデル（第 2 学習済モデル 322）の入力とされる。これらの複数の学習済モデルが制御対象装置 10B の動作制御において階層的に用いられることにより、膨大な教師データを必要とすることなく、制御対象装置 10B の動作予測を安定させることができる。

[0108] また、センシング周期がそれぞれ異なる複数のセンサを用いてロボットの状態を検知する場合、複数のセンサの検出値をひとつの学習済モデルの入力

としてマージすることは困難な場合があった。それに対し本実施形態によれば、学習済モデルが階層化されているため、各層の学習済モデルをそれぞれ独立して設計することができる。そのため、センシング周期がそれぞれ異なるセンサを複数用いる場合であっても、各センサに対応する学習済モデルをそれぞれ個別に設計することができ、学習済モデルの生成が容易になる。

[0109] 〔変形例〕

上述の実施形態では、制御対象である制御対象装置 10B がスレーブ装置 10 である場合を説明したが、制御対象装置 10B はスレーブ装置 10 と異なる他の装置であってもよい。

[0110] また、上述の実施形態では、制御対象装置 10B と情報処理装置 30B とが別体である場合の構成を説明したが、制御対象装置 10B と情報処理装置 30B とが一体の装置として構成されていてもよい。

[0111] 〔実施形態 3〕

本発明の他の実施形態について、以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0112] 図 9 は、本発明の実施形態 3 に係る情報処理装置 30C を含むシステム 1C の構成を概略的に示す図である。システム 1C は、スレーブ装置 10、マスタ装置 20、及び情報処理装置 30C を備える。情報処理装置 30C は、スレーブ装置 10 を動作させるための学習済モデルを生成するとともに、学習させた学習済モデルを用いて、スレーブ装置 10 を動作させる。

[0113] 情報処理装置 30C は、学習フェーズ実行部 31、記憶部 32、及び推定フェーズ実行部 33 を備える。これらの各構成要素については、上述の実施形態 1 又は実施形態 2 で説明したため、その説明を省略する。

[0114] 本実施形態によれば、情報処理装置 30C は、状態情報 $SLAVE(t)$ から指令情報 $MASTER(t+T1)$ を導出する第 1 学習済モデル 321 と、第 1 学習済モデルの出力に応じた情報 $DATA(t, t+T1)$ から指令情報 $(t+T2)$ を導出する第 2 学習済モデル 322 とを生成するとともに

に、生成した第1学習済モデル321及び第2学習済モデル322を用いて、スレーブ装置10を動作させる。すなわち、本実施形態では、複数の学習済モデルが階層化され、上位層の学習済モデル（第1学習済モデル321）の出力が、下位層の学習済モデル（第2学習済モデル322）の入力とされる。これらの複数の学習済モデルがスレーブ装置10の動作制御において階層的に用いられることにより、膨大な教師データを必要とすることなく、スレーブ装置10の動作予測を安定させることができる。

[0115] 〔実施形態4〕

本発明の他の実施形態について、以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0116] 図10は、本発明の実施形態4に係る情報処理装置30Dを含むシステム1Dの構成を概略的に示す図である。システム1Dは、スレーブ装置10、マスタ装置20、情報処理装置30D、及び入力装置40を備える。情報処理装置30Dは、スレーブ装置10を動作させるための学習済モデルを生成する学習フェーズを実行するとともに、学習させた学習済モデルを用いて、スレーブ装置10を動作させる推定フェーズを実行する。

[0117] 入力装置40は、学習フェーズにおけるスレーブ装置10の動作を規定する、動作速度を含む情報（以下、「速度情報」という）を、情報処理装置30Dに入力する。本実施形態において、スレーブ装置10は動作速度が可変である。入力装置40は例えば、操作者により操作されるキーボード、マウス、又はタッチパネルである。また、入力装置40は、インターネット等の通信ネットワークを介して情報処理装置30Dに速度情報を送信する装置であってもよい。

[0118] 本実施形態において、第1学習済モデル321Dに入力される入力データは、時刻 t における状態情報 $SLAVE(t)$ に加えて、時刻 t における速度情報 $v(t)$ を含む。すなわち、第1学習済モデル321Dは、状態情報 $SLAVE(t)$ と、速度情報 $v(t)$ とを含む入力データから、指令情報

($t + T_1$) を導出する学習済モデルである。

[0119] 例えば、スレーブ装置 10 の動作が周期的な動作である場合、速度情報 $v(t)$ は、スレーブ装置 10 の動作周期を表す周波数であってもよい。周期的な動作とは、例えば、スレーブ装置 10 のロボットアームで筆記具を用いて繰り返し線を引く動作、又は、掃除道具で行う周期的な掃除動作である。なお、速度情報 $v(t)$ は、動作周期の周波数に限られず、他の情報であってもよい。速度情報は、例えば、周波数と相関の高い情報である一つの動作の実行時間、スレーブ装置 10 の目標値、操作対象物、人間との会話情報などであっても良い。

[0120] 教師データの収集フェーズにおいて、マスタ装置 20 の操作者は、マスタ装置 20 の動作速度を変えながらマスタ装置 20 を動作させる。これにより、スレーブ装置 10 は、マスタ装置 20 から供給される指令情報 $MASTER(t)$ に応じて、自装置の動作速度を変えながら動作する。すなわち、この動作例では、スレーブ装置 10 の動作速度を変えながらスレーブ装置 10 を動作させることにより、指令情報 $MASTER(t)$ 、状態情報 $SLAVE(t)$ 、及び速度情報 $v(t)$ が収集される。

[0121] 学習フェーズにおいて、第 1 生成部 311D は、収集された指令情報 $MASTER(t)$ 、状態情報 $SLAVE(t)$ 、及び速度情報 $v(t)$ を用いて、第 1 学習済モデル 321D を生成する。第 1 学習済モデル 321D の生成に用いられる教師データは、状態情報 $SLAVE(t)$ 及び速度情報 $v(t)$ と、指令情報 $MASTER(t + T_1)$ との組である。

[0122] 推定フェーズにおいては、第 1 推定部 331 は、状態情報 $SLAVE(t)$ 及び速度情報 $v(t)$ を入力データとして第 1 学習済モデル 321D に入力することにより導出される指令情報 $MASTER(t + T_1)$ を推定する。また、第 2 推定部 332 は、上述の実施形態 3 と同様に、指令情報 $MASTER(t + T_1)$ 及び指令情報 $MASTER(t + T_1)$ に応じた情報 $DATA(t, t + T_1)$ を入力データとして第 2 学習済モデル 322 に入力することにより導出される指令情報 ($t + T_2$) を推定する。動作制御部 3

33は、推定した指令情報（ $t + T_2$ ）を、時刻（ $t + T_2$ ）においてスレーブ装置10に供給することにより、スレーブ装置10を動作させる。

[0123] ところで、スレーブ装置10の動作速度を、教師データを収集した際のスレーブ装置10の動作速度から異ならせようとする、教師データを時間軸方向に短縮又は伸張させて用いることが考えられる。しかしながら、教師データを時間軸方向に圧縮又は伸張すると、スレーブ装置10の動作が、教師データの収集時の動作と異なってしまい、動作の再現は困難である。これは、スレーブ装置10のロボットアーム等の動作において慣性力や摩擦の影響が生じるためである。

[0124] 図11及び図12は、従来の手法により制御されたロボットアームの状態値の時間的な変化を表すグラフを例示した図である。図11は、従来の情報処理装置において、教師データを時間軸方向に短縮させてスレーブ装置10を動作させた場合における、教師データと実際の状態値（スレーブ装置10から出力される状態情報SLAVE（ t ））とを表すグラフを例示した図である。また、図12は、従来の情報処理装置において、教師データを時間軸方向に伸張させてスレーブ装置10を動作させた場合における、教師データと実際の状態値（スレーブ装置10から出力される状態情報SLAVE（ t ））とを表すグラフを例示した図である。

[0125] 図11及び図12において、横軸は時刻を示し、縦軸はロボットアームの所定軸方向における角度を示す。破線が教師データの応答を示し、実線が実行時の応答を表す。図11の例では、スレーブ装置10の動作速度を、教師データを収集した際の動作速度よりも速くすると、スレーブ装置10の周期的な動作における振幅が小さくなってしまっていることがわかる。また、図12の例では、スレーブ装置10の動作速度を、教師データを収集した際の動作速度よりも遅くした場合、スレーブ装置10の周期的な動作における振幅が大きくなってしまっていることがわかる。

[0126] 図13は、本実施形態に係る情報処理装置30Dによりスレーブ装置10を動作させた場合のロボットアームの状態値の時間的な変化を示した図であ

る。図において、横軸は時刻を示し、縦軸はスレーブ装置 10 のロボットアームの所定軸方向における関節の角度を示す。図において、グラフ f 1 1 は、記憶部 3 2 に記憶された教師データの表す時間と角速度との関係を表すグラフである。グラフ f 1 2 は、教師データをスレーブ装置 10 に供給した場合におけるスレーブ装置 10 の実際の状態（ロボットアームの角度）を表すグラフである。マスタ装置 2 0 動作を予測したグラフ f 1 3 と、スレーブ装置 10 の動作を予測したグラフ f 1 2 は、元の教師データを示すグラフ f 1 1 から、差異（誤差）を持つことが明確に示されている。

[0127] グラフ f 1 3 は、本実施形態に係る情報処理装置 3 0 D によりスレーブ装置 10 を動作させた場合の状態値の時間的な変化を表すグラフである。グラフ f 1 2 の波形の振幅は、グラフ f 1 1 の波形の振幅よりも小さくなってしまっているのに対し、グラフ f 1 3 の波形の振幅は、グラフ f 1 1 の波形の振幅とほぼ同程度となっている。本実施形態に係る情報処理装置 3 0 D によれば、教師データの示すスレーブ装置 10 の動作と実際の動作との差分を小さくすることができる。すなわち、本実施形態によれば、教師データの示すスレーブ装置 10 の動作を、異なる動作速度で再現することができる。

[0128] 以上説明したように本実施形態によれば、情報処理装置 3 0 D は、スレーブ装置 10 の動作速度を示す速度情報 $v(t)$ を含む教師データを用いて学習された第 1 学習済モデル 3 2 1 D を用いて、指令情報 MASTER $(t + T_1)$ を推定する。これにより、動作速度が可変であるスレーブ装置 10 の動作予測を安定させることができる。

[0129] 〔変形例〕

上述の実施形態 4 では、スレーブ装置 10 の速度情報が入力装置 4 0 により情報処理装置 3 0 D に入力される場合を説明した。速度情報は、スレーブ装置 10 が自装置の動作速度を計測することにより得られる情報であってもよい。この場合、スレーブ装置 10 は、自装置の状態情報 SLAVE (t) を出力するとともに、自装置の動作速度を表す速度情報 $v(t)$ を出力する。情報処理装置 3 0 D は、スレーブ装置 10 から出力される状態情報 SLA

$VE(t)$ 及び速度情報 $v(t)$ を教師データとして、第1学習済モデル321Dを学習させる。

[0130] 本態様によれば、新しい教師データをスレーブ装置10が自律的に生成できるようになるため、人間が教示する教師データの数を大幅に削減できる。

[0131] 〔実施形態5〕

本発明の他の実施形態について、以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0132] 図14は、実施形態5に係るシステム1Eの構成を概略的に示す図である。システム1Eは、スレーブ装置10、マスタ装置20、情報処理装置30E、及び入力装置40を備える。情報処理装置30Eは、第2生成部312、第2学習済モデル322、及び第2推定部332を有していない点において、上述の実施形態4に係る情報処理装置30Dと異なる。

[0133] 上述の実施形態4では、情報処理装置30Dは、上述の実施形態3と同様に、階層化された複数の学習済モデル（第1学習済モデル321D及び第2学習済モデル322）を用いて、スレーブ装置10の動作制御を行った。それに対し、本実施形態では、情報処理装置30Eは、階層化された複数の学習済モデルでなく、ひとつの第1学習済モデル321Dを用いてスレーブ装置10の動作制御を行う。すなわち、本実施形態では、第2学習済モデル322の出力である指令情報 $(t+T2)$ ではなく、第1学習済モデル321Dの出力である指令情報 $(t+T1)$ が、スレーブ装置10の動作制御に用いられる。

[0134] 動作制御部333は、第1推定部331により推定された指令情報MASTER $(t+T1)$ を時刻 $(t+T1)$ においてスレーブ装置10に供給することにより、スレーブ装置10を動作させる。

[0135] 上記実施形態に係る情報処理装置30Eが実装する機能は、複数の装置により分担されて実装されてもよい。例えば、学習フェーズ実行部31を備える第1の装置と、推定フェーズ実行部33を備える第2の装置と、記憶部3

2を備える第3の装置とを含むシステムにより、情報処理装置30Eが実現されてもよい。

[0136] 本実施形態に係る情報処理装置30Eは、指令情報に従い動作する制御対象装置が出力した、当該制御対象装置の状態を表す情報と、前記制御対象装置に指示する動作を規定する、動作速度を含む情報とを含む入力データから、前記制御対象装置が当該状態情報を出力した後に前記制御対象装置に供給される前記指令情報を導出する学習済モデルを用いて、前記指令情報を推定する推定部と、前記推定部が推定した指令情報を用いて前記制御対象装置を動作させる動作制御部と、を備える装置として特定され得る。

[0137] また、本変形例に係る情報処理装置30Eは、指令情報に従い動作する制御対象装置が出力した、当該制御対象装置の状態を表す情報と、前記制御対象装置に指示する動作を規定する、動作速度を含む情報とを含む入力情報と、前記制御対象装置が当該状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報との組を用いて、前記入力情報と前記指令情報との相関関係を学習させた学習済モデルを生成する生成部、を備える装置として特定され得る。

[0138] 〔実施形態6〕

本発明の他の実施形態について、以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0139] 図15は、本発明の実施形態6に係るシステム1Fの構成を概略的に示す図である。システム1Fは、スレーブ装置10、マスタ装置20、及び情報処理装置30Fを備える。情報処理装置30Fは、第1生成部311に代えて第1生成部311Fを備えている点、及び、第3学習済モデル324を備えている点において、実施形態3の情報処理装置30Cと異なる。

[0140] 第3学習済モデル324は、指令情報MASTER(t)を入力データとし、状態情報SLAVE(t+T3)を導出する学習済モデルである。第3学習済モデル324により導出される状態情報SLAVE(t+T3)は、

第1学習済モデル321Fの入力情報として用いられる。第3学習済モデル324は例えば、CNN、RNN、LSTM、DNN、又はこれらの組み合わせで実現可能である。

[0141] 図16は、第1学習済モデル321Fと第3学習済モデル324との関係を例示する図である。図において、指令情報MASTER(t)を入力データとして第3学習済モデル324に入力した場合、第3学習済モデル324は、状態情報SLAVE($t+T3$)を導出する。第1生成部311Fは、第3学習済モデル324により導出された状態情報SLAVE($t+T3$)と、指令情報MASTER($t+T1+T3$)との組を教師データとして、第1学習済モデル321Fを学習させる。第3学習済モデル324により導出される状態情報SLAVE($t+T3$)は、本実施形態に係る第2の状態を表す情報（以下「第2状態情報」という）の一例である。

[0142] 学習フェーズにおいて、第1生成部311Fは、記憶部32Fに記憶された教師データ（指令情報MASTER(t)と状態情報SLAVE($t+T3$)との組）を学習させた第3学習済モデル324を生成する。また、第1生成部311Fは、記憶部32Fに記憶された教師データ（状態情報SLAVE($t+T3$)と指令情報MASTER($t+T1+T3$)との組）を学習させた第1学習済モデル321Fを生成する。このとき、第1学習済モデル321Fの生成に用いられる状態情報SLAVE($t+T3$)は、スレーブ装置10から出力された実際の状態値（状態情報SLAVE($t+T3$)）ではなく、第3学習済モデル324により導出された状態情報SLAVE($t+T3$)である。

[0143] また、第1学習済モデル321Fと第3学習済モデル324を連結し、両モデルを同時に学習させることも可能である。その場合、下記A～Eの5通りの方法が考えられる。

[0144] A. 状態情報SLAVE(t)から状態情報SLAVE($t+T1+T3$)を出力させる。つまり、状態情報SLAVE(t)→第1学習済モデル321F→第3学習済モデル324→状態情報SLAVE($t+T1+T3$)

とする。

[0145] B. 指令情報MASTER (t) から状態情報SLAVE (t + T₁ + T₃) を出力させる。つまり、指令情報MASTER (t) → 第3学習済モデル324 → 第1学習済モデル321F → 状態情報SLAVE (t + T₁ + T₃) とする。

[0146] C. 上記AやBを複数回 (n回) 繰り返す。例えば、指令情報MASTER (t) → 第3学習済モデル324 → 第1学習済モデル321F → … → 第3学習済モデル324 → 第1学習済モデル321F → 指令情報MASTER (t + nT₁ + nT₃) や、状態情報SLAVE (t) → 第1学習済モデル321F → 第3学習済モデル324 → … → 第1学習済モデル321F → 第3学習済モデル324 → 状態情報SLAVE (t + nT₁ + nT₃) とする。

[0147] D. 上記Aをn回繰り返し、最後の指令情報MASTERから状態情報SLAVEを予測するステップを省略する。つまり、状態情報SLAVE (t) から開始して、指令情報MASTER (t + nT₁ + (n - 1)T₃) を予測する。

[0148] E. 上記Bをn回繰り返し、最後の状態情報SLAVEから指令情報MASTERを予測するステップを省略する。つまり、指令情報MASTER (t) から開始して、状態情報SLAVE (t + nT₃ + (n - 1)T₁) を予測する。

[0149] 図17は、T₁ = T₃のときの第1学習済モデル321Fと第3学習済モデル324との関係を例示する図である。図において、矢印d11は、第1学習済モデル321Fの入力と出力の関係を示す。矢印d12は、第3学習済モデル324の入力と出力の関係を示す。図17の例では、第1学習済モデル321Fにより、状態情報SLAVE (t) から指令情報MASTER (t + T₁) が導出される。また、第3学習済モデル324により、指令情報MASTER (t + T₁) から状態情報 (t + 2T₁) が導出される。第3学習済モデル324がない場合、第1学習済モデル321Fは次のステッ

プを予測する開ループモデルである。状態情報（ t ）から指令情報を予測する第1学習済モデル321Fの応答と状態情報（ $t + 2T_1$ ）から指令情報を予測する第1学習済モデル321Fの応答とは何も関係がない。一方、第3学習済モデル324が介在する場合、状態情報（ t ）から指令情報を予測する第1学習済モデル321Fの出力が状態情報（ $t + 2T_1$ ）へ伝播するため、状態情報（ t ）における第1学習済モデル321Fの応答が時刻（ $t + T_2$ ）における第1学習済モデル321Fの応答に影響を与える閉ループモデルとなる。

[0150] すなわち、本実施形態では、第1学習済モデル321Fの学習フェーズにおいて、第3学習済モデル324の出力が第1学習済モデル321Fの入力として用いられる。

[0151] 推定フェーズにおいては、第3学習済モデル324は用いられず、第1学習済モデル321F（及び第2学習済モデル322）が指令情報の推定に用いられる。第1推定部331及び第2推定部332が行う推定処理は、上述の実施形態2で説明した処理と同様であり、ここではその詳細な説明を省略する。

[0152] 本実施形態によれば、情報処理装置30Fは、第3学習済モデル324が導出する状態情報SLAVE（ $t + T_3$ ）を、第1学習済モデル321Fの入力データとして学習を行わせ、第1学習済モデル321Fを生成する。この処理は学習フェーズにおいて閉ループとして学習をモデル化していることに相当する。一方、推定フェーズにおいては、情報処理装置30Fが第1学習済モデル321Fを用いてスレーブ装置10の動作制御を行うため閉ループをなす。よって、学習フェーズにおいて推定フェーズと同じ閉ループモデルとして学習できるため、制御対象装置の動作予測を安定させることができる。

[0153] 〔実施形態7〕

本発明の他の実施形態について、以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符

号を付記し、その説明を繰り返さない。

[0154] 図18は、実施形態7に係るシステム1Gの構成を概略的に示す図である。システム1Gは、スレーブ装置10、マスタ装置20、及び情報処理装置30Gを備える。情報処理装置30Gは、第2生成部312、第2学習済モデル322、及び第2推定部332を有していない点において、上述の実施形態6に係る情報処理装置30Fと異なる。

[0155] 上述の実施形態6では、情報処理装置30Fは、上述の実施形態3と同様に、階層化された複数の学習済モデル（第1学習済モデル321F及び第2学習済モデル322）を用いて、スレーブ装置10の動作制御を行った。それに対し、本実施形態では、情報処理装置30Gは、階層化された複数の学習済モデルでなく、第1学習済モデル321F及び第3学習済モデル324を用いてスレーブ装置10の動作制御を行う。すなわち、本実施形態では、第2学習済モデル322の出力である指令情報MASTER ($t + T2$)ではなく、第1学習済モデル321Fの出力である指令情報MASTER ($t + T1$)が、スレーブ装置10の動作制御に用いられる。

[0156] 動作制御部333は、第1推定部331により推定された指令情報MASTER ($t + T1$)を時刻($t + T1$)においてスレーブ装置10に供給することにより、スレーブ装置10を動作させる。

[0157] 上記実施形態に係る情報処理装置30Gが実装する機能は、複数の装置により分担されて実装されてもよい。例えば、学習フェーズ実行部31を備える第1の装置と、推定フェーズ実行部33を備える第2の装置と、記憶部32を備える第3の装置とを含むシステムにより、情報処理装置30Gが実現されてもよい。

[0158] 本実施形態に係る情報処理装置30Gは、指令情報に従い動作する制御対象装置が出力した、当該制御対象装置の状態を表す情報から、前記制御対象装置が当該状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報を導出する第1学習済モデルを用いて、前記指令情報を推定する推定部と、前記推定部が推定した前記指令情報を用いて前記制御対象装置を

動作させる動作制御部と、を備え、前記第 1 学習済モデルは、前記指令情報から、前記制御対象装置に当該指令情報が供給された後に前記制御対象装置が出力する第 2 の状態を表す情報を導出する第 3 学習済モデルにより導出される第 2 の状態を表す情報を、前記状態を表す情報として学習させた学習済モデルである、と特定され得る。

[0159] また、本変形例に係る情報処理装置 30G は、指令情報に従い動作する制御対象装置が出力した、当該制御対象装置の第 1 の状態を表す情報と、前記制御対象装置が当該第 1 の状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報との組を用いて、前記第 1 の状態を表す情報と前記指令情報との相関関係を学習させた第 1 学習済モデルを生成する第 1 生成部と、前記指令情報と、前記制御対象装置に当該指令情報が供給された後に前記制御対象装置が出力する、当該制御対象装置の状態を表す第 2 の状態を表す情報との組を用いて、前記指令情報と前記第 2 の状態を表す情報との相関関係を学習させた第 3 学習済モデルを生成する第 3 生成部と、を備え、前記第 1 生成部は、前記第 3 学習済モデルに導出される前記第 2 状態情報を、前記第 1 学習済モデルを学習させるための前記状態情報として用いる、と特定され得る。本実施形態に係る第 1 生成部 311F は、上記第 3 生成部の一例である。

[0160] [実施形態 8]

本発明の他の実施形態について以下に説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0161] 図 19 は、本発明の実施形態 8 に係る情報処理装置 30H を含むシステム 1H の構成を概略的に示す図である。システム 1H は、スレーブ装置 10、マスタ装置 20、情報処理装置 30H、及び硬軟センサ 50 を備える。情報処理装置 30H は、スレーブ装置 10 を動作させるための学習済モデルを生成する学習フェーズを実行するとともに、学習された学習済モデルを用いて、スレーブ装置 10 を動作させる推定フェーズを実行する。

[0162] 情報処理装置 30H の推定フェーズ実行部 33H は、第 1 推定部 331、第 2 推定部 332、動作制御部 333 に加えて、変更制御部 334 を備える。変更制御部 334 は、単位時間 τ 1 及び単位時間 τ 2 の一方又は両方を、スレーブ装置 10 が行う動作の対象となる対象物の硬軟の程度に応じて変更する。対象物は例えば、把持動作の対象となる物品、調理動作の対象となる食材、描画動作の対象となる筆記具及び筆記用紙、研磨動作の対象となる物品、である。

[0163] また、柔らかいもの等の動作周波数が低い環境との相互作用の制御や操作については単位時間の長い第 1 生成部が主に動作を推定し、硬いもの等の動作周波数が高い環境との相互作用の制御や操作については単位時間の短い第 1 生成部が主に動作を推定することもできる。

[0164] 硬軟センサ 50 は、スレーブ装置 10 に設けられ、対象物の硬軟の程度を検知し、検知した信号を情報処理装置 30H に出力する。

[0165] 変更制御部 334 は、硬軟センサ 50 が検知した対象物の硬軟の程度に応じて、単位時間 τ 1 及び単位時間 τ 2 の一方又は両方を変更する。

[0166] 変更制御部 334 は例えば、対象物が硬いほど、単位時間 τ 1 及び単位時間 τ 2 を短くする一方、対象物が軟らかいほど、単位時間 τ 1 及び単位時間 τ 2 を長くする制御を行ってもよい。

[0167] 〔変形例〕

上記実施形態に係る情報処理装置 30H が実装する機能は、複数の装置により分担されて実装されてもよい。例えば、学習フェーズ実行部 31 を備える第 1 の装置と、推定フェーズ実行部 33H を備える第 2 の装置と、記憶部 32 を備える第 3 の装置とを含むシステムにより、情報処理装置 30J が実現されてもよい。

[0168] 〔実施形態 9〕

本発明の他の実施形態について説明する。なお、説明の便宜上、上記実施形態にて説明した部材と同じ機能を有する部材については、同じ符号を付記し、その説明を繰り返さない。

[0169] 図20は、本発明の実施形態9に係る情報処理装置30Jを含むシステム1Jの構成を概略的に示す図である。システム1Jは、スレーブ装置10、マスタ装置20、及び情報処理装置30Jを備える。情報処理装置30Jは、第1学習済モデル321F、第3学習済モデル324、第1生成部311F、及び第1推定部331に代えて、第1学習済モデル321J、第1生成部311J、及び第1推定部331Jを備えている点において、上述の実施形態7と異なっている。

[0170] 図21は、第1学習済モデル321Jの内容を例示する図である。第1学習済モデル321Jは、状態情報SLAVE(k)を入力とし、状態情報SLAVE(k+n)及び指令情報SLAVE(k+m)を導出する学習済モデルである。

[0171] 学習フェーズにおいて、第1生成部311Jは、状態情報SLAVE(k)と、状態情報SLAVE(k+n)及び指令情報MASTER(k+m)との組を教師データとして、第1学習済モデル321Jを学習させる。

[0172] また、第1生成部311Jは、第1学習済モデル321Jの出力である状態情報SLAVE(k+n)を教師データとして、第1学習済モデル321Jを学習させる。すなわち、第1学習済モデル321Jにより導出された状態情報SLAVE(k+n)は、第1学習済モデル321Jの入力情報として用いられる。

[0173] 図22は、第1学習済モデル321Jを説明するための図である。図において、矢印d21は、第1学習済モデル321Jの入力と出力の関係を示す。図22の例では、上述したように、状態情報SLAVE(k)から状態情報SLAVE(t+n)及び指令情報MASTER(k+m)が導出される。

[0174] 推定フェーズにおいては、第1推定部331Jは、第1学習済モデル321Jを用いて、状態情報SLAVE(k)から、状態情報SLAVE(k+n)、及び、指令情報MASTER(k+m)を、単位時間 τ 1毎に推定する。

[0175] 動作制御部 333 は、第 1 推定部 331 が推定した指令情報 MASTER (k + m) を用いて制御対象装置（スレーブ装置 10）を動作させる。

[0176] 本実施形態では、スレーブ装置 10 の動作予測とマスタ装置 20 の動作予測とをあわせて実行することにより、これらに共通する特徴量が抽出され得る。また、本実施形態では、ひとつの第 1 学習済モデル 321 J を学習させるだけでよく、複数の学習済モデルを学習させる場合と比べて、学習が容易である。

[0177] 本実施形態に係る情報処理装置 30 J は、指令情報に従い動作する制御対象装置が出力する、当該制御対象装置の状態を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第 1 学習済モデルを用いて、前記指令情報を推定する推定部と、前記推定部が推定した前記指令情報を用いて前記制御対象装置を動作させる動作制御部と、を備え、前記第 1 学習済モデルは、前記状態を表す情報から、所定の時間が経過した後の前記制御対象装置の第 2 の状態を表す情報、及び、前記指令情報を導出する学習済モデルであって、当該第 1 学習済モデルが導出した前記第 2 の状態を表す情報を入力情報として学習させた学習済モデルである、と特定され得る。

[0178] 〔変形例〕

〔変形例 1〕

第 1 学習済モデル 321 J は、上述の実施形態 9 で説明した構成に限らず、他の構成であってもよい。

[0179] 図 23 は、本変形例に係る第 1 学習済モデル 321 K の内容を例示する図である。第 1 学習済モデル 321 K は、状態情報 SLAVE (k) 及び指令情報 MASTER (k) を入力とし、状態情報 SLAVE (k + n) 及び指令情報 SLAVE (k + m) を導出する学習済モデルである。

[0180] 学習フェーズにおいて、第 1 生成部 311 J は、状態情報 SLAVE (k) 及び指令情報 SLAVE (k) と、状態情報 SLAVE (k + n) 及び指令情報 MASTER (k + m) との組を教師データとして、第 1 学習済モデル 321 K を学習させる。

[0181] また、第1生成部311Jは、第1学習済モデル321Jの出力である状態情報SLAVE ($k+n$)、及び指令情報MASTER ($k+m$)を教師データとして、第1学習済モデル321Jを学習させる。すなわち、第1学習済モデル321Jにより導出された状態情報SLAVE ($k+n$)、及び指令情報MASTER ($k+m$)は、第1学習済モデル321Jの入力情報として用いられる。

[0182] 推定フェーズにおいては、第1推定部331Jは、第1学習済モデル321Jを用いて、状態情報SLAVE (k)及び指令情報SLAVE (k)から、状態情報SLAVE ($k+n$)、及び、指令情報MASTER ($k+m$)を、単位時間 τ 1毎に推定する。

[0183] 動作制御部333は、第1推定部331Jが推定した指令情報MASTER ($k+m$)を用いて制御対象装置(スレーブ装置10)を動作させる。また、このとき推定したSLAVE ($k+n$)と実際に計測されたSLAVE ($k+n$)の相違に関わる情報、例えば差や比など、をMASTER ($k+m$)の予測値にフィードバックすることでMASTER ($k+m$)の予測誤差を低減することができる。

[0184] 本変形例に係る情報処理装置30Jは、前記第1学習済モデルは、前記指令情報及び前記状態を表す情報から、所定の時間が経過した後の前記制御対象装置の第2の状態を表す情報、及び、前記指令情報を導出する学習済モデルであって、当該第1学習済モデルが導出した前記指令情報及び前記第2の状態を表す情報を、入力情報として学習させた学習済モデルである、と特定され得る。

[0185] [変形例2]

上記実施形態に係る情報処理装置30Jが実装する機能は、複数の装置により分担されて実装されてもよい。例えば、学習フェーズ実行部31Jを備える第1の装置と、推定フェーズ実行部33Jを備える第2の装置と、記憶部32Jを備える第3の装置とを含むシステムにより、情報処理装置30Jが実現されてもよい。

[0186] 〔変形例 3〕

実施形態 9 に係るシステム 1 J において、上述の実施形態 6 に係る第 2 生成部 3 1 2、第 2 学習済モデル 3 2 2、及び第 2 推定部 3 3 2 がシステムに含まれる構成であってもよい。

[0187] 図 2 4 は、本変形例に係るシステム 1 H の構成を概略的に示す図である。システム 1 H の情報処理装置 3 0 H は、第 2 生成部 3 1 2、第 2 学習済モデル 3 2 2、及び第 2 推定部 3 3 2 を有する点において、上述の実施形態 9 に係る情報処理装置 3 0 J と異なる。

[0188] 本変形例では、情報処理装置 3 0 H は、上述の実施形態 3 と同様に、階層化された複数の学習済モデル（第 1 学習済モデル 3 2 1 J 及び第 2 学習済モデル 3 2 2）を用いて、スレーブ装置 1 0 の動作制御を行う。すなわち、本変形例では、第 1 学習済モデル 3 2 1 J の出力である指令情報 M A S T E R（ $k + m$ ）ではなく、第 2 学習済モデル 3 2 2 の出力である指令情報（ $t + T 2$ ）が、スレーブ装置 1 0 の動作制御に用いられる。

[0189] 本変形例では、第 1 推定部 3 3 1 が推定した指令情報 M A S T E R（ $k + m$ ）、及び、指令情報 M A S T E R（ $k + m$ ）に基づき計算された D A T A（ $k + m$ ）が、第 2 学習済モデル 3 2 2 の入力として用いられ、第 2 学習済モデル 3 2 2 により指令情報 M A S T E R（ $k + T 2$ ）が導出される。第 2 学習済モデル 3 2 2 が導出した指令情報 M A S T E R（ $k + T 2$ ）を用いて動作制御部 3 3 3 がスレーブ装置 1 0 の動作制御を行う。

[0190] 〔実施形態 1 0〕

図 2 5 は、本発明の実施形態 1 0 に係る情報処理装置 3 0 L を含むシステム 1 L の構成を概略的に示す図である。システム 1 L は、スレーブ装置 1 0、マスタ装置 2 0、及び情報処理装置 3 0 L を備える。情報処理装置 3 0 L は、図 3 を参照しながら説明した情報処理装置 3 0 C の変形例である。図 2 5 を参照しながら、情報処理装置 3 0 L について説明する。上述の実施形態 3 において説明した構成については、同様の符号を付すことにより説明を省略する場合がある。

- [0191] 第1生成部311は、時刻 t におけるスレーブ装置10の状態情報SLAVE(t)と、時刻 t におけるマスタ装置20の指令情報MASTER(t)と、時刻 $t+1$ におけるマスタ装置20の指令情報MASTER($t+1$)とに基づき、第1学習済モデル321を生成する。なお、第1生成部311は、教師データセット323に基づいて、第1学習済モデル321を生成してもよい。
- [0192] 第2生成部312は、時刻 t におけるスレーブ装置10の状態情報SLAVE(t)と、時刻 t におけるマスタ装置20の指令情報MASTER(t)と、時刻 $t+1$ におけるマスタ装置20の指令情報MASTER($t+1$)とに基づき、第2学習済モデル322を生成する。なお、第1生成部311は、教師データセット323に基づいて、第1学習済モデル321を生成してもよい。
- [0193] 第1推定部331は、学習された第1学習済モデル321と、時刻 t におけるスレーブ装置10の状態情報SLAVE(t)に基づいて、指令情報MASTER($t+T1$)を推定する。
- [0194] 第2推定部332は、学習された第2学習済モデル322と、時刻 t におけるスレーブ装置10の状態情報SLAVE(t)に基づいて、指令情報MASTER($t+T1$)を推定する。
- [0195] 図26は、本発明の実施形態10に係る情報処理装置30Lを含むシステム1Lの機能構成を説明する図である。同図を参照しながら、情報処理装置30Lの機能構成について説明する。情報処理装置30Lは、動作生成器610-1と、動作生成器610-2と、動作生成器610-3とを備える。以降の説明において、動作生成器610-1、動作生成器610-2及び動作生成器610-3を区別しない場合には、単に動作生成器610と記載する場合がある。また、動作生成器610-3に対し動作生成器610-1を上位層と記載し、動作生成器610-1に対し動作生成器610-3を下位層と記載する場合がある。
- [0196] 動作生成器610-1は、第1生成部311と、第1推定部331とを含

む。動作生成器610-1は、スレーブ装置10の状態情報SLAVE(t)と、時刻tにおけるマスタ装置20の指令情報MASTER(t)と、時刻t+1におけるマスタ装置20の指令情報MASTER(t+1)とに基づき、第1学習済モデル321を生成する。動作生成器610-1に入力されるスレーブ装置10の状態情報SLAVE(t)と、時刻tにおけるマスタ装置20の指令情報MASTER(t)と、時刻t+1におけるマスタ装置20の指令情報MASTER(t+1)とを区別しない場合には、入力信号IN1と記載する場合がある。

動作生成器610-1は、学習された第1学習済モデル321と、時刻tにおけるスレーブ装置10の状態情報SLAVE(t)に基づいて、指令情報MASTER(t+T1)を推定し、スレーブ装置20に出力する。また、動作生成器610-1は、学習された第1学習済モデル321と、時刻tにおけるスレーブ装置10の状態情報SLAVE(t)に基づいて、下位層の動作生成器610-2に対し、割り込みINT1を発生させる。動作生成器610-1が出力する信号を、出力信号OUT1と記載する場合がある。

[0197] 動作生成器610-2は、第2生成部312-1と、第2推定部332-2とを含む。動作生成器610-2は、スレーブ装置10の状態情報SLAVE(t)と、時刻tにおけるマスタ装置20の指令情報MASTER(t)と、時刻t+1におけるマスタ装置20の指令情報MASTER(t+1)とに基づき、第2学習済モデル322-1を生成する。動作生成器610-2に入力されるスレーブ装置10の状態情報SLAVE(t)と、時刻tにおけるマスタ装置20の指令情報MASTER(t)と、時刻t+1におけるマスタ装置20の指令情報MASTER(t+1)とを区別しない場合には、入力信号IN2と記載する場合がある。

動作生成器610-2は、学習された第2学習済モデル322-1と、時刻tにおけるスレーブ装置10の状態情報SLAVE(t)に基づいて、指令情報MASTER(t+T1)を推定し、スレーブ装置20に出力する。また、動作生成器610-2は、学習された第2学習済モデル322と、時

刻 t におけるスレーブ装置 10 の状態情報 $SLAVE(t)$ に基づいて、下位層の動作生成器 610-3 に対し、割り込み $INT2$ を発生させる。動作生成器 610-2 が出力する信号を、出力信号 $OUT2$ と記載する場合がある。

[0198] 動作生成器 610-3 は、第 2 生成部 312-2 と、第 2 推定部 332-2 とを含む。動作生成器 610-3 は、スレーブ装置 10 の状態情報 $SLAVE(t)$ と、時刻 t におけるマスタ装置 20 の指令情報 $MASTER(t)$ と、時刻 $t+1$ におけるマスタ装置 20 の指令情報 $MASTER(t+1)$ とに基づき、第 2 学習済モデル 322-2 を生成する。動作生成器 610-3 に入力されるスレーブ装置 10 の状態情報 $SLAVE(t)$ と、時刻 t におけるマスタ装置 20 の指令情報 $MASTER(t)$ と、時刻 $t+1$ におけるマスタ装置 20 の指令情報 $MASTER(t+1)$ とを区別しない場合には、入力信号 $IN3$ と記載する場合がある。

動作生成器 610-3 は、学習された第 2 学習済モデル 322-2 と、時刻 t におけるスレーブ装置 10 の状態情報 $SLAVE(t)$ に基づいて、指令情報 $MASTER(t+T1)$ を推定し、スレーブ装置 20 に出力する。動作生成器 610-3 が出力する信号を、出力信号 $OUT3$ と記載する場合がある。

[0199] 図 27 は、本発明の実施形態 10 に係る階層化された動作生成器 610 をより詳細に説明するための模式図である。同図を参照しながら、実施形態 10 に係る階層化された動作生成器 610 について、より詳細に説明する。

[0200] 動作生成器 610-1 は、例えば自己位置推定を行う。自己位置推定とは、例えば、スレーブ装置 20 が筆記具により文字を書く場合の一例において、スレーブ装置 20 が把持する筆記具の位置を推定することであってもよい。動作生成器 610-1 は、最上位層の動作生成器 610 のため、その他、下位層の動作生成器 610 に比べ、遅い周波数によりある程度の正確さを有していれば足りる動作についての推定を行う。

[0201] 動作生成器 610-2 は、例えば画像認識を行う。画像認識とは、例えば

、スレーブ装置 20 が筆記具により文字を書く場合の一例において、用紙に書かれた文字及び、文字を書こうとしている用紙についての画像認識をであってもよい。動作生成器 610-2 は、上位層の動作生成器 610 及び下位層の動作生成器 610 の間に位置するため、上位層の動作生成器 610 よりも速い周波数であって、下位層の動作生成器 610 よりも遅い周波数を要する動作についての推定を行う。

[0202] 動作生成器 610-3 は、例えばリアルタイム制御を行う。リアルタイム制御とは、例えば、スレーブ装置 20 が筆記具により文字を書く場合の一例において、筆記具を把持する力や、筆記具を動かす早さ等のリアルタイムな制御を必要とする動作についての推定であってもよい。動作生成器 610-3 は、最下位層の動作生成器 610 のため、その他、上位層の動作生成器 610 に比べ、速い周波数により正確さを要する動作についての推定を行う。

[0203] すなわち、動作生成器 610-1 の動作周波数は、動作生成器 610-2 の動作周波数よりも低速であり、動作生成器 610-2 の動作周波数は、動作生成器 610-3 の動作周波数より低速である。また、動作生成器 610-1 は、動作生成器 610-2 よりも複雑な動作を制御し、動作生成器 610-2 は、動作生成器 610-3 よりも複雑な動作を制御する。すなわち、本実施形態によれば、上位層の動作生成器 610 は、下位層の動作生成器 610 よりも、より低速で複雑な動作を制御し、下位層の動作生成器 610 は、上位層の動作生成器 610 よりも、より高速で単純な動作を制御する。

[0204] 本実施形態によれば、上位層の動作生成器 610 から下位層の動作生成器 610 に割り込みを発生させることにより、それぞれの動作生成器 610 は、それぞれの動作に応じて、個別に学習されることができる。例えば自己位置推定のような低速で複雑な動作を制御する場合、動作生成器 610 は、低い周波数で学習される。また、リアルタイム制御のような高速で単純な動作を制御する場合、動作生成器 610 は、高い周波数で学習される。ここで、低速で複雑な動作を学習する場合に、高い周波数で学習すると、記憶容量を膨大に消費する場合がある、また、高速で単純な動作を学習する場合に、低

い周波数で学習すると、サンプリング周波数が不足することによる誤差が大きくなってしまう場合がある。本実施形態においては、動作に応じて学習済モデルを階層化することにより、それぞれの動作に好適な周波数で学習されることができるため、無駄なメモリ容量を要することがない。

[0205] 図28は、本発明の実施形態10に係るスレーブ装置が筆記具により文字を書く場合の一例を模式的に示す図である。本実施形態の効果を説明するための具体的な一例として、システム2がスレーブ装置10に、筆記具Wにより文字を書かせる場合の一例について説明する。図28におけるスレーブ装置10は、消しゴムEに代えて筆記具Wを備える点において、図5に記載のスレーブ装置10とは異なる。スレーブ装置10は、筆記具Wにより、紙面Pに文字を書く。

[0206] 図29は、本発明の実施形態10に係るスレーブ装置により紙面に書かれた文字の一例を模式的に示す図である。本実施形態におけるスレーブ装置10は、例えば、図29(A)に示すような文字Cを書いてもよいし、図29(B)に示すような文字列CSを書いてもよい。同図に示す一例において、文字列CSが並ぶ方向をx軸方向と、x軸方向に直行する方向をy方向と記載する場合がある。

[0207] 図29(A)に示す一例では、スレーブ装置10は、文字Cとしてアルファベットの“A”を書く。図29(B)に示す一例では、スレーブ装置10は、文字列CSとしてアルファベットの“ABC”を書く。文字列“ABC”を書く場合には、文字Cを書いてから次の文字Cを書く場所まで筆記具Wを移動させる必要がある。図29(B)に示す一例においては、スレーブ装置10は、x軸上の範囲L1においてアルファベット“A”を書いた後、距離L2の分だけ筆記具Wをx軸方向に移動させ、次のアルファベット“B”を書く。また、スレーブ装置10は、範囲L3においてアルファベット“B”を書いた後、距離L4の分だけ筆記具Wをx軸方向に移動させ、次のアルファベット“C”を書く。

[0208] ここで、範囲L1、L2又はL3に文字を記載する場合は、動作周波数は

高い方が好適である。しかしながら、距離 L_2 又は L_4 を x 軸方向に移動させる場合は、動作周波数は低くてもよいため、メモリの消費量や学習に要する時間等を考慮すると、動作周波数が低い方が好適である。すなわち、スレーブ装置10により文字列CSを書く場合、動作周波数が高い方が好適な場合と、動作周波数が低い方が好適な場合が混在する。本実施形態によれば、動作生成器610を階層化する事により、それぞれの動作に好適な動作周波数で学習させることができる。

[0209] 〔実施形態11〕

図30は、本発明の実施形態11に係るシステムを用いた場合における理想値と応答値について説明するための図である。実施形態11は、実施形態4の変形例である。同図を参照しながら、実施形態11に係るシステムについてより詳細に説明する。本実施形態において、スレーブ装置10は、動作の結果、自装置が動作する予定の速度情報（すなわち、指令速度）と計測された自装置の動作速度（すなわち、動作速度）とを比較することにより、第1学習済モデル321Dを学習させる。図30に示す一例は、指令速度と動作速度との対応関係を示す。

[0210] 図30には、理想値と、応答値が示されている。理想値とは、指令速度と動作速度とが同一となる場合を指す。例えば、指令速度が速度 v_{12} である場合における動作速度の理想値は速度 v_{22} である。しかしながら、この一例においては、指令速度が速度 v_{12} である場合における実際の動作速度は、速度 v_{21} である。この場合、第1学習済モデル321Dは、指令速度が速度 v_{12} であると学習されていた速度を、実際の動作速度 v_{21} に対応する指令速度の理想値である速度 v_{11} であると学習する。

[0211] 本実施形態において、第1学習済モデル321の入力情報及び出力情報は、速度である。すなわち、第1学習済モデル321の入力情報及び出力情報は、定量的に表すことが可能な値である。したがって、第1学習済モデル321は、入力情報と出力情報との関係から、自律的に学習することができる。また、本実施形態のように構成することにより、情報処理装置30Dは、容

易に第1学習済モデル321を学習させることができる。

[0212] 〔実施形態6及び実施形態9に係る状態の予測〕

図31から図34を参照しながら、実施形態6において図17を参照しながら説明した状態の予測と、実施形態9において図22を参照しながら説明した状態の予測について、より詳細に説明する。

[0213] 図31は、従来技術に係る状態の予測について説明する図である。同図を参照しながら、制御対象装置の状態と、当該状態における未来の状態の予測について説明する。なお、同図における説明において横軸を時間 t として示す。従来技術においては、スレーブ装置とマスタ装置とを備えていないため、入力と出力とが同一の値となり、制御対象装置は、自装置の未来の状態を予測できた。すなわち、従来技術によれば、時刻 $t = k$ の状態において時刻 $t = k + 1$ の状態を予測し、時刻 $t = k + 1$ の状態においては、自装置の予測した結果を初期値として時刻 $t = k + 2$ の状態を予測することができた。すなわち、従来技術によれば、自装置の状態に応じて未来の状態を予測するため、遠い未来を予測する場合であっても、ずれが蓄積することがなかった。

[0214] 図32は、本発明に係るずれの蓄積の問題点について説明する図である。同図を参照しながら、スレーブ装置10と、マスタ装置20とを用いた場合の、制御対象装置の状態と、当該状態における未来の状態の予測について説明する。図32(A)は、学習時における一例についての図であり、図32(B)は、実行時における一例の図である。

[0215] まず、図32(A)を参照しながら学習時における未来の予測について説明する。学習時、学習済モデルは、スレーブ装置10がある状態（例えば時刻 $t = k$ におけるスレーブ装置10の状態）に基づいて、未来のマスタ装置20の状態（例えば時刻 $t = k + 1$ におけるマスタ装置20の状態）を予測する。また、学習済モデルは、時刻 $t = k + 1$ におけるスレーブ装置10の状態に基づいて、さらに未来のマスタ装置20の状態（例えば時刻 $t = k + 2$ におけるマスタ装置20の状態）を予測する。すなわち、学習済モデルが

時刻 $t = k$ において予測した状態と、時刻 $t = k + 1$ において予測した状態とは、それぞれ独立しているため、開ループを構成する。

[0216] 次に、図 3 2 (B) を参照しながら実行時における未来の予測について説明する。実行時において、学習済モデルは、スレーブ装置 1 0 の状態と、マスタ装置 2 0 からの応答とに基づいて、未来のマスタ装置 2 0 の状態を予測する。例えば、時刻 $t = k + 1$ において、学習済モデルは、時刻 $t = k + 1$ におけるスレーブ装置 1 0 の状態と、マスタ装置 2 0 からの応答に基づいて、時刻 $t = k + 2$ におけるマスタ装置 2 0 の状態を予測する。すなわち、学習済モデルが時刻 $t = k$ において予測した状態に基づいて動作した結果に応じて、時刻 $t = k + 1$ において未来の状態を予測するため、閉ループを構成する。よって、図 3 2 に示したモデルによれば、学習時と実行時でモデルが異なっている点において問題がある。図 3 2 に示したモデルによれば、学習時には開ループであるにもかかわらず、実行時には閉ループが構成されるため、学習時には考慮されていないずれが実行時には生じており、遠い未来を予測する場合には当該ずれが蓄積してしまうことによって、未来の状態の予測が正確にできないといった問題があった。なお、ずれが大きい場合には、制御が不安定になる恐れがある。

[0217] 図 3 3 は、本発明の実施形態 6 に係る状態の予測について説明する図である。本発明の実施形態 6 に係る状態の予測は、図 3 2 を参照しながら説明した問題点を解決しようとするものである。図 3 3 を参照しながら、本発明の実施形態 6 における、制御対象装置の状態と、当該状態における未来の状態の予測について説明する。図 3 3 (A) は、学習時における一例についての図であり、図 3 3 (B) は、実行時における一例の図である。

[0218] まず、図 3 3 (A) を参照しながら学習時における未来の予測について説明する。学習時、学習済モデルは、時刻 $t = k$ におけるスレーブ装置 1 0 の状態に基づいて、時刻 $t = k + 1$ におけるマスタ装置 2 0 の状態を予測する（以降、スレーブ装置 1 0 の状態に基づくマスタ装置 2 0 の状態の予測を“人間の行動予測”と記載する。）。また、学習済モデルは、時刻 $t = k + 1$

におけるマスタ装置 20 の状態に基づいて、時刻 $t = k + 2$ におけるスレーブ装置 10 の状態を予測する（以降、マスタ装置 20 の状態に基づくスレーブ装置 10 の状態の予測を“環境の応答予測”と記載する。）。すなわち、本実施形態においては、学習済モデルが時刻 $t = k$ において予測した状態と、当該予想に基づいた応答を予測する。

[0219] 次に、図 33（B）を参照しながら実行時における未来の予測について説明する。実行時において、学習済モデルは、スレーブ装置 10 の状態と、マスタ装置 20 からの応答とに基づいて、未来のマスタ装置 20 の状態を予測する。例えば、時刻 $t = k + 1$ において、学習済モデルは、時刻 $t = k + 1$ におけるスレーブ装置 10 の状態と、マスタ装置 20 からの応答に基づいて、時刻 $t = k + 2$ におけるマスタ装置 20 の状態を予測する。すなわち、学習済モデルが時刻 $t = k$ において予測した状態に基づいて動作した結果に応じて、時刻 $t = k + 1$ において未来の状態を予測するため、閉ループを構成する。しかしながら本実施形態においては、時刻 $t = k + 1$ におけるマスタ装置 20 からの応答は、“環境の応答予測”によって、既に予測されているため、ずれが蓄積することがない。よって、図 31 に示したモデルによれば、未来の状態を、より正確に予測することができる。

[0220] ここで、図 33 に示したモデルによれば、学習時に学習された“環境の応答予測”は、実行時には用いられないことがない。したがって、図 33 に示したモデルによれば、実行時は“環境の応答予測”を使用しないため、学習時に時間とメモリ容量を消費して収集したデータを、実行時には活用できないといった問題点があった。

[0221] 図 34 は、本発明の実施形態 9 に係る状態の予測について説明する図である。本発明の実施形態 6 に係る状態の予測は、図 33 を参照しながら説明した方法により新たに生じた問題点を解決しようとするものである。図 34 を参照しながら、本発明の実施形態 9 における、制御対象装置の状態と、当該状態における未来の状態の予測について説明する。図 34（A）は、学習時における一例についての図であり、図 34（B）は、実行時における一例の

図である。

[0222] まず、図34（A）を参照しながら学習時における未来の予測について説明する。学習時、学習済モデルは、時刻 $t = k$ におけるスレーブ装置10の状態に基づいて、時刻 $t = k + 1$ におけるマスタ装置20の状態と、時刻 $t = k + 2$ におけるスレーブ装置10の状態とを予測する。すなわち、本実施形態においては、学習済モデルは、スレーブ装置10の状態に基づいて所定時間経過後のマスタ装置20の状態と、所定時間よりさらに先の時刻におけるスレーブ装置10の状態とを予測する。本実施形態においては、スレーブ装置10の状態に基づいて“環境の応答予測”を行う点において、図33における一例とは異なる。

[0223] 次に、図34（B）を参照しながら実行時における未来の予測について説明する。実行時において、学習済モデルは、スレーブ装置10の現在の状態と、マスタ装置20からの応答と、スレーブ装置10の過去の状態から予測された状態とに基づいて、未来のスレーブ装置10の状態と、未来のマスタ装置20の状態とを予測する。例えば、時刻 $t = k + 1$ において、学習済モデルは、時刻 $t = k + 1$ におけるスレーブ装置10の状態と、マスタ装置20からの応答と、時刻 $t = k$ の時点で予測されたスレーブ装置10の状態とに基づいて、時刻 $t = k + 2$ におけるマスタ装置20の状態と、スレーブ装置10の状態とを予測する。すなわち、本実施形態によれば、実際の応答と、過去に予測した情報に基づいて制御をすることができるため、未来の状態を、より正確に予測することができる。

[0224] 〔実施形態12〕

図35は、本発明の実施形態12に係る階層化された学習済モデルをより詳細に説明するための模式図である。実施形態12は、実施形態8の変形例である。同図を参照しながら、実施形態12に係るシステムについて説明する。実施形態8においては、システム1Hが硬軟センサ50を備えることにより、把持動作の対象となる物品の硬軟の程度に応じて単位時間 τ_1 及び単位時間 τ_2 の一方又は両方を変更する場合の一例について説明したが、実施

形態１２においては、硬軟センサ５０を備えない点において実施形態８とは異なる。

[0225] 実施形態１２においては、情報処理装置３０Ｈは、複数の学習済みモデルを備え、把持動作の対象となる物品の硬軟に応じて、異なる学習済みモデルに学習される。具体的には、情報処理装置３０Ｈは、複数の動作生成器６１１－ n （ n は、１以上の整数）により構成される。図３５に示す一例においては、動作生成器６１１－１と、動作生成器６１１－２とにより構成される。動作生成器６１１－１及び動作生成器６１１－２は、いずれも把持動作を行う。動作生成器６１１－１は、動作生成器６１１－２に比べて動作周波数が低く、動作生成器６１１－２は、動作生成器６１１－１に比べて動作周波数が低い。

[0226] ここで、物体を把持する動作においては、物体の硬軟に応じて学習に好適なサンプリングタイムは異なる。例えば、柔らかい物体を把持する場合のサンプリングタイムは、硬い物体を把持する場合のサンプリングタイムに比べて、長くてもよい。本実施形態において、スレーブ装置１０は、柔らかい物体を把持する動作は、動作周波数の低い動作生成器６１１－１に学習され、硬い物体を把持する動作は、動作周波数の高い動作生成器６１１－２に学習される。すなわち、本実施形態において、スレーブ装置１０は、物体の硬軟に応じて、学習される学習済みモデルが異なる。

[0227] なお、実施形態１２における構成は、実施形態１０で説明した階層構造と同様の構成を有するが、実施形態１２においては自装置の動きに応じて学習される学習済みモデルが異なり、実施形態１２においては、把持対象となる物体に応じて学習される学習済みモデルが異なる点において、両者は異なる。実施形態１０の構成と、実施形態１２の構成とを組み合わせた構成を用いてもよい。

[0228] 実施形態１２のように構成することにより、より好適なサンプリングタイムで学習済みモデルに学習させることができる。実施形態１２によれば、より好適なサンプリングタイムで学習済みモデルに学習させることができるため、

学習に要する時間を低減させ、学習済モデルの記憶に要するメモリ容量を低減させることができる。

[0229] 〔実施形態 13〕

図 36 は、本発明の実施形態 13 に係る状態の予測について説明する図である。実施形態 13 は、上位層の第 1 推定部 331 が複数の異なる時点における指令情報を推定し、第 2 推定部 332 が推定された複数の時点における指令情報ごとに指令情報を推定する点において上述した実施形態とは異なる。複数の異なる時点とは、過去の時点であってもよいし、現在であってもよいし、未来の時点であってもよい。

同図において、上位層の周期が下位層の 5 倍である場合の一例について説明する。上位層の第 1 推定部 331 は、時刻 $k + 5$ 、時刻 $k + 10$ 及び時刻 $k + 15$ の 3 つの時点における状態を予測する。

[0230] 図 36 (A) から図 36 (C) は、時刻 k から時刻 $k + 5$ の間に、下位層である第 2 推定部 332 の推定について説明するための図である。

図 36 (A) において、第 2 推定部 332 は、時刻 $k + 1$ における状態を予測する。図 36 (B) において、第 2 推定部 332 は、時刻 $k + 2$ における状態を予測する。図 36 (C) において、第 2 推定部 332 は、時刻 $k + 3$ における状態を予測する。

第 2 推定部 332 は、上位層である第 1 推定部 331 からの入力を更新しないことにより、上位層と下位層とを非干渉化させることができる。

[0231] 上位層の第 1 推定部 331 が複数の異なる時点における指令情報を推定し、第 2 推定部 332 が推定された複数の時点における指令情報ごとに指令情報を推定することにより、指令情報の時間変化に関わる情報を下位層に伝達することができる。そのため例えば、上位層の出力が位置情報のみである場合でも、等価的に速度に関わる情報を下位層が得ることができる。よって、上位層が画像情報や GPS など、速度情報を生成しづらいつセンサを用いて指令情報を出力する場合でも適切な指令情報を階層へ伝達できるようになる。

[0232] また、第 1 生成部 311 が第 1 学習済モデル 321 を生成する際に用いる

入力情報と出力情報との組のうち、出力情報とは、スレーブ装置 10 に供給される指令情報 MASTER ($t + T1$) である場合の一例に限定されず、状態情報 SLAVE ($t + T1$) 等、指令情報 MASTER ($t + T1$) と相関の高い情報を出力するモデルとしてもよい。

特に、状態情報 SLAVE ($t + T1$) を予測することは有効である。なぜならば、指令情報 MASTER ($t + T1$) を予測する場合、その予測結果が正しかったかどうかを自律動作時にはスレーブしか存在せず、マスタが存在しないため確かめることができない。一方、状態情報 SLAVE ($t + T1$) を予測する場合、自律動作時でもスレーブが存在するため一ステップ後の応答値と比較することで予測誤差を評価することができる。よって、予測誤差を低減させることや、予測を安定させることが容易になる。

[0233] また、第 1 学習済モデル 321 は、状態情報 SLAVE (t) から指令情報 MASTER ($t + T1$) を生成可能な任意の学習モデルである場合の一例に限定されず、状態情報 SLAVE ($t + T1$) など指令情報 MASTER ($t + T1$) と相関の高い情報を出力するモデルとしてもよい。上述した説明における第 1 学習済モデル 321 の出力する指令情報 MASTER ($t + T1$) は、指令情報 MASTER ($t + T1$) と相関の高い情報に代替可能である。

特に、状態情報 SLAVE ($t + T1$) を予測することは有効である。なぜならば、指令情報 MASTER ($t + T1$) を予測する場合、その予測結果が正しかったかどうかを自律動作時にはスレーブしか存在せず、マスタが存在しないため確かめることができない。一方、状態情報 SLAVE ($t + T1$) を予測する場合、自律動作時でもスレーブが存在するため一ステップ後の応答値と比較することで予測誤差を評価することができる。よって、予測誤差を低減させることや、予測を安定させることが容易になる。

[0234] また、第 2 推定部 332 が推定に用いる n 個の第 2 学習済モデル 322 は、 n が 2 以上の自然数である場合に限定されず、1 個、すなわち n が 1 以上の自然数であってもよい。

本実施形態において、第1推定部331と、第2推定部332とは、それぞれ異なる学習済モデルを用いる。したがって、学習段階において、第1推定部331が用いる第1学習済モデル321と、第2推定部332が用いる第2学習済モデル322とを、それぞれ別個独立に学習させることができるため設計コストを大幅に低減できる。また、新たな動作を学習させる必要が生じた場合にも、第2学習済モデル322を新たに学習させ、第1学習済モデル321を流用することにより、学習する量を少なくすることができる。同様に、第1学習済モデル321を新たに学習させ、第2学習済モデル322を流用することにより、学習する量を少なくすることもできる。

[0235] また、情報処理装置30Hの推定フェーズ実行部33Hにおいて、変更制御部334は、把持動作の対象となる物品、調理動作の対象となる食材、描画動作の対象となる筆記具及び筆記用紙、研磨動作の対象となる物品に加え、単に接触する接触対象物の硬軟の程度に応じて、単位時間 τ_1 及び単位時間 τ_2 の一方又は両方を変更してもよい。すなわち、制御対象装置（マスタ装置及びスレーブ装置）の動作は把持動作に限定されず、単に接触する動作も含まれる。

[0236] [ソフトウェアによる実現例]

情報処理装置30、30B～30Lの制御ブロック（特に学習フェーズ実行部31、推定フェーズ実行部33）は、集積回路（ICチップ）等に形成された論理回路（ハードウェア）によって実現してもよいし、ソフトウェアによって実現してもよい。

[0237] 後者の場合、情報処理装置30、30B～30Lは、各機能を実現するソフトウェアであるプログラムの命令を実行するコンピュータを備えている。このコンピュータは、例えば1つ以上のプロセッサを備えていると共に、上記プログラムを記憶したコンピュータ読み取り可能な記録媒体を備えている。そして、上記コンピュータにおいて、上記プロセッサが上記プログラムを上記記録媒体から読み取って実行することにより、本発明の目的が達成される。上記プロセッサとしては、例えばCPU（Central Processing Unit）を

用いることができる。上記記録媒体としては、「一時的でない有形の媒体」、例えば、ROM (Read Only Memory) 等の他、テープ、ディスク、カード、半導体メモリ、プログラマブルな論理回路などを用いることができる。また、上記プログラムを展開するRAM (Random Access Memory) などをさらに備えていてもよい。また、上記プログラムは、該プログラムを伝送可能な任意の伝送媒体（通信ネットワークや放送波等）を介して上記コンピュータに供給されてもよい。なお、本発明の一態様は、上記プログラムが電子的な伝送によって具現化された、搬送波に埋め込まれたデータ信号の形態でも実現され得る。

[0238] 本発明は上述した各実施形態に限定されるものではなく、請求項に示した範囲で種々の変更が可能であり、異なる実施形態にそれぞれ開示された技術的手段を適宜組み合わせて得られる実施形態についても本発明の技術的範囲に含まれる。

符号の説明

[0239] 1、1 B、1 C、1 D、1 E、1 F、1 G、1 H、1 I、1 J、1 K、1 L
、2 システム
10 スレーブ装置
10 B 制御対象装置
20 マスタ装置
30、30 B、30 C、30 D、30 E、30 F、30 G、30 H、30 I
、30 K、30 K、30 L 情報処理装置
31 学習フェーズ実行部
32 記憶部
33 推定フェーズ実行部
40 入力装置
111、211 ベース部
112、212 回転部
113、213 第1のアーム部

- 1 1 4、2 1 4 第2のアーム部
- 3 1 1、3 1 1 D、3 1 1 F 第1生成部
- 3 1 2 第2生成部
- 3 2 1、3 2 1 D、3 2 1 F 第1学習済モデル
- 3 2 2 第2学習済モデル
- 3 2 3 教師データセット
- 3 2 4 第3学習済モデル
- 3 3 1 第1推定部
- 3 3 2 第2推定部
- 3 3 3 動作制御部
- 6 1 0、6 1 0-1、6 1 0-2、6 1 0-3 動作生成器

請求の範囲

- [請求項1] 指令情報に従い動作する制御対象装置が出力する、当該制御対象装置の状態を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第1学習済モデルを用いて、第1の単位時間毎に前記指令情報を推定する第1推定部と、
- 前記制御対象装置が出力した前記状態を表す情報と、当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報とから、前記制御対象装置に供給される前記指令情報を導出する第2学習済モデルを用いて、前記指令情報を、前記第1の単位時間よりも短い第2の単位時間毎に推定する第2推定部と、
- 前記第2推定部が推定した前記指令情報を用いて前記制御対象装置を動作させる動作制御部と、
- を備える情報処理装置。
- [請求項2] 前記指令情報は、前記指令情報を出力する装置から発行される情報であり、
- 前記第2学習済モデルに入力される情報は、前記第1学習済モデルが導出した指令情報、及び、前記状態を表す情報を含む、
- 請求項1に記載の情報処理装置。
- [請求項3] 前記第2学習済モデルの入力は、前記制御対象装置が出力した前記状態を表す情報と前記第1学習済モデルの出力の情報の組を相関が小さくなるように変換したデータである、
- 請求項1又は2に記載の情報処理装置。
- [請求項4] 前記指令情報を出力する装置、及び前記制御対象装置はそれぞれ、ロボットアームを備え、
- 前記指令情報及び前記状態を表す情報はそれぞれ、前記ロボットアームの動作を示す情報を含む、
- 請求項1～3までの何れか1項に記載の情報処理装置。
- [請求項5] 前記指令情報を出力する装置は、ユーザ操作に応じて動作するマス

タ装置であり、

前記制御対象装置は、前記マスタ装置から出力される前記指令情報に従い動作するスレーブ装置であり、

前記第1学習済モデル及び前記第2学習済モデルは、前記マスタ装置及び前記スレーブ装置を備えるマスタ・スレーブシステムにおいて収集される前記指令情報及び前記状態を表す情報を用いて学習される

、

請求項1～4までのいずれか1項に記載の情報処理装置。

[請求項6]

指令情報に従い動作する制御対象装置が出力した、当該制御対象装置の状態を表す情報と、前記制御対象装置に供給された前記指令情報との組を用いて、前記状態を表す情報と前記指令情報との相関関係を学習させた第1学習済モデルを生成する第1生成部と、

前記制御対象装置が出力した前記状態を表す情報と当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報との関係を表す情報を含む入力情報と、前記制御対象装置に供給された前記指令情報との組を用いて、前記入力情報と前記指令情報との相関関係を学習させた第2学習済モデルを生成する第2生成部と、

を備える情報処理装置。

[請求項7]

前記制御対象装置が出力した前記状態を表す情報を第1の単位時間毎に前記第1学習済モデルに入力することにより前記指令情報を推定する第1推定部と、

前記制御対象装置が出力した前記状態を表す情報と当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報との関係を表す情報を含む入力データを、前記第1の単位時間よりも短い第2の単位時間毎に前記第2学習済モデルに入力することにより前記指令情報を推定する第2推定部と、

前記第2推定部が推定した前記指令情報を用いて前記制御対象装置

を動作させる動作制御部と、

を更に備える、請求項 6 に記載の情報処理装置。

[請求項8] 前記指令情報を出力する装置は、ユーザ操作に応じて動作するマスタ装置であり、

前記制御対象装置は、前記マスタ装置から出力される前記指令情報に従い動作するスレーブ装置であり、

前記第 1 生成部は、前記マスタ装置及び前記スレーブ装置を備えるマスタ・スレーブシステムにおいて収集される前記指令情報及び前記状態を表す情報を用いて、前記第 1 学習済モデルを生成し、

前記第 2 生成部は、前記マスタ・スレーブシステムにおいて収集される前記指令情報及び前記状態を表す情報を用いて、前記第 2 学習済モデルを生成する、

請求項 6 又は 7 に記載の情報処理装置。

[請求項9] 前記第 1 学習済モデルは、前記状態を表す情報と、前記制御対象装置の動作を規定する、動作速度を含む情報とを含む入力データから、前記制御対象装置が当該状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報を導出する学習済モデルである、

請求項 1 ～ 8 までのいずれか 1 項に記載の情報処理装置。

[請求項10] 前記動作速度を含む情報は、前記制御対象装置の動作周期を表す周波数又は動作実行時間を示す情報を含む、

請求項 9 に記載の情報処理装置。

[請求項11] 前記第 1 学習済モデルは、前記入力データから、前記制御対象装置が前記状態を表す情報を出力してから予め定められた単位時間が経過した後に前記制御対象装置に供給される前記指令情報を導出する学習済モデルである、

請求項 9 又は 10 に記載の情報処理装置。

[請求項12] 前記第 1 学習済モデルは、前記状態を表す情報と、前記制御対象装置に指示する動作を規定する、動作速度を含む情報とを含む入力デー

タから、前記制御対象装置が当該状態を表す情報を出力した後に前記制御対象装置に供給される前記指令情報を出力データとして導出する学習済モデルであり、

前記第1生成部は、前記制御対象装置の動作速度を変えながら前記制御対象装置を動作させることにより収集される前記指令情報、前記状態を表す情報、及び前記動作速度を含む情報を用いて、前記第1学習済モデルを生成する、

請求項6又は7に記載の情報処理装置。

[請求項13]

前記第1学習済モデルは、前記指令情報から、前記制御対象装置に当該指令情報が供給された後に前記制御対象装置が出力する第2の状態を表す情報を導出する第3学習済モデルにより導出される第2の状態を表す情報を入力情報として学習させた学習済モデルである、

請求項1～12までのいずれか1項に記載の情報処理装置。

[請求項14]

指令情報に従い動作する制御対象装置が出力する、当該制御対象装置の状態を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第1学習済モデルを用いて、前記指令情報を、第1の単位時間毎に推定する第1ステップと、

前記制御対象装置が出力した前記状態を表す情報と、当該状態を表す情報を前記第1学習済モデルに入力することにより導出される前記指令情報との関係を表す情報から、前記制御対象装置に供給される前記指令情報を導出する第2学習済モデルを用いて、前記指令情報を、前記第1の単位時間よりも短い第2の単位時間毎に推定する第2ステップと、

前記第2ステップにおいて推定した前記指令情報を用いて前記制御対象装置を動作させる第3ステップと、

を情報処理装置が実行する方法。

[請求項15]

前記第2推定部は、 n 個（ n は2以上の自然数）の前記第2学習済モデルを用いて前記指令情報を推定し、

n 個の前記第 2 学習済モデルは階層的に用いられ、各層の前記第 2 学習済モデルから導出される指令情報に応じた関係を表す情報が、当該第 2 学習済モデルの下層の第 2 学習済モデルの入力情報となる、

請求項 1 ～ 5 までのいずれか 1 項に記載の情報処理装置。

[請求項16]

前記制御対象装置は、対象物に対する作業を行う装置であり、

前記第 1 の単位時間及び前記第 2 の単位時間の一方又は両方を、前記対象物の硬軟の程度に応じて変更する変更制御部、

を更に備える請求項 1 ～ 5 までのいずれか 1 項に記載の情報処理装置。

[請求項17]

前記第 1 学習済モデルは、前記状態を表す情報から、所定の時間が経過した後の前記制御対象装置の第 2 の状態を表す情報、及び、前記指令情報を導出する学習済モデルであって、当該第 1 学習済モデルが導出した前記第 2 の状態を表す情報を入力情報として学習させた学習済モデルである、

請求項 1 ～ 12 までのいずれか 1 項に記載の情報処理装置。

[請求項18]

前記第 1 学習済モデルは、前記指令情報及び前記状態を表す情報から、所定の時間が経過した後の前記制御対象装置の第 2 の状態を表す情報、及び、前記指令情報を導出する学習済モデルであって、当該第 1 学習済モデルが導出した前記指令情報及び前記第 2 の状態を表す情報を、入力情報として学習させた学習済モデルである、

請求項 1 ～ 12 までのいずれか 1 項に記載の情報処理装置。

[請求項19]

前記 n 個の前記第 2 学習済モデルは、それぞれ動作周波数が異なる請求項 15 に記載の情報処理装置。

[請求項20]

前記入力情報とは、定量的に表される値である

請求項 13 及び 15 ～ 18 までのいずれか 1 項に記載の情報処理装置。

[請求項21]

前記制御対象装置は、把持対象物を把持又は接触対象物に接触し、

前記第 1 学習済モデルと、前記第 2 学習済モデルとには、前記把持

対象物を把持又は前記接触対象物に接触する動作が記憶され、

前記把持対象物を把持又は前記接触対象物に接触する動作は、前記把持対象物又は前記接触対象物の硬軟の程度に応じて前記第 1 学習済モデル又は前記第 2 学習済モデルのうちいずれか学習済モデルに学習される

請求項 1 ～ 5 までのいずれか 1 項に記載の情報処理装置。

[請求項22] 前記第 1 学習済モデルは、前記把持対象物が柔らかい物体である場合の動作を学習し、

前記第 2 学習済モデルは、前記把持対象物が硬い物体である場合の動作を学習し、

前記第 1 学習済モデルと前記第 2 学習済モデルとは、動作周波数が異なる

請求項 2 1 に記載の情報処理装置。

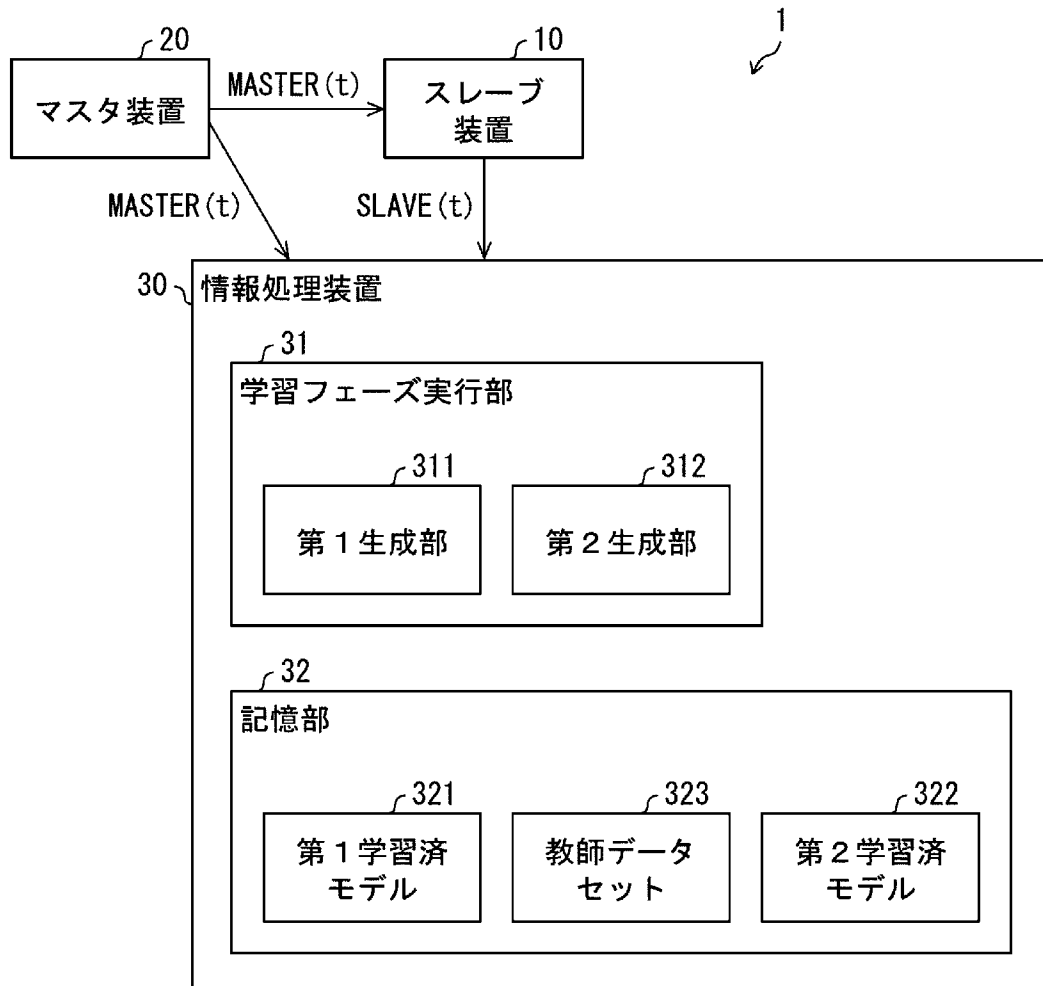
[請求項23] 請求項 1 ～ 1 3、1 5 ～ 2 2 のいずれか 1 項に記載の情報処理装置としてコンピュータを機能させるためのプログラムであって、前記各部としてコンピュータを機能させるためのプログラム。

[請求項24] 前記第 1 推定部は、複数の時点における前記指令情報を推定し、

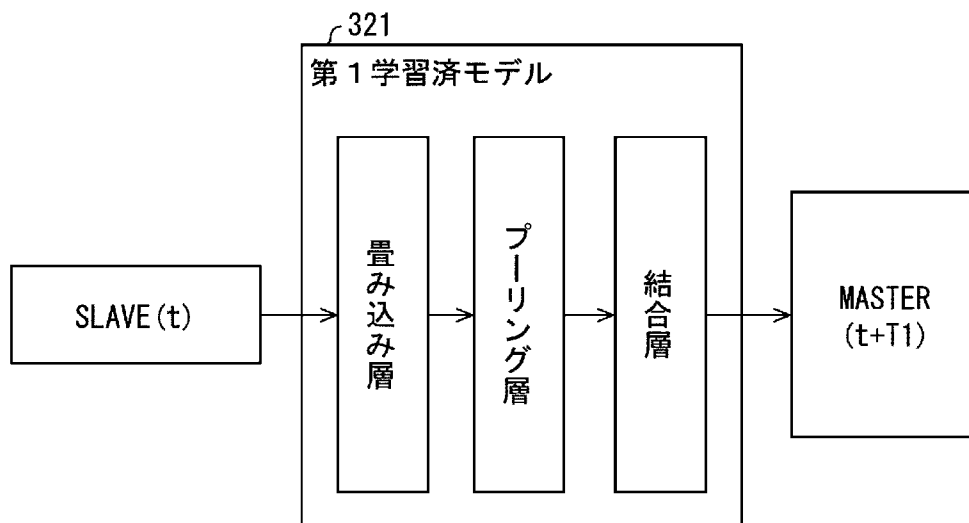
前記第 2 推定部は、推定された複数の時点における前記指令情報ごとに、前記指令情報を推定する

請求項 1 ～ 5 及び請求項 7 までのいずれか 1 項に記載の情報処理装置。

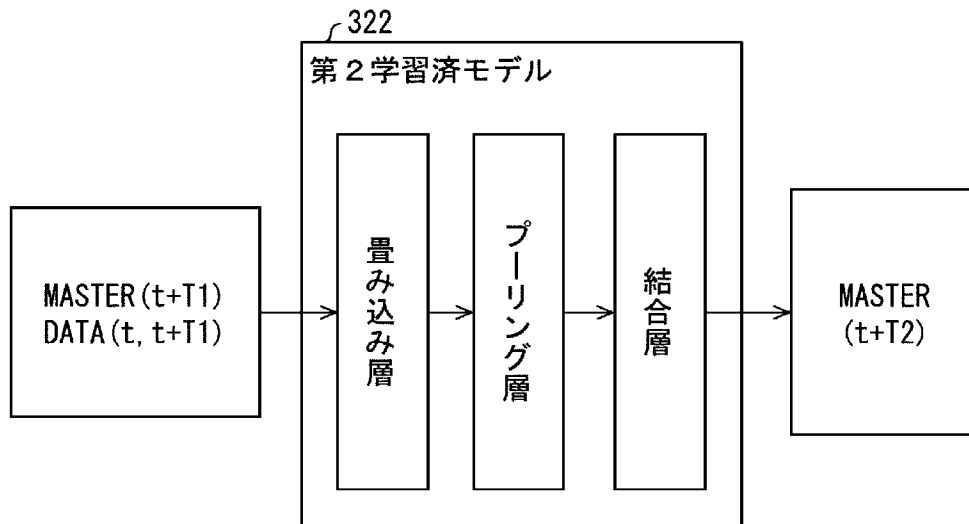
[図1]



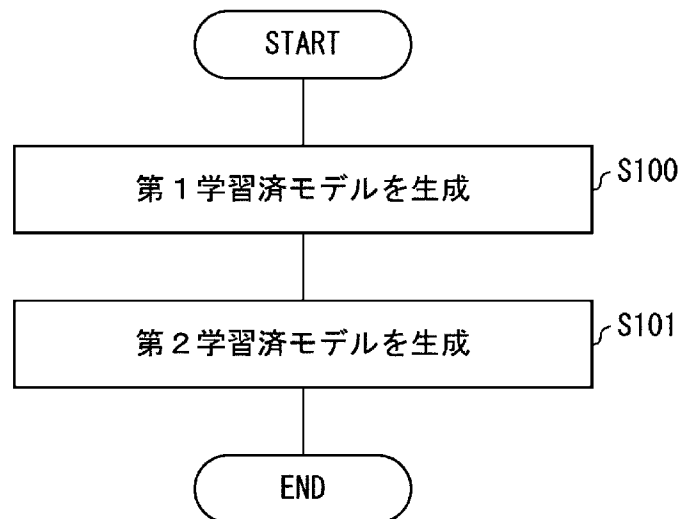
[図2]



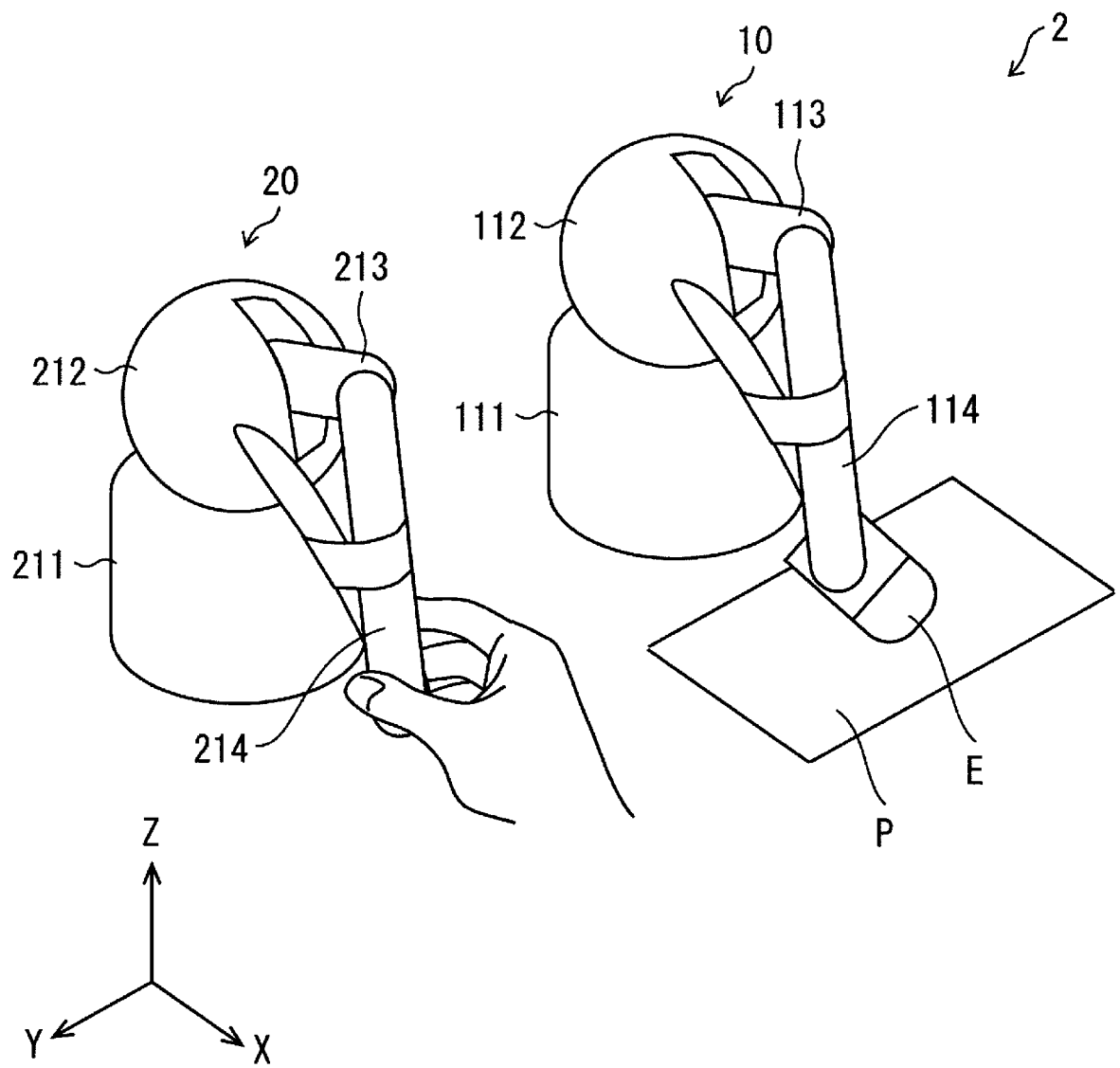
[図3]



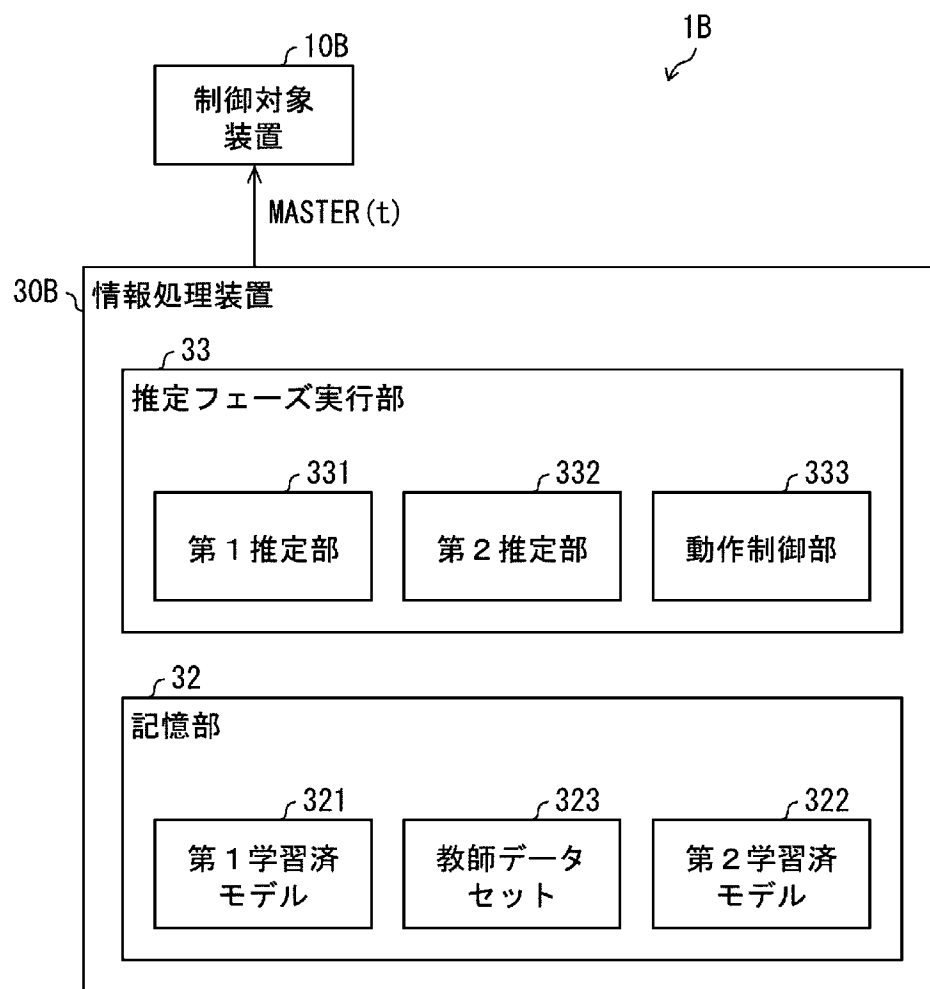
[図4]



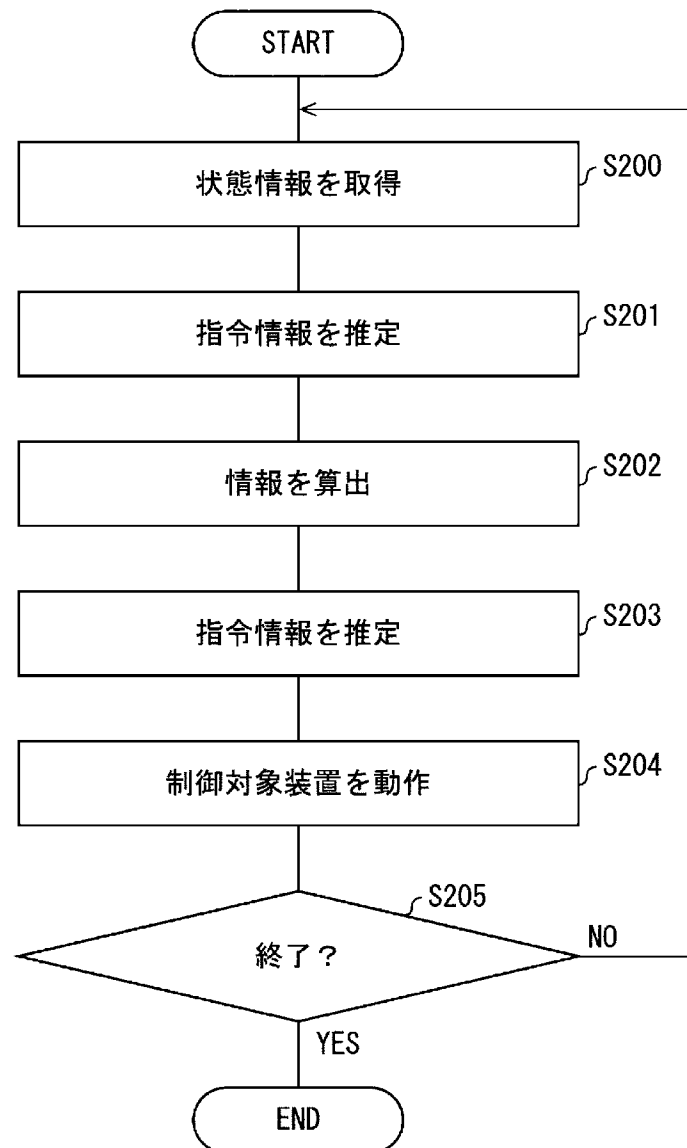
[図5]



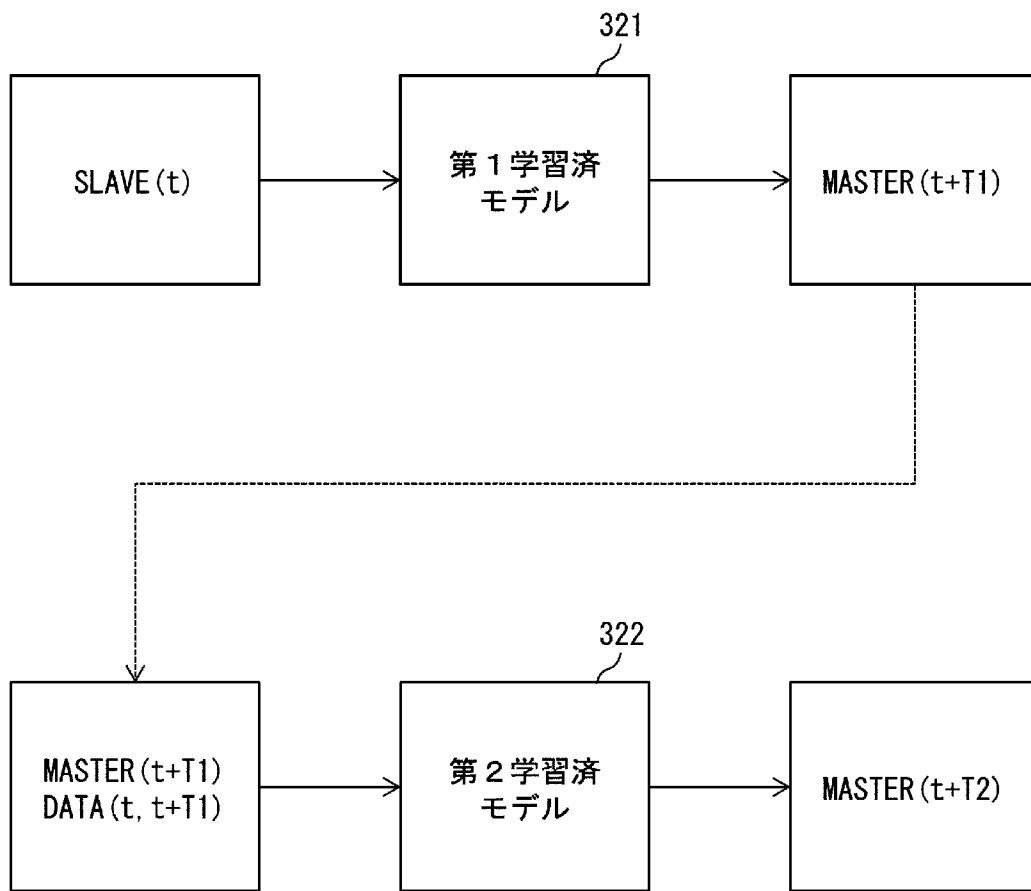
[図6]



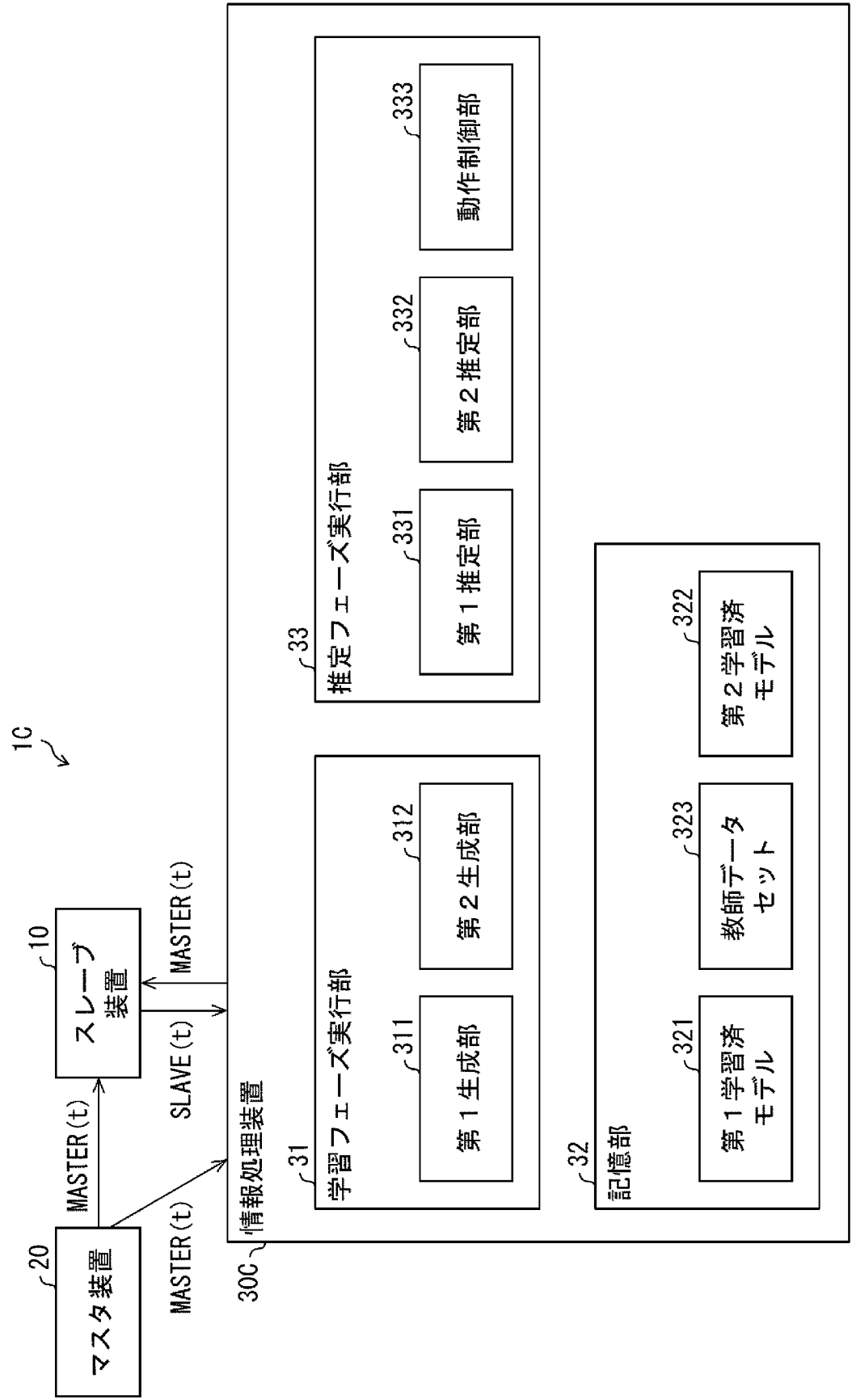
[図7]



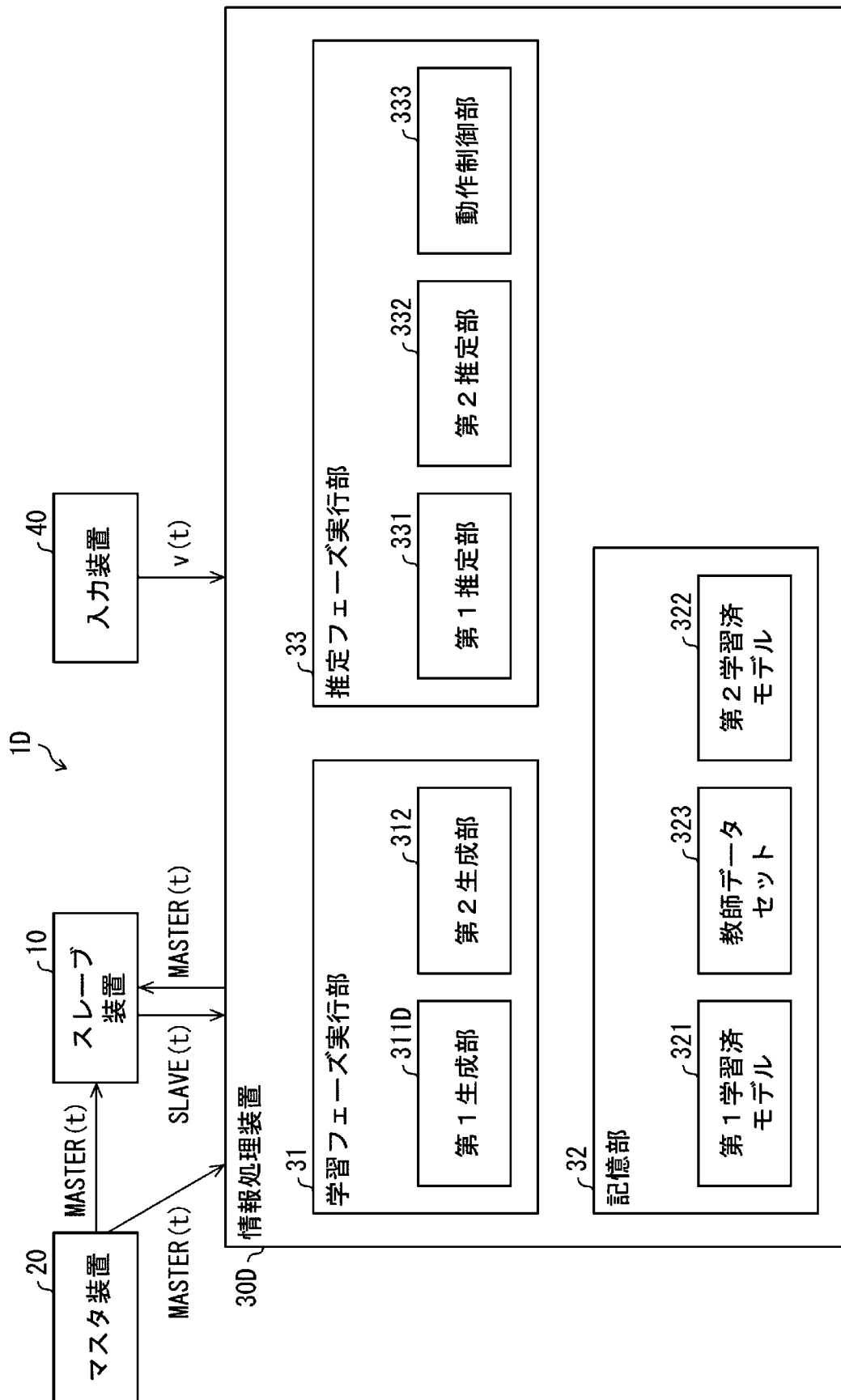
[図8]



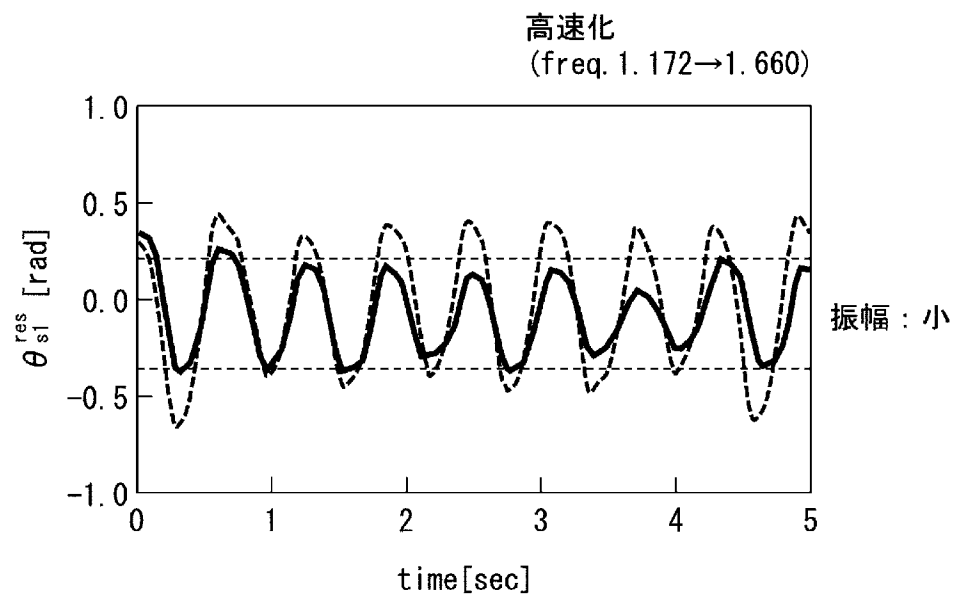
[図9]



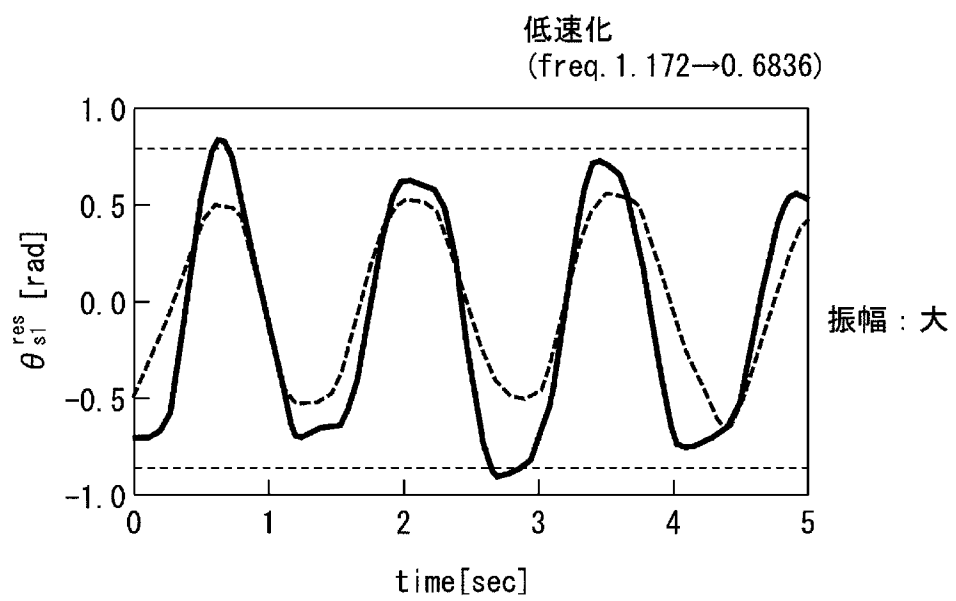
[図10]



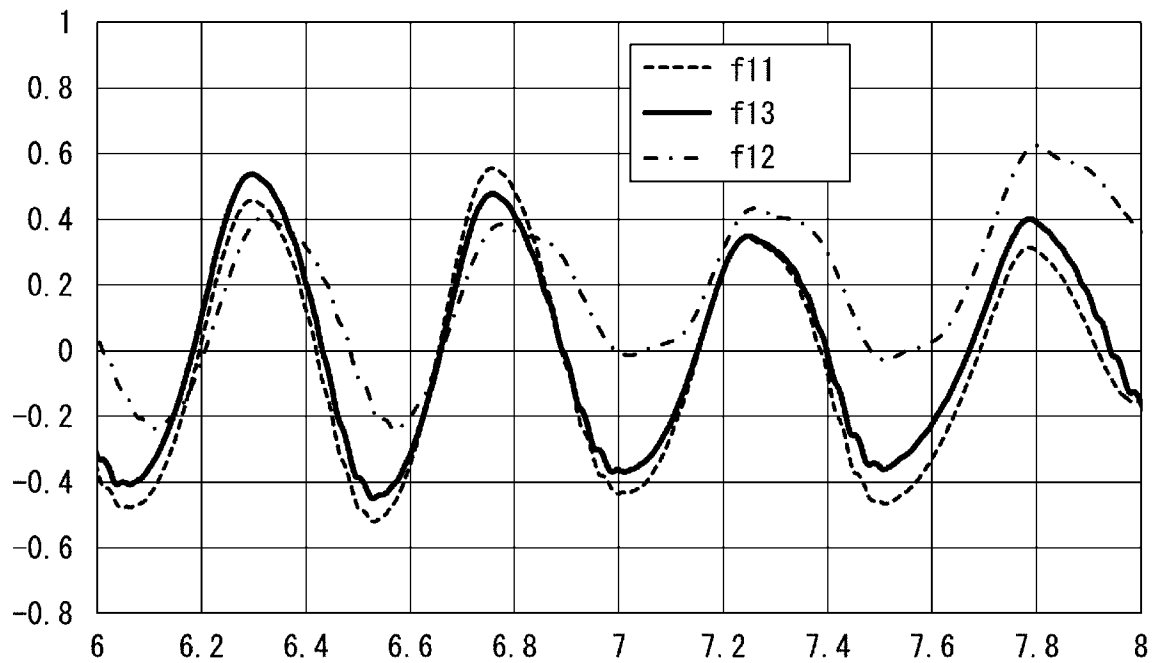
[図11]



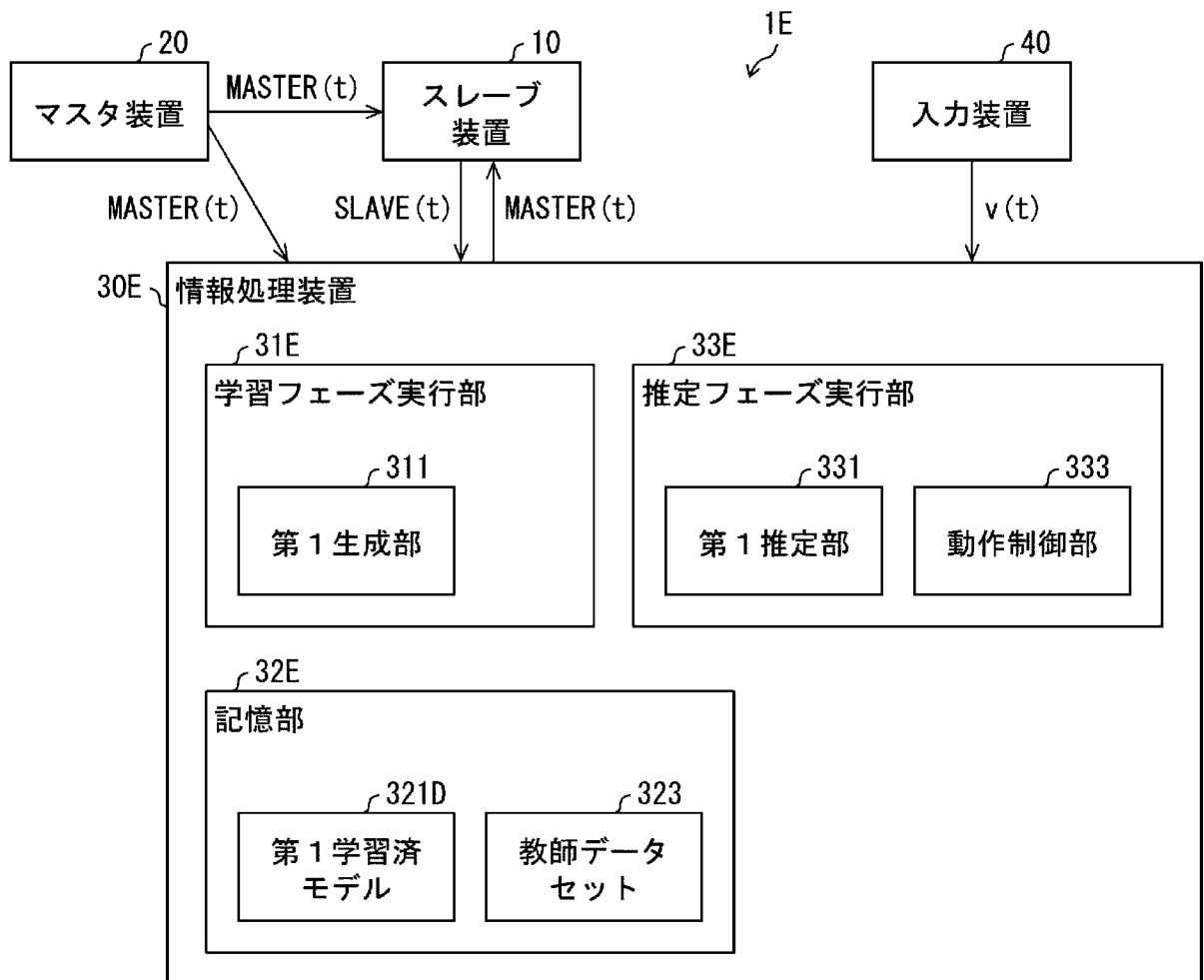
[図12]



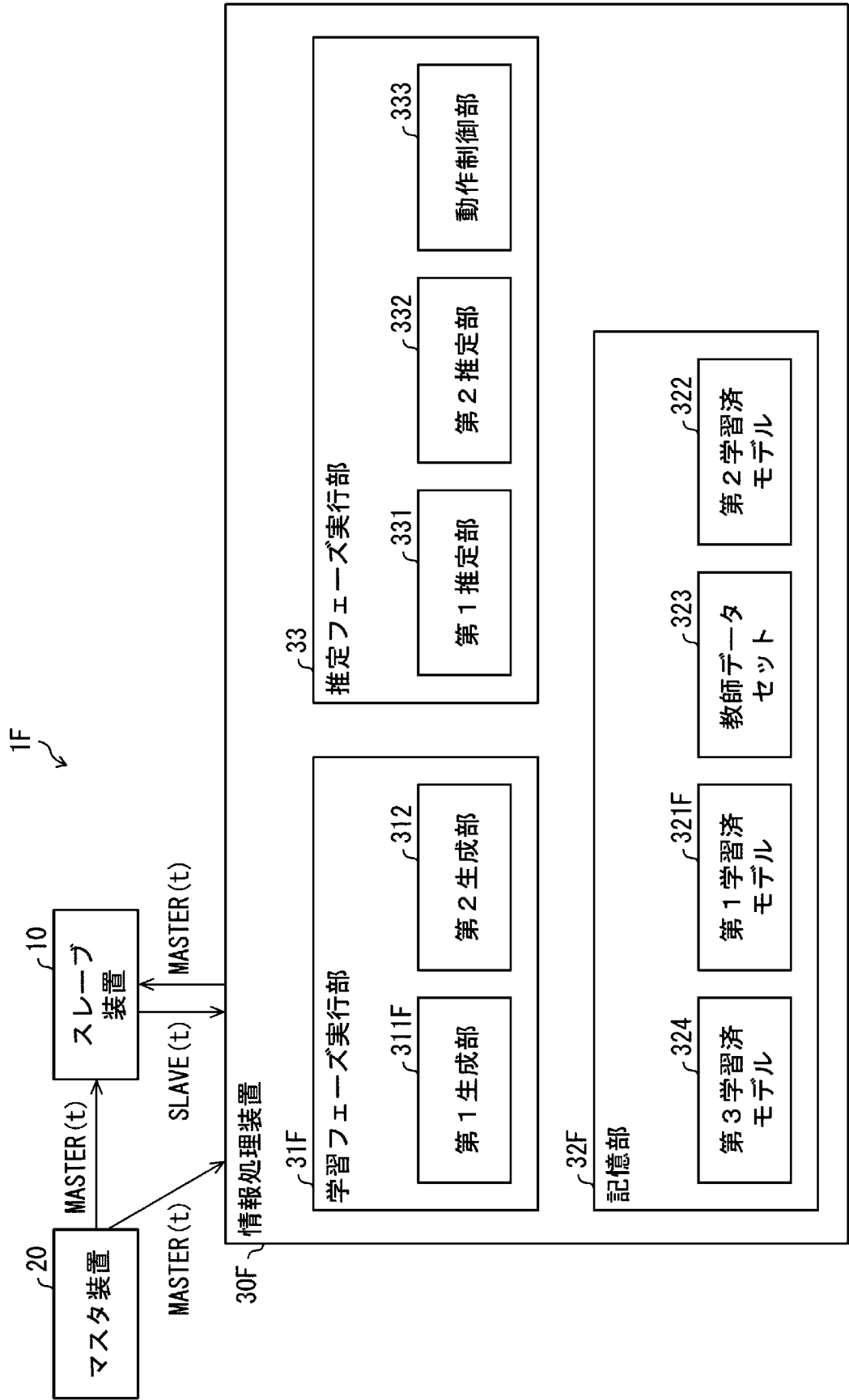
[図13]



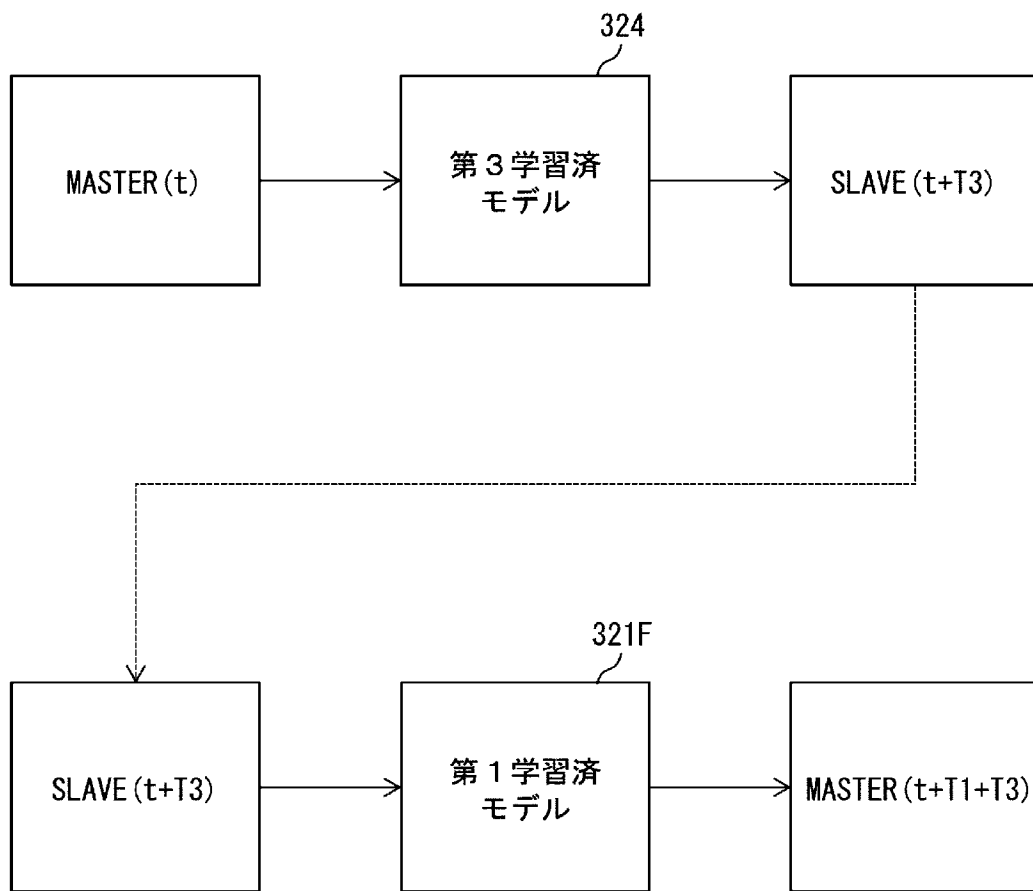
[図14]



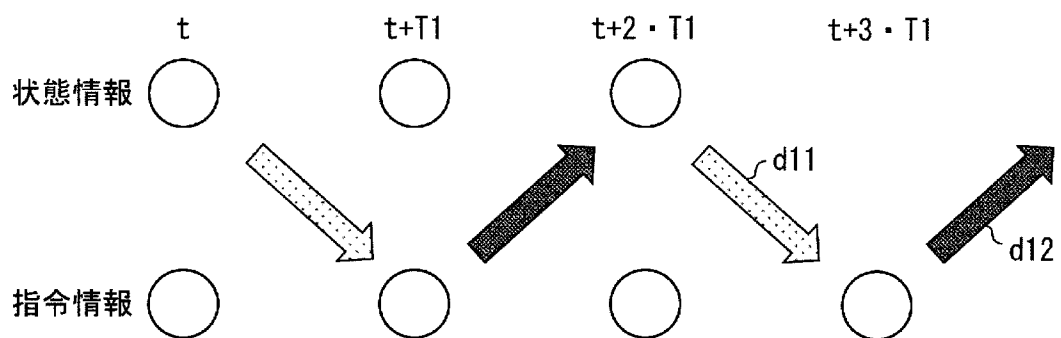
[図15]



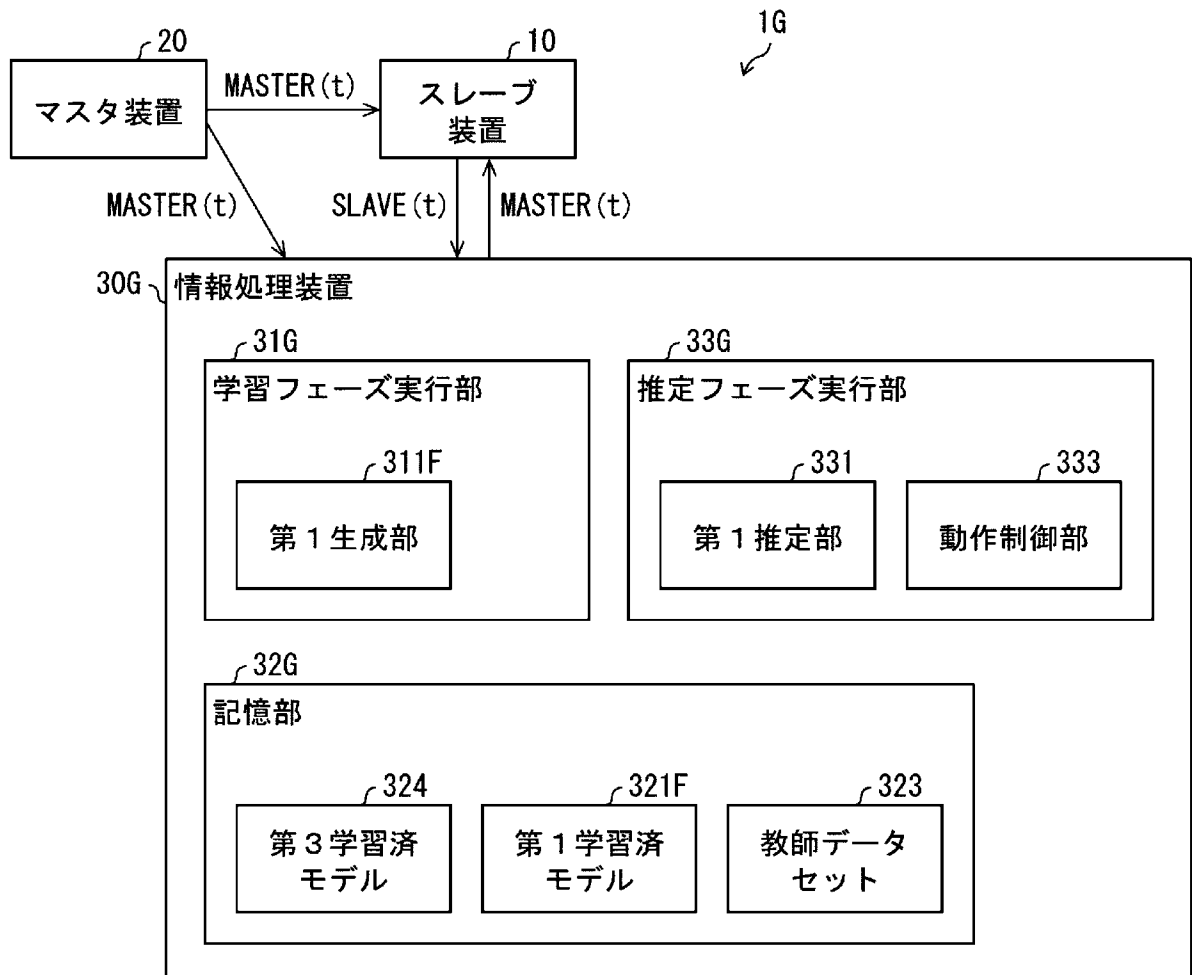
[図16]



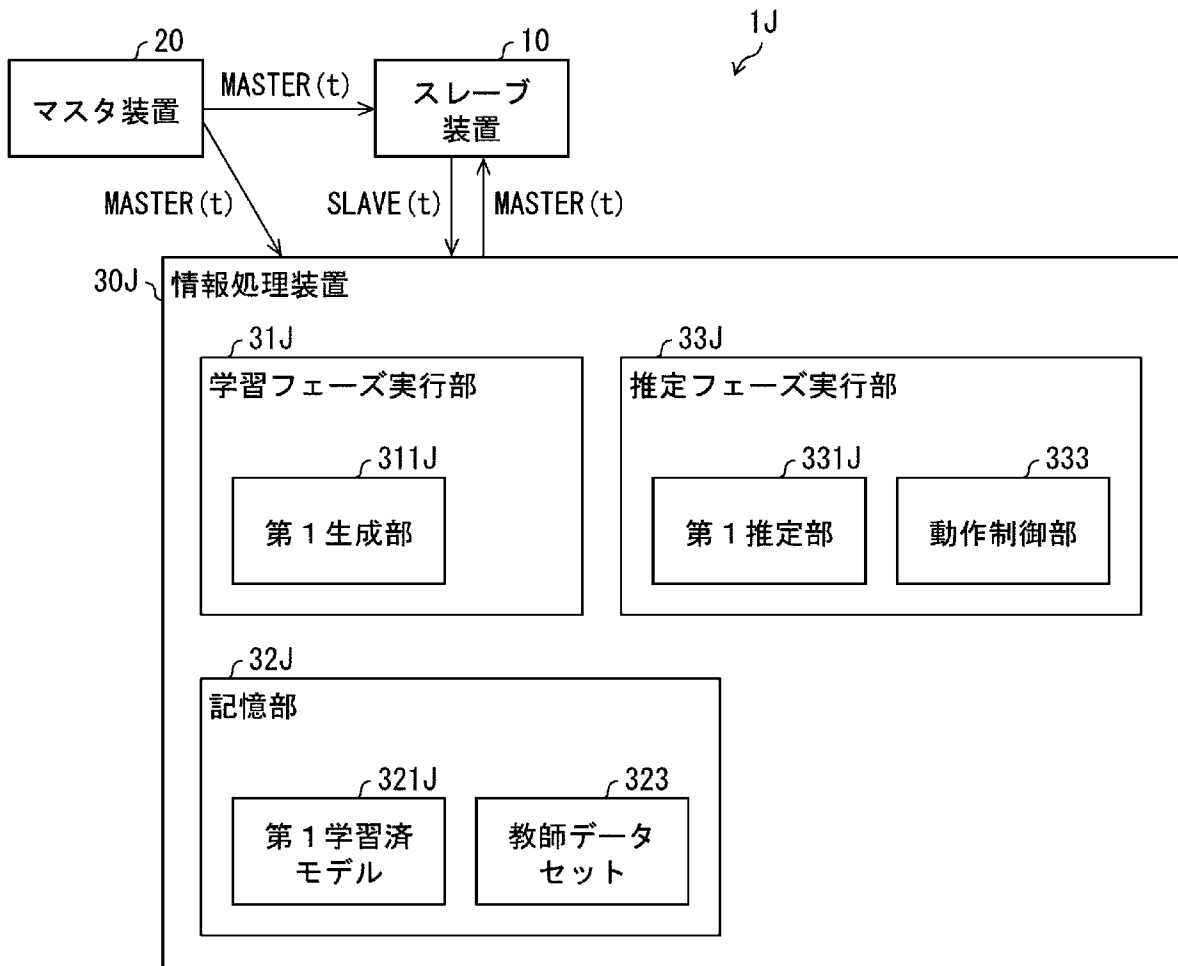
[図17]



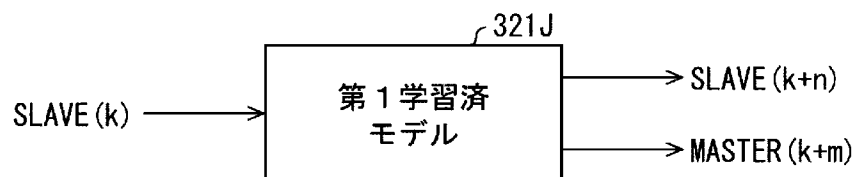
[図18]



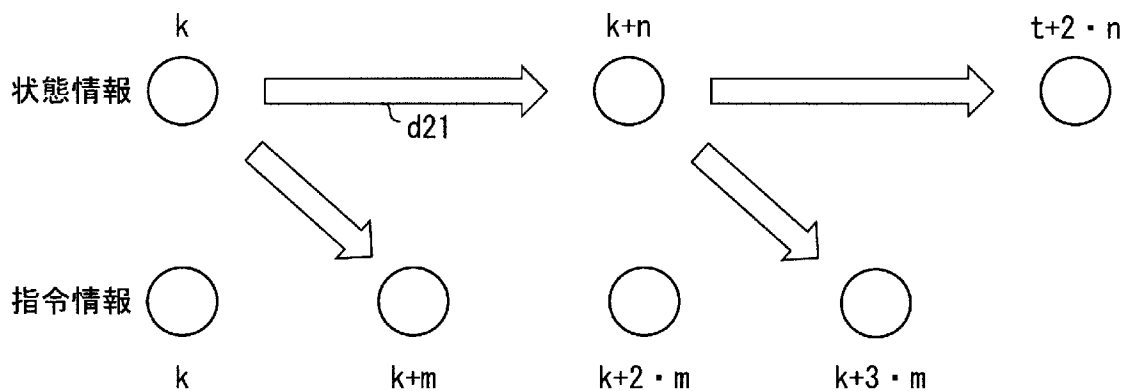
[図20]



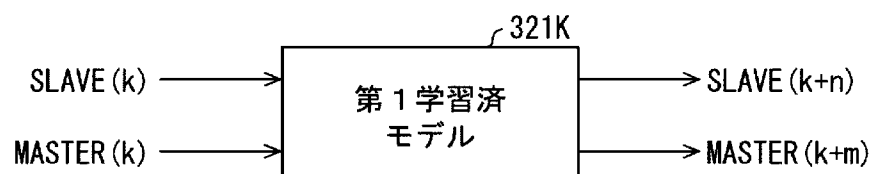
[図21]



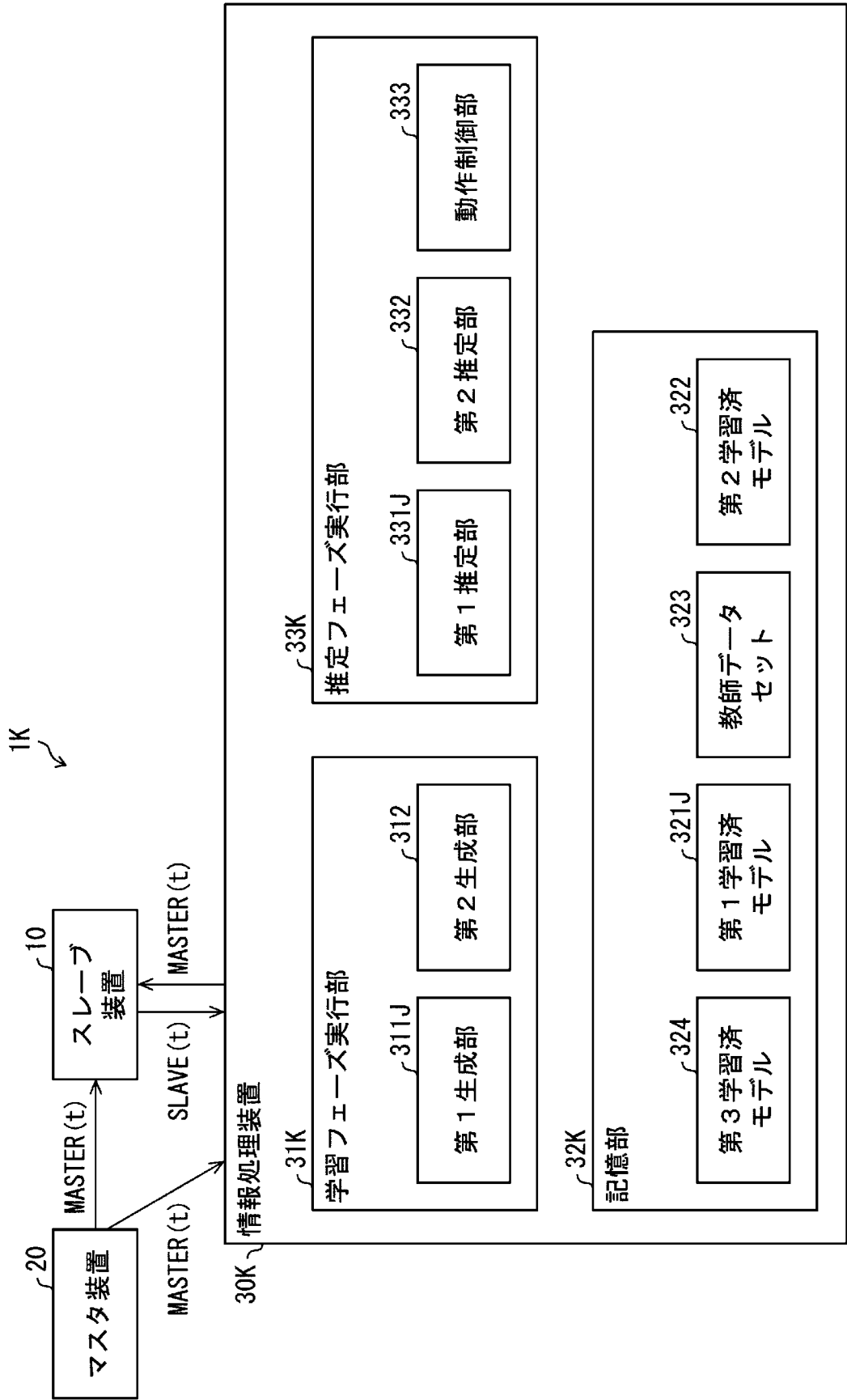
[図22]



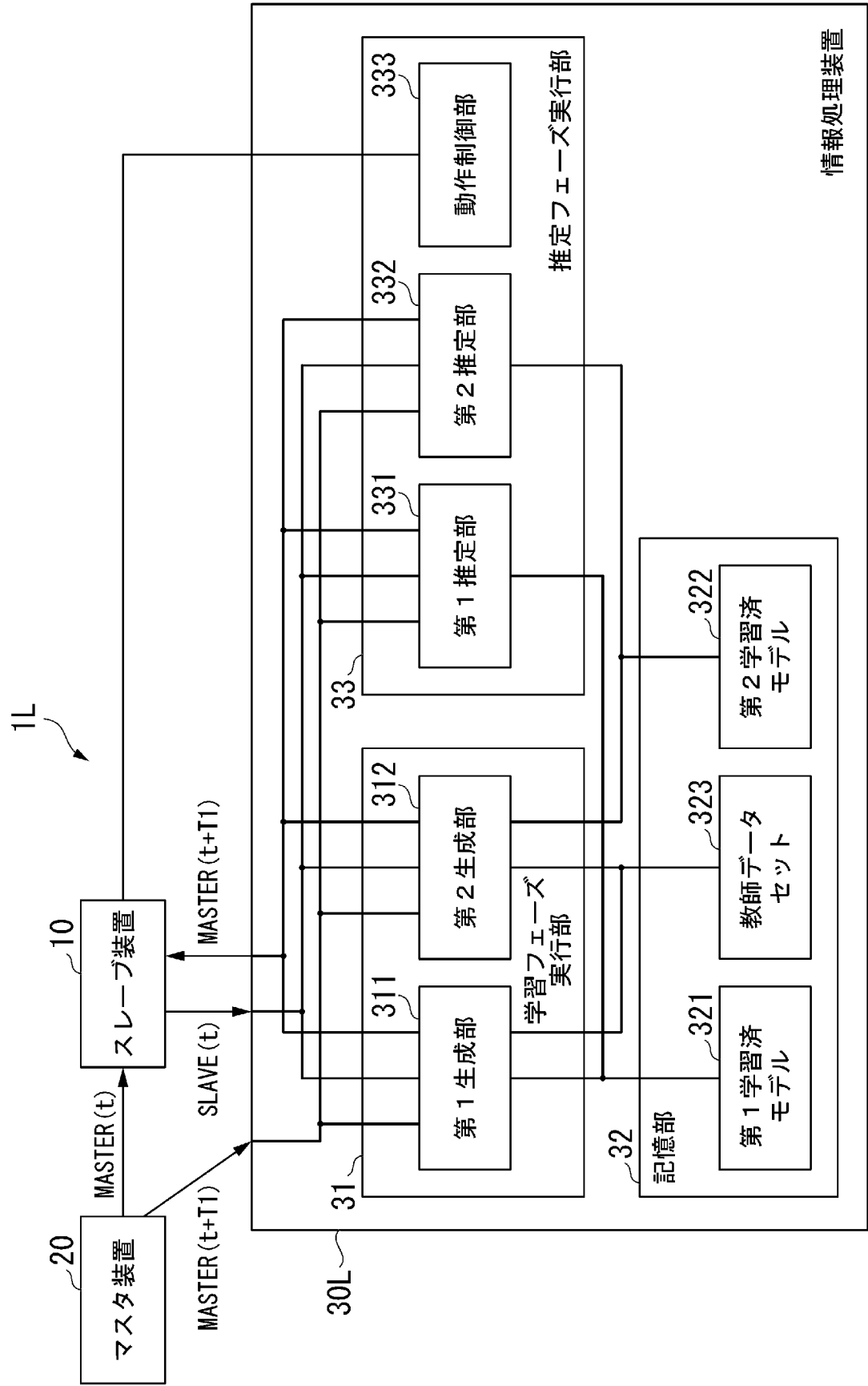
[図23]



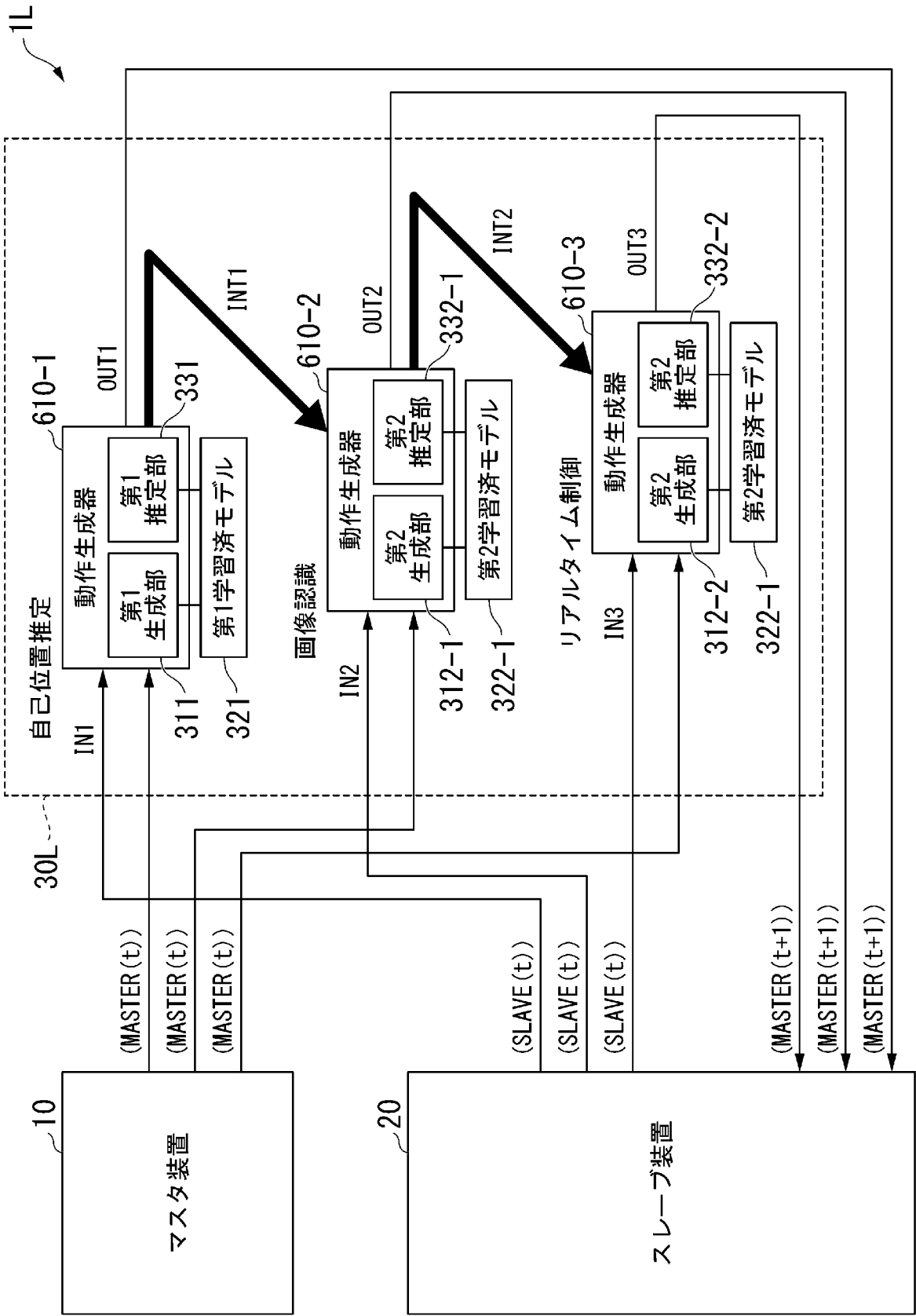
[図24]



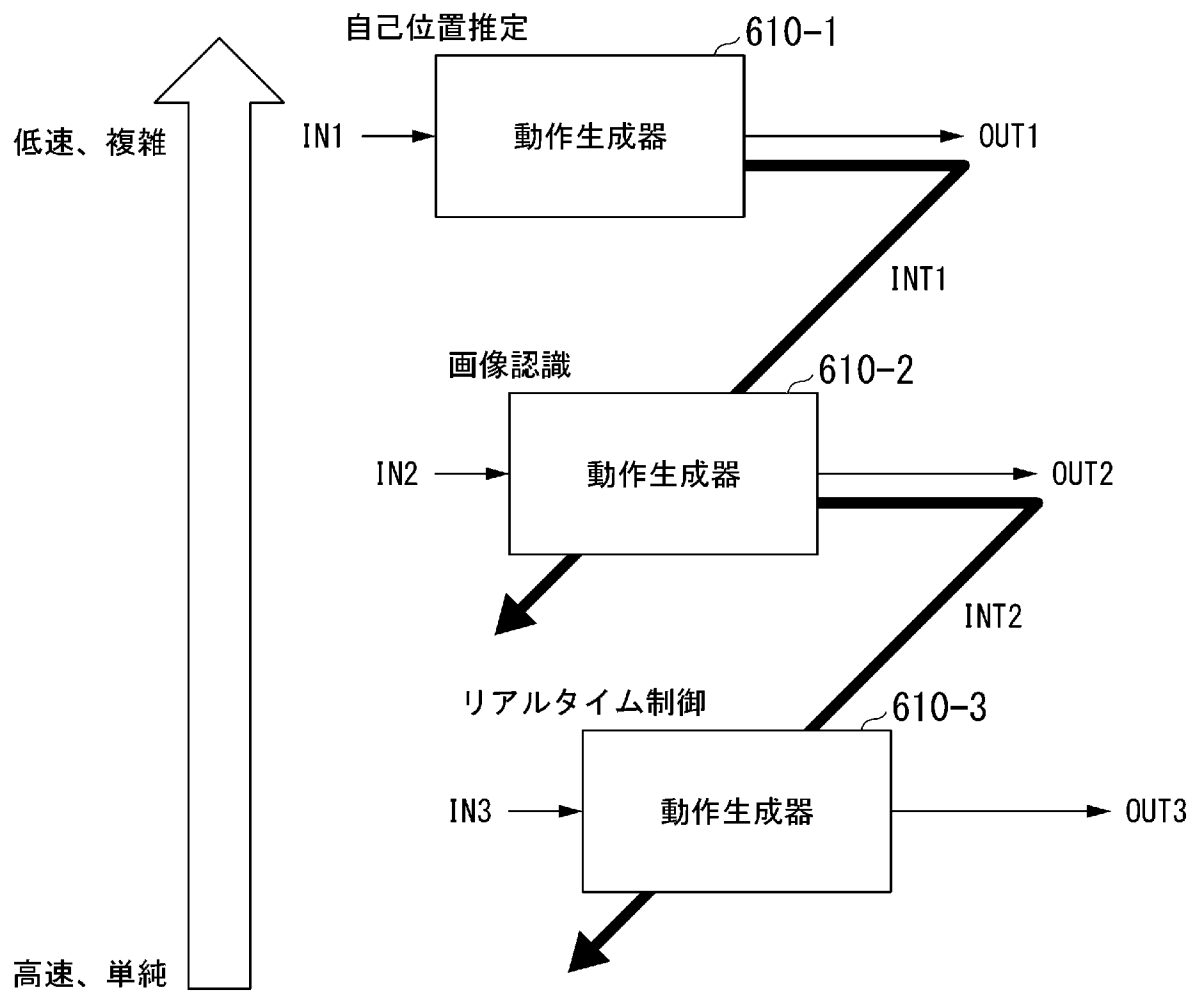
[図25]



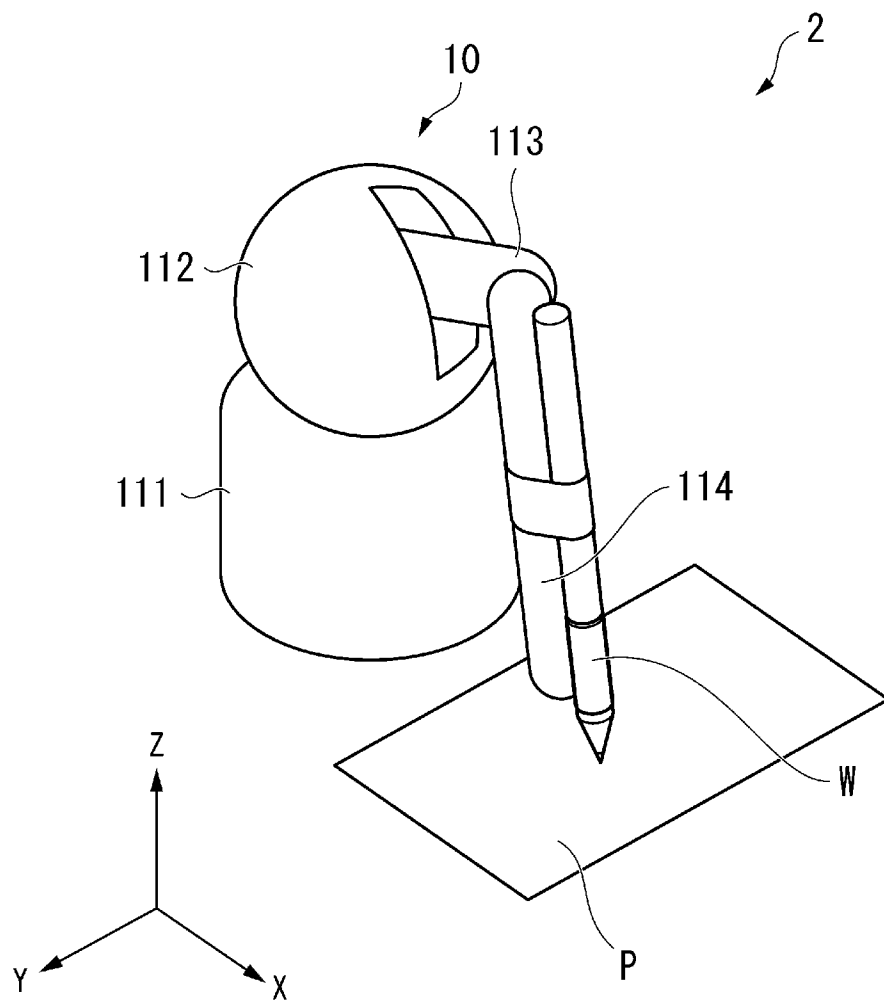
[図26]



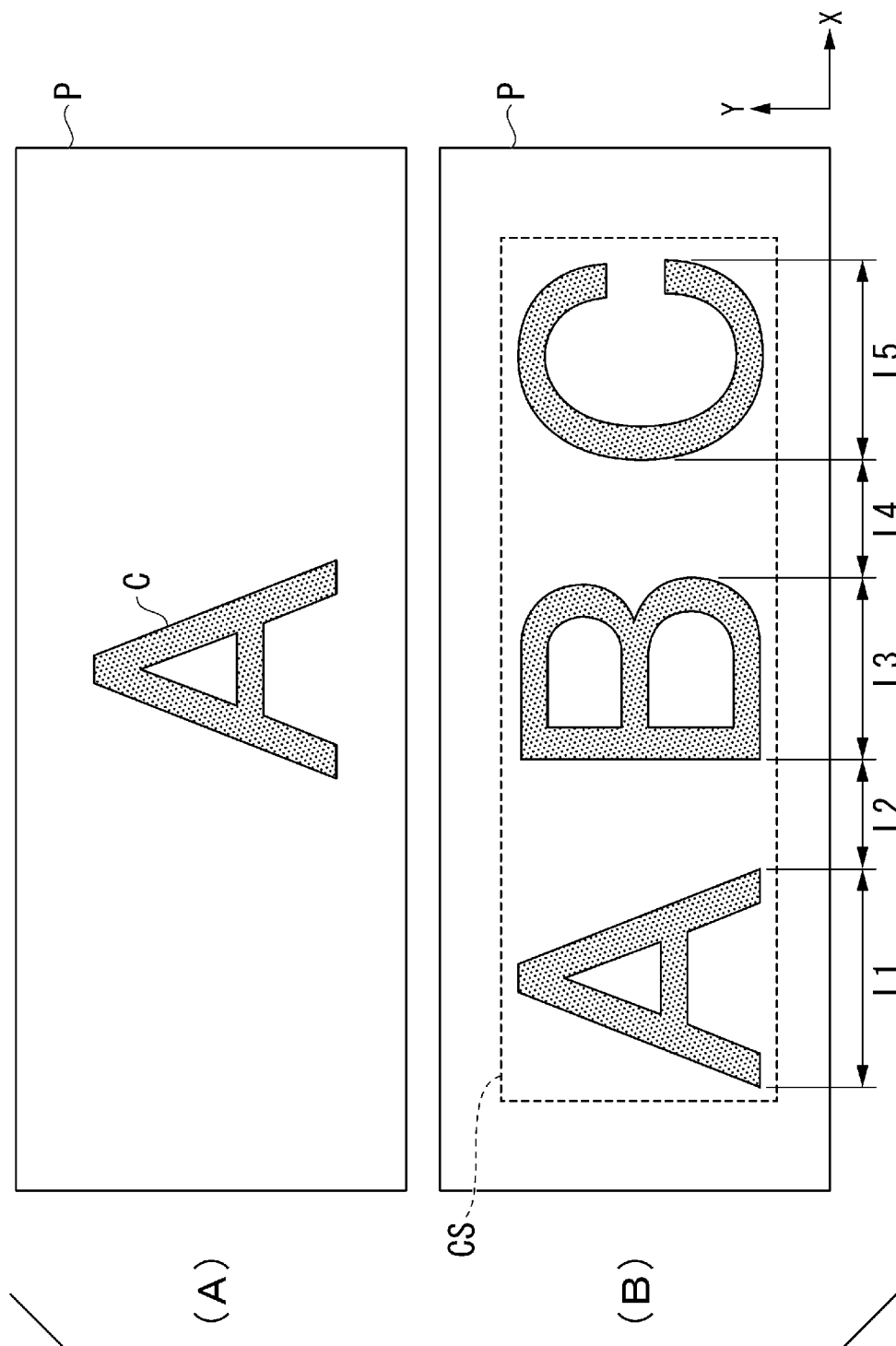
[図27]



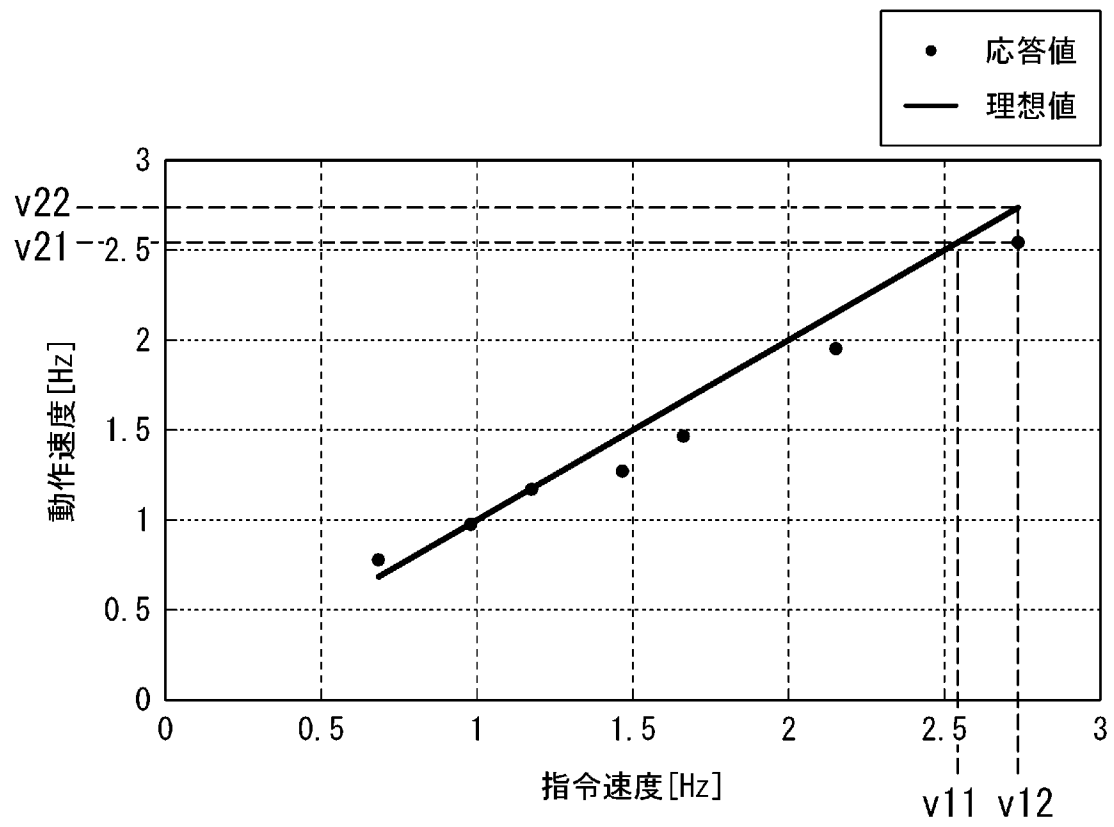
[図28]



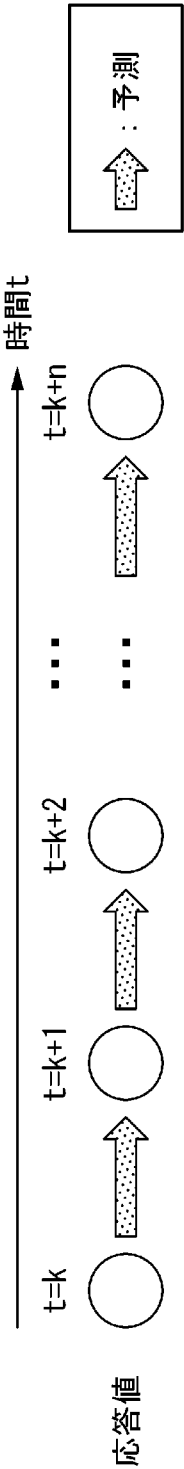
[図29]



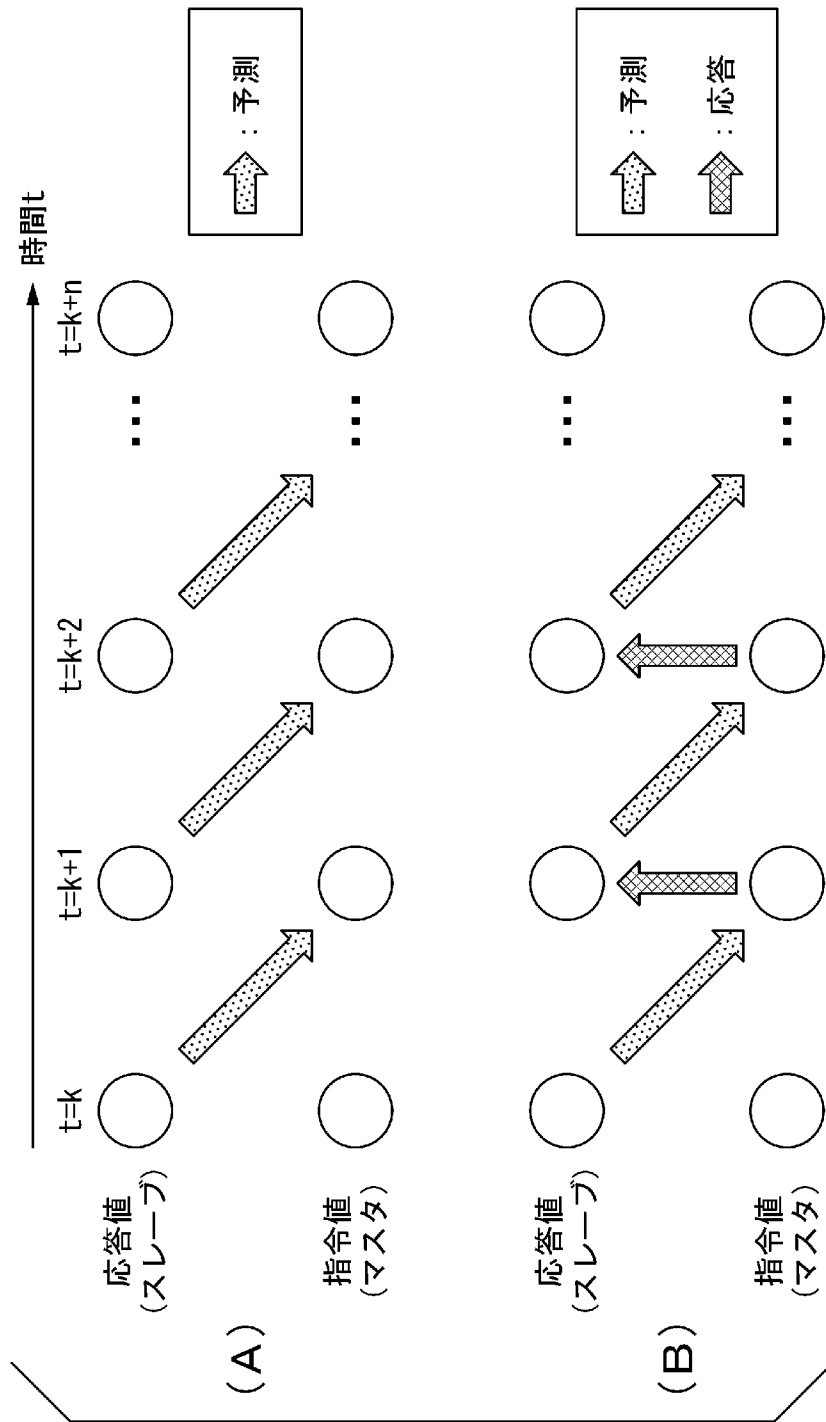
[図30]



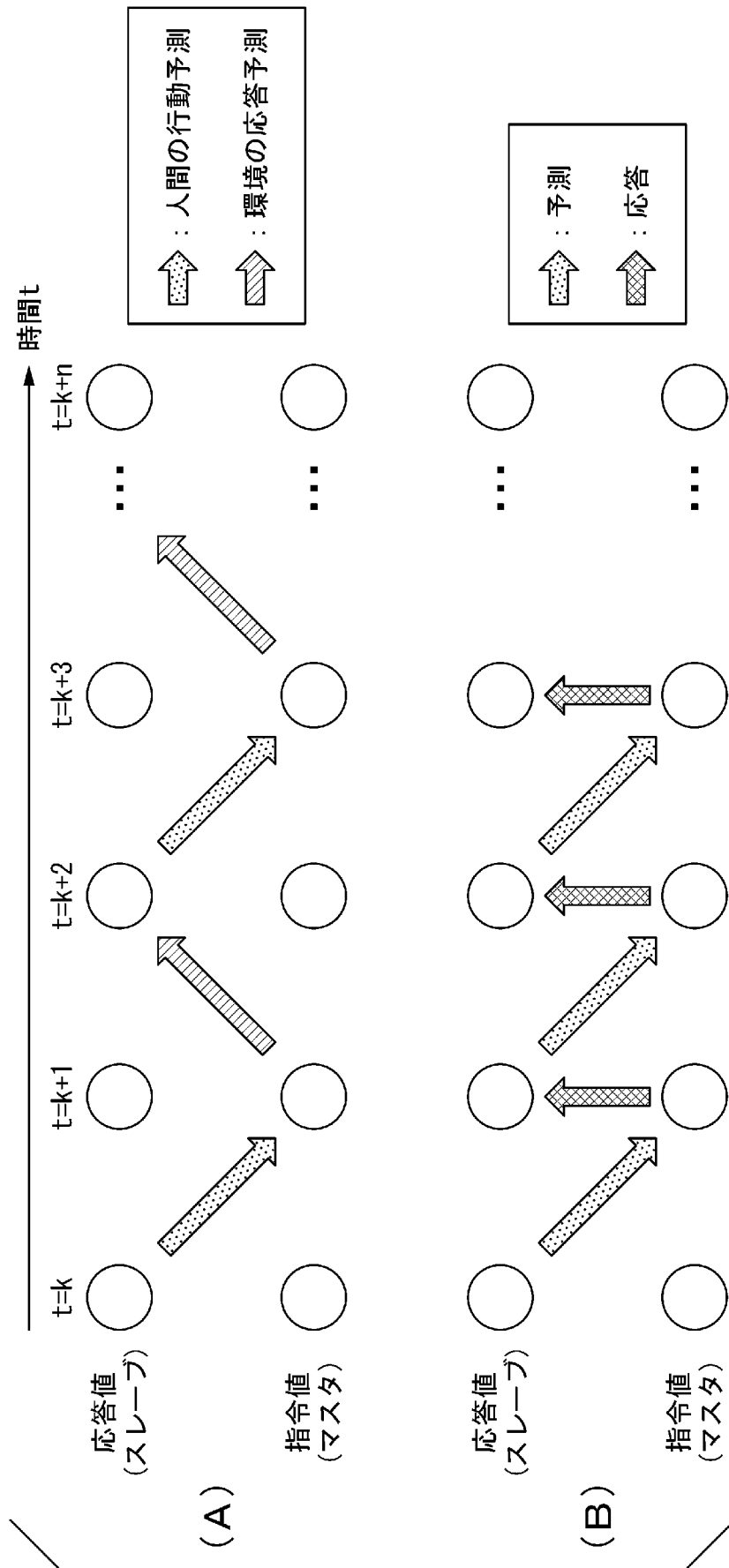
[図31]



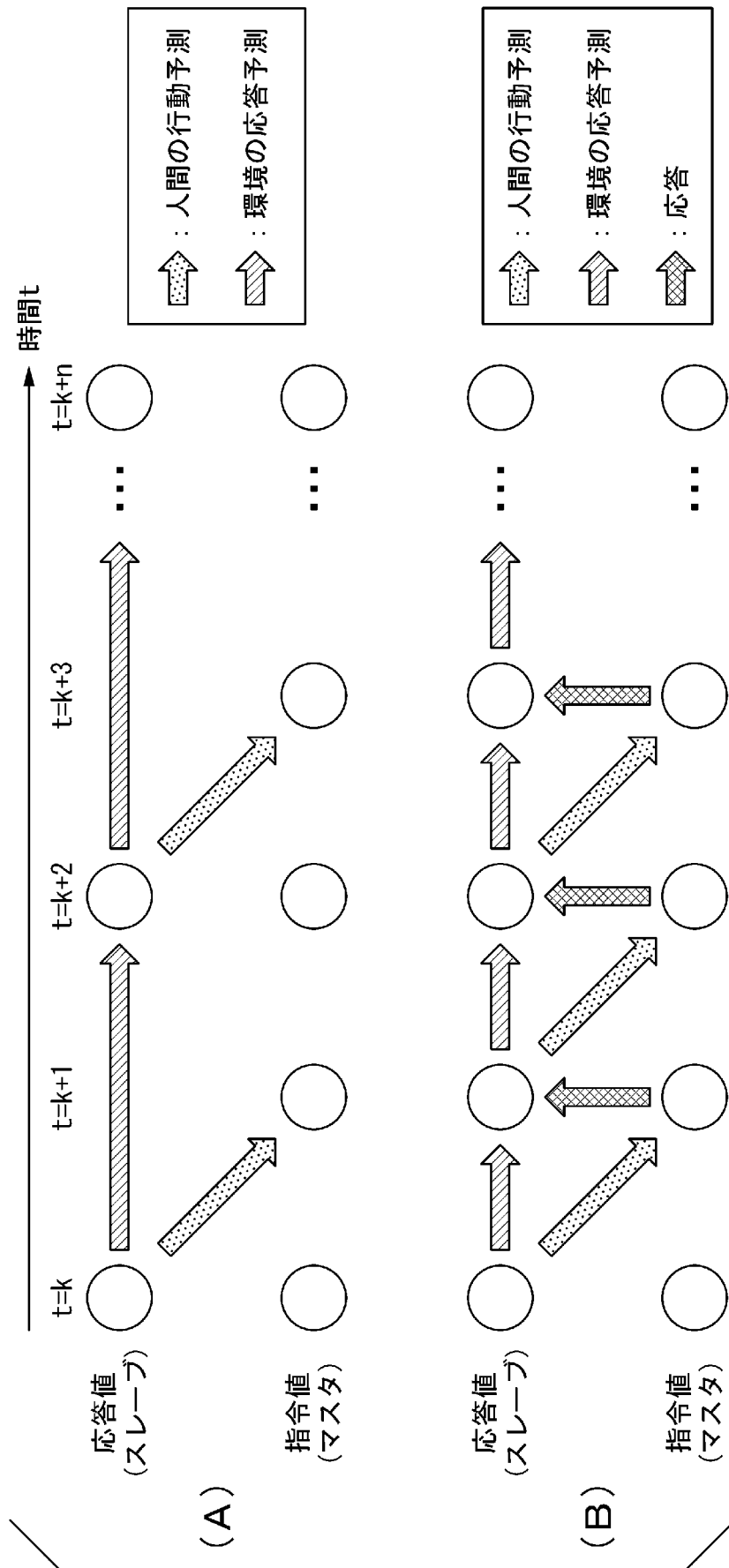
[図32]



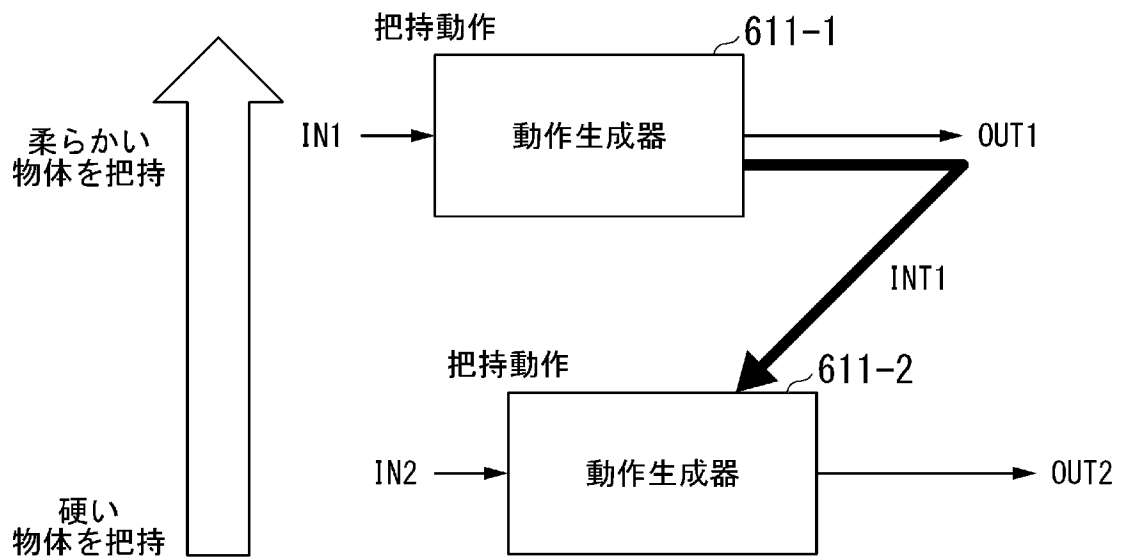
[図33]



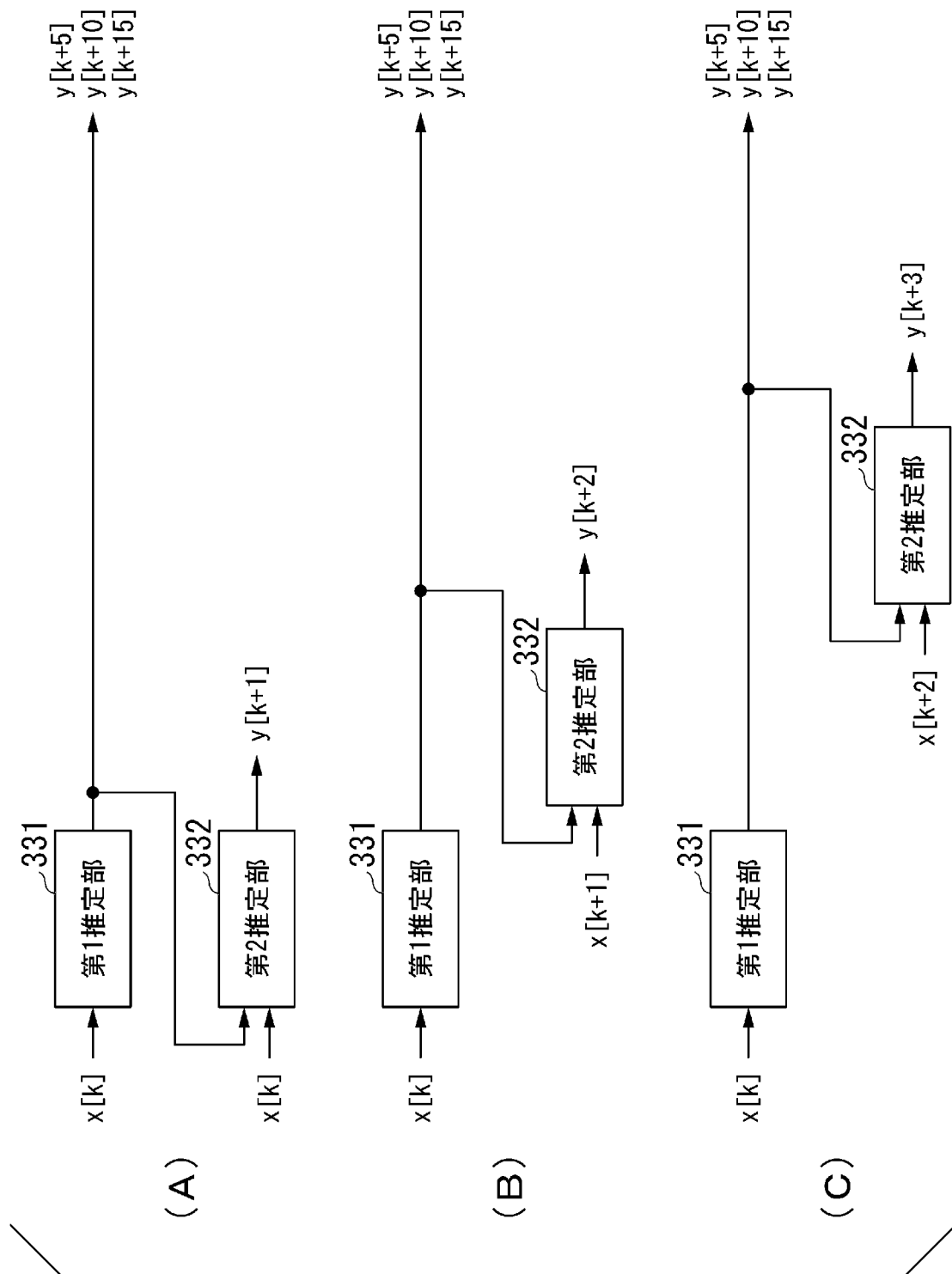
[図34]



[図35]



[図36]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/025932

A. CLASSIFICATION OF SUBJECT MATTER

B25J 3/00 (2006.01) i; B25J 13/00 (2006.01) i

FI: B25J13/00 Z; B25J3/00 Z

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)

B25J3/00; B25J13/00

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	1971-2021
Registered utility model specifications of Japan	1996-2021
Published registered utility model applications of Japan	1994-2021

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 2020-82332 A (OMRON CORP.) 04 June 2020 (2020-06-04) claims	1-24
P, A	JP 2021-94677 A (HONDA MOTOR CO., LTD.) 24 June 2021 (2021-06-24) claims	1-24
P, A	JP 2021-16908 A (KYOCERA DOCUMENT SOLUTIONS INC.) 15 February 2021 (2021-02-15) claims	1-24
P, A	JP 2021-64295 A (AISIN SEIKI CO., LTD.) 22 April 2021 (2021-04-22) claims	1-24
P, A	JP 2021-49841 A (JAPAN AEROSPACE EXPLORATION AGENCY) 01 April 2021 (2021-04-01) claims	1-24



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

30 August 2021 (30.08.2021)

Date of mailing of the international search report

07 September 2021 (07.09.2021)

Name and mailing address of the ISA/

Japan Patent Office
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/JP2021/025932

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
JP 2020-82332 A	04 Jun. 2020	US 2021/0252714 A1 claims CN 112584990 A	
JP 2021-94677 A	24 Jun. 2021	US 2021/0187737 A1 claims CN 113001539 A	
JP 2021-16908 A	15 Feb. 2012	US 2021/0016439 A1 claims (Family: none)	
JP 2021-64295 A	22 Apr. 2021	WO 2021/059787 A1	
JP 2021-49841 A	01 Apr. 2021	claims	

A. 発明の属する分野の分類（国際特許分類（IPC）） B25J 3/00(2006.01)i; B25J 13/00(2006.01)i FI: B25J13/00 Z; B25J3/00 Z		
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）） B25J3/00; B25J13/00		
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 日本国実用新案公報 1922-1996年 日本国公開実用新案公報 1971-2021年 日本国実用新案登録公報 1996-2021年 日本国登録実用新案公報 1994-2021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 2020-82332 A（オムロン株式会社）04.06.2020（2020-06-04） 特許請求の範囲	1-24
P, A	JP 2021-94677 A（本田技研工業株式会社）24.06.2021（2021-06-24） 特許請求の範囲	1-24
P, A	JP 2021-16908 A（京セラドキュメントソリューションズ株式会社）15.02.2021 （2021-02-15） 特許請求の範囲	1-24
P, A	JP 2021-64295 A（アイシン精機株式会社）22.04.2021（2021-04-22） 特許請求の範囲	1-24
P, A	JP 2021-49841 A（国立研究開発法人宇宙航空研究開発機構）01.04.2021（2021-04-01） 特許請求の範囲	1-24
☐ 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。		
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的な技術水準を示すもの “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 “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） “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 “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 “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 “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 “&” 同一パテントファミリー文献		
国際調査を完了した日 30.08.2021		国際調査報告の発送日 07.09.2021
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） 牧 初 3U 9064 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/JP2021/025932

引用文献			公表日	パテントファミリー文献			公表日
JP	2020-82332	A	04.06.2020	US	2021/0252714	A1	
				Claims			
				CN	112584990	A	
JP	2021-94677	A	24.06.2021	US	2021/0187737	A1	
				Claims			
				CN	113001539	A	
JP	2021-16908	A	15.02.2021	US	2021/0016439	A1	
				Claims			
JP	2021-64295	A	22.04.2021	(ファミリーなし)			
JP	2021-49841	A	01.04.2021	WO	2021/059787	A1	
				請求の範囲			