

(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
2024年12月12日(12.12.2024)



(10) 国際公開番号

WO 2024/253192 A1

- (51) 国際特許分類:
G06N 3/094 (2023.01) B25J 13/00 (2006.01)
- (21) 国際出願番号: PCT/JP2024/020908
- (22) 国際出願日: 2024年6月7日(07.06.2024)
- (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 2023-095038 2023年6月8日(08.06.2023) JP
- (71) 出願人: オムロン株式会社 (OMRON CORPORATION) [JP/JP]; 〒6008530 京都府京都市下京区塩小路通堀川東入南不動産町801番地 Kyoto (JP).
- (72) 発明者: 小林 聖人 (KOBAYASHI, Masato); 〒1130033 東京都文京区本郷五丁目24番5

号 ナガセ本郷ビル3F オムロンサイニックス株式会社内 Tokyo (JP). 山田 潤 (YAMADA, Jun); 〒1130033 東京都文京区本郷五丁目24番5号 ナガセ本郷ビル3F オムロンサイニックス株式会社内 Tokyo (JP). ▲浜▼屋 政志 (HAMAYA, Masashi); 〒1130033 東京都文京区本郷五丁目24番5号 ナガセ本郷ビル3F オムロンサイニックス株式会社内 Tokyo (JP). 田中 一敏 (TANAKA, Kazutoshi); 〒1130033 東京都文京区本郷五丁目24番5号 ナガセ本郷ビル3F オムロンサイニックス株式会社内 Tokyo (JP).

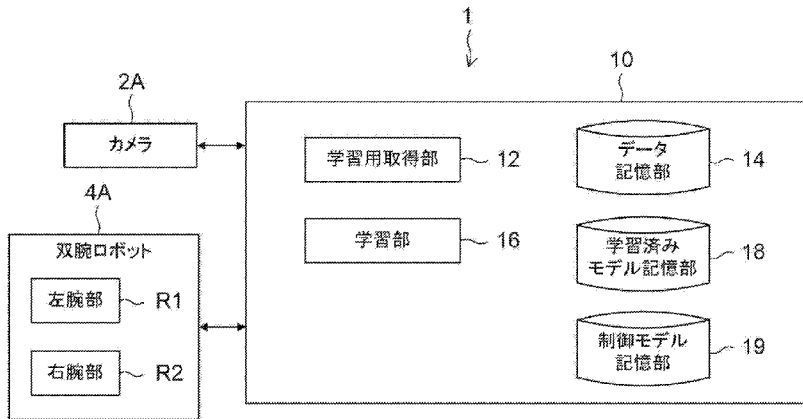
(74) 代理人: 弁理士法人太陽国際特許事務所 (TAIYO, NAKAJIMA & KATO); 〒1600022 東京都新宿区新宿4丁目3番17号 Tokyo (JP).

(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,

(54) Title: TRAINED MODEL GENERATION DEVICE, CONTROL DEVICE, TRAINED MODEL GENERATION METHOD, AND TRAINED MODEL GENERATION PROGRAM

(54) 発明の名称: 学習済みモデル生成装置、制御装置、学習済みモデル生成方法、及び学習済みモデル生成プログラム

[図5]



2A... CAMERA
4A... TWO-ARMED ROBOT
12... ACQUISITION UNIT FOR TRAINING
14... DATA STORAGE UNIT
16... TRAINING UNIT
18... TRAINED MODEL STORAGE UNIT
19... CONTROL MODEL STORAGE UNIT
R1... LEFT ARM PART
R2... RIGHT ARM PART

(57) Abstract: This trained model generation device acquires training data that includes human behavior data for training. On the basis of the training data, the trained model generation device causes a generative adversarial network model, which includes a generator and a discriminator, to carry out machine learning, thereby generating a trained generator in which behavior data of a controlled robot is output when behavior data representing behavior of a target person is input.

(57) 要約: 学習済みモデル生成装置は、学習用の人間の動作データを含む学習用データを取得する。学習済みモデル生成装置は、学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する。

BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, 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, ME, 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：

学習済みモデル生成装置、制御装置、学習済みモデル生成方法、及び学習済みモデル生成プログラム

技術分野

[0001] 本開示は、学習済みモデル生成装置、制御装置、学習済みモデル生成方法、及び学習済みモデル生成プログラムに関する。

背景技術

[0002] 従来、2つの腕部を有する双腕ロボットに対して動作を教示させる技術が知られている（例えば、Rohan Chitnis, Shubham Tulsiani, Saurabh Gupta, Abhinav Gupta, "Intrinsic Motivation for Encouraging Synergistic Behavior", ICLR 2020.を参照）。この技術では、双腕ロボットが試行錯誤をすることにより所定の動作を学習する。

[0003] また、複数の腕部を有するロボットに対して動作を教示する際に、複数の腕部の各々毎に異なる教示者によって動作を教示する技術が知られている（例えば、Albert Tung, Josiah Wong, Ajay Mandlekar, Roberto Martin, Yuk e Zhu, Li Fei-Fei, Silvio Savarese, "Learning Multi-Arm Manipulation Through Collaborative Teleoperation", ICRA, 2021.を参照）。この技術では、教示者がロボットを遠隔操作することにより、ロボットに対して動作を教示させる。

発明の概要

発明が解決しようとする課題

[0004] ところで、人間がロボットに対して動作を教示する際には、ロボットの動作に関する物理的な制約も考慮する必要がある。例えば、ロボットの可動域が人間の可動域とは異なる場合には、人間が簡単に行える動作であっても、ロボットはその動作を実行することができない場合もあり得る。また、例えば、双腕ロボットのように、ロボットの可動部位が複数である場合には、そ

の複数の可動部位を協調動作させる必要がある。人間がロボットに対して動作を教示する際に、このような複数の可動部位を協調動作させつつ、動作を教示させることは難しい。

[0005] このため、人間がロボットに対して動作を教示することは難しい、という課題がある。

[0006] 本開示は、上記の点に鑑みてなされたものであり、人間の動作データからロボットの動作データを簡易に生成することを目的とする。

課題を解決するための手段

[0007] 上記目的を達成するために、本開示に係る学習済みモデル生成装置は、学習用の人間の動作データを含む学習用データを取得する学習用取得部と、前記学習用取得部により取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する学習部と、を含む学習済みモデル生成装置である。

[0008] また、本開示の学習済みモデル生成方法は、学習用の人間の動作データを含む学習用データを取得し、取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する、処理をコンピュータが実行する学習済みモデル生成方法である。

[0009] また、本開示の学習済みモデル生成プログラムは、学習用の人間の動作データを含む学習用データを取得し、取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する、処理をコンピュータに実行させるための学習済みモデル生成

プログラムである。

発明の効果

- [0010] 本開示の学習済みモデル生成装置、制御装置、学習済みモデル生成方法、及び学習済みモデル生成プログラムによれば、人間の動作データからロボットの動作データを簡易に生成することができる。

図面の簡単な説明

- [0011] [図1]本実施形態の概要を説明するための図である。
[図2]本実施形態の概要を説明するための図である。
[図3]本実施形態のフレームワークの概要を説明するための図である。
[図4]本実施形態における敵対的生成ネットワークモデルを説明するための図である。
[図5]本実施形態の学習済みモデル生成システムの概略構成を表すブロック図である。
[図6]本実施形態に係る学習済みモデル生成装置のハードウェア構成を示すブロック図である。
[図7]本実施形態の制御システムの概略構成を表すブロック図である。
[図8]本実施形態に係る制御装置のハードウェア構成を示すブロック図である。
[図9]本実施形態における学習済みモデル生成処理の流れを示すフローチャートである。
[図10]本実施形態における学習済みモデル生成処理の流れを示すフローチャートである。
[図11]本実施形態における制御処理の流れを示すフローチャートである。
[図12]本実施例を説明するための図である。
[図13]本実施例の結果を表す図である。

発明を実施するための形態

- [0012] 以下、本開示の実施形態の一例を、図面を参照しつつ説明する。本実施形態では、本開示に係る制御装置を搭載した制御システムを例に説明する。な

お、各図面において同一又は等価な構成要素及び部分には同一の参照符号を付与している。また、図面の寸法及び比率は、説明の都合上誇張されており、実際の比率とは異なる場合がある。

[0013] <実施形態の概要>

図1及び図2は、本実施形態の概要を説明するための図である。図1に示されるように、本実施形態では、人間Hの動作を2つの腕部R1、R2を有する双腕ロボットへ教示する。図1に示されている例は、対象物bを矢印Cの方向へ移動させるタスクを実現する動作を双腕ロボットへ教示させる場合の例である。

[0014] 具体的には、図1のT1に示されているように、双腕ロボットの第1腕部R1と人間Hの腕とを協調動作させるデモンストレーションを実行することにより、2次元バーコードQ1、Q2の位置をカメラ（図示省略）によって撮像する。また、図1のT2に示されているように、双腕ロボットの第2腕部R2と人間Hの腕とを協調動作させるデモンストレーションを実行することにより、2次元バーコードQ1、Q2の位置をカメラ（図示省略）によって撮像する。

[0015] デモンストレーションにおける人間Hの腕と対象物bの動作データは、2次元バーコードQ1、Q2の位置に基づき生成される。また、デモンストレーションにおける双腕ロボットの第1腕部R1と第2腕部R2の動作データは、双腕ロボットを制御する制御装置（図示省略）から得られる。

[0016] 本実施形態では、このようにして得られた動作データに基づいて、双腕ロボットの第1腕部R1と第2腕部R2に対して教示する動作データを生成する。

[0017] そして、図2のEに示されているように、実行フェーズでは、双腕ロボットの第1腕部R1と第2腕部R2とは、対象物bを矢印Cの方向へ移動させるタスクを実行する。

[0018] この点、例えば、双腕ロボットに対して動作を教示する際には、上述したように片腕ずつ動作を教示するのではなく、第1腕部R1と第2腕部R2と

の双腕に対して一度に教示するという方法も考えられる。

[0019] 例えば、人間が両腕を用いて対象物bを矢印Cの方向へ移動させるタスクを実行し、その際の動作データを得るという方法が考えられる。しかし、この場合には、人間は、双腕ロボットが実行することが不可能な動作をしてしまう場合もある。例えば、人間が対象物bを矢印Cの方向へ移動させるタスクを実行する際に、人間は双腕ロボットの腕部の移動速度の上限を超えたような速度で対象物bを移動させてしまう場合などもあり得る。また、上述したようなタスクよりも更に複雑なタスクを実行しようとする場合には、人間は、双腕ロボットの腕部の可動域を考慮しないような動作を実行してしまう場合もあり得る。

[0020] このため、人間が両腕を用いた際の動作データが得られたとしても、その動作データをそのまま双腕ロボットへ教示させるということは難しい。

[0021] また、例えば、人間が双腕ロボットの第1腕部R1と第2腕部R2とを遠隔操作することにより、対象物bを矢印Cの方向へ移動させるタスクを実行し、その際の動作データを得るという方法が考えられる。しかし、この場合には、第1腕部R1に対する遠隔操作と第2腕部R2に対する遠隔操作とを協調させる必要がある。

[0022] 第1腕部R1を遠隔操作する人間と第2腕部R2を遠隔操作する人間とが異なる場合、タスクが複雑になるほどその協調は困難となり、適切な動作データを得ることが難しくなる。また、第1腕部R1を遠隔操作する人間と第2腕部R2を遠隔操作する人間とが同一である場合には、その遠隔操作を実現するための機器を複雑なものとする必要が出てくるため、機器を用意するコストが膨大になる。

[0023] このため、人間が双腕ロボットの第1腕部R1と第2腕部R2とを遠隔操作することにより動作データを得るということも難しい。

[0024] そこで、本実施形態では、上述したように、人間と双腕ロボットとを協調動作させるデモンストレーションを実行することにより動作データを取得し、その動作データに基づいて双腕ロボットへ教示する動作データを生成する

。

[0025] 具体的には、本実施形態では、既知の敵対的生成ネットワークモデルを利用することにより、人間の動作データから双腕ロボットの動作データを生成する。なお、以下では、図1のT1のように、第1腕部R1と人間Hとが動作する場面を「Robot1-Human」と表記する。また、図1のT2のように、第2腕部R2と人間とが動作する場面を「Human-Robot2」と表記する。また、単に人間Hと双腕ロボットの腕部とが動作する場面を「Human-Robot」と表記する。また、双腕ロボットの第1腕部R1と第2腕部R2とが動作する場面を「Robot1-Robot2」又は「Robot-Robot」と表記する。

[0026] なお、以下では、本実施形態において提案される手法を、learning from demonstrations by human and robotic arms (LfD-HR)とも称する。

[0027] 図3は、本実施形態のフレームワークの概要を説明するための図である。以下では、図3を参照しつつ、本実施形態のフレームワークを説明する。

[0028] (A. 問題定式化)

本実施形態では、マルコフ決定プロセス (S, A, P, ρ_0) により問題が定義される。なお、状態 $s \in S$ であり、行動 $a \in A$ であり、遷移関数 $P(s' \in S | s, a)$ であり、 ρ_0 は初期状態である。また、本実施形態では、2つのドメインが定義される。

[0029] ドメインXは、人間とロボットとによるタスクのデモンストレーションに属するドメインである。また、ドメインYは、双腕ロボットによるタスク動作に属するドメインである。

[0030] ドメインXに属する動作データには、人間Hの腕の状態 x^H と、双腕ロボットの第1腕部R1及び第2腕部R2の状態 x^{Ri} ($i = 1, 2$) と、対象物bの状態 x^b と、双腕ロボットの第1腕部R1及び第2腕部R2の行動 a^{Ri} とが含まれている。

[0031] ドメインYに属する動作データには、双腕ロボットの第1腕部R1及び第2腕部R2の状態 y^{Ri} と、対象物bの状態 y^b と、双腕ロボットの第1腕部R

1 及び第2腕部R 2の行動 a^{Ri} が含まれている。

[0032] 本実施形態では、ドメインXに属する状態 $x = (x^H, x^{Ri}, x^b)$ と、ドメインYに属する状態 $y = (y^{R1}, y^{R2}, y^b)$ と、行動 (a^{R1}, a^{R2}) とが定義される。また、本実施形態では、ドメインXとドメインYとにおけるマルコフ決定プロセスにおいて、 $M_x = (S_x, A_x, P_x, \rho_{0x})$ と $M_y = (S_y, A_y, P_y, \rho_{0y})$ とが定義される。また、本実施形態では、後述するように双腕ロボットの制御モデルに対して方針 $\pi : S_y \rightarrow A_y$ を学習させる。

[0033] (B. データ収集)

まず、本実施形態では、図3に示されているように、ランダムデータが収集される。具体的には、図3に示されているように、双腕ロボットの第1腕部R 1と人間Hとをランダムに動作させた状況下「Robot 1-Human」において得られるランダムデータ $\{x^{R1}, x^H, x^b, a^{R1}\}_{|D|}$ が収集される。なお、 $|D|$ はデータセットの数を表す。

[0034] また、図3に示されているように、双腕ロボットの第2腕部R 2と人間Hとをランダムに動作させた状況下「Human-Robot 2」において得られるランダムデータ $\{x^H, x^{R2}, x^b, a^{R2}\}_{|D|}$ が収集される。

[0035] また、図3に示されているように、双腕ロボットの第1腕部R 1と第2腕部R 2とをランダムに動作させた状況下「Robot 1-Robot 2」のランダムデータ $\{y^{R1}, y^{R2}, y^b, a^{R1}, a^{R2}\}_{|D|}$ が収集される。

[0036] 次に、図3に示されているように、デモンストレーションデータが収集される。デモンストレーションデータは、人間Hと双腕ロボットの腕部（第1腕部R 1又は第2腕部R 2）とを協調動作させた際に得られるデータである。

[0037] 具体的には、図3に示されているように、双腕ロボットの第1腕部R 1と人間Hとを協調動作させた状況下「Robot 1-Human」において得られるデモンストレーションデータ $\{x^{R1}, x^H, x^b, a^{R1}\}_{|D|}$ が収集される。

[0038] また、図3に示されているように、人間Hと双腕ロボットの第2腕部R 2

とを協調動作させた状況下「Human-Robot 2」において得られるデモンストレーションデータ $\{x^H, x^{R2}, x^b, a^{R2}\}_{1:D}$ が収集される。

[0039] (C. フォワードダイナミクスモデルとインバースダイナミクスモデル)

本実施形態では、双腕ロボットの腕部に関するフォワードダイナミクスモデル F とインバースダイナミクスモデル F^{-1} とが定義される。

[0040] フォワードダイナミクスモデル F は、以下の式 (1 A) によって定義される。

[0041] [数1]

$$y_{t+1} = F(y_t, a_t)$$

(1 A)

[0042] また、インバースダイナミクスモデル F^{-1} は、以下の式 (1 B) によって定義される。

[0043] [数2]

$$a_t = F^{-1}(y_t, y_{t+1})$$

(1 B)

[0044] なお、上記各式内の t は時刻を表す。フォワードダイナミクスモデル F 及びインバースダイナミクスモデル F^{-1} は、既知のダイナミクスモデル又は機械学習モデルを利用することにより実現される。

[0045] フォワードダイナミクスモデル F 及びインバースダイナミクスモデル F^{-1} は、以下の式 (1) 及び式 (2) が満たされるように、予めパラメータが設定される。

[0046] [数3]

$$\min_F L_{\text{fwd}}(F) = \mathbb{E}[\|y_{t+1} - F(y_t, a_t)\|_2]; \quad (1)$$

$$\min_{F^{-1}} L_{\text{inv}}(F^{-1}) = \mathbb{E}[\|a_t - F^{-1}(y_t, y_{t+1})\|_2]. \quad (2)$$

[0047] なお、上記各式内の $|| \cdot ||_2$ はL2ノルムを表し、 E は期待値を表す。

[0048] 本実施形態では、上記式(1)の損失関数 L_{fwd} が最小となるようなフォワードダイナミクスモデル F が予め生成される。また、上記式(2)の損失関数 L_{inv} が最小となるようなインバースダイナミクスモデル F^{-1} が予め生成される。フォワードダイナミクスモデル F とインバースダイナミクスモデル F^{-1} とは、後述するドメイントランスレーションフレームワークにおいて利用される。

[0049] (D. ドメイントランスレーションフレームワーク)

図3に示されているように、本実施形態のドメイントランスレーションフレームワークでは、人間と双腕ロボットとの協調動作のドメイン X から双腕ロボットの動作のドメイン Y への変換を実現するフレームワークである。以下、具体的に説明する。

[0050] (1) ドメイントランスレーション

本実施形態では、敵対的学習を用いて、動作データの状態 x から動作データの状態 y への写像を実行する状態遷移関数 G を生成する。具体的には、敵対的生成ネットワークモデルのうちのジェネレータを生成し、そのジェネレータを状態遷移関数とする。敵対的学習により、ジェネレータ G は、動作データの状態 x が入力されると以下の式に従って疑似的な動作データの状態 $\hat{y}$ を生成する。なお、ジェネレータ G は、敵対的生成ネットワークモデルのうちのディスクリミネータ D_y を欺くような動作データの状態 $\hat{y}$ を生成する。

[0051] [数4]

$$\hat{y} = G(x)$$

[0052] 一方で、敵対的生成ネットワークモデルのうちのディスクリミネータ D_y は、入力された動作データが、ジェネレータ G によって生成された動作データの状態 $\hat{y}$ であるのか又は現実の動作データの状態 y であるのかを見分けることを試みる。

[0053] 図4は、本実施形態における敵対的生成ネットワークモデルを説明するた

めの図である。図4に示されているように、動作データの状態 x がジェネレータ G へ入力されると、ジェネレータ G は動作データの状態 x に応じた疑似的な動作データの状態 $y^{\wedge}$ を出力する。この動作データの状態 $y^{\wedge}$ は、双腕ロボットの状態を模擬したデータである。ディスクリミネータ D_y は、この動作データ $y^{\wedge}$ が双腕ロボットの現実の状態を表す動作データであるのか否かを判定する。

[0054] 図4に示されているように、敵対的学習では、ジェネレータ G は、ディスクリミネータ D_y を欺くような動作データの状態 $y^{\wedge}$ を出力するように学習される。また、敵対的学習では、ディスクリミネータ D_y は、ジェネレータ G が出力した動作データの状態 $y^{\wedge}$ を偽物であると判定することができるように学習される。

[0055] 具体的には、本実施形態の敵対的学習では、以下の式(3)が満たされるように、敵対的生成ネットワークモデルのうちのジェネレータ G とディスクリミネータ D_y とが生成される。なお、以下の式における $p(x)$ は x が出現する確率分布を表し、 $p(y)$ は y が出現する確率分布を表す。

[0056] [数5]

$$\min_G \max_{D_y} L_{adv}(G, D_y) = \mathbb{E}_{y \sim p(y)} [\log D_y(y)] + \mathbb{E}_{x \sim p(x)} [\log(1 - D_y(G(x)))] \quad (3)$$

[0057] このため、本実施形態では、上記式(3)の敵対的生成ネットワークモデルに関する損失関数 $L_{adv}(G, D_y)$ が最小となるようにジェネレータ G が生成され、損失関数 $L_{adv}(G, D_y)$ が最大となるようにディスクリミネータ D_y が生成される。なお、本実施形態では、ジェネレータ G が、ドメイン X のデモンストレーションデータとドメイン Y のランダムデータとにオーバーフィッティングすることを防ぐために、ドメイン X のランダムデータも学習用データとして用いられる。

[0058] (2) ダイナミクス一貫性

本実施形態では、双腕ロボットのダイナミクスの一貫性も考慮して、敵対的生成ネットワークモデルを学習させる。仮に、何の制約も無くジェネレータ G を生成させた場合、ジェネレータ G によって変換された動作データの状態 $\hat{y}$ は、ドメイン Y における状態遷移 P_y との一貫性が無いものとなる場合がある。

[0059] そこで、本実施形態では、時刻 t の状態 $\hat{y}_t$ と時刻 $t + 1$ の状態 $\hat{y}_{t+1}$ との間における状態遷移 P_y の一貫性が保たれるようにジェネレータ G を生成する。

[0060] 具体的には、上述したインバースダイナミクスモデル F^{-1} によって、以下の式に従い、時刻 t に双腕ロボットが取った行動 $\tilde{a}_t$ が計算される。

[0061] [数6]

$$\tilde{a}_t = F^{-1}(\hat{y}_t, \hat{y}_{t+1})$$

[0062] また、ダイナミクスモデル F によって、以下の式に従い、時刻 $t + 1$ の状態 $\tilde{y}_{t+1}$ が計算される。

[0063] [数7]

$$\tilde{y}_{t+1} = F(\hat{y}_t, \tilde{a}_t)$$

[0064] ダイナミクスモデル F とインバースダイナミクスモデル F^{-1} とによって計算された時刻 $t + 1$ の状態 $\tilde{y}_{t+1}$ は、ジェネレータ G から出力された時刻 $t + 1$ の状態 $\hat{y}_{t+1}$ と一致している必要がある。

[0065] そこで、本実施形態では、ダイナミクスの一貫性に関する損失関数 $L_{dyn}(G)$ を設定する。ダイナミクスの一貫性に関する損失関数 $L_{dyn}(G)$ は、以下の式 (4) によって表される。

[0066] [数8]

$$\min_G L_{dyn}(G) = \mathbb{E}[\|\tilde{y}_{t+1} - \hat{y}_{t+1}\|_1] \quad (4)$$

[0067] なお、上記式内の $\|\cdot\|_1$ は L_1 ノルムを表す。本実施形態では、ダイナミクスの一貫性に関する損失関数 $L_{dyn}(G)$ が最小となるようにジェネレータ

Gが生成される。

[0068] (3) パーシャルアイデンティティマッピング

ドメイントランスレーションとダイナミクスの一貫性を考慮したとしても、ジェネレータGから、現実のデータである状態 x^{Ri} とはかけ離れた状態 y^{Ri} が出力されてしまう場合もある。例えば、ジェネレータGから双腕ロボットの腕部が取り得ない状態 y^{Ri} が出力されてしまう場合もある。

[0069] そこで、本実施形態では、パーシャルアイデンティティマッピングに関する損失関数 $L_{id}(G)$ を設定する。パーシャルアイデンティティマッピングに関する損失関数 $L_{id}(G)$ は、以下の式(5)によって表される。

[0070] [数9]

$$\min_G L_{id}(G) = E[\|\hat{y}^{Ri} - x_t^{Ri}\|_2], \quad (5)$$

[0071] 本実施形態では、パーシャルアイデンティティマッピングに関する損失関数 $L_{id}(G)$ が最小となるようにジェネレータGが生成される。

[0072] (4) 統合された損失関数

本実施形態では、上述した各損失関数を統合することにより、以下の損失関数 L_{full} が設定される。以下の式(6)の損失関数 L_{full} が最小となるように、敵対的生成ネットワークモデルのジェネレータGとディスクリミネータ D_y とが学習される。以下の式(6)における λ_{adv} , λ_{dyn} , λ_{id} は重みであり、予め設定される。

[0073] [数10]

$$L_{full} = \lambda_{adv} L_{adv}(G, D_y) + \lambda_{dyn} L_{dyn}(G) + \lambda_{id} L_{id}(G), \quad (6)$$

[0074] なお、上述したように、フォワードダイナミクスモデルFとインバースダイナミクスモデル F^{-1} とは、ランダムデータに基づいて予め学習される。また、フォワードダイナミクスモデルFとインバースダイナミクスモデル F^{-1} とは、敵対的生成ネットワークモデルを学習させる際に併せて更新するようにしてもよい。例えば、フォワードダイナミクスモデルFが有するパラメー

タとインバースダイナミクスモデル F^{-1} が有するパラメータとが、敵対的生成ネットワークモデルを学習させる際に併せて更新されるようにしてもよい。

[0075] (学習済みモデル生成システム1)

図5は、本実施形態の学習済みモデル生成システム1の概略構成を表すブロック図である。図5に示されるように、学習済みモデル生成システム1は、カメラ2Aと、双腕ロボット4Aと、学習済みモデル生成装置10とを備えている。本実施形態に係る学習済みモデル生成装置10は、人間の動作データから双腕ロボットの動作データを生成するための学習済みモデルを生成する。

[0076] カメラ2Aは、制御対象である双腕ロボット4Aの第1腕部R1及び第2腕部R2と人間Hとが動作をしている際の画像を逐次撮像する。例えば、カメラ2Aは、図1に示されているような、デモンストレーションT1とデモンストレーションT2とが行われている際の、画像を逐次撮像する。また、カメラ2Aは、双腕ロボット4Aの第1腕部R1及び第2腕部R2がランダムな動作をしている際の画像を逐次撮像する。また、カメラ2Aは、人間がランダムな動作をしている際の画像を逐次撮像する。そして、カメラ2Aは、得られた画像データを学習済みモデル生成装置10へ出力する。

[0077] 双腕ロボット4Aは、図1に示されているようなロボットであり、第1腕部R1と第2腕部R2とを備えている。

[0078] 図6は、本実施形態に係る学習済みモデル生成装置10のハードウェア構成を示すブロック図である。図6に示されるように、学習済みモデル生成装置10は、CPU (Central Processing Unit) 42、メモリ44、記憶装置46、入出力I/F (Interface) 48、記憶媒体読取装置50、及び通信I/F 52を有する。各構成は、バス54を介して相互に通信可能に接続されている。

[0079] 記憶装置46には、後述する各処理を実行するための学習済みモデル生成プログラムが格納されている。CPU 42は、中央演算処理ユニットであり

、各種プログラムを実行したり、各構成を制御したりする。すなわち、CPU 42は、記憶装置46からプログラムを読み出し、メモリ44を作業領域としてプログラムを実行する。CPU 42は、記憶装置46に記憶されているプログラムに従って、上記各構成の制御及び各種の演算処理を行う。

[0080] メモリ44は、RAM (Random Access Memory) により構成され、作業領域として一時的にプログラム及びデータを記憶する。記憶装置46は、ROM (Read Only Memory)、及びHDD (Hard Disk Drive)、SSD (Solid State Drive) 等により構成され、オペレーティングシステムを含む各種プログラム、及び各種データを格納する。

[0081] 入出力I/F 48は、カメラ2Aと双腕ロボット4Aとからのデータの入力、及びカメラ2Aと双腕ロボット4Aとへのデータの出力を行うインタフェースである。また、例えば、キーボードやマウス等の、各種の入力を行うための入力装置、及び、例えば、ディスプレイやプリンタ等の、各種の情報を出力するための出力装置が接続されてもよい。出力装置として、タッチパネルディスプレイを採用することにより、入力装置として機能させてもよい。

[0082] 記憶媒体読取装置50は、CD (Compact Disc) -ROM、DVD (Digital Versatile Disc) -ROM、ブルーレイディスク、USB (Universal Serial Bus) メモリ等の各種記憶媒体に記憶されたデータの読み込みや、記憶媒体に対するデータの書き込み等を行う。

[0083] 通信I/F 52は、他の機器と通信するためのインタフェースであり、例えば、イーサネット (登録商標)、FDDI、Wi-Fi (登録商標) 等の規格が用いられる。

[0084] 次に、学習済みモデル生成装置10の機能構成について説明する。図5に示されるように、学習済みモデル生成装置10は、機能的には、学習用取得部12と、学習部16とを含む。また、学習済みモデル生成装置10の所定

の記憶領域には、データ記憶部 14 と、
学習済みモデル記憶部 18 と、制御モデル記憶部 19 とが設けられている。
各機能構成は、CPU 42 が記憶装置 46 に記憶された各プログラムを読み出し、メモリ 44 に展開して実行することにより実現される。

[0085] データ記憶部 14 には、カメラ 2A によって撮像された画像データが格納される。また、データ記憶部 14 には、双腕ロボット 4A が動作した際の制御データが格納される。

[0086] 学習済みモデル記憶部 18 には、後述する処理によって生成された学習済みの敵対的生成ネットワークモデルが格納される。

[0087] 制御モデル記憶部 19 には、双腕ロボット 4A を制御するための制御モデルが格納される。

[0088] 学習用取得部 12 は、学習用データを取得する。具体的には、データ記憶部 14 に格納されている画像データ及び制御データを加工するなどして、後述する敵対的生成ネットワークモデルを学習させる際に利用する学習用データを計算する。本実施形態の学習用データは、学習用の人間 H の動作データと、学習用の双腕ロボット 4A の動作データとを含むデータである。

[0089] 具体的には、本実施形態の学習用データは、第 1 腕部 R1 と人間 H の腕とによる協調動作を表すデモンストレーションデータと、第 2 腕部 R2 と人間 H の腕とによる協調動作を表すデモンストレーションデータと、双腕ロボット 4A のランダムな動作を表すランダムデータと、人間 H のランダムな動作を表すランダムデータとを含んで構成されている。

[0090] 第 1 腕部 R1 と人間 H の腕とによる協調動作を表すデモンストレーションデータは、図 3 に示されている「Robot 1-Human」のデモンストレーションデータ $\{x^{R1}, x^H, x^b, a^{R1}\} \mid D\}$ である。第 2 腕部 R2 と人間 H の腕とによる協調動作を表すデモンストレーションデータは、図 3 に示されている「Human-Robot 2」のデモンストレーションデータ $\{x^H, x^{R2}, x^b, a^{R2}\} \mid D\}$ である。

[0091] 双腕ロボット 4A のランダムな動作を表すランダムデータ及び人間 H のラ

ランダムな動作を表すランダムデータは、図3に示されている「Robot 1-Human」のランダムデータ $\{x^{R1}, x^H, x^b, a^{R1}\}_{|D|}$ と、「Human-Robot 2」のランダムデータ $\{x^H, x^{R2}, x^b, a^{R2}\}_{|D|}$ と、「Robot 1-Robot 2」のランダムデータ $\{y^{R1}, y^{R2}, y^b, a^{R1}, a^{R2}\}_{|D|}$ とである。

[0092] 学習用取得部12は、データ記憶部14に格納されている画像データ及び制御データを解析することにより、図1に示されている2次元バーコードQ1、Q2の位置及び移動速度等を特定する。そして、学習用取得部12は、2次元バーコードQ1、Q2の位置及び移動速度と双腕ロボット4Aの制御データとを組み合わせることにより、上述したようなデモンストレーションデータ及びランダムデータを取得する。

[0093] 学習部16は、学習用取得部12により取得された学習用データに基づいて、ジェネレータGとディスクリミネータD_yとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間Hの動作を表す動作データxが入力されると双腕ロボット4Aの動作データyが出力される学習済みのジェネレータGを生成する。

[0094] 具体的には、学習部16は、上記式(6)に示されている統合損失関数 L_{full} が最小となるように、敵対的生成ネットワークモデルを機械学習させる。なお、上述したように、統合損失関数 L_{full} は、敵対的生成ネットワークモデルに関する損失関数 L_{adv} と、ダイナミクスの一貫性に関する損失関数 L_{dyn} と、パーシャルアイデンティティマッピングに関する損失関数 L_{id} とを含む関数である。以下、各損失関数の最小化について説明する。

[0095] (敵対的生成ネットワークモデルに関する損失関数 L_{adv})

学習部16は、実質的に、上記式(3)に示されている敵対的生成ネットワークモデルに関する損失関数 L_{adv} が最小になるようにジェネレータGを学習させつつ、敵対的生成ネットワークモデルに関する損失関数 L_{adv} が最大になるようにディスクリミネータD_yを学習させる。

[0096] (ダイナミクスの一貫性に関する損失関数 L_{dyn})

学習部 16 は、実質的に、上記式 (4) に示されているダイナミクスの一貫性に関する損失関数 L_{dyn} が最小になるようにジェネレータ G を学習させる。

[0097] 上述したように、本実施形態では、双腕ロボット 4 A の時刻 t の状態 y_t と双腕ロボット 4 A の時刻 t の行動 a_t とが入力されると、双腕ロボット 4 A の時刻 $t + 1$ の状態 y_{t+1} を出力するフォワードダイナミクスモデル F が予め用意される。また、本実施形態では、双腕ロボット 4 A の時刻 t の状態 y_t と双腕ロボット 4 A の時刻 $t + 1$ の状態 y_{t+1} とが入力されると、双腕ロボット 4 A が時刻 t に取ったであろう行動 a_t が出力されるインバースダイナミクスモデル F^{-1} とが予め用意される。

[0098] このため、学習部 16 は、敵対的生成ネットワークモデルを機械学習させる際に、インバースダイナミクスモデル F^{-1} に対して、ジェネレータ G から出力された双腕ロボット 4 A の時刻 t の状態 $y_t^{\wedge}$ と時刻 $t + 1$ の状態 $y_{t+1}^{\wedge}$ とを入力することにより、双腕ロボット 4 A が時刻 t に取ったと推定される行動 $a_t^{\sim}$ を計算する。

[0099] また、学習部 16 は、敵対的生成ネットワークモデルを機械学習させる際に、ダイナミクスモデル F に対して、ジェネレータ G から出力された双腕ロボット 4 A の状態 $y_t^{\wedge}$ と、インバースダイナミクスモデル F^{-1} により計算された行動 $a_t^{\sim}$ とを入力することにより、双腕ロボット 4 A の時刻 $t + 1$ の状態 $y_{t+1}^{\sim}$ を計算する。

[0100] そして、学習部 16 は、ジェネレータ G から出力された双腕ロボット 4 A の時刻 t の状態 $y_{t+1}^{\wedge}$ と、計算された双腕ロボット 4 A の時刻 $t + 1$ の状態 $y_{t+1}^{\sim}$ との間の差分が小さくなるようにジェネレータ G を学習させる。

[0101] (パーシャルアイデンティティマッピングに関する損失関数 L_{id})

学習部 16 は、実質的に、上記式 (5) に示されているパーシャルアイデンティティマッピングに関する損失関数 L_{id} が最小になるようにジェネレータ G を学習させる。

- [0102] 具体的には、学習部 16 は、学習用の双腕ロボット 4 A の動作データ x とジェネレータ G から出力された双腕ロボット 4 A の状態 y^* との間の差分が小さくなるように、ジェネレータ G を学習させることにより、学習済みのジェネレータ G を生成する。
- [0103] そして、学習部 16 は、学習済みのジェネレータ G と学習済みのディスクリミネータ D_y とを含む学習済みの敵対的生成ネットワークモデルを、学習済みモデル記憶部 18 へ格納する。
- [0104] 次に、学習部 16 は、学習済みのジェネレータ G に対して、対象となる人間の動作を表す対象動作データを入力することにより、双腕ロボット 4 A のロボットの動作データを生成する。ここで、対象となる人間の動作とは、双腕ロボット 4 A へ教示したい動作に相当する。
- [0105] 学習部 16 は、生成された双腕ロボット 4 A の動作データに基づいて、双腕ロボット 4 A を制御するための制御用モデルを学習させることにより、動作データのうちの状態 y が入力されると動作データのうちの行動 a を出力する制御用学習済みモデルを生成する。
- [0106] 例えば、学習部 16 は、既知の模倣学習を用いて制御用学習済みモデルを生成する。これにより、状態 y から行動 a を生成する方策が反映された制御用学習済みモデルを得ることができる。なお、制御用モデルとしては、既知の関数又は機械学習モデルを採用することができる。
- [0107] そして、学習部 16 は、制御用学習済みモデルを制御モデル記憶部 19 へ格納する。
- [0108] (制御システム 20)

図 7 は、本実施形態の制御システム 20 の概略構成を表すブロック図である。図 7 に示されるように、制御システム 20 は、カメラ 2 B と、双腕ロボット 4 B と、制御装置 30 とを備えている。本実施形態に係る制御装置 30 は、学習済みモデル生成装置 10 によって生成された制御用学習済みモデルを用いて双腕ロボット 4 B の動作を制御する。

- [0109] カメラ 2 B は、上述したカメラ 2 A と同様の構成であり、制御対象である双腕ロボット 4 B の第 1 腕部 R 1 及び第 2 腕部 R 2 が動作をしている際の画像を逐次撮像する。そして、カメラ 2 B は、得られた画像データを制御装置 3 0 へ出力する。
- [0110] 双腕ロボット 4 B は、上述した双腕ロボット 4 A と同様の構成であり、図 1 に示されているようなロボットである。
- [0111] 図 8 は、本実施形態に係る制御装置 3 0 のハードウェア構成を示すブロック図である。図 8 に示されるように、制御装置 3 0 は、CPU (Central Processing Unit) 6 2、メモリ 6 4、記憶装置 6 6、入出力 I / F (Interface) 6 8、記憶媒体読取装置 7 0、及び通信 I / F 7 2 を有する。各構成は、バス 7 4 を介して相互に通信可能に接続されている。
- [0112] 記憶装置 6 6 には、後述する各処理を実行するための制御プログラムが格納されている。CPU 6 2 は、中央演算処理ユニットであり、各種プログラムを実行したり、各構成を制御したりする。すなわち、CPU 6 2 は、記憶装置 6 6 からプログラムを読み出し、メモリ 6 4 を作業領域としてプログラムを実行する。CPU 6 2 は、記憶装置 6 6 に記憶されているプログラムに従って、上記各構成の制御及び各種の演算処理を行う。
- [0113] メモリ 6 4 は、RAM (Random Access Memory) により構成され、作業領域として一時的にプログラム及びデータを記憶する。記憶装置 6 6 は、ROM (Read Only Memory)、及び HDD (Hard Disk Drive)、SSD (Solid State Drive) 等により構成され、オペレーティングシステムを含む各種プログラム、及び各種データを格納する。
- [0114] 入出力 I / F 6 8 は、カメラ 2 B と双腕ロボット 4 B とからのデータの入力、及びカメラ 2 B と双腕ロボット 4 B とへのデータの出力を行うインタフェースである。また、例えば、キーボードやマウス等の、各種の入力を行うための入力装置、及び、例えば、ディスプレイやプリンタ等の、各種の情報

を出力するための出力装置が接続されてもよい。出力装置として、タッチパネルディスプレイを採用することにより、入力装置として機能させてもよい。

[0115] 記憶媒体読取装置70は、CD (Compact Disc) -ROM、DVD (Digital Versatile Disc) -ROM、ブルーレイディスク、USB (Universal Serial Bus) メモリ等の各種記憶媒体に記憶されたデータの読み込みや、記憶媒体に対するデータの書き込み等を行う。

[0116] 通信I/F72は、他の機器と通信するためのインタフェースであり、例えば、イーサネット (登録商標)、FDDI、Wi-Fi (登録商標) 等の規格が用いられる。

[0117] 次に、制御装置30の機能構成について説明する。図7に示されるように、制御装置30は、機能的には、取得部34と、生成部36と、制御部38とを含む。また、制御装置30の所定の記憶領域には、制御モデル記憶部32が設けられている。各機能構成は、CPU62が記憶装置66に記憶された各プログラムを読み出し、メモリ64に展開して実行することにより実現される。

[0118] 制御モデル記憶部32には、学習済みモデル生成装置10によって生成された制御用学習済みモデルが格納される。

[0119] 取得部34は、双腕ロボット4Bの状態と対象物の状態とを取得する。具体的には、取得部34は、カメラ2Bによって撮像された画像データ及び双腕ロボット4Bの制御データに基づいて、双腕ロボット4Bと対象物bとの状態 (y^{R1} , y^{R2} , y^b) を計算する。

[0120] 生成部36は、制御モデル記憶部32に格納されている制御用学習済みモデルに対して、取得部34により取得された状態 (y^{R1} , y^{R2} , y^b) を入力することにより、当該状態 (y^{R1} , y^{R2} , y^b) に応じた双腕ロボット4Bの行動 (a^{R1} , a^{R2}) を生成する。

[0121] 制御部38は、生成部36により生成された行動 (a^{R1} , a^{R2}) を取るよ

うに、双腕ロボット4 Bを制御する。具体的には、制御部3 8は、生成部3 6により生成された行動(a^{R1} , a^{R2})を取るように、双腕ロボット4 Bの第1腕部R 1と第2腕部R 2とに対して制御指令を出力する。

[0122] 次に、本実施形態に係る学習済みモデル生成システム1の作用について説明する。

[0123] まず、人間Hの動作及び双腕ロボット4 Aの動作に関するデータが収集され、学習済みモデル生成装置1 0へ入力される。人間Hの動作及び双腕ロボット4 Aの動作に関するデータは、データ記憶部1 4へ格納される。そして、学習済みモデル生成システム1が所定の指示信号を受け付けると、学習済みモデル生成装置1 0のCPU 4 2は記憶装置4 6から学習済みモデル生成プログラムを読み出して、メモリ4 4に展開して実行する。これにより、CPU 4 2が学習済みモデル生成装置1 0の各機能構成として機能し、図9及び図1 0に示す学習済みモデル生成処理が実行される。

[0124] ステップS 1 0 0において、学習用取得部1 2は、データ記憶部1 4へ格納されているデータから学習用データを取得する。

[0125] ステップS 1 0 2において、学習部1 6は、ステップS 1 0 0で取得された学習用データに基づいて、ジェネレータGとディスクリミネータD_yとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間Hの動作を表す動作データxが入力されると双腕ロボット4 Aの動作データyが出力される学習済みのジェネレータGを生成する。

[0126] ステップS 1 0 4において、学習済みのジェネレータGと学習済みのディスクリミネータD_yとを含む学習済みの敵対的生成ネットワークモデルを、学習済みモデル記憶部1 8へ格納する。

[0127] 次に、学習済みモデル生成システム1が所定の指示信号を受け付けると、学習済みモデル生成装置1 0は、図1 0に示す学習済みモデル生成処理を実行する。

[0128] ステップS 2 0 0において、学習部1 6は、対象となる人間Hの動作を表

す対象動作データを取得する。例えば、学習部 16 は、学習用データに含まれる人間 H の動作データのうち、双腕ロボット 4 A に教示したい動作データを対象動作データとして取得する。なお、対象動作データとしては学習用データとは異なるデータが用いられてもよい。

[0129] ステップ S 202 において、学習部 16 は、学習済みモデル記憶部 18 から学習済みのジェネレータ G を読み出す。

[0130] ステップ S 204 において、学習部 16 は、ステップ S 202 で読み出された学習済みのジェネレータ G に対して、ステップ S 200 で取得された対象動作データを入力することにより、双腕ロボット 4 A の動作データを生成する。

[0131] ステップ S 206 において、学習部 16 は、ステップ S 204 で得られた双腕ロボット 4 A の動作データに基づいて、双腕ロボット 4 A を制御するための制御用モデルを学習させることにより、動作データのうちの状態 y が入力されると動作データのうちの行動 a を出力する制御用学習済みモデルを生成する。

[0132] ステップ S 208 において、学習部 16 は、ステップ S 206 で生成された制御用学習済みモデルを制御モデル記憶部 19 へ格納する。

[0133] 次に、本実施形態に係る制御システム 20 の作用について説明する。

[0134] 学習済みモデル生成システム 1 によって生成された制御用学習済みモデルが、制御装置 30 へ入力されると、その制御用学習済みモデルは制御モデル記憶部 19 へ格納される。そして、制御システム 20 が所定の指示信号を受け付けると、制御装置 30 の CPU 62 は記憶装置 66 から制御プログラムを読み出して、メモリ 64 に展開して実行する。これにより、CPU 62 が制御装置 30 の各機能構成として機能し、図 11 に示す制御処理が実行される。

[0135] ステップ S 300 において、取得部 34 は、カメラ 2 B によって撮像された画像データ及び双腕ロボット 4 B の制御データから、双腕ロボット 4 B と対象物 b との状態 (y^{R1} , y^{R2} , y^b) を取得する。

- [0136] ステップS302において、生成部36は、制御モデル記憶部32に格納されている制御用学習済みモデルを読み出す。
- [0137] ステップS304において、生成部36は、ステップS302で読み出した制御用学習済みモデルに対して、ステップS300で取得された状態 (y^{R1} , y^{R2} , y^b) を入力することにより、当該状態 (y^{R1} , y^{R2} , y^b) に応じた双腕ロボット4Bの行動 (a^{R1} , a^{R2}) を生成する
- [0138] ステップS306において、制御部38は、ステップS304で生成された行動 (a^{R1} , a^{R2}) を取るように、双腕ロボット4Bを制御する。
- [0139] 図11に示される制御処理が繰り返され、双腕ロボット4Aに対して制御信号が繰り返し出力されることにより、対象物bに対するタスクが実行される。
- [0140] 以上説明したように、本実施形態に係る学習済みモデル生成装置は、学習用の人間の動作データを含む学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する。これにより、人間の動作データからロボットの動作データを簡易に生成することができる学習済みのジェネレータを得ることができる。
- [0141] また、本実施形態に係る学習済みモデル生成装置は、ダイナミクスの一貫性に関する損失関数 L_{dyn} を含む損失関数 L_{full} が最小になるように、学習済みのジェネレータを生成することにより、双腕ロボット4Aのダイナミクスの一貫性を考慮した動作データを生成することができる学習済みのジェネレータを得ることができる。これにより、双腕ロボット4Aのダイナミクスを無視したような動作データの生成は抑制される。
- [0142] また、本実施形態に係る学習済みモデル生成装置は、パーシャルアイデンティティマッピングに関する損失関数 L_{id} を含む損失関数 L_{full} が最小になるように、学習済みのジェネレータを生成することにより、双腕ロボット4Aの現実の動きに適合するような動作データを生成することができる学習済

みのジェネレータを得ることができる。これにより、双腕ロボット4 Aの現実の動きを無視したような動作データの生成が抑制される。

[0143] ダイナミクスの一貫性に関する損失関数 L_{dyn} とパーシャルアイデンティティマッピングに関する損失関数 L_{id} とを考慮して学習済みのジェネレータを生成することにより、例えば、双腕ロボット4 Aの可動域を無視した動作データの生成等が抑制される。

[0144] また、双腕ロボット4 Aのランダムな動作を表すランダムデータを学習用データに含ませることにより、制御対象ロボットはどの程度の可動域を有しているのかといった点を考慮して、学習済みのジェネレータを生成することができる。

実施例

[0145] 次に、実施例について説明する。本実施例では、提案したLfD-HRの有効性を検証するためのシミュレーションを行う。

[0146] 図12は、本実施例を説明するための図である。図12に示されるように、本実施例では、3つのシミュレーションを行う。図12に示されている「Demonstration」は、上述

した図1のデモンストレーションT1, T2に相当する。また、図12に示されている「Execution」は、上述した図2の実行フェーズEに相当する。図12 (a) Dual-arm block

pushは、Robot1とRobot2とがBlockを押し出すような協調動作を対象とするものである。また、図12 (b) Dual-arm peg insertionは、Robot1が把持する穴にRobot2が把持するpegを挿入するような協調動作を対象とするものである。

[0147] 図13は、本実施例の結果を表す図である。図13に示されているように、提案したLfD-HRによれば、Robot1とRobot2とが協調動作することにより目的のタスクを完遂することができている。

[0148] なお、上記実施形態では、制御対象ロボットが双腕ロボットである場合を例に説明した

がこれに限定されるものではなく、どのようなロボットを対象としてもよい。例えば、1つの腕を有するロボットを対象としてもよい。制御対象ロボットとしては、例えば、複数の指を有するロボット等も対象になり得る。この場合には、複数の指の動きに関する動作データが生成される。

[0149] また、上記実施形態でCPUがソフトウェア（プログラム）を読み込んで実行した各処理を、CPU以外の各種のプロセッサが実行してもよい。この場合のプロセッサとしては、FPGA（Field-Programmable Gate Array）等の製造後に回路構成を変更可能なPLD（Programmable Logic Device）、及びASIC（Application Specific Integrated Circuit）等の特定の処理を実行させるために専用に設計された回路構成を有するプロセッサである専用電気回路等が例示される。また、各処理を、これらの各種のプロセッサのうちの1つで実行してもよいし、同種又は異種の2つ以上のプロセッサの組み合わせ（例えば、複数のFPGA、及びCPUとFPGAとの組み合わせ等）で実行してもよい。また、これらの各種のプロセッサのハードウェア的な構造は、より具体的には、半導体素子等の回路素子を組み合わせた電気回路である。

[0150] また、上記実施形態では、各プログラムが記憶装置に予め記憶（インストール）されている態様を説明したが、これに限定されない。プログラムは、CD-ROM、DVD-ROM、ブルーレイディスク、USBメモリ等の記憶媒体に記憶された形態で提供されてもよい。また、プログラムは、ネットワークを介して外部装置からダウンロードされる形態としてもよい。

[0151] （付記）

以下、本開示の態様について付記する。

[0152] （付記1）

学習用の人間の動作データを含む学習用データを取得する学習用取得部と

、

前記学習用取得部により取得された学習用データに基づいて、ジェネレー

タとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する学習部と、

を含む学習済みモデル生成装置。

[0153] (付記2)

前記動作データは、前記制御対象ロボットの状態 y と行動 a とを含んで構成されており、

前記制御対象ロボットの時刻 t の状態 y_t と前記制御対象ロボットの時刻 t の行動 a_t とが入力されると、前記制御対象ロボットの時刻 $t + 1$ の状態 y_{t+1} を出力するフォワードダイナミクスモデル F と、

前記制御対象ロボットの時刻 t の状態 y_t と前記制御対象ロボットの時刻 $t + 1$ の状態 y_{t+1} とが入力されると、前記制御対象ロボットが時刻 t に取った行動 a_t が出力されるインバースダイナミクスモデル F^{-1} とが予め用意されており、

前記学習部は、

前記インバースダイナミクスモデル F^{-1} に対して、前記ジェネレータから出力された前記制御対象ロボットの時刻 t の状態 $y_t^{\wedge}$ と時刻 $t + 1$ の状態 $y_{t+1}^{\wedge}$ とを入力することにより、前記制御対象ロボットが時刻 t に取ったであろう行動 $a_t^{\sim}$ を計算し、

前記フォワードダイナミクスモデル F に対して、前記状態 $y_t^{\wedge}$ と前記行動 $a_t^{\sim}$ とを入力することにより、前記制御対象ロボットの時刻 $t + 1$ の状態 $y_{t+1}^{\sim}$ を計算し、

前記ジェネレータから出力された前記制御対象ロボットの時刻 t の状態 $y_t^{\wedge}$ と、

計算された前記制御対象ロボットの時刻 $t + 1$ の状態 $y_{t+1}^{\sim}$ との間の差分が小さくなるように、前記ジェネレータを学習させることにより、前記学習済みのジェネレータを生成する、

付記 1 に記載の学習済みモデル生成装置。

[0154] (付記 3)

前記学習用データには、学習用の制御対象ロボットの動作データ x が更に含まれており、

前記学習部は、学習用の制御対象ロボットの動作データ x と、前記ジェネレータから出力された前記制御対象ロボットの状態 $y^{\wedge}$ との間の差分が小さくなるように、前記ジェネレータを学習させることにより、前記学習済みのジェネレータを生成する、

付記 1 又は付記 2 に記載の学習済みモデル生成装置。

[0155] (付記 4)

前記制御対象ロボットは、少なくとも 1 つ以上の腕部を備えたロボットである、

付記 1 ～付記 3 の何れか 1 項に記載の学習済みモデル生成装置。

[0156] (付記 5)

前記制御対象ロボットは、第 1 腕部と第 2 腕部とを備える双腕ロボットであり、

前記学習用データは、

前記第 1 腕部と前記人間の腕とによる協調動作を表すデモンストラーションデータと、

前記第 2 腕部と前記人間の腕とによる協調動作を表すデモンストラーションデータと、を更に含んで構成されている、

付記 4 に記載の学習済みモデル生成装置。

[0157] (付記 6)

前記学習用データは、

前記制御対象ロボットのランダムな動作を表すランダムデータと、

前記人間のランダムな動作を表すランダムデータと、を更に含んで構成されている、

付記 1 ～付記 3 の何れか 1 項に記載の学習済みモデル生成装置。

[0158] (付記 7)

前記学習部は、

前記学習済みのジェネレータに対して、対象となる人間の動作を表す対象動作データを入力することにより、前記制御対象ロボットの動作データを生成し、

生成された前記制御対象ロボットの動作データに基づいて、前記制御対象ロボットを制御するための制御用学習済みモデルであって、前記動作データのうちの状態が入力されると、前記動作データのうちの行動を出力するための制御用学習済みモデルを生成する、

付記 1 ～付記 3 の何れか 1 項に記載の学習済みモデル生成装置。

[0159] (付記 8)

制御対象ロボットの状態を取得する取得部と、

付記 7 に記載の学習済みモデル生成装置によって生成された前記制御用学習済みモデルに対して、前記取得部により取得された前記状態を入力することにより、前記状態に応じた前記制御対象ロボットの行動を生成する生成部と、

前記生成部により生成された前記行動を取るように、前記制御対象ロボットを制御する制御部と、

を備えた制御装置。

[0160] (付記 9)

学習用の人間の動作データを含む学習用データを取得し、

取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する、

処理をコンピュータが実行する学習済みモデル生成方法。

[0161] (付記 10)

学習用の人間の動作データを含む学習用データを取得し、

取得された学習用データに基づいて、ジェネレータとディスクリミネータを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する、

処理をコンピュータに実行させるための学習済みモデル生成プログラム。

[0162] 2023年6月8日に出願された日本国特許出願2023-095038号の開示は、その全体が参照により本明細書に取り込まれる。本明細書に記載された全ての文献、特許出願、および技術規格は、個々の文献、特許出願、および技術規格が参照により取り込まれることが具体的かつ個々に記された場合と同程度に、本明細書中に参照により取り込まれる。

請求の範囲

[請求項1] 学習用の人間の動作データを含む学習用データを取得する学習用取得部と、

前記学習用取得部により取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する学習部と、

を含む学習済みモデル生成装置。

[請求項2] 前記動作データは、前記制御対象ロボットの状態 y と行動 a とを含んで構成されており、

前記制御対象ロボットの時刻 t の状態 y_t と前記制御対象ロボットの時刻 t の行動 a_t とが入力されると、前記制御対象ロボットの時刻 $t + 1$ の状態 y_{t+1} を出力するフォワードダイナミクスモデル F と、

前記制御対象ロボットの時刻 t の状態 y_t と前記制御対象ロボットの時刻 $t + 1$ の状態 y_{t+1} とが入力されると、前記制御対象ロボットが時刻 t に取った行動 a_t が出力されるインバースダイナミクスモデル F^{-1} とが予め用意されており、

前記学習部は、

前記インバースダイナミクスモデル F^{-1} に対して、前記ジェネレータから出力された前記制御対象ロボットの時刻 t の状態 $y_t^{\wedge}$ と時刻 $t + 1$ の状態 $y_{t+1}^{\wedge}$ とを入力することにより、前記制御対象ロボットが時刻 t に取ったであろう行動 $a_t^{\sim}$ を計算し、

前記フォワードダイナミクスモデル F に対して、前記状態 $y_t^{\wedge}$ と前記行動 $a_t^{\sim}$ とを入力することにより、前記制御対象ロボットの時刻 $t + 1$ の状態 $y_{t+1}^{\sim}$ を計算し、

前記ジェネレータから出力された前記制御対象ロボットの時刻 t の状態 $y_t^{\wedge}$ と、計算された前記制御対象ロボットの時刻 $t + 1$ の状

態 y_{t+1} との間の差分が小さくなるように、前記ジェネレータを学習させることにより、前記学習済みのジェネレータを生成する、

請求項 1 に記載の学習済みモデル生成装置。

[請求項3] 前記学習用データには、学習用の制御対象ロボットの動作データ x が更に含まれており、

前記学習部は、学習用の制御対象ロボットの動作データ x と、前記ジェネレータから出力された前記制御対象ロボットの状態 y^* との間の差分が小さくなるように、前記ジェネレータを学習させることにより、前記学習済みのジェネレータを生成する、

請求項 1 又は請求項 2 に記載の学習済みモデル生成装置。

[請求項4] 前記制御対象ロボットは、少なくとも 1 つ以上の腕部を備えたロボットである、

請求項 1 又は請求項 2 に記載の学習済みモデル生成装置。

[請求項5] 前記制御対象ロボットは、第 1 腕部と第 2 腕部とを備える双腕ロボットであり、

前記学習用データは、

前記第 1 腕部と前記人間の腕とによる協調動作を表すデモンストラーションデータと、

前記第 2 腕部と前記人間の腕とによる協調動作を表すデモンストラーションデータと、を更に含んで構成されている、

請求項 4 に記載の学習済みモデル生成装置。

[請求項6] 前記学習用データは、

前記制御対象ロボットのランダムな動作を表すランダムデータと、

前記人間のランダムな動作を表すランダムデータと、を更に含んで構成されている、

請求項 1 又は請求項 2 に記載の学習済みモデル生成装置。

[請求項7] 前記学習部は、

前記学習済みのジェネレータに対して、対象となる人間の動作を表

す対象動作データを入力することにより、前記制御対象ロボットの動作データを生成し、

生成された前記制御対象ロボットの動作データに基づいて、前記制御対象ロボットを制御するための制御用学習済みモデルであって、前記動作データのうちの状態が入力されると、前記動作データのうちの行動を出力するための制御用学習済みモデルを生成する、

請求項 1 又は請求項 2 に記載の学習済みモデル生成装置。

[請求項8]

制御対象ロボットの状態を取得する取得部と、

請求項 7 に記載の学習済みモデル生成装置によって生成された前記制御用学習済みモデルに対して、前記取得部により取得された前記状態を入力することにより、前記状態に応じた前記制御対象ロボットの行動を生成する生成部と、

前記生成部により生成された前記行動を取るように、前記制御対象ロボットを制御する制御部と、

を備えた制御装置。

[請求項9]

学習用の人間の動作データを含む学習用データを取得し、

取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する、

処理をコンピュータが実行する学習済みモデル生成方法。

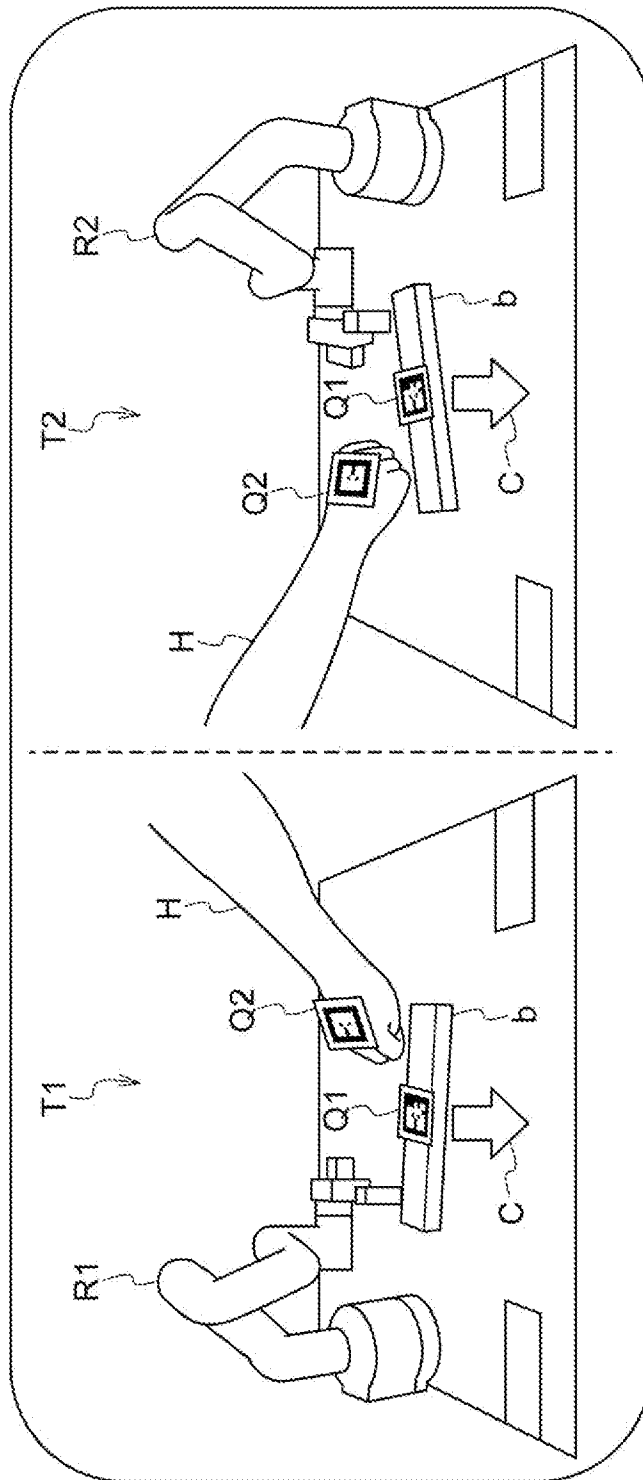
[請求項10]

学習用の人間の動作データを含む学習用データを取得し、

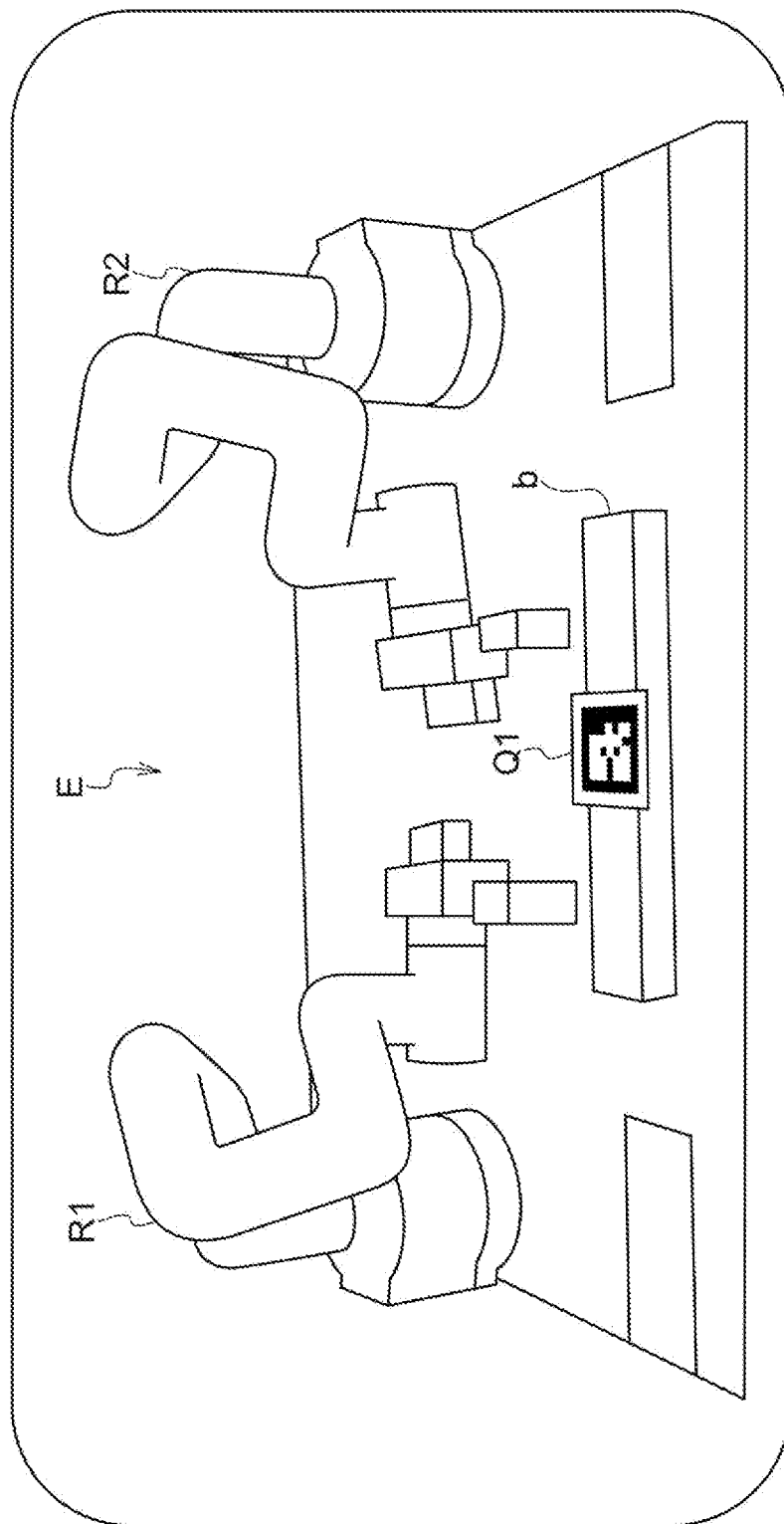
取得された学習用データに基づいて、ジェネレータとディスクリミネータとを含む敵対的生成ネットワークモデルを機械学習させることにより、対象となる人間の動作を表す動作データが入力されると制御対象ロボットの動作データが出力される学習済みのジェネレータを生成する、

処理をコンピュータに実行させるための学習済みモデル生成プログラム。

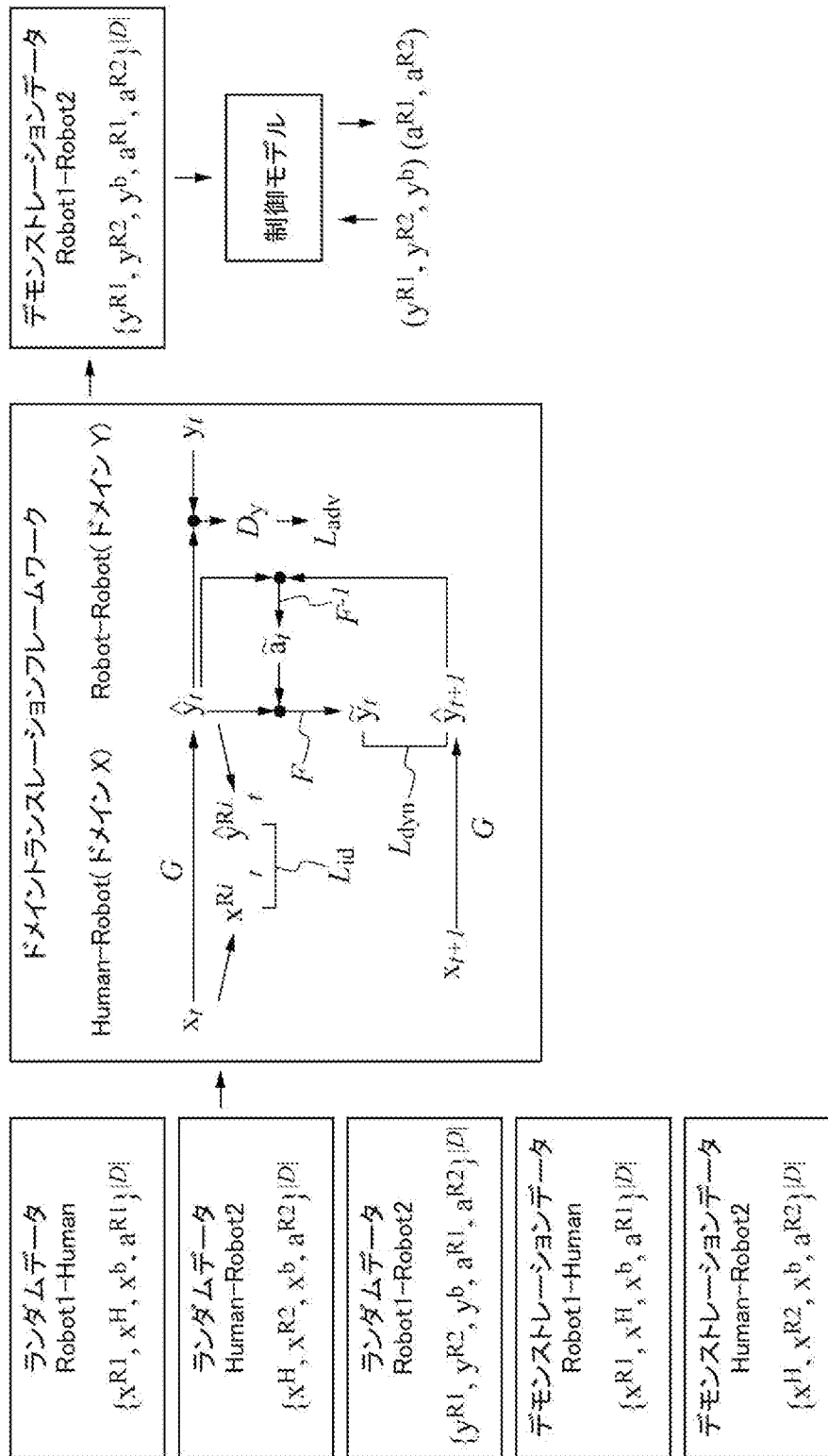
[図1]



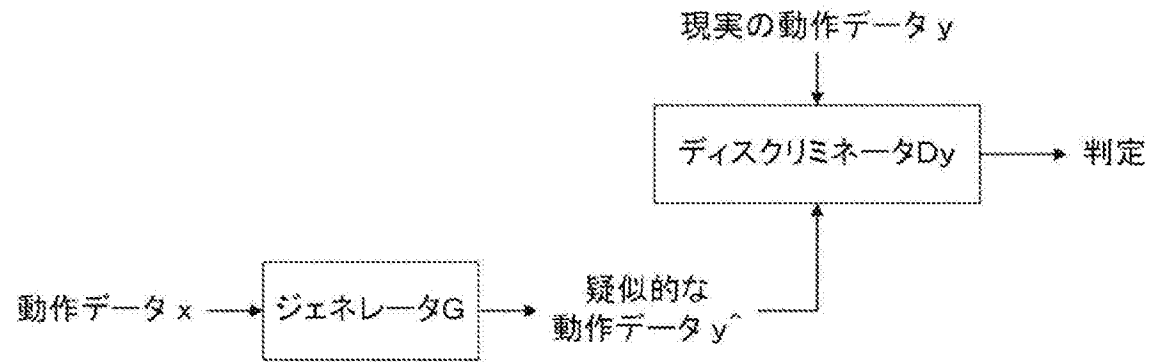
[図2]



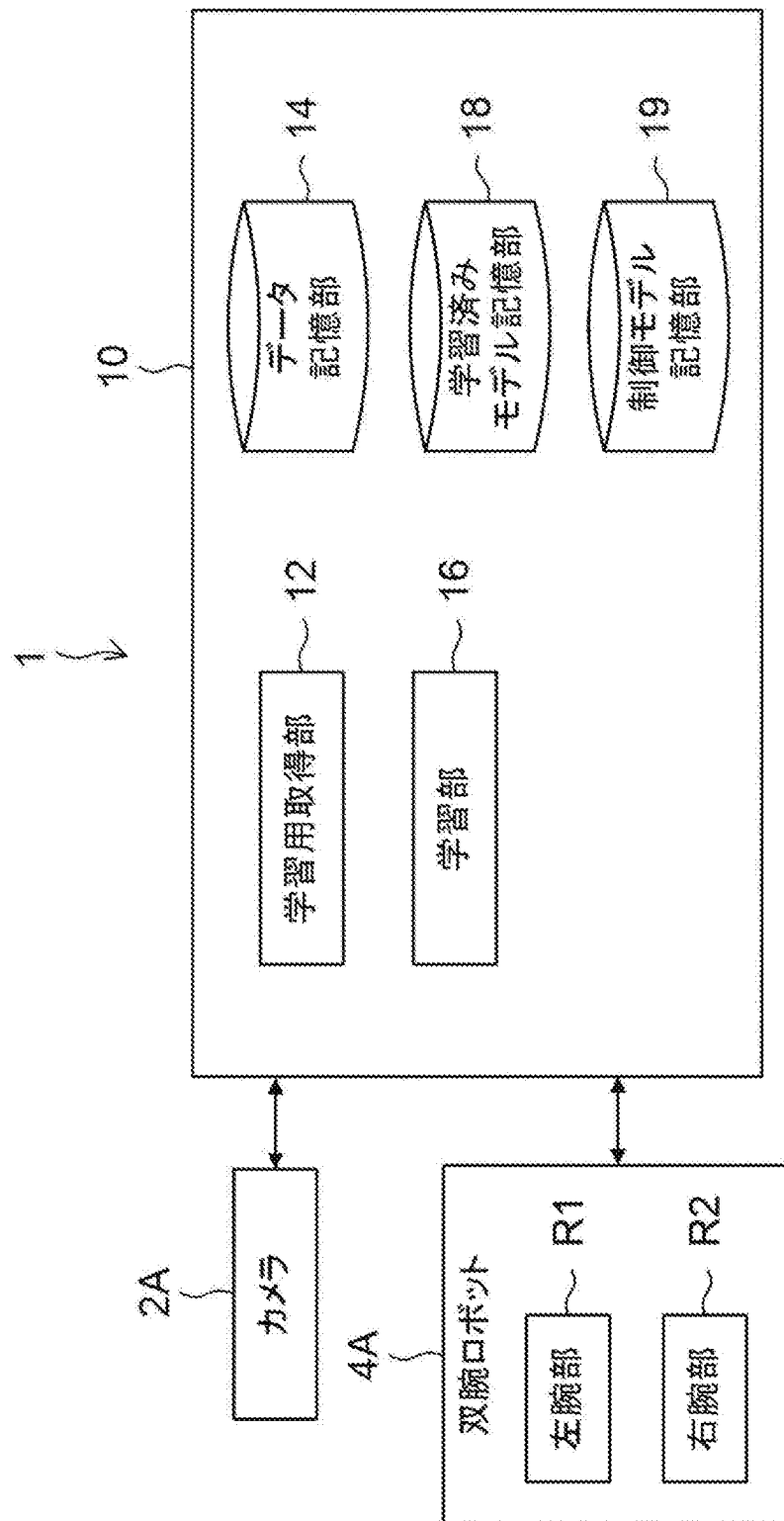
[図3]



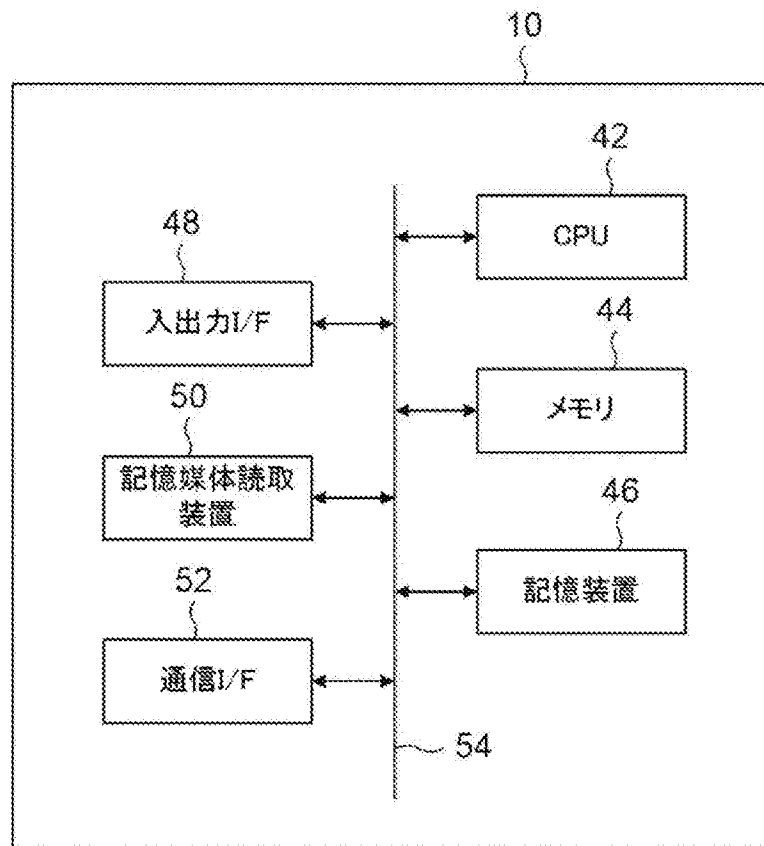
[図4]



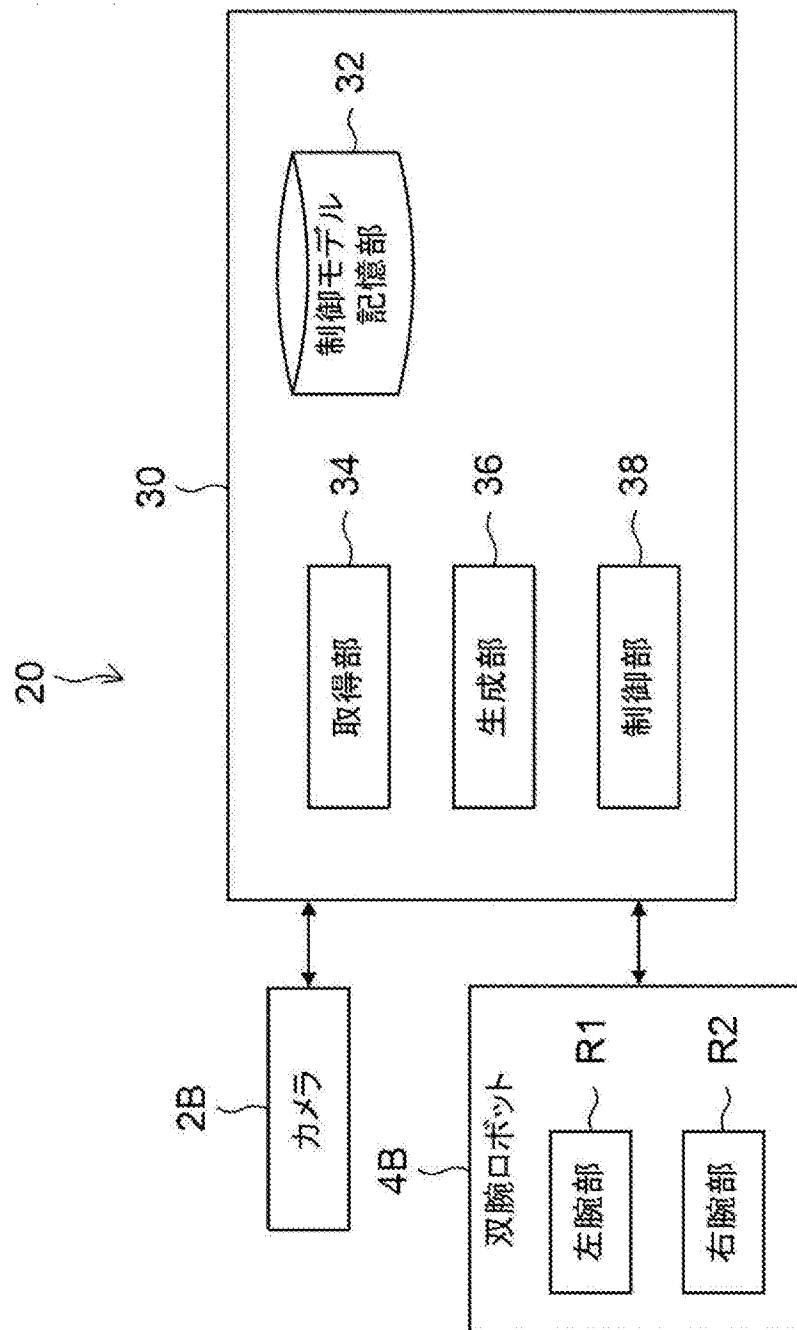
[図5]



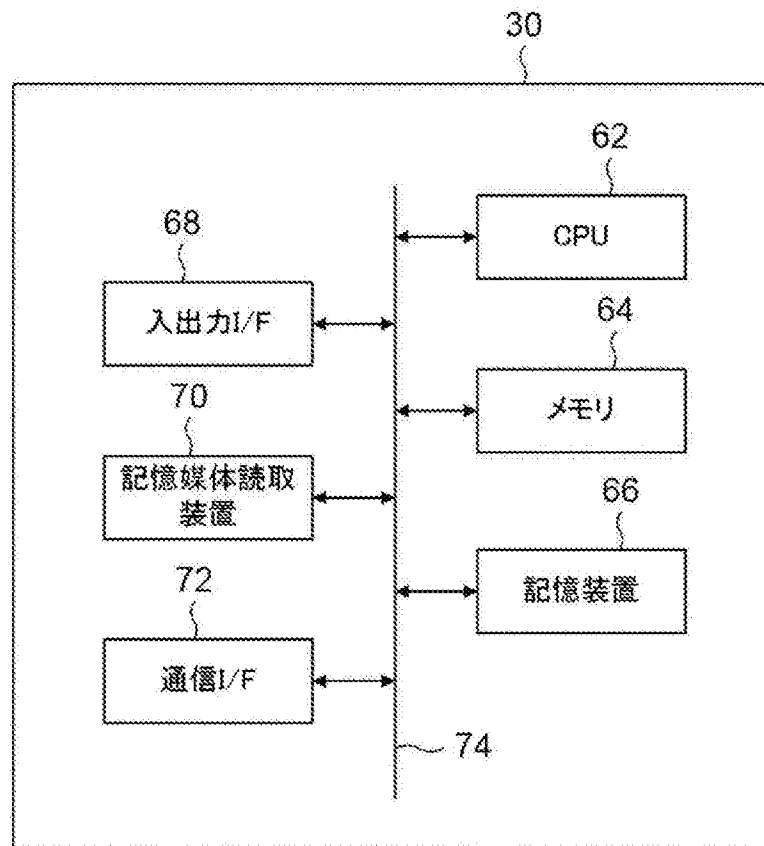
[図6]



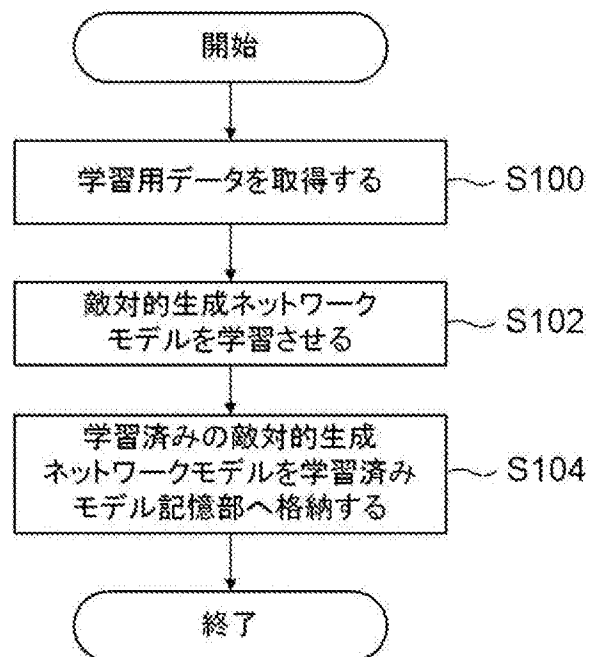
[図7]



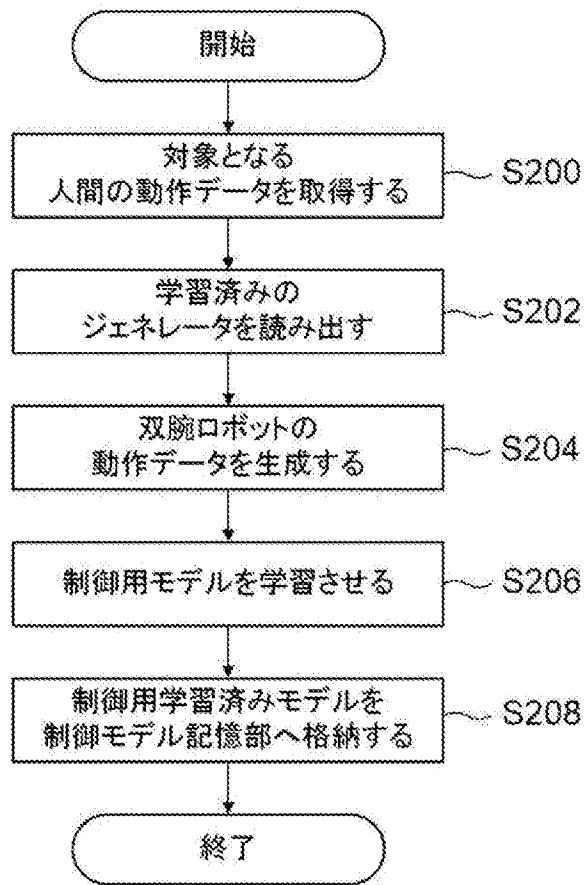
[図8]



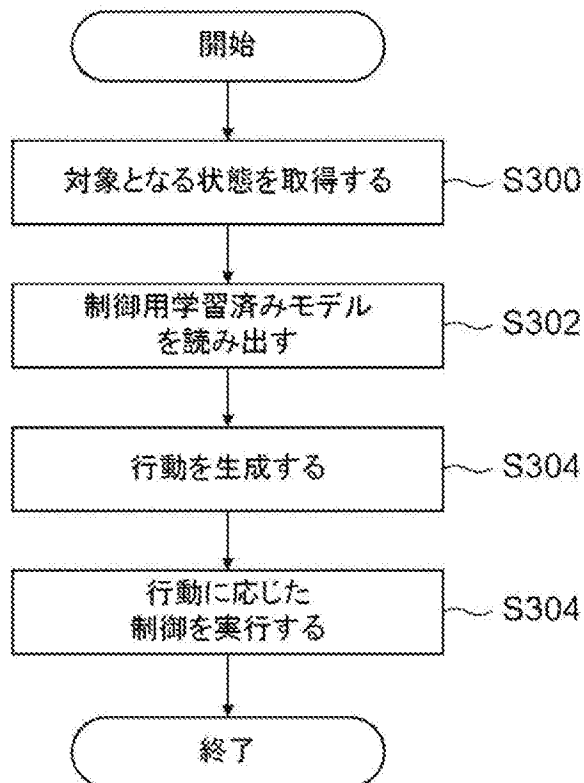
[図9]



[図10]



[図11]

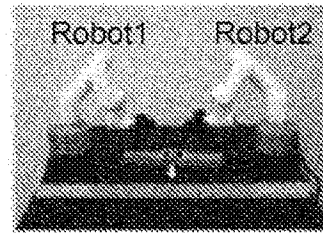
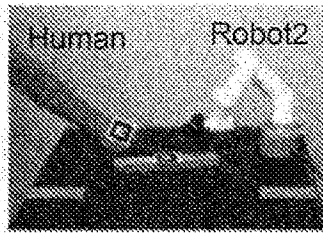
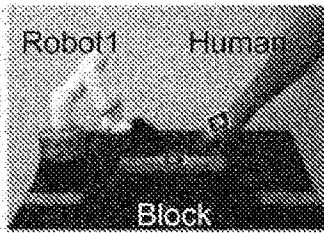


[図12]

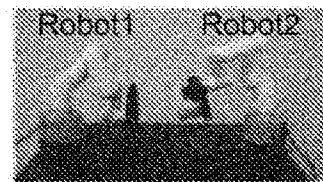
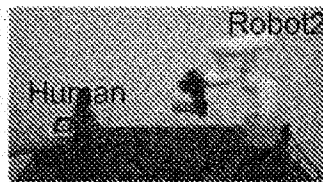
Demonstration

Execution

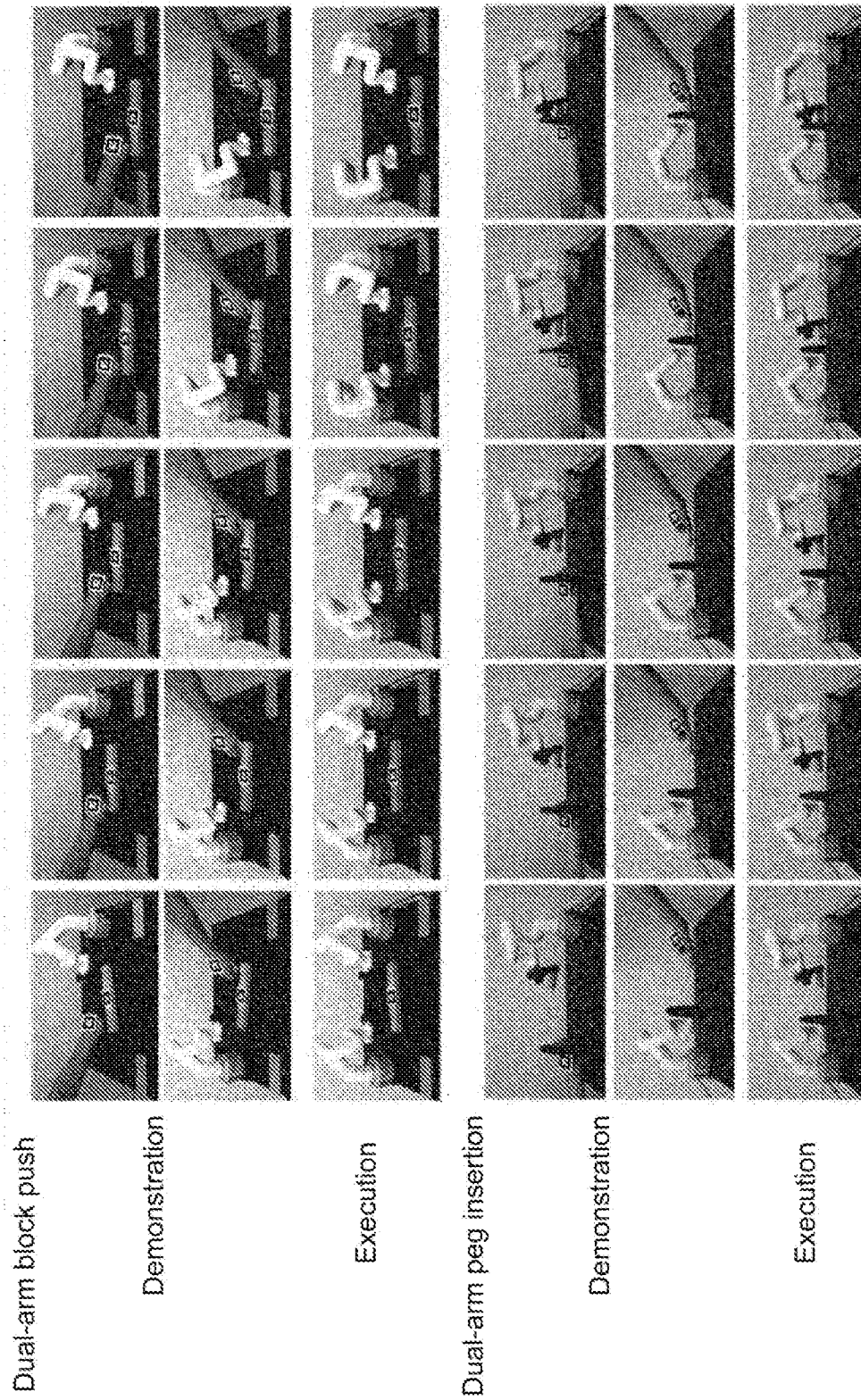
(a) Dual-arm block push



(b) Dual-arm peg insertion



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2024/020908

A. CLASSIFICATION OF SUBJECT MATTER

G06N 3/094(2023.01)i; **B25J 13/00**(2006.01)i
FI: G06N3/094; B25J13/00 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)

G06N3/094; 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 1971-2024
Registered utility model specifications of Japan 1996-2024
Published registered utility model applications of Japan 1994-2024

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.
X	西村 優佑 ほか, 敵対的生成ネットワークを用いた対話中の人間の振る舞いのモデル化, 日本ロボット学会誌, 2019, vol. 37, no. 7, pp. 632-638, 特に, "2. インタラクション動作のモデル化," "4. 対面相手の振る舞いに応じたインタラクション動作の生成," (NISHIMURA, Yusuke et al. Human Behavior Modeling During Dialogue by Using Generative Adversarial Networks. Journal of the Robotics Society of Japan.) non-official translation (in particular, "2. Modeling of interaction motion," "4. Generation of interaction motion according to behavior of adversary")	1, 3-5, 9-10
A		2, 6-8
A	松原 崇充 ほか, 動的方策計画法を用いた敵対的模倣学習とその応用, 日本ロボット学会誌, 15 July 2020, vol. 38, no. 6, pp. 5-8, (MATSUBARA, Takamitsu et al. Generative Adversarial Imitation Learning using Dynamic Policy Programming and Its Applications. Journal of the Robotics Society of Japan.) entire text, all drawings	1-10



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
"D" document cited by the applicant in the international application
"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

07 August 2024

Date of mailing of the international search report

20 August 2024

Name and mailing address of the ISA/JP

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

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC））

G06N 3/094(2023.01)i; B25J 13/00(2006.01)i

FI: G06N3/094; B25J13/00 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））

G06N3/094; 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 - 2024年

日本国実用新案登録公報

1996 - 2024年

日本国登録実用新案公報

1994 - 2024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
X	西村 優佑 ほか，敵対的生成ネットワークを用いた対話中の人間の振る舞いのモデル化，日本ロボット学会誌，2019，Vol.37，No.7，pp.632-638 特に，“2. インタラクション動作のモデル化”及び“4. 対面相手の振る舞いに応じたインタラクション動作の生成”	1, 3-5, 9-10
A		2, 6-8
A	松原 崇充 ほか，動的方策計画法を用いた敵対的模倣学習とその応用，日本ロボット学会誌，2020.07.15，Vol.38，No.6，pp.5-8 全文，全図	1-10

☐ 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、一般的技术水準を示すもの

“D” 国際出願で出願人が先行技術文献として記載した文献

“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

07.08.2024

国際調査報告の発送日

20.08.2024

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）

渡辺 順哉 5H 1157

電話番号 03-3581-1101 内線 3529

様式 PCT/ISA/210（第2ページ）（2022年7月）