

(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
2021年6月3日(03.06.2021)



(10) 国際公開番号

WO 2021/106646 A1

(51) 国際特許分類:

G05B 23/02 (2006.01) G06N 20/00 (2019.01)

目 大通バスセンタービル 1 号館 2 F 東京エレクトロン株式会社内 Hokkaido (JP).

(21) 国際出願番号: PCT/JP2020/042564

(22) 国際出願日: 2020年11月16日(16.11.2020)

(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 2019-217439 2019年11月29日(29.11.2019) JP

(71) 出願人: 東京エレクトロン株式会社 (TOKYO ELECTRON LIMITED) [JP/JP]; 〒1076325 東京都港区赤坂五丁目 3 番 1 号 Tokyo (JP).

(72) 発明者: 筒井 拓郎 (TSUTSUI, Takuro); 〒0600051 北海道札幌市中央区南 1 条東 1 丁

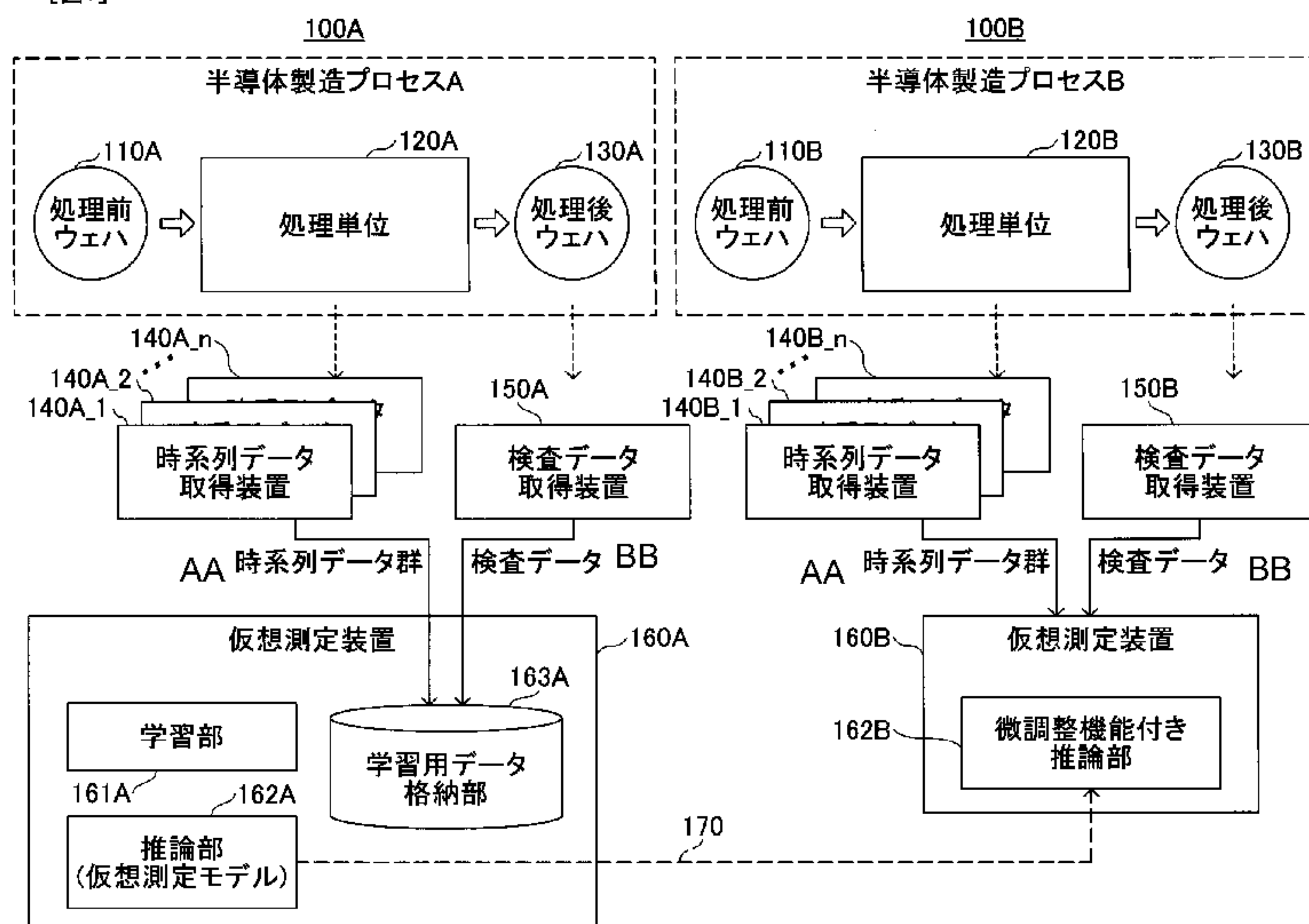
(74) 代理人: 伊東 忠重, 外 (ITOH, Tadashige et al.); 〒1000005 東京都千代田区丸の内二丁目 1 番 1 号 丸の内 M Y P L A Z A (明治安田生命ビル) 1 6 階 Tokyo (JP).

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

(54) Title: INFERENCE DEVICE, INFERENCE METHOD, AND INFERENCE PROGRAM

(54) 発明の名称: 推論装置、推論方法及び推論プログラム

[図1]



100A Semiconductor manufacturing process A
100B Semiconductor manufacturing process B
110A, 110B Wafer before processing
120A, 120B Processing unit
130A, 130B Wafer after processing
140A_1, 140B_1 Time series data acquisition device
150A, 150B Inspection data acquisition device
160A, 160B Virtual measurement device
161A Training unit
162A Inference unit (Virtual measurement model)
162B Fine-adjustment function-mounted inference unit
163A Training data storage unit
AA Time series data group
BB Inspection data

(57) Abstract: Provided are an inference device, an inference method, and an inference program by which inference is possible with high accuracy regardless of an object to be applied. This inference device (160B) comprises: acquisition units (140B_1, 140B_2, ...) which acquire time series data groups measured according to processes for an object (110B) in a prescribed process unit (120B) of a manufacturing process (100B); and an inference unit (162B) which adjusts each piece of output data output by processing the acquired time series data groups by using a plurality of pre-machine-trained network units, and outputs an inference result by synthesizing each piece of the output data adjusted, wherein the inference unit adjusts each piece of the output data by using correction parameters according to errors included in the inference result.

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

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

添付公開書類 :
一 国際調査報告 (条約第21条(3))

(57) 要約：適用対象によらず高精度な推論が可能な推論装置、推論方法及び推論プログラムを提供する。推論装置（1 6 0 B）は、製造プロセス（1 0 0 B）の所定の処理単位（1 2 0 B）において、対象物（1 1 0 B）の処理に伴い測定された時系列データ群を取得する取得部（1 4 0 B__1, 1 4 0 B__2, . . .）と、取得した前記時系列データ群を、予め機械学習された複数のネットワーク部を用いて処理することで出力された各出力データを調整し、調整した後の各出力データを合成することで、推論結果を出力する推論部（1 6 2 B）と、を有し、前記推論部は、前記推論結果に含まれる誤差に応じた補正パラメータを用いて、前記各出力データを調整する。

明 細 書

発明の名称：推論装置、推論方法及び推論プログラム

技術分野

[0001] 本開示は、推論装置、推論方法及び推論プログラムに関する。

背景技術

[0002] 従来より、各種製造プロセスの分野では、対象物の処理中に測定された測定データ（複数種類の時系列データのデータセット、以下、時系列データ群と称す）から、処理後の対象物の状態や処理中のプロセス内の事象等を推論する推論技術が知られている。

[0003] 一例として、半導体製造プロセスにおいては、処理後のウェハの状態を推論する仮想測定技術や、処理中のプロセス内の異常の有無を推論する異常検知技術等が知られている。

[0004] 一方で、これらの推論技術において用いられるモデル（例えば、仮想測定モデル、異常検知モデル）が、より高精度な推論を実現するには、プロセスごとにモデルを生成し、最適化する必要があり、コストと時間がかかる。

[0005] これに対して、特定のプロセスを対象として高精度な推論を実現したモデルを、同種の他のプロセスにも適用することができれば、モデルの最適化にかかるコストと時間を削減することができる。

先行技術文献

特許文献

[0006] 特許文献1：特開2006-163517号公報

発明の概要

発明が解決しようとする課題

[0007] 本開示は、適用対象によらず高精度な推論が可能な推論装置、推論方法及び推論プログラムを提供する。

課題を解決するための手段

[0008] 本開示の一態様による推論装置は、例えば、以下のような構成を有する。

即ち、

製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群を取得する取得部と、

取得した前記時系列データ群を、予め機械学習された複数のネットワーク部を用いて処理し、該複数のネットワーク部より出力された各出力データを調整した後に合成することで、推論結果を出力する推論部と、を有し、

前記推論部は、前記推論結果に含まれる誤差に応じた補正パラメータを用いて、前記各出力データを調整する。

発明の効果

[0009] 本開示によれば、適用対象によらず高精度な推論が可能な推論装置、推論方法及び推論プログラムを提供することができる。

図面の簡単な説明

[0010] [図1]図1は、仮想測定装置が適用されるシステムの全体構成の一例を示す図である。

[図2]図2は、半導体製造プロセスの所定の処理単位の一例を示す第1の図である。

[図3]図3は、半導体製造プロセスの所定の処理単位の一例を示す第2の図である。

[図4]図4は、取得される時系列データ群の一例を示す図である。

[図5]図5は、仮想測定装置のハードウェア構成の一例を示す図である。

[図6]図6は、仮想測定装置の学習部の機能構成の一例を示す図である。

[図7]図7は、分岐部の処理の具体例を示す第1の図である。

[図8]図8は、分岐部の処理の具体例を示す第2の図である。

[図9]図9は、分岐部の処理の具体例を示す第3の図である。

[図10]図10は、各ネットワーク部に含まれる正規化部の処理の具体例を示す図である。

[図11]図11は、分岐部の処理の具体例を示す第4の図である。

[図12]図12は、仮想測定装置の推論部の機能構成の一例を示す図である。

[図13]図 1 3 は、仮想測定装置による仮想測定処理の流れを示すフローチャートである。

[図14]図 1 4 は、仮想測定装置の微調整機能付き推論部の機能構成の一例を示す第 1 の図である。

[図15]図 1 5 は、仮想測定装置による微調整処理の流れを示すフローチャートである。

[図16]図 1 6 は、仮想測定装置の微調整機能付き推論部の機能構成の一例を示す第 2 の図である。

発明を実施するための形態

[0011] 以下、各実施形態について添付の図面を参照しながら説明する。なお、以下の各実施形態では、特定の半導体製造プロセスを対象として、ウェハの処理に伴い測定された時系列データ群を用いて、

- ・ 処理後のウェハの状態を推論する仮想測定モデル、または、
- ・ プロセス内の異常の有無を推論する異常検知モデル、

を生成する場合について説明する。このとき、以下の各実施形態では、時系列データ群を複数のネットワーク部を用いて処理することで、多面的な解析を行い、高精度な推論を実現するモデルを生成する。

[0012] また、以下の各実施形態では、生成したモデルに微調整機能を付加することで、同種の他の半導体製造プロセスに当該モデルを適用した際に、プロセス間の個体差に起因する誤差（推論結果に含まれる誤差）を、該微調整機能を用いて低減させる。

[0013] これにより、以下の各実施形態によれば、適用対象によらず高精度な推論が可能な推論装置、推論方法及び推論プログラムを提供することができる。この結果、他の半導体製造プロセスを対象として新たにモデルを生成し、最適化する場合と比較して、コストと時間を削減することができる。

[0014] なお、以下の各実施形態のうち、第 1 の実施形態では、時系列データ群に基づくモデルとして仮想測定モデルを生成し、微調整機能として補正行列を用いる場合について説明する。また、第 2 の実施形態では、微調整機能とし

て、補正行列に代えてニューラルネットワークを用いる場合について説明する。更に、第3の実施形態では、時系列データ群に基づくモデルとして、仮想測定モデルに代えて異常検知モデルを生成する場合について説明する。

[0015] なお、各実施形態及び添付の図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複した説明を省略する。

[0016] [第1の実施形態]

＜推論装置の適用例＞

はじめに、仮想測定モデルに微調整機能が付加された仮想測定装置（推論装置）の適用例について説明する。図1は、仮想測定装置が適用されるシステムの全体構成の一例を示す図である。

[0017] 図1に示すように、システム100Aは、半導体製造プロセスAと、時系列データ取得装置140A__1～140A__nと、検査データ取得装置150Aと、仮想測定装置160Aとを有する。システム100Aでは、特定のプロセスである半導体製造プロセスAを対象として、高精度な推論を実現する仮想測定モデルが生成される。

[0018] システム100Bは、半導体製造プロセスBと、時系列データ取得装置140B__1～140B__nと、検査データ取得装置150Bと、仮想測定装置160Bとを有する。システム100Bにおいて、半導体製造プロセスBは、半導体製造プロセスAと同種の他のプロセスであり、本実施形態では、システム100Aにおいて生成された仮想測定モデルに微調整機能が付加された仮想測定装置（推論装置）が適用される適用対象である。

[0019] システム100Aにおいて、半導体製造プロセスAは、所定の処理単位120Aにおいて、対象物（処理前ウェハ110A）を処理し、結果物（処理後ウェハ130A）を生成する。なお、ここでいう処理単位120Aとは、抽象的な概念であり、詳細は後述する。また、処理前ウェハ110Aとは、処理単位120Aにおいて処理される前のウェハ（基板）を指し、処理後ウェハ130Aとは、処理単位120Aにおいて処理された後のウェハ（基板

)を指す。

[0020] また、システム100Aにおいて、時系列データ取得装置140A__1～140A__nは、それぞれ、処理前ウェハ110Aの処理に伴い時系列データを測定する。時系列データ取得装置140A__1～140A__nは、互いに異なる種類の測定項目について測定を行うものとする。なお、時系列データ取得装置140A__1～140A__nそれぞれが測定する測定項目の数は1つであっても、複数であってもよい。また、処理前ウェハ110Aの処理に伴い測定した時系列データには、処理前ウェハ110Aの処理中に測定した時系列データのほか、処理前ウェハ110Aの処理の前後に行われる前処理、後処理の際に測定した時系列データも含まれる。これらの処理にはウェハ（基板）が無い状態で行われる前処理、後処理が含まれていてもよい。

[0021] 時系列データ取得装置140A__1～140A__nにより測定された時系列データ群は、学習用データ（入力データ）として、仮想測定装置160Aの学習用データ格納部163Aに格納される。

[0022] また、システム100Aにおいて、検査データ取得装置150Aは、処理単位120Aにおいて処理された処理後ウェハ130Aの所定の検査項目（例えば、ER（Etch Rate））を検査し、検査データを取得する。検査データ取得装置150Aにより取得された検査データは、学習用データ（正解データ）として、仮想測定装置160Aの学習用データ格納部163Aに格納される。

[0023] また、システム100Aにおいて、仮想測定装置160Aには、学習プログラム及び推論プログラムを含む仮想測定プログラムがインストールされている。仮想測定プログラムが実行されることで、仮想測定装置160Aは、学習部161A及び推論部162Aとして機能する。

[0024] 学習部161Aは、時系列データ取得装置140A__1～140A__nにより測定された時系列データ群と、検査データ取得装置150Aにより取得された検査データとを用いて機械学習を行う。

[0025] 具体的には、学習部161Aが有する複数のネットワーク部を用いて時系

列データ群を処理し、複数のネットワーク部より出力された各出力データの合成結果が、検査データに近づくよう、当該複数のネットワーク部について機械学習を行う。

[0026] 推論部 162A は、新たな対象物（処理前ウェハ）の処理に伴い測定された時系列データ群を取得し、機械学習が行われた複数のネットワーク部に入力する。これにより、推論部 162A は、新たな処理前ウェハの処理に伴い取得された時系列データに基づき、処理後ウェハの検査データを推論し、推論結果（仮想測定データ）を出力する。

[0027] このように、対象物の処理に伴い測定された時系列データ群を、複数のネットワーク部を用いて処理することで、仮想測定装置 160A によれば、時系列データ群を多面的に解析することが可能となる。この結果、時系列データ群を 1 のネットワーク部を用いて処理する場合と比較して、高精度な推論を実現する仮想測定モデル（推論部 162A）を生成することができる。

[0028] 一方、システム 100B において、半導体製造プロセス B は、システム 100A の半導体製造プロセス A と同種のプロセスである。また、システム 100B において、時系列データ取得装置 140B__1 ～ 140B__n、検査データ取得装置 150B は、それぞれ、システム 100A の時系列データ取得装置 140A__1 ～ 140A__n、検査データ取得装置 150A と対応している。

[0029] 更に、システム 100B において、仮想測定装置 160B（推論装置）は、システム 100A の仮想測定装置 160A と対応している。ただし、システム 100B の仮想測定装置 160B の場合、学習部 161A を有していない。また、推論部 162A に代えて、微調整機能付き推論部 162B を有している（学習プログラムを含まず、仮想測定装置 160A にインストールされた推論プログラムとは同様の推論プログラムを含む仮想測定プログラムがインストールされている）。

[0030] システム 100B の仮想測定装置 160B の場合、新たに仮想測定モデルを生成し、時系列データ群を用いて機械学習を行うことで最適化するのは

なく、システム 100A の仮想測定装置 160A において生成された仮想測定モデル（推論部 162A）を適用する。

[0031] ここで、半導体製造プロセス A と半導体製造プロセス B とは、上述したように同種のプロセスであるが、個体差を有している。このため、仮想測定装置 160A において生成された仮想測定モデル（推論部 162A）をそのまま適用しても、推論結果（仮想測定データ）には誤差が含まれる。

[0032] そこで、仮想測定装置 160B（推論装置）の場合、仮想測定装置 160A において生成された仮想測定モデル（推論部 162A）に微調整機能を付加した推論部を生成する。図 1 において、仮想測定装置 160B が有する微調整機能付き推論部 162B は、仮想測定装置 160A において生成された仮想測定モデル（推論部 162A）に微調整機能を付加した推論部の一例である。

[0033] 微調整機能付き推論部 162B は、仮想測定装置 160A において生成された仮想測定モデル（推論部 162A）が適用される一方で（破線 170 参照）、個体差に起因する誤差（推論結果に含まれる誤差）を低減する微調整機能が付加されている。

[0034] 具体的には、微調整機能付き推論部 162B は、

- ・ 生成された仮想測定モデルに含まれる複数のネットワーク部を用いて、時系列データ群を処理し、複数のネットワーク部より出力された各出力データを調整した後に合成することで出力される推論結果（仮想測定データ）と、
- ・ 検査データ取得装置 150B により取得される検査データと、

の誤差が低減されるように、補正パラメータ（各出力データを調整する際に用いる補正行列に含まれるパラメータ。詳細は後述）を更新する。

[0035] これにより、仮想測定装置 160B では、

- ・ 仮想測定装置 160A において生成された、高精度な推論を実現する仮想測定モデル（推論部 162A）が適用されたモデルであって、
- ・ 適用対象である半導体製造プロセス B においても、高精度な推論が可能なモデル、

を実現することができる。

[0036] <半導体製造プロセスの所定の処理単位>

次に、半導体製造プロセス A、B の所定の処理単位 1 2 0 A、1 2 0 B について説明する。図 2 は、半導体製造プロセスの所定の処理単位の一例を示す第 1 の図である。図 2 に示すように、基板処理装置の一例である半導体製造装置 2 0 0 は、複数のチャンバ（複数の処理空間の一例。図 2 の例では、“チャンバ A”～“チャンバ C”）を有しており、各チャンバにおいて、ウェハが処理される。

[0037] このうち、図 2 の 2 a は、複数のチャンバを処理単位 1 2 0 A、1 2 0 B と定義した場合を示している。この場合、処理前ウェハ 1 1 0 A、1 1 0 B とは、チャンバ A において処理される前のウェハを指し、処理後ウェハ 1 3 0 A、1 3 0 B とは、チャンバ C において処理された後のウェハを指す。

[0038] また、図 2 の 2 a の処理単位 1 2 0 A、1 2 0 B において、処理前ウェハ 1 1 0 A、1 1 0 B の処理に伴い測定される時系列データ群には、

- ・チャンバ A（第 1 の処理空間）における処理に伴い測定される時系列データ群と、
- ・チャンバ B（第 2 の処理空間）における処理に伴い測定される時系列データ群と、
- ・チャンバ C（第 3 の処理空間）における処理に伴い測定される時系列データ群と、

が含まれる。

[0039] 一方、図 2 の 2 b は、1 のチャンバ（図 2 の 2 b の例では、“チャンバ B”）を処理単位 1 2 0 A、1 2 0 B と定義した場合を示している。この場合、処理前ウェハ 1 1 0 A、1 1 0 B とは、チャンバ B において処理される前のウェハ（チャンバ A において処理された後のウェハ）を指す。また、処理後ウェハ 1 3 0 A、1 3 0 B とは、チャンバ B において処理された後のウェハ（チャンバ C において処理される前のウェハ）を指す。

[0040] また、図 2 の 2 b の処理単位 1 2 0 A、1 2 0 B において、処理前ウェハ

110A、110Bの処理に伴い測定される時系列データ群には、チャンバBにおいて、処理前ウェハ110A、110Bの処理に伴い測定される時系列データ群が含まれる。

[0041] 図3は、半導体製造プロセスの所定の処理単位の一例を示す第2の図である。図2と同様、半導体製造装置200は、複数のチャンバを有しており、各チャンバにおいて、ウェハが処理される。

[0042] このうち、図3の3aは、チャンバBにおける処理内容のうち、前処理と後処理とを除いた処理（“ウェハ処理”と称す）を処理単位120A、120Bと定義した場合を示している。この場合、処理前ウェハ110A、110Bとは、ウェハ処理が行われる前のウェハ（前処理が行われた後のウェハ）を指し、処理後ウェハ130A、130Bとは、ウェハ処理が行われた後のウェハ（後処理が行われる前のウェハ）を指す。

[0043] また、図3の3aの処理単位120A、120Bにおいて、処理前ウェハ110A、110Bの処理に伴い測定される時系列データ群には、チャンバBにおいて、処理前ウェハ110A、110Bのウェハ処理に伴い測定される時系列データ群が含まれる。

[0044] なお、図3の3aの例では、同一チャンバ内（チャンバB内）において、前処理、ウェハ処理（本処理）、後処理が行われる場合の、ウェハ処理を処理単位120A、120Bとした場合について示した。しかしながら、異なるチャンバで各処理が行われる場合（例えば、チャンバA内において前処理が、チャンバB内においてウェハ処理が、チャンバC内において後処理が行われる場合）にあっては、チャンバごとの各処理を処理単位120A、120Bとしてもよい。

[0045] 一方、図3の3bは、チャンバBにおける処理内容のうち、ウェハ処理に含まれる1のレシピ（図3の3bの例では“レシピ111”）の処理を処理単位120A、120Bと定義した場合を示している。この場合、処理前ウェハ110A、110Bとは、レシピ111の処理が行われる前のウェハ（レシピ111の処理が行われた後のウェハ）を指す。また、処理後ウェハ13

0 A、1 3 0 Bとは、レシピ111の処理が行われた後のウェハ（レシピ1 V（不図示）の処理が行われる前のウェハ）を指す。

[0046] また、図3の3 aの処理単位1 2 0 A、1 2 0 Bにおいて、処理前ウェハ1 1 0 A、1 1 0 Bの処理に伴い測定される時系列データ群には、チャンバBにおいて、レシピ111によるウェハ処理に伴い測定される時系列データ群が含まれる。

[0047] <時系列データ群の具体例>

次に、時系列データ取得装置1 4 0 A__1～1 4 0 A__n、1 4 0 B__1～1 4 0 B__nにおいて取得される時系列データ群の具体例について説明する。図4は、取得される時系列データ群の一例を示す図である。なお、図4の例では、説明の簡略化のため、時系列データ取得装置1 4 0 A__1～1 4 0 A__n、1 4 0 B__1～1 4 0 B__nがそれぞれ1次元のデータを測定するものとしている。ただし、1の時系列データ取得装置が2次元のデータ（複数種類の1次元データのデータセット）を測定してもよい。

[0048] このうち、図4の4 aは、処理単位1 2 0 A、1 2 0 Bが、図2の2 b、図3の3 a、図3の3 bのいずれかで定義された場合の時系列データ群を表している。この場合、時系列データ取得装置1 4 0 A__1～1 4 0 A__n、1 4 0 B__1～1 4 0 B__nは、それぞれ、チャンバBにおける処理に伴い測定した時系列データを取得する。また、時系列データ取得装置1 4 0 A__1～1 4 0 A__nは、互いに、同一時間帯に測定した時系列データを、時系列データ群として取得する。同様に、時系列データ取得装置1 4 0 B__1～1 4 0 B__nは、互いに、同一時間帯に測定した時系列データを、時系列データ群として取得する。

[0049] 一方、図4の4 bは、処理単位1 2 0 A、1 2 0 Bが、図2の2 aで定義された場合の時系列データ群を表している。この場合、時系列データ取得装置1 4 0 A__1～1 4 0 A__3、1 4 0 B__1～1 4 0 B__3は、例えば、チャンバAにおけるウェハの処理に伴い測定した時系列データ群1を取得する。また、時系列データ取得装置1 4 0 A__n-2、1 4 0 B__n-2は、

例えば、チャンバBにおける当該ウェハの処理に伴い測定した時系列データ群2を取得する。また、時系列データ取得装置140A__ $n-1$ ～140A__ n 、140B__ $n-1$ ～140B__ n は、例えば、チャンバCにおける当該ウェハの処理に伴い測定した時系列データ群3を取得する。

[0050] なお、図4の4aにおいて、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、チャンバBにおいて処理前ウェハの処理に伴い測定した、同一の時間範囲の時系列データを、時系列データ群として取得する場合について示した。しかしながら、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、チャンバBにおいて処理前ウェハの処理に伴い測定した、異なる時間範囲の時系列データを時系列データ群として取得してもよい。

[0051] 具体的には、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、前処理を実行中に測定した複数の時系列データを、時系列データ群1として取得してもよい。また、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、ウェハ処理を実行中に測定した複数の時系列データを、時系列データ群2として取得してもよい。更に、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、後処理を実行中に測定した複数の時系列データを、時系列データ群3として取得してもよい。

[0052] 同様に、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、レシピIを実行中に測定した複数の時系列データを、時系列データ群1として取得してもよい。また、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、レシピIIを実行中に測定した複数の時系列データを、時系列データ群2として取得してもよい。更に、時系列データ取得装置140A__1～140A__ n 、140B__1～140B__ n は、レシピIIIを実行中に測定した複数の時系列データを、時系列データ群3として取得してもよい。

[0053] <仮想測定装置のハードウェア構成>

次に、仮想測定装置 160A、160B のハードウェア構成について説明する。図 5 は、仮想測定装置のハードウェア構成の一例を示す図である。図 5 に示すように、仮想測定装置 160A、160B は、CPU (Central Processing Unit) 501、ROM (Read Only Memory) 502、RAM (Random Access Memory) 503 を有する。また、仮想測定装置 160 は、GPU (Graphics Processing Unit) 504 を有する。なお、CPU 501、GPU 504 などのプロセッサ (処理回路、Processing Circuit、Processing Circuitry) と、ROM 502、RAM 503 などのメモリは、いわゆるコンピュータを形成する。

[0054] 更に、仮想測定装置 160 は、補助記憶装置 505、表示装置 506、操作装置 507、I/F (Interface) 装置 508、ドライブ装置 509 を有する。なお、仮想測定装置 160 の各ハードウェアは、バス 510 を介して相互に接続される。

[0055] CPU 501 は、補助記憶装置 505 にインストールされた各種プログラム (例えば、仮想測定プログラム等) を実行する演算デバイスである。

[0056] ROM 502 は、不揮発性メモリであり、主記憶装置として機能する。ROM 502 は、補助記憶装置 505 にインストールされた各種プログラムを CPU 501 が実行するために必要な各種プログラム、データ等を格納する。具体的には、ROM 502 は BIOS (Basic Input/Output System) や EFI (Extensible Firmware Interface) 等のブートプログラム等を格納する。

[0057] RAM 503 は、DRAM (Dynamic Random Access Memory) や SRAM (Static Random Access Memory) 等の揮発性メモリであり、主記憶装置として機能する。RAM 503 は、補助記憶装置 505 にインストールされた各種プログラムが CPU 501 によって実行される際に展開される、作業領域を提供する。

[0058] GPU 504 は、画像処理用の演算デバイスであり、CPU 501 により仮想測定プログラムが実行される際に、各種画像データ (本実施形態では、

時系列データ群) について、並列処理による高速演算を行う。なお、GPU 504は、内部メモリ (GPUメモリ) を搭載しており、各種画像データについて並列処理を行う際に必要な情報を一時的に保持する。

[0059] 補助記憶装置505は、各種プログラムや、各種プログラムがCPU 501によって実行される際に用いられる各種データ等を格納する。

[0060] 表示装置506は、仮想測定装置160A、160Bの内部状態を表示する表示デバイスである。操作装置507は、仮想測定装置160A、160Bの管理者が仮想測定装置160A、160Bに対して各種指示を入力する際に用いる入力デバイスである。I/F装置508は、不図示のネットワークと接続し、通信を行うための接続デバイスである。

[0061] ドライブ装置509は記録媒体520をセットするためのデバイスである。ここでいう記録媒体520には、CD-ROM、フレキシブルディスク、光磁気ディスク等のように情報を光学的、電気的あるいは磁氣的に記録する媒体が含まれる。また、記録媒体520には、ROM、フラッシュメモリ等のように情報を電気的に記録する半導体メモリ等が含まれていてもよい。

[0062] なお、補助記憶装置505にインストールされる各種プログラムは、例えば、配布された記録媒体520がドライブ装置509にセットされ、該記録媒体520に記録された各種プログラムがドライブ装置509により読み出されることでインストールされる。あるいは、補助記憶装置505にインストールされる各種プログラムは、ネットワークを介してダウンロードされることで、インストールされてもよい。

[0063] <学習部の機能構成>

次に、システム100Aにおける、仮想測定装置160Aの学習部161Aの機能構成について説明する。図6は、仮想測定装置の学習部の機能構成の一例を示す図である。学習部161Aは、分岐部610と、第1のネットワーク部620__1～第Mのネットワーク部620__Mと、連結部630と、比較部640とを有する。

[0064] 分岐部610は、学習用データ格納部163Aより時系列データ群を読み

出す。また、分岐部610は、読み出した時系列データ群が、第1のネットワーク部620__1から第Mのネットワーク部620__Mまでの複数のネットワーク部を用いて処理されるよう、当該時系列データ群を処理する。

[0065] 第1のネットワーク部620__1～第Mのネットワーク部620__Mは、畳み込みニューラルネットワーク（CNN：Convolutional Neural Network）をベースに構成されており、複数の層を有する。

[0066] 具体的には、第1のネットワーク部620__1は、第1層620__11～第N層620__1Nを有する。同様に、第2のネットワーク部620__2は、第1層620__21～第N層620__2Nを有する。以下、同様の構成を有しており、第Mのネットワーク部620__Mは、第1層620__M1～第N層620__MNを有する。

[0067] 第1のネットワーク部620__1の第1層620__11～第N層620__1Nの各層では、正規化処理や、畳み込み処理、活性化処理、プーリング処理等の各種処理が行われる。また、第2のネットワーク部620__2～第Mのネットワーク部620__Mの各層でも同様の各種処理が行われる。

[0068] 連結部630は、第1のネットワーク部620__1の第N層620__1Nから出力された出力データから、第Mのネットワーク部620__Mの第N層620__MNから出力された出力データまでの各出力データを合成し、合成結果を比較部640に出力する。

[0069] 比較部640は、連結部630より出力された合成結果と、学習用データ格納部163Aより読み出した検査データ（正解データ）とを比較し、誤差を算出する。学習部161Aでは、比較部640により算出された誤差が所定の条件を満たすように、第1のネットワーク部620__1～第Mのネットワーク部620__M及び連結部630について機械学習を行う。

[0070] これにより、第1のネットワーク部620__1～第Mのネットワーク部620__Mの第1層～第N層それぞれのモデルパラメータ及び連結部630のモデルパラメータが最適化される。

[0071] ＜学習部の各部の処理の詳細＞

次に、システム 1 0 0 A における、仮想測定装置 1 6 0 A の学習部 1 6 1 A の各部（ここでは、特に、分岐部 6 1 0）の処理の詳細について、具体例を挙げて説明する。

[0072] （１）分岐部の処理の詳細 1

図 7 は、分岐部の処理の具体例を示す第 1 の図である。図 7 の場合、分岐部 6 1 0 は、時系列データ取得装置 1 4 0 A__1 ～ 1 4 0 A__n により測定された時系列データ群を、第 1 の基準に応じて加工することで、時系列データ群 1（第 1 の時系列データ群）を生成し、第 1 のネットワーク部 6 2 0__1 に入力する。

[0073] また、分岐部 6 1 0 は、時系列データ取得装置 1 4 0 A__1 ～ 1 4 0 A__n により測定された時系列データ群を、第 2 の基準に応じて加工することで、時系列データ群 2（第 2 の時系列データ群）を生成し、第 2 のネットワーク部 6 2 0__2 に入力する。

[0074] このように、時系列データ群を、異なる基準に応じて加工し、それぞれ異なるネットワーク部に分けて処理する構成としたうえで機械学習を行うことで、時系列データ群を多面的に解析することが可能となる。この結果、時系列データ群を 1 のネットワーク部に入力して機械学習を行う場合と比較して、高精度な推論を実現する仮想測定モデル（推論部 1 6 2 A）を生成することが可能となる。

[0075] なお、図 7 の例では、2 種類の基準に応じて時系列データ群を加工することで、2 種類の時系列データ群を生成する場合について示したが、3 種類以上の基準に応じて時系列データ群を加工することで、3 種類以上の時系列データ群を生成してもよい。

[0076] （２）分岐部による処理の詳細 2

次に、分岐部 6 1 0 の他の処理の詳細について説明する。図 8 は、分岐部の処理の具体例を示す第 2 の図である。図 8 の場合、分岐部 6 1 0 は、時系列データ取得装置 1 4 0 A__1 ～ 1 4 0 A__n により測定された時系列データ群を、データ種類に応じてグループ分けする。これにより、分岐部 6 1 0

は、時系列データ群 1（第 1 の時系列データ群）と時系列データ群 2（第 2 の時系列データ群）とを生成する。また、分岐部 6 1 0 は、生成した時系列データ群 1 を第 3 のネットワーク部 6 2 0 __ 3 に入力し、生成した時系列データ群 2 を第 4 のネットワーク部 6 2 0 __ 4 に入力する。

[0077] このように、時系列データ群を、データ種類に応じて複数のグループに分け、異なるネットワーク部を用いて処理する構成としたうえで機械学習を行うことで、時系列データ群を多面的に解析することが可能となる。この結果、時系列データ群を 1 のネットワーク部に入力して機械学習を行う場合と比較して、高精度な推論を実現する仮想測定モデル（推論部 1 6 2 A）を生成することが可能となる。

[0078] なお、図 8 の例では、時系列データ取得装置 1 4 0 A __ 1 ~ 1 4 0 A __ n の違いに基づくデータ種類の違いに応じて、時系列データ群をグループ分けしたが、データが取得された時間範囲に応じて、時系列データ群をグループ分けしてもよい。例えば、時系列データ群が複数のレシピによる処理に伴い測定された時系列データ群であった場合には、レシピごとの時間範囲に応じて、時系列データ群をグループ分けしてもよい。

[0079] （3）分岐部による処理の詳細 3

次に、分岐部 6 1 0 による他の処理の詳細について説明する。図 9 は、分岐部の処理の具体例を示す第 3 の図である。図 9 の場合、分岐部 6 1 0 は、時系列データ取得装置 1 4 0 A __ 1 ~ 1 4 0 A __ n により取得された時系列データ群を、第 5 のネットワーク部 6 2 0 __ 5 と第 6 のネットワーク部 6 2 0 __ 6 の両方に入力する。そして、第 5 のネットワーク部 6 2 0 __ 5 と第 6 のネットワーク部 6 2 0 __ 6 とで、同じ時系列データ群に対して、異なる処理（正規化処理）を施す。

[0080] 図 1 0 は、各ネットワーク部に含まれる正規化部の処理の具体例を示す図である。図 1 0 に示すように、第 5 のネットワーク部 6 2 0 __ 5 の各層には、正規化部と、畳み込み部と、活性化関数部と、プーリング部とが含まれる。

[0081] 図10の例は、第5のネットワーク部620__5に含まれる各層のうち、第1層620__51に、正規化部1001と、畳み込み部1002と、活性化関数部1003と、プーリング部1004とが含まれることを示している。

[0082] このうち、正規化部1001では、分岐部610により入力された時系列データ群に対して第1の正規化処理を行い、正規化時系列データ群1（第1の時系列データ群）を生成する。

[0083] 同様に、図10の例は、第6のネットワーク部620__6に含まれる各層のうち、第1層620__61には、正規化部1011と、畳み込み部1012と、活性化関数部1013と、プーリング部1014とが含まれることを示している。

[0084] このうち、正規化部1011では、分岐部610により入力された時系列データ群に対して第2の正規化処理を行い、正規化時系列データ群2（第2の時系列データ群）を生成する。

[0085] このように、異なる手法で正規化処理を行う正規化部をそれぞれに含む複数のネットワーク部を用いて時系列データ群を処理する構成としたうえで機械学習を行うことで、時系列データ群を多面的に解析することが可能となる。この結果、時系列データ群を1の正規化処理を行う1のネットワーク部に入力して機械学習を行う場合と比較して、高精度な推論を実現する仮想測定モデル（推論部162A）を生成することが可能となる。

[0086] （4）分岐部による処理の詳細4

次に、分岐部610による他の処理の詳細について説明する。図11は、分岐部の処理の具体例を示す第4の図である。図11の場合、分岐部610は、時系列データ取得装置140A__1～140A__nにより測定された時系列データ群のうち、チャンバAにおける処理に伴い測定された時系列データ群1（第1の時系列データ群）を第7のネットワーク部620__7に入力する。

[0087] また、分岐部610は、時系列データ取得装置140A__1～140A__

nにより測定された時系列データ群のうち、チャンバBにおける処理に伴い測定された時系列データ群2（第2の時系列データ群）を第8のネットワーク部620__8に入力する。

[0088] このように、異なるチャンバ（第1の処理空間、第2の処理空間）における処理に伴い測定されたそれぞれの時系列データ群を、異なるネットワーク部を用いて処理する構成としたうえで機械学習を行うことで、時系列データ群を多面的に解析することが可能となる。この結果、それぞれの時系列データ群を1のネットワーク部に入力して機械学習を行う場合と比較して、高精度な推論を実現する仮想測定モデル（推論部162A）を生成することが可能となる。

[0089] ＜仮想測定装置の推論部の機能構成＞

次に、システム100Aにおける、仮想測定装置160Aの推論部162Aの機能構成について説明する。図12は、仮想測定装置の推論部の機能構成の一例を示す図である。図12に示すように、仮想測定装置160Aの推論部162Aは、分岐部1210と、第1のネットワーク部1220__1から第Mのネットワーク部1220__Mと、連結部1230とを有する。

[0090] 分岐部1210は、時系列データ取得装置140A__1～140A__Nにより新たに測定された時系列データ群を取得する。また、分岐部1210は、取得した時系列データ群が、第1のネットワーク部1220__1～第Mのネットワーク部1220__Mを用いて処理されるよう制御する。

[0091] 第1のネットワーク部1220__1～第Mのネットワーク部1220__Mは、学習部161Aにより機械学習が行われ、第1のネットワーク部620__1～第Mのネットワーク部620__Mの各層のモデルパラメータが最適化されることで形成される。

[0092] 連結部1230は、学習部161Aにより機械学習が行われ、モデルパラメータが最適化された連結部630により形成される。連結部1230は、第1のネットワーク部1220__1の第N層1220__1Nから出力された出力データから、第Mのネットワーク部1220__Mの第N層1220__M

Nから出力された出力データまでの各出力データを合成し、仮想測定データを出力する。

[0093] <仮想測定処理の流れ>

次に、システム100Aにおける、仮想測定装置160Aによる仮想測定処理全体の流れについて説明する。図13は、仮想測定装置による仮想測定処理の流れを示すフローチャートである。

[0094] ステップS1301において、学習部161Aは、学習用データとして時系列データ群及び検査データを取得する。

[0095] ステップS1302において、学習部161Aは、取得した学習用データのうち、時系列データ群を入力データ、検査データを正解データとして機械学習を行う。

[0096] ステップS1303において、学習部161Aは、機械学習を継続するかどうかを判定する。更なる学習用データを取得して機械学習を継続する場合には（ステップS1303においてYESの場合には）、ステップS1301に戻る。一方、機械学習を終了する場合には（ステップS1303においてNOの場合には）、ステップS1304に進む。

[0097] ステップS1304において、推論部162Aは、機械学習により最適化されたモデルパラメータを反映することで、第1のネットワーク部1220__1～第Mのネットワーク部1220__Mを生成する。

[0098] ステップS1305において、推論部162Aは、新たな処理前ウェハ110Aの処理に伴い測定された時系列データ群を入力し、仮想測定データを推論する。

[0099] ステップS1306において、推論部162Aは、推論した仮想測定データを出力する。

[0100] <仮想測定装置の微調整機能付き推論部の機能構成>

次に、システム100Bにおける、仮想測定装置160Bの微調整機能付き推論部162Bの機能構成について説明する。図14は、仮想測定装置の微調整機能付き推論部の機能構成の一例を示す図である。

- [0101] 図14に示すように、仮想測定装置160Bの微調整機能付き推論部162Bは、取得部として機能する分岐部1210を有する。また、仮想測定装置160Bの微調整機能付き推論部162Bは、推論部として機能する、第1のネットワーク部1220__1から第Mのネットワーク部1220__M、連結部1410、個体調整部1420、微調整部1430、比較部1440を有する。
- [0102] このうち、分岐部1210は、推論部162Aの分岐部1210と同じであり、図12を用いて説明済みであるため、ここでは説明を省略する。また、第1のネットワーク部1220__1から第Mのネットワーク部1220__Mも、推論部162Aの第1のネットワーク部1220__1から第Mのネットワーク部1220__Mと同じである。
- [0103] 具体的には、第1のネットワーク部1220__1～第Mのネットワーク部1220__Mは、学習部161Aにより機械学習が行われ、第1のネットワーク部620__1～第Mのネットワーク部620__Mの各層のモデルパラメータが最適化されることで形成される。
- [0104] 連結部1410は、学習部161Aにより機械学習が行われ、モデルパラメータが最適化された連結部630により形成される。ただし、連結部1410の場合、第1のネットワーク部1220__1の第N層1220__1Nから出力された出力データから、第Mのネットワーク部1220__Mの第N層1220__MNから出力された出力データまでの各出力データを合成することなく出力する。
- [0105] 個体調整部1420は、連結部1410より出力された各出力データに、半導体製造プロセスAの処理単位120Aと、半導体製造プロセスBの処理単位120Bとの個体差に応じた係数（“個体感度”と称す）をかけ合わせる。
- [0106] 微調整部1430は、個体調整部1420により個体感度がかけ合わされた各出力データに、補正行列をかけ合わせることで、スカラ量である仮想測定データを算出する。

[0107] 比較部 1440 は、微調整部 1430 により出力された仮想測定データを取得するとともに、処理後ウェハ 130B についての検査データを取得する。また、比較部 1440 は、取得した仮想測定データと検査データとの差分を算出し、微調整部 1430 に通知する。

[0108] このように、微調整機能付き推論部 162B では、半導体製造プロセス B において、所定の期間、処理後ウェハ 130B についての検査データに基づいて、微調整部 1430 が補正パラメータ ($P_1 \sim P_M$) を更新する。そして、微調整機能付き推論部 162B の微調整部 430 では、仮想測定データと検査データとの差分が所定の閾値以下となるまで、補正パラメータ ($P_1 \sim P_M$) の更新を継続する。

[0109] これにより、微調整部 1430 では、半導体製造プロセス A の処理単位 120A と、半導体製造プロセス B の処理単位 120B との個体差に起因する誤差（推論結果に含まれる誤差）を低減させることが可能になる。

[0110] なお、微調整機能付き推論部 162B の場合、半導体製造プロセス B において測定された時系列データ群を追加データとして、仮想測定モデルを再学習することで最適化する場合と比較して、コストと時間を削減することができる。

[0111] <微調整処理の流れ>

次に、システム 100B における、仮想測定装置 160B による微調整処理の流れについて説明する。図 15 は、仮想測定装置による微調整処理の流れを示すフローチャートである。

[0112] ステップ S1501 において、微調整機能付き推論部 162B の分岐部 1210 は、半導体製造プロセス B の処理単位 120B において、新たな処理前ウェハ 110B の処理に伴い測定された時系列データ群を取得する。また、微調整機能付き推論部 162B の第 1 ～ 第 M のネットワーク部 1220__1 ～ 1220__M は、取得された時系列データ群を処理する。これにより、第 1 ～ 第 M のネットワーク部 1220__1 ～ 1220__M の最終層から各出力データが出力される。

- [0113] ステップS 1 5 0 2において、微調整機能付き推論部 1 6 2 Bの個体調整部 1 4 2 0は、第1～第Mのネットワーク部 1 2 2 0__1～1 2 2 0__Mの最終層より出力された各出力データに、個体感度をかけ合わせることで、各出力データを調整する。
- [0114] ステップS 1 5 0 3において、微調整機能付き推論部 1 6 2 Bの微調整部 1 4 3 0は、個体感度がかけあわされた各出力データに、補正行列をかけ合わせることで、仮想測定データを算出する。
- [0115] ステップS 1 5 0 4において、微調整機能付き推論部 1 6 2 Bは、処理後ウェハ 1 3 0 Bについての検査データを取得し、比較部 1 4 4 0に通知する。また、比較部 1 4 4 0は、微調整部 1 4 3 0より出力された仮想測定データと、通知された検査データとを比較し、差分（推論結果に含まれる誤差）を算出する。
- [0116] ステップS 1 5 0 5において、微調整機能付き推論部 1 6 2 Bの比較部 1 4 4 0は、比較結果に基づいて、差分が所定の閾値以下か否かを判定することで、補正パラメータの更新が必要か否かを判定する。
- [0117] ステップS 1 5 0 5において、差分が所定の閾値を超えており、補正パラメータの更新が必要と判定した場合には（ステップS 1 5 0 5においてY E Sの場合には）、ステップS 1 5 0 6に進む。
- [0118] ステップS 1 5 0 6において、微調整機能付き推論部 1 6 2 Bの微調整部 1 4 3 0は、比較部 1 4 4 0により算出された差分（推論結果に含まれる誤差）に応じて、補正行列の補正パラメータ（ $P_1 \sim P_M$ ）を更新する。その後、ステップS 1 5 0 7に進む。
- [0119] 一方、ステップS 1 5 0 5において、差分が所定の閾値以下であり、補正パラメータの更新が必要でないと判定した場合には（ステップS 1 5 0 5においてN Oの場合には）、直接、ステップS 1 5 0 7に進む。
- [0120] ステップS 1 5 0 7において、微調整機能付き推論部 1 6 2 Bは、微調整処理を終了するか否かを判定する。ステップS 1 5 0 7において、微調整処理を終了しないと判定した場合には（ステップS 1 5 0 7においてN Oの場

合には)、ステップS 1 5 0 1に戻る。

[0121] 一方、ステップS 1 5 0 7において、微調整処理を終了すると判定した場合には(ステップS 1 5 0 7においてY E Sの場合には)、微調整処理を終了する。

[0122] <まとめ>

以上の説明から明らかなように、仮想測定装置1 6 0 Aは、

- ・製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群を取得する。

- ・取得した時系列データ群を複数のネットワーク部を用いて処理することで各ネットワーク部より出力される各出力データの合成結果が、対象物を処理することで得られる結果物の検査データに近づくように、各ネットワーク部について機械学習を行う。

[0123] このように、時系列データ群を複数のネットワーク部を用いて処理することで、多面的な解析を行うことが可能となる。この結果、仮想測定装置1 6 0 Aでは、高精度な推論を実現する仮想測定モデルを生成することが可能となる。

[0124] また、仮想測定装置1 6 0 B(推論装置)は、

- ・他の製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群を、生成された仮想測定モデルに含まれる複数のネットワーク部を用いて処理し、各出力データを出力する。

- ・出力した各出力データを補正パラメータを用いて微調整した後に合成することで、仮想測定データを推論する。

- ・推論した仮想測定データに含まれる誤差に応じて補正パラメータを更新する。

[0125] このように、製造プロセスの所定の処理単位において、時系列データ群を用いて生成された仮想測定モデルを、他の製造プロセスに適用する際、仮想測定装置1 6 0 Bでは、複数のネットワーク部より出力される各出力データを微調整する機能を付加する。

[0126] これにより、他の製造プロセスに仮想測定モデルを適用する際に、プロセス間の個体差に起因する誤差（推論結果に含まれる誤差）を低減させることが可能になる。つまり、第1の実施形態によれば、適用対象によらず高精度な推論が可能な推論装置、推論方法及び推論プログラムを提供することができる。

[0127] [第2の実施形態]

上記第1の実施形態では、各ネットワーク部の最終層から出力された各出力データを、個体感度及び補正行列を用いて微調整するものとして説明した。しかしながら、微調整機能付き推論部による各出力データの微調整の方法はこれに限定されず、例えば、微調整用のネットワーク部を用いて、各出力データを微調整してもよい。

[0128] 図16は、仮想測定装置の微調整機能付き推論部の機能構成の一例を示す第2の図である。図14との相違点は、図16に示す微調整機能付き推論部1600Bの場合、微調整ネットワーク部1610を有する点である。

[0129] 微調整ネットワーク部1610は、畳み込みニューラルネットワークをベースに構成され、連結部1410より出力された各出力データを入力することで、仮想測定データを出力する。

[0130] また、微調整ネットワーク部1610は、仮想測定データを出力したことに応じて、比較部1440から通知される差分に基づいて、微調整ネットワーク部1610のモデルパラメータである補正パラメータを更新する。

[0131] このように、微調整機能付き推論部1600Bでは、半導体製造プロセスBにおいて、所定の期間、処理後ウェハ130Bについての検査データに基づいて、微調整ネットワーク部1610が補正パラメータを更新する。なお、このとき、第1のネットワーク部1220__1から第Mのネットワーク部1220__Mのモデルパラメータは固定した状態を維持するものとする。そして、微調整機能付き推論部1600Bの微調整ネットワーク部1610では、仮想測定データと検査データとの差分が所定の閾値以下になるまで、補正パラメータの更新を継続する。

[0132] これにより、微調整ネットワーク部 1 6 1 0 では、半導体製造プロセス A の処理単位 1 2 0 A と、半導体製造プロセス B の処理単位 1 2 0 B との個体差に起因する誤差（推論結果に含まれる誤差）を低減させることが可能になる。

[0133] なお、微調整機能付き推論部 1 6 0 0 B の場合、新たに仮想測定モデルを生成し、半導体製造プロセス B において測定された時系列データ群を用いて最適化する場合と比較して、オーバフィッティングの可能性を低減させることができる。

[0134] [第 3 の実施形態]

上記第 1 及び第 2 の実施形態では、仮想測定装置 1 6 0 A が生成した仮想測定モデルを他の半導体製造プロセス B に適用する場合について説明したが、他の半導体製造プロセス B に適用されるモデルは、仮想測定モデルに限定されない。

[0135] 第 3 の実施形態では、第 1 及び第 2 の実施形態において説明した仮想測定装置 1 6 0 A 及び 1 6 0 B を、異常検知装置 1 6 0 A 及び 1 6 0 B と読み替え、異常検知装置 1 6 0 A が生成した異常検知モデルを他の半導体製造プロセス B に適用する場合について説明する。

[0136] 異常検知装置 1 6 0 A の場合、学習部 1 6 1 A は、時系列データ群を入力データ、イベント（異常の有無を示す情報）を正解データとして、異常検知モデル（推論部 1 6 2 A）について機械学習を行う。異常検知モデル（推論部 1 6 2 A）は、仮想測定モデル（推論部 1 6 2 A）と同様の構成を有しているものとし、機械学習に用いる学習用データだけが異なっているものとする。

[0137] なお、異常検知装置 1 6 0 A の場合、機械学習に用いる時系列データ群を出力する時系列データ取得装置 1 4 0 A__1 ～ 1 4 0 A__n には、例えば、
・時系列データ群である O E S（Optical Emission Spectrometry）データを出力する、発光分光分析装置、
・時系列データ群である温度データ、圧力データ等のプロセスデータを出力

する、プロセスデータ取得装置、

- ・時系列データであるRFデータを出力する、プラズマ用高周波電源装置、等が含まれる。

[0138] また、異常検知装置160B（推論装置）の場合、微調整機能付き推論部1600Bは、時系列データ群を入力し、異常の有無を示す情報を推論する。

[0139] なお、異常検知装置160Bの場合、推論に用いる時系列データ群を出力する時系列データ取得装置140A__1～140A__nには、例えば、

- ・時系列データ群であるOES（Optical Emission Spectrometry）データを出力する、発光分光分析装置、
- ・時系列データ群である温度データ、圧力データ等のプロセスデータを出力する、プロセスデータ取得装置、
- ・時系列データであるRFデータを出力する、プラズマ用高周波電源装置、等が含まれる。

[0140] ＜まとめ＞

以上の説明から明らかなように、異常検知装置160Aは、

- ・製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群（OESデータ、プロセスデータ、RFデータ）を取得する。

- ・取得した時系列データ群を複数のネットワーク部を用いて処理することで各ネットワーク部より出力された各出力データの合成結果が、対象物の処理に伴い発生したイベント（異常の有無を示す情報）に近づくように、各ネットワーク部について機械学習を行う。

[0141] このように、時系列データ群を複数のネットワーク部を用いて処理することで、多面的な解析を行うことが可能となる。この結果、異常検知装置160Aでは、高精度な推論を実現する異常検知モデルを生成することが可能となる。

[0142] また、異常検知装置160B（推論装置）は、

- ・他の製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群（O E S データ、プロセスデータ、R F データ）を、生成された異常検知モデルに含まれる複数のネットワーク部を用いて処理し、各出力データを出力する。
- ・出力した各出力データを補正パラメータを用いて微調整した後に合成することで、異常の有無を示す情報を推論する。
- ・推論した異常の有無を示す情報に含まれる誤差に応じて補正パラメータを更新する。

[0143] このように、製造プロセスの所定の処理単位において、時系列データ群を用いて生成された異常検知モデルを、他の製造プロセスに適用する際、異常検知装置 1 6 0 B では、複数のネットワーク部より出力される各出力データを微調整する機能を付加する。

[0144] これにより、他の製造プロセスに異常検知モデルを適用する際に、プロセス間の個体差に起因する誤差（推論結果に含まれる誤差）を低減させることが可能になる。つまり、第 3 の実施形態によれば、適用対象によらず高精度な推論が可能な推論装置、推論方法及び推論プログラムを提供することができる。

[0145] 〔その他の実施形態〕

上記第 1 及び第 2 の実施形態では、各出力データの微調整の方法として、個体感度及び補正行列、または、微調整用のネットワーク部を用いる場合について説明した。しかしながら、各出力データの微調整の方法は、これに限定されず、例えば、一般化線形混合モデルや、ガウス過程回帰分析、カルマンフィルタ等を用いてもよい。

[0146] また、上記第 3 の実施形態では、異常検知装置が、対象物の処理に伴い発光分光分析装置、プロセスデータ取得装置、プラズマ用高周波電源装置より出力された O E S データ、プロセスデータ、R F データを取得するものとして説明した。しかしながら、異常検知装置が取得するデータの組み合わせは、これに限定されず、いずれか 1 のデータを取得してもよいし、いずれか 2

のデータの組み合わせを取得してもよい。

[0147] また、上記各実施形態では、微調整機能付き推論部 1 6 2 B、1 6 0 0 B が、第 1 ～第 M のネットワーク部 1 2 2 0 __ 1 ～1 2 2 0 __ M を有するものとして説明した。しかしながら、微調整機能付き推論部 1 6 2 B、1 6 0 0 B は、第 1 ～第 M のネットワーク部 1 2 2 0 __ 1 ～1 2 2 0 __ M を全て有している必要はなく、少なくとも 2 以上のいずれかのネットワーク部を有しているものとする。

[0148] また、上記各実施形態では、学習部 1 6 1 A の各ネットワーク部の機械学習アルゴリズムを、畳み込みニューラルネットワークをベースに構成するものとして説明した。しかしながら、学習部 1 6 1 A の各ネットワーク部の機械学習アルゴリズムは、畳み込みニューラルネットワークに限定されず、他の機械学習アルゴリズムをベースに構成してもよい。

[0149] また、上記各実施形態では、仮想測定装置または異常検知装置 1 6 0 A が学習部 1 6 1 A 及び推論部 1 6 2 A として機能するものとして説明した。しかしながら、学習部 1 6 1 A として機能する装置と、推論部 1 6 2 A として機能する装置とは一体である必要はなく、別体により構成されてもよい。つまり、仮想測定装置または異常検知装置 1 6 0 A は、推論部 1 6 2 A を有していない学習部 1 6 1 A として機能させてもよいし、学習部 1 6 1 A を有していない推論部 1 6 2 A として機能させてもよい。

[0150] また、上記各実施形態では、システム 1 0 0 A において生成された仮想測定モデル（または異常検知モデル）に微調整機能が付加された仮想測定装置（または異常検知装置）を、システム 1 0 0 B に適用するものとして説明した。しかしながら、微調整機能が付加された仮想測定装置（または異常検知装置）が適用される適用対象は、他のシステムに限定されず、自システムであってもよい。

[0151] 例えば、プロセスレシピの一部を変更した場合等、変更の程度が小さい場合において、自システムで生成した仮想測定モデル（または異常検知モデル）に、微調整機能を付加して適用してもよい。

[0152] あるいは、自システム内の装置において、部品交換などの保守作業を行った場合や、自システム内の装置のパーツ消耗等により装置内の環境が変化した場合等、自システムで生成した仮想測定モデル（または異常検知モデル）の精度が低下した際に適用してもよい。

[0153] なお、上記実施形態に挙げた構成等に、その他の要素との組み合わせ等、ここで示した構成に本発明が限定されるものではない。これらの点に関しては、本発明の趣旨を逸脱しない範囲で変更することが可能であり、その応用形態に応じて適切に定めることができる。

[0154] 本出願は、2019年11月29日に出願された日本国特許出願第2019-217439号に基づきその優先権を主張するものであり、同日本国特許出願の全内容を参照することにより本願に援用する。

符号の説明

[0155]	100A、100B	: システム
	110A、110B	: 処理前ウェハ
	120A、120B	: 処理単位
	130A、130B	: 処理後ウェハ
	140A__1～140A__n	: 時系列データ取得装置
	140B__1～140B__n	: 時系列データ取得装置
	150A、150B	: 検査データ取得装置
	160A、160B	: 仮想測定装置
	161A	: 学習部
	162A	: 推論部
	162B	: 微調整機能付き推論部
	200	: 半導体製造装置
	610	: 分岐部
	620__1	: 第1のネットワーク部
	620__11～620__1N	: 第1層～第N層
	620__2	: 第2のネットワーク部

6 2 0 __ 2 1 ~ 6 2 0 __ 2 N	: 第 1 層 ~ 第 N 層
6 2 0 __ M	: 第 M のネットワーク部
6 2 0 __ M 1 ~ 6 2 0 __ M N	: 第 1 層 ~ 第 N 層
6 3 0	: 連結部
6 4 0	: 比較部
1 0 0 1、1 0 1 1	: 正規化部
1 0 0 4、1 0 1 4	: プーリング部
1 2 1 0	: 分岐部
1 2 2 0 __ 1	: 第 1 のネットワーク部
1 2 2 0 __ 1 1 ~ 1 2 2 0 __ 1 N	: 第 1 層 ~ 第 N 層
1 2 2 0 __ 2	: 第 2 のネットワーク部
1 2 2 0 __ 2 1 ~ 1 2 2 0 __ 2 N	: 第 1 層 ~ 第 N 層
1 2 2 0 __ M	: 第 M のネットワーク部
1 2 2 0 __ M 1 ~ 1 2 2 0 __ M N	: 第 1 層 ~ 第 N 層
1 2 4 0	: 連結部
1 4 1 0	: 連結部
1 4 2 0	: 個体調整部
1 4 3 0	: 微調整部
1 4 4 0	: 比較部
1 6 0 0 B	: 微調整機能付き推論部
1 6 1 0	: 微調整ネットワーク部

請求の範囲

- [請求項1] 製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群を取得する取得部と、
- 取得した前記時系列データ群を、予め機械学習された複数のネットワーク部を用いて処理することで出力された各出力データを調整し、調整した後の各出力データを合成することで、推論結果を出力する推論部と、を有し、
- 前記推論部は、前記推論結果に含まれる誤差に応じた補正パラメータを用いて、前記各出力データを調整する、推論装置。
- [請求項2] 前記推論部は、前記複数のネットワーク部のモデルパラメータを固定した状態で、前記推論結果に含まれる誤差が低減されるように、前記補正パラメータを更新する、請求項1に記載の推論装置。
- [請求項3] 前記推論部は、
- 取得した前記時系列データ群を第1の基準及び第2の基準に応じてそれぞれ処理することで、第1の時系列データ群と第2の時系列データ群とを生成し、生成した前記第1の時系列データ群及び前記第2の時系列データ群を、前記複数のネットワーク部を用いて処理する、請求項2に記載の推論装置。
- [請求項4] 前記推論部は、
- 取得した前記時系列データ群をデータ種類または時間範囲に応じてグループ分けし、各グループを、前記複数のネットワーク部を用いて処理する、請求項2に記載の推論装置。
- [請求項5] 前記推論部は、
- 取得した前記時系列データ群を、異なる手法で正規化を行う正規化部をそれぞれに含む前記複数のネットワーク部に入力し、前記複数のネットワーク部を用いて処理する、請求項2に記載の推論装置。
- [請求項6] 前記推論部は、
- 取得した前記時系列データ群を、前記所定の処理単位の第1の処理

空間における前記対象物の処理に伴い測定された第 1 の時系列データ群と、第 2 の処理空間における前記対象物の処理に伴い測定された第 2 の時系列データ群とに分け、前記第 1 の時系列データ群及び前記第 2 の時系列データ群を、前記複数のネットワーク部を用いて処理する、請求項 2 に記載の推論装置。

[請求項7] 前記時系列データ群は、基板処理装置における処理に伴い測定されたデータである、請求項 1 に記載の推論装置。

[請求項8] 製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群を取得する取得工程と、

取得した前記時系列データ群を、予め機械学習された複数のネットワーク部を用いて処理することで出力された各出力データを調整し、調整した後の各出力データを合成することで、推論結果を出力する推論工程と、を有し、

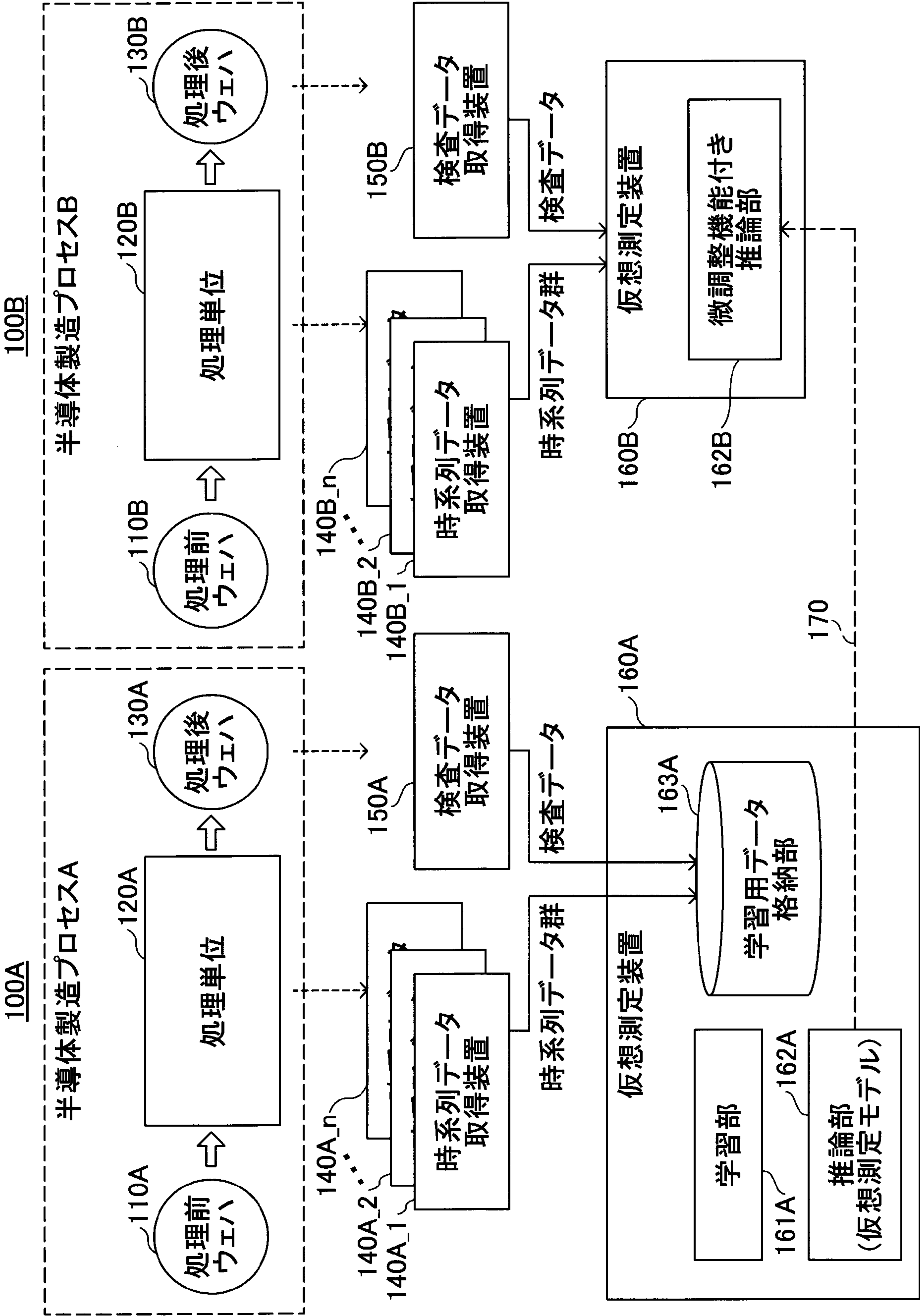
前記推論工程は、前記推論結果に含まれる誤差に応じた補正パラメータを用いて、前記各出力データを調整する、推論方法。

[請求項9] 製造プロセスの所定の処理単位において、対象物の処理に伴い測定された時系列データ群を取得する取得工程と、

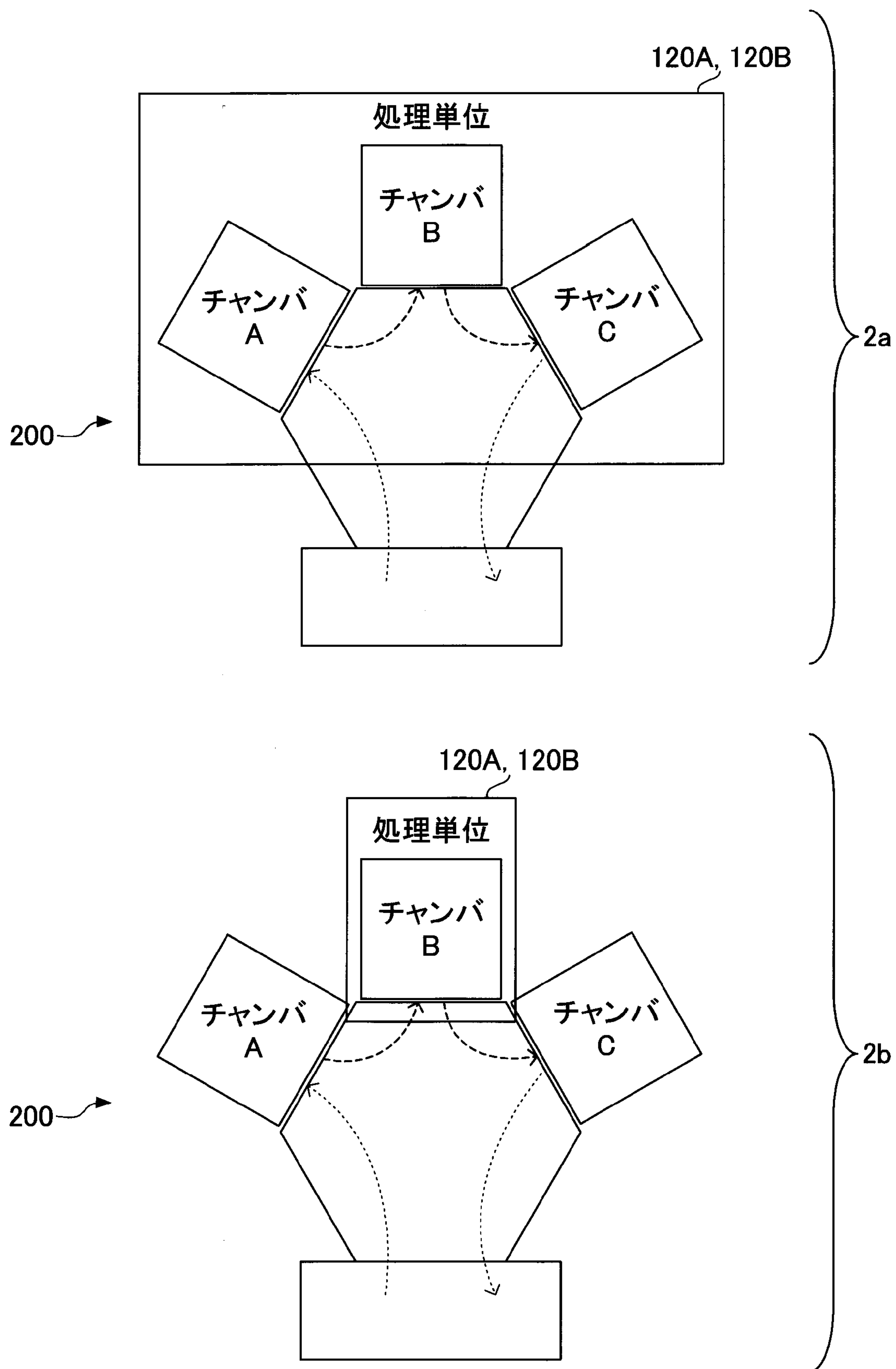
取得した前記時系列データ群を、予め機械学習された複数のネットワーク部を用いて処理することで出力された各出力データを調整し、調整した後の各出力データを合成することで、推論結果を出力する推論工程と、をコンピュータに実行させるための推論プログラムであって、

前記推論工程は、前記推論結果に含まれる誤差に応じた補正パラメータを用いて、前記各出力データを調整する、推論プログラム。

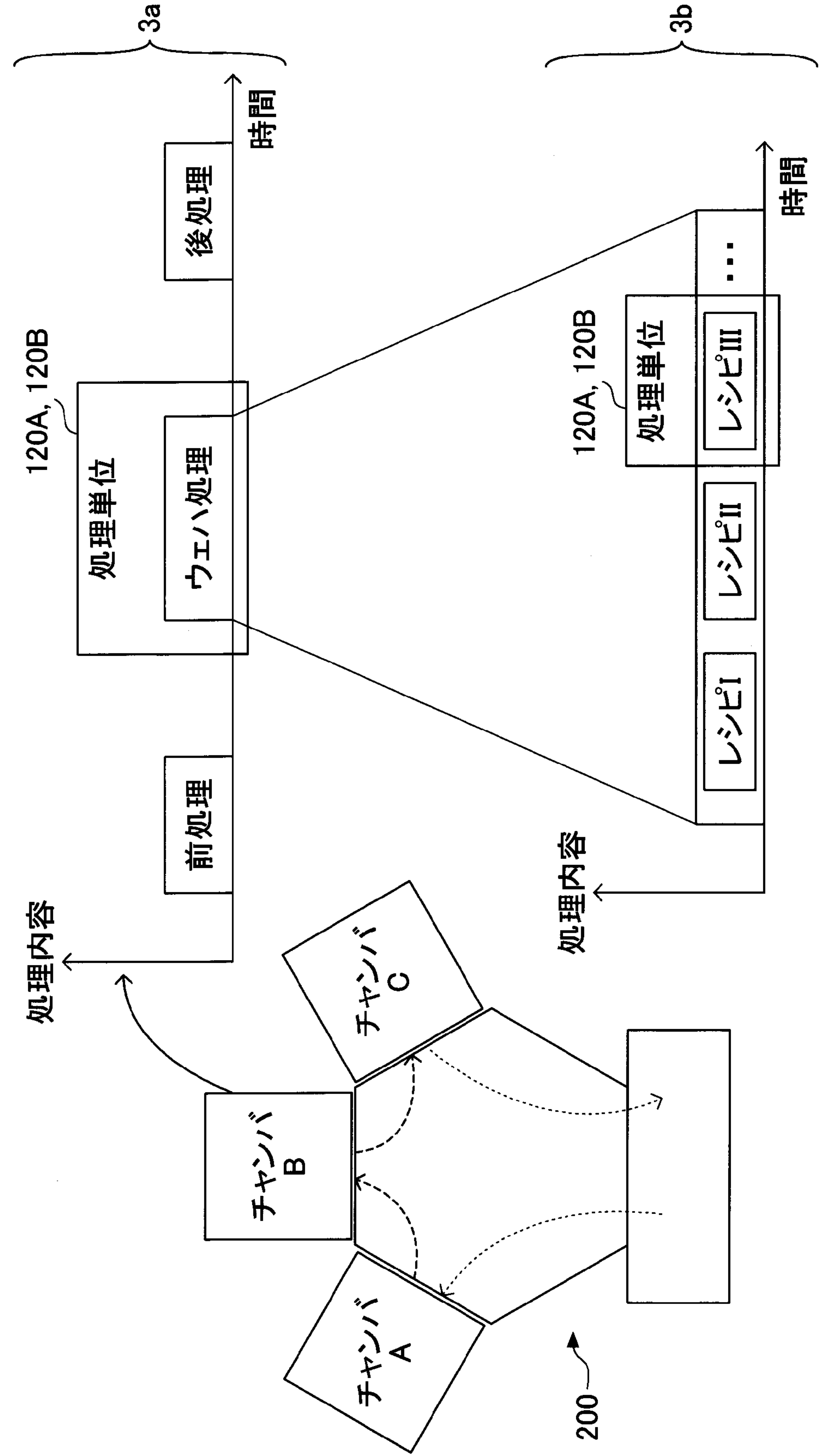
[図1]



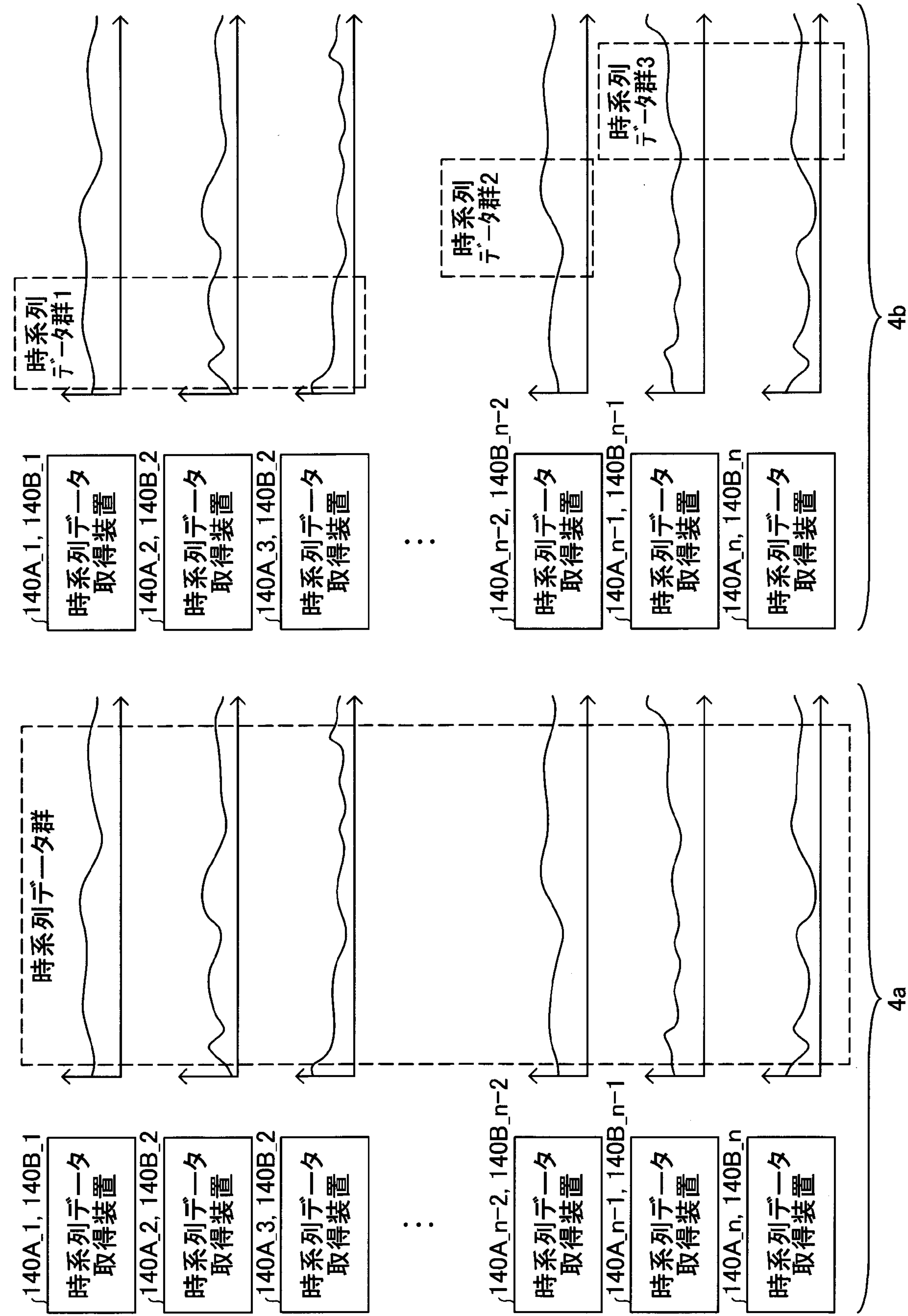
[図2]



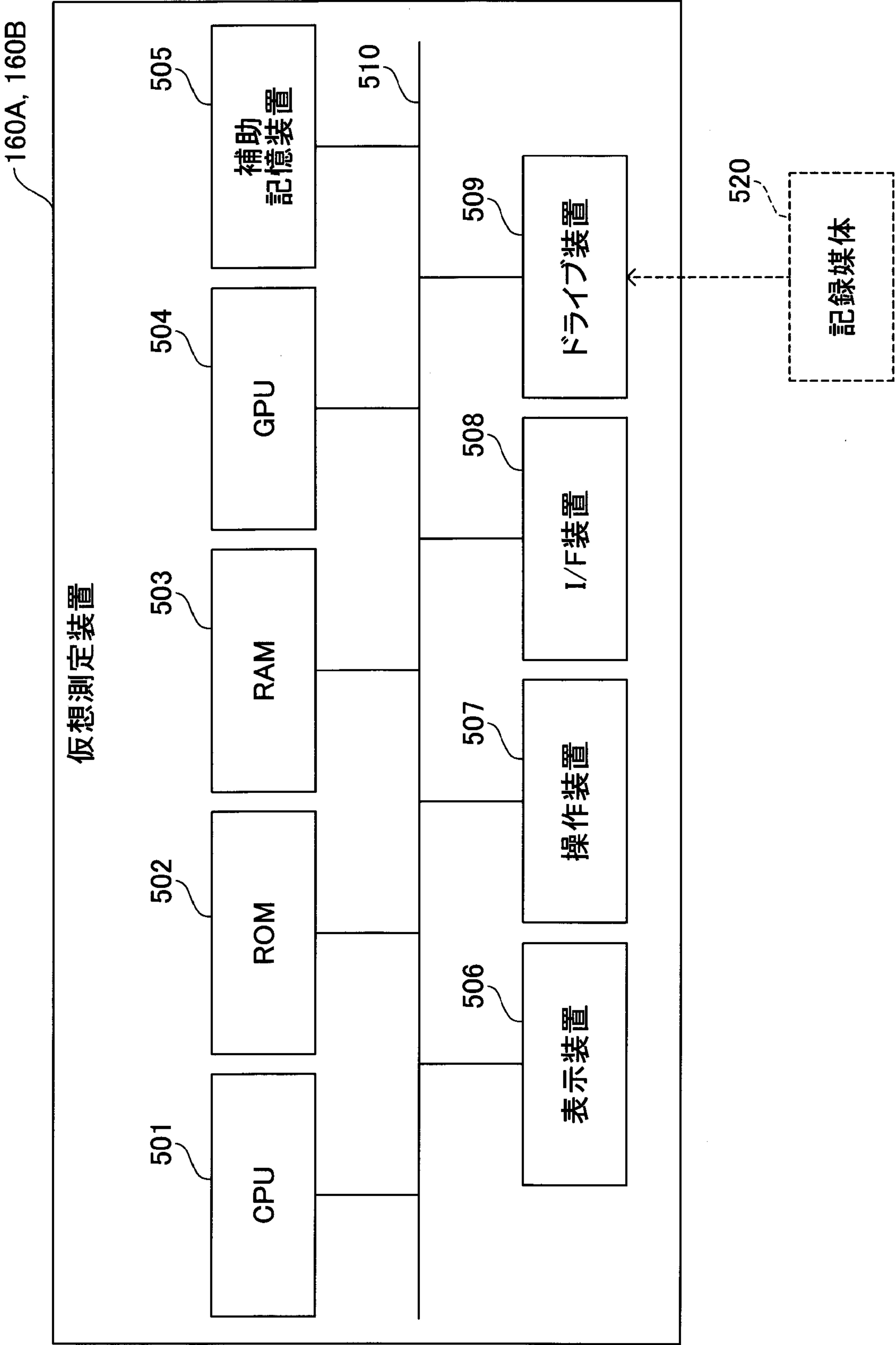
[図3]



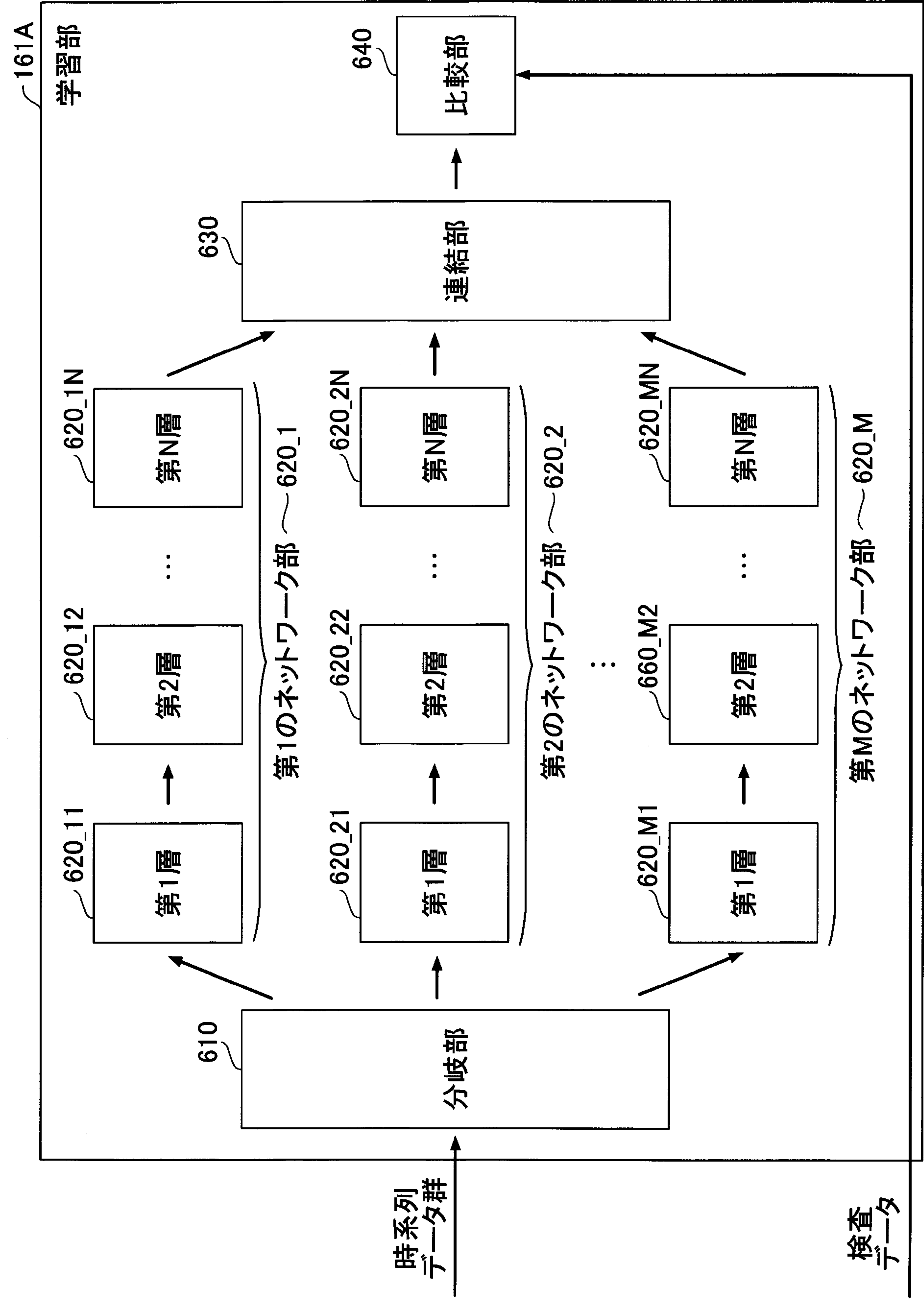
[図4]



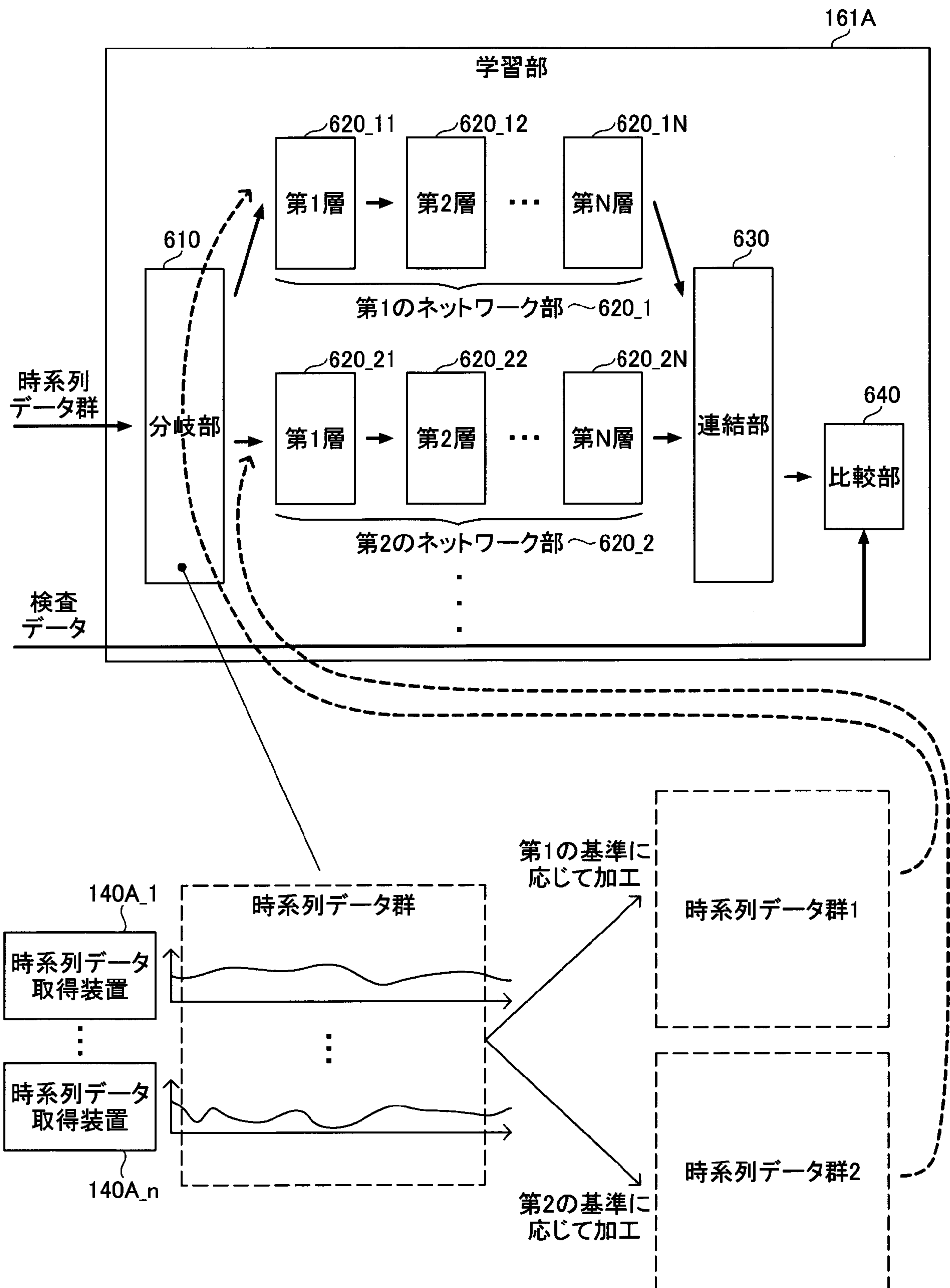
[図5]



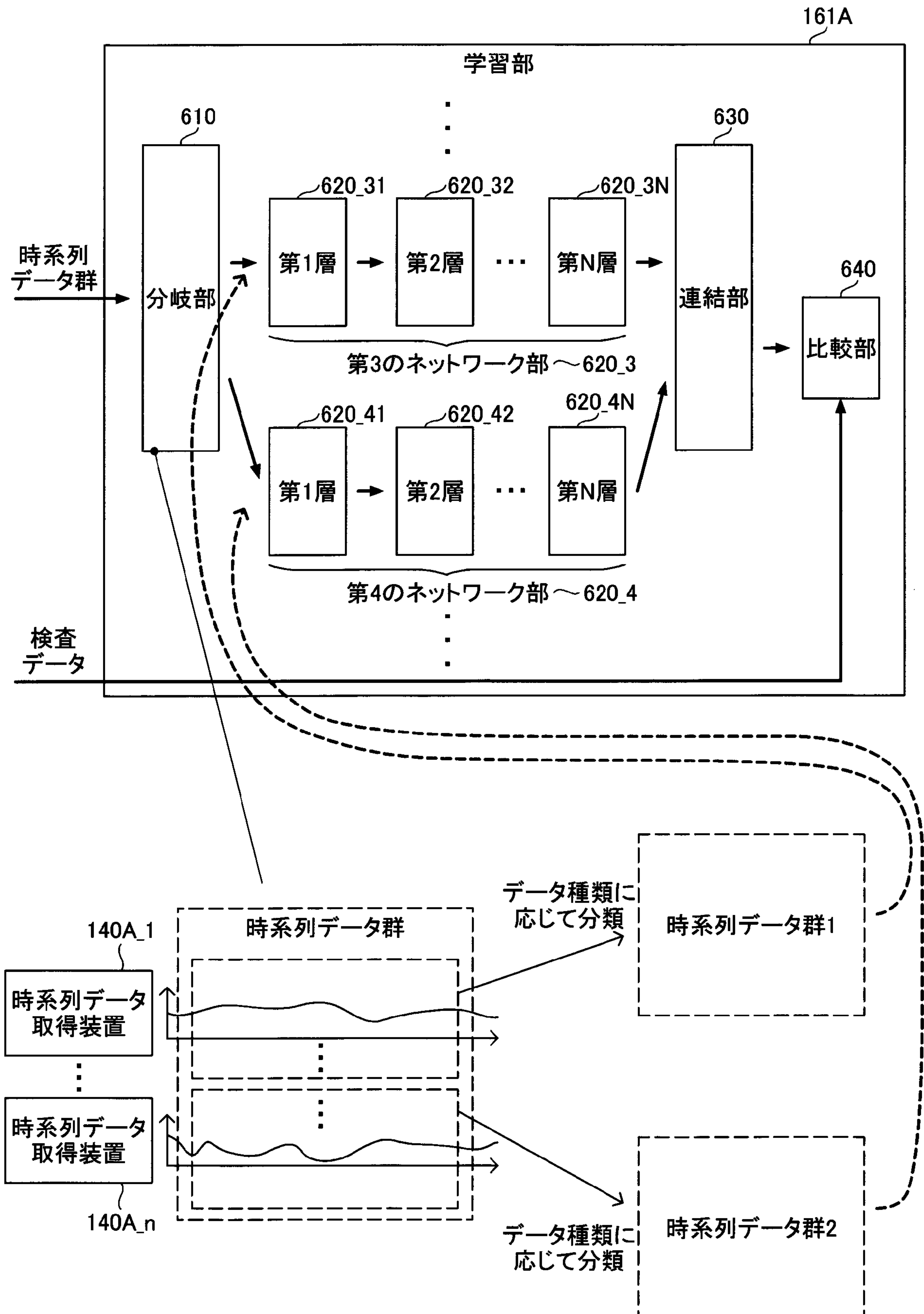
[図6]



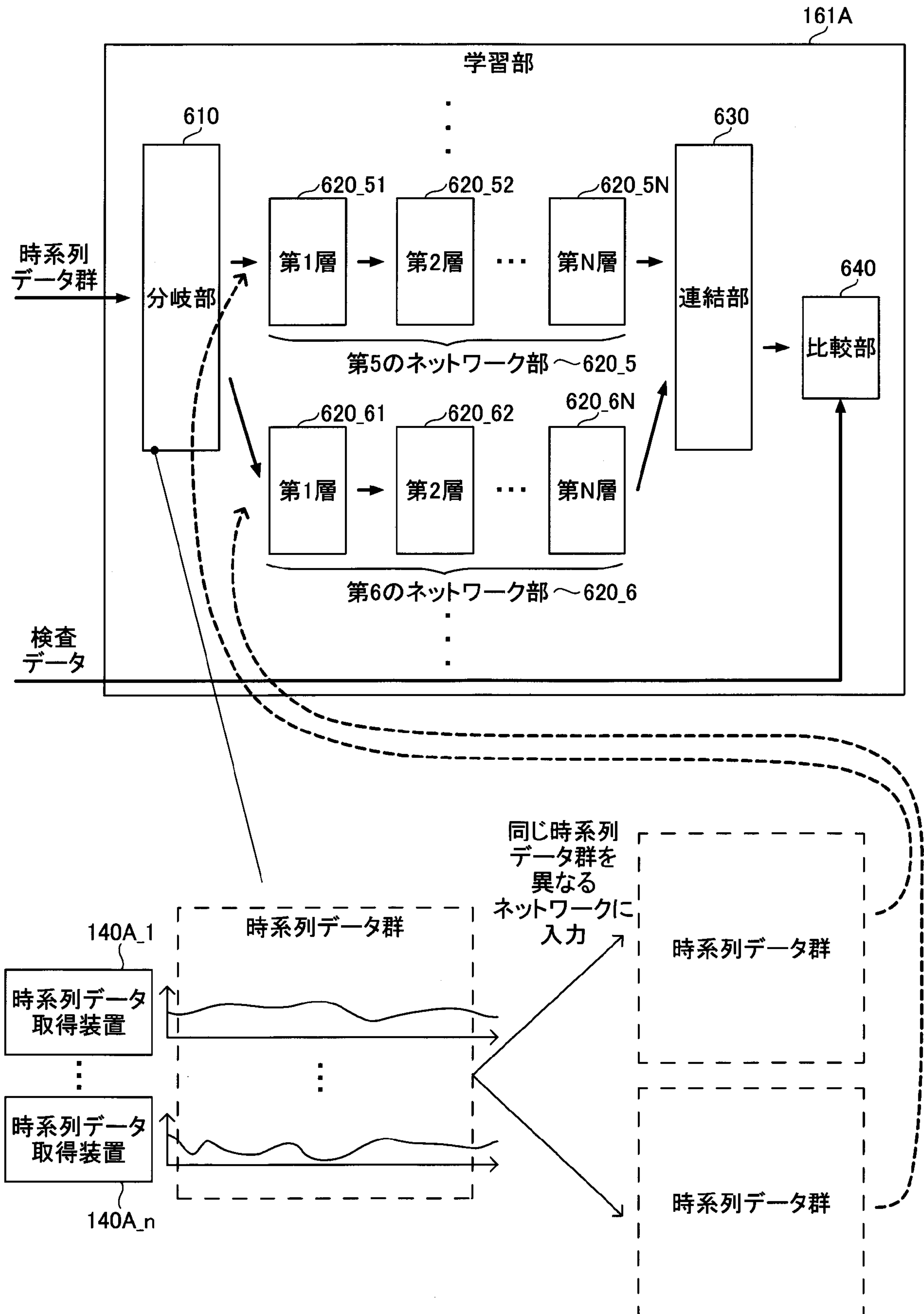
[図7]



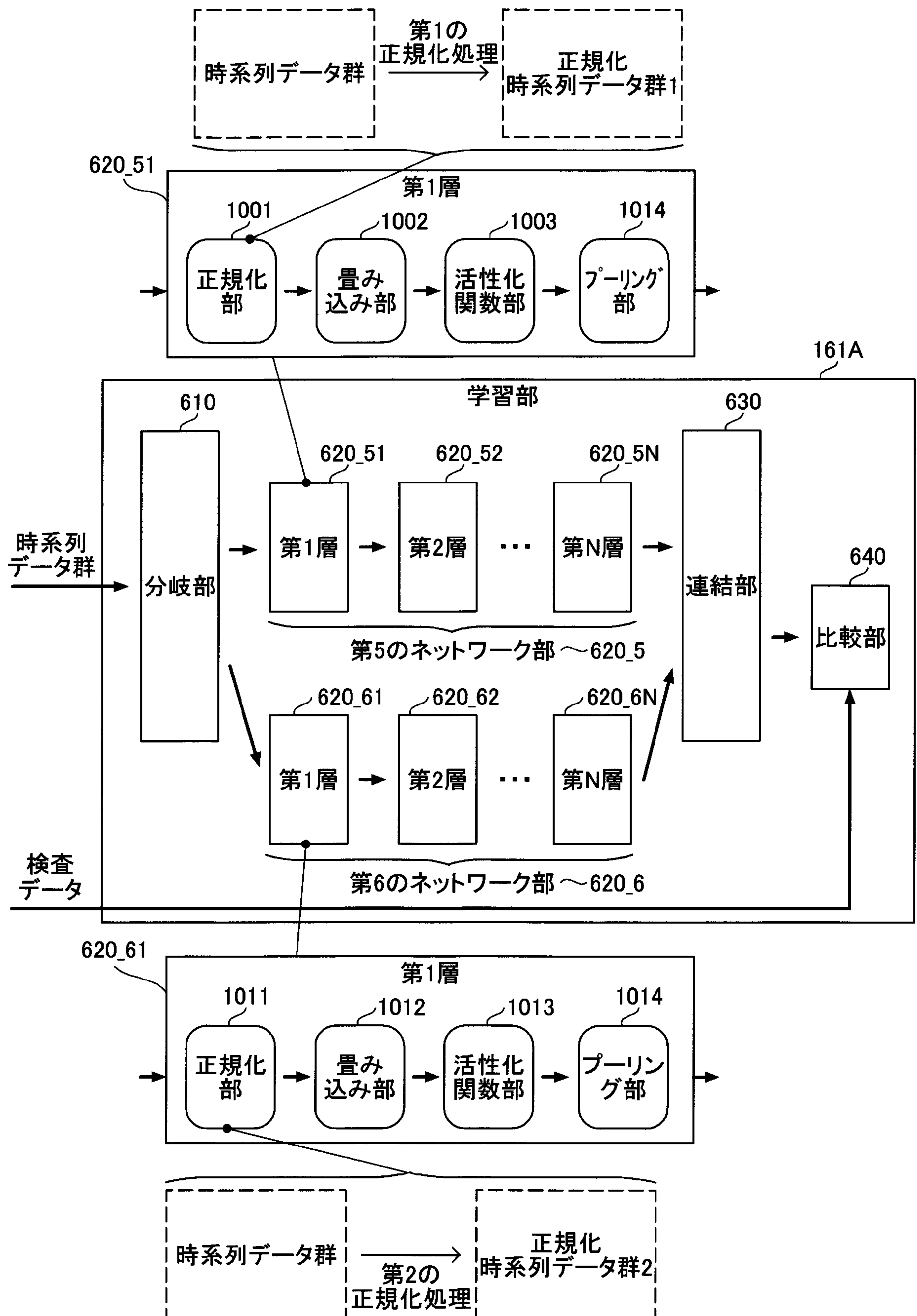
[図8]



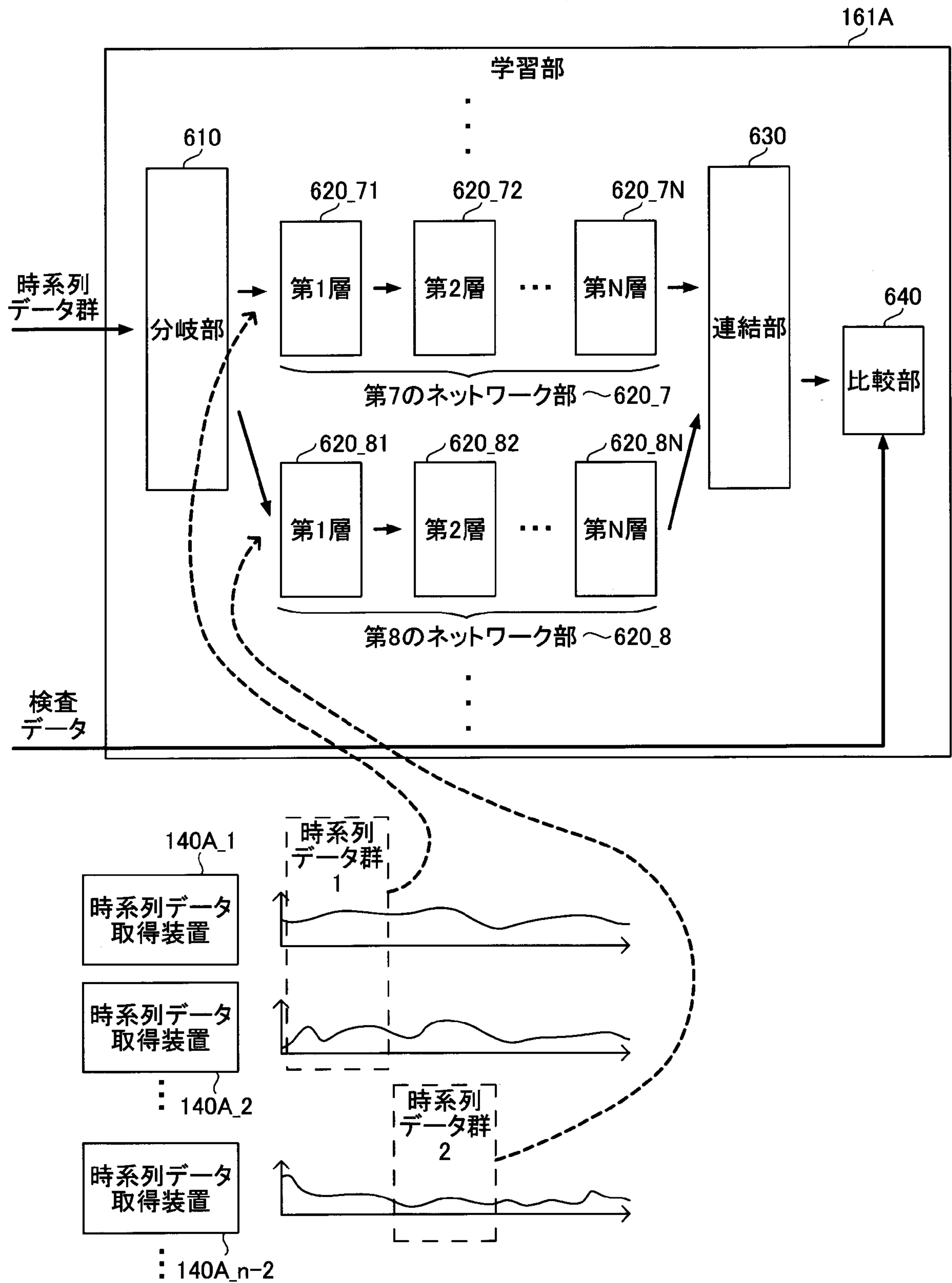
[図9]



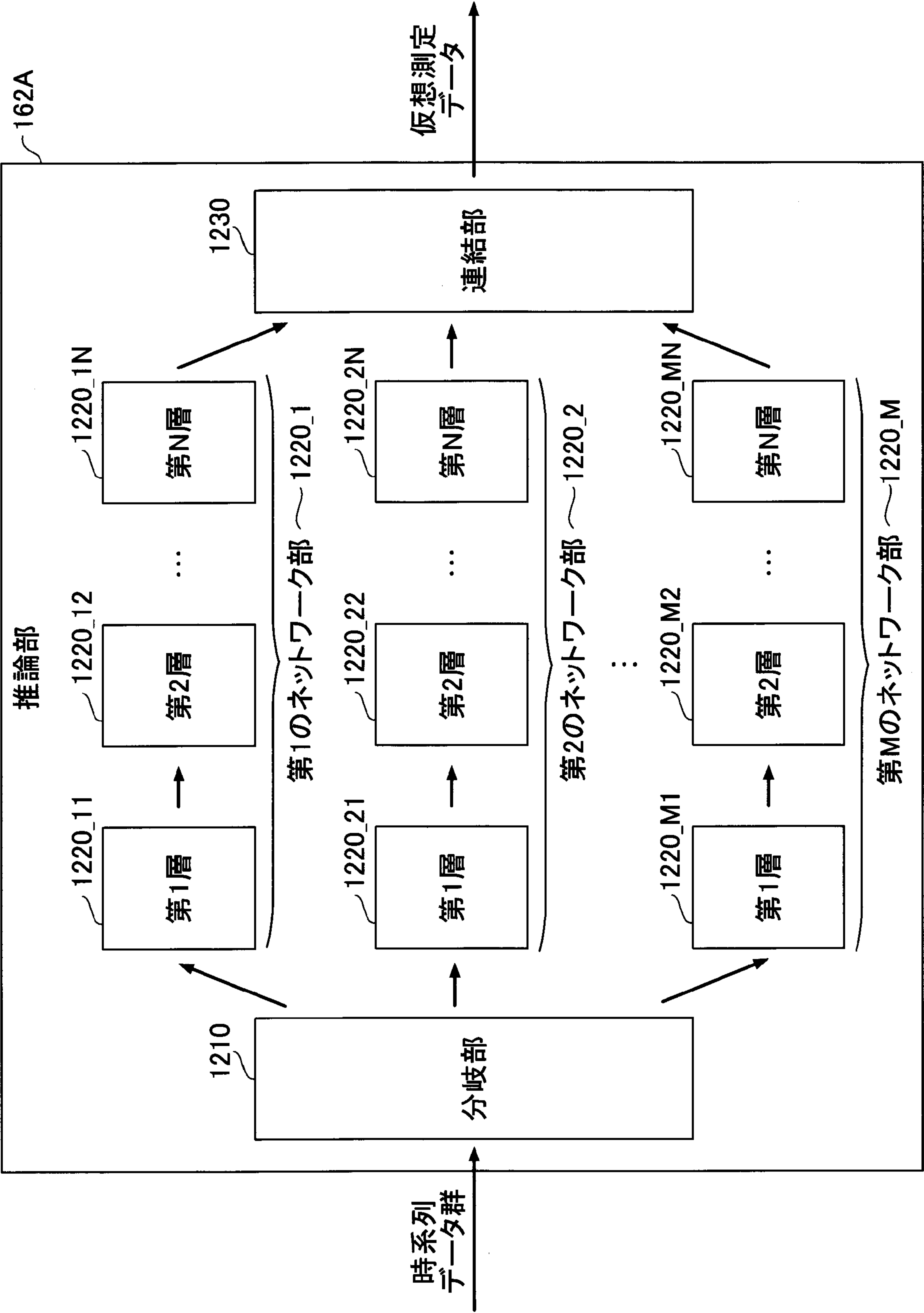
[図10]



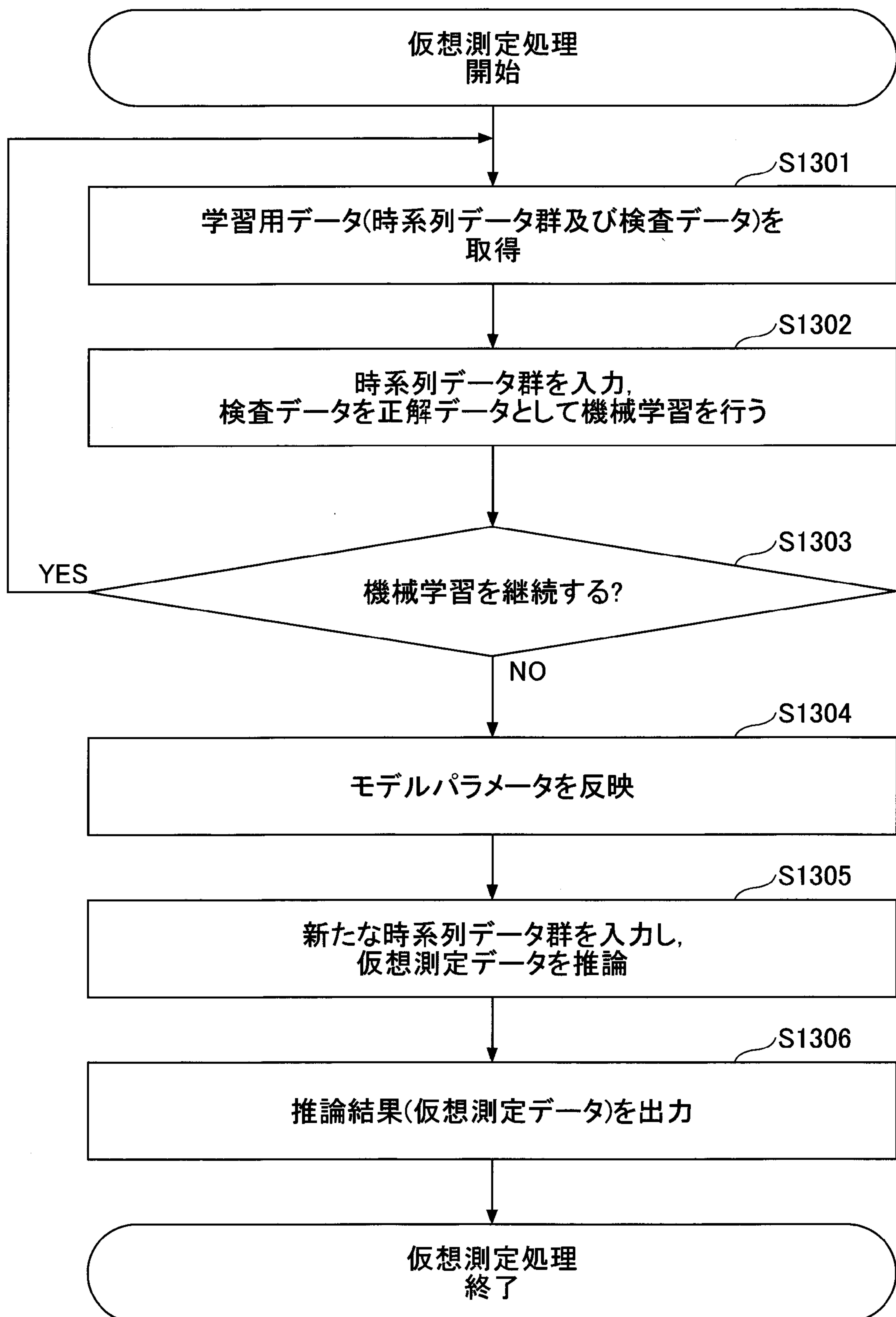
[図11]



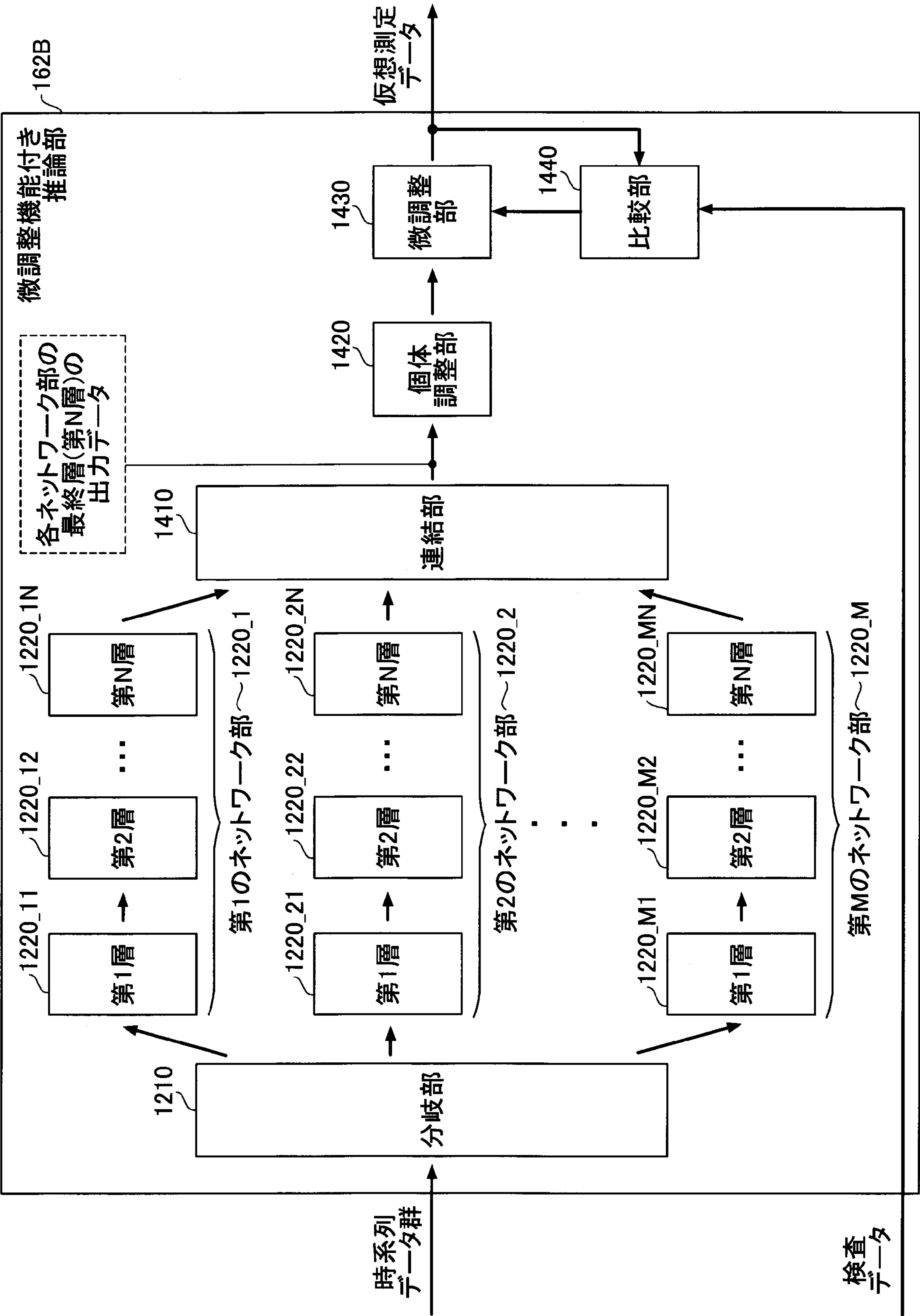
[図12]



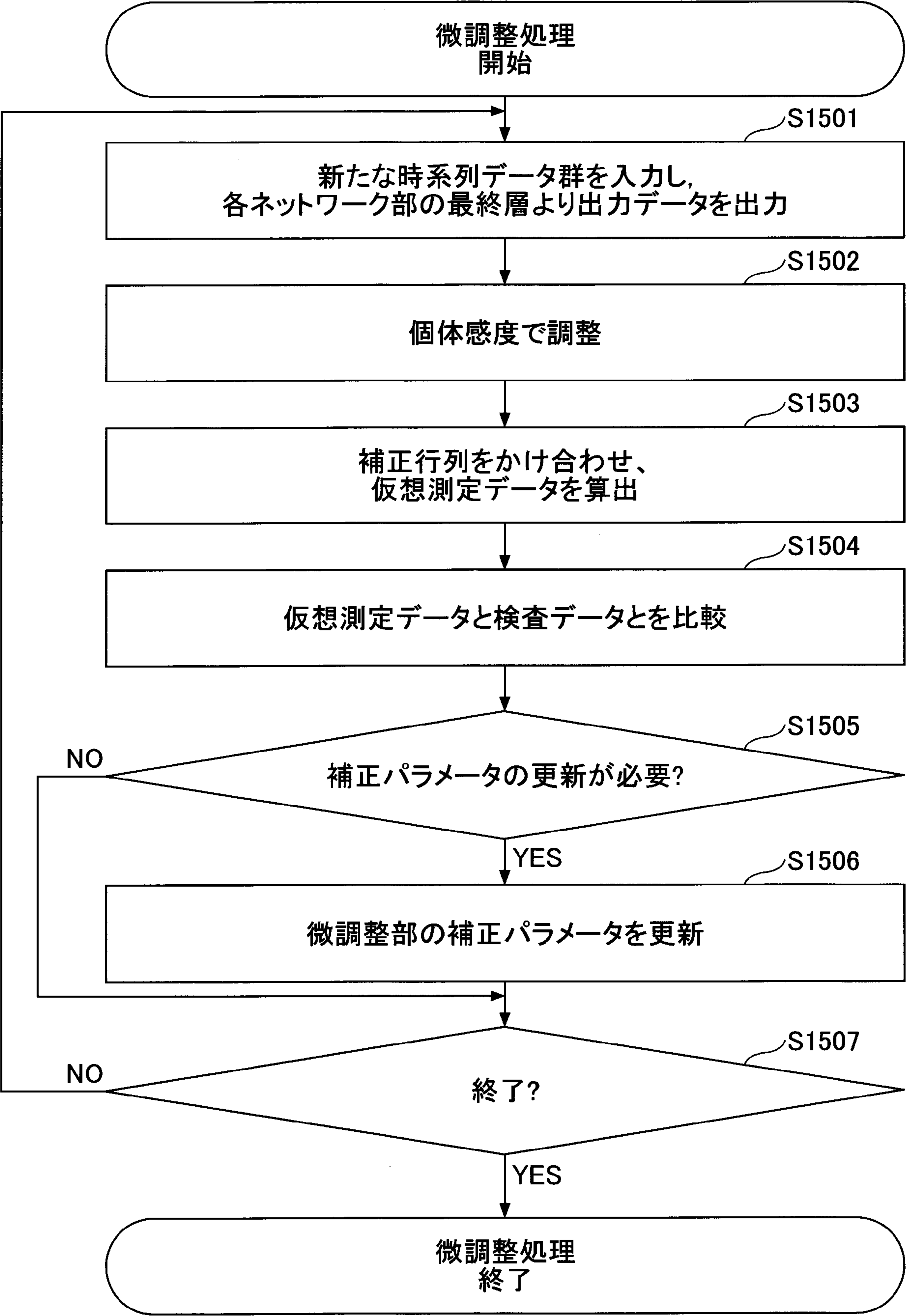
[図13]



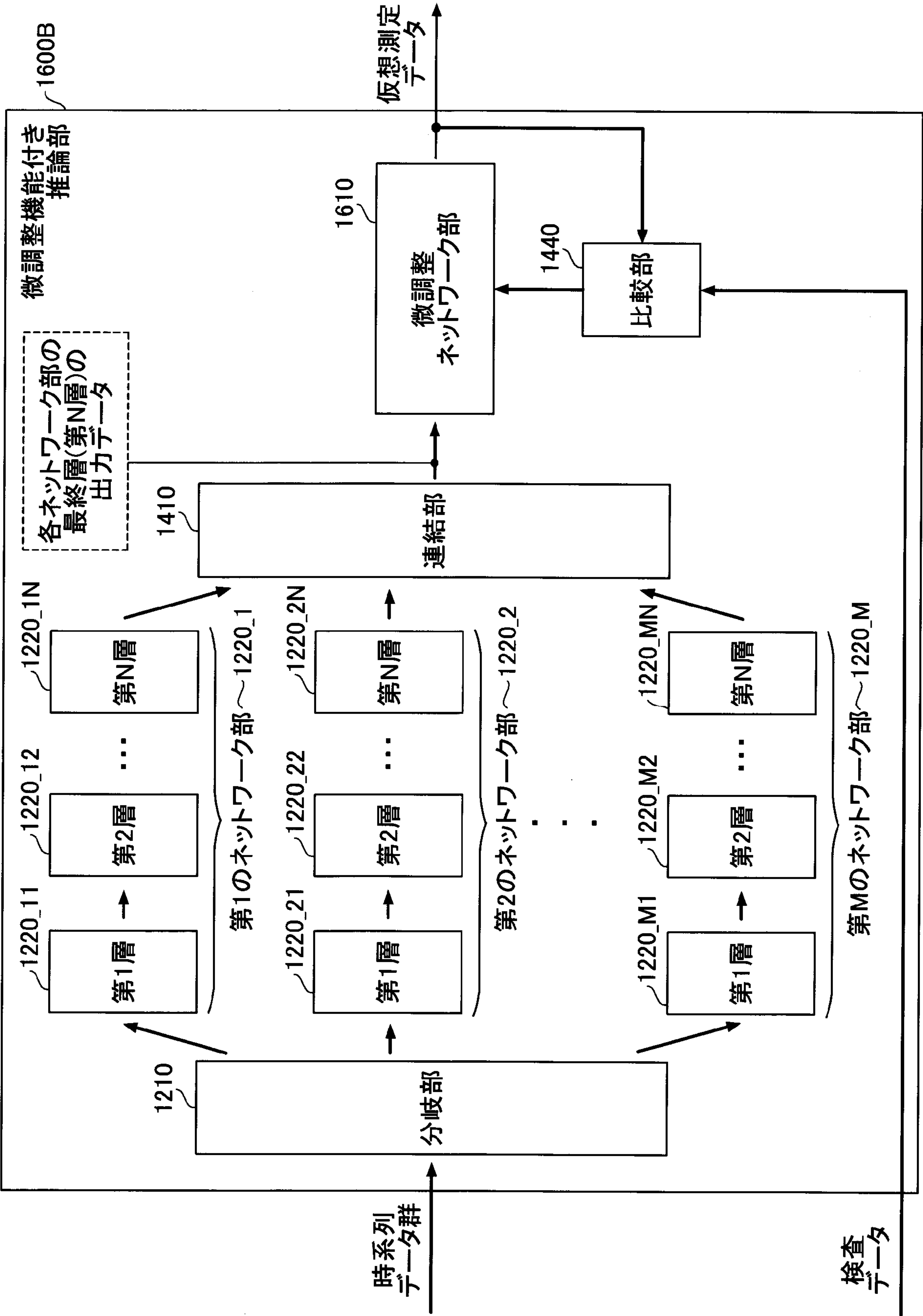
[図14]



[図15]



[図16]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2020/042564

A. CLASSIFICATION OF SUBJECT MATTER

G05B 23/02 (2006.01) i; G06N 20/00 (2019.01) i

FI: G05B23/02 Z; G06N20/00

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)

G05B23/02; G06N20/00

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2020

Registered utility model specifications of Japan 1996-2020

Published registered utility model applications of Japan 1994-2020

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.
Y	WO 2010/119894 A1 (SHARP CORP.) 21 October 2010 (2010-10-21) paragraphs [0063]-[0086], fig. 1-4	1-9
Y	JP 2015-127914 A (MITSUBISHI HEAVY INDUSTRIES, LTD.) 09 July 2015 (2015-07-09) paragraphs [0027]-[0051], fig. 1-4	1-9
Y	JP 7-36505 A (NKK CORP.) 07 February 1995 (1995-02-07) paragraphs [0020]-[0093], fig. 1-11	3-6
Y	WO 2019/159483 A1 (OMRON CORP.) 22 August 2019 (2019-08-22) paragraphs [0020]-[0044], fig. 1-7	3-6
A	US 2007/0100487 A1 (TAIWAN SEMICONDUCTOR MANUFACTURING COMPANY LTD.) 03 May 2007 (2007-05-03) paragraphs [0011]-[0035], fig. 1-6	1-9
A	JP 10-260718 A (FUJI ELECTRIC CO., LTD.) 29 September 1998 (1998-09-29) paragraphs [0002]-[0049], fig. 1-13	1-9

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

* Special categories of cited documents:	"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
"A" document defining the general state of the art which is not considered to be of particular relevance	"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
"E" earlier application or patent but published on or after the international filing date	"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
"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)	"&" document member of the same patent family
"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	

Date of the actual completion of the international search 24 December 2020 (24.12.2020)	Date of mailing of the international search report 12 January 2021 (12.01.2021)
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/JP2020/042564

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
WO 2010/119894 A1	21 Oct. 2010	JP 2010-267242 A	
JP 2015-127914 A	09 Jul. 2015	(Family: none)	
JP 7-36505 A	07 Feb. 1995	(Family: none)	
WO 2019/159483 A1	22 Aug. 2019	JP 2019-140586 A	
US 2007/0100487 A1	03 May 2007	CN 101067742 A	
JP 10-260718 A	29 Sep. 1998	(Family: none)	

国際調査報告		国際出願番号 PCT/JP2020/042564
A. 発明の属する分野の分類（国際特許分類（I P C）） G05B 23/02(2006.01)i; G06N 20/00(2019.01)i FI: G05B23/02 Z; G06N20/00		
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（I P C）） G05B23/02; G06N20/00		
最小限資料以外の資料で調査を行った分野に含まれるもの 日本国実用新案公報 1 9 2 2 - 1 9 9 6 年 日本国公開実用新案公報 1 9 7 1 - 2 0 2 0 年 日本国実用新案登録公報 1 9 9 6 - 2 0 2 0 年 日本国登録実用新案公報 1 9 9 4 - 2 0 2 0 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
Y	WO 2010/119894 A1（シャープ株式会社）21.10.2010（2010 - 10 - 21） [0 0 6 3] - [0 0 8 6]、図 1 - 4	1-9
Y	JP 2015-127914 A（三菱重工業株式会社）09.07.2015（2015 - 07 - 09） [0 0 2 7] - [0 0 5 1]、図 1 - 4	1-9
Y	JP 7-36505 A（日本鋼管株式会社）07.02.1995（1995 - 02 - 07） [0 0 2 0] - [0 0 9 3]、図 1 - 1 1	3-6
Y	WO 2019/159483 A1（オムロン株式会社）22.08.2019（2019 - 08 - 22） [0 0 2 0] - [0 0 4 4]、図 1 - 7	3-6
A	US 2007/0100487 A1（TAIWAN SEMICONDUCTOR MANUFACTURING COMPANY, LTD.） 03.05.2007（2007 - 05 - 03） [0 0 1 1] - [0 0 3 5]、図 1 - 6	1-9
A	JP 10-260718 A（富士電機株式会社）29.09.1998（1998 - 09 - 29） [0 0 0 2] - [0 0 4 9]、図 1 - 1 3	1-9
☐ 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 24. 12. 2020		国際調査報告の発送日 12. 01. 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） 中田 善邦 3U 3225 電話番号 03-3581-1101 内線 3364

国際出願番号
PCT/JP2020/042564

引用文献			公表日	パテントファミリー文献			公表日
WO	2010/119894	A1	21.10.2010	JP	2010-267242	A	

JP	2015-127914	A	09.07.2015	(ファミリーなし)			

JP	7-36505	A	07.02.1995	(ファミリーなし)			

WO	2019/159483	A1	22.08.2019	JP	2019-140586	A	

US	2007/0100487	A1	03.05.2007	CN	101067742	A	

JP	10-260718	A	29.09.1998	(ファミリーなし)			
