

(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4月22日(22.04.2021)



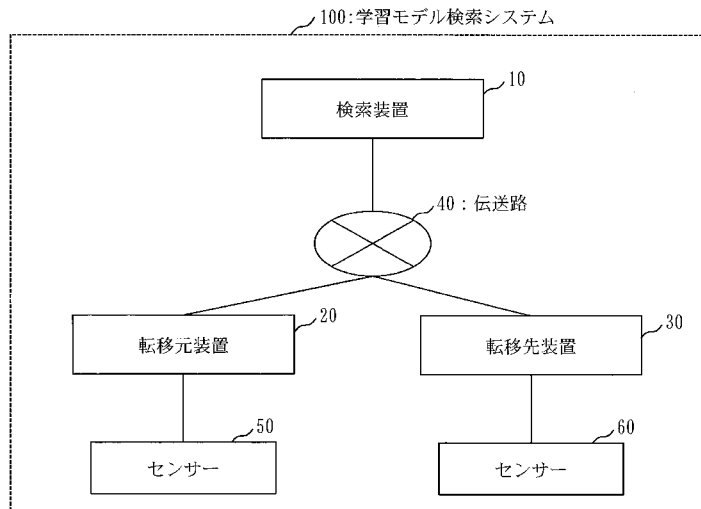
(10) 国際公開番号
WO 2021/074990 A1

- (51) 国際特許分類: *G06N 20/00* (2019.01) *G06F 16/90* (2019.01)
- (21) 国際出願番号: PCT/JP2019/040614
- (22) 国際出願日: 2019年10月16日(16.10.2019)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 三菱電機株式会社(MITSUBISHI ELECTRIC CORPORATION) [JP/JP]; 〒1008310 東京都千代田区丸の内二丁目7番3号 Tokyo (JP).
- (72) 発明者: 森 郁海(MORI, Ikumi); 〒1008310 東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内 Tokyo (JP).
- (74) 代理人: 溝井 国際特許業務法人(MIZOI INTERNATIONAL PATENT FIRM); 〒2470056 神奈川県鎌倉市大船二丁目17番10号3階 Kanagawa (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, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

(54) Title: SEARCH DEVICE, SEARCH METHOD, SEARCH PROGRAM, AND LEARNING MODEL SEARCH SYSTEM

(54) 発明の名称: 検索装置、検索方法、検索プログラム及び学習モデル検索システム

[図1]



10... SEARCH DEVICE
 20... TRANSFER SOURCE DEVICE
 30... TRANSFER DESTINATION DEVICE
 40... TRANSMISSION ROUTE
 50, 60... SENSOR
 100... LEARNING MODEL SEARCH SYSTEM

(57) Abstract: A search device (10) that obtains first data obtained by basis conversion of a feature value vector in a transfer source device (20), said basis conversion being performed on the basis of the information amount for each feature axis. In addition, the search device (10) obtains second data obtained by basis conversion of a feature value vector in a transfer destination device (30), said basis conversion being performed on the basis of the information amount for each feature axis. The search device (10) determines whether or not the transfer source device (20) is suitable as a transfer



WO 2021/074990 A1

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,
SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, 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))

source, by determining whether or not the first data and the second data are similar.

(57) 要約 : 検索装置 (10) は、転移元装置 (20) における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第1データを取得する。また、検索装置 (10) は、転移先装置 (30) における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第2データを取得する。検索装置 (10) は、第1データと第2データとが類似するか否かを判定することにより、転移元装置 (20) が転送元として相応しいか否かを判定する。

明 細 書

発明の名称：

検索装置、検索方法、検索プログラム及び学習モデル検索システム

技術分野

[0001] この発明は、転移学習における転移元の検索技術に関する。

背景技術

[0002] IoT (Internet of Things) 機器上でAI (Artificial Intelligence) を利用するソリューションが増加している。例えば、以下のようなアプリケーションが挙げられる。(1) 空調及び照明といったIoT家電の制御、(2) 製造装置の故障診断、(3) 製造ラインにおける製品の画像による検品、(4) ビル等の入門時の動画による不審者侵入検知、(4) EMS (エネルギー管理システム) におけるエネルギー需要予測、(5) プラントの故障診断。

[0003] IoT機器毎にAIを利用する場合、学習処理に用いる十分な数の訓練データを確保することが困難になる。そこで、少ない訓練データで効率的に学習を行う必要がある。少ない訓練データで学習する方法として、訓練データが収集された環境とは異なる環境の訓練データ及び学習モデルを転移させる、転移学習と呼ばれる方法がある。

転移学習では、転移元を特定するために、転移元の全データに対して転移元になり得るかを1つずつ評価する。評価の結果、転移が有効であることを示す「正の転移」であることを確認できた場合、転移元データとして決定する。この評価は、自動で行うことが望ましいが、何らかの形で人手が関わることもある。

[0004] 特許文献1には、転移元になり得るかを評価する技術について記載されている。具体的には、特許文献1には、転移元の訓練データを用いて学習を試行し、転移先のデータを入力とした推論結果と転移元のデータを入力とした推論結果との違いを利用して転移の有効性を判定することが記載されている

。

先行技術文献

特許文献

[0005] 特許文献1：特開2016-191975号公報

発明の概要

発明が解決しようとする課題

[0006] 特許文献1に記載された技術では、転移元になり得るかを評価する際、転移元の訓練データを用いて学習を試行せねばならず、転移元の探索空間が大きい場合は、処理時間を要してしまう。

この発明は、短い処理時間で適切な転移元を特定可能にすることを目的とする。

課題を解決するための手段

[0007] この発明に係る検索装置は、

転移元装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第1データを取得する第1取得部と、

転移先装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第2データを取得する第2取得部と、

前記第1取得部によって取得された前記第1データと、前記第2取得部によって取得された前記第2データとが類似するか否かを判定する類似判定部と

を備える。

発明の効果

[0008] この発明では、特徴量ベクトルを特徴軸毎の情報量を基準として基底変換して得られたデータが類似するか否かを判定する。データが類似するか否かにより、転移元になり得るかを評価することが可能である。データが類似するか否かを判定する処理は、転移元の訓練データを用いて学習を試行する処理に比べ処理時間がかからない。したがって、短い処理時間で適切な転移元

を特定可能になる。

図面の簡単な説明

- [0009] [図1]実施の形態1に係る学習モデル検索システム100の構成図。
- [図2]実施の形態1に係る検索装置10の構成図。
- [図3]実施の形態1に係る転移元装置20の構成図。
- [図4]実施の形態1に係る転移先装置30の構成図。
- [図5]実施の形態1に係る学習モデル検索システム100の全体的な処理の説明図。
- [図6]実施の形態1に係る転移元装置20の第1データ送信処理のフローチャート。
- [図7]実施の形態1に係る基底変換処理の説明図。
- [図8]実施の形態1に係る正規化処理の説明図。
- [図9]実施の形態1に係るベクトル \vec{z} の説明図。
- [図10]実施の形態1に係る2次元画像の説明図。
- [図11]実施の形態1に係る軸の対応関係の説明図。
- [図12]実施の形態1に係る転移先装置30の第2データ送信処理のフローチャート。
- [図13]実施の形態1に係る検索装置10の検索処理のフローチャート。
- [図14]実施の形態1に係る無相関ではないと判定された場合における類似度計算処理のフローチャート。
- [図15]実施の形態1に係る軸の対応関係の説明図。
- [図16]実施の形態1に係る転移先装置30の分析処理のフローチャート。
- [図17]実施の形態1に係る学習モデル検索システム100を用いた転移元特定処理の説明図。
- [図18]転移元の候補となる転移元装置20が2つ以上である場合における転移先装置30の分析処理のフローチャート。
- [図19]実施の形態1に係る2次元画像の例の説明図。
- [図20]実施の形態2に係る類似判定処理のフローチャート。

[図21]実施の形態3に係る類似判定処理のフローチャート。

[図22]実施の形態3に係る検定方法選定の説明図。

[図23]実施の形態4に係る類似判定処理のフローチャート。

発明を実施するための形態

[0010] 実施の形態1.

構成の説明

図1を参照して、実施の形態1に係る学習モデル検索システム100の構成を説明する。

学習モデル検索システム100は、検索装置10と、1つ以上の転移元装置20と、転移先装置30とを備える。検索装置10と転移元装置20と転移先装置30とは、インターネットといった伝送路40を介して接続されている。

各転移元装置20には、1つ以上のセンサ50が接続されている。転移先装置30には、1つ以上のセンサ60が接続されている。

[0011] 図2を参照して、実施の形態1に係る検索装置10の構成を説明する。

検索装置10は、クラウドコンピューティングにおけるサーバといったコンピュータである。

検索装置10は、コンピュータである。

検索装置10は、プロセッサ11と、メモリ12と、ストレージ13と、通信インタフェース14とのハードウェアを備える。プロセッサ11は、信号線を介して他のハードウェアと接続され、これら他のハードウェアを制御する。

[0012] 検索装置10は、機能構成要素として、第1取得部111と、第2取得部112と、類似判定部113と、写像生成部114と、データ送信部115とを備える。検索装置10の各機能構成要素の機能はソフトウェアにより実現される。

ストレージ13には、検索装置10の各機能構成要素の機能を実現するプログラムが格納されている。このプログラムは、プロセッサ11によりメモ

り12に読み込まれ、プロセッサ11によって実行される。これにより、検索装置10の各機能構成要素の機能が実現される。

[0013] また、ストレージ13は、学習モデル記憶部131と、統計量記憶部132とを実現する。

[0014] 図3を参照して、実施の形態1に係る転移元装置20の構成を説明する。

転移元装置20は、IoT機器といったコンピュータである。

転移元装置20は、プロセッサ21と、メモリ22と、ストレージ23と、通信インタフェース24とのハードウェアを備える。プロセッサ21は、信号線を介して他のハードウェアと接続され、これら他のハードウェアを制御する。

[0015] 転移元装置20は、機能構成要素として、基底変換部211と、正規化部212と、統計量計算部213と、データ送信部214とを備える。転移元装置20の各機能構成要素の機能はソフトウェアにより実現される。

ストレージ23には、転移元装置20の各機能構成要素の機能を実現するプログラムが格納されている。このプログラムは、プロセッサ21によりメモリ22に読み込まれ、プロセッサ21によって実行される。これにより、転移元装置20の各機能構成要素の機能が実現される。

[0016] また、ストレージ23は、学習モデル記憶部231と、訓練データ記憶部232とを実現する。

[0017] 図4を参照して、実施の形態1に係る転移先装置30の構成を説明する。

転移先装置30は、IoT機器といったコンピュータである。

転移先装置30は、プロセッサ31と、メモリ32と、ストレージ33と、通信インタフェース34とのハードウェアを備える。プロセッサ31は、信号線を介して他のハードウェアと接続され、これら他のハードウェアを制御する。

[0018] 転移先装置30は、機能構成要素として、基底変換部311と、正規化部312と、統計量計算部313と、データ送信部314と、データ取得部315と、学習モデル生成部316と、入力データ変換部317と、出力ラベ

ル変換部318とを備える。転移先装置30の各機能構成要素の機能はソフトウェアにより実現される。

ストレージ33には、転移先装置30の各機能構成要素の機能を実現するプログラムが格納されている。このプログラムは、プロセッサ31によりメモリ32に読み込まれ、プロセッサ31によって実行される。これにより、転移先装置30の各機能構成要素の機能が実現される。

[0019] また、ストレージ33は、学習モデル記憶部331と、観測データ記憶部332とを実現する。

[0020] プロセッサ11, 21, 31は、プロセッシングを行うIC (Integrated Circuit) である。プロセッサ11, 21, 31は、具体例としては、CPU (Central Processing Unit)、DSP (Digital Signal Processor)、GPU (Graphics Processing Unit) である。

[0021] メモリ12, 22, 32は、データを一時的に記憶する記憶装置である。メモリ12, 22, 32は、具体例としては、SRAM (Static Random Access Memory)、DRAM (Dynamic Random Access Memory) である。

[0022] ストレージ13, 23, 33は、データを保管する記憶装置である。ストレージ13, 23, 33は、具体例としては、HDD (Hard Disk Drive) である。また、ストレージ13, 23, 33は、SD (登録商標, Secure Digital) メモリカード、CF (Compact Flash, 登録商標)、NANDフラッシュ、フレキシブルディスク、光ディスク、コンパクトディスク、ブルーレイ (登録商標) ディスク、DVD (Digital Versatile Disk) といった可搬記録媒体であってもよい。

[0023] 通信インタフェース14, 24, 34は、外部の装置と通信するためのインタフェースである。通信インタフェース14, 24, 34は、具体例としては、Ethernet (登録商標)、USB (Universal Se

rial Bus)、HDMI(登録商標, High-Definition Multimedia Interface)のポートである。

[0024] ***動作の説明***

図5から図16を参照して、実施の形態1に係る学習モデル検索システム100の動作を説明する。

実施の形態1に係る学習モデル検索システム100の検索装置10の動作手順は、実施の形態1に係る検索方法に相当する。また、実施の形態1に係る学習モデル検索システム100の検索装置10の動作を実現するプログラムは、実施の形態1に係る検索プログラムに相当する。

[0025] 図5を参照して、実施の形態1に係る学習モデル検索システム100の全体的な処理について説明する。

(1) 各転移元装置20は、訓練データから類似性の比較に必要な統計量とを生成する。訓練データは、各転移元装置20がセンサ50から取得したデータに対して教師データ(ラベル)の付与を行い生成されたデータである。(2) 各転移元装置20は、学習モデルと統計量とを検索装置10に送信する。(3) 転移先装置30は、観測データから類似性の比較に必要な統計量を生成し、検索装置10に送信する。観測データは、転移先装置30がセンサ60から取得したデータに対して教師データ(ラベル)の付与を行い生成されたデータである。

(4) 検索装置10は、各転移元装置20によって生成された統計量と、転移先装置30によって生成された統計量とが類似するか否かを判定する。これにより、検索装置10は、転送元の候補になる転移元装置20を特定する。(5) 検索装置10は、転移元の候補になる転移元装置20について、データ写像 f 及びラベル写像 g を生成する。データ写像 f は、転移先から転移元への入力変換である。ラベル写像 g は、転移元から転移先への出力変換である。

(6) 転移先装置30は、転移元の候補になった転移元装置20の学習モデルを入力とし、転移先装置30の学習器を生成する。(7) 転移先装置3

0は、観測データをデータ写像 f で変換した後、生成された学習器に入力する。(8) 転移先装置 30は、学習器から出力されたラベルをラベル写像 g で変換する。(9) 転移先装置 30は、変換されたラベルを出力する。

[0026] 図6を参照して、実施の形態1に係る転移元装置20の第1データ送信処理(図5の(1)(2)の処理に相当)を説明する。

(ステップS11: 基底変換処理)

基底変換部211は、訓練データ記憶部232に記憶された訓練データの特徴量ベクトルの座標系を変換する。訓練データの特徴量ベクトルは、訓練データからラベルを除いたデータである。この処理は、転移元装置20の訓練データの特徴量ベクトルの分布と、転移先装置30の観測データの特徴量ベクトルの分布とを比較するために、座標系を揃える処理である。

具体的には、基底変換部211は、特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換する。ここでは、図7に示すように、基底変換部211は、主成分分析を用いて、特徴量ベクトルの各要素の情報量が大きい特徴軸から順にベクトル $z \rightarrow$ の要素 z_i を割り当てることにより、正規直交基底を得る。ここで、情報量は、分散値又は固有値と言い換えることができる。図7では、最も情報量が大きい特徴軸に基底要素 z_1 が割り当てられ、次に情報量が大きい特徴軸に要素 z_2 が割り当てられている。つまり、基底変換部211は、 p 次元ユークリッド空間 R^p 上の特徴量ベクトル $x \rightarrow$ を、 m 次元主成分空間 Z^m 上のベクトル $z \rightarrow$ に変換する。

ここで、ベクトル $z \rightarrow$ の第 i 主成分を要素 z_i 、要素 z_i の寄与率を $P V_i$ 、累積寄与率を $C P V_m$ と表記する。この変換により、主成分同士が無相関化される。ベクトル $z \rightarrow$ の次元数を m とすると、 $1 \leq m \leq p$ 、 $0 < C P V_m \leq 1$ を満たす。特に、 $m < p$ の場合、次元削減と呼ばれる。主成分分析により、転移元装置20及び転移先装置30の特徴量ベクトル空間の軸が寄与率の降順にソートされる。

[0027] (ステップS12: 正規化処理)

正規化部212は、ステップS11で座標系が変換された後のベクトル $z \rightarrow$

を、定義域が一定の範囲に収まるように変換する。この処理は、転移元装置 20 の訓練データの特徴量ベクトルの分布と、転移先装置 30 の観測データの特徴量ベクトルの分布とをスケールにとらわれずに比較をするために、特徴量ベクトルを正規化する処理である。

具体的には、図 8 に示すように、正規化部 212 は、ベクトル $z \rightarrow$ の要素 z_i のスケールを $z_{\min} \leq z_i \leq z_{\max}$ となるよう、数 1 により正規化する。ベクトル $z \rightarrow$ を正規化したベクトルを $z^{\wedge} \rightarrow$ と表記する。

[数1]

$$\hat{z}_i = \mathcal{C}(z_i, z_{\min}, z_{\max})$$

$$s. t. \mathcal{C}(x, C_{\min}, C_{\max}) = \frac{x - \min(x)}{\max(x) - \min(x)} (C_{\max} - C_{\min}) + C_{\min}$$

[0028] (ステップ S 13 : 統計量計算処理)

統計量計算部 213 は、ステップ S 12 で変換されたデータに対して統計量を計算する。この処理は、転移元装置 20 の訓練データの特徴量ベクトルの分布と、転移先装置 30 の観測データの特徴量ベクトルの分布とを比較する際に使用する統計量を計算する処理である。

[0029] 具体的には、まず、統計量計算部 213 は、正規化されたベクトル $z^{\wedge} \rightarrow$ を 2次元画像化する。図 9 に示すように、統計量計算部 213 は、この処理を、ラベル y_k 毎の正規化されたベクトル $z^{\wedge} \rightarrow$ に対して実行する。なお、MDS (多次元尺度構成法) と、SOM (自己組織化写像) と、t-SNE (t-distributed Stochastic Neighbor Embedding) といったデータ可視化 (次元削減) の技術が存在する。しかし、データ数を変化させると、出力画像の様相が大きく異なる場合がある。この場合、正常に類似性を判定することができない恐れがある。

そこで、統計量計算部 213 は、正規化されたベクトル $z^{\wedge} \rightarrow$ を以下の手順で 2次元画像化する。ここでは、正規化されたベクトル $z^{\wedge} \rightarrow$ は、 $z_{\min} = 0$, $z_{\max} = 255$ で正規化されているとする。

まず、統計量計算部 213 は、数 2 に示すように、正規化されたベクトル

\hat{z}_{y_k} の天井関数を計算して8ビットに量子化する。ここで、 y_k は、 y_k を意味する。以下同様に、 i_j は、 i の下付きとして j が付された i_j を意味する。

[数2]

$$\left[\hat{z}_{y_k} \right]$$

次に、統計量計算部 213 は、量子化されたデータに対して寄与率 PV で重み付けしたグレースケール画像に変換する。グレースケール画像は、ユニット U と呼ぶ小さな領域の集合からなる。 i 行 j 列のユニットを $U(i, j)$ と表記する。すると、図 10 に示すように、ユニット $U(i, j)$ の画素値は数 3 に示すように正規化されたベクトル \hat{z} の要素 \hat{z}_j の天井関数を計算した値になり、高さは 1 になり、幅 w_j は数 4 に示す値になる。

[数3]

$$\left[\hat{z}_j \right]$$

[数4]

$$w_j = \begin{cases} \left[PV_j \times 100 + 0.5 \right] & , w_j > 0 \\ 1 & , w_j \leq 0 \end{cases}$$

以降、グレースケール画像の i 行 j 列の画素値を $g_{i,j} \in G$ ($1 \leq i \leq N$, $1 \leq j \leq \sum_{j=1}^m w_j$) と表記する。ここで、図 9 に示すように、 N は、ラベル毎の特徴量ベクトル数である。図 9 では、例えば、 N_{y_1} は、ラベル y_1 の特徴量ベクトル数なので 10 になっている。

[0030] 次に、統計量計算部 213 は、転移元装置 20 と転移先装置 30 との画素値の集合 G が類似しているか否かの判定を容易にするために、各ラベルについてヒストグラムを計算する。但し、特徴量ベクトルから生成されたヒストグラムは、本来の母集団の特徴を反映していない可能性がある。そこで、統計量計算部 213 は、母集団の確率密度関数の推定を行う。ここでは、集合 G を母集団の標本としてカーネル密度推定量 $\hat{f}_n(x)$ を数 5 により定義す

る。

[数5]

$$\hat{f}_h(x) = \frac{1}{|\mathbb{G}|h} \sum_{g_{i,j} \in \mathbb{G}} K\left(\frac{x - g_{i,j}}{h}\right)$$

数5において、 $|\mathbb{G}|$ は画素数であり、 h は平滑化パラメータであるバンド幅であり、 K はカーネル関数である。

統計量計算部213は、各ラベルについて計算されたカーネル密度推定量 $\hat{f}_h(x)$ の集合を、類似しているか否かの判定に用いられる統計量を表す第1データに設定する。

[0031] (ステップS14：統計量送信処理)

データ送信部214は、ステップS11で座標系が変換された前のデータと後のデータとの軸の対応関係と、ステップS12で正規化される前の各軸 i の最小値 $_{min}(x_i)$ 及び最大値 $_{max}(x_i)$ と、ステップS13で計算された統計量を表す第1データとを検索装置10に送信する。すると、検索装置10の第1取得部111は、送信された軸の対応関係と、最小値 $_{min}(x_i)$ 及び最大値 $_{max}(x_i)$ と、第1データとを取得して、統計量記憶部132に書き込む。

図11に示すように、軸の対応関係は、軸の大小関係から特定される。図11の場合には、軸の対応関係は、数6に示すように表される。

[数6]

$$\left(z_1^{(S)}, z_2^{(S)}\right) \leftrightarrow \left(x_1^{(S)}, x_2^{(S)}\right)$$

[0032] (ステップS15：学習モデル送信処理)

データ送信部214は、訓練データ記憶部232に記憶された訓練データから生成された学習モデルを学習モデル記憶部231から読み出して、検索装置10に送信する。すると、検索装置10の第1取得部111は、送信された学習モデルを、ステップS14で送信された第1データと関連付けて、

学習モデル記憶部 131 に書き込む。

[0033] 図 12 を参照して、実施の形態 1 に係る転移先装置 30 の第 2 データ送信処理（図 5 の（3）の処理に相当）を説明する。

（ステップ S 2 1 : 基底変換処理）

基底変換部 311 は、観測データ記憶部 332 に記憶された観測データの特徴量ベクトルの座標系を変換する。座標系の変換方法は、図 6 のステップ S 1 1 と同じである。

[0034] （ステップ S 2 2 : 正規化処理）

正規化部 312 は、ステップ S 2 1 で座標系が変換された後のベクトル z を、定義域が一定の範囲に収まるように変換する。データの変換方法は、図 6 のステップ S 1 2 と同じである。なお、正規化部 312 は、図 6 のステップ S 1 2 と同じ定義域（最小値 z_{min} 及び最大値 z_{max} ）を用いる。

[0035] （ステップ S 2 3 : 統計量計算処理）

統計量計算部 313 は、ステップ S 2 2 で変換されたデータに対して統計量を計算する。統計量の計算方法は、図 6 のステップ S 1 3 と同じである。統計量計算部 313 は、各ラベルについて計算されたカーネル密度推定量 $f^*_h(x)$ の集合を、類似しているか否かの判定に用いられる統計量を表す第 2 データに設定する。

[0036] （ステップ S 2 4 : 統計量送信処理）

データ送信部 314 は、ステップ S 2 1 で座標系が変換された前のデータと後のデータとの軸の対応関係と、ステップ S 2 2 で正規化される前の各軸 i の最小値 $_{min}(x_i)$ 及び最大値 $_{max}(x_i)$ と、ステップ S 2 3 で計算された統計量を表す第 2 データとを検索装置 10 に送信する。すると、検索装置 10 の第 2 取得部 112 は、送信された軸の対応関係と、最小値 $_{min}(x_i)$ 及び最大値 $_{max}(x_i)$ と、第 2 データとを取得して、メモリ 12 に書き込む。

[0037] 図 13 を参照して、実施の形態 1 に係る検索装置 10 の検索処理（図 5 の（4）（5）の処理に相当）を説明する。

(ステップS31：類似判定処理)

類似判定部113は、第1取得部111によって1つ以上の転移元装置20から取得された第1データそれぞれを対象として、対象の第1データと、第2取得部112によって取得された第2データとが類似するか否かを判定する。つまり、類似判定部113は、第1データであるカーネル密度推定量 $f_{h_n}^{(S)}(x)$ の集合と、第2データであるカーネル密度推定量 $f_{h_n}^{(T)}(x)$ の集合とが類似するか否かを判定する。ここで、上付きで示された(S)(T)は転移元装置20と転移先装置30とを区別するための情報であり、(S)は転移元装置20を表し、(T)は転移先装置30を表す。

具体的には、類似判定部113は、カーネル密度推定量 $f_{h_n}^{(S)}(x)$ の集合と、カーネル密度推定量 $f_{h_n}^{(T)}(x)$ の集合とについて、ピアソン相関係数による類似度比較を行う。なお、非特許文献「杉山 将, 山田 誠, ドウ・プレシ マーティヌス・クリストフェル, リウ ソン: 非定常環境下での学習: 共変量シフト適応, クラスバランス変化適応, 変化検知, 日本統計学会誌, vol. 44, no. 1, pp. 113-136 (2014).」には、類似度の評価にカルバック・ライブラー距離、ピアソン距離、又は L^2 距離を用いる方法が記載されている。しかし、IoTでの転移を考える場合、転移先のデータ数が転移元のデータ数に比べ少ない状況($N_{y_i}^{(T)} < N_{y_j}^{(S)}$)が多いと考えられる。そのため、各画素値に対する出現頻度分布に差異が生じ、前述の距離では、類似性を正しく判断できない。そこで、ここでは、類似判定部113は、2データ間の増減の關係に着目し、ピアソン相関係数を使用する。つまり、ここでは、類似判定部113は、対象の第1データと第2データとの増減の關係の類似性に基づき、第1データと第2データとが類似するか否かを判定する。

[0038] まず、類似判定部113は、ピアソン無相関検定で対象の第1データと第2データとの間に相関があるかを検定する。そして、類似判定部113は、検定の結果、無相関ではないと判定した場合、数7に示すように、ピアソン相関係数を類似度とする。一方、類似判定部113は、検定の結果、無相関

であるとは言えない（帰無仮説を棄却できない）場合、類似度を0と定義する。ピアソン無相関検定及び相関係数の計算に用いる標本は、ヒストグラムのbin幅で十分であるので、 $x = 1, \dots, 255$ を代入したときのカーネル密度推定量 $\hat{f}_h^{(T)}(x)$ 及びカーネル密度推定量 $\hat{f}_h^{(S)}(x)$ の値を使用する。

[数7]

$$\text{score}_{(y_k^{(T)}, y_l^{(S)})} = \text{pearsonr}(\hat{f}_h^{(T)}(x)_{y_k}, \hat{f}_h^{(S)}(x)_{y_l})$$

数7では、ラベル y_k に対応する $\hat{f}_h^{(T)}(x)$ が $\hat{f}_h^{(T)}(x)_{y_k}$ 、ラベル y_l に対応する $\hat{f}_h^{(S)}(x)$ が $\hat{f}_h^{(S)}(x)_{y_l}$ と表記されている。ラベル $y_k^{(T)}$ に対応するラベル $y_l^{(S)}$ は、 $\text{score}(y_k^{(T)}, y_l^{(S)})$ が最も高いものとする。

具体的には、検定の結果、無相関ではないと判定した場合、類似判定部113は、第2データにおけるラベル $y_k^{(T)}$ の探索始点を変更しながら、第2データにおける各ラベル $y_k^{(T)}$ と相関係数が高い第1データにおけるラベル $y_l^{(S)}$ を順次特定する。これにより、類似判定部113は、第2データにおける各ラベル $y_k^{(T)}$ に対応する対象の第1データにおけるラベル $y_l^{(S)}$ を特定する。そして、類似判定部113は、対象の第1データと第2データとについて、対応するラベル y_l とラベル y_k との間の最大の相関係数を、対象の第1データと第2データとの類似度とする。なお、類似判定部113は、対応するラベル y_l とラベル y_k との間の相関係数の平均値又は合計値を、対象の第1データと第2データとの類似度としてもよい。

[0039] 類似判定部113は、類似度が閾値Tよりも高い第1データの取得元である転移元装置20のみを転移元の候補とする。あるいは、類似判定部113は、第1データを類似度の高い順にソートして、類似度の高い基準個の第1データの取得元の転移元装置20のみを転移元の候補とする。これにより、類似判定部113は、転移元の候補となる転移元装置20を絞り込む。

[0040] 図14を参照して、実施の形態1に係る無相関ではないと判定された場合

における類似度計算処理を説明する。

ステップS311では、類似判定部113は、 $score_{max}$ に初期値として0を設定する。

ループ1では、類似判定部113は、変数 r を0から $q^{(T)} - 1$ まで1ずつずらしながらステップS312からステップS317までの処理を繰り返し実行する。ここで、 $q^{(T)}$ は、転移先装置30におけるラベル $y^{(T)}$ の種類の数である。つまり、転移先装置30におけるラベル $y^{(T)}$ には、 $\{y_0^{(T)}, \dots, y_{q^{(T)}-1}^{(T)}\}$ の $q^{(T)}$ 種類のラベルが存在する。また、ループ2では、類似判定部113は、 $y_r^{(T)}, y_{1+r}^{(T)}, \dots, y_{(q^{(T)}-1+r) \bmod q^{(T)}}^{(T)}$ の順に、ステップS312からステップS314までの処理を繰り返し実行する。ここで、下付きで表記された $q^{(T)}$ は $q^{(T)}$ を意味する。つまり、ループ1及びループ2では、探索順序を $y_r^{(T)}, y_{1+r}^{(T)}, \dots, y_{(q^{(T)}-1+r) \bmod q^{(T)}}^{(T)}$ の順とし、探索の始点を表す変数 r を0から $q^{(T)} - 1$ まで1ずつずらしながら探索を行うことを意味する。

ステップS312では、類似判定部113は、使用済のラベルの集合 $used$ に初期値として空集合を設定する。

[0041] ループ3では、類似判定部113は、変数 l を0から $q^{(S)}$ まで1ずつずらしながらステップS313の処理を繰り返し実行する。ステップS313では、類似判定部113は、第2データのラベル $y_k^{(T)}$ と、対象の第1データのラベル $y_l^{(S)}$ との間のピアソン相関係数を計算して、 $score(y_k^{(T)}, y_l^{(S)})$ に設定する。

[0042] ステップS314では、類似判定部113は、ラベル $y_l^{(S)}$ が集合 $used$ に含まれていないラベル $y_l^{(S)}$ のうち、 $score(y_k^{(T)}, y_l^{(S)})$ が最大のラベル $y_l^{(S)}$ を、対象のラベル $y_l^{(S)}$ に設定する。類似判定部113は、対象のラベル $y_l^{(S)}$ を集合 $used$ に加える。また、類似判定部113は、処理対象になっているラベル $y_k^{(T)}$ と、対象のラベル $y_l^{(S)}$ との間の $score(y_k^{(T)}, y_l^{(S)})$ を、 $score_{tmp}$ に設定する。また、類似判定部113は、処理対象になっているラベル $y_k^{(T)}$ と、対象のラベル $y_l^{(S)}$

s) との組 $(y_k^{(T)}, y_l^{(S)})$ を集合 g_{tmp} に加える。

[0043] ループ2及びループ3の処理を実行することにより、ループ1で設定された探索順序において、相関係数の高い順に、各ラベル $y_k^{(T)}$ に対応するラベル $y_l^{(S)}$ が特定される。そして、各ラベル $y_k^{(T)}$ と対応するラベル $y_l^{(S)}$ との間の相関係数のうち最も高い相関係数が $score_{tmp}$ に設定される。また、各ラベル $y_k^{(T)}$ と対応するラベル $y_l^{(S)}$ との組が集合 g_{tmp} に設定される。

[0044] ステップS315では、類似判定部113は、 $score_{max}$ よりも $score_{tmp}$ が高いか否かを判定する。類似判定部113は、 $score_{max}$ よりも $score_{tmp}$ が高い場合には処理をステップS316に進め、 $score_{max}$ よりも $score_{tmp}$ が高くない場合には処理をステップS317の後に進める。

ステップS316では、類似判定部113は、 $score_{tmp}$ を $score_{max}$ に設定する。ステップS317では、類似判定部113は、集合 g_{tmp} を集合 g に設定する。

[0045] ループ1からループ3の処理を実行することにより、全ての探索順序で特定された相関係数 $score_{tmp}$ のうち最も高い相関係数 $score_{tmp}$ が相関係数 $score_{max}$ に設定される。ここでは、この相関係数 $score_{max}$ を対象の第1データと第2データとの類似度とする。また、相関係数 $score_{max}$ が計算された探索順序で特定された各ラベル $y_k^{(T)}$ と対応するラベル $y_l^{(S)}$ との組が集合 g に設定される。

[0046] ステップS31で絞り込まれた転移元の候補となる転移元装置20から取得された第1データそれぞれを対象として、ステップS32からステップS34の処理が実行される。

[0047] (ステップS32：ラベル写像生成処理)

写像生成部114は、対象の第1データの元になった訓練データにおけるラベルと、第2データの元になった観測データにおけるラベルとの対応関係を示すラベル写像 g を生成する。

具体的には、写像生成部 114 は、ステップ S31 で特定された各ラベル $y_k^{(T)}$ に対応するラベル $y_l^{(S)}$ を示す集合 g をラベル写像 g として生成する。

[0048] (ステップ S33 : データ写像生成処理)

写像生成部 114 は、対象の第 1 データの元になった訓練データの特徴量ベクトルと、第 2 データの元になった観測データの特徴量ベクトルとの対応関係を示すデータ写像 f を生成する。

具体的には、まず、写像生成部 114 は、対象の第 1 データとともに取得された軸の対応関係と、第 2 データとともに取得された軸の対応関係とに基づき、対象の第 1 データの元になった訓練データの特徴量ベクトルと、第 2 データの元になった観測データの特徴量ベクトルとの対応関係を特定する。対象の第 1 データの元になった訓練データの特徴量ベクトルと、第 2 データの元になった観測データの特徴量ベクトルとの対応関係は、転移先装置 30 の元の座標系 → 転移先装置 30 の基底変換後の座標系 → 転移元装置 20 の基底変換後の座標系 → 転移元装置 20 の元の座標系の順に対応関係を特定することで特定される。

具体例としては、図 15 に示すように、対象の第 1 データとともに取得された軸の対応関係が数 8 に示す関係であり、第 2 データとともに取得された軸の対応関係が数 9 に示す関係であるとする。また、図 15 に示すように、対象の第 1 データの元になった訓練データの特徴量ベクトルの基底変換後のデータと、第 2 データの元になった観測データの特徴量ベクトル基底変換後のデータとの対応関係が数 10 に示す関係であったとする。

[数 8]

$$\left(z_1^{(S)}, z_2^{(S)} \right) \leftrightarrow \left(x_1^{(S)}, x_2^{(S)} \right)$$

[数 9]

$$\left(x_2^{(T)}, x_1^{(T)} \right) \leftrightarrow \left(z_1^{(T)}, z_2^{(T)} \right)$$

[数10]

$$\left(z_1^{(T)}, z_2^{(T)} \right) \leftrightarrow \left(z_1^{(S)}, z_2^{(S)} \right)$$

すると、対象の第1データの元になった訓練データの特徴量ベクトルと、第2データの元になった観測データの特徴量ベクトルとの対応関係Rは、数11に示すようになる。

[数11]

$$\begin{aligned} \left(x_2^{(T)}, x_1^{(T)} \right) &\leftrightarrow \left(z_1^{(T)}, z_2^{(T)} \right) \leftrightarrow \left(z_1^{(S)}, z_2^{(S)} \right) \leftrightarrow \left(x_1^{(S)}, x_2^{(S)} \right) \\ &\Rightarrow \left(x_2^{(T)}, x_1^{(T)} \right) \leftrightarrow \left(x_1^{(S)}, x_2^{(S)} \right) \end{aligned}$$

この対応関係を $R(i) = j$ と表すと、図15の場合には、 $R(2) = 1$ 、 $R(1) = 2$ になる。ここで、変数 i は転移先装置30の軸の添え字 ($x_1^{(T)}$ であれば1) であり、変数 j は転移元装置20の軸の添え字 ($x_2^{(S)}$ であれば2) である。

[0049] そして、写像生成部114は、特定された対応関係Rと、対象の第1データとともに取得された各軸 i の最小値 $\min(x_i^{(S)})$ 及び最大値 $\max(x_i^{(S)})$ と、第2データとともに取得された各軸 i の最小値 $\min(x_i^{(T)})$ 及び最大値 $\max(x_i^{(T)})$ とに基づき、数12に示すように、データ写像 f を生成する。

[数12]

$$\begin{aligned} f = \begin{cases} \mathcal{D}(x_i^{(T)}) = \mathcal{C}(x_i^{(T)}, \min(x_i^{(S)}), \max(x_i^{(S)})) \\ \mathcal{R}(i) = j \end{cases} \\ \left(x_1^{(T)}, \dots, x_{p^{(T)}}^{(T)} \right) \rightarrow \left(\mathcal{D}(x_{\mathcal{R}(1)}^{(T)}), \dots, \mathcal{D}(x_{\mathcal{R}(p^{(T)})}^{(T)}) \right) \end{aligned}$$

数12において、 $p^{(T)}$ は、第2データの元になった観測データの特徴量ベクトル $x^{(T)}$ の次元数である。Cは、数1に定義された通りである。

[0050] (ステップS34：データ送信処理)

データ送信部 115 は、対象の第 1 データについてステップ S 3 2 で生成されたラベル写像 g と、対象の第 1 データについてステップ S 3 3 で生成されたデータ写像 f と、対象の第 1 データの取得元である転移元装置 20 から取得された学習モデルとを、転移先装置 30 に送信する。

すると、データ取得部 315 は、ラベル写像 g とデータ写像 f と学習モデルとを取得する。データ取得部 315 は、ラベル写像 g を出力ラベル変換部 318 に設定し、データ写像 f を入力データ変換部 317 に設定し、学習モデルを学習モデル記憶部 331 に書き込む。

[0051] 図 16 を参照して、実施の形態 1 に係る転移先装置 30 の分析処理（図 5 の（6）～（9）の処理に相当）を説明する。

ここでは、ステップ S 3 1 で絞り込まれた転移元の候補となる転移元装置 20 が 1 つである場合を説明する。

[0052] （ステップ S 4 1：学習モデル生成処理）

学習モデル生成部 316 は、転移先装置 30 用の学習モデルを生成する。ここでは、転移元の候補となる転移元装置 20 が 1 つであるため、学習モデル生成部 316 は、ステップ S 3 4 で取得された学習モデルをそのまま転移先装置 30 用の学習モデルに設定する。

[0053] （ステップ S 4 2：データ変換処理）

入力データ変換部 317 は、センサ 60 から取得された観測データを、ステップ S 3 4 で設定されたデータ写像 f によって変換する。これにより、入力データ変換部 317 は、転移元の候補である転移元装置 20 のデータの形式に観測データの形式を合わせる。つまり、観測データの形式は、転移元装置 20 から取得された学習モデルの入力形式に変換される。

具体例としては、転移先装置 30 の観測データと、各軸の関係が図 15 に示す関係であったとする。この場合には、入力データ変換部 317 は、数 1 1 に示す対応関係 R に従い、数 1 3 に示すように、 $x_1^{(T)}$ 軸を $x_2^{(T)}$ 軸に入れ替え、 $x_2^{(T)}$ 軸を $x_1^{(T)}$ 軸に入れ替えた上で、スケールの変換を行う。

[数13]

$$\left(x_1^{(T)}, x_2^{(T)}\right) \rightarrow \left(\mathcal{D}(x_2^{(T)}), \mathcal{D}(x_1^{(T)})\right)$$

$$\text{s. t. } \mathcal{R}(1) = 2, \mathcal{R}(2) = 1$$

[0054] (ステップS 4 3 : データ入力処理)

入力データ変換部 3 1 7 は、ステップS 4 2 で変換された観測データを、ステップS 4 1 で生成された学習モデルに入力する。すると、学習モデルで推論された結果として出力ラベルが出力される。

[0055] (ステップS 4 4 : 出力ラベル変換処理)

出力ラベル変換部 3 1 8 は、ステップS 4 3 で出力された出力ラベルを、ステップS 3 4 で設定されたラベル写像 g によって変換する。これにより、出力ラベル変換部 3 1 8 は、転移先装置 3 0 のラベルに出力ラベルを変換する。そして、出力ラベル変換部 3 1 8 は、変換された出力ラベルを、観測データから推論された結果として出力する。

具体例としては、ラベル写像 g は $\{(y_k^{(T)}, y_l^{(S)})\}$ で表され、ラベル写像 $g = \{(りんご, 自動車), (みかん, バイク), (ばなな, 自転車)\}$ であったとする。このとき、ステップS 4 3 で出力された出力ラベルがバイクであった場合には、バイクがみかんに変換される。

[0056] つまり、図 1 7 に示すように、実施の形態 1 に係る学習モデル検索システム 1 0 0 は、転移元装置 2 0 が学習モデルを生成した際に使用した訓練データと、転移先装置 3 0 によって得られた少数の観測データとの類似性を判定して、転移元の候補となる転移元装置 2 0 を絞り込む（フェーズ 1）。その後、転移元の候補となる転移元装置 2 0 から、転移元として採用する転移元装置 2 0 が、自動的あるいは人手により抽出される（フェーズ 2）。

[0057] ***実施の形態 1 の効果***

以上のように、実施の形態 1 に係る学習モデル検索システム 1 0 0 は、転移元装置 2 0 の訓練データから生成された統計量と、転移先装置 3 0 の観測データから生成された統計量とに基づき、転送元の候補となる転移元装置 2

0を絞り込む。これにより、短い処理時間で適切な転移元を特定可能である。その結果、短い処理時間で転移先装置30の学習モデルを生成可能である。

[0058] 特に、実施の形態1に係る学習モデル検索システム100は、訓練データとの特徴量ベクトルと観測データの特徴量ベクトルとのそれぞれを、特徴軸毎の情報量を基準として基底変換して得られたデータが類似するか否かを判定することにより、転送元の候補となる転移元装置20を絞り込む。データが類似するか否かを判定する処理は、転移元の訓練データを用いて学習を試行する処理に比べ処理時間がかからない。したがって、短い処理時間で適切な転移元を特定可能になる。

[0059] また、実施の形態1に係る学習モデル検索システム100は、特徴量ベクトルが基底変換された後に、特徴量ベクトルのスケールが正規化されて得られたデータが類似するか否かを判定することにより、転送元の候補となる転移元装置20を絞り込む。これにより、データのスケールの影響を除いた上で比較され、適切な判定が可能になる。

[0060] また、実施の形態1に係る学習モデル検索システム100は、データの増減の関係の類似性に基つき、データが類似するか否かを判定する。これにより、転移先のデータ数が転移元のデータ数に比べ少ない状況であっても、適切な判定が可能になる。

[0061] また、実施の形態1に係る学習モデル検索システム100は、データが類似するか否かの判定に用いる統計量はカーネル密度推定量 $f_n^{\wedge}(x)$ であり、ピアソン相関係数の計算時には、 $x = 1, \dots, 255$ を固定的に用いる。そのため、転移元装置20の訓練データの数に依存せず、計算量を一定にすることが可能である。

[0062] また、実施の形態1に係る学習モデル検索システム100は、統計量である第1データ及び第2データと、転移元装置20の学習モデルとだけが検索装置10に与えられる。そのため、例えば、検索装置10をクラウドコンピューティングのサーバによって実現するような場合であっても、検索装置1

0によって転移元装置20の訓練データが推測されるといったことがなく、安全性が高い。

[0063] ***他の構成***

<変形例1>

実施の形態1では、転移先装置30の分析処理について、ステップS31で絞り込まれた転移元の候補となる転移元装置20が1つである場合を説明した。しかし、ステップS31で絞り込まれた転移元の候補となる転移元装置20が2つ以上である場合もある。

[0064] 図18を参照して、ステップS31で絞り込まれた転移元の候補となる転移元装置20が2つ以上である場合における転移先装置30の分析処理を説明する。

ここでは、一対他分類器 (one-versus-the-rest classifier) の考え方に基づいた処理を説明する。

[0065] (ステップS51:学習モデル生成処理)

学習モデル生成部316は、転移元の候補となる各転移元装置20から取得された学習モデルを、弱学習モデルとして生成する。そして、学習モデル生成部316は、弱学習モデルの組を転移先装置30用の学習モデルとして生成する。

つまり、各転移元装置20から取得された学習モデルは、転移先装置30の一部のラベル識別が可能であると考えられる。そこで、学習モデル生成部316は、各転移元装置20から取得された学習モデルを弱学習モデルとし、弱学習モデルの組を転移先装置30用の学習モデルに設定する。

[0066] (ステップS52:学習モデル選択処理)

入力データ変換部317は、ステップS51で設定された転移先装置30の学習モデルを構成する弱学習モデルのうち、未選択の1つの弱学習モデルを対象の弱学習モデルとして選択する。

なお、入力データ変換部317は、未選択の弱学習モデルが存在しない場合には、観測データを分類不可と判定する。

[0067] (ステップS 5 3 : 入力データ変換処理)

入力データ変換部 3 1 7 は、センサ 6 0 から取得された観測データを、ステップS 5 2 で選択された弱学習モデルの取得元の転移元装置 2 0 についてのデータ写像 f によって変換する。

[0068] (ステップS 5 4 : データ入力処理)

入力データ変換部 3 1 7 は、ステップS 5 3 で変換された観測データを、ステップS 5 2 で選択された弱学習モデルに入力する。すると、学習モデルで推論された結果として出力ラベル又は推論不可との結果が出力される。

[0069] (ステップS 5 5 : 出力判定処理)

入力データ変換部 3 1 7 は、ステップS 5 4 で出力ラベルが出力されたか否かを判定する。

入力データ変換部 3 1 7 は、出力ラベルが出力された場合には、処理をステップS 5 6 に進める。一方、入力データ変換部 3 1 7 は、推論不可との結果が出力された場合には、処理をステップS 5 2 に戻して、他の弱学習モデルを選択する。

[0070] (ステップS 5 6 : 出力ラベル変換処理)

出力ラベル変換部 3 1 8 は、ステップS 5 4 で出力された出力ラベルを、ステップS 5 2 で選択された弱学習モデルの取得元の転移元装置 2 0 についてのラベル写像 g によって変換する。

[0071] 上記処理は、一対他分類器の考え方に基づいた処理である。しかし、これに限らず、一対一分類器 (one-versus-one classifier)、又は、誤り訂正出力符号 (error correcting output code) の考え方に基づく処理としてもよい。

[0072] <変形例 2>

実施の形態 1 では、類似度が閾値よりも高いか否かといった方法により転移元の候補となる転移元装置 2 0 を絞り込んだ。しかし、転移元の候補とするか否かを最終的に人が判断してもよい。この場合には、検索装置 1 0 は、ステップS 1 3 で訓練データが 2 次元画像化された画像データと、ステップ

S 2 3で観測データが2次元画像化された画像データとを表示すればよい。そして、2次元画像化された画像データ同士を目視により比較することで、データが類似しているか否かを人が判定すればよい。

2次元画像化された画像データ同士の比較であるため、人が容易に行うことが可能である。例えば、図19に示すような2次元画像化された画像データが得られる。図19では、転移先装置30のラベル9.0と、転移元装置20のラベル6.0とが類似しており、転移先装置30のラベル10.0と、転移元装置20のラベル9.0とが類似していることが分かる。

[0073] <変形例3>

実施の形態1では、統計量の比較をする際、ピアソン相関係数を用いた。しかし、統計量の比較をする際、画像識別の技術を使用してもよい。具体例としては、類似判定部113は、訓練データが2次元画像化された画像データと、観測データが2次元画像化された画像データとのそれぞれから特徴点を抽出する。そして、類似判定部113は、訓練データが2次元画像化された画像データにおける特徴点同士の距離と、観測データが2次元画像化された画像データにおける特徴点同士の距離とを比較することが考えられる。

[0074] <変形例4>

実施の形態1では、転移元装置20が第1データを生成した上で、第1データを検索装置10に送信した。しかし、転移元装置20は訓練データを検索装置10に送信して、検索装置10が第1データを生成してもよい。この場合には、転移元装置20が備える基底変換部211と正規化部212と統計量計算部213との機能構成要素を検索装置10が備えるようにすればよい。

同様に、実施の形態1では、転移先装置30が第2データを生成した上で、第2データを検索装置10に送信した。しかし、転移先装置30は観測データを検索装置10に送信して、検索装置10が第2データを生成してもよい。この場合には、転移先装置30が備える基底変換部311と正規化部312と統計量計算部313との機能構成要素を検索装置10が備えるように

すればよい

[0075] なお、検索装置10に訓練データを送信する場合には、検索装置10に訓練データが漏洩することになる。同様に、検索装置10に観測データを送信する場合には、検索装置10に観測データが漏洩することになる。したがって、訓練データ又は観測データを外部に漏らしたくない場合には、実施の形態1の構成とすることが望ましい。

[0076] <変形例5>

実施の形態1では、各機能構成要素がソフトウェアで実現された。しかし、変形例5として、各機能構成要素はハードウェアで実現されてもよい。この変形例5について、実施の形態1と異なる点を説明する。

[0077] 各機能構成要素がハードウェアで実現される場合には、検索装置10は、プロセッサ11とメモリ12とストレージ13とに代えて、電子回路15を備える。電子回路15は、各機能構成要素と、メモリ12と、ストレージ13との機能とを実現する専用の回路である。

[0078] 同様に、各機能構成要素がハードウェアで実現される場合には、転移元装置20は、プロセッサ21とメモリ22とストレージ23とに代えて、電子回路25を備える。電子回路25は、各機能構成要素と、メモリ22と、ストレージ23との機能とを実現する専用の回路である。

[0079] 同様に、各機能構成要素がハードウェアで実現される場合には、転移先装置30は、プロセッサ31とメモリ32とストレージ33とに代えて、電子回路35を備える。電子回路35は、各機能構成要素と、メモリ32と、ストレージ33との機能とを実現する専用の回路である。

[0080] 電子回路15, 25, 35としては、単一回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ロジックIC、GA (Gate Array)、ASIC (Application Specific Integrated Circuit)、FPGA (Field-Programmable Gate Array) が想定される。

検索装置10と転移元装置20と転移先装置30との各装置は、各機能構

成要素を1つの電子回路15, 25, 35で実現してもよいし、各機能構成要素を複数の電子回路15に分散させて実現してもよい。

[0081] <変形例6>

変形例6として、検索装置10と転移元装置20と転移先装置30との各装置は、一部の各機能構成要素がハードウェアで実現され、他の各機能構成要素がソフトウェアで実現されてもよい。

[0082] プロセッサ11, 21, 31とメモリ12, 22, 32とストレージ13, 23, 33と電子回路15, 25, 35とを処理回路という。つまり、各機能構成要素の機能は、処理回路により実現される。

[0083] 実施の形態2.

実施の形態2は、2次元画像化された画像データに代えて、m次元主成分空間のベクトル \hat{z} の要素 \hat{z}_i 毎の確率密度推定量を統計量として用いる点の実施の形態1と異なる。実施の形態2では、この異なる点を説明して、同一の点については説明を省略する。

[0084] ***動作の説明***

図6を参照して、実施の形態2に係る転移元装置20の第1データ送信処理を説明する。

ステップS12では、正規化部212は、 $z_{min}=0$, $z_{max}=1$ として、ベクトル z を正規化してベクトル \hat{z} を生成する。

ステップS13では、統計量計算部213は、数14に示すように、ベクトル \hat{z} の要素 \hat{z}_i 毎のカーネル密度推定量 $\hat{f}_h(x)$ を用いて、確率密度関数を推定する。

[数14]

$$\hat{f}_h(x) = \frac{1}{|\hat{z}_i|h} \sum_{\hat{x} \in \hat{z}_i} K\left(\frac{x - \hat{x}}{h}\right)$$

数14において、 $|\hat{z}_i|$ は、ベクトル \hat{z} の第i主成分軸のデータ総数である。

[0085] 図12を参照して、実施の形態2に係る転移先装置30の第2データ送信処理を説明する。

ステップS22では、図6のステップS12と同様に、正規化部312は、 $z_{min}=0$ 、 $z_{max}=1$ として、ベクトル z を正規化してベクトル z^{\wedge} を生成する。

ステップS23では、図6のステップS13と同様に、統計量計算部313は、ベクトル z^{\wedge} の要素 z^{\wedge}_i 毎のカーネル密度推定量 $f^{\wedge}_h(x)$ を用いて、確率密度関数を推定する。

[0086] 図13を参照して、実施の形態2に係る検索装置10の検索処理を説明する。

ステップS31では、類似判定部113は、数15に示すように、要素 z^{\wedge}_i の寄与率 PV_i で重み付けしたピアソン相関係数を類似度とする。ここで、ピアソン無相関検定及び相関係数の計算に用いる標本は、 $x=0, 0.001, \dots, 1$ を代入したときのカーネル密度推定量 $f^{\wedge}_h^{(T)}(x)$ 及びカーネル密度推定量 $f^{\wedge}_h^{(S)}(x)$ の値を使用する。

[数15]

$$\begin{aligned} & score(y_k^{(T)}, y_l^{(S)}) \\ &= \sum_{i=1}^{\min(m^{(T)}, m^{(S)})} PV_i^{(T)} \times pearsonr(\hat{f}_h^{(T)}(x)_{y_k}, \hat{f}_h^{(S)}(x)_{y_l}) \end{aligned}$$

言い換えると、類似判定部113は、各特徴軸を対象として、対象の特徴軸についての第1データと第2データとの増減の関係の類似性（ピアソン相関係数）に対して、対象の特徴軸の情報量に応じて重み付け（寄与率 PV_i で重み付け）して得られた結果を線形結合して、第1データと第2データとが類似するか否かを判定する。

[0087] 図20を参照して実施の形態2に係る類似判定処理を説明する。

類似度判定処理は、ループ3の処理が図14に示す処理と異なる。ループ3では、ループ4の処理が実行される。ループ4では、類似判定部113は

、変数 i が 1 から $\min(m^{(T)}, m^{(S)})$ まで 1 ずつずらしながらステップ S 3 1 3 の処理を繰り返し実行する。ステップ S 3 1 3 では、類似判定部 1 1 3 は、要素 z_i^{\wedge} の寄与率 $P V_i^{(T)}$ で重み付けした、第 2 データのラベル $y_k^{(T)}$ と対象の第 1 データの $y_l^{(S)}$ との間のピアソン相関係数を計算して、 $score(y_k^{(T)}, y_l^{(S)})$ に加算する。

[0088] ***実施の形態 2 の効果***

以上のように、実施の形態 2 に係る学習モデル検索システム 1 0 0 は、特徴量ベクトルを基底変換して無相関化させた上で、ベクトルの要素毎の類似性を線形結合して類似するか否かを判定する。これにより、実施の形態 1 に比べて計算量を減らすことが可能になる。

[0089] また、実施の形態 2 に係る学習モデル検索システム 1 0 0 は、ベクトルの要素毎の類似性に対して、寄与率で重み付けする。これにより、機械学習における出力に重要な影響を与える要素が似ているほど、類似性が高いと判定されるようになり、適切な判定が可能になる。

[0090] また、実施の形態 2 に係る学習モデル検索システム 1 0 0 は、ベクトルの要素毎に外挿（確率密度推定）を行うことにより、適切な判定が可能になる。

[0091] ***他の構成***

<変形例 7>

実施の形態 2 では、確率密度関数の推定にカーネル密度推定量を用いた。しかし、より計算量の少ない線形外挿又は直線外挿のような線形補間の技術を使用したアルゴリズムを用いてもよい。想定される定義域内のデータを万遍なく採取可能な場合のように、共変量シフトやクラスバランス変化を考慮しなくてよい場合、外挿ではなく線形補間又は多項式補間を用いてもよい。

[0092] 実施の形態 3.

実施の形態 3 は、 m 次元主成分空間のベクトル z^{\wedge} の要素 z_i^{\wedge} 毎に統計的仮説検定を用いる点の実施の形態 2 と異なる。実施の形態 3 では、この異なる点を説明して、同一の点については説明を省略する。

[0093] ***動作の説明***

図6を参照して、実施の形態3に係る転移元装置20の第1データ送信処理を説明する。

ステップS12では、正規化部212は、実施の形態2と同様に、 $z_{min}=0$ 、 $z_{max}=1$ として、ベクトル z を正規化してベクトル z^{\wedge} を生成する。

ステップS13では、統計量計算部213は、統計量を計算しない。統計量計算部213は、統計的仮説検定における検定精度の低下を抑制するために、外れ値又はノイズの除去と、データの補間又は外挿とを行う。

[0094] 図12を参照して、実施の形態3に係る転移先装置30の第2データ送信処理を説明する。

ステップS22では、図6のステップS12と同様に、正規化部312は、 $z_{min}=0$ 、 $z_{max}=1$ として、ベクトル z を正規化してベクトル z^{\wedge} を生成する。

ステップS23では、図6のステップS13と同様に、統計量計算部313は、統計的仮説検定における検定精度の低下を抑制するために、外れ値又はノイズの除去と、データの補間又は外挿とを行う。

[0095] 図13を参照して、実施の形態3に係る検索装置10の検索処理を説明する。

ステップS31では、類似判定部113は、統計的仮説検定により類似度を計算する。統計的仮説検定では、帰無仮説 H_0 及び対立仮説 H_1 が定められ、 H_0 を棄却することにより H_1 が採択される。ここでは、類似判定部113は、検定結果から類似度を計算するため、 H_0 が棄却された場合を0、棄却できない場合を1と定義し、検定結果を2値化する。但し、検定結果が1であっても H_0 を採択しないことに注意する。検定の標本は、 $(z^{\wedge}_i^{(T)})_{y_k}$ 、 $(z^{\wedge}_i^{(S)})_{y_l}$ を用いる。下付きの y_k 及び y_l は、ラベル y_k 及びラベル y_l に対応する特徴量ベクトル z^{\wedge} の要素 z^{\wedge}_i であることを示す。

類似判定部113は、数16に示すように、実施の形態2と同様に、検定結果に寄与率 PV_i で重み付けすることにより類似度を算出する。

[数16]

$$\begin{aligned} & score(y_k^{(T)}, y_l^{(S)}) \\ &= \sum_{i=1}^{\min(m^{(T)}, m^{(S)})} \{PV_i^{(T)} \cdot Test\left(\left(\hat{z}_i^{(T)}\right)_{y_k}, \left(\hat{z}_i^{(S)}\right)_{y_l}\right)\} \\ Test &= \begin{cases} 1 & , H_0 \text{を棄却できない場合} \\ 0 & , H_0 \text{を棄却する場合} \end{cases} \end{aligned}$$

数16において、Testは検定結果を2値化した値である。

言い換えると、類似判定部113は、各特徴軸を対象として、対象の特徴軸についての第1データと第2データとの類似性を統計的仮説検定により特定する。そして、類似判定部113は、特定された類似性に対して、対象の特徴軸の情報量に応じて重み付けして得られた結果を線形結合して、第1データと第2データとが類似するか否かを判定する。

[0096] 図21を参照して実施の形態3に係る類似判定処理を説明する。

類似度判定処理は、ステップS313の処理が図20と異なる。ステップS313では、類似判定部113は、ラベル $y_k^{(T)}$ に対応する要素 $\hat{z}_i^{(T)}$ とラベル $y_l^{(S)}$ に対応する要素 $\hat{z}_i^{(S)}$ との間の統計的仮説検定の検定結果に、要素 \hat{z}_i の寄与率 $PV_i^{(T)}$ で重み付けして、 $score(y_k^{(T)}, y_l^{(S)})$ に加算する。

[0097] なお、検定方法を選定する場合には、転移元装置20及び転移先装置30の特定によって、以下のような条件を考慮する必要がある。

- (1) 正規性を仮定できない
- (2) サンプル数が異なる（独立2標本，対応のない標本）

(1) (2)の条件が成立する場合は、図22に示す対応のないノンパラメトリック検定を使用する。対応のないノンパラメトリック検定には、マンホイットニのU検定と2標本コルモゴロフスミルノフ検定とがある。マンホイットニのU検定では、帰無仮説 H_0 を「両標本が同じ母集団から抽出された」とし、対立仮説 H_1 を「両標本が異なる母集団から抽出された」とする。

2標本コルモゴロフスミルノフ検定では、帰無仮説 H_0 を「両標本の母集団の確率分布が等しい」とし、対立仮説 H_1 を「両標本の母集団の確率分布が等しくない」とする。

[0098] 転移元装置20及び転移先装置30の特性によっては、データに対応があること、又は、正規分布のようなある分布に従うことが想定できる場合がある。このような場合は、パラメトリック検定を利用してもよい。

[0099] ***実施の形態3の効果***

以上のように、実施の形態3に係る学習モデル検索システム100は、統計的仮説検定により、類似するか否かを判定する。これにより、入力サンプルではなく、入力サンプルの母集団の類似性を厳密に判定することができるため、適切な判定が可能になる。

[0100] また、実施の形態3に係る学習モデル検索システム100は、基底変換及び正規化を行って得られたベクトル \hat{z} を用いて統計的仮説検定を行う。これにより、入力ベクトルの要素毎に検定を行うことができるため、高次元の入力ベクトルに対しても、既存の低次元の統計的仮説検定の手法を用いることが可能である。

[0101] 実施の形態4.

実施の形態4は、2次元画像化された画像データに代えて、 m 次元主成分空間のベクトル \hat{z} の平均ベクトルのコサイン類似度を統計量として用いる点の実施の形態1と異なる。実施の形態4では、この異なる点を説明して、同一の点については説明を省略する。

[0102] ***動作の説明***

図6を参照して、実施の形態4に係る転移元装置20の第1データ送信処理を説明する。

ステップS12では、正規化部212は、 $z_{min}=0$ 、 $z_{max}=1$ として、ベクトル z を正規化してベクトル \hat{z} を生成する。

ステップS13では、統計量計算部213は、数17に示すように、ベクトル \hat{z} の代表値として相加平均ベクトル \hat{z} を計算する。

[数17]

$$\vec{\hat{z}} = \frac{\sum \vec{z}}{|\vec{z}|}$$

数17において、 $|\vec{z}|$ は、特徴量ベクトル \vec{z} の総数 (N_{y_x}) である。

[0103] 図12を参照して、実施の形態4に係る転移先装置30の第2データ送信処理を説明する。

ステップS22では、図6のステップS12と同様に、正規化部312は、 $z_{min}=0$ 、 $z_{max}=1$ として、ベクトル \vec{z} を正規化してベクトル $\vec{\hat{z}}$ を生成する。

ステップS23では、図6のステップS13と同様に、統計量計算部313は、ベクトル $\vec{\hat{z}}$ の代表値として相加平均ベクトル $\vec{\hat{z}}$ を計算する。

[0104] 図13を参照して、実施の形態2に係る検索装置10の検索処理を説明する。

ステップS31では、類似判定部113は、数18に示すように、相加平均ベクトル $\vec{\hat{z}}^{(T)}$ と相加平均ベクトル $\vec{\hat{z}}^{(S)}$ とのコサイン類似度を計算する。

[数18]

$$\begin{aligned} score_{(y_k^{(T)}, y_l^{(S)})} &= \cos \left(\left(\frac{\vec{\hat{z}}^{(T)}}{y_k} \right), \left(\frac{\vec{\hat{z}}^{(S)}}{y_l} \right) \right) \\ &= \frac{\sum_{i=1}^{\min(m^{(T)}, m^{(S)})} \left\{ \left(\frac{\vec{\hat{z}}_i^{(T)}}{y_k} \right) \cdot \left(\frac{\vec{\hat{z}}_i^{(S)}}{y_l} \right) \right\}}{\sqrt{\sum_{i=1}^{\min(m^{(T)}, m^{(S)})} \left\{ \left(\frac{\vec{\hat{z}}_i^{(T)}}{y_k} \right)^2 \right\}} \cdot \sqrt{\sum_{i=1}^{\min(m^{(T)}, m^{(S)})} \left\{ \left(\frac{\vec{\hat{z}}_i^{(S)}}{y_l} \right)^2 \right\}}} \end{aligned}$$

言い換えると、類似判定部113は、第1データ及び第2データについて代表値を計算して、代表値により第1データと第2データとが類似するか否かを判定する。特に、類似判定部113は、第1データについての代表値と、第2データについての代表値とのコサイン類似度を計算することにより、

第1データと第2データとが類似するか否かを判定する。

[0105] 図23を参照して実施の形態4に係る類似判定処理を説明する。

類似度判定処理は、ステップS313の処理が図14に示す処理と異なる。ステップS313では、類似判定部113は、相加平均ベクトル $\vec{z}^{(T)}$ と相加平均ベクトル $\vec{z}^{(S)}$ とのコサイン類似度を計算して、 $score(y_k^{(T)}, y_l^{(S)})$ に設定する。

[0106] ***実施の形態4の効果***

以上のように、実施の形態4に係る学習モデル検索システム100は、ベクトル \vec{z} の平均ベクトルのコサイン類似度により、類似するか否かを判定する。これにより、入力サンプル数によらず一度の比較で類似するか否かを判定することができるため、検索速度を一定に維持することが可能になる。

[0107] ***他の構成***

<変形例8>

実施の形態4では、代表値として相加平均ベクトルを用いた。しかし、代表値として、刈込平均と、中央値と、四分位点と、重心と、最頻値と、k近傍といった値を用いてもよい。

[0108] なお、上記説明において、数19に示すベクトルを本文中で \vec{z} と表記する。また、数20に示す正規化されたベクトルを本文中で $\hat{\vec{z}}$ と表記する。また、数21に示す相加平均ベクトルを \vec{z} と表記する。また、本文中で x_y と表記した場合には、 x_y を意味する。

[数19]

\vec{z}

[数20]

$\hat{\vec{z}}$

[数21]

\vec{z}

[0109] 以上、この発明の実施の形態及び変形例について説明した。これらの実施の形態及び変形例のうち、いくつかを組み合わせ実施してもよい。また、いずれか1つ又はいくつかを部分的に実施してもよい。なお、この発明は、以上の実施の形態及び変形例に限定されるものではなく、必要に応じて種々の変更が可能である。

符号の説明

[0110] 100 学習モデル検索システム、10 検索装置、11 プロセッサ、12 メモリ、13 ストレージ、14 通信インタフェース、15 電子回路、111 第1取得部、112 第2取得部、113 類似判定部、114 写像生成部、115 データ送信部、131 学習モデル記憶部、132 統計量記憶部、20 転移元装置、21 プロセッサ、22 メモリ、23 ストレージ、24 通信インタフェース、25 電子回路、211 基底変換部、212 正規化部、213 統計量計算部、214 データ送信部、231 学習モデル記憶部、232 訓練データ記憶部、30 転移先装置、31 プロセッサ、32 メモリ、33 ストレージ、34 通信インタフェース、35 電子回路、311 基底変換部、312 正規化部、313 統計量計算部、314 データ送信部、315 データ取得部、316 学習モデル生成部、317 入力データ変換部、318 出力レベル変換部、40 伝送路、50 センサ、60 センサ。

請求の範囲

- [請求項1] 転移元装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第1データを取得する第1取得部と、
転移先装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第2データを取得する第2取得部と、
前記第1取得部によって取得された前記第1データと、前記第2取得部によって取得された前記第2データとが類似するか否かを判定する類似判定部と
を備える検索装置。
- [請求項2] 前記第1データ及び前記第2データは、前記特徴量ベクトルが基底変換された後に、前記特徴量ベクトルのスケールが正規化されて得られた
請求項1に記載の検索装置。
- [請求項3] 前記第1データ及び前記第2データは、前記特徴量ベクトルが正規化された後に、2次元画像化された画像データの画素値の分布の統計量が計算されて得られた
請求項2に記載の検索装置。
- [請求項4] 前記類似判定部は、前記第1データと前記第2データとの増減の関係の類似性に基づき、前記第1データと前記第2データとが類似するか否かを判定する
請求項3に記載の検索装置。
- [請求項5] 前記第1データ及び前記第2データは、前記特徴量ベクトルが正規化された後に、前記特徴軸毎の値の分布の統計量が計算されて得られた
請求項2に記載の検索装置。
- [請求項6] 前記類似判定部は、各特徴軸を対象として、対象の特徴軸についての前記第1データと前記第2データとの増減の関係の類似性に対して、前記対象の特徴軸の情報量に応じて重み付けして得られた結果を線

形結合して、前記第1データと前記第2データとが類似するか否かを判定する

請求項5に記載の検索装置。

[請求項7] 前記類似判定部は、各特徴軸を対象として、対象の特徴軸についての前記第1データと前記第2データとの類似性を統計的仮説検定により特定して、前記類似性に対して、前記対象の特徴軸の情報量に応じて重み付けして得られた結果を線形結合して、前記第1データと前記第2データとが類似するか否かを判定する

請求項2に記載の検索装置。

[請求項8] 前記類似判定部は、前記第1データ及び前記第2データについて代表値を計算して、前記代表値により前記第1データと前記第2データとが類似するか否かを判定する

請求項2に記載の検索装置。

[請求項9] 前記類似判定部は、前記第1データについての前記代表値と、前記第2データについての前記代表値とのコサイン類似度を計算することにより、前記第1データと前記第2データとが類似するか否かを判定する

請求項8に記載の検索装置。

[請求項10] 前記検索装置は、さらに、

前記第1データと前記第2データとが類似すると前記類似判定部によって判定された場合に、前記第1データを生成した際の基底変換と、前記第2データが生成された際の基底変換とに基づき、前記転移先装置における特徴量ベクトルを前記転移元装置における特徴量ベクトルに合わせるためのデータ写像を生成する写像生成部

を備える請求項1から9までのいずれか1項に記載の検索装置。

[請求項11] 前記転移元装置における特徴量ベクトルと、前記転移先装置における特徴量ベクトルとは、要素毎にラベルが付されており、

前記写像生成部は、前記第1データと前記第2データとの類似度に

に基づき、前記第1データのラベルと前記第2データのラベルとの間の対応関係を示すラベル写像を生成する

請求項10に記載の検索装置。

- [請求項12] 第1取得部が、転移元装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第1データを取得し、
- 第2取得部が、転移先装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第2データを取得し、
- 類似判定部が、前記第1データと前記第2データとが類似するか否かを判定する検索方法。

- [請求項13] 転移元装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第1データを取得する第1取得処理と、
- 転移先装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第2データを取得する第2取得処理と、
- 前記第1取得処理によって取得された前記第1データと、前記第2取得処理によって取得された前記第2データとが類似するか否かを判定する類似判定処理と
- を行う検索装置としてコンピュータを機能させる検索プログラム。

- [請求項14] 検索装置と転移先装置とを備える学習モデル検索システムであり、
- 前記検索装置は、
- 転移元装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第1データを取得する第1取得部と、
- 前記転移先装置における特徴量ベクトルを、特徴軸毎の情報量を基準として基底変換して得られた第2データを取得する第2取得部と、
- 前記第1取得部によって取得された前記第1データと、前記第2取得部によって取得された前記第2データとが類似するか否かを判定する類似判定部と
- を備え、
- 前記転移先装置は、前記類似判定部によって前記第1データと前記

第2データとが類似すると判定された場合に、前記転移元装置の学習モデルに基づき学習モデルを生成する学習モデル生成部を備える学習モデル検索システム。

[請求項15] 前記第1取得部は、複数の転移元装置それぞれを対象として、対象の転移元装置についての前記第1データを取得し、

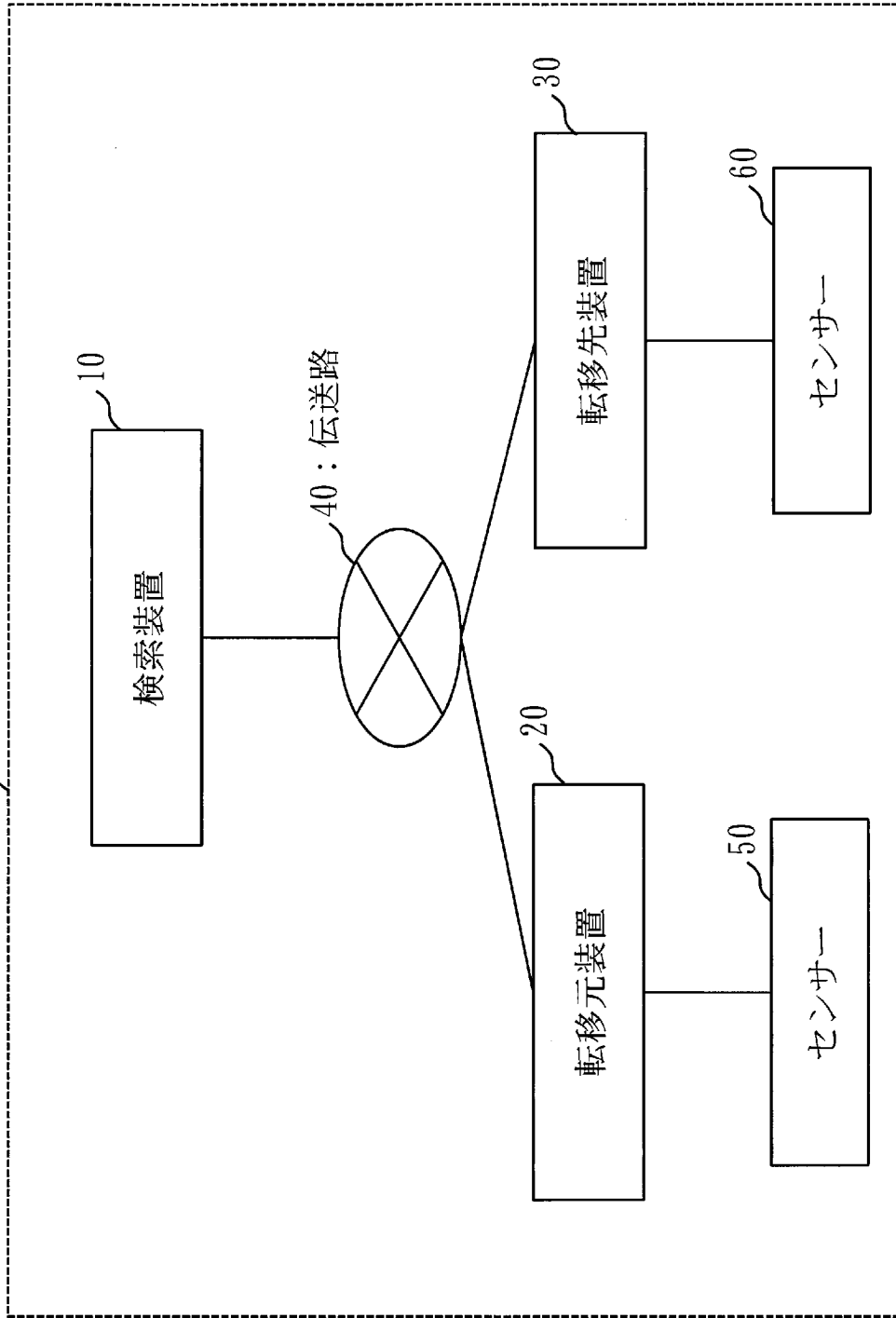
前記類似判定部は、前記複数の転移元装置それぞれを対象として、対象の転移元装置についての前記第1データと、前記第2データとが類似するか否かを判定し、

前記学習モデル生成部は、2つ以上の転移元装置についての前記第1データと前記第2データとが類似すると判定された場合には、前記2つ以上の転移元装置の学習モデルに基づき学習モデルを生成する請求項14に記載の学習モデル検索システム。

[図1]

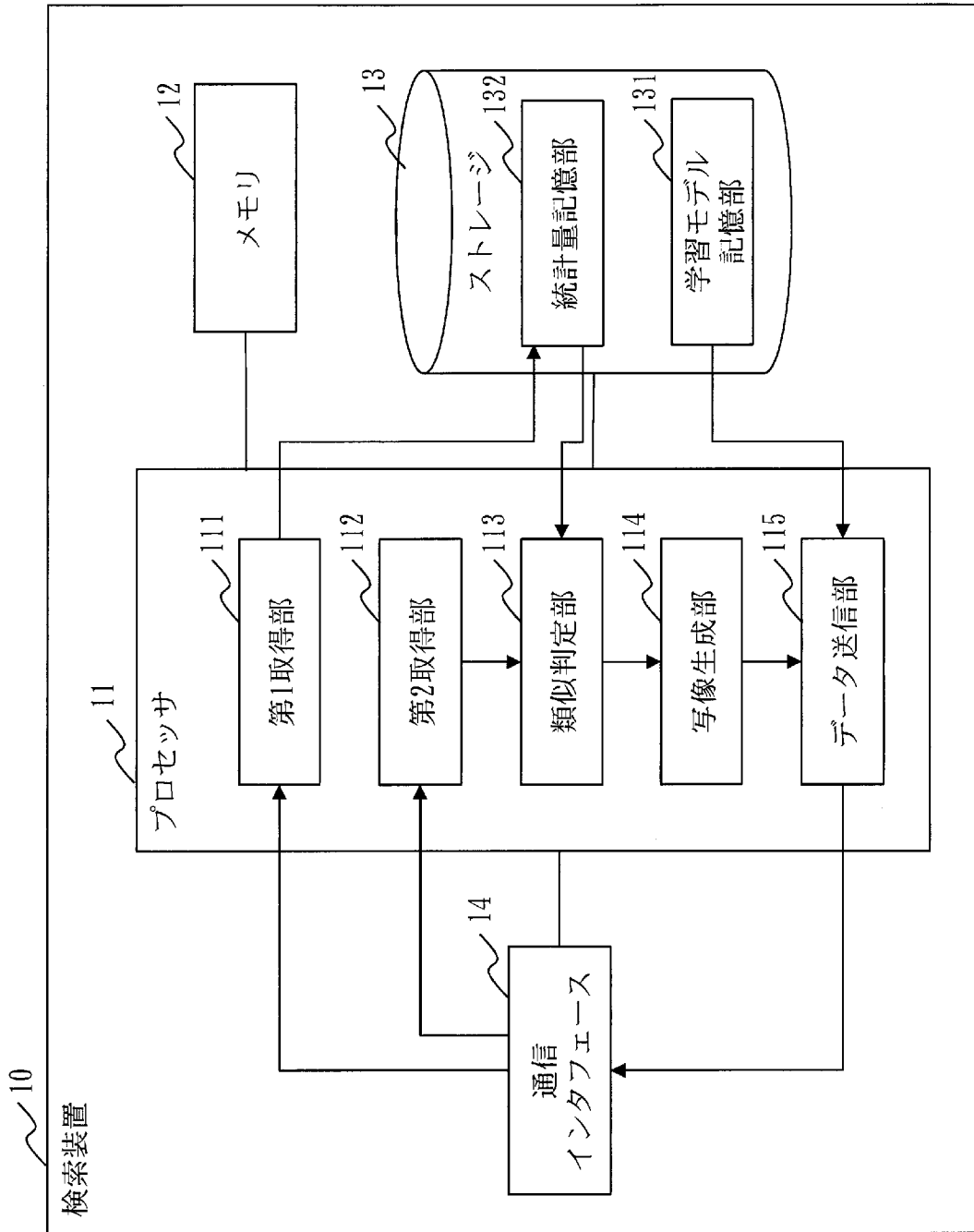
図1

100:学習モデル検索システム



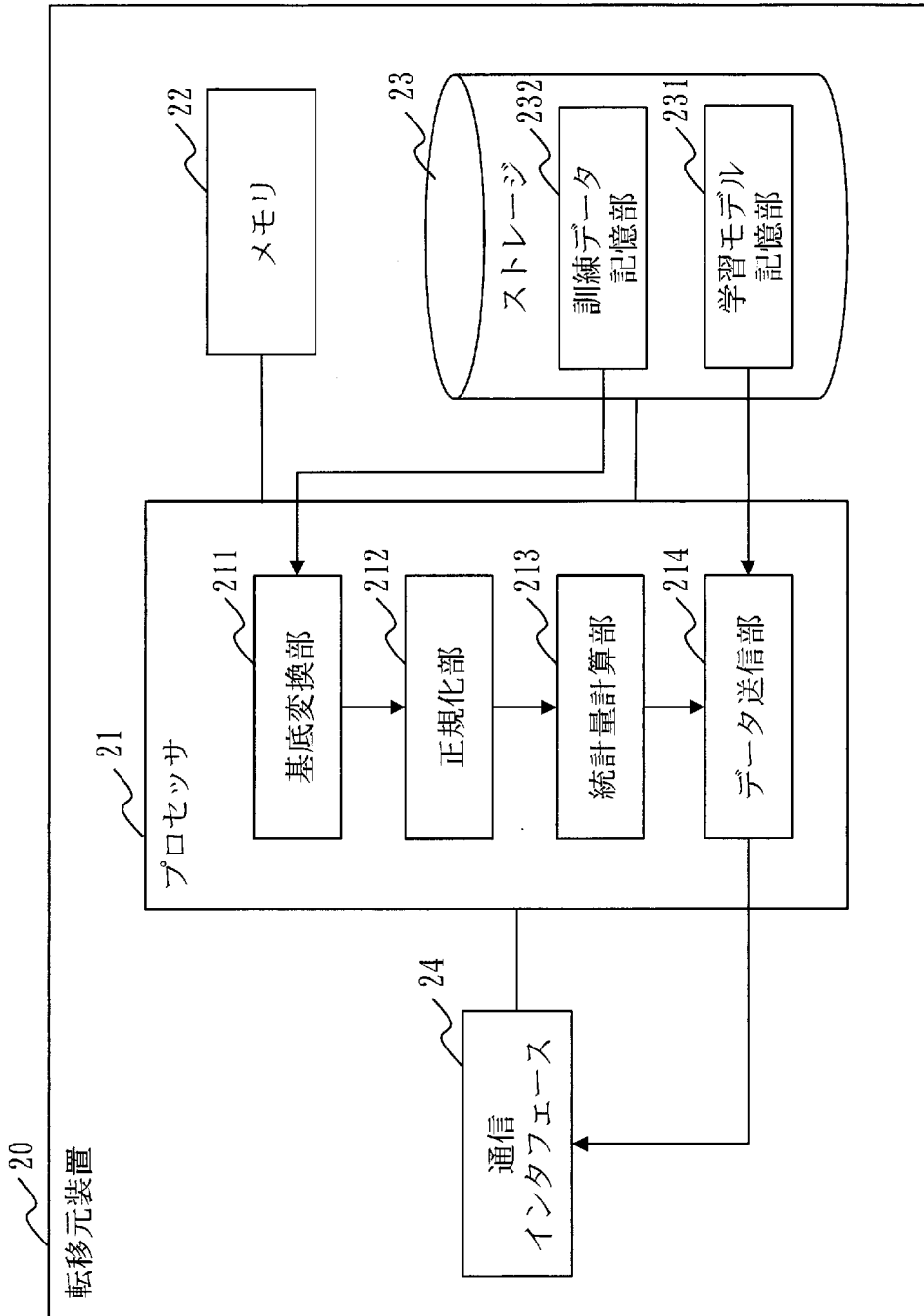
[図2]

図2



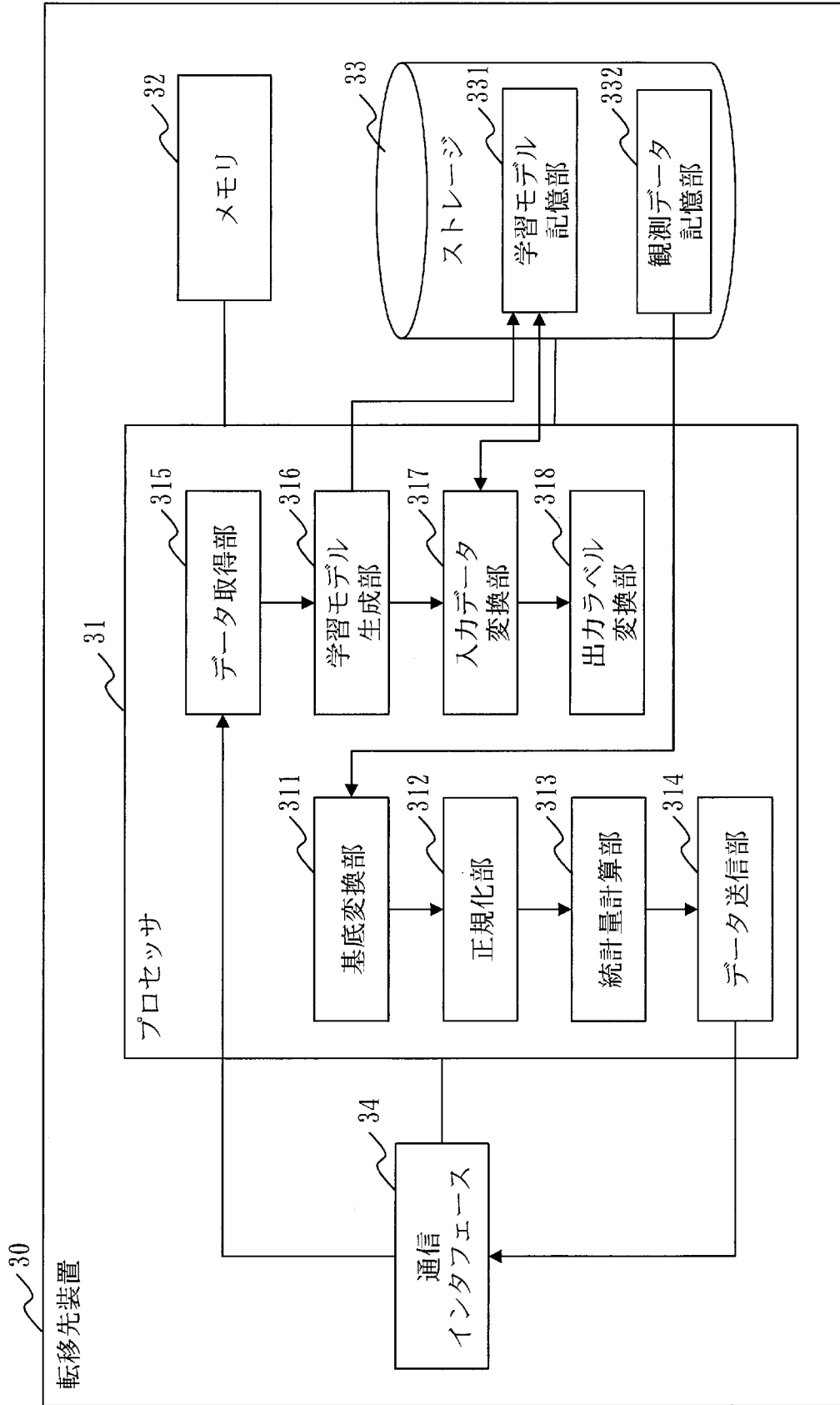
[図3]

図3



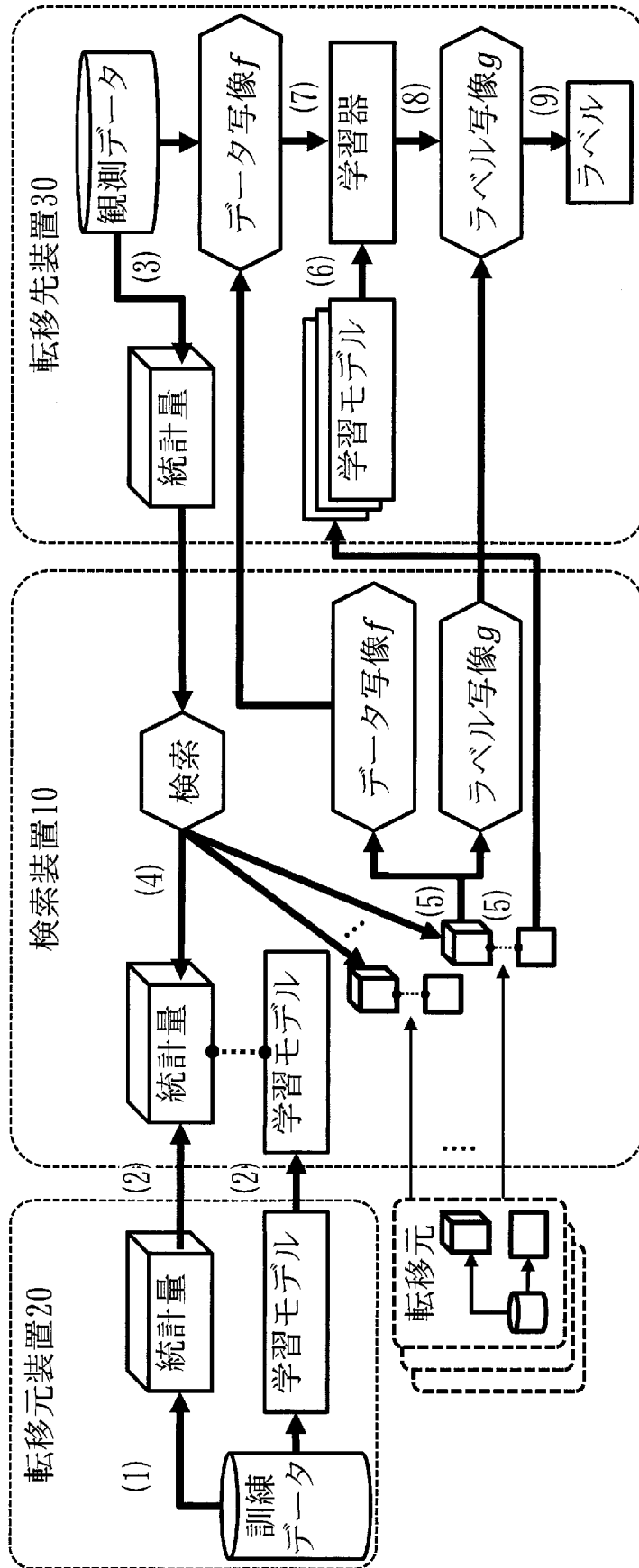
[図4]

図4



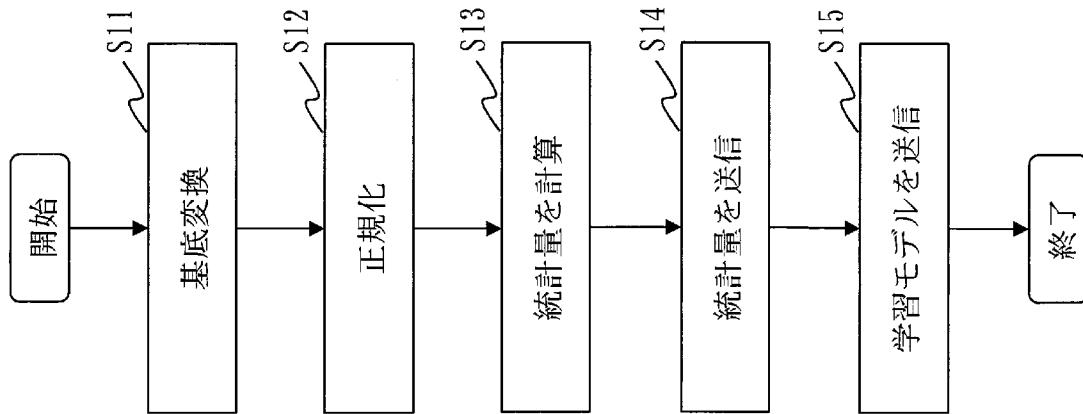
[図5]

図5



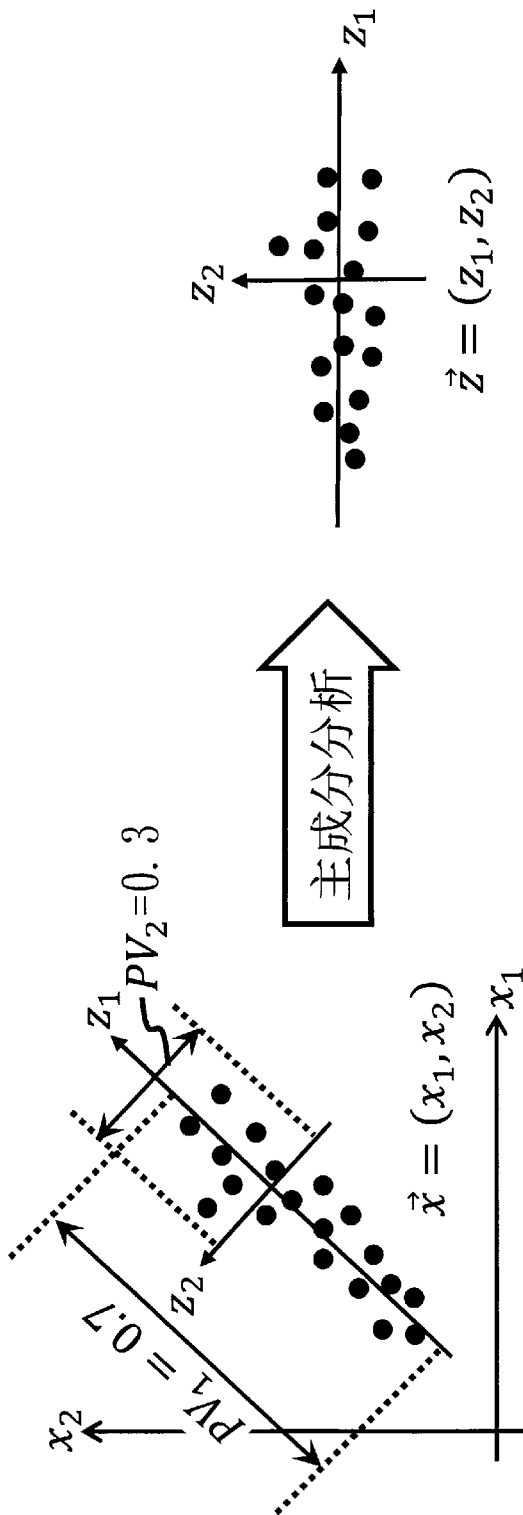
[図6]

図6



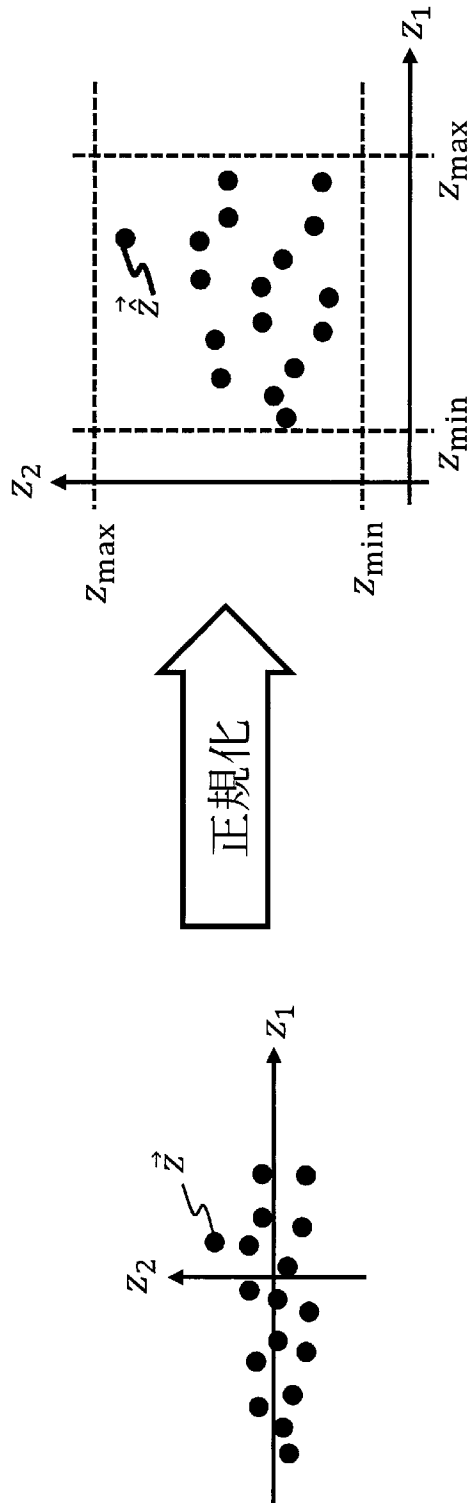
[図7]

図7



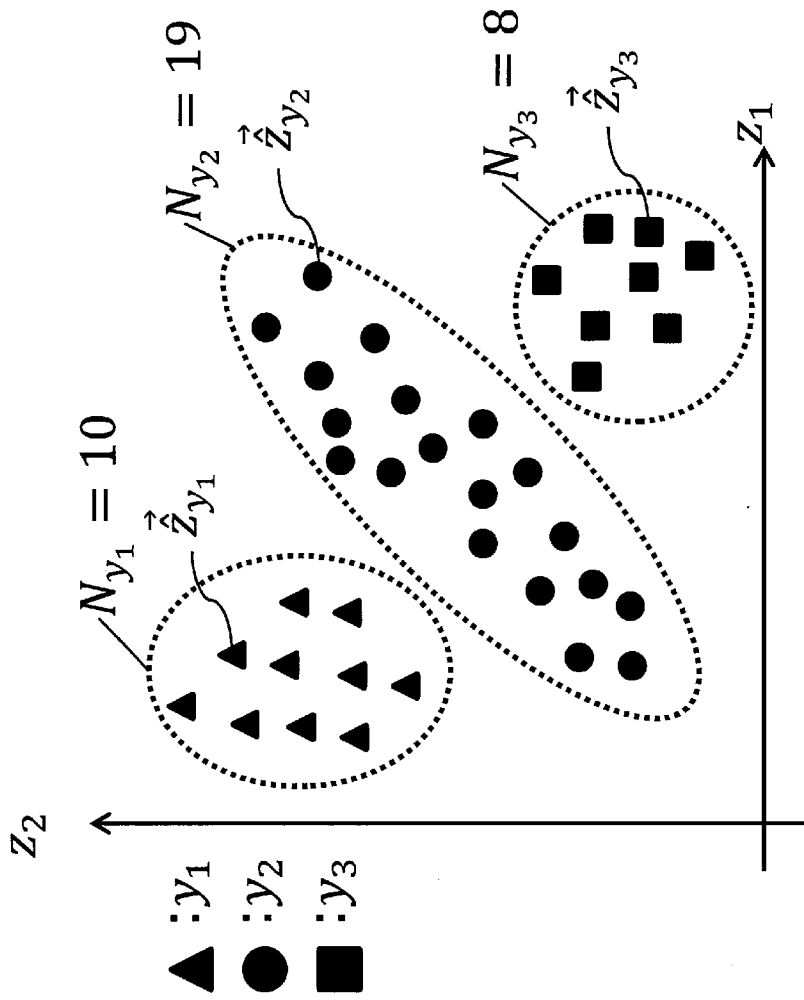
[図8]

図8



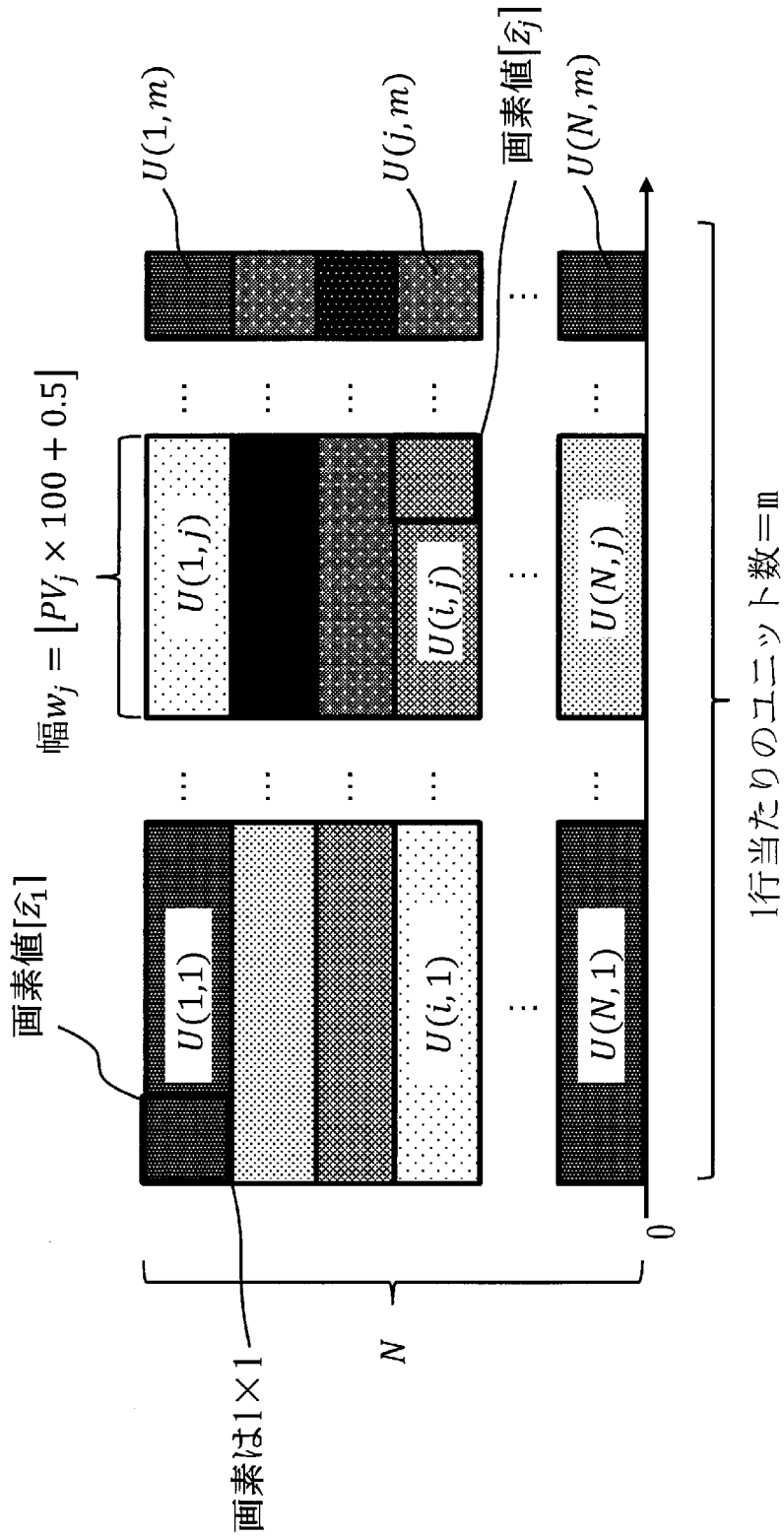
[図9]

図9



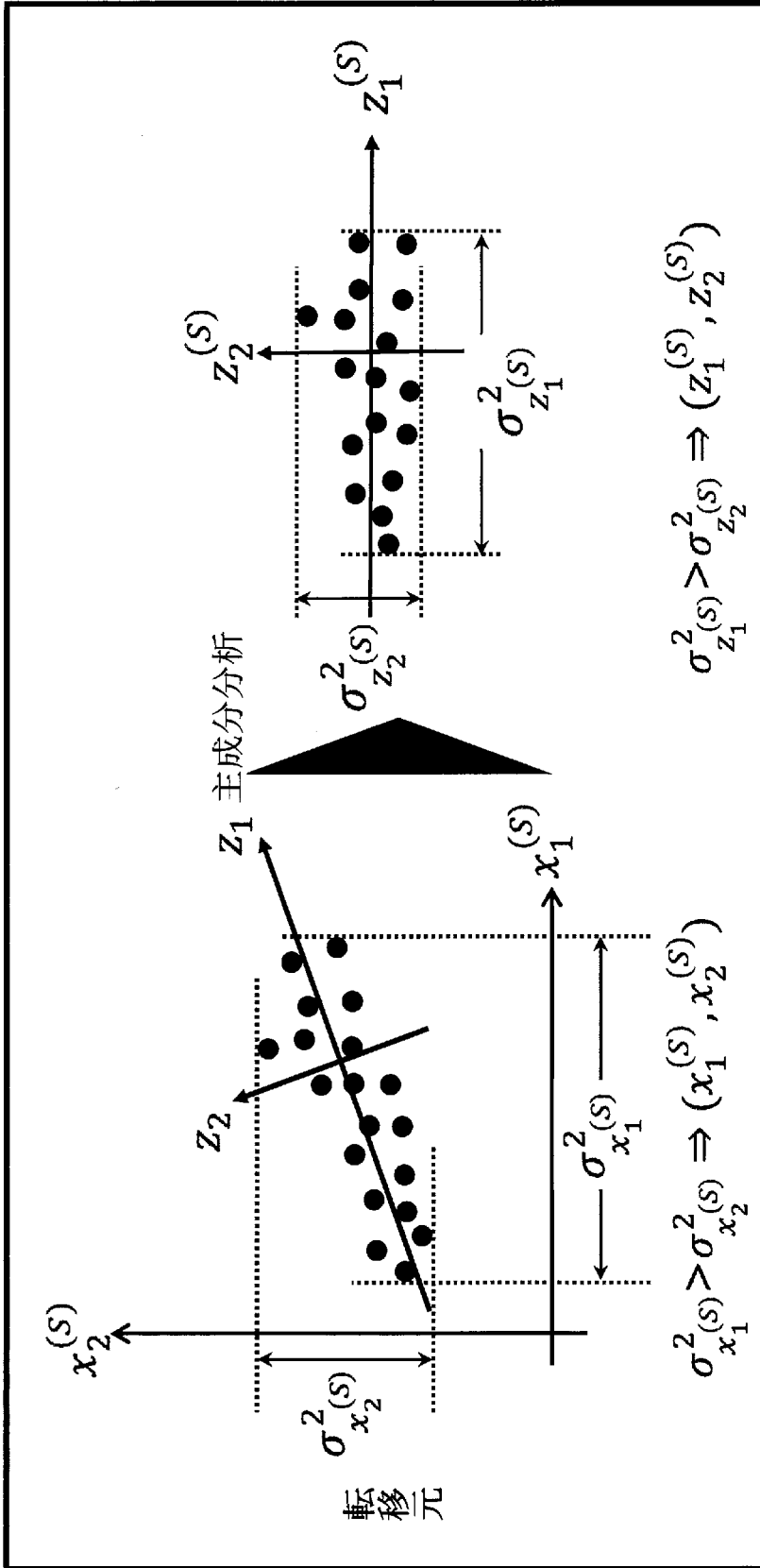
[図10]

図10



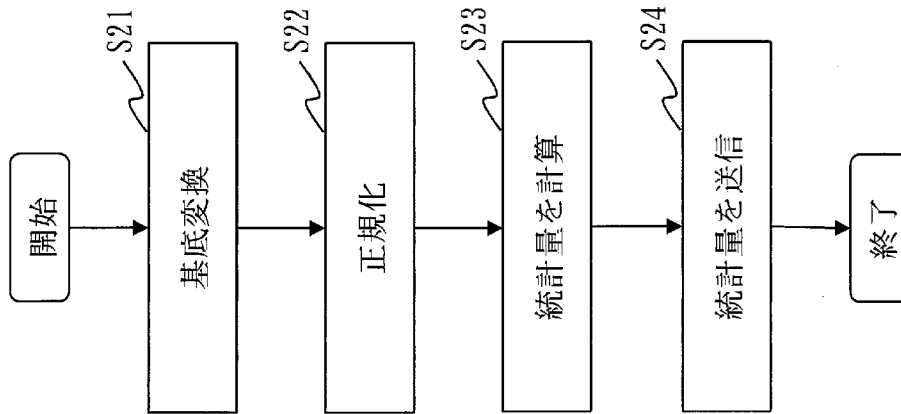
[図11]

図11



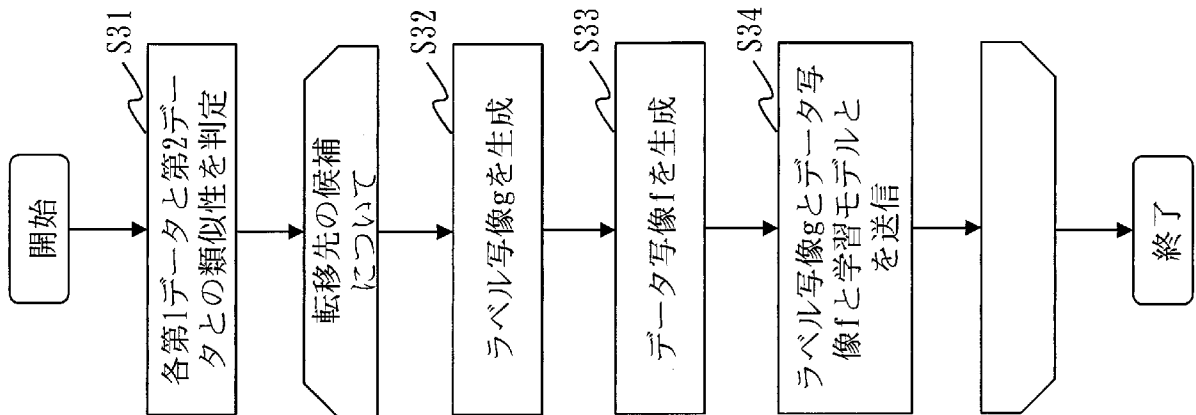
[図12]

図12



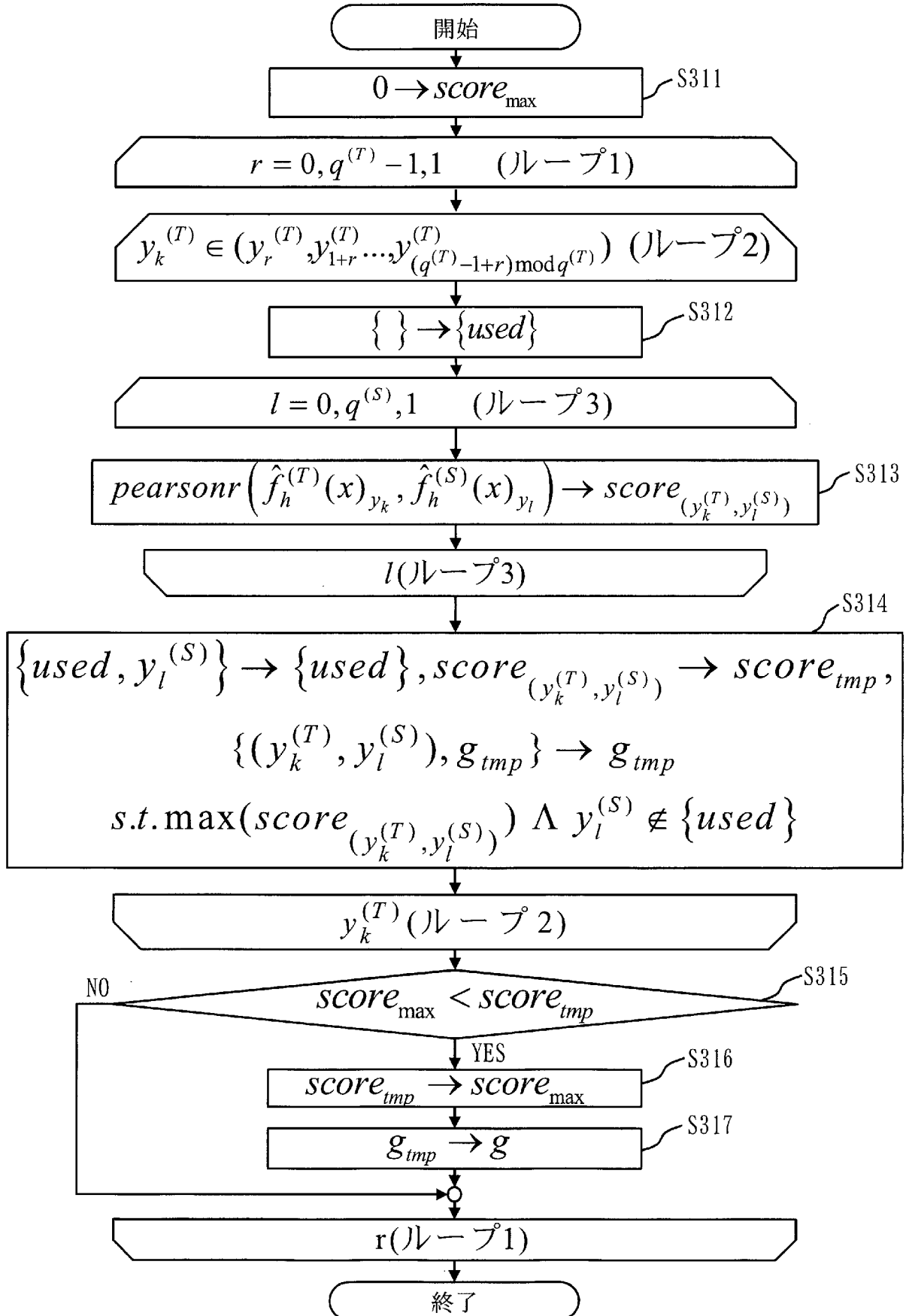
[図13]

図13



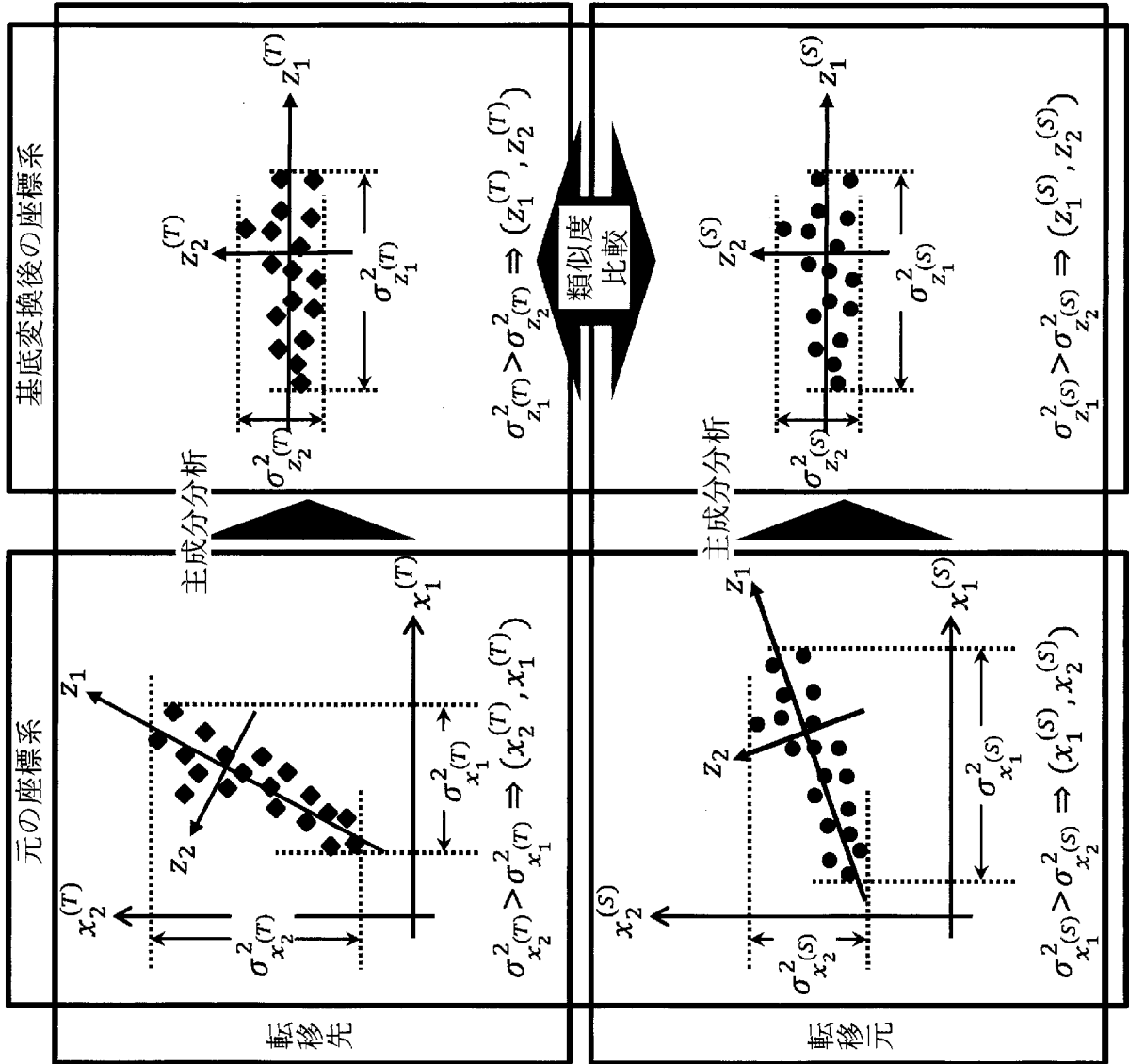
[図14]

図 14



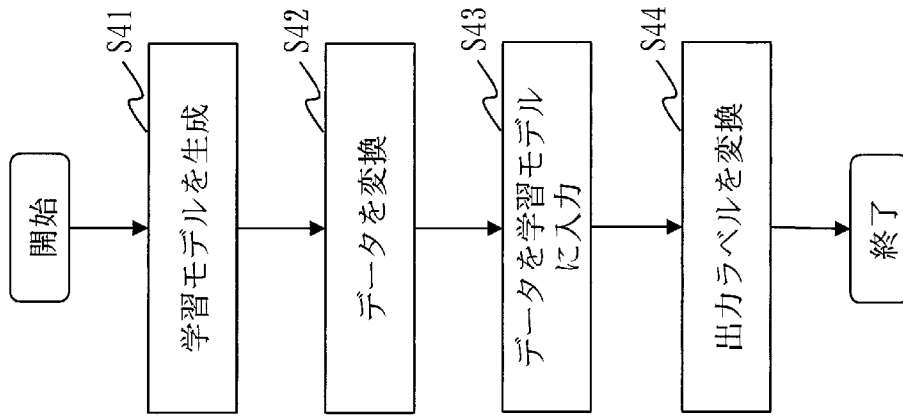
[図15]

図15



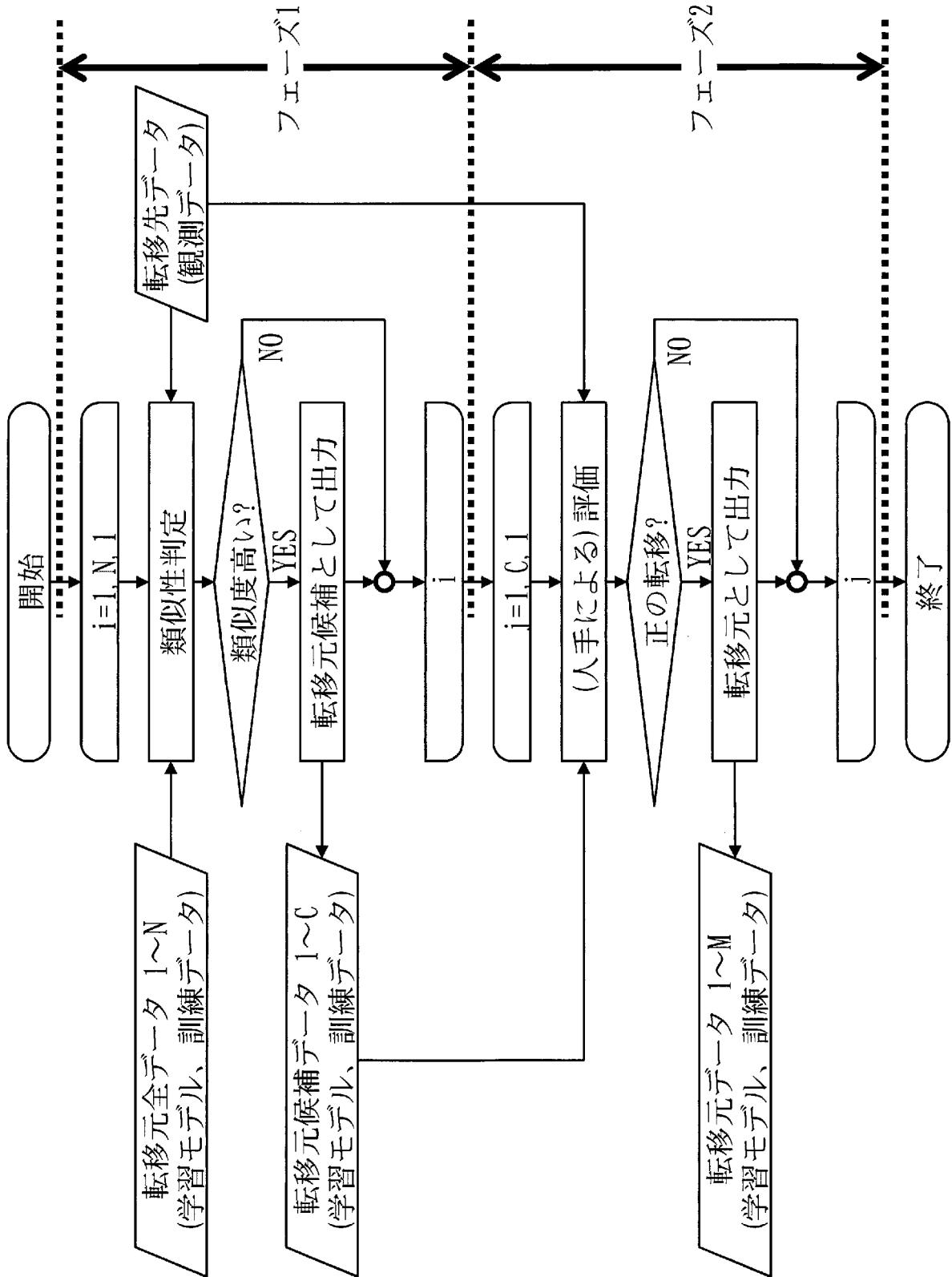
[図16]

図16



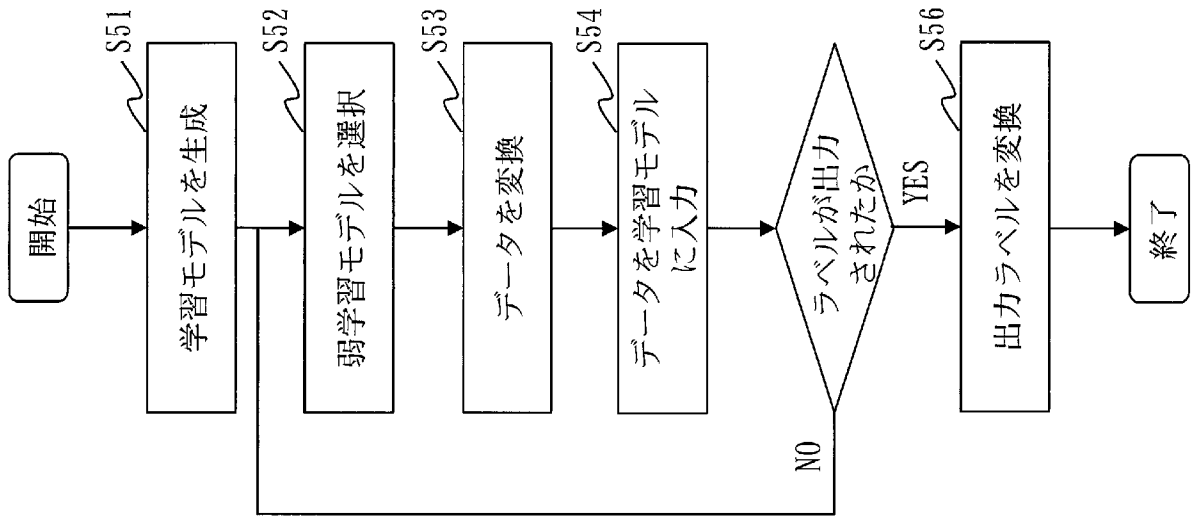
[図17]

図17



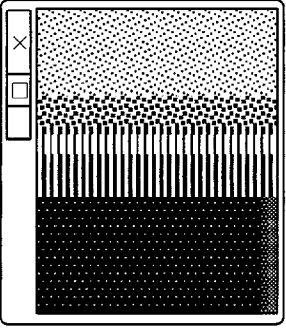
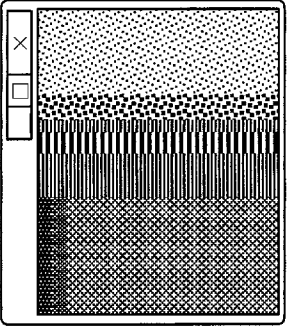
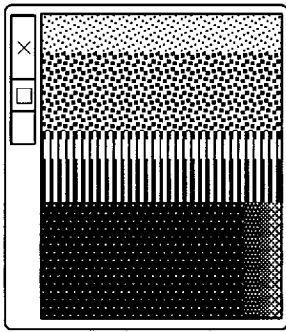
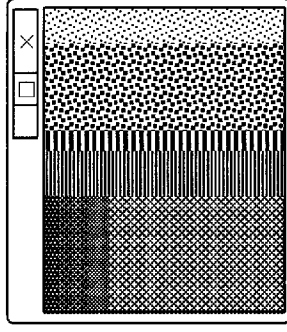
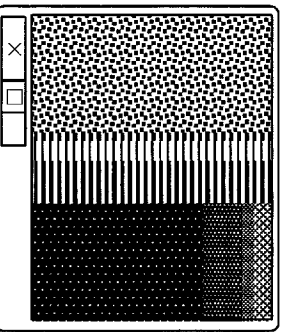
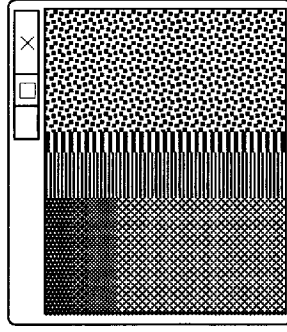
[図18]

図18



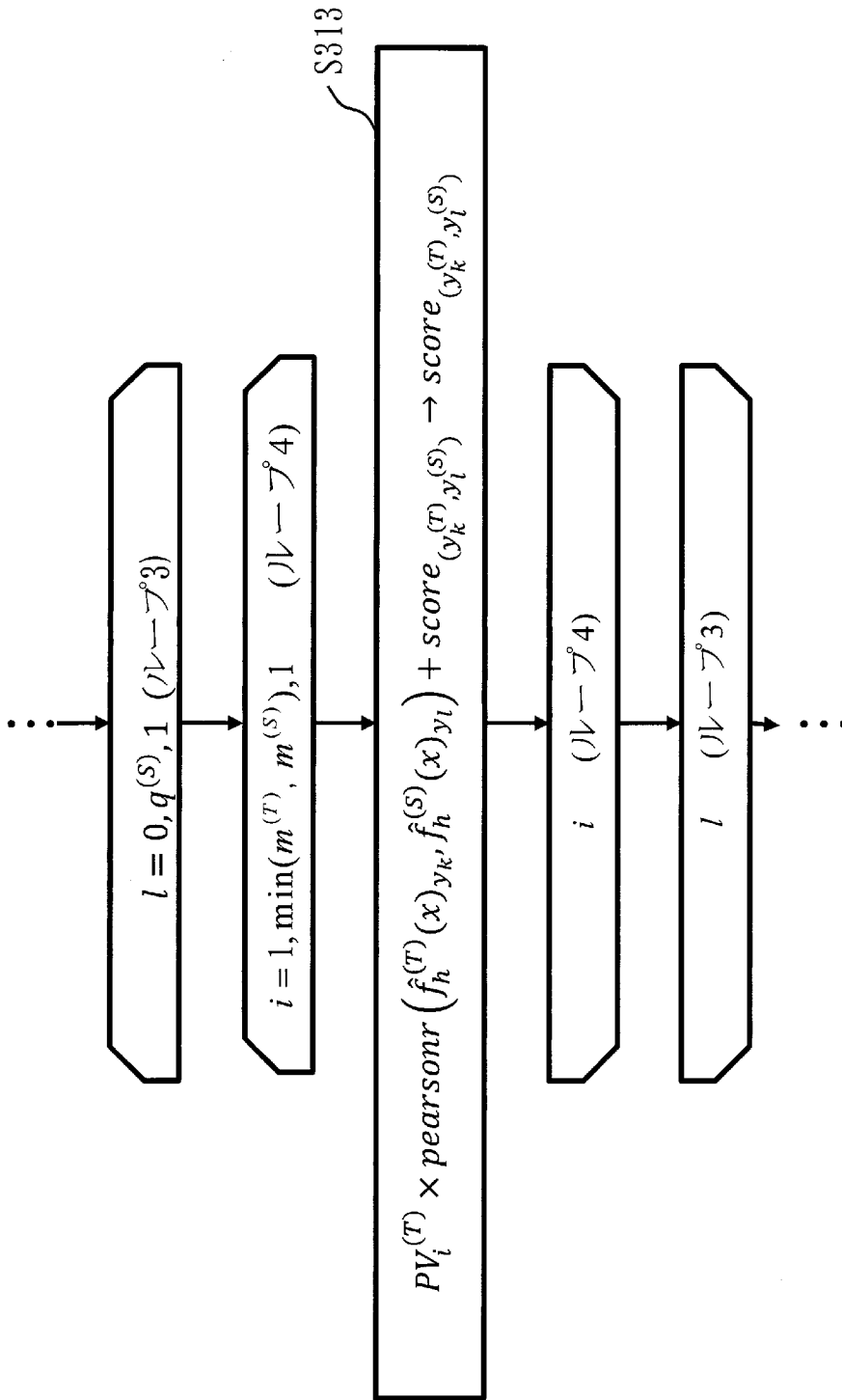
[図19]

図19

転移元	$N^{(S)} = 952$		
		ラベル6.0	ラベル9.0
転移先	$N^{(T)} = 570$		
		ラベル9.0	ラベル10.0
	$N^{(T)} = 353$		
		ラベル9.0	ラベル10.0

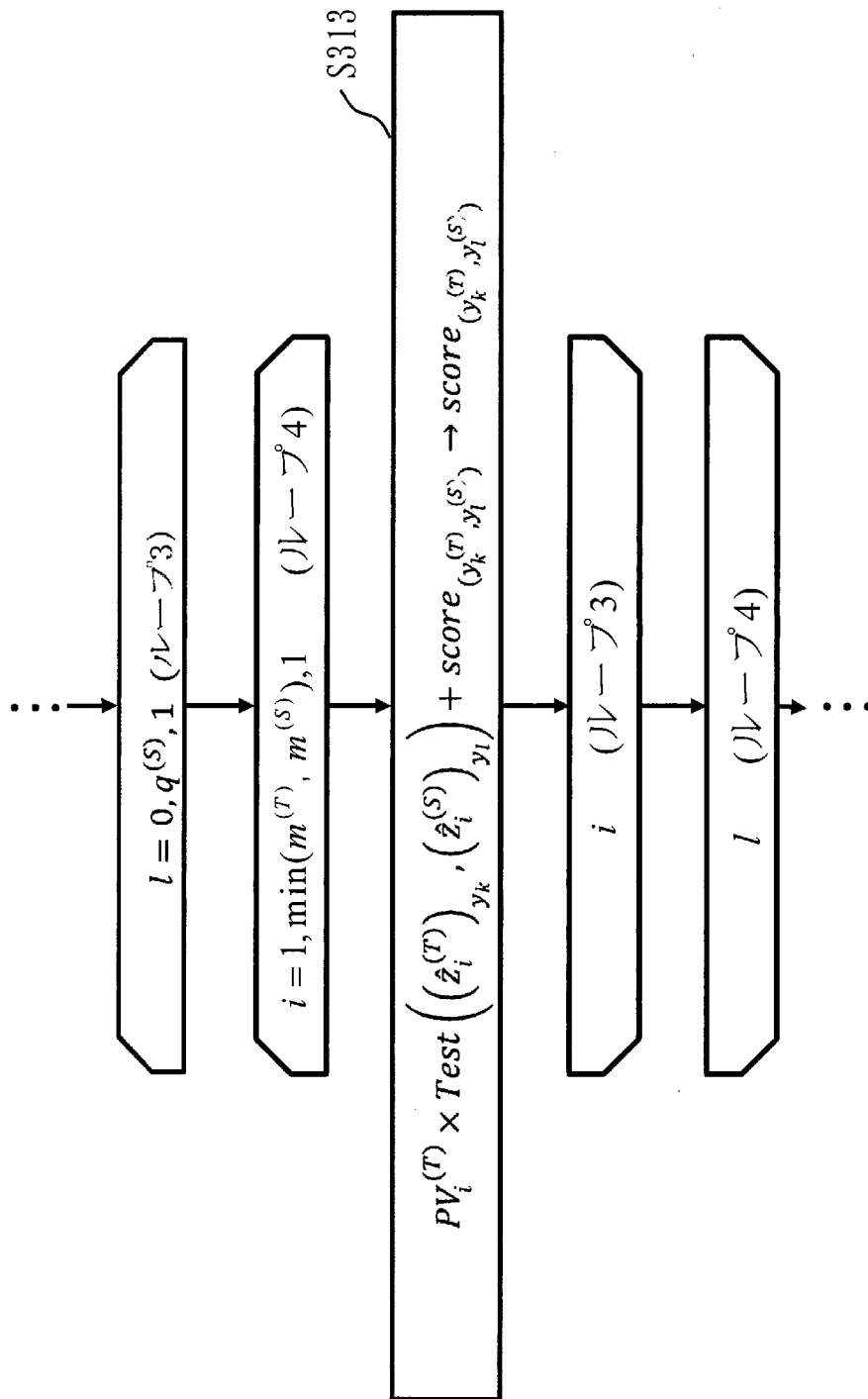
[図20]

図20



[図21]

図21



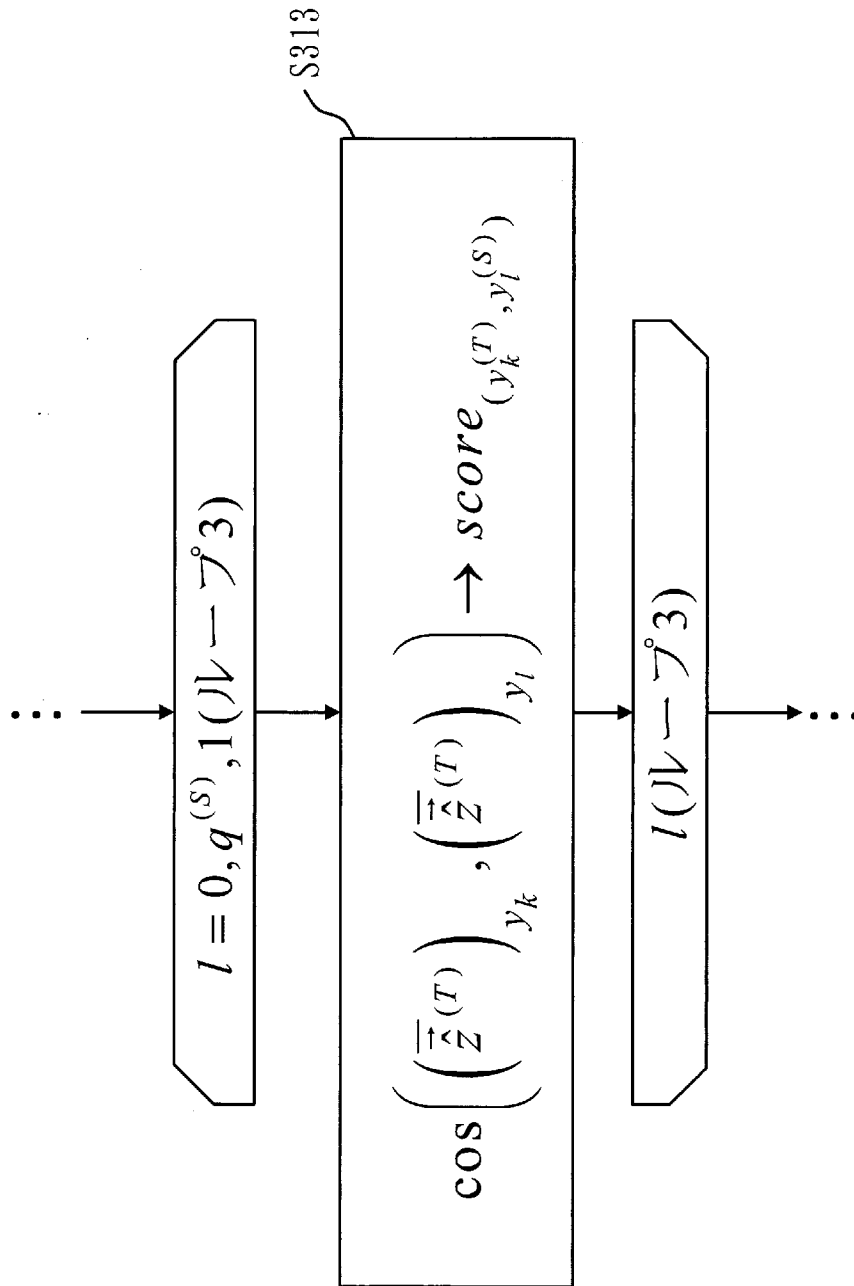
[図22]

図22

		検定法 (例)	
パラメトリック検定	2群の平均値の差の検定	対応あり	対応のあるt検定
		対応なし	ウェルチのt検定
ノンパラメトリック検定	2群の代表値の差の検定	対応あり	ウィルコクソンの符号順位和検定 符号検定
		対応なし	マンホイットニのU検定 (ウィルコクソンの順位和検定)
			2標本コルモゴロフ-スミルノフ (KS) 検定

[図23]

図23



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2019/040614

A. CLASSIFICATION OF SUBJECT MATTER

Int.Cl. G06N20/00 (2019.01) i, G06F16/90 (2019.01) i

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)

Int.Cl. G06N3/00-99/00, G06F16/00-16/958, G06T7/00, G10L15/00-17/26

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	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

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 2017-224156 A (CANON INC.) 21 December 2017, abstract, paragraphs [0056]-[0061], fig. 8 & US 2017/0364826 A1, abstract, [paragraphs [0100]-[0110], fig. 8 & EP 3258425 A1 & CN 107527063 A	1-15
A	US 2016/0253597 A1 (XEROX CORPORATION) 01 September 2016, entire text, all drawings (in particular, paragraphs [0043], [0050], [0072]) (Family: none)	1-15

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
20 November 2019 (20.11.2019)

Date of mailing of the international search report
03 December 2019 (03.12.2019)

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

International application No.

PCT/JP2019/040614

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	森 郁海 ほか, 分散エッジ環境における機械学習実現最適化の検討, 情報処理学会研究報告 マルチメディア通信と分散処理(DPS) [online], 24 January 2019, vol. 2019-DPS-177, no. 2, pp. 1-8, [retrieval date 28 January 2019], Internet: <URL:https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_uri&item_id=194230&file_id=1&file_no=1>, ISSN: 2188-8906, in particular, chapter 6, (MORI, Ikumi et al., "A Study on Optimisation of Machine Learning Realization In Edge Computing", IPSJ SIG technical reports: Distributed processing system (DPS) [online])	1-15

A. 発明の属する分野の分類 (国際特許分類 (IPC))

Int.Cl. G06N20/00(2019.01)i, G06F16/90(2019.01)i

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))

Int.Cl. G06N3/00-99/00, G06F16/00-16/958, G06T7/00, G10L15/00-17/26

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	1971-2019年
日本国実用新案登録公報	1996-2019年
日本国登録実用新案公報	1994-2019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 2017-224156 A (キヤノン株式会社) 2017.12.21, [要約], 段落[0056]-[0061], 図8 & US 2017/0364826 A1, 要約, 段落[0100]-[0110], 図8 & EP 3258425 A1 & CN 107527063 A	1-15
A	US 2016/0253597 A1 (XEROX CORPORATION) 2016.09.01, 全文, 全図 (特に段落[0043], [0050], [0072]) (ファミリーなし)	1-15

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矛盾するものではなく、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

20.11.2019

国際調査報告の発送日

03.12.2019

国際調査機関の名称及びあて先

日本国特許庁 (ISA/J P)
 郵便番号100-8915
 東京都千代田区霞が関三丁目4番3号

特許庁審査官 (権限のある職員)

多賀 実

電話番号 03-3581-1101 内線 3545

5B

9367

C (続き) . 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	森 郁海 ほか, 分散エッジ環境における機械学習実現最適化の検討, 情報処理学会研究報告 マルチメディア通信と分散処理 (D P S) [online], 2019.01.24, 第 2019-DPS-177 巻, 第 2 号, pp.1-8, [検索日 2019.01.28], インターネット: <URL:https://ipsj.ixsq.nii.ac.jp/ej/?action=repository_uri &item_id=194230&file_id=1&file_no=1> ISSN:2188-8906, 特に第 6 章	1-15