

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7613488号
(P7613488)

(45)発行日 令和7年1月15日(2025.1.15)

(24)登録日 令和7年1月6日(2025.1.6)

(51)国際特許分類 F I
G 0 6 N 7/00 (2023.01) G 0 6 N 7/00

請求項の数 7 (全17頁)

| | | | |
|-------------|-----------------------------|----------|---|
| (21)出願番号 | 特願2022-577874(P2022-577874) | (73)特許権者 | 000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号 |
| (86)(22)出願日 | 令和3年1月27日(2021.1.27) | (74)代理人 | 110001519 弁理士法人太陽国際特許事務所 |
| (86)国際出願番号 | PCT/JP2021/002817 | (72)発明者 | 藤本 純也 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 |
| (87)国際公開番号 | WO2022/162782 | (72)発明者 | 中山 收文 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 |
| (87)国際公開日 | 令和4年8月4日(2022.8.4) | 審査官 | 渡辺 順哉 |
| 審査請求日 | 令和5年6月29日(2023.6.29) | | |

最終頁に続く

(54)【発明の名称】 行動区間推定モデル構築装置、行動区間推定モデル構築方法及び行動区間推定モデル構築プログラム

(57)【特許請求の範囲】

【請求項1】

各々が人の動作の種類を状態とする複数の第1隠れマルコフモデルを含む第2隠れマルコフモデルを複数含む隠れセミマルコフモデルであって、複数の第2隠れマルコフモデルの各々が複数の前記動作を組み合わせて定まる行動を状態とする隠れセミマルコフモデルにおいて、前記複数の第1隠れマルコフモデルの前記動作の種類ごとの観測確率を教師なし学習で学習する観測確率学習部と、

前記観測確率学習部にて学習した前記観測確率を固定し、入力された第1教師ありデータを水増しすることで第2教師ありデータとし、前記第1隠れマルコフモデルの前記動作の遷移確率を前記第2教師ありデータを使用した教師あり学習で学習する遷移確率学習部と、

10

前記観測確率学習部で学習した前記観測確率及び前記遷移確率学習部で学習した前記遷移確率を使用して前記行動の区間を推定するモデルである前記隠れセミマルコフモデルを構築する構築部と、

を含み、

前記遷移確率学習部は、前記第1教師ありデータに、時間方向のオーバーサンプリング及び特徴空間上のオーバーサンプリングの少なくとも一方を行うことで生成したデータの各々に前記第1教師ありデータの教師情報を付加することで水増しする、

行動区間推定モデル構築装置。

【請求項2】

20

前記時間方向のオーバーサンプリングは、各時刻でランダムに設定されたオリジナルパラメータを減衰させながら前後の時刻に伝播し、

各時刻において、前記オリジナルパラメータ、前後の時刻から伝播されたパラメータの内、最大のパラメータの時刻に対応する前記動作の特徴値を前記各時刻の特徴値として選択する、

請求項 1 に記載の行動区間推定モデル構築装置。

【請求項 3】

前記オリジナルパラメータは、所定数離れた時刻で 0 となるように減衰される、

請求項 2 に記載の行動区間推定モデル構築装置。

【請求項 4】

前記特徴空間上のオーバーサンプリングは、前記第 1 教師ありデータに、動作を行う人の身体部位ごとの速度に関連するノイズを前記身体部位ごとの前記動作の特徴値に付加する、

請求項 1 ~ 請求項 3 の何れか 1 項に記載の行動区間推定モデル構築装置。

【請求項 5】

前記身体部位ごとの速度に関連するノイズの大きさは、前記身体部位ごとの角速度の各々が大きいほど大きい、

請求項 4 に記載の行動区間推定モデル構築装置。

【請求項 6】

コンピュータが、

各々が人の動作の種類を状態とする複数の第 1 隠れマルコフモデルを含む第 2 隠れマルコフモデルを複数含む隠れセミマルコフモデルであって、複数の第 2 隠れマルコフモデルの各々が複数の前記動作を組み合わせて定まる行動を状態とする隠れセミマルコフモデルにおいて、前記複数の第 1 隠れマルコフモデルの前記動作の種類ごとの観測確率を教師なし学習で学習し、

学習した前記観測確率を固定し、入力された第 1 教師ありデータを水増しすることで第 2 教師ありデータとし、前記第 1 隠れマルコフモデルの前記動作の遷移確率を前記第 2 教師ありデータを使用した教師あり学習で学習し、

学習した前記観測確率及び前記遷移確率を使用して前記行動の区間を推定するモデルである前記隠れセミマルコフモデルを構築する、

行動区間推定モデル構築方法であって、

前記第 1 教師ありデータに、時間方向のオーバーサンプリング及び特徴空間上のオーバーサンプリングの少なくとも一方を行うことで生成したデータの各々に前記第 1 教師ありデータの教師情報を付加することで水増しする、

行動区間推定モデル構築方法。

【請求項 7】

各々が人の動作の種類を状態とする複数の第 1 隠れマルコフモデルを含む第 2 隠れマルコフモデルを複数含む隠れセミマルコフモデルであって、複数の第 2 隠れマルコフモデルの各々が複数の前記動作を組み合わせて定まる行動を状態とする隠れセミマルコフモデルにおいて、前記複数の第 1 隠れマルコフモデルの前記動作の種類ごとの観測確率を教師なし学習で学習し、

学習した前記観測確率を固定し、入力された第 1 教師ありデータを水増しすることで第 2 教師ありデータとし、前記第 1 隠れマルコフモデルの前記動作の遷移確率を前記第 2 教師ありデータを使用した教師あり学習で学習し、

学習した前記観測確率及び前記遷移確率を使用して前記行動の区間を推定するモデルである前記隠れセミマルコフモデルを構築する、

処理であって、

前記第 1 教師ありデータに、時間方向のオーバーサンプリング及び特徴空間上のオーバーサンプリングの少なくとも一方を行うことで生成したデータの各々に前記第 1 教師ありデータの教師情報を付加することで水増しする、

10

20

30

40

50

処理をコンピュータに実行させる行動区間推定モデル構築プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、行動区間推定モデル構築装置、行動区間推定モデル構築方法及び行動区間推定モデル構築プログラムに関する。

【背景技術】

【0002】

ディープラーニング技術の発展により通常のRGBカメラで撮影した人の映像から姿勢を高精度に認識できるようになり、この認識情報を利用して人の行動を推定する様々な研究開発が行われている。当該状況下において、人の映像から検出した姿勢の時系列データから指定した行動が発生した時間区間を推定する取り組みが行われている。

10

【先行技術文献】

【非特許文献】

【0003】

【文献】山本龍一、酒向慎司、北村正、「隠れセミマルコフモデルと線形動的システムを組み合わせた音楽音響信号と楽譜の実時間アライメント手法」、研究報告音楽情報科学(MUS)、2012年

【文献】Shun-Zheng Yu、「Hidden semi-Markov models」、Artificial Intelligence、Volume 174、Issue 2、2010年2月、215~243頁

20

【文献】若林啓、三浦孝夫、「階層型隠れマルコフモデルの高速パラメータ推定」、電子情報通信学会論文誌、2011年

【文献】「映像から人の様々な行動を認識するAI技術「行動分析技術 Actlyzer」を開発」、[online]、2019年11月25日、富士通株式会社、[2020年1月19日検索]、インターネット(URL: <https://pr.fujitsu.com/jp/news/2019/11/25.html>)

【発明の概要】

【発明が解決しようとする課題】

【0004】

行動の時間区間を推定するモデルを学習させる際の教師ありデータの教師情報を作成するコストが高い。

30

【0005】

本開示は、1つの側面として、行動区間推定モデルを効率的に構築することを目的とする。

【課題を解決するための手段】

【0006】

1つの実施形態では、隠れセミマルコフモデルにおいて、複数の第1隠れマルコフモデルの動作の種類ごとの観測確率を教師なし学習で学習する。隠れセミマルコフモデルは、各々が人の動作の種類を状態とする複数の第1隠れマルコフモデルを含む第2隠れマルコフモデルを複数含み、複数の第2隠れマルコフモデルの各々は複数の動作を組み合わせることで定まる行動を状態とする。学習した観測確率を固定し、入力された第1教師ありデータを水増しすることで第2教師ありデータとし、第1隠れマルコフモデルの動作の遷移確率を第2教師ありデータを使用した教師あり学習で学習する。学習した観測確率及び遷移確率を使用して行動の区間を推定するモデルである隠れセミマルコフモデルを構築する。第1教師ありデータに、時間方向のオーバーサンプリング及び特徴空間上のオーバーサンプリングの少なくとも一方を行うことで生成したデータの各々に第1教師ありデータの教師情報を付加することで水増しする。

40

【発明の効果】

【0007】

50

本開示は、1つの側面として、行動区間推定モデルを効率的に構築することができる。

【図面の簡単な説明】

【0008】

【図1】本実施形態の隠れセミマルコフモデルを例示する概念図である。

【図2】本実施形態の機能構成を例示するブロック図である。

【図3】本実施形態の第1隠れマルコフモデルの状態を例示する概念図である。

【図4】教師ありデータの増しを説明する概念図である。

【図5】教師ありデータの増しを説明する概念図である。

【図6】教師ありデータの増しを説明する概念図である。

【図7】教師ありデータの増しを説明する概念図である。

10

【図8】教師ありデータの増しを説明する概念図である。

【図9】教師ありデータの増しを説明する概念図である。

【図10】本実施形態のハードウェア構成を例示するブロック図である。

【図11】行動区間推定モデル構築処理の流れを例示するフローチャートである。

【図12】特徴ベクトル抽出処理の流れを例示するフローチャートである。

【図13】教師ありデータ増し処理の流れを例示するフローチャートである。

【図14】行動区間推定処理の流れを例示するフローチャートである。

【図15】関連技術の行動を説明する概念図である。

【図16】関連技術の階層型隠れマルコフモデルを例示する概念図である。

【図17】関連技術の概要を例示する概念図である。

20

【図18】本実施形態の概要を例示する概念図である。

【図19】観測データの揺らぎを例示する概念図である。

【発明を実施するための形態】

【0009】

本実施形態では、人の行動が発生した時間区間を推定する行動区間推定モデルの一例として、図1に例示するような隠れセミマルコフモデル(以下、HSMM(Hidden semi-Markov model)という。)を構築する。HSMMは、隠れマルコフモデル(以下、HMM(Hidden Markov model)という。)のパラメータに加え、状態ごとの継続時間の確率分布をパラメータとしてもつ。

【0010】

30

本実施形態のHSMMは、人の動作の各々を状態とする複数の第1HMMと、各々が複数の動作を組み合わせて定まる行動を状態とする第2HMMと、を含む。m1、m2、m3は動作の一例であり、a1、a2、a3は行動の一例である。行動は、複数の動作の組合せであり、動作は、複数の姿勢の組合せである。

【0011】

パラメータを設定することで構築されたHSMMに人の姿勢を検知することで生成された時系列センサデータが与えられると、HSMMは最適な行動の時間区間(以下、行動区間という。)を推定する。d1、d2、d3は行動区間の一例である。

【0012】

HMMのパラメータには、観測確率及び遷移確率が存在する。O1、...、O8は観測確率の一例であり、遷移確率は状態をつなぐ矢印に対応する確率である。観測確率とは、各状態において、ある特徴が観測される確率であり、遷移確率とは、ある状態から別の状態に遷移する確率である。遷移の順番が定まっている場合は、遷移確率は不要である。なお、動作の数、行動の数、即ち、第1HMM、第2HMMの数の数は例示であり、図1に例示される数に限定されない。

40

【0013】

図2は、本実施形態の行動区間推定モデル構築装置10の機能ブロック図の一例である。行動区間推定モデル構築装置10は、観測確率学習部11、遷移確率学習部12、構築部13を有する。観測確率学習部11は、以下に説明するように、教師なしデータで行動区間推定モデルの一例であるHSMMの観測確率を学習する。

50

【 0 0 1 4 】

本実施形態では、ある作業目標を達成するための限定された行動を対象とする。このような行動は、例えば、工場のラインで行われる定型作業での行動であり、以下の性質を有する。

性質 1：作業を構成する各行動の違いは、限定された複数の動作の組合せの違いである。

性質 2：同じ作業を行う際に観測される複数の姿勢は類似している。

【 0 0 1 5 】

本実施形態では、性質 1 に基づいて、全ての行動が 1 つの動作群に含まれる動作で構成される。図 3 に例示するように、動作群には、例えば、3 つの動作 m_{11} 、 m_{12} 、 m_{13} が含まれている。

10

【 0 0 1 6 】

例えば、動作 m_{11} は「腕を上げる」、動作 m_{12} は「腕を降ろす」、動作 m_{13} は「腕を前に伸ばす」であってよい。動作群に含まれる動作の数は図 3 の例に限定されない。また、各行動に含まれる動作の数も図 3 の例に限定されない。

【 0 0 1 7 】

図 3 の HMM において、破線矢印に対応する各動作の観測確率は行動には依存しないため、行動区間の教師なしデータで学習することができる。学習は、例えば、機械学習、ニューラルネットワーク、ディープラーニングなどを使用して行う。

【 0 0 1 8 】

詳細には、観測確率の教師なし学習に使用するモデルは混合ガウス分布（以下、GMM (Gaussian Mixture Model) という。）であってよい。各観測は動作のうち 1 つの動作が確率的に選択され、その動作についてのガウス分布により生成されると仮定する。これは、観測の時系列的な依存関係を使用しない教師あり学習とは異なる仮定である。学習した GMM の各ガウス分布のパラメータを各動作における観測確率の確率分布であるガウス分布に割り当てる。

20

【 0 0 1 9 】

遷移確率学習部 12 は、以下に説明するように、教師情報をもつ学習データ（以下、教師ありデータという。）で、第 1 HMM の動作の遷移確率を学習する。教師情報は、姿勢の時系列データに対して各行動が発生している時間区間の正解を与える情報である。学習は、例えば、最尤推定や EM アルゴリズム (Expectation - Maximization algorithm) などを使用して行う（その他の機械学習、ニューラルネットワーク、ディープラーニングなどの方式を使用してもよい）。

30

【 0 0 2 0 】

教師ありデータの生成には、時間及び労力がかかる。したがって、本実施形態では、観測確率学習部 11 で学習した観測確率を固定し、既存の教師ありデータから遷移確率を学習する。

【 0 0 2 1 】

詳細には、図 4 に例示するように、第 1 教師ありデータの一例である既存の教師ありデータのデータを種データ SD とし、オーバーサンプリングすることでデータを水増しする。本実施形態では、例えば、時間方向にオーバーサンプリングを行い、その後、特徴空間上でオーバーサンプリングを行う。

40

【 0 0 2 2 】

時間方向のオーバーサンプリングについて説明する。時間方向のオーバーサンプリングでは、例えば、人によって異なる動作にかかる時間の長さに関連する時間的な伸び縮みを考慮する。詳細には、以下の通りである。

(1) 図 5 に例示するように、人の動作の観測系列の各時刻について、当該時刻の特徴の伸びの強さを表す乱数を生成する。図 5 の各時刻の縦線は、オリジナルパラメータに相当する乱数で生成した伸びの強さを表す。

(2) 各時刻の伸びの強さを減衰させながら当該時刻の前後の時刻に伝播させる。伸びの強さは、所定の数離れた時刻で 0 になるように減衰される。図 5 の例では、破線で表すよ

50

うに、3時刻離れた時刻で0になるように減衰されている。減衰は、直線的な減衰でなくともよい。

(3) 各時刻のオリジナルの伸びの強さ、前後の時刻から伝播されたパラメータに相当する伝播された伸びの強さの内、最大の強さに対応する時刻の特徴値を、当該時刻の特徴値として選択する。図5の例では、時刻1では、オリジナルの伸びの強さが最大であるため、オリジナルの特徴値である時刻1の特徴値を選択し、時刻2では、時刻1から伝播された伸びの強さが最大であるため、時刻1の特徴値を選択する。時刻3では、時刻1から伝播された伸びの強さが最大であるため、時刻1の特徴値を選択し、時刻4では、オリジナルの伸びの強さが最大であるため、オリジナルの特徴値である時刻4の特徴値を選択する。

【0023】

特徴空間上でのオーバーサンプリングについて説明する。上記性質2によれば、同じ作業の姿勢は類似しているため、ノイズを付加することで、図6に例示するように実際の観測ごとのばらつきに類似したばらつきをもつデータを生成することができる。

【0024】

種データSDの教師情報TIを、水増ししたデータの各々に共通に適用することで教師ありデータを水増しする。第2教師ありデータの一例である水増しした教師ありデータを使用して、第1HMMの複数の動作の遷移確率を教師あり学習で学習する。

【0025】

オーバーサンプリングでは、各時刻の特徴値にノイズを生成して付加する。例えば、特定した動作のサンプル群の共分散の定数倍の共分散の多変量ガウス分布から生成したノイズを付加してもよい。また、特定した動作のサンプル群から最も中心距離が近い動作のサンプル群までの中心距離dを算出し、特徴空間の各軸方向の標準偏差がdの定数倍となる等方性のガウス分布(共分散行列が対角行列である)から生成したノイズを付加してもよい。

【0026】

本実施形態では、動作を行う人の身体部位ごとの速度に関連するノイズを身体部位ごとの動作の特徴値に付加する。例えば、ガウス分布の共分散行列のうち分散成分である対角成分を、動作を行う人の身体部位ごとに変更する。詳細には、身体部位i(iは自然数)の特徴ベクトルの姿勢成分である特徴値の標準偏差 σ_i (分散 σ_i^2)を、身体部位iの角速度 ω_i 、ベースとなる標準偏差の値 σ_i (分散 σ_i^2)、定数係数kを用いて式(1)で算出する。

$$\sigma_i' = \sigma_i + k \omega_i \dots (1)$$

【0027】

σ_i 及びkは、実験的に事前に決定される定数であり、身体部位ごとには変更しない。式(1)の第2項に示されるように、角速度の大きさに比例してノイズ、即ち、姿勢のばらつきを大きくする。例えば、図7の横軸は、身体部位1の姿勢成分である特徴値1を表し、縦軸は、身体部位2の姿勢成分である特徴値2を表す。

【0028】

図7では、特徴空間を2次元で表現しているが、次元数は2より多くてもよい。図7において、楕円は動作m21、m22、m23の特徴空間上の点で表されるサンプルが観測される確率分布(ガウス分布)の等高線を表す。楕円の中心に近いほど、確率が高い。

【0029】

身体部位1の動きの角速度成分と身体部位2の動きの角速度成分が略同様である場合、図7の左に示されるように、縦軸方向及び横軸方向の双方に、略同様の大きさのノイズを付加する。一方、身体部位1の動きの角速度成分が身体部位2の動きの角速度成分より大きい場合、図7の右に示されるように、縦軸方向に比較して横軸方向に大きいノイズを付加する。

【0030】

時間方向のオーバーサンプリングによれば、時間方向の変化に対応することができる。即ち、同じ作業を行っている場合でも、速く動いたり、遅く動いたりすることで、ある動

10

20

30

40

50

作（動き特徴）が短く観測されたり、長く観測されたりする。速い動きでは、ある動作が観測されない場合もある。

【0031】

図8の左に例示するように、例えば、作業者Aは、動作2に略3時刻使用しているが、図8の右上に例示するように、作業者Bは動作2に略4時刻使用し、図8の右下に例示するように、作業者Cは動作2に略1時刻使用している。時間方向のオーバーサンプリングを行うことで、このように、時間的な伸び縮みのあるサンプルを増やすことができる。

【0032】

特徴空間上でのオーバーサンプリングによれば、姿勢を表す特徴値のばらつきに対応することができる。例えば、図9の左に例示するように、第1の腕の移動速度が大きく、第2の腕の移動速度が小さい場合、図9の右に例示するように、第1の腕の姿勢変化も速度に比例して大きく、したがって、特徴値のばらつきも大きい。

10

【0033】

一方、第2の腕の姿勢変化は速度に比例して小さく、したがって、特徴値のばらつきも小さい。特徴空間上でのオーバーサンプリングを行うことで、このように、身体部位によって特徴値のばらつきが異なるサンプルを増やすことができる。

【0034】

時間方向のオーバーサンプリング及び特徴方向のオーバーサンプリングは両方行われてもよいし、何れか一方だけが行われてもよい。特徴方向のオーバーサンプリングだけが行われる場合、各時刻の身体部位ごとの当該時刻の特徴値に動作を行う人の身体部位ごとの速度に関連するノイズが付加される。

20

【0035】

構築部13は、観測確率学習部11で学習した観測確率、及び遷移確率学習部12で学習した状態遷移確率を使用して、図1に例示するようなHSMを構築する。O1、O2、...、O8は、観測確率学習部11で学習した観測確率を表し、行動a1、a2、a3の各々に含まれる動作m1、m2、及びm3間の矢印は、遷移確率学習部12で学習した状態遷移確率に対応する。d1、d2、d3は、各行動の継続時間を表し、継続時間の確率分布は、教師情報の行動の継続時間から決定される。例えば、継続時間の確率分布は、一定範囲の一様分布であってよい。構築したHSMに、センサで人の姿勢を検知して生成したセンサデータを適用して、各行動の時間区間である行動区間を推定する。推定についての詳細は、後述する。

30

【0036】

本実施形態の行動区間推定モデル構築装置10は、以下の特徴を有する。

1. 第1HMMの全行動で共通な動作の観測確率は教師なし学習で学習する。
2. 第1HMMの動作間の遷移確率は、教師あり種データから増強した教師ありデータを使用して、教師あり学習で学習する。

【0037】

行動区間推定モデル構築装置10は、一例として、図10に示すように、CPU(Central Processing Unit)51、一次記憶装置52、二次記憶装置53、及び、外部インターフェイス54を含む。CPU51は、ハードウェアであるプロセッサの一例である。CPU51、一次記憶装置52、二次記憶装置53、及び、外部インターフェイス54は、バス59を介して相互に接続されている。CPU51は、単一のプロセッサであってもよいし、複数のプロセッサであってもよい。また、CPU51に代えて、例えば、GPU(Graphics Processing Unit)が使用されてもよい。

40

【0038】

一次記憶装置52は、例えば、RAM(Random Access Memory)などの揮発性のメモリである。二次記憶装置53は、例えば、HDD(Hard Disk Drive)、又はSSD(Solid State Drive)などの不揮発性のメモリである。

50

【 0 0 3 9 】

二次記憶装置 5 3 は、プログラム格納領域 5 3 A 及びデータ格納領域 5 3 B を含む。プログラム格納領域 5 3 A は、一例として、行動区間推定モデル構築プログラムなどのプログラムを記憶している。データ格納領域 5 3 B は、一例として、教師ありデータ、教師なしデータ、学習した観測確率、及び遷移確率などを記憶する。

【 0 0 4 0 】

C P U 5 1 は、プログラム格納領域 5 3 A から行動区間推定モデル構築プログラムを読み出して一次記憶装置 5 2 に展開する。C P U 5 1 は、行動区間推定モデル構築プログラムをロードして実行することで、図 2 の観測確率学習部 1 1、遷移確率学習部 1 2、及び構築部 1 3 として動作する。

10

【 0 0 4 1 】

なお、行動区間推定モデル構築プログラムなどのプログラムは、外部サーバに記憶され、ネットワークを介して、一次記憶装置 5 2 に展開されてもよい。また、行動区間推定モデル生成プログラムなどのプログラムは、D V D (D i g i t a l V e r s a t i l e D i s c) などの非一時的記録媒体に記憶され、記録媒体読込装置を介して、一次記憶装置 5 2 に展開されてもよい。

【 0 0 4 2 】

外部インターフェイス 5 4 には外部装置が接続され、外部インターフェイス 5 4 は、外部装置と C P U 5 1 との間の各種情報の送受信を司る。図 1 0 では、外部インターフェイス 5 4 に、ディスプレイ 5 5 A 及び外部記憶装置 5 5 B が接続されている例を示している。外部記憶装置 5 5 B には、例えば、教師ありデータ、教師なしデータ、及び、構築した H S M M などを記憶する。ディスプレイ 5 5 A は、例えば、構築した H S M M モデルを視認可能に表示する。

20

【 0 0 4 3 】

行動区間推定モデル構築装置 1 0 は、例えば、パーソナルコンピュータ、サーバ、及び、クラウド上のコンピュータなどであってよい。

【 0 0 4 4 】

図 1 1 に、行動区間推定モデル構築処理の流れを例示する。C P U 5 1 は、ステップ 1 0 1 で、後述するように、学習データから人の姿勢の連鎖である運動を表す特徴ベクトルを抽出する。C P U 5 1 は、ステップ 1 0 2 で、ステップ 1 0 1 で抽出した特徴ベクトルのクラスタリング (G M M のパラメータ推定) により、要素となる動作に分類し、各動作の観測確率を教師なし学習で学習する。

30

【 0 0 4 5 】

C P U 5 1 は、ステップ 1 0 3 で、後述するように、教師あり種データをオーバーサンプリングして生成したデータに教師あり種データの教師情報を付与することで、教師ありデータを水増しする。C P U 5 1 は、ステップ 1 0 4 で、教師ありデータについて、教師情報で与えられた各行動の時間区間ごとに特徴ベクトルを振り分ける。

【 0 0 4 6 】

C P U 5 1 は、ステップ 1 0 5 で、ステップ 1 0 4 で振り分けた時間区間内の特徴ベクトルの系列を観測データとして、ステップ 1 0 3 で水増しした教師ありデータを使用し、第 1 H M M の動作の遷移確率を教師あり学習で学習する。

40

【 0 0 4 7 】

C P U 5 1 は、ステップ 1 0 6 で、各行動の継続時間の確率分布として、教師情報で与えられた各行動の継続時間に対して所定の範囲の一様分布を設定する。C P U 5 1 は、ステップ 1 0 2 で学習した観測確率及びステップ 1 0 5 で学習した遷移確率を使用して、H S M M を構築する。ステップ 1 0 6 の設定で一定時間継続後に教師情報で与えられた各行動の順番に第 2 H M M の行動が遷移する H S M M を構築する。構築した H S M M は、例えば、データ格納領域 5 3 B に格納されてもよい。

【 0 0 4 8 】

図 1 2 は、図 1 1 のステップ 1 0 1 の特徴ベクトル抽出処理の詳細を例示する。C P U

50

51は、ステップ151で、学習に使用するデータから人を検出し、追跡することで、人の姿勢情報を取得する。CPU51は、ステップ152で、ステップ151で取得した姿勢情報が複数人の姿勢情報を含む場合、姿勢情報の時系列データから分析対象とする姿勢情報の時系列データを取得する。分析対象とする姿勢情報は、人を囲むバウンディングボックスの大きさ、及び時間などから選択する。

【0049】

CPU51は、ステップ153で、ステップ152で取得した姿勢情報の時系列データから身体の一部位についての運動情報の時系列データを取得する。運動情報の時系列とは、例えば、各部位の曲げの程度、曲げの速度などであってよい。各部位とは、例えば、肘、膝などであってよい。

【0050】

CPU51は、ステップ154で、スライディングタイムウィンドウにより一定の時間間隔ごとにウィンドウ内のステップ153の運動情報を時間方向で平均化して特徴ベクトルを算出する。

【0051】

図13に、図11のステップ103の教師ありデータ水増し処理の流れを例示する。CPU51は、ステップ251で、観測データ(人の動作の観測時系列)の時刻の各々で、当該時刻の特徴の伸びの強さを表す乱数を生成する。CPU51は、ステップ252で、時刻の各々で生成した伸びの強さの値を減衰させながら当該時刻の前後の時間に伝播させる。

【0052】

CPU51は、ステップ253で、当該時刻の伸びの強さ、及び、他の時刻から伝播された伸びの強さの値の内、最大の伸びの強さに該当する時刻の観測データの特徴値を、当該時刻の特徴値として選択する。CPU51は、ステップ254で、ガウス分布の共分散行列を、身体部位の各々の角速度の値に基づいて算出する。

【0053】

CPU51は、ステップ255で、ステップ253で選択した特徴値の各々に、ステップ254で算出した共分散行列のガウス分布で生成したノイズを付加する。教師ありデータの水増しを繰り返すことで、教師ありデータを水増しする。

【0054】

ステップ254及びステップ255の処理だけを繰り返し行ってもよい。この場合、各時刻のオリジナルの特徴値にノイズを付加する。また、ステップ251～ステップ253の処理だけを繰り返し行ってもよい。

【0055】

図14に、本実施形態で構築したHSMMを使用した行動区間推定処理の流れを例示する。図10の行動区間推定モデル構築装置10は、構築したHSMMをデータ格納領域53Bに格納することで行動区間推定装置として機能してもよい。

【0056】

CPU51は、ステップ201で、センサで人の姿勢を検知することにより生成されたセンサデータから特徴ベクトルを抽出する。センサは、人の姿勢を検知するデバイスであり、例えば、カメラ、赤外線センサ、モーションキャプチャデバイスなどであってよい。図14のステップ201は、図11のステップ101と同様であるため、詳細な説明は省略する。

【0057】

CPU51は、ステップ202で、ステップ201で抽出した特徴ベクトルの系列を観測データとして、行動区間推定モデル構築処理で構築したHSMMと照合して各行動状態の継続時間を推定する。CPU51は、ステップ203で、ステップ202で推定した各行動状態の継続時間から各行動の時間区間を推定する。

【0058】

例えば、映像を入力として、映像における特定の行動を認識するような技術では、基本

10

20

30

40

50

動作認識、要素行動認識、及び上位行動認識を行う。映像における特定の行動とは、要素行動の組合せで、さらに複雑な上位行動であり、基本動作認識とは、フレームごとの姿勢認識であり、要素行動認識とは、時間的空間的認識を行い、ある程度の時間長における単純行動を認識することである。上位行動認識とは、ある程度の時間長における複雑行動の認識である。当該技術において、本実施形態の行動区間推定モデル構築処理及び構築した行動区間推定モデルを適用し、行動区間を推定することができる。

【 0 0 5 9 】

関連技術では、行動に含まれる動作が特に限定されないH S M Mが使用され得る。当該関連技術では、図 1 5 に例示するように、例えば、以下の動作が存在すると仮定する。

(1) 腕を上げる、(2) 腕を降ろす、(3) 腕を前に伸ばす、(4) 両手を身体の前で近づける、(5) 前に移動する、(6) 横に移動する、(7) しゃがむ、(8) 立つ

10

【 0 0 6 0 】

行動の例は、例えば、以下の通りである。

行動 a 3 1 : (1) 腕を上げる (3) 腕を前に伸ばす (1) 腕を上げる (4) 両手を身体の前で近づける (7) しゃがむ、

行動 a 3 2 : (7) しゃがむ (4) 両手を身体の前で近づける (8) 立つ (5) 前に移動する (3) 腕を前に伸ばす、など

【 0 0 6 1 】

上記のように、一般的な行動の動作、即ち、推定する行動が制限されない複数の動作をH M Mが含む場合、動作の観測確率を1つの単純な確率分布で表すことは困難である。この問題に対処するために、階層型隠れマルコフモデルを使用する技術が存在する。階層型隠れマルコフモデルは、図 1 6 に例示するように、上位階層H M Mが複数の下位階層H M Mを状態として含む。行動 a 5 1、a 5 2、及びa 5 3は、下位階層H M Mの例である。下位階層H M Mの各々は、動作を状態として含み、m 5 1、m 5 2、m 5 3、m 6 1、m 6 2、m 6 3、m 7 1、及びm 7 2は、動作の例である。

20

【 0 0 6 2 】

階層型H M Mでは、図 1 7 に例示するように、教師情報T I Lをもつ学習データL Dを使用して、各行動の動作の観測確率及び遷移確率を教師あり学習で学習する。図 1 7 では、行動 a 5 1の観測確率p 1 1、遷移確率p 2 1、行動 a 5 2の観測確率p 1 2、遷移確率p 2 2、行動 a 5 3の観測確率p 1 3、遷移確率p 2 3を例示する。しかしながら、階層型H M Mでは、パラメータの数が多く、パラメータの自由度が高いため、パラメータの学習のために教師ありデータを多数使用する。教師ありデータの教師情報を作成するには、時間及び労力を要する。

30

【 0 0 6 3 】

一方、本開示では、図 1 8 に例示するように、H S M Mの行動に対応する第1 H M Mの各々で共通の観測確率p 1は教師なしデータL D Nを使用して教師なし学習で学習する。学習した観測確率p 1を固定して、第1 H M Mの各々の動作の遷移確率p 2 1 D、p 2 2 D、p 2 3 Dを教師ありデータを使用して教師あり学習で学習する。本開示では、既存の教師ありデータL D Dをオーバーサンプリングし生成したデータに、教師ありデータL D Dの教師情報T I Lを付加することで、教師ありデータを水増しして教師あり学習に使用する。したがって、本実施形態では、既存の教師ありデータが少ない場合でも、行動区間推定モデルを効率的に構築することができる。

40

【 0 0 6 4 】

図 1 9 の左に例示するように、例えば、時刻 t 1 で動作m 3 1、時刻 t 2 で動作m 3 1、時刻 t 3 で動作m 3 3、時刻 t 4 で動作m 3 2が高確率な動作の並びとなる場合の観測データの揺らぎについて例示する。図 1 9 の右上に例示するように、動作の動きが変化し、時刻 t 2 の観測が動作m 3 2の近くに变化した場合、時刻 t 1 で動作m 3 1、時刻 t 2 で動作m 3 2、時刻 t 3 で動作m 3 3、時刻 t 4 で動作m 3 2が高確率な動作の並びとなる。

【 0 0 6 5 】

50

図19の右下に例示するように、動作の速度が上がると、図19の左の時刻t3のサンプルが観測されず、時刻t1で動作m31、時刻t2で動作m31、時刻t3で動作m32が高確率な動作の並びとなる。このような揺らぎに対し、どのような揺らぎが生じ得るかを事前に学習して遷移確率としてモデルに反映することができる。

【0066】

しかしながら、教師ありデータが少ない場合、多様な揺らぎを直接学習することができず、観測データの揺らぎへの対応が弱い。本実施形態では、時間方向のオーバーサンプリング及び特徴空間上でのオーバーサンプリングを行うことで、観測データの揺らぎに対応することができる適切な教師ありデータを水増しすることができる。

【0067】

本実施形態では、これにより、既存の教師ありデータが少ない場合であっても、観測データの揺らぎを想定した動作の並び方のモデル化が可能となる。したがって、観測データに揺らぎがある場合であっても高精度に時間区間を推定することができる。

【0068】

本実施形態では、隠れセミマルコフモデルにおいて、複数の第1隠れマルコフモデルの動作の種類ごとの観測確率を教師なし学習で学習する。隠れセミマルコフモデルは、各々が人の動作の種類を状態とする複数の第1隠れマルコフモデルを含む第2隠れマルコフモデルを複数含み、複数の第2隠れマルコフモデルの各々は複数の動作を組み合わせて定まる行動を状態とする。学習した観測確率を固定し、入力された第1教師ありデータを水増しすることで第2教師ありデータとし、第1隠れマルコフモデルの動作の遷移確率を第2教師ありデータを使用した教師あり学習で学習する。学習した観測確率及び遷移確率を使用して行動の区間を推定するモデルである隠れセミマルコフモデルを構築する。第1教師ありデータに、時間方向のオーバーサンプリング及び特徴空間上のオーバーサンプリングの少なくとも一方を行うことで生成したデータの各々に第1教師ありデータの教師情報を付加することで水増しする。

【0069】

本開示によれば、行動区間推定モデルを効率的に構築することができる。即ち、例えば、工場での定型作業、ダンスの振り付け、武道の型などのように決まった順序で動作を行う複数の行動について、発生する順序に制約があるという条件の下で各行動の時間区間を正確に推定することができる。

【符号の説明】

【0070】

- 10 行動区間推定モデル構築装置
- 11 観測確率学習部
- 12 遷移確率学習部
- 13 構築部
- 51 CPU
- 52 一次記憶装置
- 53 二次記憶装置

10

20

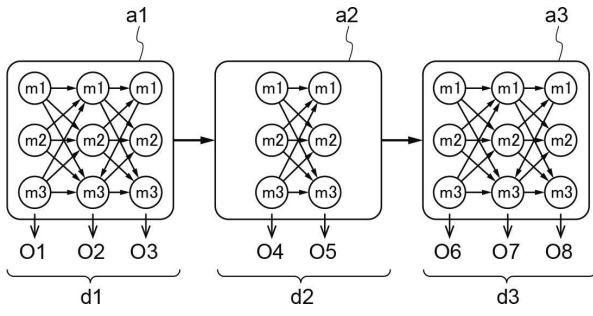
30

40

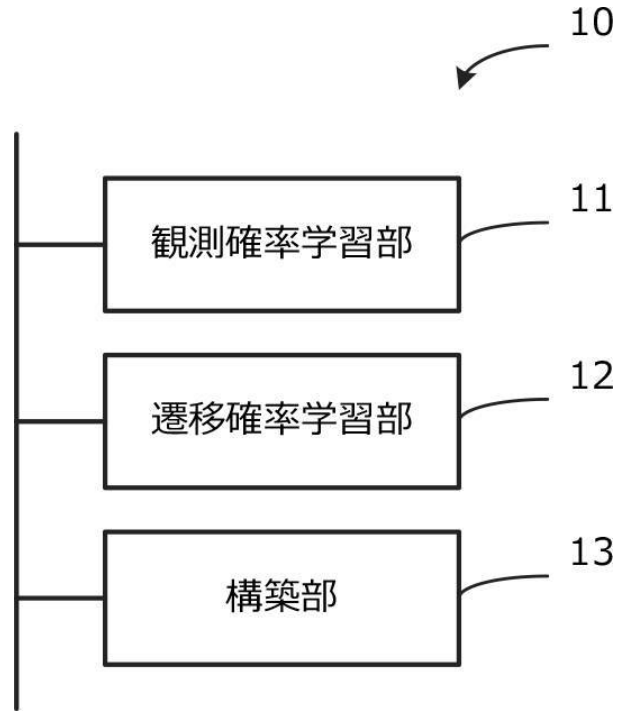
50

【図面】

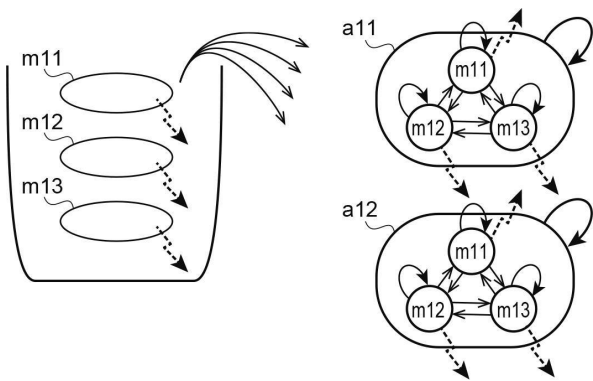
【図 1】



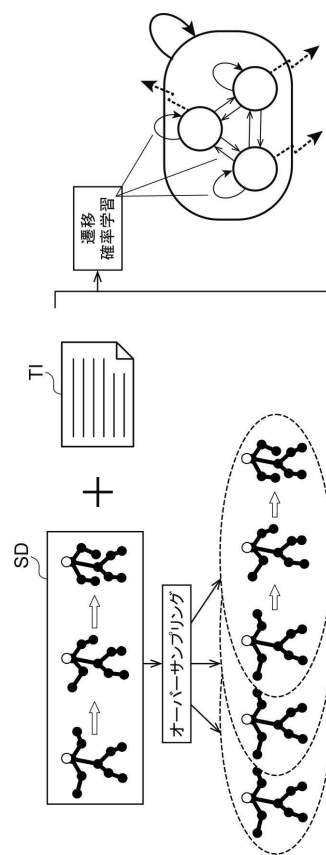
【図 2】



【図 3】



【図 4】



10

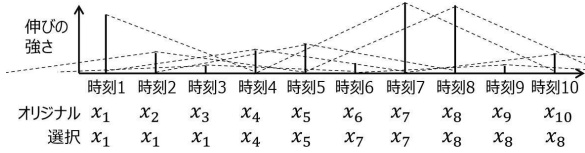
20

30

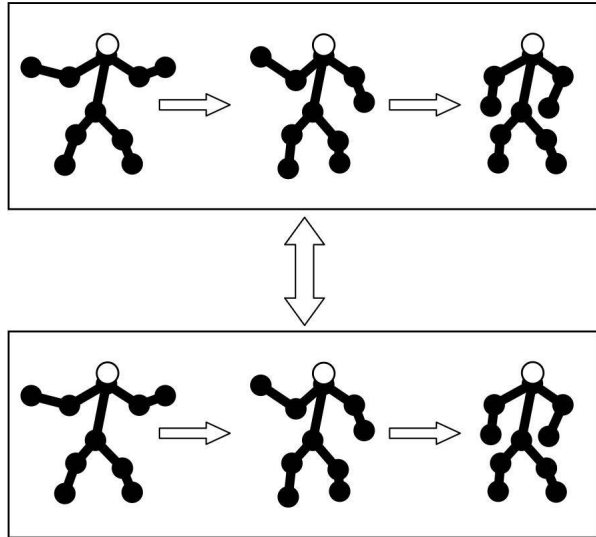
40

50

【 図 5 】

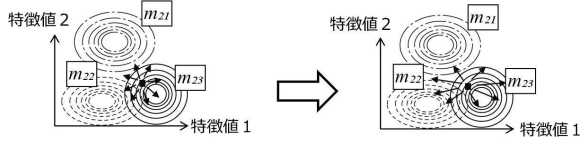


【 図 6 】

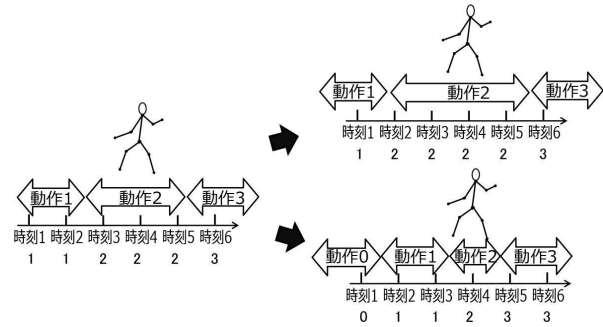


10

【 図 7 】



【 図 8 】



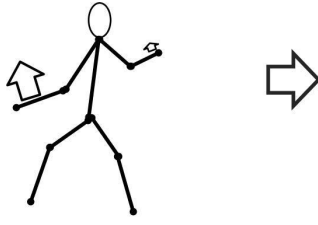
20

30

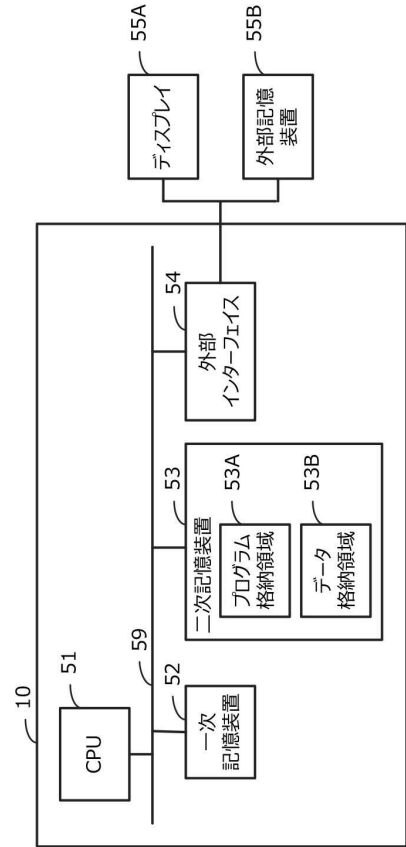
40

50

【図9】



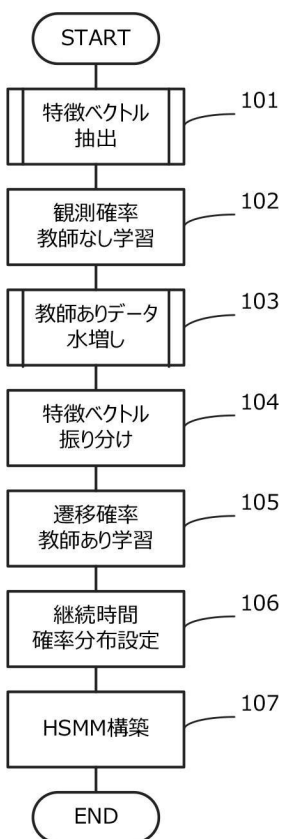
【図10】



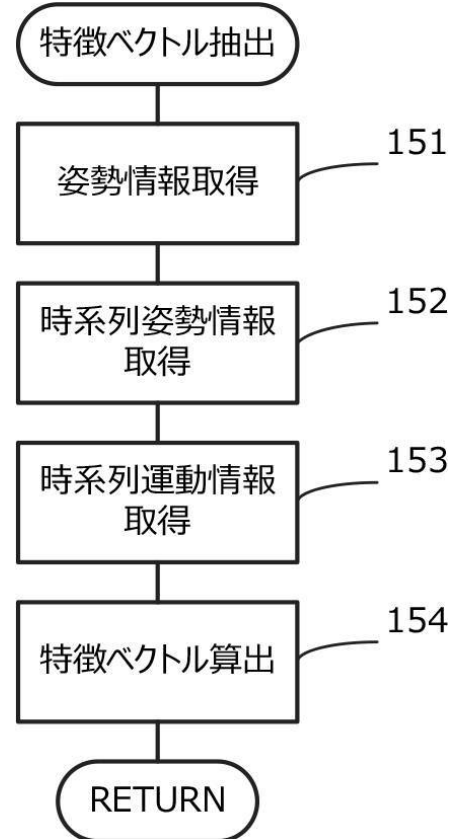
10

20

【図11】



【図12】

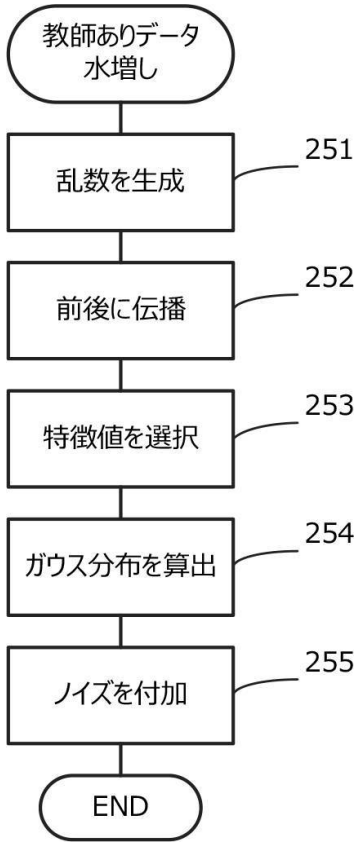


30

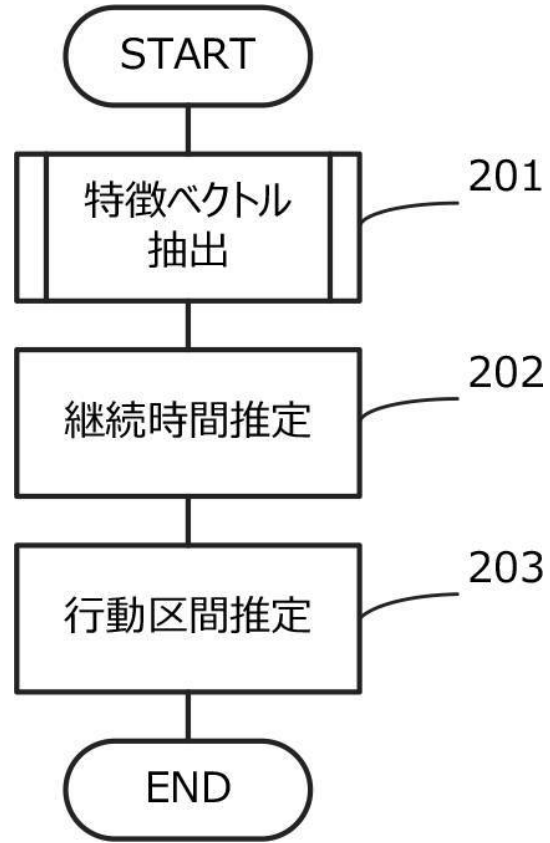
40

50

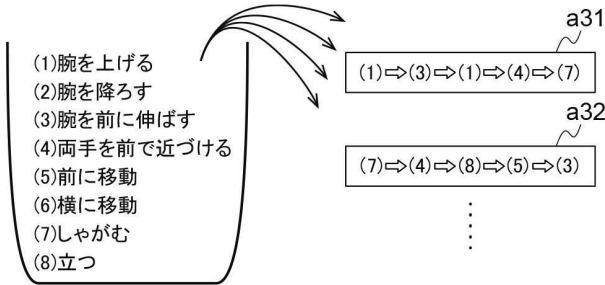
【図13】



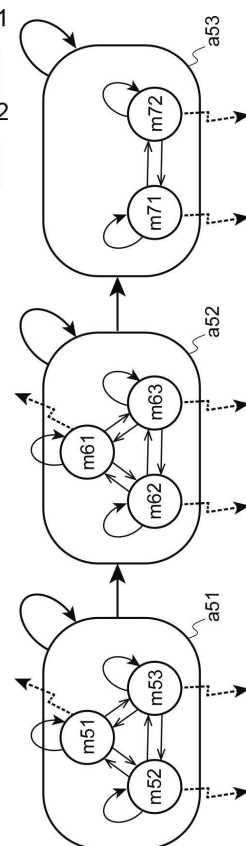
【図14】



【図15】



【図16】



10

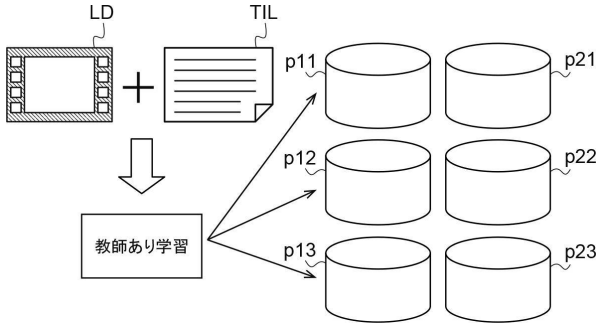
20

30

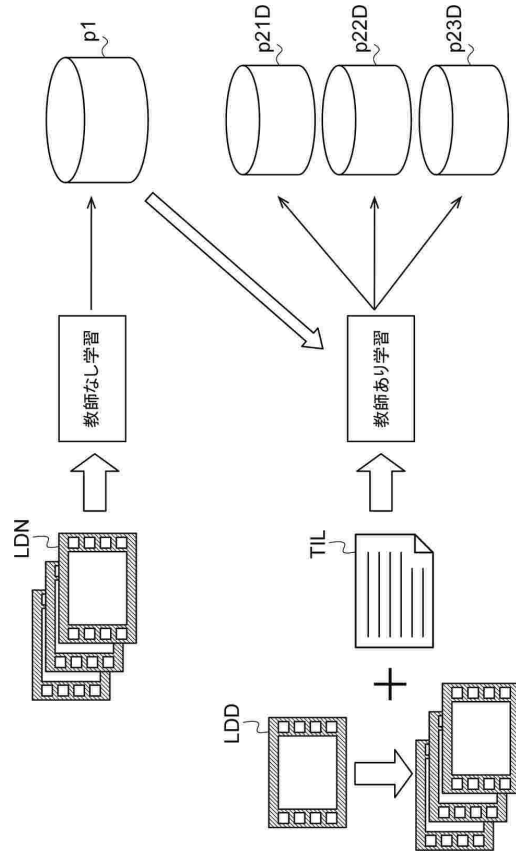
40

50

【図 17】



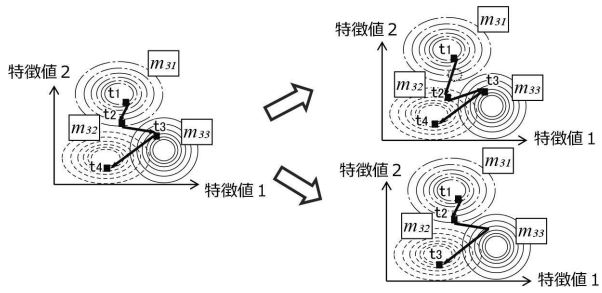
【図 18】



10

20

【図 19】



30

40

50

フロントページの続き

- (56)参考文献 国際公開第2020/050111(WO, A1)
杉野原 和也 ほか, レーザセンシングシステムと階層的隠れマルコフモデルによる居住者の行動推定, ロボティクス・メカトロニクス 講演会2015 講演論文集, No.15-2, 日本, 一般社団法人 日本機械学会, 2015年05月16日
- (58)調査した分野 (Int.Cl., DB名)
G06N 3/00 - 99/00