

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年11月17日(17.11.2022)



(10) 国際公開番号

WO 2022/239231 A1

(51) 国際特許分類:

G06T 7/00 (2017.01)

G06N 3/08 (2006.01)

G06T 7/20 (2017.01)

G06N 20/00 (2019.01)

(21) 国際出願番号: PCT/JP2021/018416

(22) 国際出願日: 2021年5月14日(14.05.2021)

(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: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 永井 隆昌 (NAGAI, Takasuke); 〒1808585 東京都武蔵野市緑町三丁目9番11号 NTT 知的財産センタ内 Tokyo (JP).
松村 誠明 (MATSUMURA, Masaaki); 〒1808585 東京都武蔵野市緑町三丁目9番11号 NTT

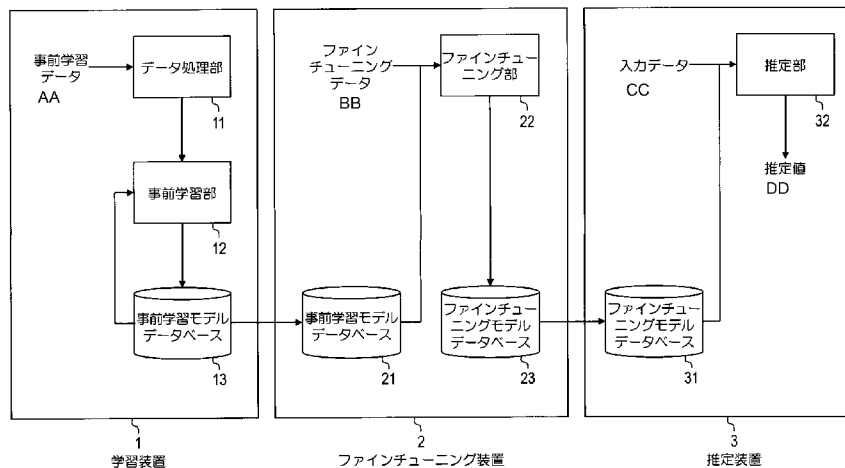
T 知的財産センタ内 Tokyo (JP). 志水 信哉 (SHIMIZU, Shinya); 〒1808585 東京都武蔵野市緑町三丁目9番11号 NTT 知的財産センタ内 Tokyo (JP). 山本 奏 (YAMAMOTO, Susumu); 〒1808585 東京都武蔵野市緑町三丁目9番11号 NTT 知的財産センタ内 Tokyo (JP).

(74) 代理人: 中尾 直樹, 外 (NAKAO, Naoki et al.); 〒1600022 東京都新宿区新宿三丁目1番22号 新宿NSOビル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,

(54) Title: LEARNING DEVICE, ESTIMATION DEVICE, LEARNING METHOD, ESTIMATION METHOD, AND PROGRAM

(54) 発明の名称: 学習装置、推定装置、学習方法、推定方法、プログラム



- 1 Learning device
- 2 Fine tuning device
- 3 Estimation device
- 11 Data processing unit
- 12 Prior learning unit
- 13, 21 Prior learning model database
- 22 Fine tuning unit
- 23, 31 Fine tuning model database
- 32 Estimation unit
- AA Prior learning data
- BB Fine tuning data
- CC Input data
- DD Estimated value

図1

(57) Abstract: This learning device comprises: a data processing unit which acquires human motion data, and masks all the human motion data at an arbitrary time group; and a prior learning unit which trains a prior learning model by computing a loss function from errors between true values and prediction data obtained by predicting data to be masked by means of the prior learning model, and updating weights of the prior learning model.

(57) 要約: 学習装置は、ヒューマンモーションデータを取得し、ヒューマンモーションデータの任意の時刻群におけるデータ全体をマスクするデータ処理部と、マスクの対象となったデータを事前学習モデルにより予測したデータである予測データと真値との誤差から損失関数を計算し、事前学習モデルの重みを更新することにより事前学習モデルを学習する事前学習部を含む。

NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

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

添付公開書類：

- 一 国際調査報告 (条約第21条(3))

明 細 書

発明の名称：

学習装置、推定装置、学習方法、推定方法、プログラム

技術分野

[0001] 本発明は、モデルを学習する学習装置、推定装置、学習方法、推定方法、プログラムに関する。

背景技術

[0002] 人間の動作を予測する従来技術の例として、非特許文献1がある。

先行技術文献

非特許文献

[0003] 非特許文献1: Yujun Cai et al., "Learning Progressive Joint Propagation for Human Motion Prediction", Computer Vision -ECCV 2020, 16th European Conference, Glasgow, UK, August 23-28, 2020, Proceedings, Part V II, pp. 226-242

発明の概要

発明が解決しようとする課題

[0004] 非特許文献1は、時間的に単方向の関係性のみを考慮しており、人間の動作を扱う場合に重要となる前後の動作の関係性を学習することができない。

[0005] そこで本発明では、前後の動作の関係性を考慮した汎用性の高い事前学習モデルを学習することができる学習装置を提供することを目的とする。

課題を解決するための手段

[0006] 本発明の学習装置は、データ処理部と、事前学習部を含む。

[0007] データ処理部は、ヒューマンモーションデータを取得し、ヒューマンモーションデータの任意の時刻群におけるデータ全体をマスクする。事前学習部は、マスクの対象となったデータを事前学習モデルにより予測したデータである予測データと真値との誤差から損失関数を計算し、事前学習モデルの重

みを更新することにより事前学習モデルを学習する。

発明の効果

[0008] 本発明の学習装置によれば、前後の動作の関係性を考慮した汎用性の高い事前学習モデルを学習することができる。

図面の簡単な説明

[0009] [図1]実施例1の学習装置とファインチューニング装置と推定装置の機能構成を示すブロック図。

[図2]実施例1の学習装置とファインチューニング装置と推定装置の動作を示すフローチャート。

[図3]実施例1の学習装置のデータ処理部の機能構成を示すブロック図。

[図4]実施例1の学習装置のデータ処理部の動作を示すフローチャート。

[図5]ヒューマンモーションデータとマスクとトークンを例示する図。

[図6]データ処理部が生成する埋め込み処理後の特徴ベクトルを例示する図。

[図7]実施例1の学習装置の事前学習部の機能構成を示すブロック図。

[図8]実施例1の学習装置の事前学習部の動作を示すフローチャート。

[図9]コンピュータの機能構成例を示す図。

発明を実施するための形態

[0010] 以下、本発明の実施の形態について、詳細に説明する。なお、同じ機能を有する構成部には同じ番号を付し、重複説明を省略する。

実施例 1

[0011] 以下、図1を参照して、実施例1の学習装置1、ファインチューニング装置2、推定装置3の機能構成を説明する。

[0012] <学習装置1>

同図に示すように本実施例の学習装置1は、データ処理部11と、事前学習部12と、事前学習モデルデータベース13を含む。

[0013] データ処理部11は、事前学習データとしてヒューマンモーションデータを取得し、ヒューマンモーションデータの任意の時刻群におけるデータ全体

をマスクし、ヒューマンモーションデータの一部分を他のヒューマンモーションデータの対応する一部分と置換する（S 1 1）。ヒューマンモーションデータとして、人間の姿勢情報（各関節の座標や関節角、ボーンの長さなど）を含むデータが一般的であるが、データの形式については限定されない。ヒューマンモーションデータが、関節の情報を含む場合、関節の数は任意に設定可能である。

[0014] 事前学習部 1 2 は、マスクの対象となったデータを事前学習モデルにより予測したデータである予測データと真値との誤差から損失関数を計算して事前学習モデルの重みを更新し、ヒューマンモーションデータが連続であるか不連続であるかを事前学習モデルにより予測した結果と真値との誤差から損失関数を計算して事前学習モデルの重みを更新し、事前学習モデルを学習する（S 1 2）。

[0015] 事前学習モデルデータベース 1 3 は、学習された事前学習モデルを記憶する（S 1 3）。

[0016] <ファインチューニング装置 2>

ファインチューニング装置 2 は、事前学習モデルデータベース 2 1 と、ファインチューニング部 2 2 と、ファインチューニングモデルデータベース 2 3 を含む。

[0017] 事前学習モデルデータベース 2 1 は、学習装置 1 により学習された事前学習モデルを記憶する（S 2 1）。

[0018] ファインチューニング部 2 2 は、学習装置 1 により学習された事前学習モデルをタスクに対応したファインチューニングデータにより更新してファインチューニングモデルを取得する（S 2 2）。ファインチューニングデータは、ファインチューニング先のタスクに依存する。例えば Human Motion の分類タスクの場合、ファインチューニングデータは、Human Motion データと各データに対するラベルのペアである。この場合、ファインチューニング部 2 2 は、ヒューマンモーションデータをモデルの入力としてクラス分布を予測し、真値ラベルとのクロスエントロピー誤差などを計算し、重みを更新し、

ネットワークの重みを出力する。ファインチューニングモデルデータベース 23 は、学習されたファインチューニングモデルを記憶する。

[0019] <推定装置 3>

推定装置 3 は、ファインチューニングモデルデータベース 31 と、推定部 32 を含む。

[0020] ファインチューニングモデルデータベース 31 は、ファインチューニング装置 2 が学習したファインチューニングモデルを記憶する。

[0021] 推定部 32 は、ラベルが未知の入力データ（例えば Human Motion の分類タスクの場合、Human Motion データのみ）を取得し、ファインチューニングモデルに基づいてクラス分布を推定し、推定されたラベル（推定値）を出力する（S32）。

[0022] 以下、図 3、図 4 を参照して学習装置 1 のデータ処理部 11 の詳細な機能構成及び詳細な動作を説明する。

[0023] <学習装置 1、データ処理部 11、S11>

図 3 に示すように、データ処理部 11 は、モーション置換部 111 と、マスク置換部 112 と、トークン挿入部 113 を含む。

[0024] モーション置換部 111 は、入力されたヒューマンモーションデータを前半部（モーション A）、後半部（モーション B）に分割し、所定の確率（例えば 50%）で後半部（モーション B）をデータセット中の別のモーションデータに置換する（S111）。モーション A とモーション B の分割例を図 5 に示す。

[0025] マスク置換部 112 は、入力されたヒューマンモーションデータのうち所定の割合のデータ（例えば 15% のデータ）をマスクに置き換える（S112）。

[0026] トークン挿入部 113 は、ヒューマンモーションデータにトークンを挿入する（S113）。トークンとして、[CLS]、[SEP] などがある。図 5 に例示するように、トークン挿入部 113 は、ヒューマンモーションデータの先頭に [CLS]、モーション A と B の間に [SEP]、ヒューマンモーションデータの末

尾に[SEP]を挿入する。Pose_tはそれぞれ、[関節数×自由度]次元で表現される。

[0027] データ処理部11は、ステップS111～S113の処理後のデータに対して埋め込みを行い、埋め込み処理後の各ベクトルを合計したものを出力する。

[0028] データ処理部11は、埋め込み処理として以下の1～4の処理を実行する。

[0029] <1. Pose Embedding>

BERTにおいてToken Embeddingに相当する処理である。Pose Embeddingは、Pose（入力）を特徴空間に埋め込む役割を果たす。データ処理部11は、データを任意の次元数の特徴ベクトルに変換する。変換の方法は任意の方法でよい。例えばデータ処理部11は、データを任意のDNNにより所望の次元数に変換してもよい。

<2. Position Embedding>

Position Embeddingは、各Poseの順番を埋め込む役割を果たす。例えばデータ処理部11は、0, 1, 2, 3, 4, 5, ..., 9というスカラー値を各PoseのPositionとして与え、これを任意の次元数の特徴ベクトルに変換する。変換の方法は任意の方法でよい。例えばデータ処理部11は、上記データを任意のDNNにより所望の次元数に変換してもよい。

<3. Segment Embedding>

Segment Embeddingは、モーションの分類（AかBか）を埋め込む役割を果たす。例えばデータ処理部11が、モーションAに対しては0、モーションBに対しては1を与える場合（0, 0, 0, 0, 0, 1, 1, 1, 1, 1）となる。データ処理部11は、このデータを任意の次元数の特徴ベクトルに変換する。変換の方法は任意の方法でよい。データ処理部11は、当該データを任意のDNNにより所望の次元数に変換してもよい。

[0030] <4. Token Embedding>

Token Embeddingは、トークンが全て同じ値の場合、どの入力かどのトーク

ンかをモデルが識別できるようにする役割を果たす。また、例えば[MASK]が線形補間の場合、[MASK]ではない通常のPoseか[MASK]かをモデルが識別できるようにする。例えばデータ処理部11は、[CLS]には0、[SEP]には1、[MASK]には2、[POSE]には3などの値を与え、このデータを任意の次元数の特徴ベクトルに変換する。変換の方法は任意の方法でよい。データ処理部11は、当該データを任意のDNNにより所望の次元数に変換してもよい。

[0031] [CLS]、[SEP]、[MASK]をHuman Motionタスクで扱うために、データ処理部11は、以下の四つのパターンの何れかによるデータ処理を行えば好適である。[CLS]、[SEP]、[MASK]の値をそれぞれ事前に定義する。

[0032] <パターン1>

例えば、Poseが1関節×3自由度の3次元で表現される場合、[CLS]=[0, 0, 0]、[SEP]=[1, 0, 2]、[MASK]=[2, 2, 2]などの数値を用いる。利用する値は任意で構わないが、各トークンで異なる値を用いる。入力データかトークンかをモデルが判別可能とするために、入力データ中にはない値であることが望ましい。例えば入力データ中に[1, 0, 2]が含まれるなら、トークンはそれ以外の値を利用する。

[0033] <パターン2>

[CLS]、[SEP]、[MASK]を全て同じ値で埋め、Token Embeddingを追加する。例えばPoseが1関節×3自由度の3次元で表現される場合、[CLS]=[0, 0, 0]、[SEP]=[0, 0, 0]、[MASK]=[0, 0, 0]などの数値を用いる。利用する値は任意で構わないが、入力データ中にはない値が望ましい。前述したToken Embeddingによりトークンの判別を行う。

[0034] <パターン3>

[CLS]、[SEP]の値をそれぞれ事前に定義し（パターン1と同様）、[MASK]は前後のPoseの値から線形補間する。さらにToken Embeddingを追加する。[CLS]、[SEP]はパターン1と同様である。例えば[MASK]の例として、Poseが1関節×3自由度の3次元で表現される場合、時刻tのPoseを[MASK]に置換し、時刻t-1のPoseが[0, 1, 0]、時刻t+1のPoseが[0, 0, 1]である場合、時

刻 $t-1$ と $t+1$ のPoseから線形補完して[MASK]=[0, 0.5, 0.5]とする。[MASK]が2つ以上連続する可能性もある。この場合にも前後のPoseから線形補完を行う。例えば、時刻 t , $t+1$ が[MASK]の場合は、時刻 $t-1$ 、 $t+2$ から線形補完を行う。前述したToken Embeddingによりトークンの判別を行う。

[0035] <パターン4>

[CLS]、[SEP]の値を全て同じ値で埋め（パターン2と同様）、[MASK]は前後のPoseの値から線形補間する（パターン3と同様）。さらにToken Embeddingを追加する。[CLS]、[SEP]はパターン2と同様、[MASK]はパターン3と同様。前述したToken Embeddingによりトークンの判別を行う。

事前学習部12は、埋め込み処理後の特徴ベクトルを受け取り、MLM (=Masked Language Modeling)、NSP (=Next Sentence Prediction) タスクを実行することでモデルの学習を行う。以下、図7、図8を参照して事前学習部12の詳細な機能構成及び詳細な動作を説明する。

[0036] <学習装置1、事前学習部12、S12>

図7に示すように事前学習部12は、マスク置換前データ予測部121と、モーション連続性予測部122と、損失関数計算部123と、重み更新部124と、真値記憶部125を含む。

[0037] マスク置換前データ予測部121は、事前学習モデルに基づき、マスクに置換されたPoseの元の値の予測を出力する（S121）。

[0038] モーション連続性予測部122は、事前学習モデルに基づき、モーションAとモーションBが連続か／不連続かの予測を行う（S122）。

[0039] 損失関数計算部123は、ステップS121の結果について、予測Poseと真値Pose間の誤差から損失関数を計算する。損失関数は何を用いてもよい。損失関数として、例えばL1 distanceやL2 distanceを用いることができる。また、損失関数計算部123は、ステップS122の結果について、連続の場合は0、不連続の場合は1などのワンホットベクトルを真値として予め用意しておき、2値分類を行い、損失関数としてクロスエントロピーを用いて、損失関数を計算する（S123）。

[0040] 重み更新部 124 は、損失関数が最小になるように、モデルの重みを更新する（S124）。

[0041] 真値記憶部 125 には、ステップ S123 で用いる真値が記憶されている。

[0042] <補記>

本発明の装置は、例えば単一のハードウェアエンティティとして、キーボードなどが接続可能な入力部、液晶ディスプレイなどが接続可能な出力部、ハードウェアエンティティの外部に通信可能な通信装置（例えば通信ケーブル）が接続可能な通信部、CPU（Central Processing Unit、キャッシュメモリやレジスタなどを備えていてもよい）、メモリであるRAMやROM、ハードディスクである外部記憶装置並びにこれらの入力部、出力部、通信部、CPU、RAM、ROM、外部記憶装置の間のデータのやり取りが可能なように接続するバスを有している。また必要に応じて、ハードウェアエンティティに、CD-ROMなどの記録媒体を読み書きできる装置（ドライブ）などを設けることとしてもよい。このようなハードウェア資源を備えた物理的実体としては、汎用コンピュータなどがある。

[0043] ハードウェアエンティティの外部記憶装置には、上述の機能を実現するために必要となるプログラムおよびこのプログラムの処理において必要となるデータなどが記憶されている（外部記憶装置に限らず、例えばプログラムを読み出し専用記憶装置であるROMに記憶しておくこととしてもよい）。また、これらのプログラムの処理によって得られるデータなどは、RAMや外部記憶装置などに適宜に記憶される。

[0044] ハードウェアエンティティでは、外部記憶装置（あるいはROMなど）に記憶された各プログラムとこの各プログラムの処理に必要なデータが必要に応じてメモリに読み込まれて、適宜にCPUで解釈実行・処理される。その結果、CPUが所定の機能（上記、…部、…手段などと表した各構成要件）を実現する。

[0045] 本発明は上述の実施形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲で適宜変更が可能である。また、上記実施形態において説明した

処理は、記載の順に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的にあるいは個別に実行されるとしてもよい。

[0046] 既述のように、上記実施形態において説明したハードウェアエンティティ（本発明の装置）における処理機能をコンピュータによって実現する場合、ハードウェアエンティティが有すべき機能の処理内容はプログラムによって記述される。そして、このプログラムをコンピュータで実行することにより、上記ハードウェアエンティティにおける処理機能がコンピュータ上で実現される。

[0047] 上述の各種の処理は、図9に示すコンピュータ10000の記録部10020に、上記方法の各ステップを実行させるプログラムを読み込ませ、制御部10010、入力部10030、出力部10040などに動作させることで実施できる。

[0048] この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体としては、例えば、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等のようなものでもよい。具体的には、例えば、磁気記録装置として、ハードディスク装置、フレキシブルディスク、磁気テープ等を、光ディスクとして、DVD（Digital Versatile Disc）、DVD-RAM（Random Access Memory）、CD-ROM（Compact Disc Read Only Memory）、CD-R（Recordable）／RW（ReWritable）等を、光磁気記録媒体として、MO（Magneto-Optical disc）等を、半導体メモリとしてEEPROM（Electrically Erasable and Programmable-Read Only Memory）等を用いることができる。

[0049] また、このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータ

にそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

[0050] このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。そして、処理の実行時、このコンピュータは、自己の記録媒体に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。また、このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。また、サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP (Application Service Provider) 型のサービスによって、上述の処理を実行する構成としてもよい。なお、本形態におけるプログラムには、電子計算機による処理の用に供する情報であってプログラムに準ずるもの（コンピュータに対する直接の指令ではないがコンピュータの処理を規定する性質を有するデータ等）を含むものとする。

[0051] また、この形態では、コンピュータ上で所定のプログラムを実行させることにより、ハードウェアエンティティを構成することとしたが、これらの処理内容の少なくとも一部をハードウェア的に実現することとしてもよい。

請求の範囲

- [請求項1] ヒューマンモーションデータを取得し、前記ヒューマンモーションデータの任意の時刻群におけるデータ全体をマスクするデータ処理部と、
- マスクの対象となったデータを事前学習モデルにより予測したデータである予測データと真値との誤差から損失関数を計算し、前記事前学習モデルの重みを更新することにより前記事前学習モデルを学習する事前学習部を含む
- 学習装置。
- [請求項2] 請求項1に記載の学習装置であって、
- 前記データ処理部は、
- 所定の確率で、前記ヒューマンモーションデータの一部分を他の前記ヒューマンモーションデータの対応する一部分と置換し、
- 前記事前学習部は、
- 前記ヒューマンモーションデータが連続であるか不連続であるかを前記事前学習モデルにより予測した結果と真値との誤差から損失関数を計算し、前記前記事前学習モデルの重みを更新する
- 学習装置。
- [請求項3] ヒューマンモーションデータの任意の時刻群におけるデータ全体をマスクして得た教師データから学習された事前学習モデルを、タスクに対応したファインチューニングデータにより更新してなるファインチューニングモデルを記憶するファインチューニングモデルデータベースと、
- 前記ヒューマンモーションデータを取得し、前記ファインチューニングモデルに基づいて、前記ヒューマンモーションデータの前記タスクに応じたラベルを推定する推定部を含む
- 推定装置。
- [請求項4] 学習装置が実行する学習方法であって、

ヒューマンモーションデータを取得し、前記ヒューマンモーションデータの任意の時刻群におけるデータ全体をマスクするデータ処理ステップと、

マスクの対象となったデータを事前学習モデルにより予測したデータである予測データと真値との誤差から損失関数を計算し、前記事前学習モデルの重みを更新することにより前記事前学習モデルを学習する事前学習ステップを含む

学習方法。

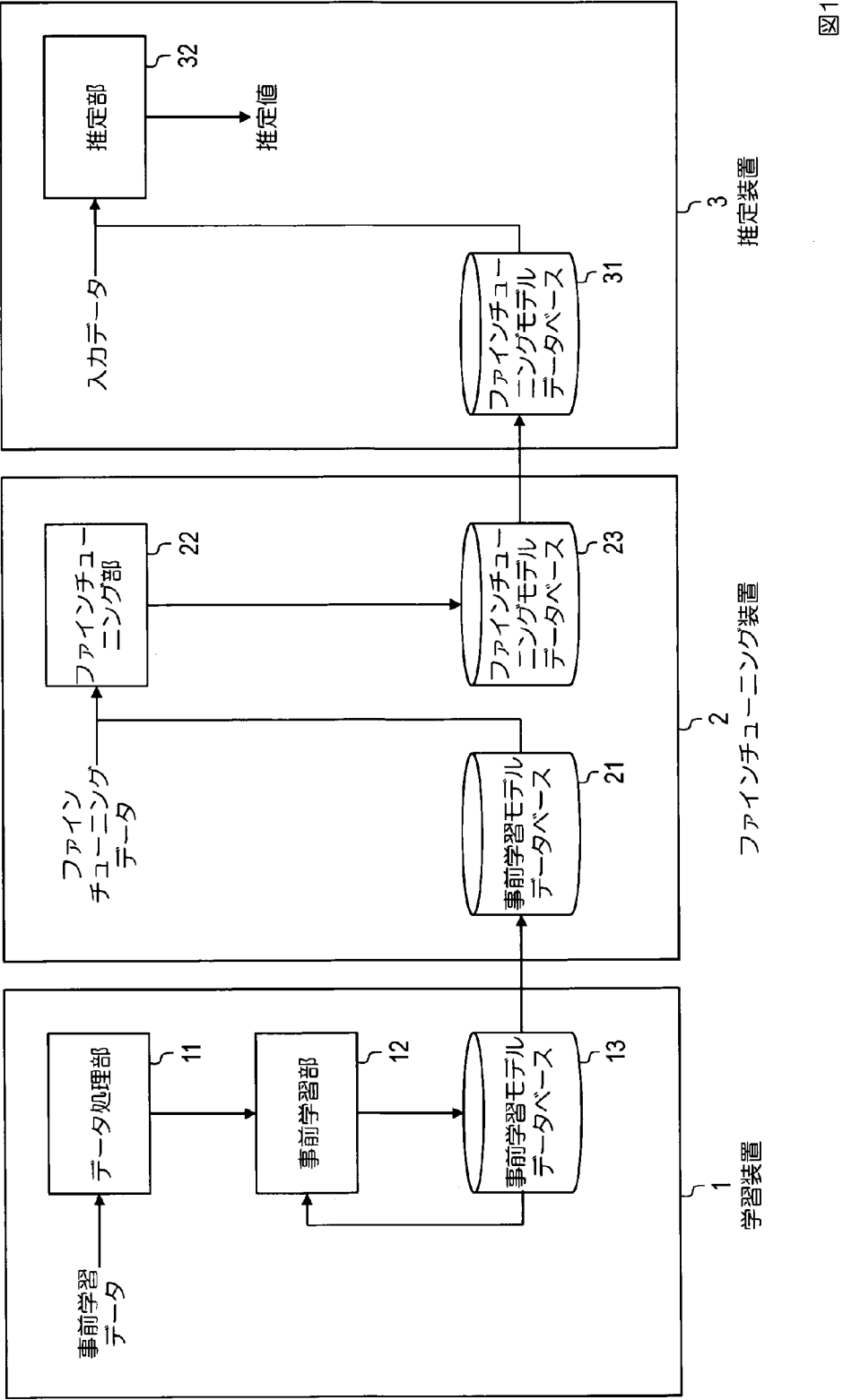
[請求項5] 推定装置が実行する推定方法であって、

ヒューマンモーションデータの任意の時刻群におけるデータ全体をマスクして得た教師データから学習された事前学習モデルを、タスクに対応したファインチューニングデータにより更新してなるファインチューニングモデルに基づいて、取得した前記ヒューマンモーションデータの前記タスクに応じたラベルを推定する推定ステップを含む推定方法。

[請求項6] コンピュータを請求項 1 または 2 に記載の学習装置として機能させるプログラム。

[請求項7] コンピュータを請求項 3 に記載の推定装置として機能させるプログラム。

[図1]



[図2]

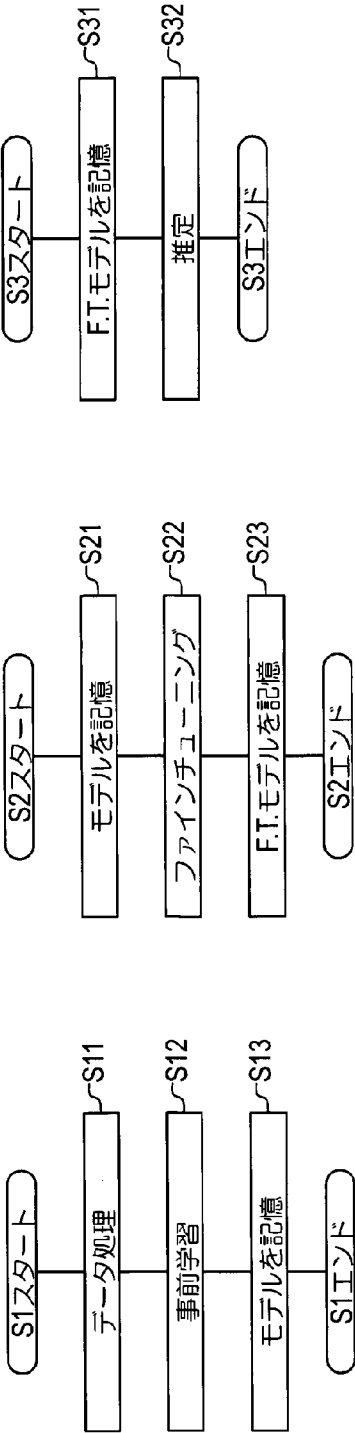


図2

[図3]

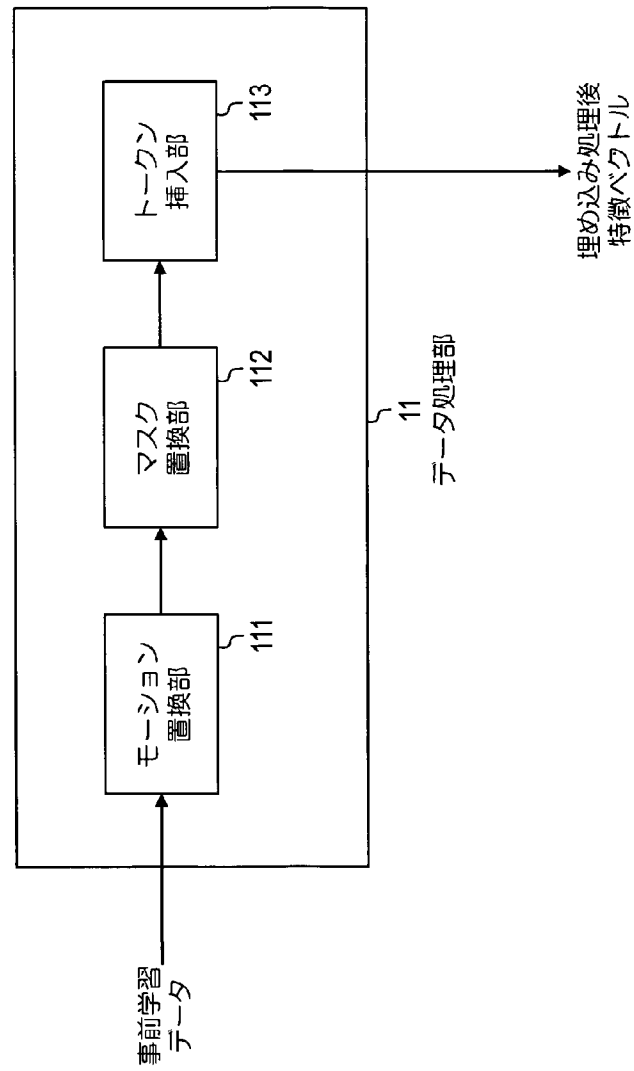
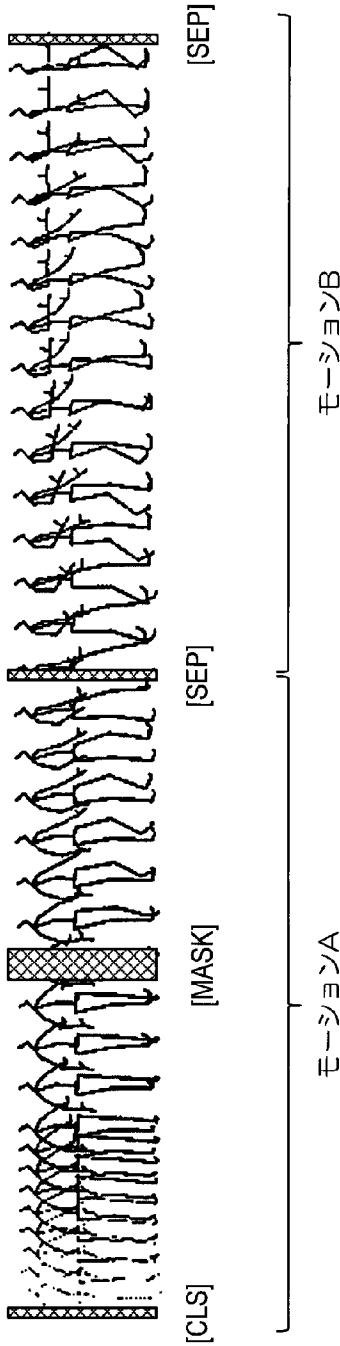


図3

[図4]

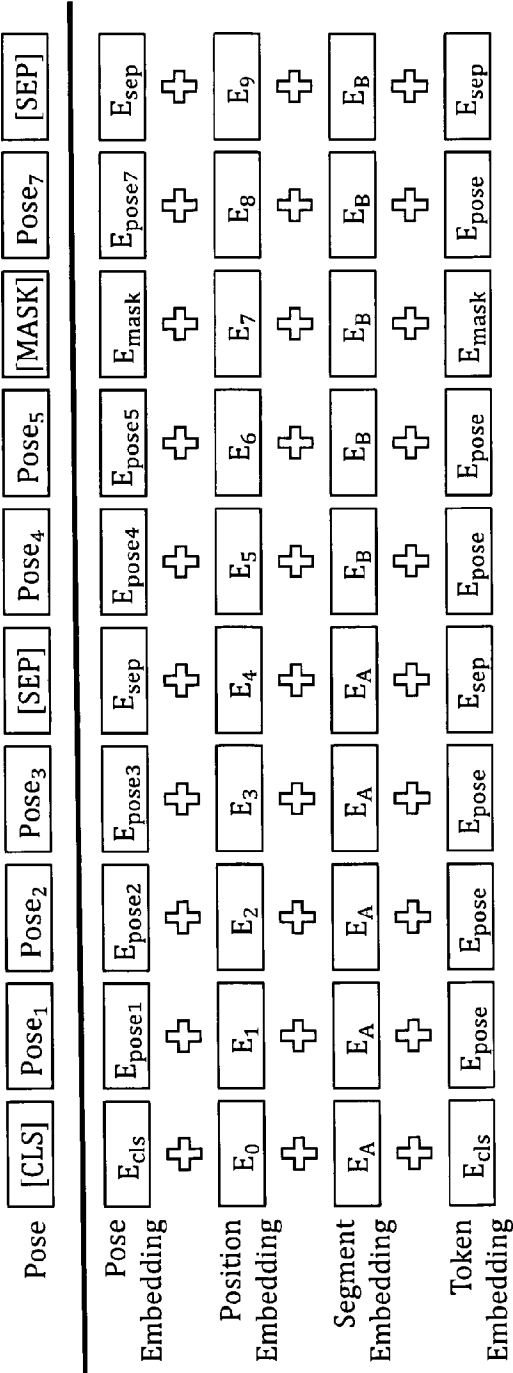


[図5]



[図5]

[図6]



[図7]

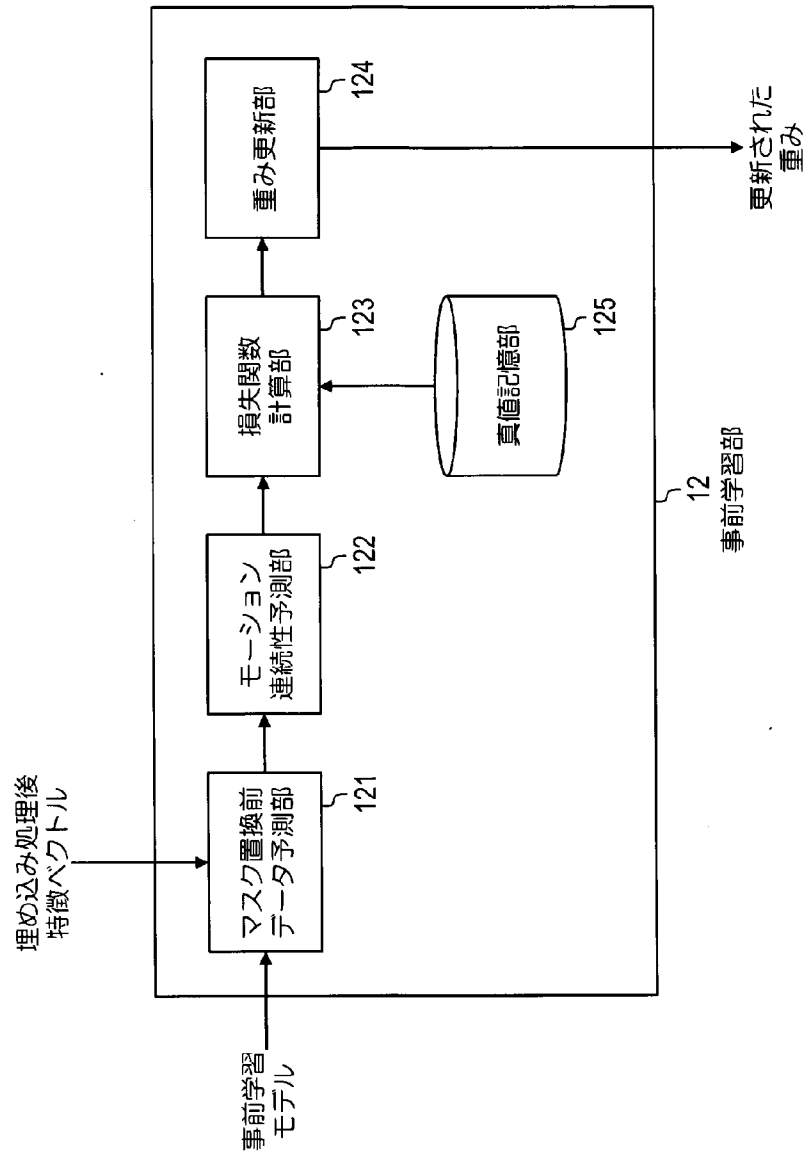
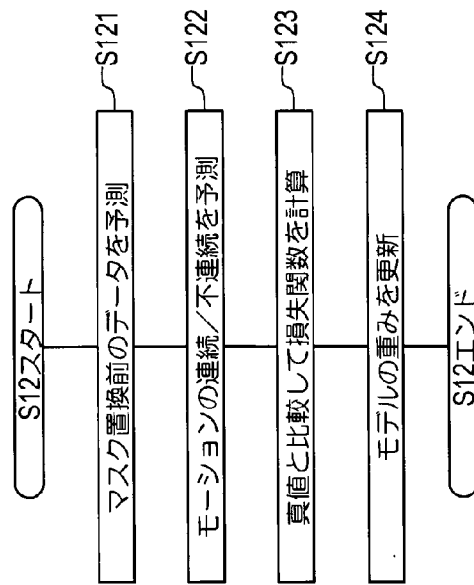


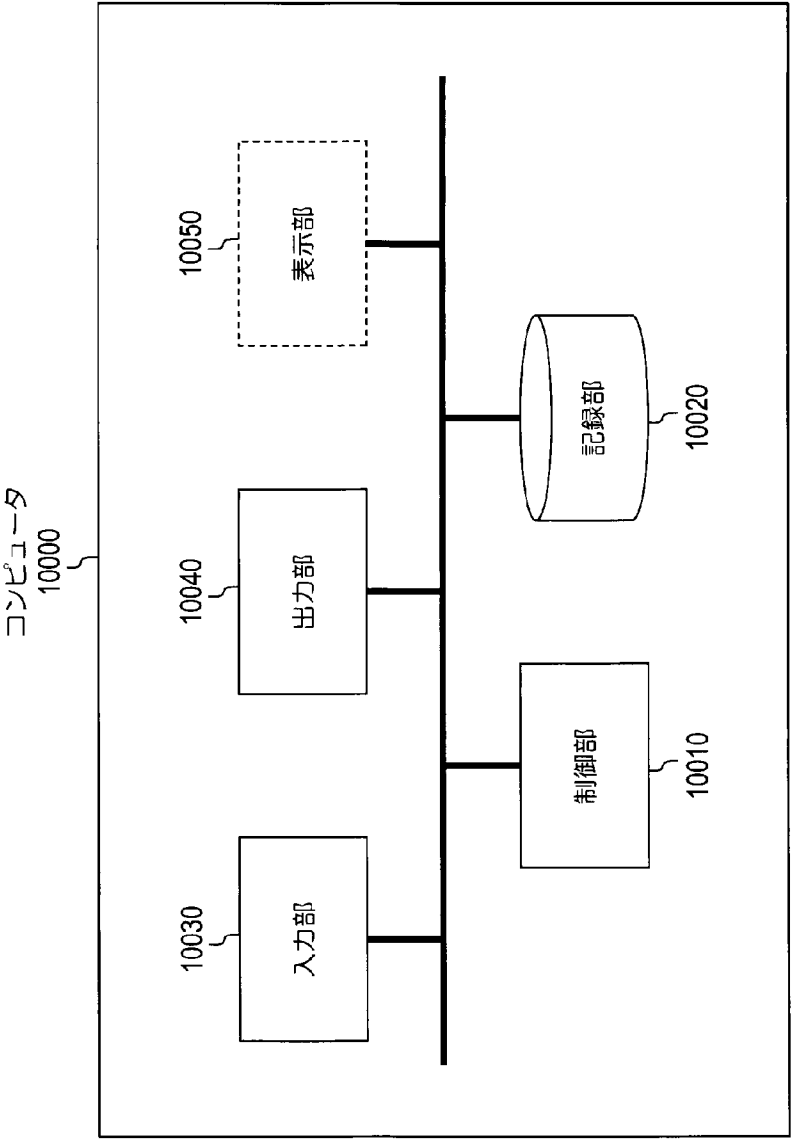
図7

[図8]



[図8]

[図9]



[図9]

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/018416

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/00(2017.01)i; G06T 7/20(2017.01)i; G06N 3/08(2006.01)i; G06N 20/00(2019.01)i

FI: G06N20/00 130; G06N3/08; G06T7/00 350B; G06T7/20 300Z

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)

G06T7/00; G06T7/20; G06N3/00-99/00

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

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2021

Registered utility model specifications of Japan 1996-2021

Published registered utility model applications of Japan 1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

IEEE Xplore; THE ACM DIGITAL LIBRARY

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	CHENG, Yi-Bin et al. Motion-Transformer: Self-supervised Pre-training for Skeleton-based Action Recognition. MMAsia'20: Proceedings of the 2nd ACM International Conference on Multimedia in Asia [online]. Association for Computing Machinery., March 2021, article no. 36, pp. 1-6, [retrieval date 28 June 2021], Internet:<URL:https://dl.acm.org/doi/abs/10.1145/3444685.3446289>, <DOI:10.1145/3444685.3446289>, abstract, chapter 1, sections 3.2-3.4, fig. 1	1, 3-7 2
X	JIANG, Tao et al. Skeletor: Skeletal Transformers for Robust Body-Pose Estimation. arXiv.org [online]. arXiv: 2104.11712v1., Cornell University. April 2021, [retrieval date 28 June 2021], Internet:<URL:https://arxiv.org/pdf/2104.11712v1>, chapter 1, sections 4.1, 5.1,	1, 4, 6
A	chapter 6	2

☐ 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

28 June 2021 (28.06.2021)

Date of mailing of the international search report

13 July 2021 (13.07.2021)

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

A. 発明の属する分野の分類（国際特許分類（IPC）） G06T 7/00(2017.01)i; G06T 7/20(2017.01)i; G06N 3/08(2006.01)i; G06N 20/00(2019.01)i FI: G06N20/00 130; G06N3/08; G06T7/00 350B; G06T7/20 300Z										
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）） G06T7/00; G06T7/20; G06N3/00-99/00										
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 <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922-1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971-2021年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2021年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2021年</td> </tr> </table>			日本国実用新案公報	1922-1996年	日本国公開実用新案公報	1971-2021年	日本国実用新案登録公報	1996-2021年	日本国登録実用新案公報	1994-2021年
日本国実用新案公報	1922-1996年									
日本国公開実用新案公報	1971-2021年									
日本国実用新案登録公報	1996-2021年									
日本国登録実用新案公報	1994-2021年									
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語） IEEE Xplore; THE ACM DIGITAL LIBRARY										
C. 関連すると認められる文献										
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号								
X	CHENG, Yi-Bin et al., Motion-Transformer: Self-supervised Pre-training for Skeleton-based Action Recognition, MMAsia '20: Proceedings of the 2nd ACM International Conference on Multimedia in Asia [online], Association for Computing Machinery, 2021.03, article no. 36, pp. 1-6, [検索日 2021.06.28], インターネット:<URL: https://dl.acm.org/doi/abs/10.1145/3444685.3446289><DOI:10.1145/3444685.3446289> ABSTRACT, 第1章, 第3.2節-第3.4節, 図1	1, 3-7								
A		2								
X	JIANG, Tao et al., Skeletor: Skeletal Transformers for Robust Body-Pose Estimation, arXiv.org [online], arXiv: 2104.11712v1, Cornell University, 2021.04, [検索日 2021.06.28], インターネット:<URL: https://arxiv.org/pdf/2104.11712v1> 第1章, 第4.1節, 第5.1節	1, 4, 6								
A	第6章	2								
☐ C欄の続きにも文献が列挙されている。 ☐ パテントファミリーに関する別紙を参照。										
* 引用文献のカテゴリー “A” 特に関連のある文献ではなく、一般的な技術水準を示すもの “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献										
国際調査を完了した日 28.06.2021		国際調査報告の発送日 13.07.2021								
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） 多賀 実 5B 9367 電話番号 03-3581-1101 内線 3545								