

(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
2023年10月5日(05.10.2023)



(10) 国際公開番号
WO 2023/188216 A1

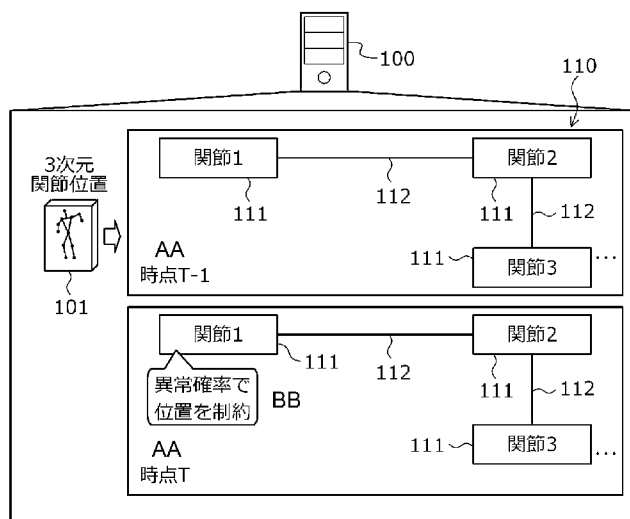
- (51) 国際特許分類:
G06T 7/00 (2017.01)
- (21) 国際出願番号: PCT/JP2022/016363
- (22) 国際出願日: 2022年3月30日(30.03.2022)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (71) 出願人: 富士通株式会社 (FUJITSU LIMITED) [JP/JP]; 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 Kanagawa (JP).
- (72) 発明者: 小田嶋 成幸 (ODASHIMA, Shigeyuki); 〒2118588 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内 Kanagawa (JP).

山尾 創輔(YAMAOKI, Sosuke); 〒2118588 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 Kanagawa (JP). 鈴木 達也(SUZUKI, Tatsuya); 〒2118588 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 Kanagawa (JP). 楊 帆(YANG, Fan); 〒2118588 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 Kanagawa (JP). 榎井 昇一(MASUI, Shoichi); 〒2118588 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 Kanagawa (JP). 姜 山(JIANG, Shan); 〒2118588 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 Kanagawa (JP).

(54) Title: INFORMATION PROCESSING PROGRAM, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING DEVICE

(54) 発明の名称: 情報処理プログラム、情報処理方法、および情報処理装置

[図1]



- 101 Three-dimensional joint position
- 111 Joint
- AA Point in time
- BB Constrain position by abnormal probability

(57) Abstract: An information processing device (100) identifies, with regard to skeleton information (101) at a first point in time in time series data, any site that is in an abnormal state in relation to position on the basis of a feature amount relating to the skeleton information (101) in the time series data. The information processing device (100) determines a probability distribution model that constrains the position of the identified any site in accordance with the magnitude of a probability that the any site that was identified with regard to the skeleton information (101) at the first point in time is



WO 2023/188216 A1

(74) 代理人: 酒井 昭徳(SAKAI, Akinori); 〒1020084
東京都千代田区二番町 1 1 - 3 相互二番町ビル
ディング別館 酒井総合特許事務所 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, JM, 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, 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, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, 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))

in an abnormal state. The information processing device (100) generates a graph (110) which includes first edges (112) and nodes (111) indicating the position of each site at each point in time, and in which the determined model has been associated with the node (111) indicating the position of the any site that was identified. On the basis of the generated graph (110), the information processing device (100) corrects the skeleton information (101) at the first point in time in the time series data.

(57) 要約: 情報処理装置 (100) は、時系列データにおける骨格情報 (101) に関する特徴量に基づいて、時系列データにおける第1時点の骨格情報 (101) について、位置に関して異常な状態であるいずれかの部位を特定する。情報処理装置 (100) は、第1時点の骨格情報 (101) において、特定したいずれかの部位の位置を、当該部位の異常な状態である確率の大きさに応じて制約する確率分布のモデルを決定する。情報処理装置 (100) は、時点ごとのそれぞれの部位の位置を示すノード (111) と、第1エッジ (112) とを含み、特定したいずれかの部位の位置を示すノード (111) に、決定したモデルを対応付けたグラフ (110) を生成する。情報処理装置 (100) は、生成したグラフ (110) に基づいて、時系列データにおける第1時点の骨格情報 (101) を修正する。

明 細 書

発明の名称：

情報処理プログラム、情報処理方法、および情報処理装置

技術分野

[0001] 本発明は、情報処理プログラム、情報処理方法、および情報処理装置に関する。

背景技術

[0002] 従来、スポーツ、ヘルスケア、または、エンターテインメントの分野において、人の動きを認識する技術が望まれる。例えば、深層学習を利用して、異なるアングルから撮像した多視点画像に基づいて、人のそれぞれの関節の3次元座標を特定する技術がある。

[0003] 先行技術としては、例えば、第1処理の結果の尤度と、第2処理の結果の尤度と、第3処理の結果の尤度とに基づいて、第1処理の結果、第2処理の結果、第3処理の結果のうち、いずれかの結果を、被写体の骨格認識結果として出力するものがある。また、例えば、被写体の距離画像から、被写体の複数の関節位置の尤度を複数方向から投影したヒートマップ画像を認識する技術がある。また、例えば、特徴点の位置候補と対象の多関節構造を用いた逆運動学に基づく最適化計算を行い、対象の各関節角を取得し、関節角を用いて順運動学計算を行い、対象の関節を含む特徴点の位置を取得する技術がある。また、例えば、リカレントニューラルネットワークを利用する行動検出の技術がある。

先行技術文献

特許文献

[0004] 特許文献1：国際公開第2021/064942号

特許文献2：国際公開第2021/002025号

特許文献3：特開2020-42476号公報

特許文献4：米国特許出願公開第2017/0344829号明細書

発明の概要

発明が解決しようとする課題

[0005] しかしながら、従来技術では、人のそれぞれの関節の3次元座標を精度よく特定することが難しい場合がある。例えば、人の右手の関節の3次元座標を、人の左手の関節の3次元座標と誤認することがある。例えば、多視点画像に映った人以外の物体の一部の3次元座標を、人の関節の3次元座標と誤認することがある。

[0006] 1つの側面では、本発明は、対象者の部位の位置を精度よく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] 1つの実施態様によれば、対象者の複数の部位のそれぞれの部位の位置を含む骨格情報の時系列データを取得し、取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、取得した前記時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定し、取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、前記第1時点の骨格情報において、特定した前記いずれかの部位の位置を、特定した前記いずれかの部位の異常な状態である確率の大きさに応じて制約する確率分布のモデルを決定し、時点ごとの前記それぞれの部位の位置を示すノードと、前記時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジとを含み、前記いずれかの部位の位置を示すノードに決定した前記モデルを対応付けたグラフを生成し、生成した前記グラフに基づいて、前記時系列データにおける前記第1時点の骨格情報を修正する情報処理プログラム、情報処理方法、および情報処理装置が提案される。

発明の効果

[0008] 一態様によれば、対象者の部位の位置を精度よく特定可能にすることが可能になる。

図面の簡単な説明

[0009] [図1]図1は、実施の形態にかかる情報処理方法の一実施例を示す説明図である。

[図2]図2は、情報処理システム200の一例を示す説明図である。

[図3]図3は、情報処理装置100のハードウェア構成例を示すブロック図である。

[図4]図4は、画像撮像装置201のハードウェア構成例を示すブロック図である。

[図5]図5は、情報処理装置100の機能的構成例を示すブロック図である。

[図6]図6は、情報処理装置100の動作の流れを示す説明図である。

[図7]図7は、異常関節を特定する具体例を示す説明図（その1）である。

[図8]図8は、異常関節を特定する具体例を示す説明図（その2）である。

[図9]図9は、Factor Graph 900を生成する具体例を示す説明図である。

[図10]図10は、3D骨格推論結果602を修正する具体例を示す説明図である。

[図11]図11は、動作例におけるデータ処理の流れの具体例を示す説明図（その1）である。

[図12]図12は、動作例におけるデータ処理の流れの具体例を示す説明図（その2）である。

[図13]図13は、全体処理手順の一例を示すフローチャートである。

発明を実施するための形態

[0010] 以下に、図面を参照して、本発明にかかる情報処理プログラム、情報処理方法、および情報処理装置の実施の形態を詳細に説明する。

[0011] （実施の形態にかかる情報処理方法の一実施例）

図1は、実施の形態にかかる情報処理方法の一実施例を示す説明図である。情報処理装置100は、対象者の関節の位置を精度よく特定可能にするためのコンピュータである。対象者は、例えば、人である。位置は、例えば、

3次元座標である。

[0012] 従来、深層学習を利用して、時点ごとに、異なるアングルから撮像した多視点画像に基づいて、人のそれぞれの関節の3次元座標を特定することにより、人のそれぞれの関節の3次元座標の時間変化を特定するための技術がある。

[0013] 具体的には、多視点画像のうち、人が映った領域を検出し、検出した領域に基づいて人のそれぞれの関節の2次元座標を特定し、アングルを考慮して、特定した2次元座標に基づいて、人のそれぞれの関節の3次元座標を特定することが考えられる。具体的には、人のそれぞれの関節の3次元座標を特定するにあたり、深層学習で学習したモデルが利用される。この技術の一例については、具体的には、下記参考文献1および下記参考文献2を参照することができる。

[0014] 参考文献1 : Iskakov, Karim, et al. "Learnable triangulation of human pose." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2019.

[0015] 参考文献2 : Moon, Gyeongsik, Ju Yong Chang, and Kyoung Mu Lee. "V2v-pose-net: Voxel-to-voxel prediction network for accurate 3d hand and human pose estimation from a single depth map." Proceedings of the IEEE conference on computer vision and pattern Recognition. 2018.

[0016] しかしながら、従来技術では、人のそれぞれの関節の3次元座標を精度よく特定することが難しい場合がある。例えば、異なる時点でも同一人のそれぞれの関節間の距離は一定であるものの、深層学習にてモデルを学習する際

、人のそれぞれの関節間の距離が一定であることを考慮していない。このため、人のそれぞれの関節の3次元座標を精度よく特定することができず、人のそれぞれの関節の3次元座標の時間変化を精度よく特定することができない。

[0017] 人の動きを解析する解析者は、特定した人のそれぞれの関節の3次元座標を参照した際、人のそれぞれの関節の3次元座標が、直感的に誤っているという印象を覚えることになり易い。具体的には、解析者は、人の腕の長さが伸び縮みしているという印象を覚えることになる。また、具体的には、解析者は、人が出し得ない速度で人の腕が動いているという印象を覚えることになる。

[0018] そこで、本実施の形態では、対象者の関節の位置を精度よく特定可能にすることができる情報処理方法について説明する。

[0019] 図1において、(1-1)情報処理装置100は、骨格情報101の時系列データを取得する。骨格情報101は、例えば、対象者の複数の部位のそれぞれの部位の位置を含む。部位は、例えば、首、頭、右肩および左肩、右肘および左肘、右手および左手、右膝および左膝、右足および左足などである。部位は、具体的には、関節である。図1の例では、部位は、具体的には、関節1と、関節2と、関節3となどである。位置は、例えば、3次元座標である。時系列データは、例えば、時点ごとの骨格情報101を含む。図1の例では、時系列データは、具体的には、時点Tの骨格情報101、時点T-1の骨格情報101などを含む。

[0020] (1-2)情報処理装置100は、取得した時系列データにおける骨格情報101に関する特徴量に基づいて、取得した時系列データにおける第1時点の骨格情報101について、位置に関して異常な状態であるいずれかの部位を特定する。特徴量は、例えば、骨格情報101が示す対象者のそれぞれの部位の位置であってもよい。特徴量は、例えば、異なる時点の骨格情報101が示す対象者のそれぞれの部位の位置の偏差であってもよい。特徴量は、例えば、骨格情報101が示す対象者の異なる部位の位置間の距離であつ

てもよい。

[0021] 情報処理装置100は、例えば、対象者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定するための第1モデルを有する。第1モデルは、例えば、骨格情報101に関する特徴量の入力に応じて、対象者の複数の部位のそれぞれの部位の位置が、異常な状態である確率の大きさを算出し、それぞれの部位の位置が、異常な状態であるか否かを判定可能にする機能を有する。情報処理装置100は、例えば、第1モデルを利用して、取得した時系列データにおける第1時点の骨格情報101について、位置に関して異常な状態であるいずれかの部位を特定する。

[0022] (1-3) 情報処理装置100は、取得した時系列データにおける骨格情報101に関する特徴量に基づいて、第2モデルを決定する。情報処理装置100は、例えば、第1時点の骨格情報101において、特定したいずれかの部位の位置を、当該部位の異常な状態である確率の大きさに応じて制約する確率分布の第2モデルを決定する。確率は、例えば、時系列データにおける骨格情報101に関する特徴量に基づいて、第1モデルによって算出される。図1の例では、情報処理装置100は、具体的には、時点Tの骨格情報101において、関節1の位置を、関節1の異常な状態である確率の大きさに応じて制約する確率分布の第2モデルを決定する。

[0023] (1-4) 情報処理装置100は、時点ごとのそれぞれの部位の位置を示すノード111と、時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジ112とを含むグラフ110を生成する。情報処理装置100は、グラフ110を生成する際、特定したいずれかの部位の位置を示すノード111に、決定した第2モデルを対応付ける。図1の例では、情報処理装置100は、具体的には、時点Tの対象者の関節1の位置を示すノード111に、決定した第2モデルを対応付けて、グラフ110を生成する。

[0024] (1-5) 情報処理装置100は、生成したグラフ110に基づいて、時系列データにおける第1時点の骨格情報101を修正する。情報処理装置1

00は、例えば、時系列データにおける時点Tの骨格情報101に含まれる対象者の関節1の位置を修正する。これにより、情報処理装置100は、対象者のそれぞれの関節の位置を精度よく特定することができる。情報処理装置100は、対象者のそれぞれの関節の位置の時間変化を精度よく特定することができる。

[0025] ここでは、情報処理装置100が、第1モデルを利用して、時系列データにおける第1時点の骨格情報101について、位置に関して異常な状態であるいずれかの部位を特定する場合について説明したが、これに限らない。例えば、情報処理装置100が、第1モデルを利用せずに、時系列データにおける第1時点の骨格情報101について、位置に関して異常な状態であるいずれかの部位を特定する場合があってもよい。

[0026] ここでは、情報処理装置100が、単独で動作する場合について説明したが、これに限らない。例えば、複数のコンピュータが協働して、情報処理装置100としての機能を実現する場合があってもよい。具体的には、位置に関して異常な状態であるいずれかの部位を特定するコンピュータと、グラフ110を生成するコンピュータと、グラフ110に基づいて、時系列データにおける第1時点の骨格情報101を修正するコンピュータとが協働する場合があってもよい。

[0027] (情報処理システム200の一例)

次に、図2を用いて、図1に示した情報処理装置100を適用した、情報処理システム200の一例について説明する。

[0028] 図2は、情報処理システム200の一例を示す説明図である。図2において、情報処理システム200は、情報処理装置100と、1以上の画像撮像装置201と、1以上のクライアント装置202とを含む。

[0029] 情報処理システム200において、情報処理装置100と画像撮像装置201とは、有線または無線のネットワーク210を介して接続される。ネットワーク210は、例えば、LAN (Local Area Network)、WAN (Wide Area Network)、インターネットな

どである。また、情報処理システム200において、情報処理装置100とクライアント装置202とは、有線または無線のネットワーク210を介して接続される。

[0030] 情報処理装置100は、時点ごとに異なるアングルから対象者を撮像した複数の画像を、1以上の画像撮像装置201から取得する。情報処理装置100は、時点ごとに、取得した複数の画像に基づいて、3次元空間における対象者のそれぞれの部位の存在確率の分布を特定し、対象者のそれぞれの部位の3次元座標を特定する。情報処理装置100は、時点ごとに、取得した複数の画像に基づいて、3次元座標に関して対象者のそれぞれの部位の異常度を算出し、対象者のいずれかの部位を、3次元座標に関して異常な状態である異常部位として特定する。

[0031] 情報処理装置100は、算出した異常度に基づいて、特定した異常部位の位置を制約する確率分布のモデルを決定する。情報処理装置100は、特定したそれぞれの時点の対象者のそれぞれの部位の3次元座標を示すノードと、ノード間を接続するエッジとを含み、異常部位の3次元座標を示すノードに、決定したモデルを対応付けたグラフを生成する。情報処理装置100は、グラフを参照して、特定した対象者のそれぞれの部位の3次元座標を修正する。

[0032] 情報処理装置100は、修正した対象者のそれぞれの部位の3次元座標を出力する。出力形式は、例えば、ディスプレイへの表示、プリンタへの印刷出力、他のコンピュータへの送信、または、記憶領域への記憶などである。情報処理装置100は、例えば、修正した対象者のそれぞれの部位の3次元座標を、クライアント装置202に送信する。情報処理装置100は、例えば、サーバ、または、PC (Personal Computer) などである。

[0033] 画像撮像装置201は、対象者を撮像するコンピュータである。画像撮像装置201は、複数の撮像素子を有するカメラを含み、カメラによって対象者を撮像する。画像撮像装置201は、対象者を撮像した画像を生成し、情

報処理装置100に送信する。画像撮像装置201は、例えば、スマートフォンなどである。画像撮像装置201は、例えば、定点カメラなどであってもよい。画像撮像装置201は、例えば、ドローンなどであってもよい。

[0034] クライアント装置202は、対象者のそれぞれの部位の3次元座標を、情報処理装置100から受信する。クライアント装置202は、受信した対象者のそれぞれの部位の3次元座標を、利用者が参照可能に出力する。クライアント装置202は、例えば、受信した対象者のそれぞれの部位の3次元座標を、ディスプレイに表示する。クライアント装置202は、例えば、PC、タブレット端末、または、スマートフォンなどである。

[0035] ここでは、情報処理装置100が、画像撮像装置201とは異なる装置である場合について説明したが、これに限らない。例えば、情報処理装置100が、画像撮像装置201としての機能を有し、画像撮像装置201としても動作する場合があってもよい。ここでは、情報処理装置100が、クライアント装置202とは異なる装置である場合について説明したが、これに限らない。例えば、情報処理装置100が、クライアント装置202としての機能を有し、クライアント装置202としても動作する場合があってもよい。

[0036] (情報処理装置100のハードウェア構成例)

次に、図3を用いて、情報処理装置100のハードウェア構成例について説明する。

[0037] 図3は、情報処理装置100のハードウェア構成例を示すブロック図である。図3において、情報処理装置100は、CPU (Central Processing Unit) 301と、メモリ302と、ネットワークI/F (Interface) 303と、記録媒体I/F 304と、記録媒体305とを有する。情報処理装置100は、さらに、ディスプレイ306と、入力装置307とを有する。また、各構成部は、バス300によってそれぞれ接続される。

[0038] ここで、CPU 301は、情報処理装置100の全体の制御を司る。メモ

リ302は、例えば、ROM (Read Only Memory)、RAM (Random Access Memory) およびフラッシュROMなどを有する。具体的には、例えば、フラッシュROMやROMが各種プログラムを記憶し、RAMがCPU301のワークエリアとして使用される。メモリ302に記憶されるプログラムは、CPU301にロードされることにより、コーディングされている処理をCPU301に実行させる。

[0039] ネットワークI/F303は、通信回線を通じてネットワーク210に接続され、ネットワーク210を介して他のコンピュータに接続される。そして、ネットワークI/F303は、ネットワーク210と内部のインターフェースを司り、他のコンピュータからのデータの入出力を制御する。ネットワークI/F303は、例えば、モデムやLANアダプタなどである。

[0040] 記録媒体I/F304は、CPU301の制御に従って記録媒体305に対するデータのリード/ライトを制御する。記録媒体I/F304は、例えば、ディスクドライブ、SSD (Solid State Drive)、USB (Universal Serial Bus) ポートなどである。記録媒体305は、記録媒体I/F304の制御で書き込まれたデータを記憶する不揮発メモリである。記録媒体305は、例えば、ディスク、半導体メモリ、USBメモリなどである。記録媒体305は、情報処理装置100から着脱可能であってもよい。

[0041] ディスプレイ306は、カーソル、アイコン、ツールボックス、文書、画像、または、機能情報などのデータを表示する。ディスプレイ306は、例えば、CRT (Cathode Ray Tube)、液晶ディスプレイ、または、有機EL (Electroluminescence) ディスプレイなどである。入力装置307は、文字、数字、または、各種指示などの入力のためのキーを有し、データの入力を行う。入力装置307は、例えば、キーボード、または、マウスなどである。入力装置307は、例えば、タッチパネル式の入力パッド、または、テンキーなどであってもよい。

[0042] 情報処理装置100は、上述した構成部の他、例えば、カメラなどを有し

てもよい。また、情報処理装置100は、上述した構成部の他、例えば、プリンタ、スキャナ、マイク、または、スピーカーなどを有してもよい。また、情報処理装置100は、記録媒体1/F304および記録媒体305を複数有していてもよい。また、情報処理装置100は、ディスプレイ306、または、入力装置307などを有していなくてもよい。また、情報処理装置100は、記録媒体1/F304および記録媒体305を有していなくてもよい。

[0043] (画像撮像装置201のハードウェア構成例)

次に、図4を用いて、画像撮像装置201のハードウェア構成例について説明する。

[0044] 図4は、画像撮像装置201のハードウェア構成例を示すブロック図である。図4において、画像撮像装置201は、CPU401と、メモリ402と、ネットワークI/F403と、記録媒体1/F404と、記録媒体405と、カメラ406とを有する。また、各構成部は、バス400によってそれぞれ接続される。

[0045] ここで、CPU401は、画像撮像装置201の全体の制御を司る。メモリ402は、例えば、ROM、RAMおよびフラッシュROMなどを有する。具体的には、例えば、フラッシュROMやROMが各種プログラムを記憶し、RAMがCPU401のワークエリアとして使用される。メモリ402に記憶されるプログラムは、CPU401にロードされることにより、コーディングされている処理をCPU401に実行させる。

[0046] ネットワークI/F403は、通信回線を通じてネットワーク210に接続され、ネットワーク210を介して他のコンピュータに接続される。そして、ネットワークI/F403は、ネットワーク210と内部のインターフェースを司り、他のコンピュータからのデータの入出力を制御する。ネットワークI/F403は、例えば、モデムやLANアダプタなどである。

[0047] 記録媒体1/F404は、CPU401の制御に従って記録媒体405に対するデータのリード/ライトを制御する。記録媒体1/F404は、例え

ば、ディスクドライブ、SSD、USBポートなどである。記録媒体405は、記録媒体1/F404の制御で書き込まれたデータを記憶する不揮発メモリである。記録媒体405は、例えば、ディスク、半導体メモリ、USBメモリなどである。記録媒体405は、画像撮像装置201から着脱可能であってもよい。カメラ406は、複数の撮像素子を有し、複数の撮像素子によって対象物を撮像した画像を生成する。カメラ406は、例えば、競技用カメラである。カメラ406は、例えば、監視カメラである。

[0048] 画像撮像装置201は、上述した構成部の他、例えば、キーボード、マウス、ディスプレイ、プリンタ、スキャナ、マイク、スピーカーなどを有してもよい。また、画像撮像装置201は、記録媒体1/F404や記録媒体405を複数有していてもよい。また、画像撮像装置201は、記録媒体1/F404や記録媒体405を有していなくてもよい。

[0049] (クライアント装置202のハードウェア構成例)

クライアント装置202のハードウェア構成例は、具体的には、図3に示した情報処理装置100のハードウェア構成例と同様であるため、説明を省略する。

[0050] (情報処理装置100の機能的構成例)

次に、図5を用いて、情報処理装置100の機能的構成例について説明する。

[0051] 図5は、情報処理装置100の機能的構成例を示すブロック図である。情報処理装置100は、記憶部500と、取得部501と、解析部502と、学習部503と、特定部504と、決定部505と、生成部506と、修正部507と、出力部508とを含む。

[0052] 記憶部500は、例えば、図3に示したメモリ302や記録媒体305などの記憶領域によって実現される。以下では、記憶部500が、情報処理装置100に含まれる場合について説明するが、これに限らない。例えば、記憶部500が、情報処理装置100とは異なる装置に含まれ、記憶部500の記憶内容が情報処理装置100から参照可能である場合があってもよい。

- [0053] 取得部501～出力部508は、制御部の一例として機能する。取得部501～出力部508は、具体的には、例えば、図3に示したメモリ302や記録媒体305などの記憶領域に記憶されたプログラムをCPU301に実行させることにより、または、ネットワークI/F303により、その機能を実現する。各機能部の処理結果は、例えば、図3に示したメモリ302や記録媒体305などの記憶領域に記憶される。
- [0054] 記憶部500は、各機能部の処理において参照され、または更新される各種情報を記憶する。記憶部500は、例えば、連続する複数の時点のそれぞれの時点において特定の者を異なるアングルから撮像した複数の画像を記憶する。アングルは、撮像位置を示す。画像は、例えば、取得部501によって取得される。
- [0055] 記憶部500は、例えば、骨格情報の時系列データを記憶する。時系列データは、連続する複数の時点のそれぞれの時点における骨格情報を含む。骨格情報は、特定の者の複数の部位のそれぞれの部位の位置を含む。部位は、例えば、関節である。部位は、例えば、首、頭、右肩および左肩、右肘および左肘、右手および左手、右膝および左膝、右足および左足などである。位置は、例えば、3次元座標である。時系列データは、例えば、取得部501によって取得される。時系列データは、例えば、解析部502によって生成されてもよい。
- [0056] 取得部501は、各機能部の処理に用いられる各種情報を取得する。取得部501は、取得した各種情報を、記憶部500に記憶し、または、各機能部へ出力する。また、取得部501は、記憶部500に記憶しておいた各種情報を、各機能部へ出力してもよい。取得部501は、例えば、利用者の操作入力に基づき、各種情報を取得する。取得部501は、例えば、情報処理装置100とは異なる装置から、各種情報を受信してもよい。
- [0057] 取得部501は、例えば、対象者の骨格情報の時系列データを取得する。対象者の骨格情報は、例えば、対象者の複数の部位のそれぞれの部位の位置を含む。取得部501は、具体的には、利用者の操作入力に基づき、対象者

の骨格情報の時系列データの入力を受け付けることにより、対象者の骨格情報の時系列データを取得する。取得部501は、具体的には、対象者の骨格情報の時系列データを、他のコンピュータから受信することにより取得してもよい。

[0058] 取得部501は、例えば、過去の被験者の骨格情報の時系列データを取得してもよい。被験者は、例えば、対象者と同一であってもよい。被験者の骨格情報は、例えば、被験者の複数の部位のそれぞれの部位の位置を含む。取得部501は、具体的には、利用者の操作入力に基づき、被験者の骨格情報の時系列データの入力を受け付けることにより、被験者の骨格情報の時系列データを取得する。取得部501は、具体的には、被験者の骨格情報の時系列データを、他のコンピュータから受信することにより取得してもよい。

[0059] 取得部501は、例えば、連続する複数の時点のそれぞれの時点において対象者を異なる角度から撮像した複数の画像を取得する。取得部501は、対象者の骨格情報の時系列データを取得せずに解析部502で生成する場合、複数の画像を取得する。これにより、取得部501は、対象者の骨格情報の時系列データを解析部502で生成可能にすることができる。

[0060] 取得部501は、例えば、連続する複数の時点のそれぞれの時点において被験者を異なる角度から撮像した複数の画像を取得してもよい。取得部501は、被験者の骨格情報の時系列データを取得せずに解析部502で生成する場合、複数の画像を取得する。これにより、取得部501は、被験者の骨格情報の時系列データを解析部502で生成可能にすることができる。

[0061] 取得部501は、いずれかの機能部の処理を開始する開始トリガーを受け付けてもよい。開始トリガーは、例えば、利用者による所定の操作入力があったことである。開始トリガーは、例えば、他のコンピュータから、所定の情報を受信したことであってもよい。開始トリガーは、例えば、いずれかの機能部が所定の情報を出力したことであってもよい。

[0062] 取得部501は、例えば、複数の画像を取得したことを、解析部502の処理を開始する開始トリガーとして受け付けてもよい。取得部501は、例

例えば、被験者の骨格情報の時系列データを取得したことを、学習部503の処理を開始する開始トリガーとして受け付けてもよい。取得部501は、例えば、対象者の骨格情報の時系列データを取得したことを、特定部504と決定部505と生成部506と修正部507との処理を開始する開始トリガーとして受け付けてもよい。

[0063] 解析部502は、所定の者の骨格情報の時系列データを生成する。解析部502は、例えば、対象者の骨格情報の時系列データを生成する。解析部502は、具体的には、複数の時点のそれぞれの時点において対象者を異なるアングルから撮像した複数の画像に基づいて、それぞれの時点における対象者のそれぞれの部位の位置を推定し、推定した位置を含む対象者の骨格情報を生成する。解析部502は、具体的には、生成した対象者の骨格情報に基づいて、対象者の骨格情報の時系列データを生成する。これにより、解析部502は、それぞれの時点における対象者のそれぞれの部位の位置を仮に特定することができ、修正対象を得ることができる。

[0064] 解析部502は、例えば、被験者の骨格情報の時系列データを生成してもよい。解析部502は、具体的には、複数の時点のそれぞれの時点において被験者を異なるアングルから撮像した複数の画像に基づいて、それぞれの時点における被験者の骨格情報を生成し、被験者の骨格情報の時系列データを生成する。解析部502は、生成した被験者の骨格情報の時系列データにノイズを加えてもよい。解析部502は、被験者の骨格情報を、学習モデルを生成するための教師情報に設定する。これにより、解析部502は、学習モデルを生成するための教師情報を得ることができる。

[0065] 学習部503は、被験者の複数の部位のそれぞれの部位の位置を含む教師情報に基づいて、学習モデルを学習する。学習モデルは、所定の者の骨格情報の時系列データにおける骨格情報に関する特徴量に応じて、所定の者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定可能にする機能を有する。学習モデルは、例えば、所定の者のそれぞれの部位が、位置に関して異常な状態であるか否かを判定可能にする機能を有する。

- [0066] 学習モデルは、具体的には、所定の者のそれぞれの部位が、位置に関して異常な状態である確率の大きさを示す指標値を算出する機能を有する。学習モデルは、より具体的には、骨格情報に関する特徴量が入力されたことに応じて、所定の者のそれぞれの部位について、当該部位が位置に関して異常な状態である確率の大きさを示す指標値を出力する。学習モデルは、具体的には、ニューラルネットワークである。これにより、学習部503は、対象者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定可能にすることができる。
- [0067] 特定部504は、取得した対象者の骨格情報の時系列データにおける骨格情報に関する特徴量に基づいて、取得した対象者の骨格情報の時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定する。特定部504は、例えば、学習した学習モデルを利用して、取得した対象者の骨格情報の時系列データにおける骨格情報に関する特徴量に基づいて、第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定する。
- [0068] 特定部504は、具体的には、学習モデルに、対象者の骨格情報の時系列データにおける骨格情報に関する特徴量を入力することにより、第1時点の骨格情報について、対象者のそれぞれの部位の異常な状態である確率の大きさを示す指標値を算出する。特定部504は、具体的には、算出した指標値に基づいて、第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定する。特定部504は、より具体的には、対象者の複数の部位のうち、算出した指標値が閾値以上であるいずれかの部位を、位置に関して異常な状態である部位として特定する。これにより、特定部504は、対象者の複数の部位のそれぞれの部位の位置を修正する指針を得ることができる。特定部504は、対象者のいずれの部位の位置を修正することが好ましいかを判別可能にすることができる。
- [0069] 特定部504は、例えば、所定のルールを参照して、取得した対象者の骨格情報の時系列データにおける骨格情報に関する特徴量に基づいて、第1時

点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定してもよい。所定のルールは、例えば、対象者の骨格情報の時系列データにおける骨格情報に関する特徴量に応じて、対象者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定可能にするルールを含む。所定のルールは、具体的には、所定の者のそれぞれの部位が、位置に関して異常な状態である確率の大きさを示す指標値を算出可能にするルールを含んでいてもよい。

[0070] 特定部504は、具体的には、所定のルールを参照して、対象者の骨格情報の時系列データにおける骨格情報に関する特徴量に基づいて、第1時点の骨格情報について、対象者のそれぞれの部位の異常な状態である確率の大きさを示す指標値を算出する。特定部504は、具体的には、算出した指標値に基づいて、第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定する。特定部504は、より具体的には、対象者の複数の部位のうち、算出した指標値が閾値以上であるいずれかの部位を、位置に関して異常な状態である部位として特定する。これにより、特定部504は、対象者の複数の部位のそれぞれの部位の位置を修正する指針を得ることができる。特定部504は、対象者のいずれの部位の位置を修正することが好ましいかを判別可能にすることができる。

[0071] 決定部505は、取得した対象者の骨格情報の時系列データにおける骨格情報に関する特徴量に基づいて、第1時点の骨格情報において、特定した部位の位置を制約する確率分布の分布モデルを決定する。分布モデルは、例えば、いずれかの部位の異常な状態である確率の大きさを示す指標値に応じて、いずれかの部位の位置を制約するためのモデルである。これにより、決定部505は、特定部504で特定した部位の位置を修正する指針を得ることができる。

[0072] 生成部506は、時点ごとのそれぞれの部位の位置を示すノードと、時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジとを含み、特定した部位の位置を示すノードに決定した分布モデルを

対応付けたグラフを生成する。これにより、生成部506は、対象者の骨格情報の時系列データにおける第1時点の骨格情報を修正可能にすることができる。

[0073] 生成部506は、ノードと、第1エッジと、異なる時点のいずれかの部位の位置を示すノード同士を接続する第2エッジとを含み、特定した部位の位置を示すノードに決定した分布モデルを対応付けたグラフを生成してもよい。これにより、生成部506は、対象者の骨格情報の時系列データにおける第1時点の骨格情報を修正可能にすることができる。

[0074] 修正部507は、生成したグラフに基づいて、対象者の骨格情報の時系列データにおける第1時点の骨格情報を修正する。修正部507は、例えば、生成したグラフを最適化することにより、対象者の骨格情報の時系列データにおける第1時点の骨格情報を修正する。これにより、修正部507は、対象者のそれぞれの部位の異常な状態である確率の大きさを考慮して、対象者のそれぞれの部位の位置を精度よく特定可能にすることができる。

[0075] 出力部508は、少なくともいずれかの機能部の処理結果を出力する。出力形式は、例えば、ディスプレイへの表示、プリンタへの印刷出力、ネットワークI/F303による外部装置への送信、または、メモリ302や記録媒体305などの記憶領域への記憶である。これにより、出力部508は、少なくともいずれかの機能部の処理結果を利用者に通知可能にし、情報処理装置100の利便性の向上を図ることができる。

[0076] 出力部508は、例えば、修正部507で修正した第1時点の骨格情報を出力する。出力部508は、具体的には、修正部507で修正した第1時点の骨格情報を、クライアント装置202に送信する。出力部508は、具体的には、修正部507で修正した第1時点の骨格情報を、ディスプレイに表示する。これにより、出力部508は、対象者のそれぞれの部位の位置を利用可能にすることができる。

[0077] (情報処理装置100の動作例)

次に、図6～図12を用いて、情報処理装置100の動作例について説明

する。まず、例えば、図6を用いて、情報処理装置100の動作の流れについて説明する。

[0078] 図6は、情報処理装置100の動作の流れを示す説明図である。図6において、情報処理装置100は、異なる時点における対象者を異なるアングルで撮像した複数の多視点画像600を取得する。情報処理装置100は、複数の多視点画像600のそれぞれの多視点画像600に対して、人検出処理を実施することにより、それぞれの多視点画像600のうち対象者が映った領域を検出する。

[0079] 情報処理装置100は、時点ごとに、それぞれの多視点画像600に対して、2D(Dimension)pose推定処理を実施する。情報処理装置100は、時点ごとに、それぞれの多視点画像600に対して、2Dpose推定処理を実施することにより、それぞれの多視点画像600における対象者のそれぞれの関節の存在確率の分布を示す2Dヒートマップ601を生成する。2Dヒートマップ601は、例えば、多視点画像600に対応する2D空間のそれぞれの点における、対象者のいずれかの関節の存在確率を示す関節尤度を含む。

[0080] 情報処理装置100は、時点ごとに、それぞれの多視点画像600における対象者のそれぞれの関節の存在確率の分布を示す2Dヒートマップ601に基づいて、当該多視点画像600における、対象者の当該関節の2D座標を特定する。2Dヒートマップ601における、対象者の関節の存在確率を示す関節尤度の分散は、特定した2D座標の精度を表す指標値として扱うことができる。

[0081] 情報処理装置100は、時点ごとに、それぞれの多視点画像600のアングルを示す配置情報を取得する。情報処理装置100は、時点ごとに、配置情報と、それぞれの多視点画像600における、対象者のそれぞれの関節の2D座標とに基づいて、3Dpose推定処理を実施することにより、3D空間における、対象者のそれぞれの関節の3D座標を特定する。情報処理装置100は、時点ごとに、特定した対象者のそれぞれの関節の3D座標を含

む3D骨格推論結果602を生成し、3D骨格推論結果602の時系列データを生成する。

[0082] 情報処理装置100は、3D骨格推論結果602の時系列データに対して、修正処理を実施することにより、3D骨格推論結果602を修正する。情報処理装置100は、修正した後の3D骨格推論結果603の時系列データを利用可能に出力する。情報処理装置100は、例えば、修正した後の3D骨格推論結果603の時系列データを、利用者が参照可能に出力する。

[0083] 利用者は、修正した後の3D骨格推論結果603の時系列データに基づいて、所定の解析処理を実施する。具体的には、対象者が、運動の大会の参加者である場合が考えられる。この場合、解析処理は、例えば、運動の大会の競技における参加者の採点である。利用者は、修正した後の3D骨格推論結果603の時系列データに基づいて、参加者を採点する解析処理を実施する。

[0084] 具体的には、対象者が、リハビリテーションを実施する医療機関受診者、または、歩行能力などの運動能力の診断を受ける医療機関受診者などである場合が考えられる。この場合、解析処理は、例えば、リハビリテーションの効果判定、あるいは、運動能力または健康状態の診断などである。利用者は、修正した後の3D骨格推論結果603の時系列データに基づいて、医療機関受診者のリハビリテーションの効果判定を実施し、または、医療機関受診者の運動能力または健康状態の診断を実施する。

[0085] 情報処理装置100は、修正した後の3D骨格推論結果603の時系列データに基づいて、上述した解析処理を実施してもよい。情報処理装置100は、解析処理を実施した結果を、利用者が参照可能に出力する。情報処理装置100は、修正した後の3D骨格推論結果603の時系列データを、上述した解析処理を実施する解析部502に出力してもよい。解析部502は、例えば、情報処理装置100以外の他のコンピュータが有する。これにより、情報処理装置100は、解析処理を精度よく実施可能にすることができる。

- [0086] 次に、図7～図9を用いて、修正処理の具体例について説明する。具体的には、まず、図7および図8を用いて、情報処理装置100が、対象者の複数の関節のうち、3D座標に関して異常な状態であると判断される異常関節を特定する具体例について説明する。
- [0087] 図7および図8は、異常関節を特定する具体例を示す説明図である。図7において、情報処理装置100は、複数の元データ700の時系列データを取得する。元データ700は、被験者の骨格情報を示す。元データ700は、被験者の複数の関節のそれぞれの関節の3D座標を示す。関節の3D座標は、例えば、図中の●である。
- [0088] 情報処理装置100は、元データ700にノイズを付与することにより、加工データ701を生成する。情報処理装置100は、例えば、元データ700が示す被験者の複数の関節の少なくともいずれかの関節の3D座標を、異常な状態であると判断される3D座標に変更することにより、加工データ701を生成する。異常な状態は、例えば、関節の3D座標を誤って推定した状態に対応する。異常な状態は、具体的には、*jitter*、*inversion*、*swap*、または、*miss*などである。これにより、情報処理装置100は、加工データ701の時系列データを取得することができる。
- [0089] 情報処理装置100は、加工データ701の時系列データを利用して、異常判定DNN (Deep Neural Network) 710を学習する。異常判定DNN 710は、例えば、3D骨格推論結果602の時系列データにおける3D骨格推論結果602の特徴量の入力に応じて、少なくともいずれかの3D骨格推論結果602における、対象者のそれぞれの関節の異常確率を出力する機能を有する。異常確率は、対象者の関節の3D座標が、位置的に異常な状態である確率の大きさを示す。異常判定DNN 710は、例えば、3D骨格推論結果602の時系列データにおける3D骨格推論結果602の特徴量の入力に応じて、当該時系列データ全体における、対象者のそれぞれの関節の異常確率を出力する機能を有していてもよい。次に、図8の説明に移行する。

[0090] 図8において、情報処理装置100は、3D骨格推論結果602の時系列データにおける3D骨格推論結果602の特徴量を、異常判定DNN710に入力する。情報処理装置100は、入力に応じて異常判定DNN710が出力した、3D骨格推論結果602の時系列データ全体における、対象者のそれぞれの関節の異常確率を取得する。情報処理装置100は、取得した対象者のそれぞれの関節の異常確率に基づいて、異常関節を特定する。情報処理装置100は、例えば、対象者の複数の関節のうち、取得した異常確率が閾値以上であるいずれかの関節を、異常関節として特定する。

[0091] ここでは、情報処理装置100が、異常判定DNN710を利用して、異常関節を特定する場合について説明したが、これに限らない。例えば、情報処理装置100が、ルールベースで異常関節を特定する場合があってもよい。具体的には、情報処理装置100は、3D骨格推論結果602において、それぞれの関節に関する特徴量と閾値との差分の大きさに応じて、当該関節の異常確率を算出するルールを記憶しておいてもよい。具体的には、情報処理装置100は、記憶したルールを参照して、それぞれの関節の異常確率を算出し、算出した異常確率が閾値以上であるいずれかの関節を、異常関節であると特定することが考えられる。次に、図9を用いて、情報処理装置100が、Factor Graph900を生成する具体例について説明する。

[0092] 図9は、Factor Graph900を生成する具体例を示す説明図である。図9において、情報処理装置100は、Factor Graph900を生成する。Factor Graph900は、例えば、時点ごとの対象者のそれぞれの関節の位置を示すノードを含む。

[0093] Factor Graph900は、具体的には、時点ごとの、対象者の頭部、上部頸椎、下部頸椎、胸椎、腰椎、左右股関節、左右膝関節、左右脚関節、左右足部、左右肩関節、左右肘関節、左右手首、左右手のそれぞれの位置を示すノードを含む。

[0094] 情報処理装置100は、Factor Graph900のうち、異なる

時点の所定の関節の位置を示すノード同士を、第2エッジで接続した上で、Factor Graph 901を生成してもよい。第2エッジは、例えば、関節の種類ごとに定められた時系列制約を示すPairwise Termが対応付けられていてもよい。時系列制約は、例えば、等位置運動、等速直線運動、または、等加速運動などに対応する確率分布によって定義される。

- [0095] 等速直線運動に対応する時系列制約を示すPairwise Termは、例えば、 $g_t(x^{j,t-1}, x^{j,t}) \sim N(\|x^{j,t-1}, x^{j,t}\| | v_j^{\wedge} \Delta t, \Sigma_{v_j}^{\wedge})$ である。以下の説明では、便宜上、上部に \wedge を付けた文字を「文字 \wedge 」と表記する場合がある。 $x^{j,t-1}$ は、時刻 $t-1$ の関節の推定位置である。 $x^{j,t}$ は、時刻 t の関節の推定位置である。 v_j^{\wedge} は、関節の平均速度である。 Δt は、単位時間幅である。 $\Sigma_{v_j}^{\wedge}$ は、関節の速度分散である。等位置運動に対応する時系列制約を示すPairwise Termの分布モデルが対応付けられる。Pairwise Termは、例えば、 $g_t(x^{j,t-1}, x^{j,t}) \sim N(\|x^{j,t-1}, x^{j,t}\| | 0, \Sigma_{x_j}^{\wedge})$ である。 $\Sigma_{x_j}^{\wedge}$ は、関節の位置分散である。
- [0096] 情報処理装置100は、Factor Graph 900のうち、異なる時点の同一の関節の位置を示す、エッジが1つだけ接続された葉ノード同士を、第3エッジで接続した上で、Factor Graph 901を生成してもよい。葉ノードは、例えば、第1エッジが1つだけ接続され、第2エッジが接続されないノードである。
- [0097] 情報処理装置100は、例えば、Factor Graph 900のうち、第1エッジが1つだけ接続され、第2エッジが接続されない葉ノードが、特定した異常関節の位置を示すノードであるか否かを判定する。情報処理装置100は、葉ノードが、特定した異常関節の位置を示すノードであれば、異なる時点の葉ノード同士を、第3エッジ910で接続する。
- [0098] 第3エッジ910は、所定の時系列制約を示すPairwise Termが対応付けられていてもよい。これにより、情報処理装置100は、異常関節の位置を精度よく修正可能にすることができる。次に、図10を用いて

、情報処理装置100が、生成したFactor Graph 901を利用して、3D骨格推論結果602を修正する具体例について説明する。

[0099] 図10は、3D骨格推論結果602を修正する具体例を示す説明図である。図10において、情報処理装置100は、生成したFactor Graph 901を利用して、3D骨格推論結果602を修正する。図10の例では、Factor Graph 901は、時刻 $t-1$ に対応するノード群1010と、時刻 t に対応するノード群1020などを含む。ノード群1010は、ノード1011~1013などを含む。ノード群1020は、ノード1021~1023などを含む。

[0100] 例えば、ノード1011とノード1012とは、第1エッジ1031で接続される。例えば、ノード1012とノード1013とは、第1エッジ1032で接続される。例えば、ノード1021とノード1022とは、第1エッジ1041で接続される。例えば、ノード1022とノード1023とは、第1エッジ1042で接続される。例えば、ノード1022とノード1023とを接続する第1エッジ1042は、骨の長さの制約を示すPairwise Termが対応付けられていてもよい。

[0101] 情報処理装置100は、例えば、ノード1012とノード1022とを、さらに、第2エッジ1051で接続してもよい。第2エッジ1051は、例えば、関節の種類ごとに定められた時系列制約を示すPairwise Termが対応付けられていてもよい。例えば、ノード1011とノード1021とは、第3エッジ1061で接続される。第3エッジ1061は、例えば、所定の時系列制約を示すPairwise Termが対応付けられていてもよい。

[0102] 情報処理装置100は、Factor Graph 901の少なくともいずれかの関節の位置を示すノードに、当該関節の異常確率に応じて当該関節の位置を制約するよう作用する、異常関節の制約を示すUnary Termを対応付ける。図10の例では、情報処理装置100は、例えば、ノード群1020のうち、関節1の位置を示すノード1021に、関節1の異常確

率を含めたUnary Termを対応付ける。Unary Termは、例えば、 $f(x^j) \sim N(x^j | x^{j^*}, \Sigma_{3D}^{j^*}) \cdot p(x^j)$ である。 $p(x^j)$ は、異常確率である。

[0103] 情報処理装置100は、Factor Graph901におけるUnary Termと、Pairwise Termとに基づいて、それぞれの時点のそれぞれの関節の位置を修正する。情報処理装置100は、例えば、Factor Graph901を最適化することにより、それぞれの時点のそれぞれの関節の位置を修正する。

[0104] これにより、情報処理装置100は、3D骨格推論結果602を精度よく修正することができる。情報処理装置100は、それぞれの時点のそれぞれの関節の位置を精度よく特定することができる。情報処理装置100は、例えば、対象者が、体操などのような比較的高速または比較複雑な動きをした場合においても、確度が比較的高い、それぞれの時点の対象者のそれぞれの関節の位置を特定することができる。

[0105] ここで、下記参考文献3を参照して、Factor GraphにおけるUnary Termとして、異常確率を含まない $f(x^j) \sim N(x^j | x^{j^*}, \Sigma_{3D}^{j^*})$ を採用する比較例が考えられる。 x^{j^*} は、複数の2Dヒートマップの関節尤度を統合した3Dヒートマップの関節尤度の重み付け和である。 $\Sigma_{3D}^{j^*}$ は、複数の2Dヒートマップの関節尤度を統合した3Dヒートマップの関節尤度の分散である。

[0106] このため、比較例は、Unary Termにより、関節尤度が比較的低いいずれかの関節の3D座標を、尤度が比較的高い他の関節の3D座標を基準に修正するよう作用することになる。しかしながら、比較例では、対象者のそれぞれの関節の3D座標を精度よく特定することが難しく、人のそれぞれの関節の3D座標の時間変化を精度よく修正することが難しい場合がある。例えば、関節尤度が比較的低いいずれかの関節の3D座標を修正することが好ましいとは限らず、関節尤度が比較的高いいずれかの関節の3D座標を基準として採用することが好ましいとも限らない。このため、比較例では、

対象者のそれぞれの関節の3D座標を適切に修正することができない場合がある。

[0107] 参考文献3 : Bultmann, Simon, and Sven Behnke. "Real-time multi-view 3D human pose estimation using semantic feedback to smart edge sensors." arXiv preprint arXiv:2106.14729 (2021).

[0108] これに対し、情報処理装置100は、Factor Graph901におけるUnary Termとして、異常確率を含む $f(x^j) \sim N(x^j | x^j \hat{,} \Sigma_{3D}^{j \hat{}}) \cdot p(x^j)$ を採用することができる。このため、情報処理装置100は、修正することが好ましい関節の3D座標を修正し易く、基準として採用することが好ましい関節の3D座標を固定し易くすることができる。従って、情報処理装置100は、対象者のそれぞれの関節3D座標を適切に修正することができる。次に、図11および図12を用いて、動作例におけるデータ処理の流れの具体例について説明する。

[0109] 図11および図12は、動作例におけるデータ処理の流れの具体例を示す説明図である。図11に示すように、情報処理装置100は、時点ごとに、複数のカメラ画像1101を取得する。情報処理装置100は、2D骨格推論モデル1110を記憶する。情報処理装置100は、例えば、2D骨格推論モデル1110となるニューラルネットワークを定義する重みパラメータを記憶する。

[0110] 情報処理装置100は、時点ごとに、2D骨格推論モデル1110を参照して、複数のカメラ画像1101のそれぞれのカメラ画像1101に対して、2D骨格推論処理を実施することにより、2D骨格推論結果1102を生成する。2D骨格推論結果1102は、例えば、関節の位置を示す2D座標 $(x[\text{pixel}], y[\text{pixel}])$ と、関節の位置の確からしさを示す尤度とを含む。

- [0111] 情報処理装置100は、3D骨格推論モデル1120を記憶する。情報処理装置100は、例えば、3D骨格推論モデル1120となるニューラルネットワークを定義する重みパラメータを記憶する。
- [0112] 情報処理装置100は、時点ごとに、3D骨格推論モデル1120を参照して、複数の2D骨格推論結果1102に対して、3D骨格推論処理を実施することにより、3D骨格推論結果1103を生成する。3D骨格推論結果1103は、例えば、関節の位置を示す3D座標（ x [mm] , y [mm] , z [mm]）を含む。情報処理装置100は、時点ごとの3D骨格推論結果1103を纏めた時系列データ1104を生成する。次に、図12の説明に移行する。
- [0113] 図12に示すように、情報処理装置100は、異常検知モデル1210を記憶する。情報処理装置100は、例えば、異常検知モデル1210となるニューラルネットワークを定義する重みパラメータを記憶する。
- [0114] 情報処理装置100は、異常検知モデル1210を参照して、時系列データ1104に対して、異常検知処理を実施することにより、関節ごとの異常確率を算出し、算出した関節ごとの異常確率を含む骨格異常検知結果1201を生成する。
- [0115] 情報処理装置100は、骨長さモデル1220と、時系列運動モデル1230とを記憶する。骨長さモデル1220は、骨の長さの制約を示す *Pairwise Term* を定義するパラメータを含む。パラメータは、例えば、骨の長さの平均と分散とである。時系列運動モデル1230は、関節ごとの時系列制約を示す *Pairwise Term* を含む。
- [0116] 情報処理装置100は、骨格異常検知結果1201と、骨長さモデル1220と、時系列運動モデル1230とを参照して、*Factor Graph* 生成処理を実施することにより、*Factor Graph* を生成する。情報処理装置100は、生成した *Factor Graph* に対して、最適化処理を実施することにより、それぞれの関節の位置を修正する。情報処理装置100は、修正したそれぞれの関節の位置を含む修正後3D骨格推論結

果1202を生成する。

[0117] (全体処理手順)

次に、図13を用いて、情報処理装置100が実行する、全体処理手順の一例について説明する。全体処理は、例えば、図3に示したCPU301と、メモリ302や記録媒体305などの記憶領域と、ネットワークI/F303とによって実現される。

[0118] 図13は、全体処理手順の一例を示すフローチャートである。図13において、情報処理装置100は、対象者の3次元骨格推論結果の時系列データを取得する(ステップS1301)。

[0119] 次に、情報処理装置100は、取得した対象者の3次元骨格推論結果の時系列データに基づいて、対象者のそれぞれの部位の異常度を算出し、算出した異常度に基づいて、対象者の複数の部位のうち、異常部位を特定する(ステップS1302)。そして、情報処理装置100は、対象者のそれぞれの部位の異常度に基づいて、対象者のそれぞれの部位の尤度を算出する(ステップS1303)。

[0120] 次に、情報処理装置100は、対象者の特定の部位にUnary Term=尤度を設定し、対象者の特定の部位について時間軸に沿ってPairwise Termを設定したFactor Graphを生成する(ステップS1304)。そして、情報処理装置100は、Factor Graphを最適化することにより、対象者の3次元骨格推論結果の時系列データを修正する(ステップS1305)。

[0121] 次に、情報処理装置100は、修正後の対象者の3次元骨格推論結果の時系列データを出力する(ステップS1306)。そして、情報処理装置100は、全体処理を終了する。これにより、情報処理装置100は、対象者の3次元骨格推論結果を精度よく修正することができる。このため、情報処理装置100は、対象者の3次元骨格推論結果の有用性の向上を図ることができる。情報処理装置100は、例えば、対象者の3次元骨格推論結果に基づく解析処理の精度の向上を図ることができる。

[0122] 以上説明したように、情報処理装置100によれば、対象者の複数の部位のそれぞれの部位の位置を含む骨格情報の時系列データを取得することができる。情報処理装置100によれば、取得した時系列データにおける骨格情報に関する特徴量に基づいて、取得した時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定することができる。情報処理装置100によれば、第1時点の骨格情報において、特定したいずれかの部位の位置を、特定したいずれかの部位の異常な状態である確率の大きさに応じて制約する確率分布のモデルを決定することができる。情報処理装置100によれば、時点ごとのそれぞれの部位の位置を示すノードを含むグラフを生成することができる。情報処理装置100によれば、グラフにおいて、時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジを付与することができる。情報処理装置100によれば、グラフにおいて、特定したいずれかの部位の位置を示すノードに、決定したモデルを対応付けることができる。情報処理装置100によれば、生成したグラフに基づいて、時系列データにおける第1時点の骨格情報を修正することができる。これにより、情報処理装置100は、第1時点の骨格情報を精度よく修正することができる。

[0123] 情報処理装置100によれば、被験者の複数の部位のそれぞれの部位の位置を含む教師情報に基づいて、対象者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定可能にするモデルを学習することができる。情報処理装置100によれば、学習したモデルを利用して、取得した時系列データにおける骨格情報に関する特徴量に基づいて、取得した時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定することができる。これにより、情報処理装置100は、位置に関して異常な状態であるいずれかの部位を精度よく特定することができる。

[0124] 情報処理装置100によれば、時系列データにおける骨格情報に関する特徴量に応じて、対象者の複数の部位のうち、位置に関して異常な状態である

いずれかの部位を特定可能にするルールを記憶することができる。情報処理装置100によれば、ルールを参照して、取得した時系列データにおける骨格情報に関する特徴量に基づいて、取得した時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定することができる。これにより、情報処理装置100は、位置に関して異常な状態であるいずれかの部位を精度よく特定することができる。

[0125] 情報処理装置100によれば、ノードと、第1エッジと、異なる時点のいずれかの部位の位置を示すノード同士を接続する第2エッジとを含むグラフを生成することができる。これにより、情報処理装置100は、第1時点の骨格情報を精度よく修正し易くすることができる。

[0126] なお、本実施の形態で説明した情報処理方法は、予め用意されたプログラムをPCやワークステーションなどのコンピュータで実行することにより実現することができる。本実施の形態で説明した情報処理プログラムは、コンピュータで読み取り可能な記録媒体に記録され、コンピュータによって記録媒体から読み出されることによって実行される。記録媒体は、ハードディスク、フレキシブルディスク、CD (Compact Disc) -ROM、MO (Magneto Optical disc)、DVD (Digital Versatile Disc) などである。また、本実施の形態で説明した情報処理プログラムは、インターネットなどのネットワークを介して配布してもよい。

符号の説明

- [0127] 100 情報処理装置
101 骨格情報
110 グラフ
111, 1011~1013, 1021~1023 ノード
112, 1031, 1032, 1041, 1042 第1エッジ
200 情報処理システム
201 画像撮像装置

202 クライアント装置
210 ネットワーク
300, 400 バス
301, 401 CPU
302, 402 メモリ
303, 403 ネットワーク I/F
304, 404 記録媒体 I/F
305, 405 記録媒体
306 ディスプレイ
307 入力装置
406 カメラ
500 記憶部
501 取得部
502 解析部
503 学習部
504 特定部
505 決定部
506 生成部
507 修正部
508 出力部
600 多視点画像
601 2Dヒートマップ
602, 603, 1103, 1202 3D骨格推論結果
700 元データ
701 加工データ
710 異常判定DNN
900, 901 Factor Graph
910, 1061 第3エッジ

- 1 0 1 0, 1 0 2 0 ノード群
- 1 0 5 1 第2エッジ
- 1 1 0 1 カメラ画像
- 1 1 0 2 2 D骨格推論結果
- 1 1 0 4 時系列データ
- 1 1 1 0 2 D骨格推論モデル
- 1 1 2 0 3 D骨格推論モデル
- 1 2 0 1 骨格異常検知結果
- 1 2 1 0 異常検知モデル
- 1 2 2 0 骨長さモデル
- 1 2 3 0 時系列運動モデル

請求の範囲

- [請求項1] 対象者の複数の部位のそれぞれの部位の位置を含む骨格情報の時系列データを取得し、
- 取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、取得した前記時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定し、
- 取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、前記第1時点の骨格情報において、特定した前記いずれかの部位の位置を、特定した前記いずれかの部位の異常な状態である確率の大きさに応じて制約する確率分布のモデルを決定し、
- 時点ごとの前記それぞれの部位の位置を示すノードと、前記時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジとを含み、前記いずれかの部位の位置を示すノードに決定した前記モデルを対応付けたグラフを生成し、
- 生成した前記グラフに基づいて、前記時系列データにおける前記第1時点の骨格情報を修正する、
- 処理をコンピュータに実行させることを特徴とする情報処理プログラム。
- [請求項2] 被験者の複数の部位のそれぞれの部位の位置を含む教師情報に基づいて、前記時系列データにおける骨格情報に関する特徴量に応じて、前記対象者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定可能にするモデルを学習し、
- 前記特定する処理は、
- 学習した前記モデルを利用して、取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、取得した前記時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定する、ことを特徴とする請求項1に記載の情報処理プログラム。

[請求項3]

前記特定する処理は、

前記時系列データにおける骨格情報に関する特徴量に応じて、前記対象者の複数の部位のうち、位置に関して異常な状態であるいずれかの部位を特定可能にするルールを参照して、取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、取得した前記時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定する、ことを特徴とする請求項1に記載の情報処理プログラム。

[請求項4]

前記生成する処理は、

前記ノードと、前記第1エッジと、異なる時点の前記いずれかの部位の位置を示すノード同士を接続する第2エッジとを含むグラフを生成する、ことを特徴とする請求項1～3のいずれか一つに記載の情報処理プログラム。

[請求項5]

対象者の複数の部位のそれぞれの部位の位置を含む骨格情報の時系列データを取得し、

取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、取得した前記時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定し、

取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、前記第1時点の骨格情報において、特定した前記いずれかの部位の位置を、特定した前記いずれかの部位の異常な状態である確率の大きさに応じて制約する確率分布のモデルを決定し、

時点ごとの前記それぞれの部位の位置を示すノードと、前記時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジとを含み、前記いずれかの部位の位置を示すノードに決定した前記モデルを対応付けたグラフを生成し、

生成した前記グラフに基づいて、前記時系列データにおける前記第1時点の骨格情報を修正する、

処理をコンピュータが実行することを特徴とする情報処理方法。

[請求項6] 対象者の複数の部位のそれぞれの部位の位置を含む骨格情報の時系列データを取得し、

取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、取得した前記時系列データにおける第1時点の骨格情報について、位置に関して異常な状態であるいずれかの部位を特定し、

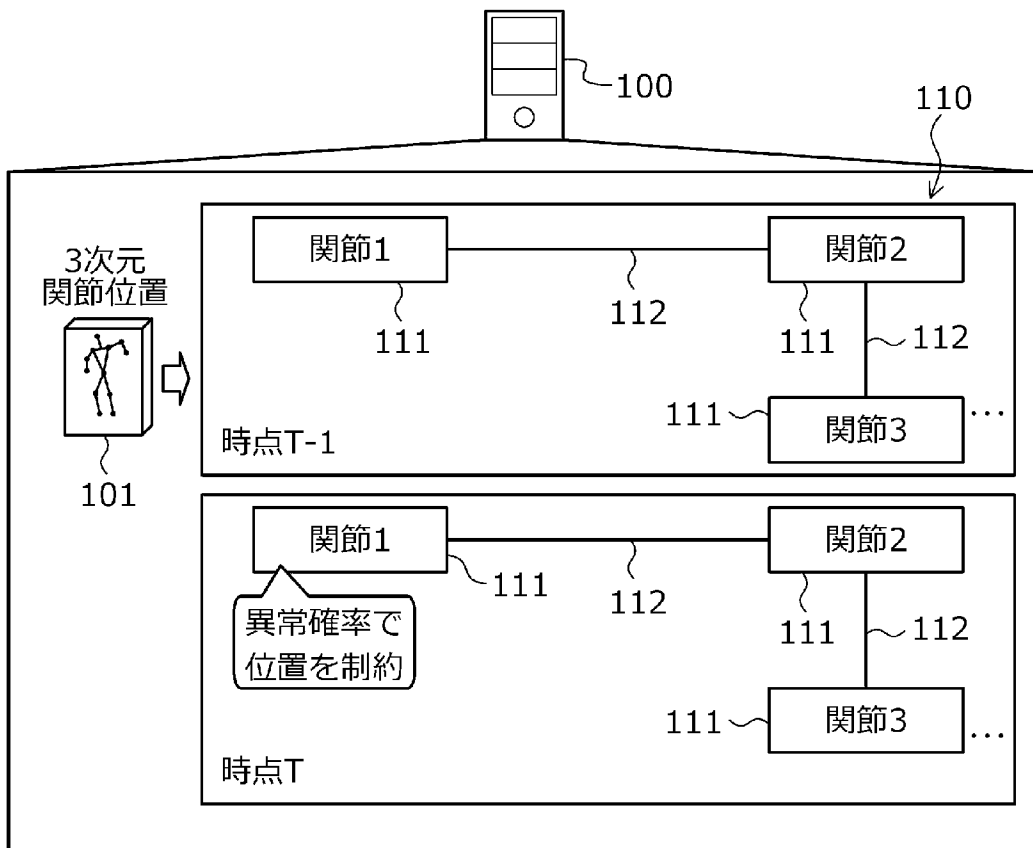
取得した前記時系列データにおける骨格情報に関する特徴量に基づいて、前記第1時点の骨格情報において、特定した前記いずれかの部位の位置を、特定した前記いずれかの部位の異常な状態である確率の大きさに応じて制約する確率分布のモデルを決定し、

時点ごとの前記それぞれの部位の位置を示すノードと、前記時点ごとの生体的に連結する異なる部位の位置を示すノード同士を接続する第1エッジとを含み、前記いずれかの部位の位置を示すノードに決定した前記モデルを対応付けたグラフを生成し、

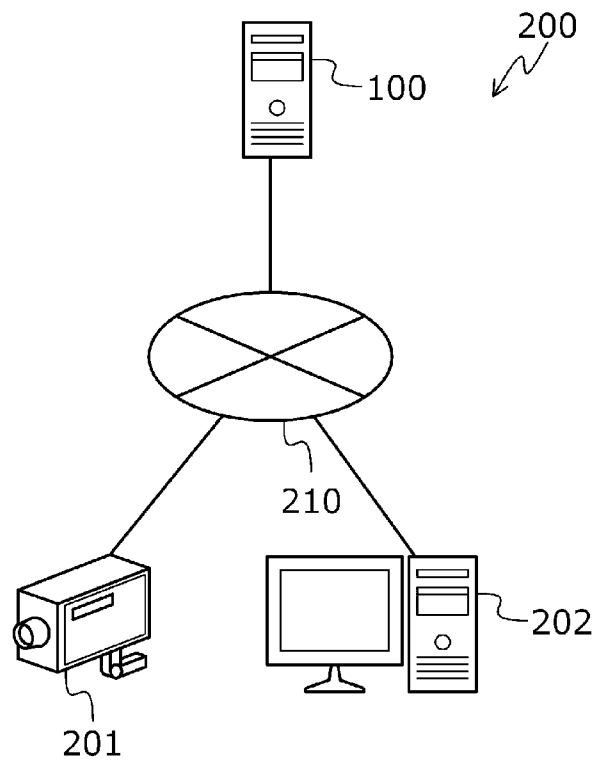
生成した前記グラフに基づいて、前記時系列データにおける前記第1時点の骨格情報を修正する、

制御部を有することを特徴とする情報処理装置。

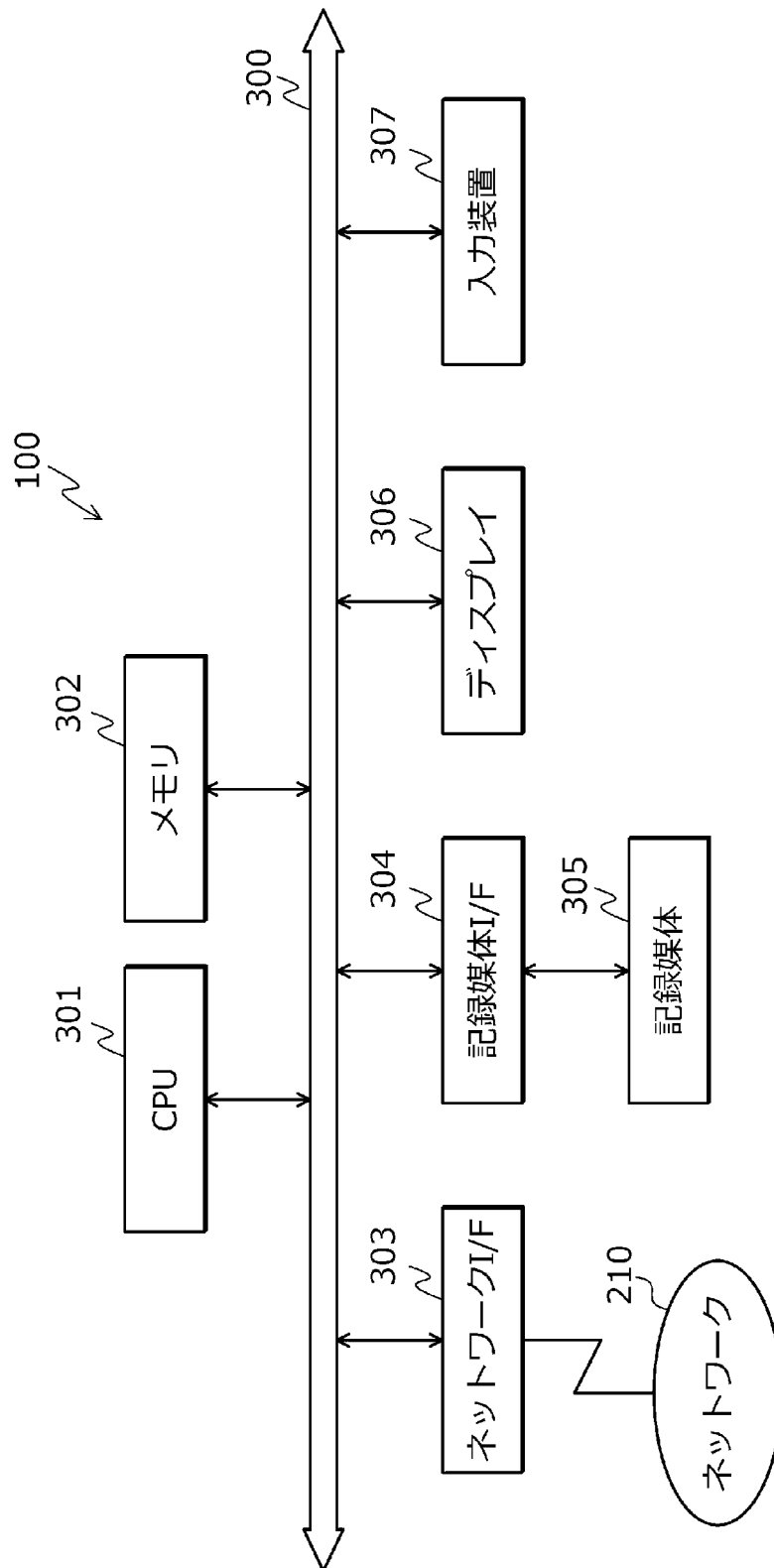
[図1]



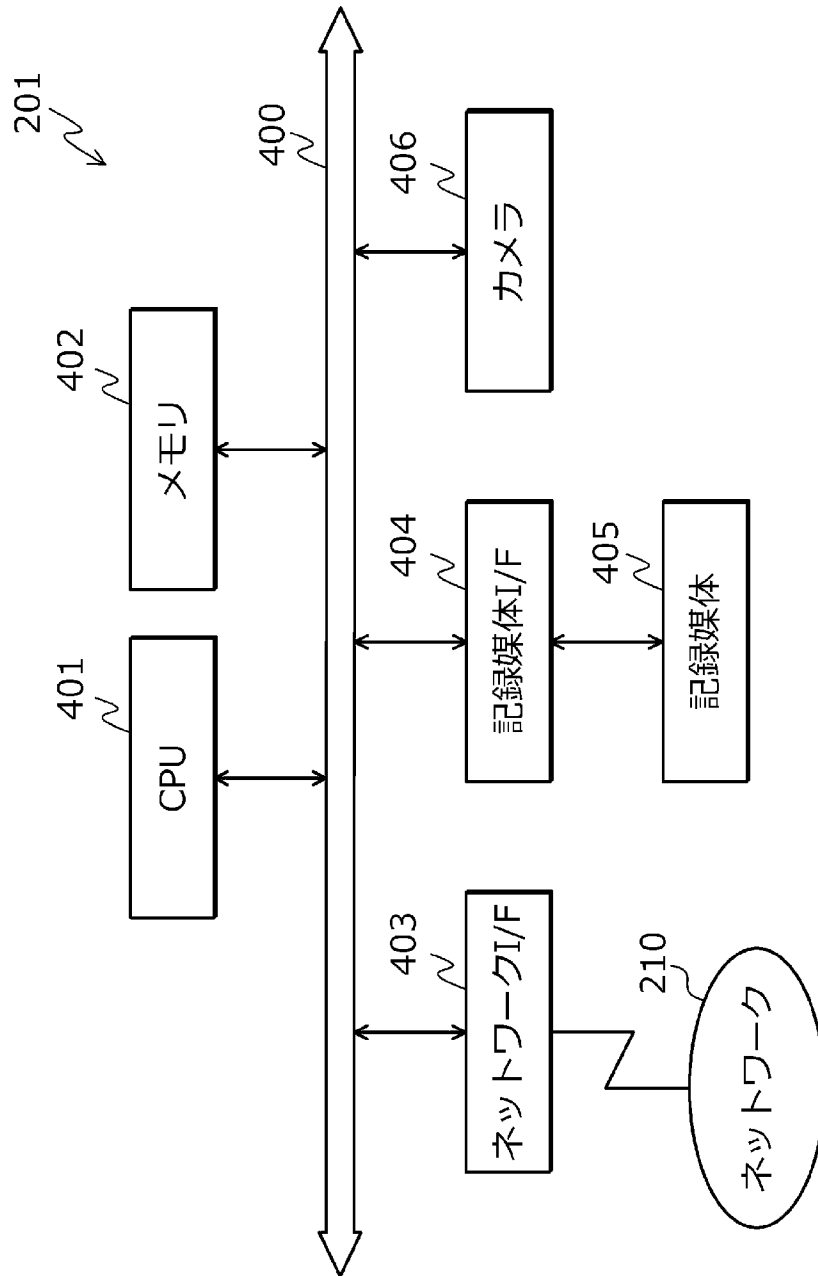
[図2]



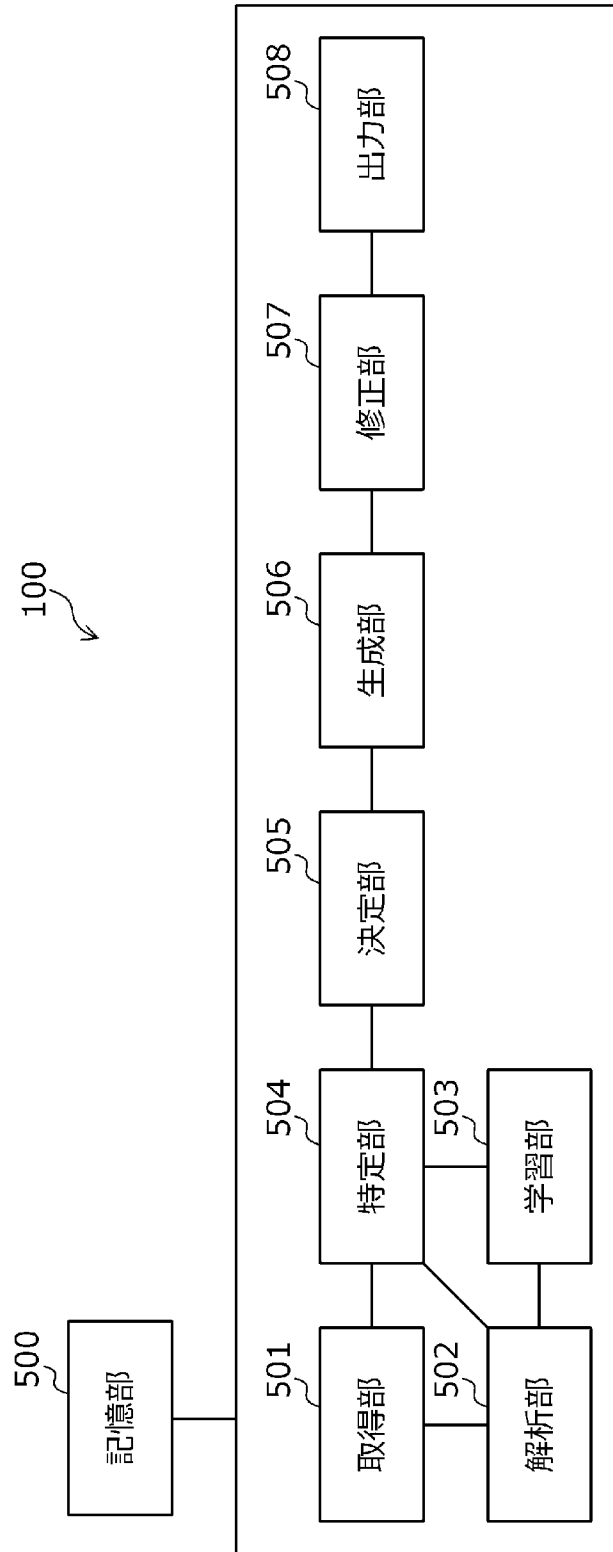
[図3]



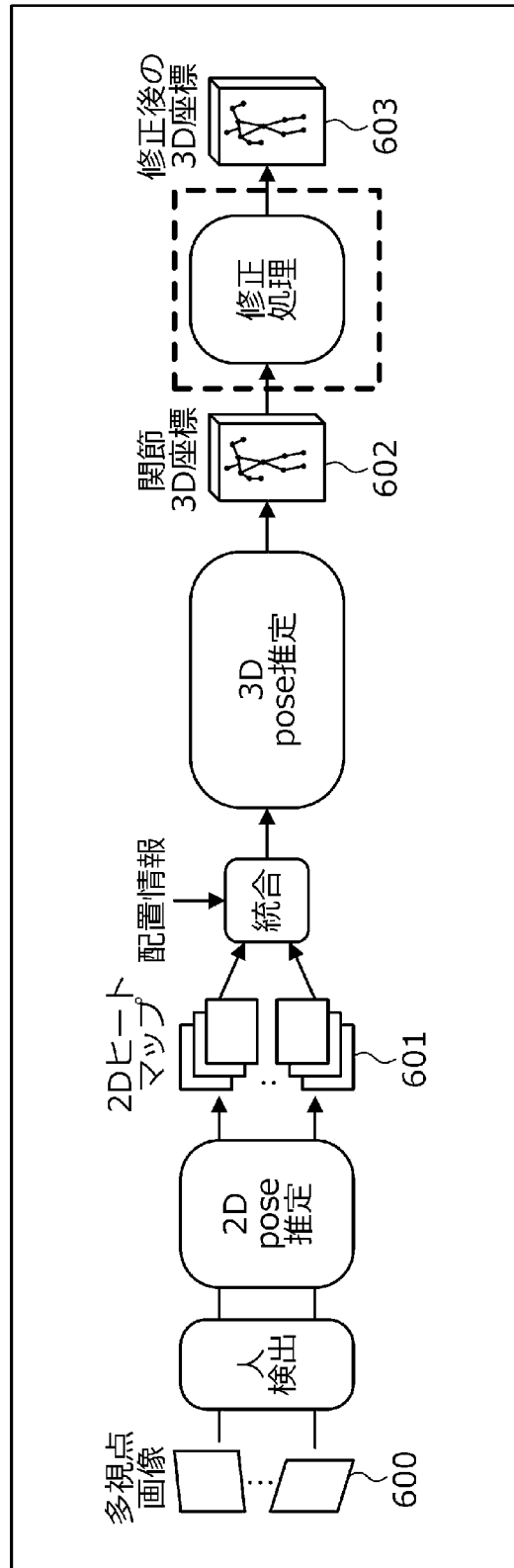
[図4]



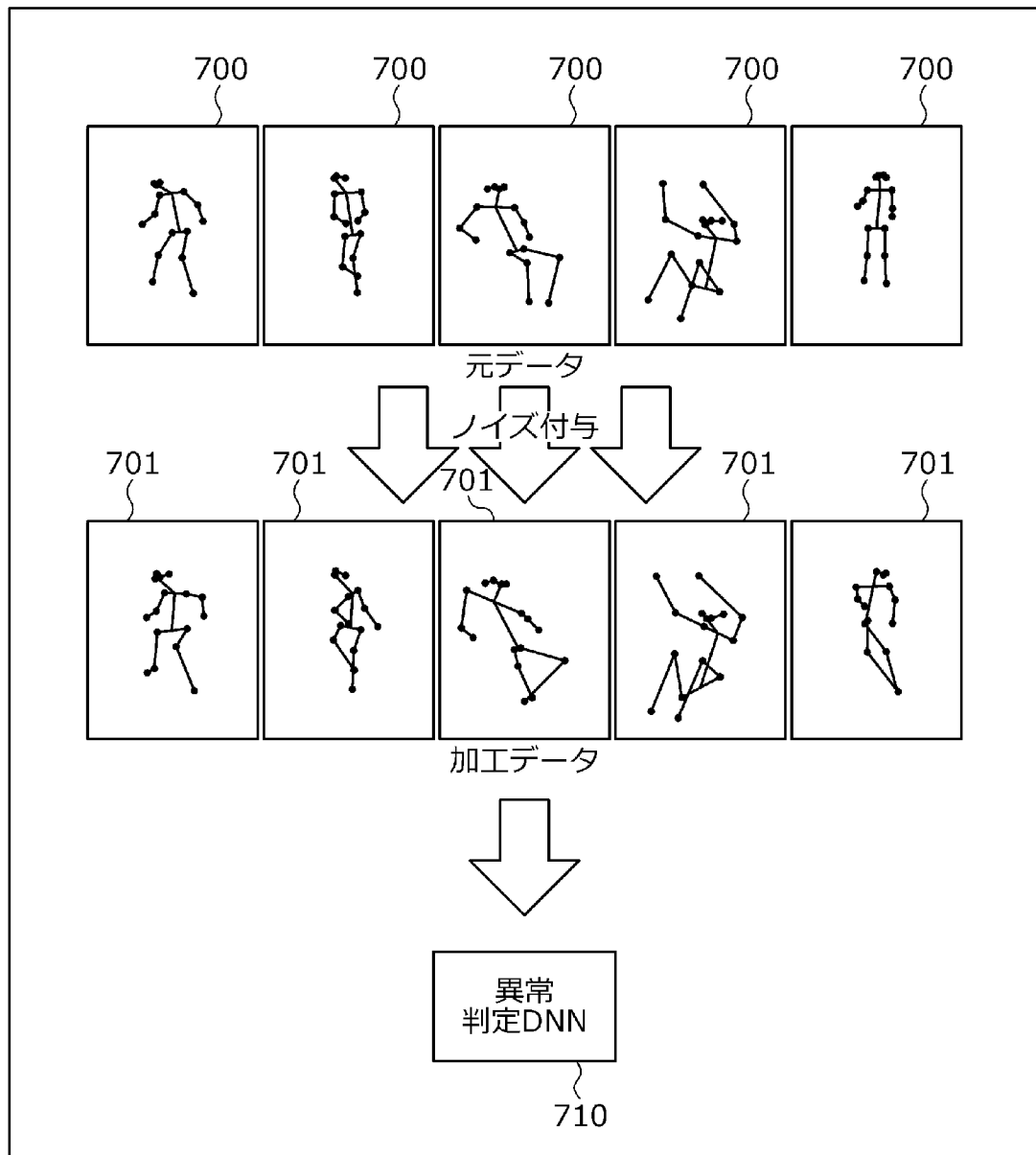
[図5]



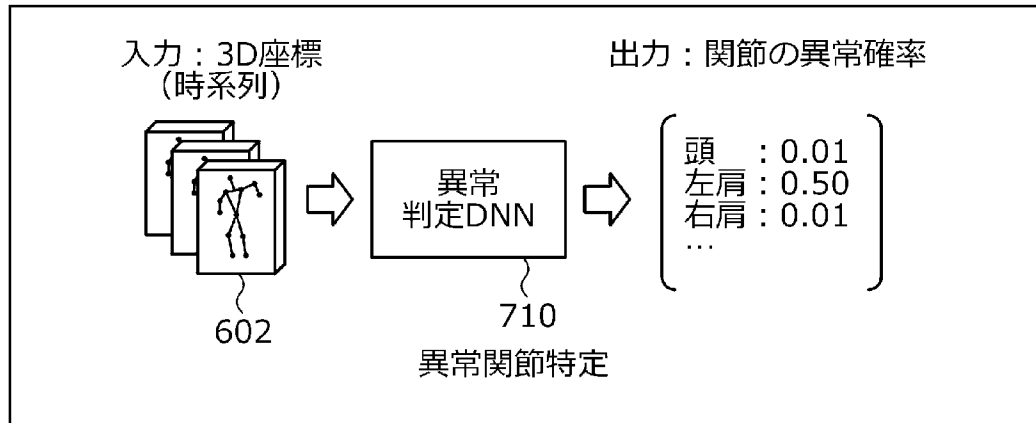
[図6]



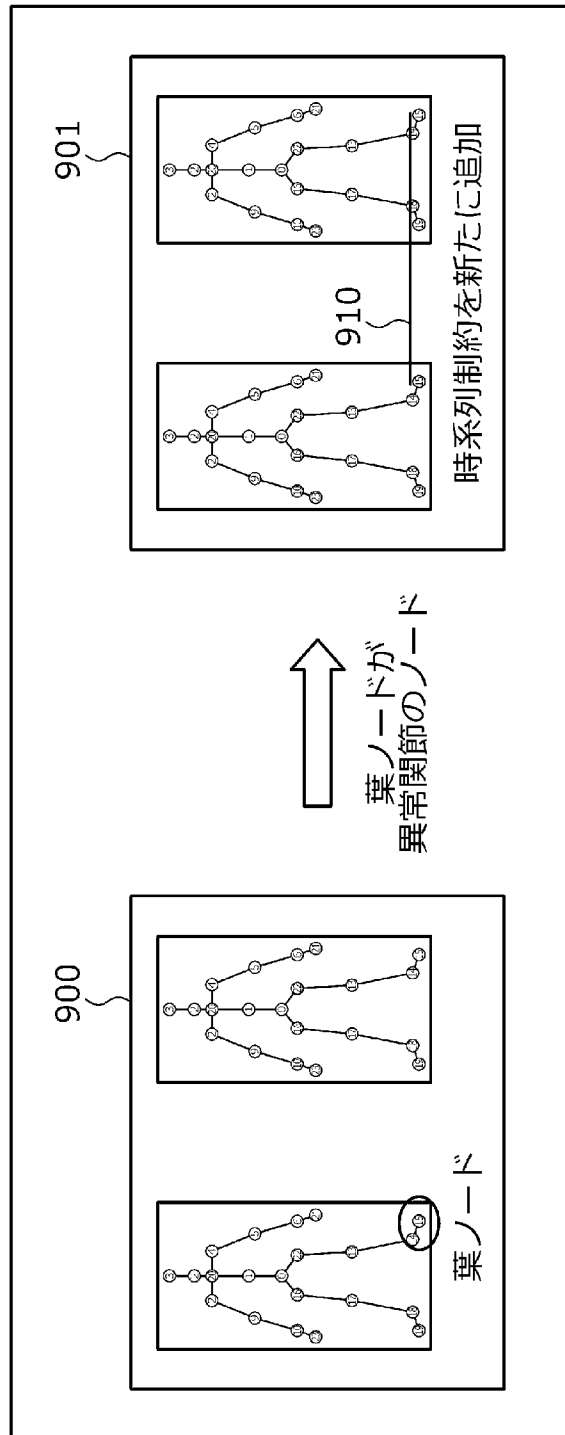
[図7]



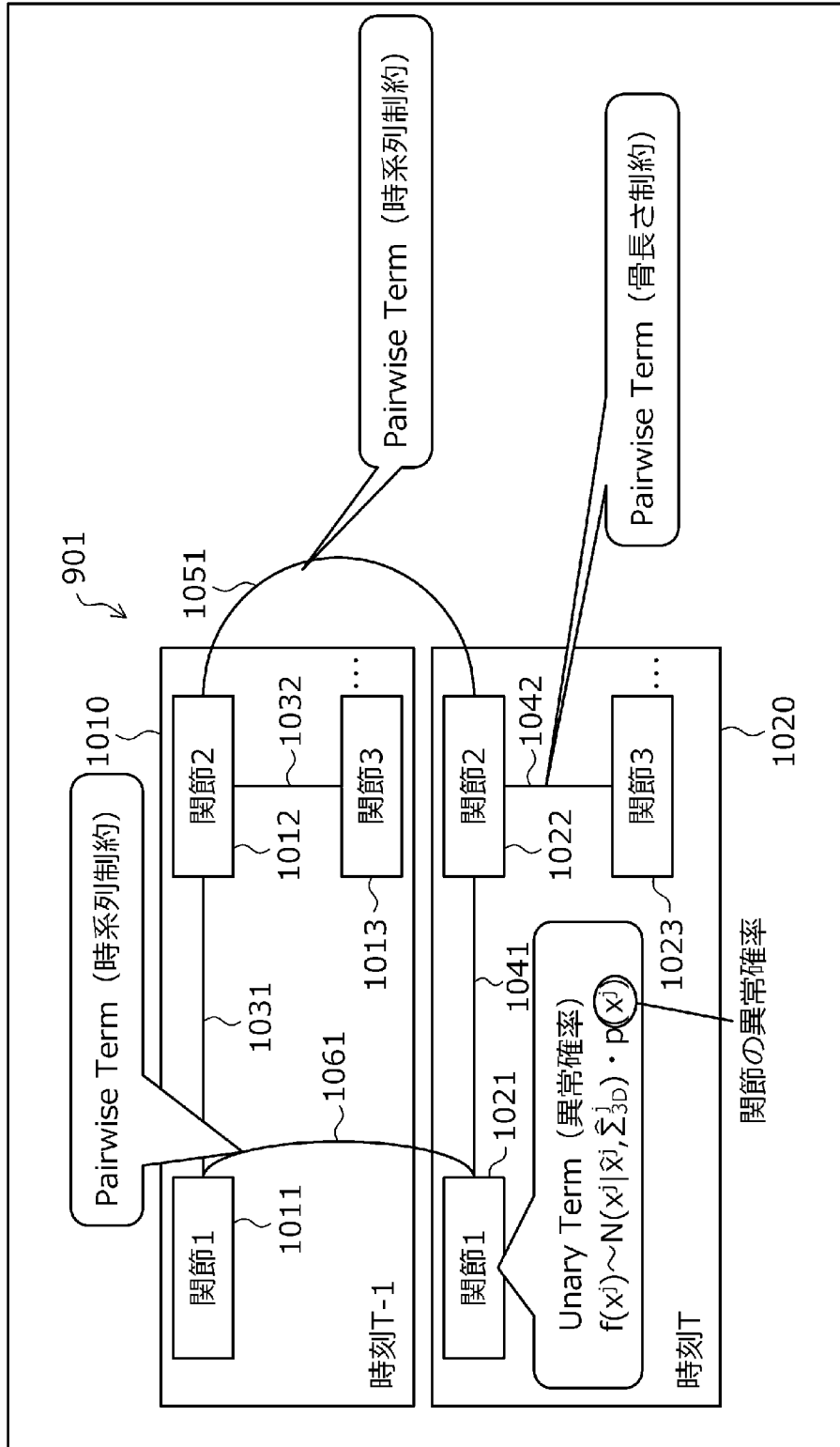
[図8]



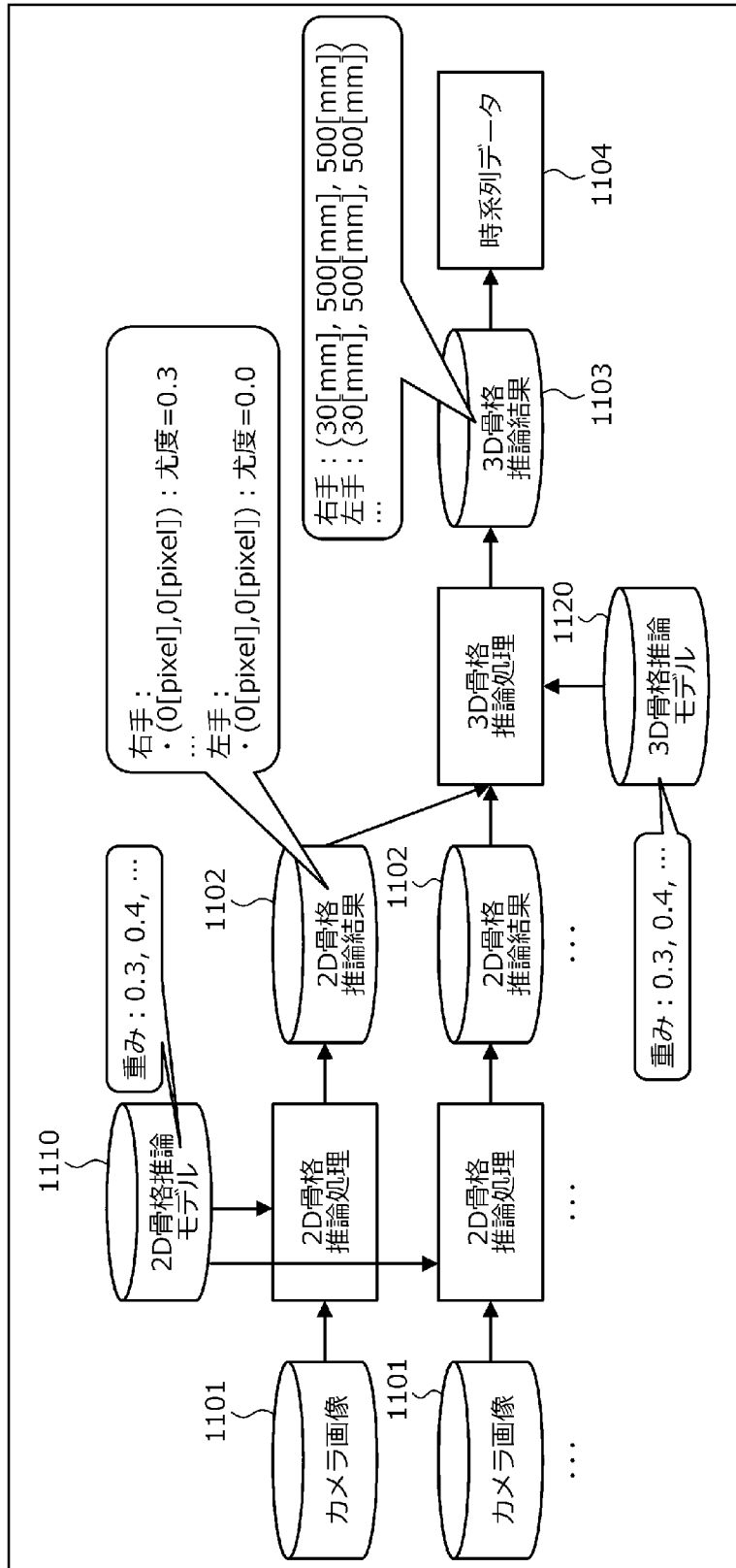
[図9]



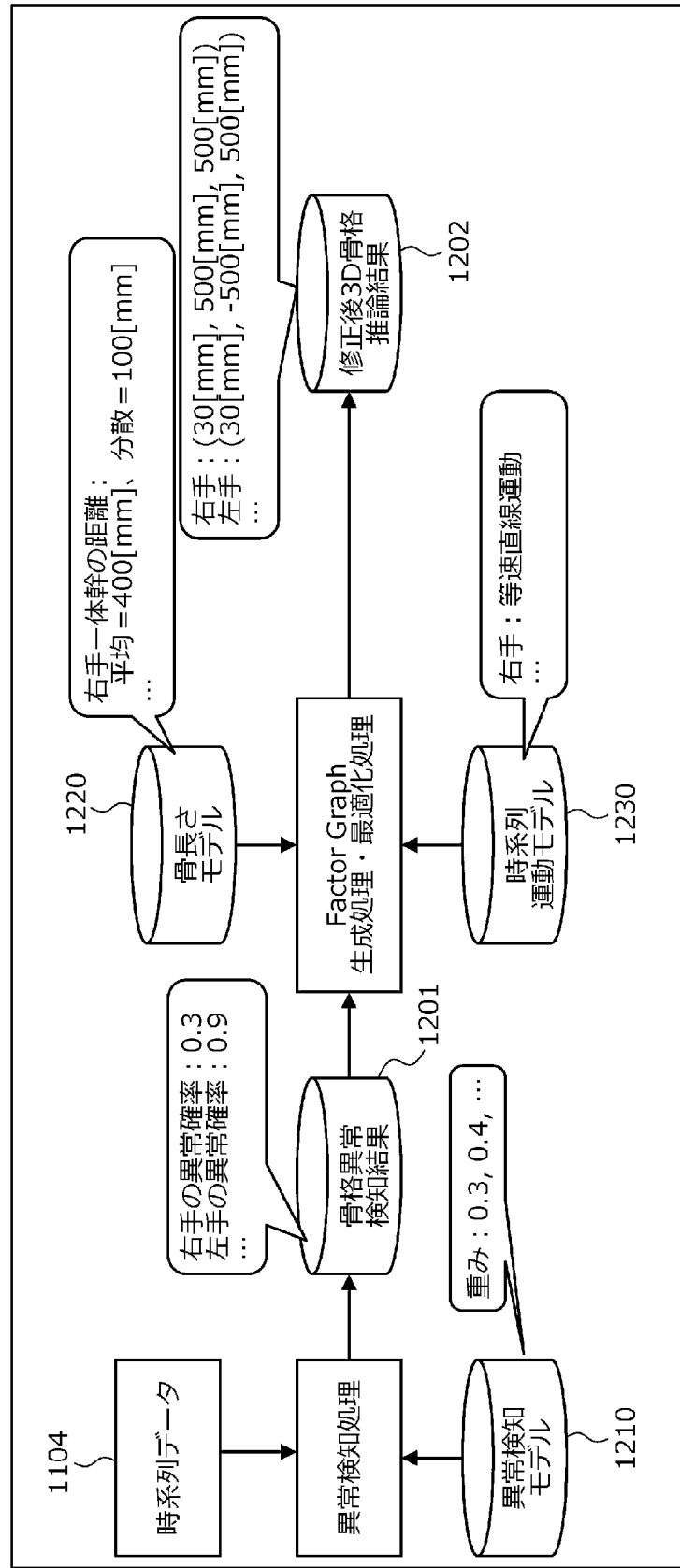
[図10]



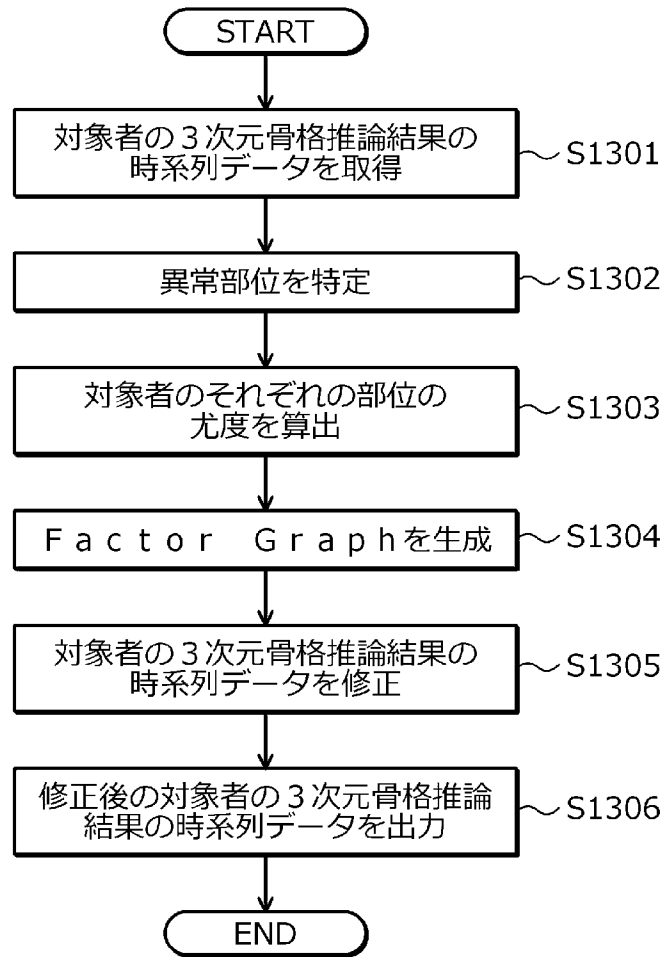
[図11]



[図12]



[図13]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/016363

A. CLASSIFICATION OF SUBJECT MATTER		
G06T 7/00(2017.01)i FI: G06T7/00 350B; G06T7/00 660B		
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		
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 1971-2022 Registered utility model specifications of Japan 1996-2022 Published registered utility model applications of Japan 1994-2022		
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 2021-105887 A (UNIV TOKYO, NTT DOCOMO INC) 26 July 2021 (2021-07-26) paragraph [0026]-[0063], fig. 1-4	1-6
A	CN 113191230 A (INNER MONGOLIA UNIVERSITY OF TECHNOLOGY) 30 July 2021 (2021-07-30) fig. 3-5	4
A	CN 112991656 A (BEIJING UNIVERSITY OF TECHNOLOGY) 18 June 2021 (2021-06-18) fig. 4	4
A	WEI, Shinen et al. Convolutional Pose Machines, [online]. 12 April 2016, https://arxiv.org/pdf/1602.00134.pdf , [retrieval date: 22 April 2022] in particular, fig. 1	1-6
<input type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> 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 22 April 2022		Date of mailing of the international search report 17 May 2022
Name and mailing address of the ISA/JP Japan Patent Office (ISA/JP) 3-4-3 Kasumigaseki, Chiyoda-ku, Tokyo 100-8915 Japan		Authorized officer Telephone No.

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.

PCT/JP2022/016363

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP	2021-105887	A	26 July 2021	(Family: none)	
CN	113191230	A	30 July 2021	(Family: none)	
CN	112991656	A	18 June 2021	(Family: none)	

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 FI: G06T7/00 350B; G06T7/00 660B		
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 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 1971-2022年 日本国実用新案登録公報 1996-2022年 日本国登録実用新案公報 1994-2022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 2021-105887 A (国立大学法人 東京大学, 株式会社NTTドコモ) 26.07.2021 (2021-07-26) 段落[0026]-[0063], [図1]-[図4]	1-6
A	CN 113191230 A (INNER MONGOLIA UNIVERSITY OF TECHNOLOGY) 30.07.2021 (2021-07-30) FIGs.3-5	4
A	CN 112991656 A (BEIJING UNIVERSITY OF TECHNOLOGY) 18.06.2021 (2021-06-18) FIG.4	4
A	Shin-En Wei et al., Convolutional Pose Machines, [online], 2016.04.12, https://arxiv.org/pdf/1602.00134.pdf , [検索日: 2022.04.22] 特に、FIG.1	1-6
<input type="checkbox"/> C欄の続きにも文献が列挙されている。 <input checked="" type="checkbox"/> 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 22.04.2022	国際調査報告の発送日 17.05.2022	
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） 片岡 利延 5H 4881 電話番号 03-3581-1101 内線 3531	

国際調査報告
パテントファミリーに関する情報

国際出願番号

PCT/JP2022/016363

引用文献	公表日	パテントファミリー文献	公表日
JP 2021-105887 A	26.07.2021	(ファミリーなし)	
CN 113191230 A	30.07.2021	(ファミリーなし)	
CN 112991656 A	18.06.2021	(ファミリーなし)	