(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1月27日(27.01.2022)

WIPO | PCT

(10) 国際公開番号
WO 2022/018811 A1

(51) 国際特許分類:
G06T 7/70 (2017.01)

(21) 国際出願番号： PCT/JP2020/028163

(22) 国際出願日： 2020年7月20日(20.07.2020)

(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).

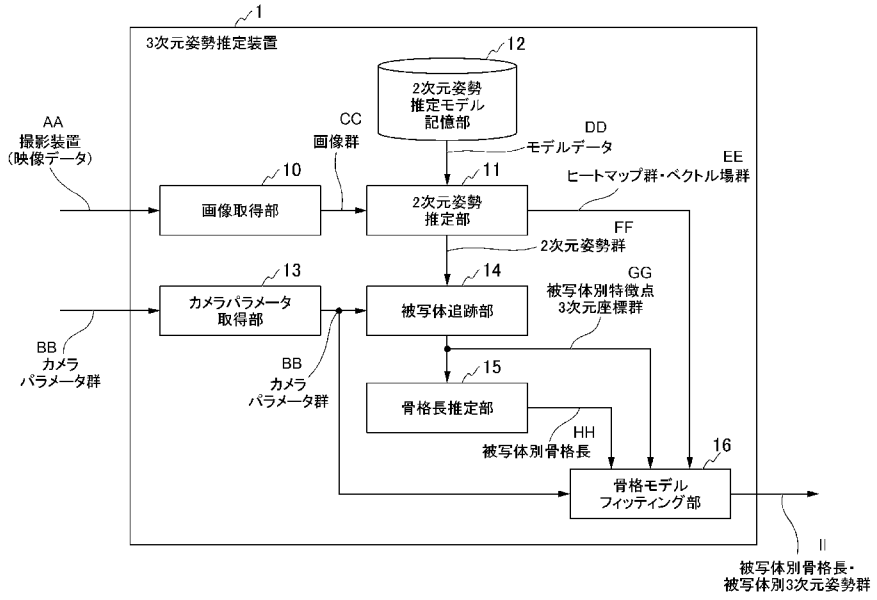
(72) 発明者: 松村 誠明 (**MATSUMURA Masaaki**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ知的財産センタ内 Tokyo (JP). 能登肇 (**NOTO Hajime**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ知的財産センタ内 Tokyo (JP). 山本奏 (**YAMAMOTO Susumu**); 〒1808585 東京都武蔵野市緑町３丁目９－１１ ＮＴＴ知的財産センタ内 Tokyo (JP).

(74) 代理人: 特許業務法人 志賀国際特許事務所 (**SHIGA INTERNATIONAL PATENT OFFICE**); 〒1006620 東京都千代田区丸の内一丁目９番２号 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,

(54) **Title:** THREE-DIMENSIONAL POSTURE OF SUBJECT ESTIMATION DEVICE, THREE-DIMENSIONAL POSTURE ESTIMATION METHOD, AND PROGRAM

(54) 発明の名称： 被写体の３次元姿勢推定装置、３次元姿勢推定方法、及びプログラム



1 Three-dimensional posture estimation device
10 Image acquisition unit
11 Two-dimensional posture estimation unit
12 Two-dimensional posture estimation model storage unit
13 Camera parameter acquisition unit
14 Subject tracking unit
15 Skeleton length estimation unit
16 Skeleton model fitting unit
AA Photographing device (Video data)
BB Camera parameter group
CC Image group
DD Model data
EE Heat map group, vector field group
FF Two-dimensional posture group
GG Three-dimensional coordinate group of feature points by subject
HH Skeleton length by subject
II Skeleton length by subject, three-dimensional posture group by subject

(57) **Abstract:** A three-dimensional posture estimation device provided with: an image acquisition unit for acquiring a plurality of images in which a subject is photographed from a plurality of different points; a two-dimensional posture estimation unit for generating a heat map that represents a two-dimensional posture by subject and existence probability per feature point and a vector field to correct the coordinates around a peak value in the heat map to the direction of the feature points concerned; a two-dimensional posture estimation model storage unit in which the model data used by the two-dimensional posture estimation unit is stored; a camera parameter acquisition unit for acquiring the parameter of a

WO 2022/018811 A1

［続葉有］

HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, 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))

photographing device; a subject tracking unit for generating a three-dimensional coordinate group of feature points by subject; a skeleton length estimation unit for estimating a skeleton length by subject; and a skeleton model fitting unit for estimating a three-dimensional posture by subject.

(57) 要約:被写体が複数の異なる地点から撮影された複数の画像を取得する画像取得部、被写体別に２次元姿勢、特徴点毎の存在確率を表すヒートマップ、及びヒートマップにおけるピーク値周辺の座標を対象特徴点方向に補正するベクトル場を生成する２次元姿勢推定部と、２次元姿勢推定部にて使用されるモデルデータが格納される２次元姿勢推定モデル記憶部と、撮影装置のパラメータを取得するカメラパラメータ取得部と、被写体別特徴点３次元座標群を生成する被写体追跡部と、被写体別の骨格長を推定する骨格長推定部と、被写体別の３次元姿勢を推定する骨格モデルフィッティング部とを備える３次元姿勢推定装置。

明　細　書

発明の名称 ：

被写体の３次元姿勢推定装置、３次元姿勢推定方法、及びプログラム

技術分野

[0001] 本発明は、被写体の３次元姿勢推定装置、３次元姿勢推定方法、及びプログラムに関する。

背景技術

[0002] カメラなどの撮影装置で撮影された画像における被写体の特徴点及び特徴点同士の接続関係を推定する２次元姿勢推定技術が提案されている。特徴点の例としては、被写体の目、鼻、関節などが挙げられる。

被写体の特徴点は図１９のようなツリー状の階層構造で記述される。図１９は、人体における特徴点の例を示す図である。各特徴点の接続関係を記述するベクトル場には階層構造における子の特徴点から親の特徴点方向へのベクトルを生成するように学習がなされる。特徴点１１０は、鼻の位置を表す特徴点である。特徴点１１１は、左目の位置を表す特徴点である。特徴点１１２は、右目の位置を表す特徴点である。図１９に示す例においては、左目の特徴点１１１及び右目の特徴点１１２は、鼻の特徴点１１０を親に持つ。特徴点１１３－１２６は、被写体に定められた他の部位の位置をそれぞれ表す特徴点である。

[0003] 非特許文献１では、Part Confidence Mapsと呼ばれる、体の各特徴点をヒートマップに基づいて予測するネットワーク及びPart Affinity Fieldと呼ばれる特徴点同士の接続関係を予測するネットワークを使用して、複数人のそれぞれの特徴点と特徴点の接続関係を検出している。

非特許文献２では、ハフ投票(Hough voting)と呼ばれる手法により特徴点を抽出する方法が開示されている。具体的には非特許文献２では体の各特徴点を予測するヒートマップと特徴点同士の接続関係を予測するMid-range offsetに加え、short-range offsetと呼ぶ近傍の特徴点を指し示すベクトルから

なるベクトル場を使用することで、複数人から特徴点と特徴点の接続関係を検出している。

[0004] また、複数の撮影装置を用いて被写体を撮影し、撮影した画像を基にして被写体特徴点の３次元座標を推定する技術が提案されている。３次元的な骨格の座標推定には撮影した複数の画像から得られる２次元平面における特徴点の２次元座標が使用される。特徴点の２次元座標を撮影装置のパラメータ（撮影装置の位置、向き、視野角、歪みなど）を基に特徴点を３次元空間に投影し、三角測量の手法を用いることで特徴点の３次元座標を計算する。しかし、特徴点の２次元座標や撮影装置のパラメータには推定誤差があり、三角測量の手法により幾何学的に３次元座標を決定することはできない場合がある。そのため、推定した特徴点の３次元座標を画像上に再度投影し、投影した点の２次元座標（以下、スクリーン座標と呼ぶ）と特徴点の２次元座標の差が最小となる点を３次元空間における特徴点に決定するという手法が一般的に行われる。

非特許文献３では、単一の被写体に対して各特徴点のヒートマップを生成し（被写体が複数名存在する場合はその数だけ実施）、他者による遮蔽の影響を考慮して補正したヒートマップを用いて２次元平面における特徴点の２次元座標を導出し、三角測量で取得した特徴点の３次元座標が人体の骨格モデル（各骨格長、最大３軸の関節回転角、及びroot関節の３軸座標で定義されるモデル）にフィットするように関節回転角及び３軸座標を調整することで３次元姿勢を推定する手法が提案されている。

先行技術文献

非特許文献

[0005] 非特許文献1：Cao, Z., Hidalgo, G., Simon, T., Wei, S.E., Sheikh, Y.: OpenPose: realtime multi-person 2D pose estimation using Part Affinity Fields. In: arXiv preprint arXiv:1812.08008, 2018.

非特許文献2：G. Papandreou, T. Zhu, L.-C. Chen, S. Gidaris, J. Tompson, and K. Murphy. PersonLab: Person pose estimation and instance segm

entation with a bottom-up, part-based, geometric embedding model. arXiv:1803.08225, 2018.

非特許文献3：T.Ohashi1,Y.Ikegami,Y.Nakamura：Synergetic Reconstruction from 2D Pose and 3D Motion for Wide-Space Multi-Person Video Motion Capture in the Wild. arXiv preprint arXiv:2001.05613v1,2020.

発明の概要

発明が解決しようとする課題

[0006] しかし、前述の特徴点の３次元座標の推定においては、各撮影装置で撮影した画像に対して得られる特徴点は被写体別に特徴点毎、最も確度が高い１点であり、三角測量では、この座標１点とスクリーン座標との差を最小化するため、この座標１点において遮蔽物等により座標の推定誤りが生じた撮影装置が存在する場合は、３次元座標の推定精度が低下する（例えば４台の撮影装置で被写体を撮影したときに、１台の撮影装置に推定誤りが生じた場合、その誤った座標方向に25%の影響力で誤差が重畳する）。非特許文献３では、この影響を軽減するため、他者による遮蔽の影響を考慮して補正したヒートマップを用いて２次元平面における特徴点の２次元座標を導出する機構を導入することで、推定誤りの影響を軽減しつつ人体の骨格モデルをフィッティングする手法を提案しているが、遮蔽の影響を排除しきれなかった際には誤差が重畳することとなるうえ、画像内に被写体が複数名存在する場合は２次元姿勢推定技術を複数回実施する必要があり、膨大な演算量を必要とする。

推定誤りによる影響力を軽減する別手法として、任意の３次元座標をスクリーン座標に再投影し、全ての撮影装置におけるヒートマップの信号合計値を対象特徴点の存在確率とし、高い値を示す方向に例えば勾配法を用いて骨格モデルをフィッティングする手法も考えられる。非特許文献１や非特許文献２のように、複数人に対応した２次元姿勢推定技術を用いて得られるヒートマップにより、画像内の被写体数に依存することなく前述の手法と比較して演算量を削減することができる。しかし、複数人に対応した２次元姿勢推

定技術におけるヒートマップ解像度は３次元座標を推定する用途には十分ではなく、遠方の被写体などではヒートマップの１画素のズレによって大きな誤差を生じることがあり、精度が著しく低下する恐れがある。

本発明の目的は、複数人に対応した２次元姿勢推定技術を使用しつつ、３次元空間における被写体の３次元姿勢をより高い精度で推定する３次元姿勢推定装置を提供することにある。

課題を解決するための手段

[0007] 本発明の一態様は、被写体が複数の異なる地点から撮影された複数の画像を取得する画像取得部と、被写体別に２次元姿勢、特徴点毎の存在確率を表すヒートマップ、及びヒートマップにおけるピーク値周辺の座標を対象特徴点方向に補正するベクトル場を生成する２次元姿勢推定部と、２次元姿勢推定部にて使用されるモデルデータが格納される２次元姿勢推定モデル記憶部と、撮影装置のパラメータを取得するカメラパラメータ取得部と、被写体別特徴点３次元座標群を生成する被写体追跡部と、被写体別の骨格長を推定する骨格長推定部と、被写体別の３次元姿勢を推定する骨格モデルフィッティング部とを備える３次元姿勢推定装置である。

発明の効果

[0008] 本発明によれば、複数人に対応した２次元姿勢推定技術を使用しつつ、３次元空間における被写体の３次元姿勢をより高い精度で推定する３次元姿勢推定装置を提供することができる。

図面の簡単な説明

[0009] [図1]モーションキャプチャシステムの構成を示す概略図である。

[図2]第１の実施形態に係る姿勢推定装置の構成を示す図である。

[図3A]２次元姿勢推定部が出力する２次元姿勢群の一例である。

[図3B]２次元姿勢推定部が出力する２次元姿勢の座標の一例である。

[図4]２次元姿勢推定部が出力するヒートマップの一例である。

[図5]２次元姿勢推定部が出力するベクトル場の一例である。

[図6A]被写体追跡部による被写体識別子の割り当ての一例である。

[図6B]被写体追跡部が出力する被写体別特徴点３次元座標群の一例である。

[図7]骨格長推定部が出力する被写体別骨格長の一例である。

[図8]第１の実施形態に係る骨格モデルフィッティング部の構成を示す図である。

[図9A]ピーク座標を含むヒートマップの一例である。

[図9B]ピーク座標出力部が出力するピーク値周辺座標群の一例である。

[図10]ハフ投票点群出力部が出力するハフ投票点群の一例である。

[図11]第１の実施形態に係る骨格モデルフィッティング部の動作を示すフローチャートである。

[図12]第２の実施形態に係る骨格モデルフィッティング部の構成を示す図である。

[図13A]ピーク座標とラベル付きピーク座標の一例である。

[図13B]ラベル付きピーク座標出力部が出力するラベル付きピーク座標群の一例である。

[図14]ラベル付きハフ投票点群を用いて計算対象から除外する手法を示す図である。

[図15]第２の実施形態に係る骨格モデルフィッティング部の動作を示すフローチャートである。

[図16]第３の実施形態に係る骨格モデルフィッティング部の構成を示す図である。

[図17A]ピーク座標と信号値付きピーク座標の一例である。

[図17B]信号値付きピーク座標出力部が出力する信号値付きピーク座標群の一例である。

[図18]第３の実施形態に係る骨格モデルフィッティング部の動作を示すフローチャートである。

[図19]人体における特徴点の例を示す図である。

発明を実施するための形態

[0010] 以下、図面を参照しながら本発明の実施形態について詳しく説明する。

[0011] 〈第１の実施形態〉

《学習済みモデルを用いたマーカーレスモーションキャプチャについて》

以下、図面を参照しながら実施形態について詳しく説明する。

図１は、モーションキャプチャシステム２の構成を示す概略図である。

マーカーレスモーションキャプチャを実現するモーションキャプチャシステム２について説明する。モーションキャプチャシステム２は、複数の撮影装置２０１と、３次元姿勢推定装置１を備える。複数の撮影装置２０１は、それぞれ異なる方向から被写体Ａ１及びＡ２を撮影するように設置される。複数の撮影装置２０１は、互いに内部時計の時刻が同期されている。

[0012] ３次元姿勢推定装置１は、複数の撮影装置２０１から映像データを取得し、各映像データのフレーム（画像）ごとに被写体Ａ１及びＡ２の各３次元姿勢を推定する。つまり、３次元姿勢推定装置１は、複数の映像データの同時刻に係る画像における特徴点の２次元座標と、現実空間（３次元空間）における複数の撮影装置２０１のパラメータに基づいて、三角測量により、被写体Ａ１及びＡ２の各特徴点の３次元座標を推定することで、被写体Ａ１及びＡ２の各３次元姿勢を推定することができる。

[0013] 《３次元姿勢推定装置の構成》

図２は、第１の実施形態に係る３次元姿勢推定装置１の構成を示す図である。３次元姿勢推定装置１は、画像取得部１０、２次元姿勢推定部１１、２次元姿勢推定モデル記憶部１２、カメラパラメータ取得部１３、被写体追跡部１４、骨格長推定部１５、骨格モデルフィッティング部１６を備える。

[0014] 画像取得部１０は、複数の撮影装置２０１それぞれから映像データ（画像群）を取得する。複数の撮影装置２０１が撮影する映像データには、１個以上の同一の被写体が写る。被写体は例えば人間である。２次元姿勢推定部１１は、画像群及びモデルデータを入力として被写体別に分離された２次元姿勢群、ヒートマップ群、及びベクトル場群を出力する。２次元姿勢推定モデル記憶部１２は、画像を入力として２次元姿勢推定部１１の出力が得られるように学習されたモデルデータを記憶する。

[0015] カメラパラメータ取得部１３は、外部から撮影装置２０１の位置、向き、視野角、歪みなどを記述したパラメータを取得し、各撮影装置のカメラパラメータ群を出力する。被写体追跡部１４は、カメラパラメータ群と２次元姿勢群を入力として被写体別特徴点３次元座標群を出力する。骨格長推定部１５は、被写体別特徴点３次元座標群を入力として被写体別骨格長を出力する。骨格モデルフィッティング部１６はカメラパラメータ群、ヒートマップ群、ベクトル場群、被写体別特徴点３次元座標群、及び被写体別骨格長を入力とし、被写体別骨格長及び被写体別３次元姿勢群を外部のデータベースなどに出力する。

[0016] 図３Ａは被写体Ａ１、Ａ２及び被写体の２次元姿勢Ｐ１、Ｐ２を示す図である。図３Ｂは、被写体別に分離された２次元姿勢の一例である。２次元姿勢推定部１１は、学習済みのモデルデータに画像群を入力することで各画像に対し、図３Ａに示すような２次元姿勢Ｐ１及びＰ２、図４に示すようなヒートマップＨ１、並びに図５に示すようなベクトル場Ｖ１をそれぞれ生成する。２次元姿勢は、図３Ｂに示すように、被写体の特徴点毎の２次元座標によって表される。つまり、２次元姿勢は、特徴点の２次元位置を示す２次元座標群であるといえる。ヒートマップＨ１及びベクトル場Ｖ１は、各画像の特徴点毎に生成される。ベクトル場Ｖ１は、ヒートマップにおけるピーク値周辺の座標を対象特徴点方向に補正するものである。

図４はヒートマップの一例である。ヒートマップＨ１には、画像に写る被写体Ａ１及びＡ２が写っており、被写体Ａ１及びＡ２の対象特徴点の近傍には、特徴点の存在確率を表す確率場が格納されている。確率場は対象特徴点近傍ほど高い値を示し、距離が離れるに従って減衰する特徴を持つ。図４に示す例においては、確率場の色が白に近いほど特徴点の存在確率が高いことを示す。２次元姿勢推定部１１は、２次元姿勢の集合を２次元姿勢群として、ヒートマップの集合をヒートマップ群として、ベクトル場の集合をベクトル場群として出力する。

[0017] 図４に示すヒートマップ及び図５に示すベクトル場は、ともに被写体Ａ１

及びＡ２の右ひじの特徴点に基づいて生成されたヒートマップ及びベクトル場である。右ひじの特徴点は図１９における特徴点１１８に対応する。また、機械学習により学習させたモデルデータを使用する場合、左右の逆方向にも微弱な反応を示す場合がある。そのため図４に示すヒートマップ及び図５に示すベクトル場においては、左ひじの特徴点（図１９において特徴点１１７に対応する）の周囲にもわずかに存在確率を表す確率場及び特徴点方向に補正するベクトル場が生成されている。

２次元姿勢群に含まれる２次元姿勢の数は入力画像の数と同じであり、ヒートマップ群に含まれるヒートマップ及びベクトル場群に含まれるベクトル場の数は、入力画像の数と被写体の数と特徴点の数の積と同じである。

[0018] 被写体追跡部１４は、カメラパラメータ群及び２次元姿勢群を入力として、前述のように同一フレームにおける特徴点の２次元座標と三角測量を用いてフレーム毎に被写体別特徴点３次元座標を求める。このとき、被写体が複数名撮像空間内に存在する場合、被写体追跡部１４は、例えば図６Ａに示すように各被写体の同一特徴点におけるフレーム間３次元距離が最も近くなる被写体同士を同一被写体としての識別子（被写体識別子）を割り当てることで、図６Ｂに示すような被写体別特徴点３次元座標群を出力する。なお、被写体追跡部１４が算出する被写体別特徴点３次元座標群は、従来手法と同じ三角測量を用いるため骨格長はフレーム毎に異なる値となる。

[0019] 図７は、被写体別骨格長の一例である。骨格長推定部１５は、被写体別特徴点３次元座標群を入力することで同一被写体のフレーム毎に異なる骨格長の中から例えば平均値を算出することで、各被写体の骨格モデルにおける骨格長を決定する。骨格モデルは、被写体を骨と関節で表すモデルである。各関節には可動域が設定される。各骨には骨格長が設定される。一部の特徴点は関節と一致する。上腕の長さなど、肩と肘の特徴点間距離により計測できる骨格長には計測した値を用い、背骨の長さなど計測困難な骨格長は、その他特徴点との距離関係から導出する推定値を用いる。なお、ここでは骨格長を決定するために平均値を用いたが、中央値や最頻値などを用いてもよい。

[0020]《骨格モデルフィッティング部の構成》

図８は、第１の実施形態に係る骨格モデルフィッティング部１６の構成を示す図である。第１の実施形態に係る骨格モデルフィッティング部１６は、ピーク座標出力部１６１０、ハフ投票点群出力部１６１１、被写体抽出部１６１２、フレーム抽出部１６１３、３次元姿勢推定部１６１４、対象フレーム３次元姿勢推定部１６１５、３次元姿勢記憶部１６１６、被写体出力部１６１７を備える。

[0021] ピーク座標出力部１６１０は、ヒートマップ群を入力としてピーク値周辺座標群を出力する。ハフ投票点群出力部１６１１は、ベクトル場群及びピーク値周辺座標群を入力としてハフ投票点群を出力する。被写体抽出部１６１２は、被写体別特徴点３次元座標群及び被写体別骨格長を入力として対象となる被写体別に被写体特徴点３次元座標群及び被写体骨格長を出力する。フレーム抽出部１６１３は、被写体特徴点３次元座標群及びハフ投票点群を入力として対象フレーム別にフレーム被写体特徴点３次元座標群及びフレームハフ投票点群を出力する。３次元姿勢推定部１６１４は、被写体骨格長、フレーム被写体特徴点３次元座標群、過去被写体３次元姿勢群を入力として推定３次元姿勢を出力する。対象フレーム３次元姿勢推定部１６１５は、カメラパラメータ群、被写体骨格長、推定３次元姿勢、及びフレームハフ投票点群を入力としてフレーム被写体３次元姿勢を３次元姿勢記憶部１６１６に出力する。３次元姿勢記憶部１６１６は、被写体骨格長及びフレーム被写体３次元姿勢を入力として記憶し、要求に応じて記憶したデータを出力する。被写体出力部１６１７は、被写体別骨格長及び被写体別３次元姿勢群を３次元姿勢記憶部１６１６から取得して出力する。

[0022] 図９Ａはピーク座標を含むヒートマップの一例である。図９Ｂはピーク座標出力部が出力するピーク値周辺座標群の一例である。

ピーク座標出力部１６１０は、ヒートマップ群を入力として各ヒートマップにおけるピークを示す座標値周辺の座標をピーク値周辺座標群として出力する。このとき、ピーク座標は、例えば図９Ａに示すように周辺信号値との

勾配が凸になる部分を探索することで検出し、ピーク座標のヒートマップ信号値と近しい値をとるピーク座標周辺の座標を例えば所定の数量もしくはピーク座標のヒートマップ信号値との差が一定値以内の座標をサンプリングし、ピーク値周辺座標群を生成する。

[0023] 図１０は、ハフ投票点群出力部１６１１が出力するハフ投票点群の一例である。ハフ投票点群出力部１６１１は、ベクトル場群及びピーク値周辺座標群を入力として、図１０に示すようにピーク値周辺座標群の各座標に対してベクトル場で補正したハフ投票点を、全ての座標に対してまとめたハフ投票点群を出力する。

つまりハフ投票点群出力部１６１１は、ベクトル場により被写体の特徴点の３次元座標群を補正する。

[0024] 過去被写体３次元姿勢群は、被写体抽出部１６１２にて対象とする被写体の過去フレームにおける３次元姿勢の集合を表す。

[0025] ３次元姿勢推定部１６１４は、例えば先頭フレームや他物体による遮蔽によって対象フレーム直前の３次元姿勢が存在しない場合は、フレーム被写体特徴点３次元座標群における特徴点の３次元座標と対応する骨格モデルにおける関節の３次元座標間の誤差が最小になるようInverse KinematicsやForward Kinematicsによって被写体骨格長に設定した骨格モデルの関節回転角群を推定し、推定３次元姿勢を生成する。それ以外のフレームにおいては、過去被写体３次元姿勢群から対象フレームの関節回転角群を、例えば物理シミュレーションなどによって推定して推定３次元姿勢を生成する。

[0026] 対象フレーム３次元姿勢推定部１６１５は、被写体骨格長と推定３次元姿勢から骨格モデルにおける全身関節の３次元座標群を求め、各関節の３次元座標を各撮影装置の画面に再投影したスクリーン座標と、対象関節と対応する特徴点におけるフレームハフ投票点群との間の距離を基にして特徴点の存在確率を算出する。例えば、対象フレーム３次元姿勢推定部１６１５は、式（１）により対象特徴点の存在確率を算出する。

[0027]

[数1]

$$V = \frac{1}{C\sum_{i=1}^{C} N_i} \sum_{i=1}^{C} \sum_{j=1}^{N_i} \left(1 - \frac{l_{i,j}}{L}\right) \quad \cdots(1)$$

[0028] ここでＣは撮影装置の数、Ｎ$_i$はｉ番目の撮影装置における対象特徴点に含まれるハフ投票点の数、ｌ$_{i,j}$はｉ番目の撮影装置におけるｊ番目のハフ投票点が指し示す座標とスクリーン座標との距離（最大値をＬとする）、Ｌはあらかじめ設定する正の定数である。すなわち、対象フレーム３次元姿勢推定部１６１５は、スクリーン座標と複数のハフ投票点が指し示す座標との距離に応じた特徴点の存在確率を用いて骨格モデルにおける各関節の回転角のフィッティングを行う。

[0029] 対象フレーム３次元姿勢推定部１６１５は、スクリーン座標と複数のハフ投票点が指し示す座標との距離に応じた特徴点の存在確率を用いて骨格モデルにおける各関節の回転角のフィッティングを行うとき、例えば各関節の回転角を微小変化させた際の勾配を求め、勾配法を用いて全ての特徴点の合計存在確率が最大となるよう骨格モデルをフィッティングすることでフレーム被写体３次元姿勢を生成し出力する。

[0030] 《骨格モデルフィッティング部の動作》

次に、第１の実施形態に係る骨格モデルフィッティング部１６の動作について説明する。

図１１は、第１の実施形態に係る骨格モデルフィッティング部１６の動作を示すフローチャートである。

初めに、ピーク座標出力部１６１０は、ヒートマップ群の各ヒートマップに対してピーク値周辺の座標群をサンプリングする（ステップＳ１００）。次に、ハフ投票点群出力部１６１１は、ピーク値周辺の各座標に対して対応するフレーム及び特徴点のベクトル場を用いて座標値を修正する（ステップＳ１０１）。被写体抽出部１６１２は、被写体別骨格長から被写体の人数を取得し、対象とする被写体を１つずつ選択し、骨格モデルフィッティング部

１６は被写体毎にステップＳ１０３からステップＳ１１１を繰り返す（ステップＳ１０２）。続いて、フレーム抽出部１６１３は、ハフ投票点群から総フレーム数を取得し、対象とするフレーム（対象フレーム）を１つずつ選択し、骨格モデルフィッティング部１６は、フレーム毎にステップＳ１０４からステップＳ１１１を繰り返す（ステップＳ１０３）。３次元姿勢推定部１６１４は、被写体骨格長、フレーム被写体特徴点３次元座標群、及び過去被写体３次元姿勢群から対象フレームにおける推定３次元姿勢を生成する（ステップＳ１０４）。

[0031] 対象フレーム３次元姿勢推定部１６１５は、被写体骨格長に設定した骨格モデルの関節回転角を推定３次元姿勢に設定し（ステップＳ１０５）、特徴点に対応する全ての関節から１つずつ選択して関節毎にステップＳ１０７からステップＳ１０９を繰り返す（ステップＳ１０６）。続いて、対象フレーム３次元姿勢推定部１６１５は対象関節を各回転軸プラスマイナス方向に各微小角度回転させた際の子関節（子特徴点）の存在確率を式（１）にてそれぞれ計算する（ステップＳ１０７）。なお、ステップＳ１０７において、対象フレーム３次元姿勢推定部１６１５は対象関節の回転角自由度や回転範囲の制限を超えて回転する方向には変位しないよう制約を設けてもよい。

次に、対象フレーム３次元姿勢推定部１６１５はステップＳ１０７で計算した各回転軸の存在確率の勾配から、子関節（子特徴点）の存在確率が高くなる方向に対象関節の回転角を修正する（ステップＳ１０８）。対象フレーム３次元姿勢推定部１６１５は、修正した対象関節の回転角を用いて子関節（子特徴点）の存在確率を式（１）にて算出する（ステップＳ１０９）。対象フレーム３次元姿勢推定部１６１５は、全ての特徴点についての算出される存在確率の総和の変化量に基づいて、３次元姿勢を確定するか否かを判定する（ステップＳ１１０）。存在確率の総和が所定の量を超える場合（ステップＳ１１０－ＹＥＳ）、対象フレーム３次元姿勢推定部１６１５は直近に定められた３次元姿勢を対象フレームの３次元姿勢として決定し、３次元姿勢記憶部１６１６に被写体骨格長及びフレーム被写体３次元姿勢を保存する

（ステップＳ１１１）。存在確率の和が所定の値以下である場合（ステップＳ１１０－ＮＯ）、対象フレーム３次元姿勢推定部１６１５は、ステップＳ１０６に処理を戻し、存在確率の総和が増加するように、骨格モデルの関節の角度を調整しながら、関節の回転角を補正する。

[0032] 被写体出力部１６１７は、３次元姿勢記憶部１６１６から被写体別に被写体骨格長及び過去被写体３次元姿勢群を取得し、被写体別骨格長及び被写体別３次元姿勢群を出力する（ステップＳ１１２）。

[0033] 《作用・効果》

このように、第１の実施形態によれば、３次元姿勢推定装置１はヒートマップ及びベクトル場からハフ投票点を生成し、骨格モデルの関節（特徴点）毎に関節の３次元座標をスクリーン座標に投影して、対象関節（特徴点）におけるハフ投票点との距離から算出される存在確率を最大化するよう被写体３次元姿勢を決定する。３次元姿勢推定装置１はハフ投票点群を用いた確率場を使用していることから、ヒートマップにおける１点を決定するよりも誤差の影響を取り除くことができ、より正確に被写体の３次元姿勢を推定することができる。

[0034] 〈第２の実施形態〉

第１の実施形態に係る３次元姿勢推定装置１は、ヒートマップ及びベクトル場からハフ投票点を生成し、骨格モデルの関節（特徴点）毎に関節の３次元座標をスクリーン座標に投影して、対象関節（特徴点）におけるハフ投票点との距離から算出される存在確率を最大化するように被写体３次元姿勢を決定する。これに対し、第２の実施形態に係る３次元姿勢推定装置１は、ヒートマップのピーク値周辺画素に対してラベリングを行うことで、計算量を低減しつつ、被写体の３次元姿勢を決定する。

[0035] 図１２は、第２の実施形態に係る骨格モデルフィッティング部１６の構成を示す図である。第２の実施形態に係る骨格モデルフィッティング部１６は、第１の実施形態の構成におけるピーク座標出力部１６１０、ハフ投票点群出力部１６１１、対象フレーム３次元姿勢推定部１６１５の代わりに、それ

ぞれラベル付きピーク座標出力部１６２０、ラベル付きハフ投票点群出力部１６２１、対象フレーム３次元姿勢高速推定部１６２２を備える。

[0036] 図１３Ａはピーク座標とラベル付きピーク座標の一例である。図１３Ｂはラベル付きピーク座標出力部１６２０が出力するラベル付きピーク座標群の一例である。ラベル付きピーク座標出力部１６２０は、ヒートマップ群を入力として各ヒートマップにおけるピークを示す座標値周辺の座標及びラベルを、ラベル付きピーク値周辺座標群として出力する。基本的な機能はピーク座標出力部１６１０とほぼ同一であるが、図１３Ｂに示すように隣接するピーク値周辺座標について同一のラベルを割り当て、出力データにラベルを含める部分が異なる。ラベル付きハフ投票点群出力部１６２１の基本的な機能はハフ投票点群出力部１６１１とほぼ同一であるが、入力がピーク値周辺座標群の代わりにラベル付きピーク値周辺座標群となり出力がラベル付きハフ投票点群となる部分が異なる。

[0037] 対象フレーム３次元姿勢高速推定部１６２２は、基本的な機能は対象フレーム３次元姿勢推定部１６１５とほぼ同一であるが、入力がハフ投票点群の代わりにラベル付きハフ投票点群となり、スクリーン座標と対象関節と対応する特徴点におけるラベル付きフレームハフ投票点群との間の距離を基にして特徴点の存在確率を算出する部分が異なる。例えば、対象フレーム３次元姿勢高速推定部１６２２は、式（２）により対象特徴点の存在確率を算出する。

[0038] [数2]

$$V = \frac{1}{C\sum_{i=1}^{C} M_i} \sum_{i=1}^{C} \sum_{j=1}^{M_i} \left(1 - \frac{l_{i,j}}{L}\right) \quad \cdots(2)$$

[0039] ここでＭ$_i$はｉ番目の撮影装置におけるスクリーン座標が属するラベルと同一のラベルを持つラベル付きハフ投票点群の数である。すなわち、図１４に示すようにスクリーン座標が属するラベルｂ以外のラベルが付与されたハフ投票点群は計算対象から除外される。

[0040] 図１５は、第２の実施形態に係る骨格モデルフィッティング部１６の動作を示すフローチャートである。

初めに、ラベル付きピーク座標出力部１６２０は、ヒートマップ群の各ヒートマップに対してピーク値周辺の座標群をサンプリングし（ステップＳ２００）、サンプリングした座標にラベルを付与する（ステップＳ２０１）。次に、ラベル付きハフ投票点群出力部１６２１は、ピーク値周辺の各座標に対して対応するフレーム及び特徴点のベクトル場を用いて座標値を修正する（ステップＳ２０２）。被写体抽出部１６１２は、被写体別骨格長から被写体の人数を取得し、対象とする被写体を１つずつ選択し、骨格モデルフィッティング部１６は被写体毎にステップＳ２０４からステップＳ２１２を繰り返す（ステップＳ２０３）。続いて、フレーム抽出部１６１３は、ラベル付きハフ投票点群から総フレーム数を取得し、対象とするフレームを１つずつ選択し、骨格モデルフィッティング部１６はフレーム毎にステップＳ２０５からステップＳ２１２を繰り返す（ステップＳ２０４）。次に３次元姿勢推定部１６１４は、被写体骨格長、フレーム被写体特徴点３次元座標群、及び過去被写体３次元姿勢群から対象フレームにおける推定３次元姿勢を生成する（ステップＳ２０５）。

[0041] 対象フレーム３次元姿勢高速推定部１６２２は、被写体骨格長に設定した骨格モデルの関節回転角を推定３次元姿勢に設定し（ステップＳ２０６）、特徴点に対応する全ての関節から１つずつ選択して関節毎にステップＳ２０８からステップＳ２１０を繰り返す（ステップＳ２０７）。続いて、対象フレーム３次元姿勢高速推定部１６２２は、対象関節を各回転軸プラスマイナス方向に各微小角度回転させた際の子関節（子特徴点）の存在確率を式（２）にてそれぞれ計算する（ステップＳ２０８）。なお、対象フレーム３次元姿勢高速推定部１６２２は、ステップＳ２０８において、対象関節の回転角自由度や回転範囲の制限を超えて回転する方向には変位しないよう制約を設けてもよい。次に、対象フレーム３次元姿勢高速推定部１６２２はステップＳ２０８で計算した各回転軸の存在確率の勾配から、子関節（子特徴点）の

存在確率が高くなる方向に対象関節の回転角を修正する（ステップＳ２０９）。対象フレーム３次元姿勢高速推定部１６２２は、修正した対象関節の回転角を用いて子関節（子特徴点）の存在確率を式（２）にて算出する（ステップＳ２１０）。対象フレーム３次元姿勢高速推定部１６２２は、全ての特徴点についての算出される存在確率の総和の変化量に基づいて、３次元姿勢を確定するか否かを判定する（ステップＳ２１１）。存在確率の総和が所定の量を超える場合（ステップＳ２１１－ＹＥＳ）、対象フレーム３次元姿勢高速推定部１６２２は直近に定められた３次元姿勢を対象フレームの３次元姿勢として決定し、３次元姿勢記憶部１６１６に被写体骨格長及びフレーム被写体３次元姿勢を保存する（ステップＳ２１２）。存在確率の和が所定の値以下である場合（ステップＳ２１１－ＮＯ）、対象フレーム３次元姿勢高速推定部１６２２は、ステップＳ２０７に処理を戻し、存在確率の総和が増加するように、骨格モデルの関節の角度を調整しながら、関節の回転角を補正する。

[0042] 被写体出力部１６１７は、３次元姿勢記憶部１６１６から被写体別に被写体骨格長及び過去被写体３次元姿勢群を取得し、被写体骨格長及び被写体別３次元姿勢群を出力する（ステップＳ１１２）。

[0043] 《作用・効果》

このように、第２の実施形態によれば、３次元姿勢推定装置１はヒートマップのピーク値周辺画素に対してラベリングを行うことで、スクリーン座標が属するラベル以外のハフ投票点群は計算対象から除外することができる。これにより、３次元姿勢推定装置１はより速く被写体の３次元姿勢を推定することができる。

[0044] 〈第３の実施形態〉

第２の実施形態に係る３次元姿勢推定装置１は、ヒートマップのピーク値周辺画素に対してラベリングを行うことで、スクリーン座標が属するラベル以外のハフ投票点群は計算対象から除外することで存在確率を求める。これに対し、第３の実施形態に係る３次元姿勢推定装置１は、ヒートマップ信号

値を用いて存在確率の推定精度向上を行う。

図１６は、第３の実施形態に係る骨格モデルフィッティング部１６の構成を示す図である。第３の実施形態に係る骨格モデルフィッティング部１６は、第１の実施形態の構成におけるピーク座標出力部１６１０、ハフ投票点群出力部１６１１、及び対象フレーム３次元姿勢推定部１６１５の代わりに、それぞれ信号値付きピーク座標出力部１６３０、信号値付きハフ投票点群出力部１６３１、及び対象フレーム３次元姿勢高精度推定部１６３２を備える。

[0045] 図１７Ａはピーク座標と信号値付きピーク座標の一例である。図１７Ｂは信号値付きピーク座標出力部が出力する信号値付きピーク座標群の一例である。信号値付きピーク座標出力部１６３０は、ヒートマップ群を入力として各ヒートマップにおけるピークを示す座標値周辺の座標及びヒートマップ信号値を、信号値付きピーク値周辺座標群として出力する。基本的な機能はピーク座標出力部１６１０とほぼ同一であるが、図１７Ｂに示すように対象座標のヒートマップ信号値を出力データに含める部分が異なる。信号値付きハフ投票点群出力部１６３１の基本的な機能はハフ投票点群出力部１６１１とほぼ同一であるが、入力がピーク値周辺座標群の代わりに信号値付きピーク値周辺座標群となり出力が信号値付きハフ投票点群となる部分が異なる。

[0046] 対象フレーム３次元姿勢高精度推定部１６３２は、基本的な機能は対象フレーム３次元姿勢推定部１６１５とほぼ同一であるが、入力がハフ投票点群の代わりに信号値付きハフ投票点群となり、スクリーン座標と対象関節と対応する特徴点における信号値付きフレームハフ投票点群との間の距離を基にして特徴点の存在確率を算出する部分が異なる。例えば、対象フレーム３次元姿勢高精度推定部１６３２は、式（３）により対象特徴点の存在確率を算出する。

[0047] [数3]

$$V = \frac{1}{C\sum_{i=1}^{C} N_i} \sum_{i=1}^{C} \sum_{j=1}^{N_i} h_{i,j} \left(1 - \frac{l_{i,j}}{L}\right) \quad \cdots(3)$$

[0048] ここで$h_{i,j}$はｉ番目のヒートマップにおけるｊ番目の画素の特徴点の存在確率の値である。すなわち、ヒートマップ信号値に応じた加重和を、存在確率として算出する。

[0049] ２次元姿勢推定技術では、遮蔽物などで特徴点が隠れている場合においても確からしい座標周辺にヒートマップの信号値は低いながらもピークが現れる特徴を持つが、これらのピーク値周辺座標におけるハフ投票点群を、当該特徴点が明確に撮影された撮影装置２０１から取得されたヒートマップのピーク値周辺座標におけるハフ投票点群と同列に扱うと誤差が重畳する。そのため、ヒートマップ信号値に応じた加重和を用いて存在確率を算出することで、誤差を軽減し３次元姿勢推定精度を高めることができる。

[0050] 図１８は、第３の実施形態に係る骨格モデルフィッティング部１６の動作を示すフローチャートである。

初めに、信号値付きピーク座標出力部１６３０は、ヒートマップ群の各ヒートマップに対してピーク値周辺の座標群をサンプリングし（ステップＳ３００）、サンプリングした座標にヒートマップ信号値を付与する（ステップＳ３０１）。次に、信号値付きハフ投票点群出力部１６３１は、ピーク値周辺の各座標に対して対応するフレーム及び特徴点のベクトル場を用いて座標値を修正する（ステップＳ３０２）。被写体抽出部１６１２は、被写体別骨格長から被写体の人数を取得し、対象とする被写体を１つずつ選択し、骨格モデルフィッティング部１６は被写体毎にステップＳ３０４からステップＳ３１２を繰り返す（ステップＳ３０３）。続いて、フレーム抽出部１６１３は、ハフ投票点群から総フレーム数を取得し、対象とするフレームを１つずつ選択し、骨格モデルフィッティング部１６はフレーム毎にステップＳ３０５からステップＳ３１２を繰り返す（ステップＳ３０４）。次に３次元姿勢推定部１６１４は、被写体骨格長、フレーム被写体特徴点３次元座標群、及び過去被写体３次元姿勢群から対象フレームにおける推定３次元姿勢を生成する（ステップＳ３０５）。

[0051] 対象フレーム３次元姿勢高精度推定部１６３２は、被写体骨格長に設定し

た骨格モデルの関節回転角を推定３次元姿勢に設定し（ステップＳ３０６）、特徴点に対応する全ての関節から１つずつ選択して関節毎にステップＳ３０８からステップＳ３１０を繰り返す（ステップＳ３０７）。続いて、対象フレーム３次元姿勢高精度推定部１６３２は対象関節を各回転軸プラスマイナス方向に各微小角度回転させた際の子関節（子特徴点）の存在確率を式（３）にてそれぞれ計算する（ステップＳ３０８）。なお、ステップＳ３０８において対象フレーム３次元姿勢高精度推定部１６３２は、対象関節の回転角自由度や回転範囲の制限を超えて回転する方向には変位しないよう制約を設けてもよい。次に、対象フレーム３次元姿勢高精度推定部１６３２はステップＳ３０８で計算した各回転軸の存在確率の勾配から、子関節（子特徴点）の存在確率が高くなる方向に対象関節の回転角を修正する（ステップＳ３０９）。対象フレーム３次元姿勢高精度推定部１６３２は、修正した対象関節の回転角を用いて子関節（子特徴点）の存在確率を式（３）にて算出する（ステップＳ３１０）。対象フレーム３次元姿勢高精度推定部１６３２は全ての特徴点についての算出される存在確率の総和の変化量に基づいて、３次元姿勢を確定するか否かを判定する（ステップＳ３１１）。存在確率の総和が所定の量を超える場合（ステップＳ３１１－ＹＥＳ）、対象フレーム３次元姿勢高精度推定部１６３２は直近に定められた３次元姿勢を対象フレームの３次元姿勢として決定し、３次元姿勢記憶部１６１６に被写体骨格長及びフレーム被写体３次元姿勢を保存する（ステップＳ３１２）。存在確率の和が所定の値以下である場合（ステップＳ３１１－ＮＯ）、対象フレーム３次元姿勢高精度推定部１６３２は、ステップＳ３０７に処理を戻し、存在確率の総和が増加するように、骨格モデルの関節の角度を調整しながら、関節の回転角を補正する。

[0052] 被写体出力部１６１７は、３次元姿勢記憶部１６１６から被写体別に被写体骨格長及び過去被写体３次元姿勢群を取得し、被写体骨格長及び被写体別３次元姿勢群を出力する（ステップＳ１１２）。

[0053] 《作用・効果》

このように、第３の実施形態によれば、３次元姿勢推定装置１はヒートマップ信号値に応じた加重和を用いて存在確率を算出することで、確からしさが低いハフ投票点群の影響を軽減することができる。これにより、３次元姿勢推定装置１はより高い精度で被写体の３次元姿勢を推定することができる。

[0054] 本明細書における第１から第３の実施形態においては、関節の回転角についてフィッティングする例の解説を行ったが、root関節の３軸座標も同様に微小変化させた際の勾配を求め、勾配法を用いてフィッティングすることができる。

[0055] 〈他の実施形態〉

以上、図面を参照してこの発明の一実施形態について詳しく説明してきたが、具体的な構成は上述のものに限られることはなく、この発明の要旨を逸脱しない範囲内において様々な設計変更等をすることが可能である。

[0056] 例えば、上述した実施形態においては、特徴点推定装置を、被写体の姿勢を推定する３次元姿勢推定装置１として用いるが、他の実施形態においてはこれに限られない。例えば、他の実施形態においては、特徴点推定装置が被写体の一部の特徴点のみを推定し、姿勢の推定までを行わないものであってもよい。

[0057] 姿勢推定装置は、バスで接続されたＣＰＵ（Central Processing Unit）やメモリや補助記憶装置などを備え、プログラムを実行することによって画像取得部、ヒートマップ生成部、特徴点推定部、骨格推定部、出力部を備える装置として機能する。なお、姿勢推定装置の各機能の全て又は一部は、ＡＳＩＣ（Application Specific Integrated Circuit）やＰＬＤ（Programmable Logic Device）やＦＰＧＡ（Field Programmable Gate Array）等のハードウェアを用いて実現されてもよい。プログラムは、コンピュータ読み取り可能な記録媒体に記録されてもよい。コンピュータ読み取り可能な記録媒体とは、例えばフレキシブルディスク、光磁気ディスク、ＲＯＭ、ＣＤ－ＲＯＭ等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶

装置である。プログラムは、電気通信回線を介して送信されてもよい。

符号の説明

[0058] １　３次元姿勢推定装置

２　モーションキャプチャシステム

１０　画像取得部

１１　２次元姿勢推定部

１２　２次元姿勢推定モデル記憶部

１３　カメラパラメータ取得部

１４　被写体追跡部

１５　骨格長推定部

１６　骨格モデルフィッティング部

１６１０　ピーク座標出力部

１６１１　ハフ投票点群出力部

１６１２　被写体抽出部

１６１３　フレーム抽出部

１６１４　３次元姿勢推定部

１６１５　対象フレーム３次元姿勢推定部

１６１６　３次元姿勢記憶部

１６１７　被写体出力部

１６２０　ラベル付きピーク座標出力部

１６２１　ラベル付きハフ投票点群出力部

１６２２　対象フレーム３次元姿勢高速推定部

１６３０　信号値付きピーク座標出力部

１６３１　信号値付きハフ投票点群出力部

１６３２　対象フレーム３次元姿勢高精度推定部

請求の範囲

[請求項1] 異なる地点に位置する複数の撮影装置により同一の被写体が撮影された画像群を入力とし、前記被写体に対し予め学習された２次元姿勢推定モデルを用いて、被写体別に、複数の特徴点の２次元位置を示す２次元座標群と、特徴点毎の存在確率を表すヒートマップ群と、ヒートマップのピーク値周辺の座標を対象特徴点方向に補正するベクトル場群とを生成する２次元姿勢推定部と、

前記撮影装置のカメラパラメータと前記２次元座標群とを用いて前記被写体別に、前記複数の特徴点の３次元位置を示す特徴点３次元座標群を生成する被写体追跡部と、

前記被写体別の特徴点３次元座標群を用いて被写体別の骨格長を推定する骨格長推定部と、

前記ヒートマップ群、前記ベクトル場群、前記撮影装置のカメラパラメータ、前記特徴点３次元座標群、及び前記骨格長を用いて被写体別の３次元姿勢を推定する骨格モデルフィッティング部と

を備える３次元姿勢推定装置。

[請求項2] 前記骨格モデルフィッティング部は、前記ヒートマップ群及び前記ベクトル場群に基づくハフ投票により前記特徴点３次元座標群を補正する

請求項１に記載の３次元姿勢推定装置。

[請求項3] 前記骨格モデルフィッティング部は、前記被写体を関節と骨の組み合わせで表す骨格モデルの前記関節の回転角を推定することで、前記被写体別の３次元姿勢を推定する

請求項１又は請求項２に記載の３次元姿勢推定装置。

[請求項4] 前記骨格モデルフィッティング部は、

前記特徴点毎に生成した前記ヒートマップの信号値がピーク値を示す座標群を前記ベクトル場群によって補正したハフ投票点群を生成し、

前記被写体を関節と骨の組み合わせで表す骨格モデルの各関節を各撮影装置の画像上に投影したスクリーン座標と、各撮影装置の対象関節と対応する特徴点を指す前記ハフ投票点群とスクリーン座標との距離に応じて特徴点の存在確率を計算し、

全ての特徴点における存在確率が最も高くなるように、前記骨格モデルの各関節における関節回転角を推定することで、被写体の３次元姿勢を推定することを特徴とする

請求項１から請求項３の何れか１項に記載の３次元姿勢推定装置。

[請求項5] 前記骨格モデルフィッティング部は、

特徴点毎に生成した前記ヒートマップの信号値がピーク値を示す座標群に対してラベルを付与し、

前記特徴点の存在確率を計算する際に、子関節の３次元座標をスクリーン上に投影したスクリーン座標が属するラベルと同一のラベルが付与されたハフ投票点群のみを用いることを特徴とする

請求項４に記載の３次元姿勢推定装置。

[請求項6] 前記骨格モデルフィッティング部は、

特徴点毎に生成した前記ヒートマップの信号値がピーク値を示す座標群に対して前記ヒートマップの信号値を付与し、

前記特徴点の存在確率の導出のため、各ハフ投票点との距離を計算する際に、前記ヒートマップの信号値を乗じることを特徴とする

請求項４に記載の３次元姿勢推定装置。

[請求項7] 異なる地点に位置する複数の撮影装置により同一の被写体が撮影された画像群を入力とし、前記被写体に対し予め学習された２次元姿勢推定モデルを用いて、被写体別に、複数の特徴点の２次元位置を示す２次元座標群と、特徴点毎の存在確率を表すヒートマップ群と、ヒートマップのピーク値周辺の座標を対象特徴点方向に補正するベクトル場群とを生成する２次元姿勢推定ステップと、

前記撮影装置のカメラパラメータと前記２次元座標群とを用いて前

記被写体別に、前記複数の特徴点の３次元位置を示す特徴点３次元座標群を生成する被写体追跡ステップと、

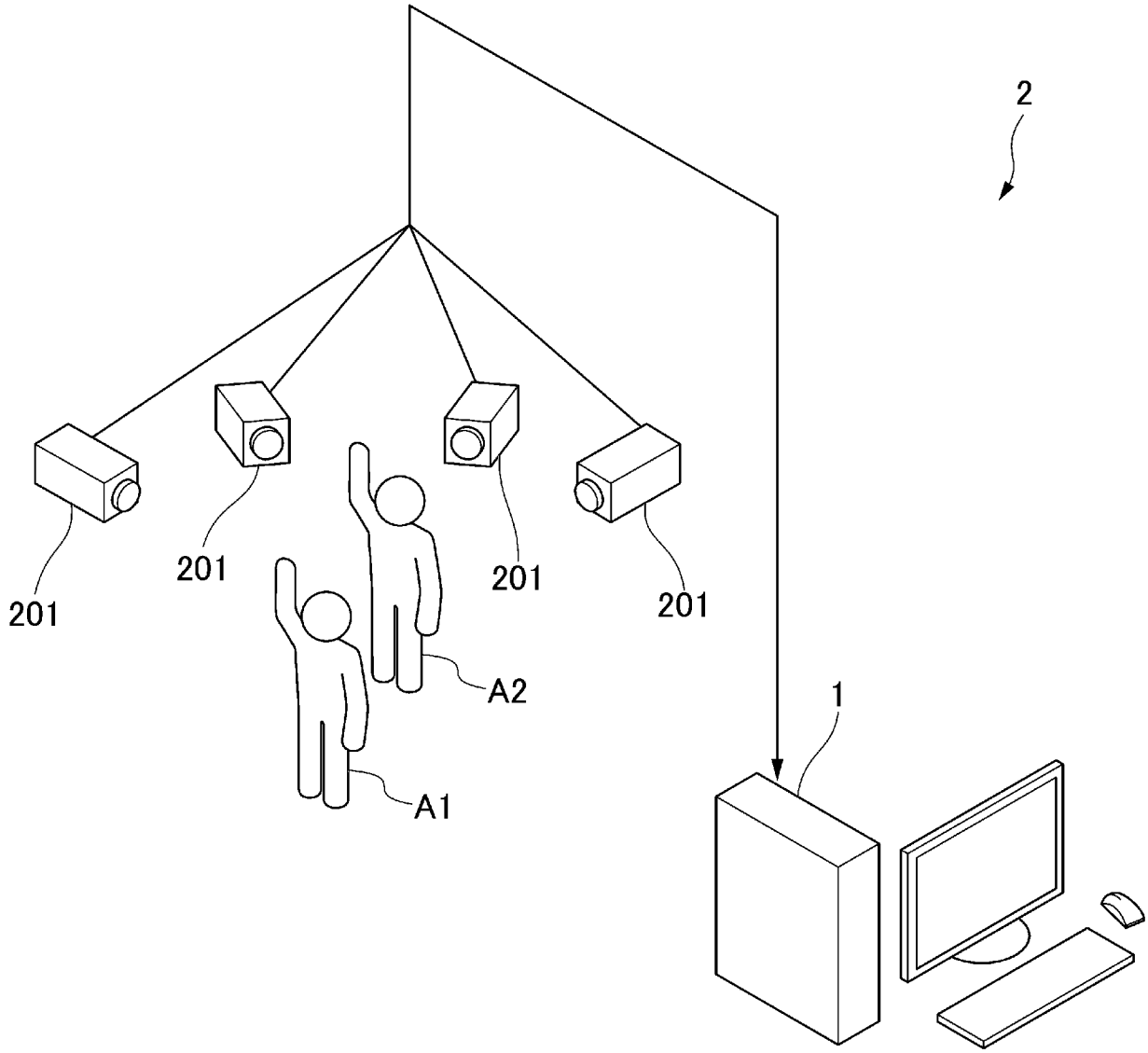
前記被写体別の特徴点３次元座標群を用いて被写体別の骨格長を推定する骨格長推定ステップと、

前記ヒートマップ群、前記ベクトル場群、前記撮影装置のカメラパラメータ、前記特徴点３次元座標群、及び前記骨格長を用いて被写体別の３次元姿勢を推定する骨格モデルフィッティングステップと

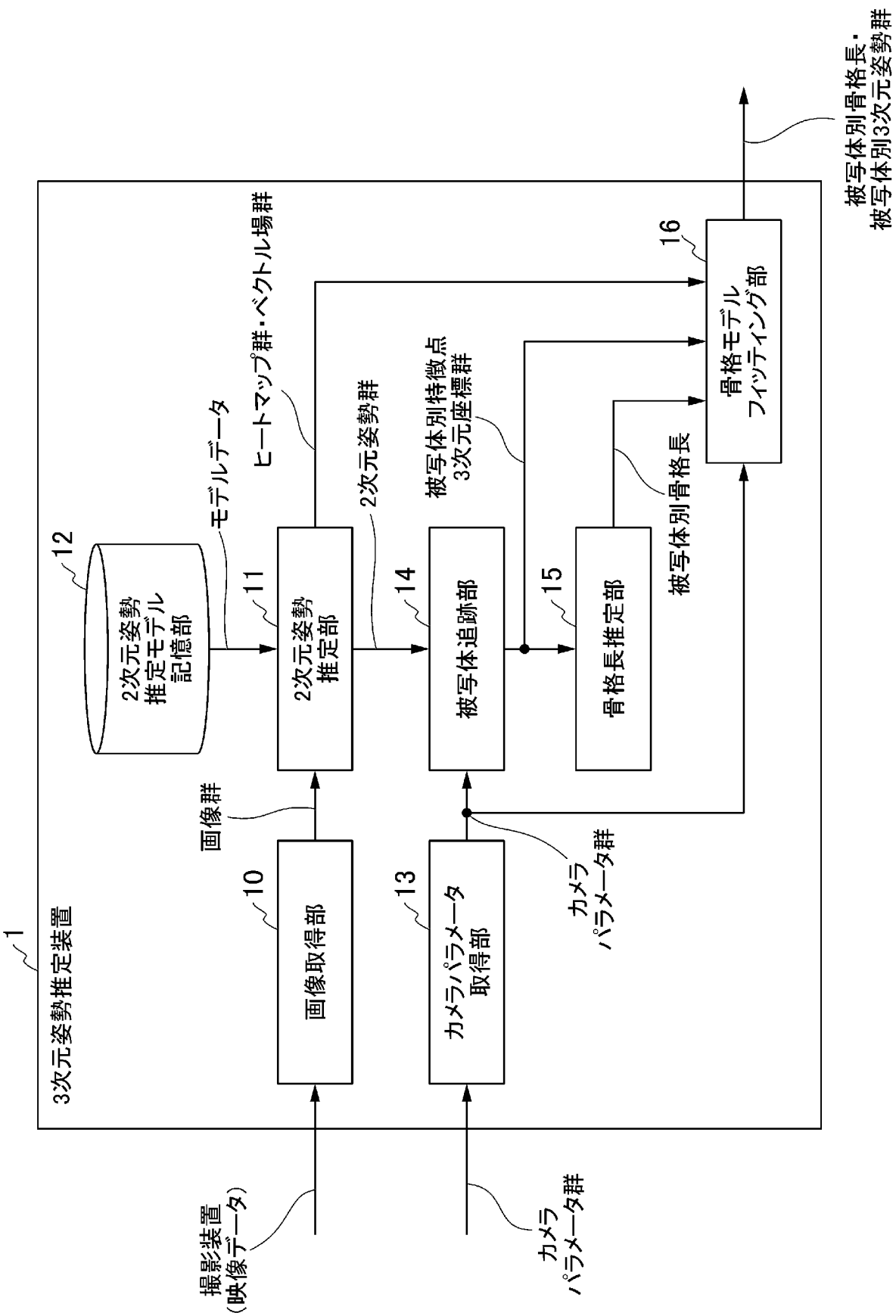
を有する３次元姿勢推定方法。

[請求項8] コンピュータを、請求項１から請求項６の何れか１項に記載の３次元姿勢推定装置として機能させるためのプログラム。

[図1]

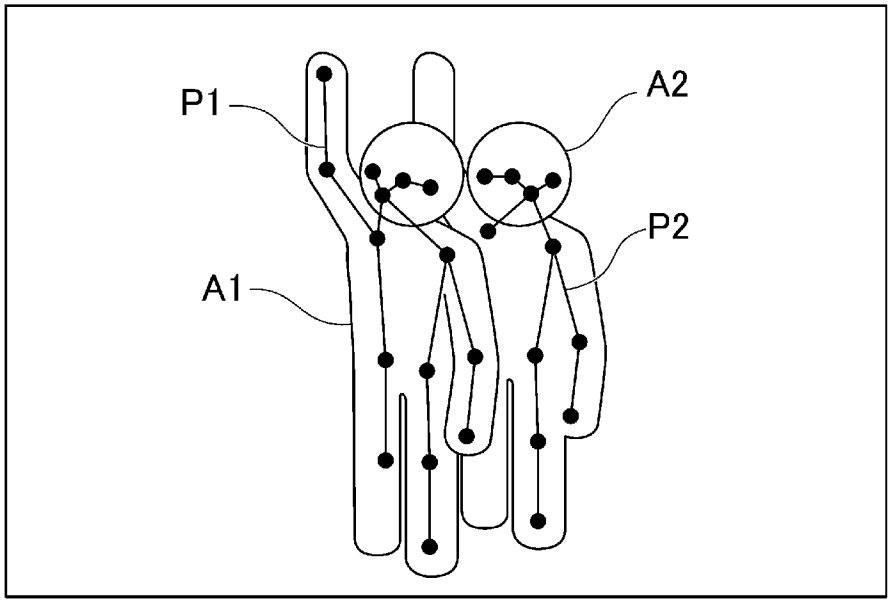


[図2]


1
3次元姿勢推定装置
撮影装置
(映像データ)
カメラ
パラメータ群
10
画像取得部
画像群
13
カメラパラメータ
取得部
カメラ
パラメータ群
12
2次元姿勢
推定モデル
記憶部
モデルデータ
11
2次元姿勢
推定部
ヒートマップ群・ベクトル場群
2次元姿勢群
14
被写体追跡部
被写体別特徴点
3次元座標群
15
骨格長推定部
被写体別骨格長
16
骨格モデル
フィッティング部
被写体別骨格長,
被写体別3次元姿勢群

[図3A]

撮影装置201によって取得された画像

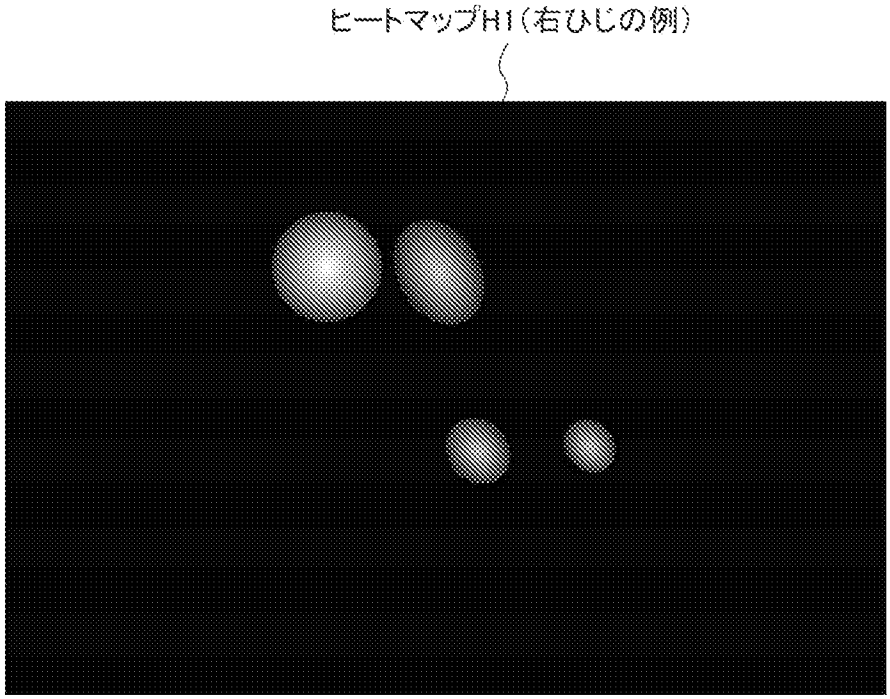


[図3B]

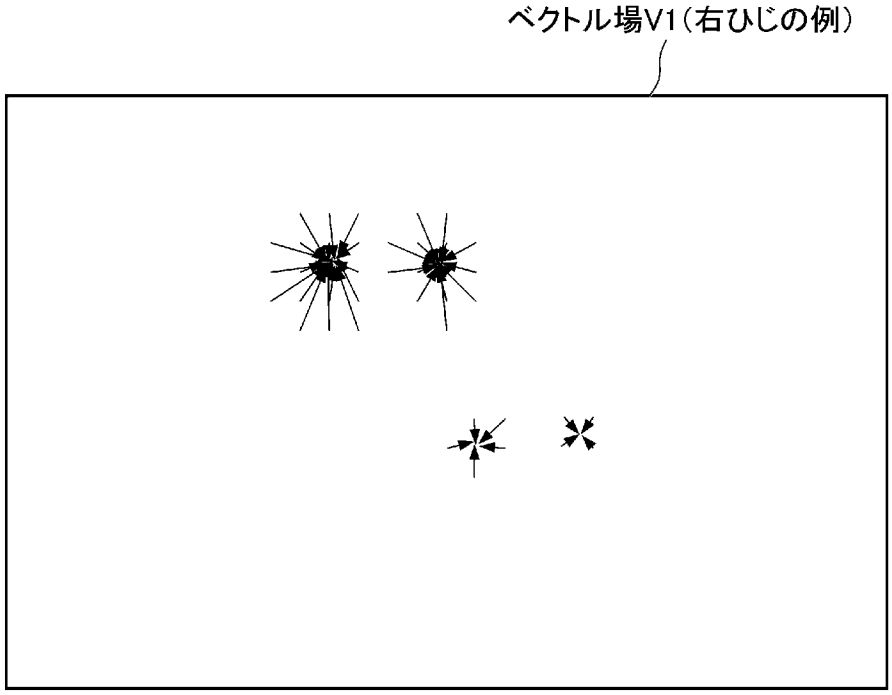
2次元姿勢の例(画像あたり1つのテーブルデータ)

撮影装置識別子	被写体	特徴点名	X座標	Y座標
201-1	A1	鼻	272.3	135.4
		左目	287.1	125.0
		右目	266.3	119.8
		左耳	308.7	130.4
		左肩	321.1	179.9
		右肩	269.6	168.8
		…	…	…
	A2	鼻	382.4	136.0
		左目	398.3	126.2
		右目	368.9	123.4
		右耳	348.7	123.1
		…	…	…

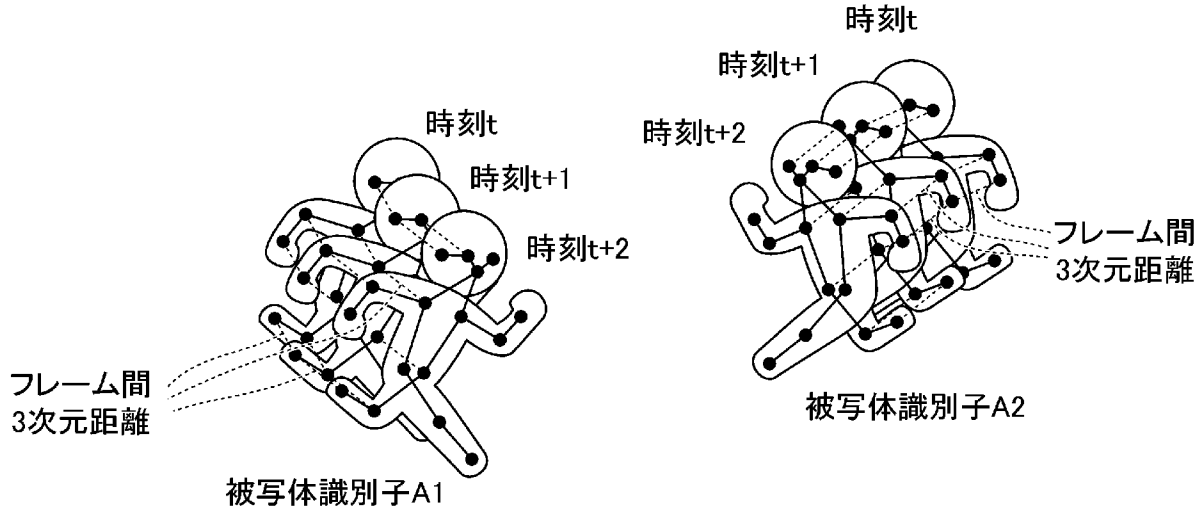
[図4]



[図5]



[図6A]

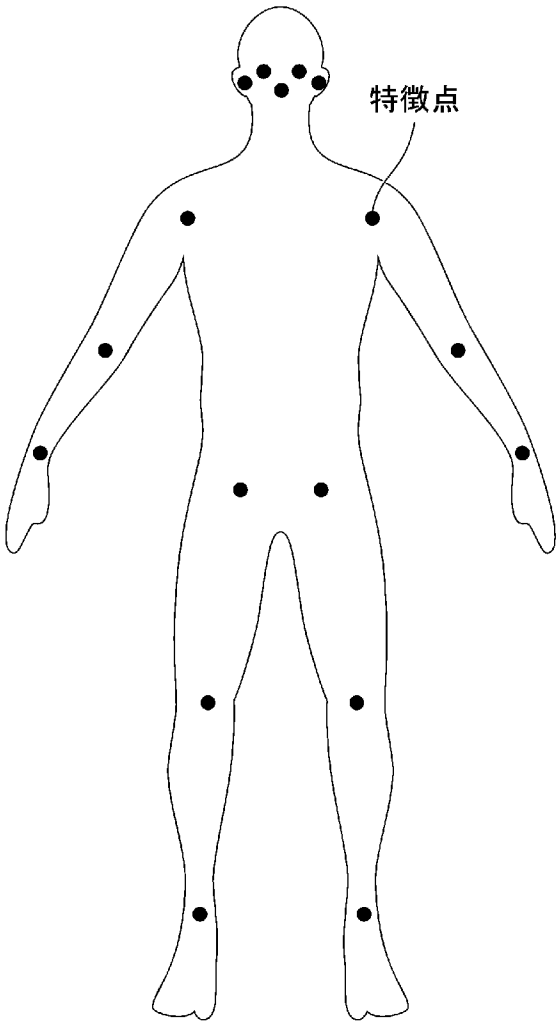


[図6B]

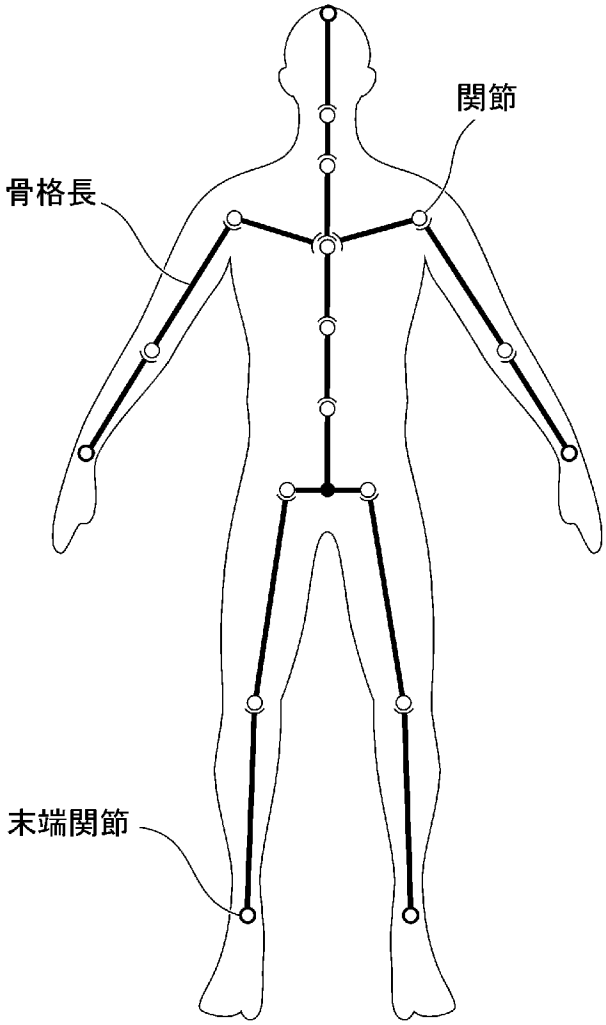
被写体識別子	フレーム（時刻）	特徴点名	X座標	Y座標	Z座標
A1	0	鼻	5.784	1.351	-2.441
		左目	5.710	1.397	-2.267
		…	…	…	…
	1	鼻	5.702	1.274	-2.273
		左目	5.683	1.312	-2.286
		…	…	…	…
	…	…	…	…	…
A2	0	鼻	0.980	1.547	-4.201
		左目	0.949	1.599	-4.227
		…	…	…	…
	1	鼻	1.052	1.618	-4.029
		左目	1.028	1.661	-4.051
		…	…	…	…
	…	…	…	…	…

[図7]

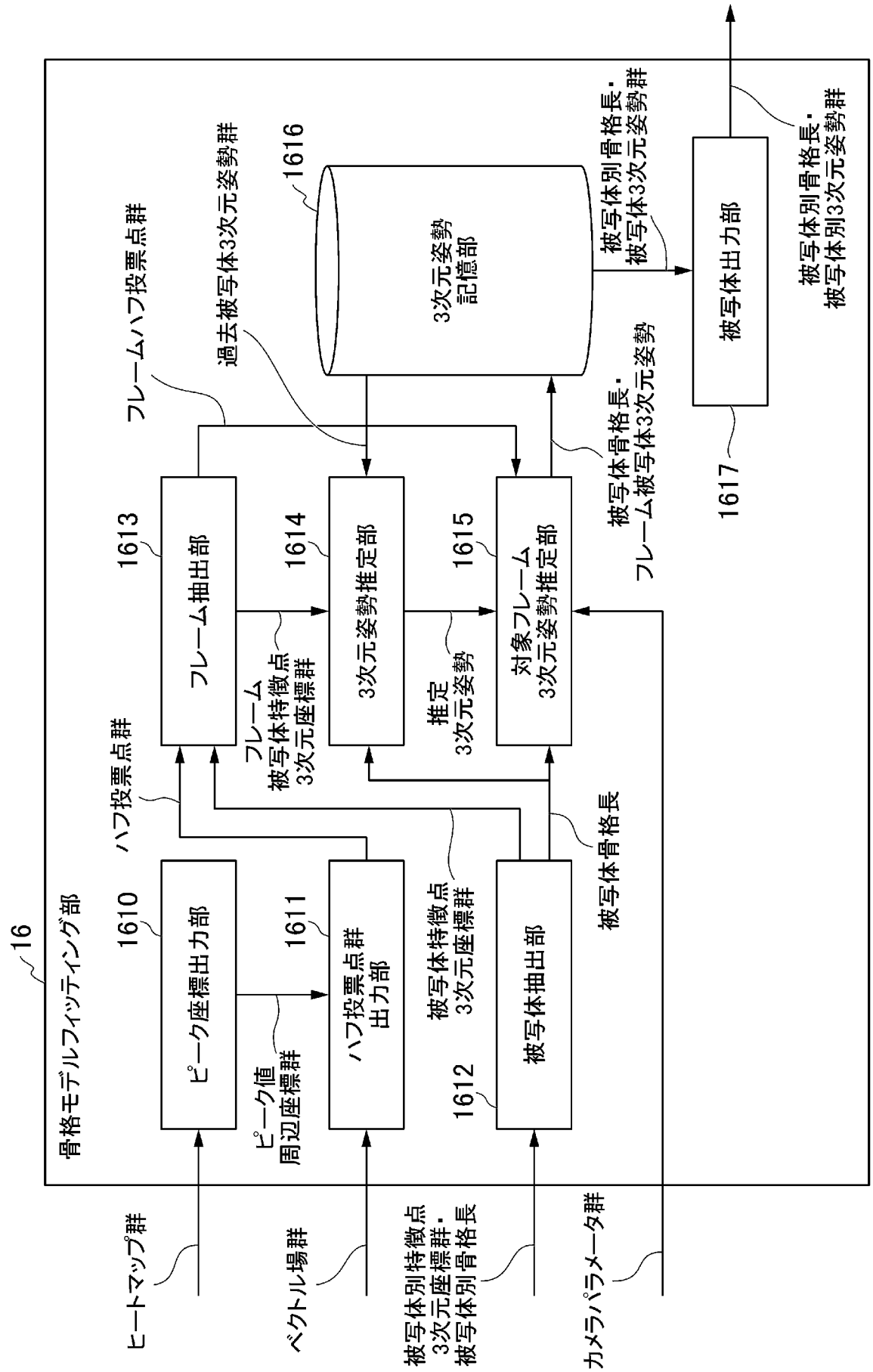
特徴点の定義位置



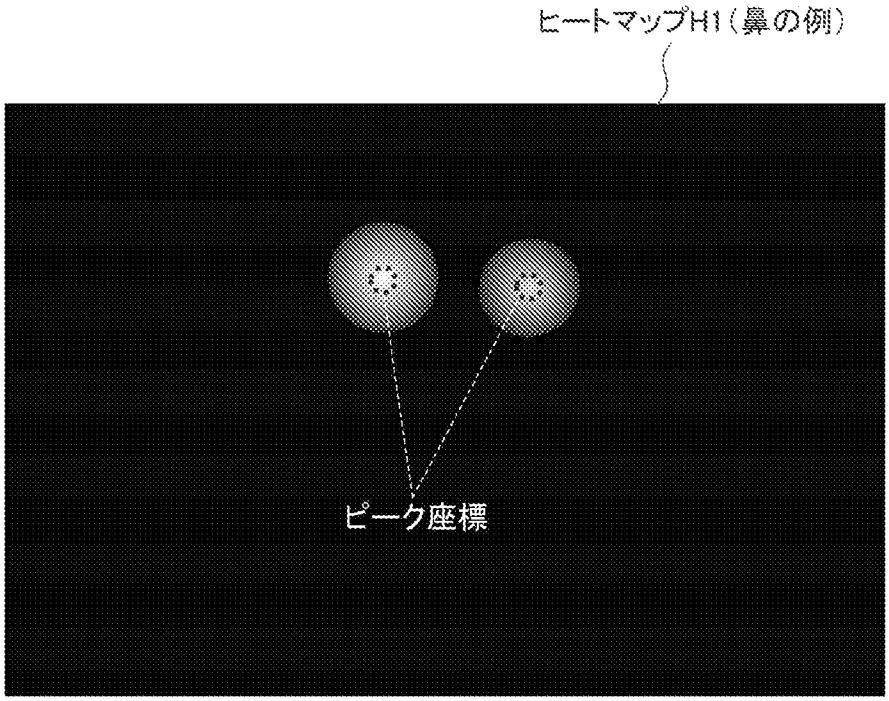
骨格モデルの関節と骨格長の定義位置



[図8]


16
骨格モデルフィッティング部
1610
ピーク座標出力部
1611
ハフ投票点群出力部
1612
被写体抽出部
1613
フレーム抽出部
1614
3次元姿勢推定部
1615
対象フレーム3次元姿勢推定部
1616
3次元姿勢記憶部
1617
被写体出力部
ヒートマップ群
ベクトル場群
被写体別特徴点3次元座標群・被写体別骨格長
カメラパラメータ群
ピーク値周辺座標群
ハフ投票点群
被写体特徴点3次元座標群
被写体骨格長
フレーム被写体特徴点3次元座標群
推定3次元姿勢
フレームハフ投票点群
過去被写体3次元姿勢群
被写体骨格長・フレーム被写体3次元姿勢
被写体別骨格長・被写体3次元姿勢群
被写体別骨格長・被写体別3次元姿勢群

[図9A]



[図9B]

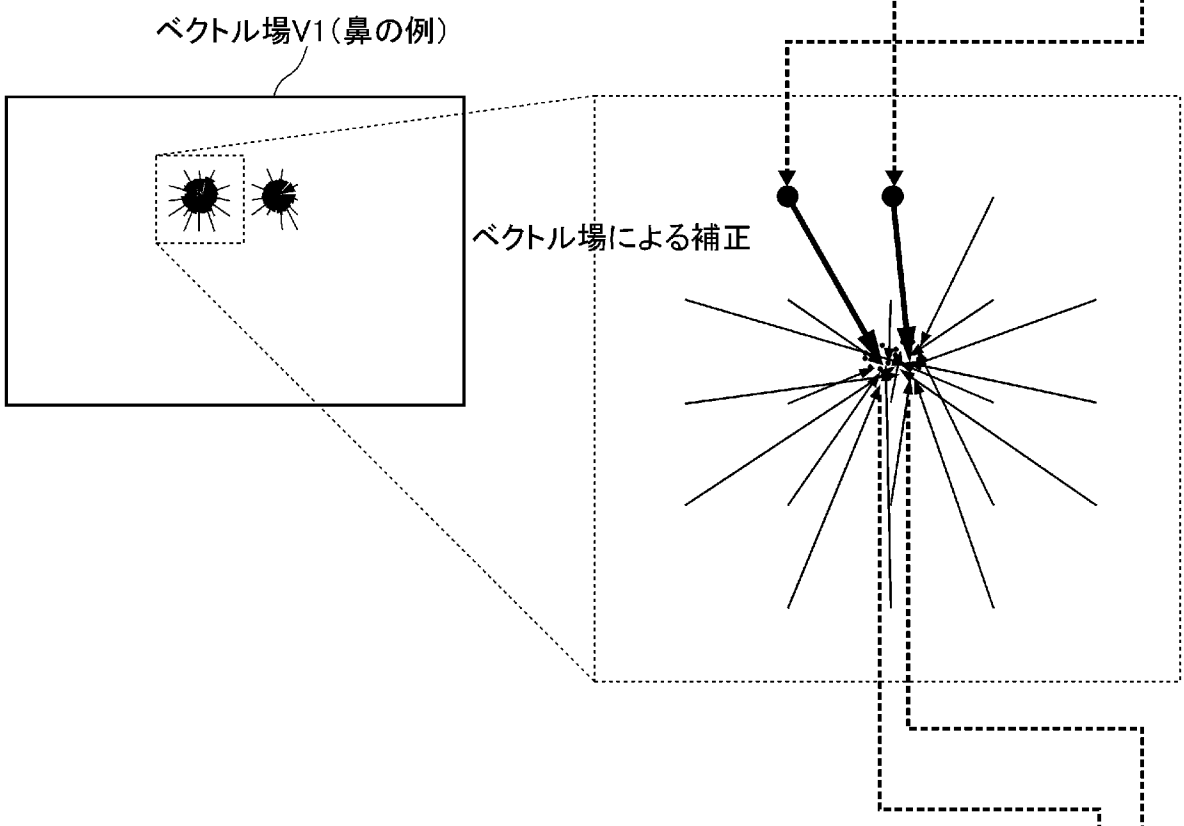
ピーク値周辺座標群の例

撮影装置識別子	フレーム（時刻）	特徴点名	X座標	Y座標
201-1	0	鼻	271	134
			272	134
			273	134
			270	135
			271	135
			272	135
			273	135
			…	…
		左目	286	124
			287	124
			288	124
			284	125
			…	…
		…	…	…
	…	…	…	…
…	…	…	…	…

[図10]

入力(ピーク値周辺座標群)の例

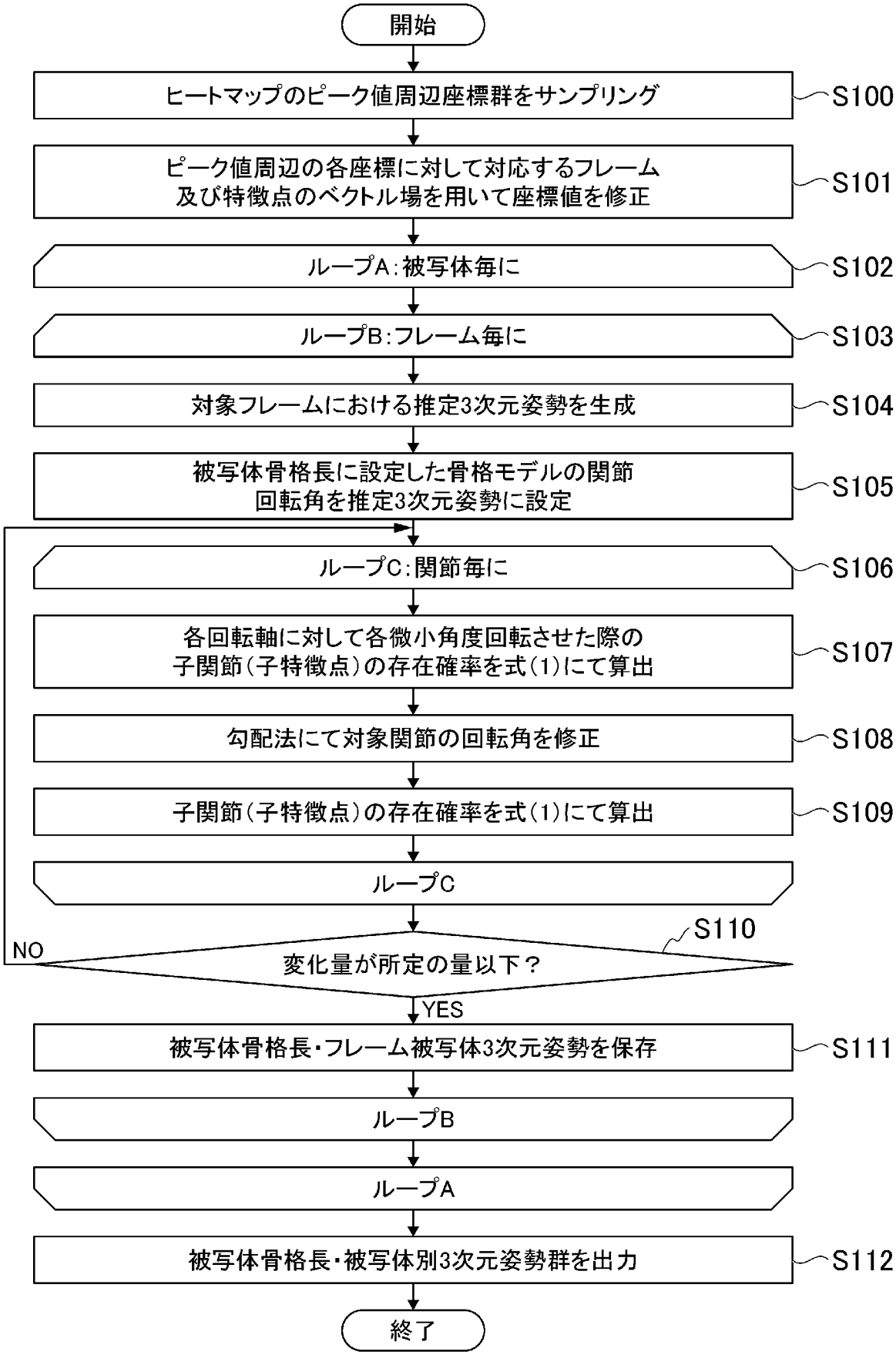
撮影装置識別子	フレーム(時刻)	特徴点名	X座標	Y座標
201-1	0	鼻	271	134
			272	134
			…	…
		…	…	…
	…	…	…	…
…	…	…	…	…



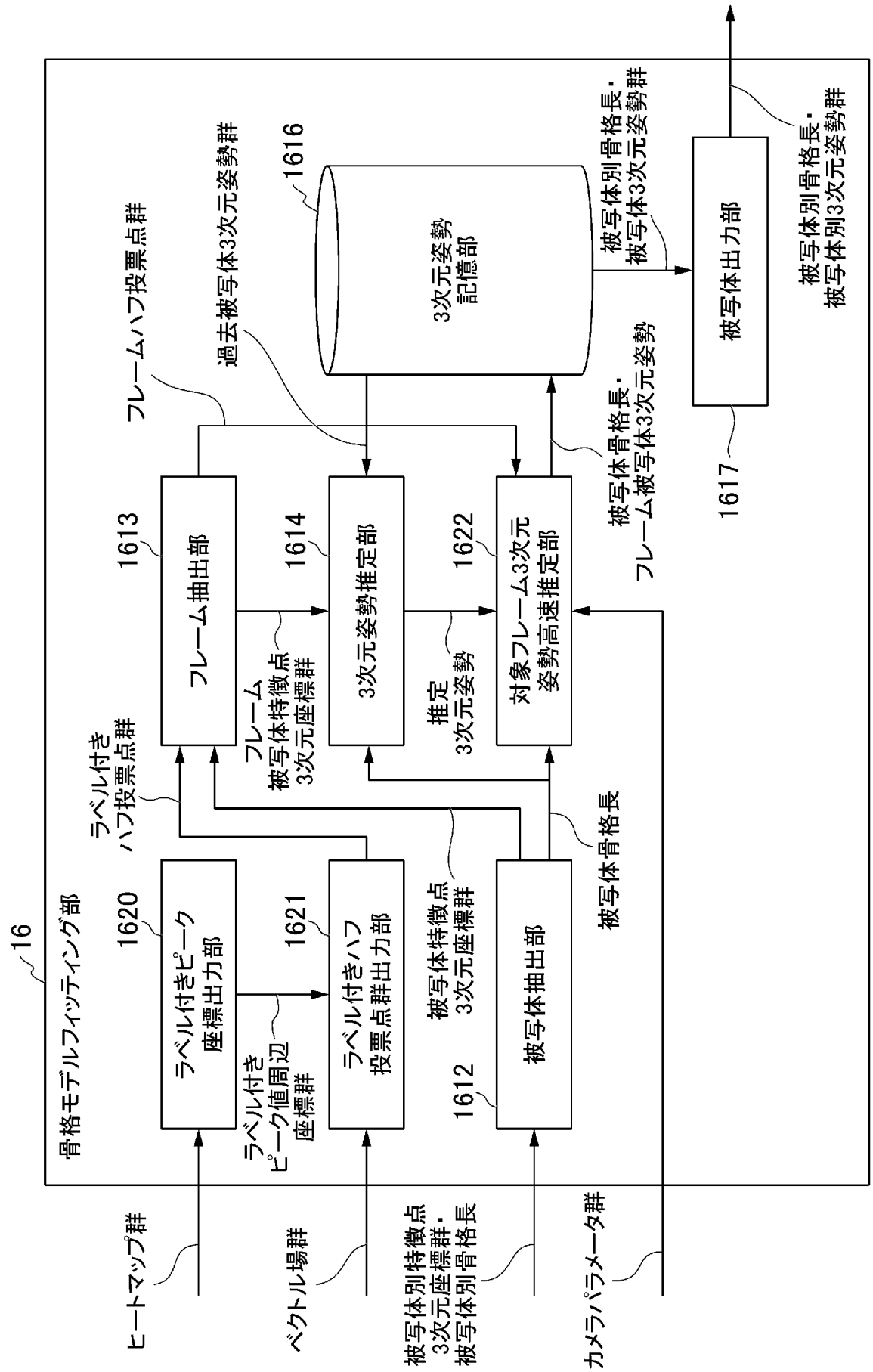
ハフ投票点群の例

撮影装置識別子	フレーム(時刻)	特徴点名	X座標	Y座標
201-1	0	鼻	271.9	135.6
			272.2	135.5
			…	…
		…	…	…
	…	…	…	…
…	…	…	…	…

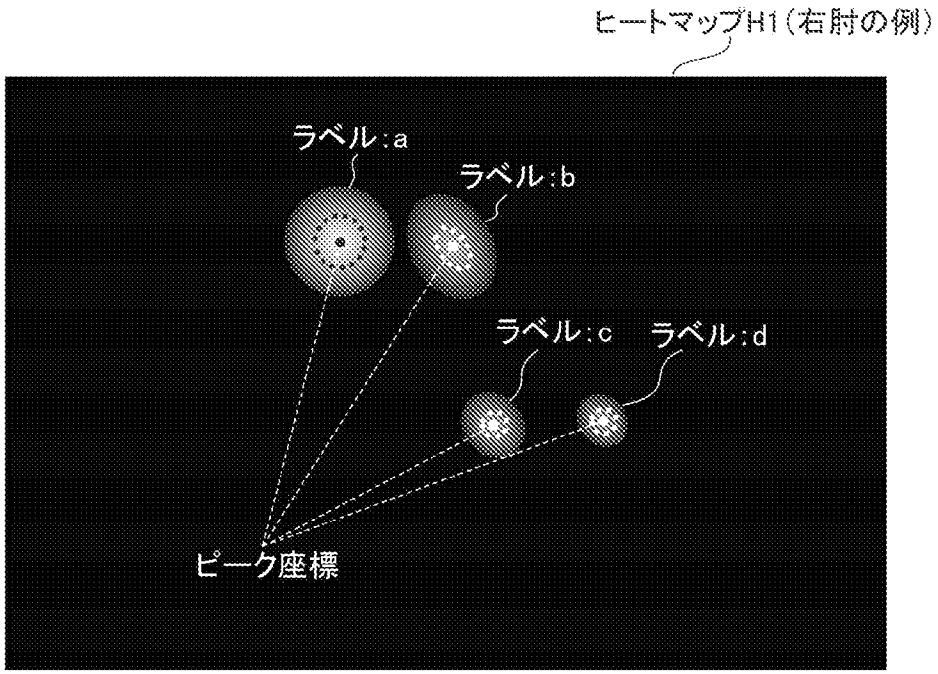
[図11]


開始
ヒートマップのピーク値周辺座標群をサンプリング
S100
ピーク値周辺の各座標に対して対応するフレーム
及び特徴点のベクトル場を用いて座標値を修正
S101
ループA:被写体毎に
S102
ループB:フレーム毎に
S103
対象フレームにおける推定3次元姿勢を生成
S104
被写体骨格長に設定した骨格モデルの関節
回転角を推定3次元姿勢に設定
S105
ループC:関節毎に
S106
各回転軸に対して各微小角度回転させた際の
子関節(子特徴点)の存在確率を式(1)にて算出
S107
勾配法にて対象関節の回転角を修正
S108
子関節(子特徴点)の存在確率を式(1)にて算出
S109
ループC
S110
NO
変化量が所定の量以下?
YES
被写体骨格長・フレーム被写体3次元姿勢を保存
S111
ループB
ループA
被写体骨格長・被写体別3次元姿勢群を出力
S112
終了

[図12]


16
骨格モデルフィッティング部
ヒートマップ群
ベクトル場群
被写体別特徴点
3次元座標群・
被写体別骨格長
カメラパラメータ群
1620
ラベル付き
ピーク座標出力部
1621
ラベル付き
投票点群出力部
1612
被写体抽出部
1613
フレーム抽出部
1614
3次元姿勢推定部
1622
対象フレーム3次元
姿勢高速推定部
1616
3次元姿勢
記憶部
1617
被写体出力部
ラベル付き
ピーク値周辺
座標群
被写体特徴点
3次元座標群
ラベル付き
ハフ投票点群
フレーム
被写体特徴点
3次元座標群
推定
3次元姿勢
被写体骨格長
フレーム被写体3次元姿勢
被写体骨格長・
フレームハフ投票点群
過去被写体3次元姿勢群
被写体別骨格長・
被写体3次元姿勢群
被写体別骨格長・
被写体別3次元姿勢群

[図13A]

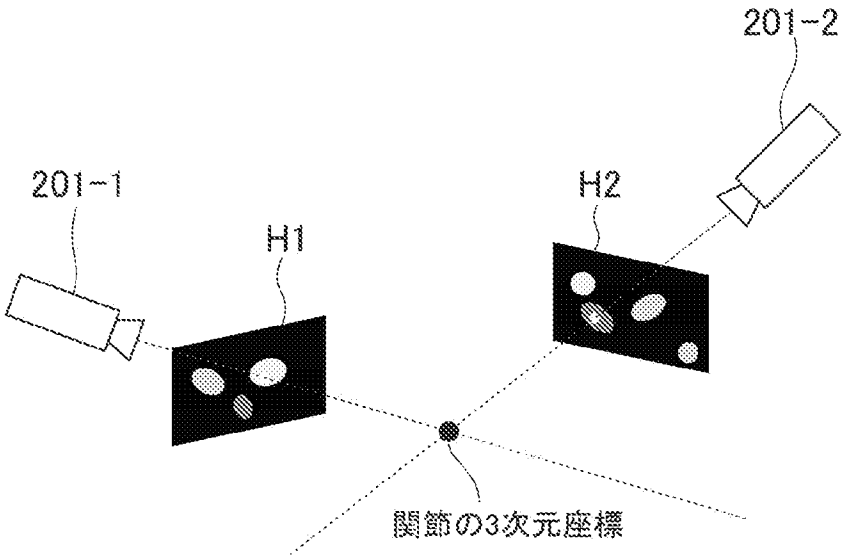


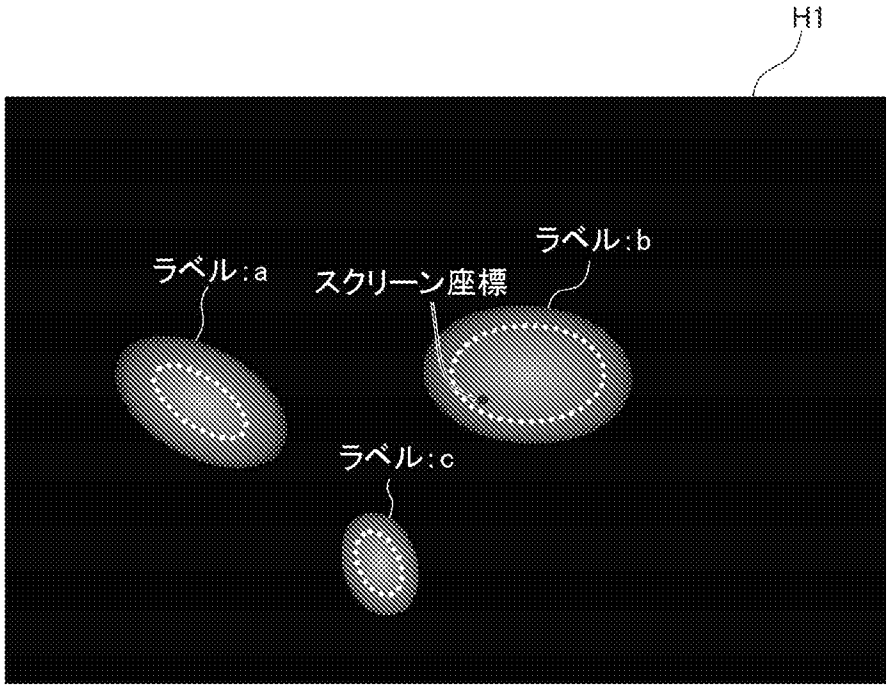
[図13B]

ラベル付きピーク値周辺座標群の例

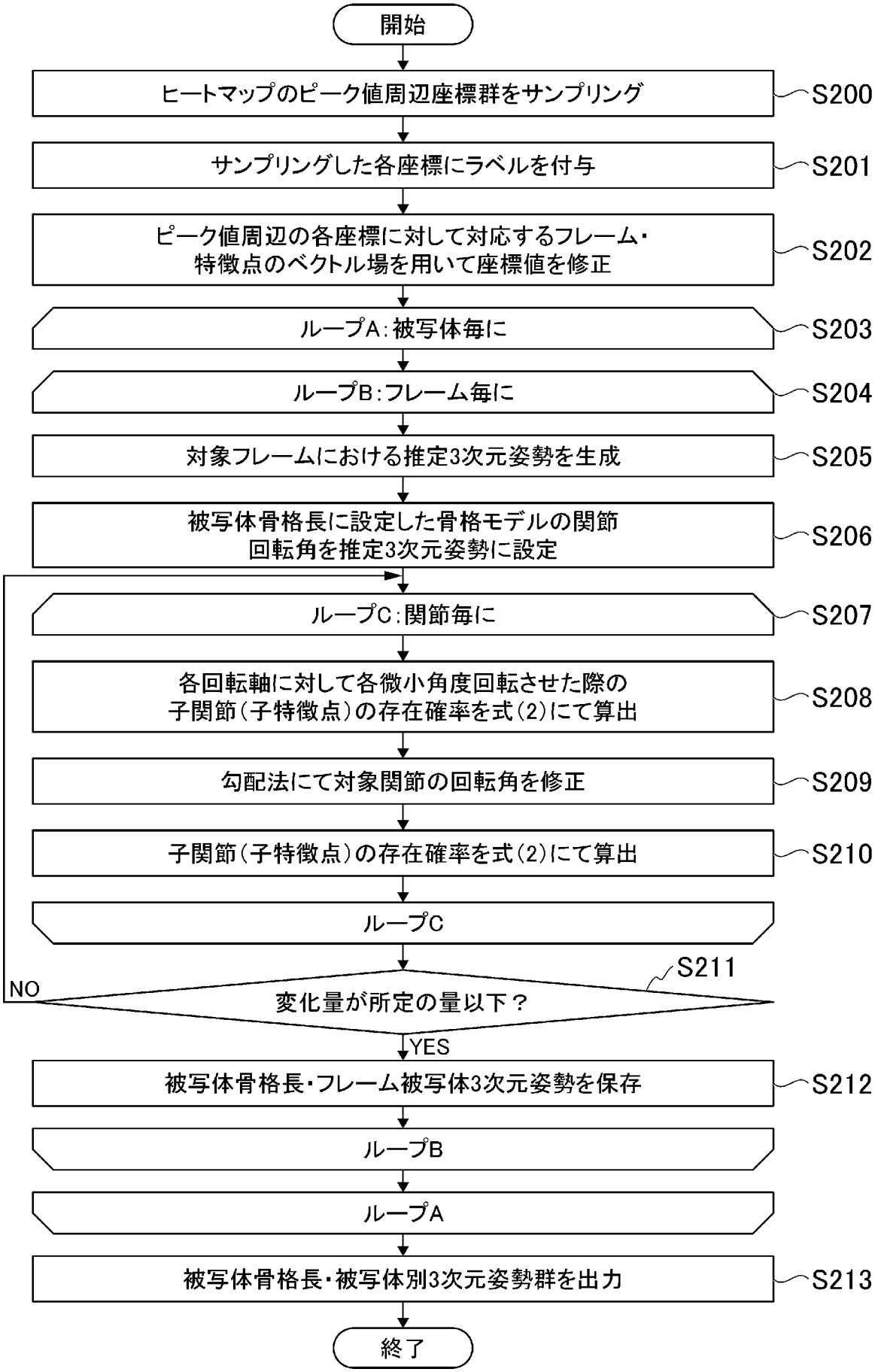
撮影装置識別子	フレーム（時刻）	特徴点名	ラベル	X座標	Y座標
201-1	0	鼻	a	271	134
				272	134
				273	134
				270	135
				271	135
				272	135
				273	135
				…	…
			…	…	…
		左目	a	286	124
				287	124
				288	124
				284	125
				…	…
			…	…	…
		…	…	…	…
	…	…	…	…	…
…	…	…	…	…	…

[図14]

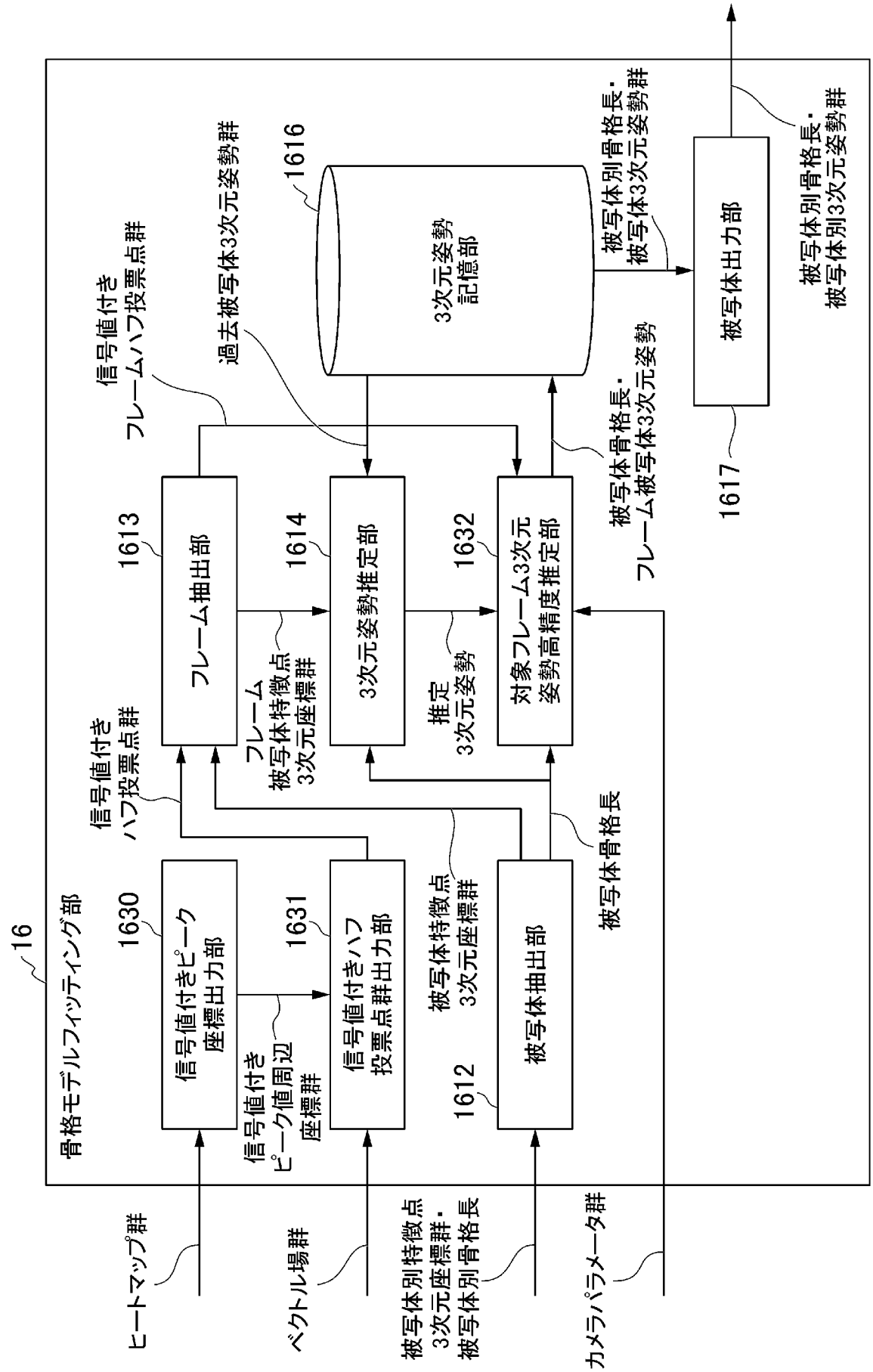




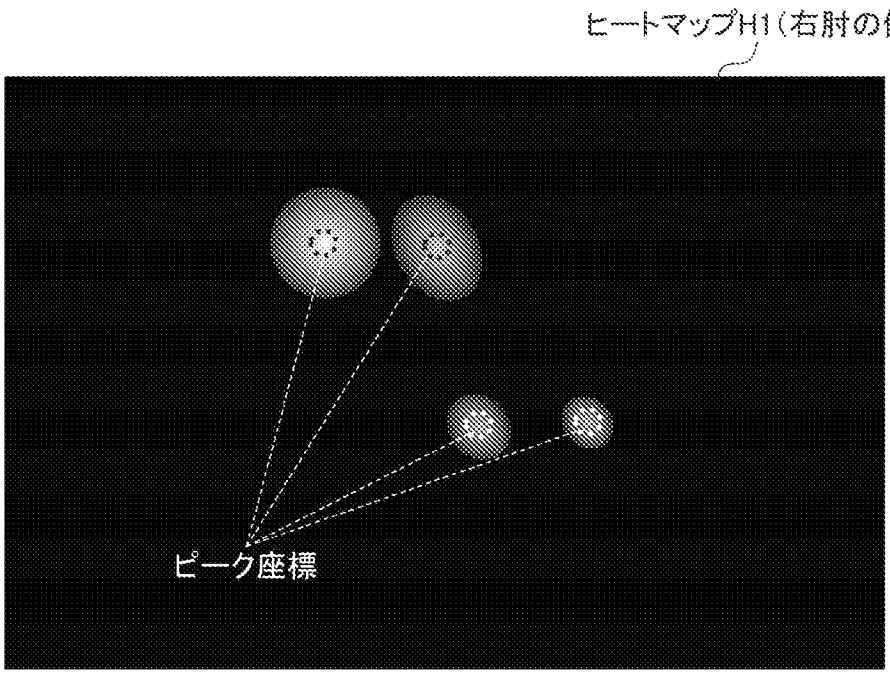
[図15]


開始
ヒートマップのピーク値周辺座標群をサンプリング
S200
サンプリングした各座標にラベルを付与
S201
ピーク値周辺の各座標に対して対応するフレーム・特徴点のベクトル場を用いて座標値を修正
S202
ループA:被写体毎に
S203
ループB:フレーム毎に
S204
対象フレームにおける推定3次元姿勢を生成
S205
被写体骨格長に設定した骨格モデルの関節回転角を推定3次元姿勢に設定
S206
ループC:関節毎に
S207
各回転軸に対して各微小角度回転させた際の子関節(子特徴点)の存在確率を式(2)にて算出
S208
勾配法にて対象関節の回転角を修正
S209
子関節(子特徴点)の存在確率を式(2)にて算出
S210
ループC
変化量が所定の量以下？
S211
NO
YES
被写体骨格長・フレーム被写体3次元姿勢を保存
S212
ループB
ループA
被写体骨格長・被写体別3次元姿勢群を出力
S213
終了

[図16]


16
骨格モデルフィッティング部
1630
信号値付きピーク座標出力部
1631
信号値付きハフ投票点群出力部
1612
被写体抽出部
1613
フレーム抽出部
1614
3次元姿勢推定部
1632
対象フレーム3次元姿勢高精度推定部
1616
3次元姿勢記憶部
1617
被写体出力部
信号値付きハフ投票点群
信号値付きピーク値周辺座標群
被写体特徴点3次元座標群
フレーム被写体特徴点3次元座標群
過去被写体3次元姿勢群
推定3次元姿勢
被写体骨格長
被写体骨格長・フレーム被写体3次元姿勢
被写体別骨格長・被写体3次元姿勢群
被写体別骨格長・被写体別3次元姿勢群
ヒートマップ群
ベクトル場群
被写体別特徴点3次元座標群・被写体別骨格長
カメラパラメータ群

[図17A]

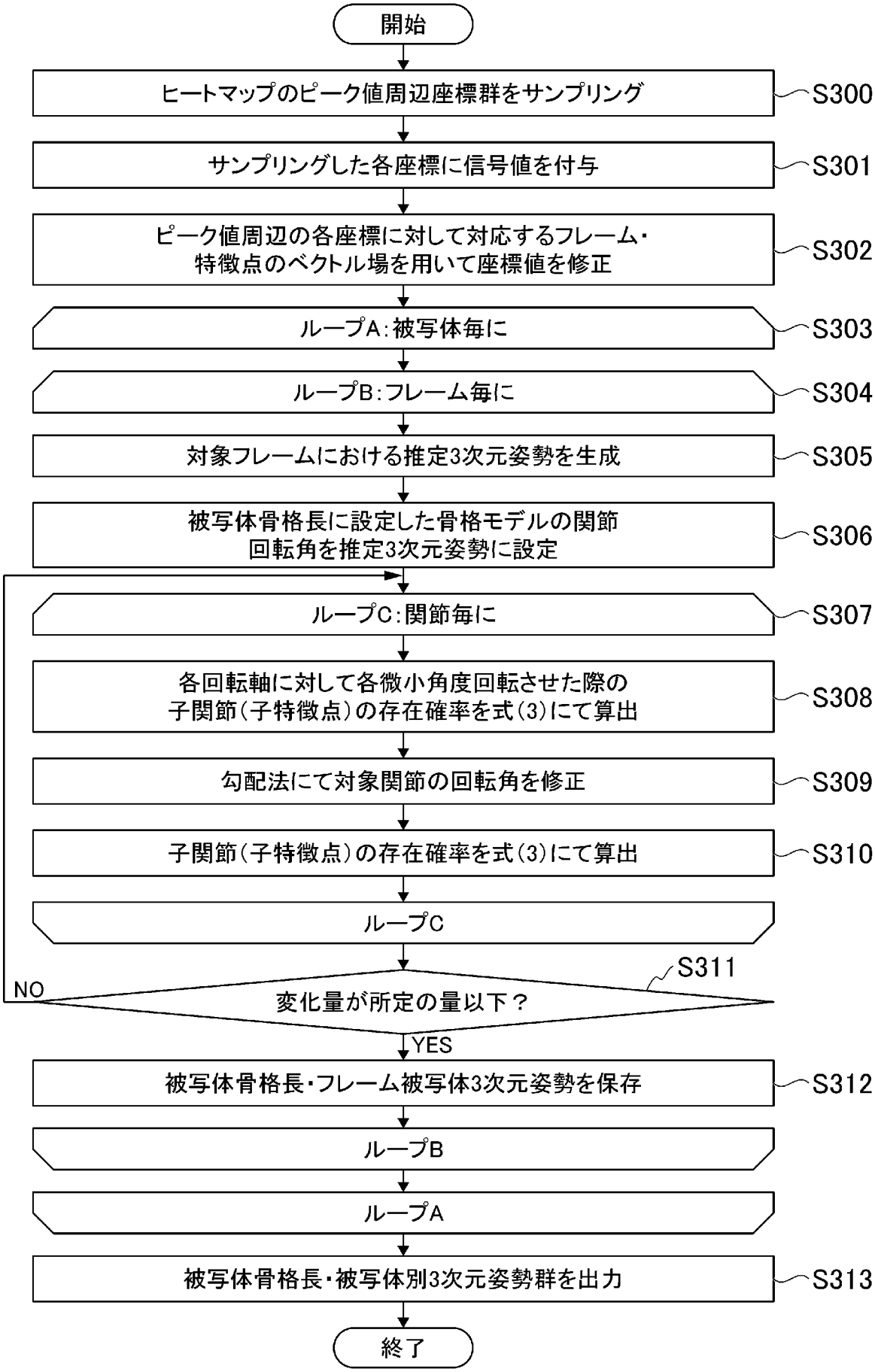


[図17B]

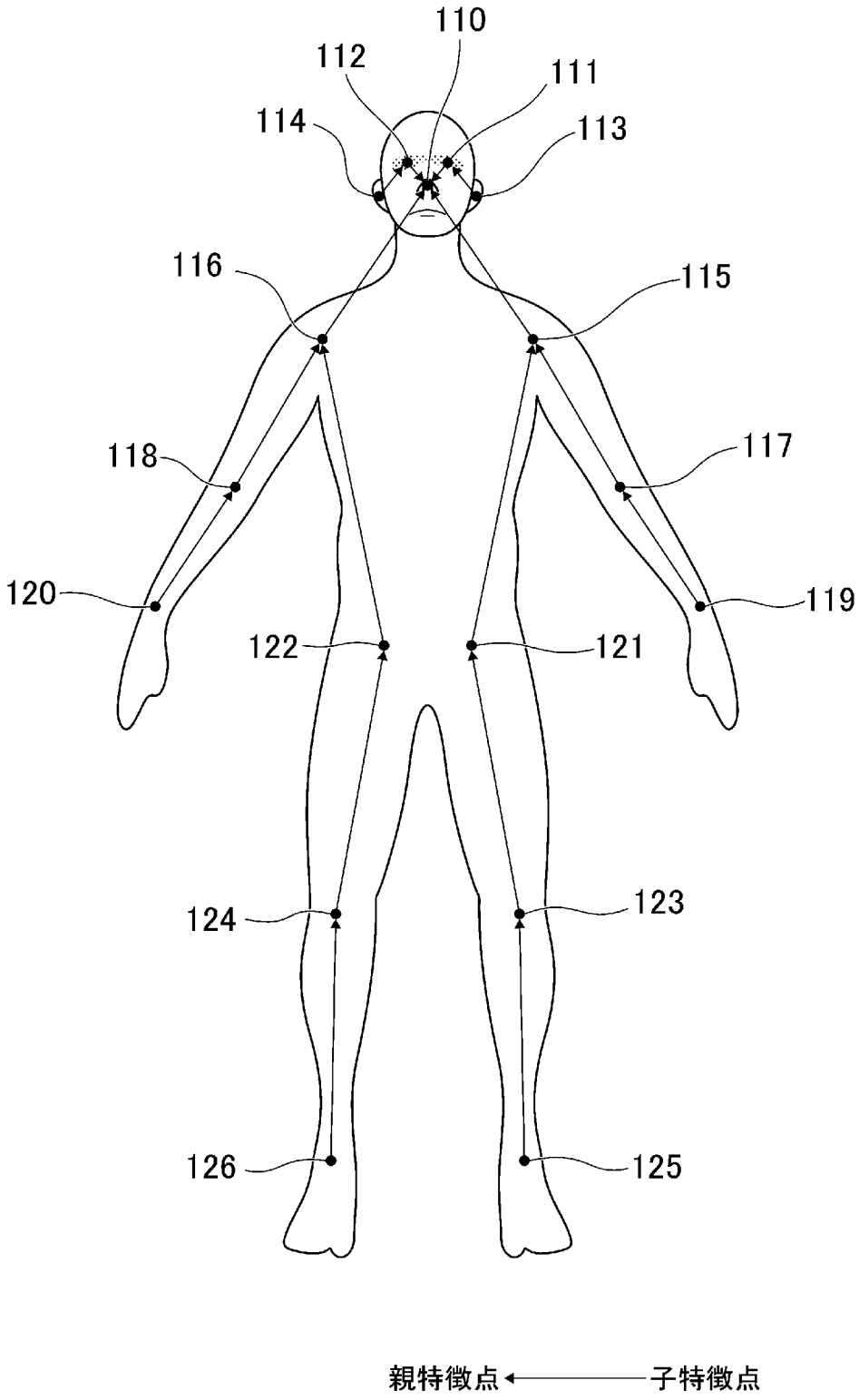
信号値付きピーク値周辺座標群の例

撮影装置識別子	フレーム（時刻）	特徴点名	X座標	Y座標	ヒートマップ信号値
Cam1	0	鼻	271	134	0.751
			272	134	0.769
			273	134	0.742
			270	135	0.750
			271	135	0.793
			272	135	0.842
			273	135	0.799
			…	…	…
		左目	286	124	0.575
			287	124	0.621
			288	124	0.569
			284	125	0.488
			…	…	…
		…	…	…	…
	…	…	…	…	…
…	…	…	…	…	…

[図18]


開始
ヒートマップのピーク値周辺座標群をサンプリング
S300
サンプリングした各座標に信号値を付与
S301
ピーク値周辺の各座標に対して対応するフレーム・特徴点のベクトル場を用いて座標値を修正
S302
ループA:被写体毎に
S303
ループB:フレーム毎に
S304
対象フレームにおける推定3次元姿勢を生成
S305
被写体骨格長に設定した骨格モデルの関節回転角を推定3次元姿勢に設定
S306
ループC:関節毎に
S307
各回転軸に対して各微小角度回転させた際の子関節(子特徴点)の存在確率を式(3)にて算出
S308
勾配法にて対象関節の回転角を修正
S309
子関節(子特徴点)の存在確率を式(3)にて算出
S310
ループC
変化量が所定の量以下？
S311
NO
YES
被写体骨格長・フレーム被写体3次元姿勢を保存
S312
ループB
ループA
被写体骨格長・被写体別3次元姿勢群を出力
S313
終了

[図19]


110
112
111
114
113
116
115
118
117
120
119
122
121
124
123
126
125
親特徴点
子特徴点

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2020/028163

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl. G06T7/70(2017.01)i
FI: G06T7/70Z

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. G06T7/70

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

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2020
Registered utility model specifications of Japan	1996-2020
Published registered utility model applications of Japan	1994-2020

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

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	JP 2019-79487 A (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) 23 May 2019 (2019-05-23) paragraphs [0016], [0019], [0021], [0025], [0028]-[0030], [0037], [0038], [0055], [0056], fig. 4	1-4, 7-8
A	entire text, all drawings	5-6
Y	PAPANDREOU, G., ZHU, T., CHEN, L. C., GIDARIS, S., TOMPSON, J., and MURPHY, K., "PersonLab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model", European Conference on Computer Vision 2018, 09 October 2018, pp. 282-299, particularly, [3 Methods, 3.1, Person detection and pose estimation Keypoint detection], fig. 1	1-4, 7-8
A	entire text, all drawings	5-6

☒ 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 29 September 2020	Date of mailing of the international search report 13 October 2020
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/JP2020/028163

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	WO 2019/030794 A1 (FUJITSU LIMITED) 14 February 2019 (2019-02-14), paragraph [0032], entire text, all drawings	3-4
A		5-6

INTERNATIONAL SEARCH REPORT
Information on patent family members

International application No.
PCT/JP2020/028163

JP 2019-79487 A	23 May 2019	(Family: none)
WO 2019/030794 A1	14 February 2019	(Family: none)

国際調査報告

国際出願番号
PCT/JP2020/028163

A. 発明の属する分野の分類（国際特許分類（ＩＰＣ））

G06T 7/70(2017.01)i
FI: G06T7/70 Z

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））

G06T7/70

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	１９７１－２０２０年
日本国実用新案登録公報	１９９６－２０２０年
日本国登録実用新案公報	１９９４－２０２０年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献のカテゴリー*	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
Y	JP 2019-79487 A (日本電信電話株式会社) 23.05.2019 (2019-05-23) 段落[0016],[0019],[0021],[0025],[0028]-[0030],[0037]-[0038],[0055]-[0056],図4	1-4,7-8
A	全文,全図	5-6
Y	George Papandreou, Tyler Zhu, Liang-Chieh Chen, Spyros Gidaris, Jonathan Tompson, and Kevin Murphy, "PersonLab: Person Pose Estimation and Instance Segmentation with a Bottom-Up, Part-Based, Geometric Embedding Model", European Conference on Computer Vision 2018, 2018.10.09, PP.282-299 特に、[3 Methods 3.1 Person detection and pose estimation Keypoint detection], 図1	1-4,7-8
A	全文,全図	5-6
Y	WO 2019/030794 A1 (富士通株式会社) 14.02.2019 (2019-02-14) 段落[0032]	3-4
A	全文,全図	5-6

☐ C欄の続きにも文献が列挙されている。

☑ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー
"A" 特に関連のある文献ではなく、一般的技術水準を示すもの
"E" 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
"L" 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
"O" 口頭による開示、使用、展示等に言及する文献
"P" 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
"T" 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの
"X" 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
"Y" 特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
"&" 同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
29.09.2020	13.10.2020

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）

粕谷　満成 5H 5587

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/JP2020/028163

引用文献	公表日	パテントファミリー文献	公表日
JP 2019-79487 A	23.05.2019	（ファミリーなし）	
WO 2019/030794 A1	14.02.2019	（ファミリーなし）	