

(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
2024年2月1日(01.02.2024)



(10) 国際公開番号

WO 2024/024055 A1

(51) 国際特許分類:

H04N 7/18 (2006.01)

(21) 国際出願番号:

PCT/JP2022/029178

(22) 国際出願日:

2022年7月28日(28.07.2022)

(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: 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).

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).

(74) 代理人: 中島 淳, 外 (NAKAJIMA, Jun et al.); 〒1600022 東京都新宿区新宿4丁目3番

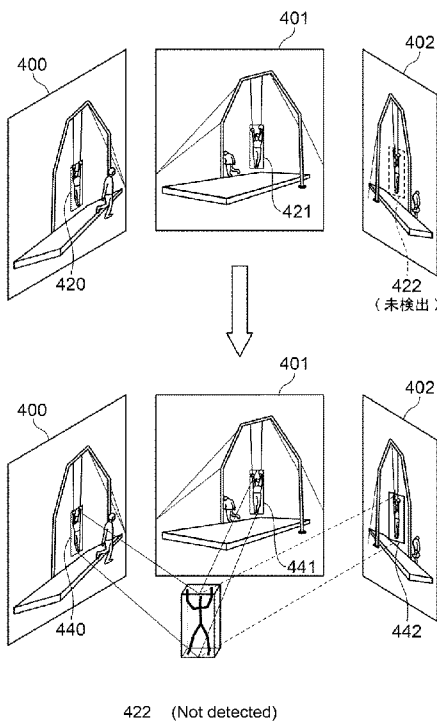
17号 H K 新宿ビル7階 太陽国際特許事務所 Tokyo (JP).

(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, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, 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,

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

(54) 発明の名称: 情報処理方法、装置、及びプログラム



(57) Abstract: This information processing device estimates three-dimensional posture information on an object on the basis of a camera parameter and two-dimensional posture information on the object within each 2D-BBOX (420, 421) detected using a detector from images (400, 401) included in a multi-view image, calculates a refined 2D-BBOX range from two-dimensional posture information on the object obtained by reprojecting the estimated three-dimensional position information on each image (400, 401, 402), and generates each calculated 2D-BBOX range as a pseudo-label (440, 441, 442).

(57) 要約: 情報処理装置は、多視点画像に含まれる画像(400、401)から検出器を用いて検出された2D-BBOX(420、421)内の対象物の2次元の姿勢情報と、カメラパラメータとに基づいて対象物の3次元の姿勢情報を推定し、推定した3次元の位置情報を各画像(400、401、402)に再投影することで得られる対象物の2次元の姿勢情報から、精巧化された2D-BBOXの範囲を算出し、算出した2D-BBOXの範囲を疑似ラベル(440、441、442)として生成する。

WO 2024/024055 A1

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： 情報処理方法、装置、及びプログラム

技術分野

[0001] 開示の技術は、情報処理方法、情報処理装置、及び情報処理プログラムに関する。

背景技術

[0002] 従来、複数の異なる視点から対象物を撮影した多視点画像から、対象物を検出することが行われている。例えば、天井カメラと壁カメラとで互いに異なる方向から撮影することで、互いに異なる背景画像中の人の像を取得する画像監視装置が提案されている。この装置は、壁カメラ画像における変化領域の画素を天井カメラ画像に射影してエピポーラ線を求め、エピポーラ線のうち変化領域の画素と特徴が共通する領域を抽出し、当該領域の存在領域に基づいて射影領域を生成する。また、この装置は、射影領域と天井カメラ画像における変化領域とを合成して合成変化領域を求め、当該合成変化領域に基づいて天井カメラ画像中の人を検出する。

[0003] 画像から対象物を検出するために、例えば、ニューラルネットワーク等の機械学習モデルが利用される。このような機械学習モデルの機械学習を実行するためには、画像内における対象物の位置情報を示す正解ラベル付きのデータが大量に必要である。しかし、正解ラベル付きのデータを大量に用意するには膨大な作業コストがかかる。そこで、機械学習モデルで検出された対象物の位置情報を疑似ラベルとして用い、予め用意された正解ラベル付きのデータに加え、疑似ラベル付きのデータも用いて機械学習モデルの機械学習を実行する半教師あり学習の手法も提案されている。

先行技術文献

特許文献

[0004] 特許文献1：特開2010-045501号公報

発明の概要

発明が解決しようとする課題

[0005] 上述したように、半教師あり学習において、機械学習モデルで検出された対象物の位置情報を疑似ラベルとして用いる場合、検出された対象物の位置情報の精度が低いと、その疑似ラベルを用いて機械学習が実行される機械学習モデルの精度も低下してしまう。特に、対象物が体操競技の選手のように、多様な姿勢を取り得る場合などには、画像から精度良く対象物の位置情報を検出することは困難である。

[0006] 一つの側面として、開示の技術は、画像内の対象物の位置情報を精度良く算出することを目的とする。

課題を解決するための手段

[0007] 一つの態様として、開示の技術は、対象物を異なる複数の視点から撮影する複数のカメラの各々で撮影された複数の画像を取得する。また、開示の技術は、前記複数の画像の各々から検出された前記対象物の2次元位置情報と、前記複数のカメラの各々のカメラパラメータとに基づいて、前記対象物の3次元位置情報を推定する。そして、開示の技術は、前記複数の画像の少なくとも1つの画像を撮影したカメラのカメラパラメータに基づいて、前記対象物の3次元位置情報を前記少なくとも1つの画像に投影し、前記少なくとも1つの画像における前記対象物の2次元位置情報を算出する。

発明の効果

[0008] 一つの側面として、画像内の対象物の位置情報を精度良く算出することができる、という効果を有する。

図面の簡単な説明

[0009] [図1]本実施形態に係る情報処理装置とカメラとの接続を示す概略図である。
[図2]2D-BBOXを検出する検出器の機械学習、及び2D-BBOXの検出を説明するための図である。
[図3]半教師あり学習を利用した検出器の機械学習を説明するための図である。

。

[図4]本実施形態に係る情報処理装置の機能ブロック図である。

[図5]2D-BBOXを説明するための図である。

[図6]対象物の2次元の姿勢情報を説明するための図である。

[図7]対象物の3次元の姿勢情報の画像への投影、及び対象物の2次元の姿勢情報の算出を説明するための図である。

[図8]3次元の姿勢情報を投影して2次元の姿勢情報を算出することの効果の説明するための図である。

[図9]疑似ラベルの空間的制限による選択を説明するための図である。

[図10]疑似ラベルの時間的制限による選択を説明するための図である。

[図11]疑似ラベルの評価に基づく選択を説明するための図である。

[図12]本実施形態に係る情報処理装置として機能するコンピュータの概略構成を示す図である。

[図13]本実施形態に係る情報処理の一例を示すフローチャートである。

[図14]本実施形態に係る情報処理装置による疑似ラベルの生成結果の一例を示す図である。

[図15]本実施形態に係る情報処理装置の、体操競技の採点システムへの適用を説明するための図である。

発明を実施するための形態

[0010] 以下、図面を参照して、開示の技術に係る実施形態の一例を説明する。

[0011] 図1に示すように、本実施形態に係る情報処理装置10は、各々異なる方向からの視点nにおいて対象物(図1の例では、対象物は人物)90を撮影する複数のカメラ30nの各々と接続される。図1の例では、 $n=0, 1, 2$ であり、視点0から撮影するカメラ300、視点1から撮影するカメラ301、及び視点2から撮影するカメラ302が情報処理装置10と接続されている。なお、情報処理装置10に接続されるカメラ30nの数は図1の例に限定されず、2台でもよいし、4台以上でもよい。

[0012] カメラ30nは、対象物90が撮影範囲に収まる角度及び位置に設置される。カメラ30nで撮影された映像は順次、情報処理装置10へ入力される

。なお、各カメラ30nには同期信号が送られ、各カメラ30nで撮影された映像の同期がとられる。

[0013] 情報処理装置10は、複数の異なる視点から撮影された複数の画像（以下、「多視点画像」という）の各々から検出される対象物90の2次元位置情報に基づいて、精巧化された対象物90の2次元位置情報を算出する。

[0014] ここで、各画像から対象物90を検出するためには、ニューラルネットワーク等の機械学習モデルである検出器が用いられる。この検出器は、例えば、図2上図に示すように、対象物90の2次元位置情報を示す正解ラベル付きの画像を大量に用いた機械学習により生成される。図2の例では、対象物90（図2の例では、体操選手）を囲む2次元のバウンディングボックス（以下、「2D-BOX」という）の左上の点の座標 $[x_1, y_1]$ 及び右下の点の座標 $[x_2, y_2]$ を正解ラベルとしている。このような正解ラベル付きの画像を用いて機械学習が実行された検出器に対して、図2下図に示すように、ラベルなしの画像を入力することで、その画像から対象物90の位置を示す2D-BOXが検出される。

[0015] 上述したように、検出器の機械学習を実行するために大量の正解付き画像を用意することは、膨大な作業コストを要する。そこで、図3に示すような半教師あり学習を利用して、検出器の機械学習を実行することが考えられる。具体的には、半教師あり学習では、正解ラベル付き画像で機械学習（図3中の「機械学習1」）が実行された検出器にラベルなし画像を入力して得られる検出結果を疑似ラベルとして生成する。そして、生成した疑似ラベルを基の画像に付与した疑似ラベル付き画像で検出器の機械学習（図3中の「機械学習2」）を実行する。これにより、大量の疑似ラベル付き画像を機械学習に利用することができ、正解付き画像が少ない場合でも検出器の機械学習を実行することができる。

[0016] しかし、単一の画像の検出結果を疑似ラベルとする場合、偽陽性及び偽陰性となる疑似ラベルを減らすことは困難である。また、偽陽性及び偽陰性ではない疑似ラベルであっても、位置のバイアスの問題が生じている場合もあ

る。位置のバイアスとは、画像上の対象物90の実際の領域に対して、疑似ラベルが示す領域の位置がシフトしている、領域のサイズが大きい、小さい等のずれが生じていることである。さらに、特に、対象物90が体操選手のように、多様な姿勢を取り得る場合、あらゆる姿勢に関してバランスの取れた疑似ラベルを生成することは困難である。例えば、体操選手の演技中の姿勢を検出した検出結果から生成された疑似ラベルに対して、単に立っているだけの姿勢を検出した検出結果から生成された疑似ラベルの数が多過ぎる場合などである。この場合、体操選手の姿勢の多様性に応じたバランスの取れた疑似ラベルが生成されているとはいえない。

[0017] また、多視点画像を用いて、多視点画像に含まれる画像間の対応関係に基づいて、生成された疑似ラベルのうち、信頼度の高い疑似ラベルを選択することも考えられる。しかし、この場合、検出されている2D-BBOXから疑似ラベルを選択するため、偽陽性の疑似ラベルを減らすことはできるが、偽陰性の疑似ラベルを減らすことはできない。また、上述した位置のバイアスの問題、及び対象物90の姿勢の多様性に応じたバランスの取れた疑似ラベルが生成できない問題も残存する。

[0018] そこで、本実施形態では、偽陽性及び偽陰性の疑似ラベルを削減することができるように、画像内の対象物90の2次元位置情報を精度良く算出する。また、本実施形態では、対象物90の実際の領域に対する位置のバイアスが修正された疑似ラベルを生成する。さらに、本実施形態では、対象物90の姿勢の多様性に応じたバランスの取れた疑似ラベルを生成する。以下、本実施形態に係る情報処理装置10について、詳細に説明する。

[0019] 図4に示すように、情報処理装置10は、機能的には、取得部11と、推定部12と、生成部13と、選択部14と、機械学習部15とを含む。また、情報処理装置10の所定の記憶領域には、検出器22と、カメラパラメータDB (database) 24とが記憶される。検出器22は、正解付き画像を訓練データとする機械学習により生成された、画像から対象物90の領域を示す2D-BBOXを検出するための機械学習モデルである。カメラパラメー

タDB24には、各カメラ30nの内部パラメータ及び外部パラメータが記憶されている。なお、生成部13は、開示の技術の「算出部」の一例である。

[0020] 取得部11は、複数のカメラ30nで撮影された時系列の多視点画像を取得する。

[0021] 推定部12は、多視点画像に含まれる各画像から検出された対象物90の領域を示す2D-BOXと、各画像を撮影した各カメラのカメラパラメータとに基づいて、対象物90の3次元位置情報を推定する。

[0022] 具体的には、推定部12は、図5に示すように、カメラ30nで撮影された画像40nから、検出器22を用いて、対象物90の領域を示す2D-BOX42nを検出する。そして、推定部12は、検出した2D-BOX42nから対象物90である人物の1以上の部位を認識するために予め機械学習により生成された認識モデル（図示省略）を用いて、対象物90の各部位の2次元位置情報を推定する。例えば、図6に示すように、認識モデルにより、対象物90である人物の各関節等の位置（図6中の黒丸）が認識される場合、推定部12は、その各関節等の位置の座標値を、対象物90の2次元位置情報として推定する。以下では、対象物90の関節等の各部位の2次元位置情報の一群を2次元の姿勢情報という。

[0023] また、推定部12は、カメラパラメータDB24に記憶されたカメラ30nのカメラパラメータと、推定した対象物90の2次元の姿勢情報とを用いて、三角測量により、対象物90の各部位の3次元位置情報を推定する。以下では、対象物90の関節等の各部位の3次元位置情報の一群を3次元の姿勢情報という。認識モデルにより、対象物90である1人の人物につきn個の関節等の部位が認識される場合、3次元の姿勢情報を $\{ [P_x^1, P_y^1, P_z^1], [P_x^2, P_y^2, P_z^2], \dots, [P_x^n, P_y^n, P_z^n] \}$ と表記する。

[0024] 生成部13は、図7のAに示すように、多視点画像に含まれる画像40nを撮影したカメラ30nのカメラパラメータに基づいて、対象物90の3次

元の姿勢情報を画像40nに投影し、画像40nにおける、精巧化した対象物90の2次元の姿勢情報を算出する。具体的には、3次元の姿勢情報 $\{ [P_x^1, P_y^1, P_z^1], [P_x^2, P_y^2, P_z^2], \dots, [P_x^n, P_y^n, P_z^n] \}$ に対応する2次元の姿勢情報を $\{ [p_x^1, p_y^1], [p_x^2, p_y^2], \dots, [p_x^n, p_y^n] \}$ とする。この場合、生成部13は、下記(1)式により、2次元の姿勢情報を算出する。なお、(1)式において、Hは、カメラ30nのカメラパラメータから決定される、3次元から2次元へ射影行列である。

[0025] [数1]

$$\begin{bmatrix} [p_x^1, p_y^1], \\ [p_x^2, p_y^2], \\ \dots \\ [p_x^n, p_y^n] \end{bmatrix} = H \begin{bmatrix} [P_x^1, P_y^1, P_z^1], \\ [P_x^2, P_y^2, P_z^2], \\ \dots \\ [P_x^n, P_y^n, P_z^n] \end{bmatrix} \quad (1)$$

[0026] 生成部13は、図7のBに示すように、算出した2次元の姿勢情報に基づいて、対象物90の領域を示す疑似ラベル44nを生成する。具体的には、生成部13は、下記(2)式に示すように、2次元の姿勢情報に含まれる各点の2次元座標の最大値及び最小値を用いて、疑似ラベル44nの左上の点の座標 $[x_1, y_1]$ 、及び右下の点の座標 $[x_2, y_2]$ を算出する。

[0027] [数2]

$$\left. \begin{aligned} w &= \max(p_x^1, p_x^2, \dots, p_x^n) - \min(p_x^1, p_x^2, \dots, p_x^n) \\ h &= \max(p_y^1, p_y^2, \dots, p_y^n) - \min(p_y^1, p_y^2, \dots, p_y^n) \\ x_1 &= \min(p_x^1, p_x^2, \dots, p_x^n) - w * \alpha \\ y_1 &= \min(p_y^1, p_y^2, \dots, p_y^n) - h * \alpha \\ x_2 &= \max(p_x^1, p_x^2, \dots, p_x^n) + w * \alpha \\ y_2 &= \max(p_y^1, p_y^2, \dots, p_y^n) + h * \alpha \end{aligned} \right\} (2)$$

[0028] なお、(2)式内のw及びhは、算出された2次元の姿勢情報が示す対象物90の外接矩形の幅及び高さである。座標 $[x_1, y_1]$ 及び $[x_2, y_2]$ の算出において、幅w又は高さhに定数 α (例えば、 $\alpha = 0.05$)を乗算した値を減算又は加算することで、 $w \times h$ の外接矩形に所定のマージンを持

たせた領域を疑似ラベル44nの範囲として算出しているものである。なお、マージンは、幅w又は高さhに定数 α を乗算した値に限定されない。予め定めた所定画素分（例えば、5画素分）をマージンとして、 $w \times h$ の範囲の上下及び左右方向に付加した範囲を疑似ラベル44nの範囲としてもよい。

[0029] このように、各画像40nの2次元の姿勢情報から3次元の姿勢情報を推定し、その3次元の姿勢情報を各画像40nに再投影して、精巧化された2次元の姿勢情報を算出することで、疑似ラベルの生成精度を向上させることができる。例えば、図8に示すように、多視点画像に画像400、401、及び402が含まれ、推定部12において、画像400及び401からは2D-BOX420及び421が検出され、画像402からは2D-BOX422が検出されていないとする。この場合でも、生成部13は、3次元の姿勢情報を画像402に再投影して、画像402から疑似ラベル442を生成することができる。すなわち、偽陰性の疑似ラベルを削減することができる。

[0030] また、生成部13は、画像400及び401に対しても3次元の姿勢情報を再投影して疑似ラベル440及び441を生成することで、2D-BOX420及び421に生じている位置のバイアスを修正することができる。

[0031] 選択部14は、生成部13で生成された疑似ラベル44nから、検出器22の機械学習に用いる疑似ラベルを空間的制限及び時間的制限に基づいて選択する。

[0032] 具体的には、選択部14は、疑似ラベル44nを生成する際の投影元である3次元の姿勢情報が示す対象物90の3次元空間における位置（以下、「3次元位置」という）が予め定めた範囲に含まれる場合に、その疑似ラベル44nを選択する。例えば、対象物90が体操選手の場合、予め定めた範囲は、競技種目に応じた競技エリアとしてよい。より具体的には、器具を使う種目の場合、その器具を含む所定範囲、種目が床の場合、規定された演技範囲を含む所定範囲を競技エリアとして定めておけばよい。

[0033] 例えば、図9に示すように、画像400から生成された疑似ラベル440

A、及び画像401から生成された疑似ラベル441Aの投影元の3次元位置46Aが競技エリア内であるとする。この場合、選択部14は、疑似ラベル440A、441Aを機械学習に用いる疑似ラベル44nとして選択する。一方、画像400から生成された疑似ラベル440B、及び画像401から生成された疑似ラベル441Bの投影元の3次元位置46Bが競技エリア外であるとする。この場合、選択部14は、疑似ラベル440B、441Bを機械学習に用いる疑似ラベル44nから除外する。これにより、選手以外の補助者、審判員等が誤って検出されている場合などに、それらの人物について生成された疑似ラベル44nを除外することができる。

[0034] また、選択部14は、疑似ラベル44nが生成された画像の撮影時刻が予め定めた時間範囲に含まれる場合に、その疑似ラベル44nを機械学習に用いる疑似ラベル44nとして選択する。例えば、対象物90が体操選手の場合、予め定めた時間範囲は、演技の開始から終了までに対応した時間範囲としてよい。

[0035] より具体的には、選択部14は、図10に示すように、一連の時系列の多視点画像の各フレームから、演技の開始に対応する開始フレーム、及び演技の終了に対応する終了フレームを特定する。器具を使う種目の場合、選択部14は、一例として、選手が競技エリアに入って、最初に足が床面から離れた瞬間の所定フレーム前のフレームを開始フレームとして特定する。また、選択部14は、選手が競技エリアから出る所定フレーム前のフレームを終了フレームとして特定する。そして、選択部14は、開始フレームから終了フレームまでを対象時間とし、対象時間に含まれるフレーム（画像40n）から生成された疑似ラベル44nを選択する。一方、選択部14は、対象時間以外の対象外のフレームから生成された疑似ラベル44nを除外する。これにより、演技開始前等に単に立っているだけの選手の姿勢に基づく疑似ラベル44nを除外することができ、選手の姿勢の多様性に応じたバランスの取れた疑似ラベル44nを選択することができる。

[0036] また、選択部14は、生成された疑似ラベル44nの品質を評価し、評価

結果が基準を満たす場合に、検出器 2 2 の機械学習に用いる疑似ラベル 4 4 n として選択する。具体的には、選択部 1 4 は、推定部 1 2 で検出器 2 2 を用いて検出された 2 D - B B O X 4 2 n と、その 2 D - B B O X 4 2 n に基づいて生成部 1 3 で生成された疑似ラベル 4 4 n との重複度を算出する。重複度は、例えば、重複部分の面積 / 疑似ラベル 4 4 n の面積としてよい。選択部 1 4 は、図 1 1 に示すように、重複度が所定の閾値以上となる疑似ラベル 4 4 n を選択し、閾値未満の疑似ラベル 4 4 n を除外する。

[0037] また、選択部 1 4 は、重複度が閾値未満の疑似ラベル 4 4 n をユーザに提示し、ユーザによる採否の判断を受け付けて、ユーザにより採用された疑似ラベル 4 4 n を、検出器 2 2 の機械学習に用いる疑似ラベル 4 4 n として選択するようにしてもよい。これにより、生成された疑似ラベル 4 4 n の採否の判断を全てユーザが行う場合に比べ、基準を満たさない疑似ラベル 4 4 n についてのみユーザに判断させるため、ユーザの負荷を軽減することができる。

[0038] 機械学習部 1 5 は、選択部 1 4 で選択された疑似ラベル 4 4 n を画像 4 0 n に付加した疑似ラベル付き画像と、正解付き画像とを訓練データとして用いて、検出器 2 2 の機械学習を実行する。機械学習部 1 5 は、取得部 1 1、推定部 1 2、生成部 1 3、及び選択部 1 4 の処理を繰り返し実行させ、得られた疑似ラベル 4 4 n を用いて、検出器 2 2 の機械学習を繰り返し実行する。処理を繰り返すことで、疑似ラベル付き画像の数が増加するため、検出器 2 2 による 2 D - B B O X 4 2 n の検出精度が向上し、疑似ラベル 4 4 n の生成精度も向上する。さらに、繰り返し処理の中で、選択部 1 4 により品質の評価結果が基準を満たす疑似ラベル 4 4 n のみを用いることで、検出器 2 2 による 2 D - B B O X 4 2 n の検出精度がより向上する。

[0039] 情報処理装置 1 0 は、例えば、図 1 2 に示すコンピュータ 5 0 で実現されてよい。コンピュータ 5 0 は、CPU (Central Processing Unit) 5 1 と、一時記憶領域としてのメモリ 5 2 と、不揮発性の記憶装置 5 3 とを備える。また、コンピュータ 5 0 は、入力装置、表示装置等の入出力装置 5 4 と、記

憶媒体 59 に対するデータの読み込み及び書き込みを制御する R/W (Read/Write) 装置 55 とを備える。また、コンピュータ 50 は、インターネット等のネットワークに接続される通信 I/F (Interface) 56 を備える。CPU 51、メモリ 52、記憶装置 53、入出力装置 54、R/W 装置 55、及び通信 I/F 56 は、バス 57 を介して互いに接続される。

[0040] 記憶装置 53 は、例えば、HDD (Hard Disk Drive)、SSD (Solid State Drive)、フラッシュメモリ等である。記憶媒体としての記憶装置 53 には、コンピュータ 50 を、情報処理装置 10 として機能させるための情報処理プログラム 60 が記憶される。情報処理プログラム 60 は、取得プロセス制御命令 61 と、推定プロセス制御命令 62 と、生成プロセス制御命令 63 と、選択プロセス制御命令 64 と、機械学習プロセス制御命令 65 とを有する。また、記憶装置 53 は、検出器 22 及びカメラパラメータ DB 24 を構成する情報が記憶される情報記憶領域 70 を有する。

[0041] CPU 51 は、情報処理プログラム 60 を記憶装置 53 から読み出してメモリ 52 に展開し、情報処理プログラム 60 が有する制御命令を順次実行する。CPU 51 は、取得プロセス制御命令 61 を実行することで、図 4 に示す取得部 11 として動作する。また、CPU 51 は、推定プロセス制御命令 62 を実行することで、図 4 に示す推定部 12 として動作する。また、CPU 51 は、生成プロセス制御命令 63 を実行することで、図 4 に示す生成部 13 として動作する。また、CPU 51 は、選択プロセス制御命令 64 を実行することで、図 4 に示す選択部 14 として動作する。また、CPU 51 は、機械学習プロセス制御命令 65 を実行することで、図 4 に示す機械学習部 15 として動作する。また、CPU 51 は、情報記憶領域 70 から情報を読み出して、検出器 22 及びカメラパラメータ DB 24 をメモリ 52 に展開する。これにより、情報処理プログラム 60 を実行したコンピュータ 50 が、情報処理装置 10 として機能することになる。なお、プログラムを実行する CPU 51 はハードウェアである。

[0042] なお、情報処理プログラム 60 により実現される機能は、例えば半導体集

積回路、より詳しくはASIC (Application Specific Integrated Circuit)、FPGA (Field-Programmable Gate Array) 等で実現されてもよい。

[0043] 次に、本実施形態に係る情報処理装置10の動作について説明する。情報処理装置10に時系列の多視点画像が入力され、検出器22の機械学習の実行が指示されると、情報処理装置10において、図13に示す情報処理が実行される。なお、情報処理は、開示の技術の情報処理方法の一例である。

[0044] ステップS11で、取得部11が、時系列の多視点画像を複数取得する。次に、ステップS12で、推定部12が、多視点画像に含まれる各画像40nから、検出器22を用いて、対象物90の領域を示す2D-BOX42nを検出する。そして、推定部12が、検出した2D-BOX42nから、認識モデルを用いて、対象物90の2次元の姿勢情報を推定する。次に、ステップS13で、推定部12が、カメラパラメータDB24に記憶されたカメラ30nのカメラパラメータと、推定した対象物90の2次元の姿勢情報とを用いて、三角測量により、対象物90の3次元の姿勢情報を推定する。

[0045] 次に、ステップS14で、生成部13が、各画像40nを撮影したカメラ30nのカメラパラメータに基づいて、対象物90の3次元の姿勢情報を各画像40nに投影し、各画像40nにおける、精巧化した対象物90の2次元の姿勢情報を算出する。そして、生成部13が、算出した2次元の姿勢情報に基づいて、疑似ラベル44nを生成する。

[0046] 次に、ステップS15で、選択部14が、上記ステップS14で生成された疑似ラベル44nから、検出器22の機械学習に用いる疑似ラベルを時空間的な制限に基づいて選択する。具体的には、選択部14が、疑似ラベル44nを生成する際の投影元である3次元の姿勢情報が示す対象物90の3次元位置が予め定めた範囲に含まれる場合に、その疑似ラベル44nを選択する。また、選択部14が、疑似ラベル44nが生成された画像の撮影時刻が予め定めた時間範囲に含まれる場合に、その疑似ラベル44nを選択する。

[0047] 次に、ステップS16で、選択部14が、上記ステップS15で選択され

た疑似ラベル44nの品質を評価し、評価結果が基準を満たす場合に、検出器22の機械学習に用いる疑似ラベル44nとして選択する。次に、ステップS17で、機械学習部15が、上記ステップS16で選択された疑似ラベル44nを画像40nに付加した疑似ラベル付き画像と、正解付き画像とを訓練データとして用いて、検出器22の機械学習を実行する。

[0048] 次に、ステップS18で、機械学習部15が、検出器22の機械学習の終了条件を満たすか否かを判定する。例えば、機械学習部15は、繰り返し回数が所定回数に到達した場合、検出器22の検出精度が所定値となった場合、検出器22の検出精度が収束した場合等に、終了条件を満たすと判定する。終了条件を満たさない場合には、ステップS11に戻り、終了条件を満たす場合には、情報処理は終了する。

[0049] 以上説明したように、本実施形態に係る情報処理装置は、多視点画像に含まれる各画像から検出された対象物の2次元位置情報と、カメラパラメータとに基づいて、対象物の3次元位置情報を推定する。そして、情報処理装置は、カメラパラメータに基づいて、対象物の3次元位置情報を各画像に投影し、精巧化された対象物の2次元位置情報を算出する。これにより、画像内の対象物の位置情報を精度良く算出することができる。また、この2次元位置情報に基づいて疑似ラベルを生成することで、疑似ラベルの偽陰性を削減し、また、疑似ラベルの位置のバイアスを修正することができる。

[0050] さらに、本実施形態に係る情報処理装置は、時空間的な制限に基づいて、生成された疑似ラベルから、検出器の機械学習に用いる疑似ラベルを選択することで、対象物の姿勢の多様性に応じたバランスの取れた疑似ラベルを生成することができる。

[0051] ここで、図14に、本実施形態に係る情報処理装置による疑似ラベルの生成結果の一例を示す。図14における左3つの図は、本実施形態における半教師あり学習を適用する前の検出器により2D-BBOXを検出する手法（以下、「比較手法」という）による検出結果の一例を概略的に示す図である。また、図14における右3つの図は、本実施形態における半教師あり学習

を適用した検出器により2D-BBOXを検出する手法（以下、「本手法」という）による検出結果の一例を概略的に示す図である。

[0052] 図14の上段の図に示すように、比較手法では不正確であった2D-BBOXが、本手法では改善されることが分かる。また、図14の中段の図に示すように、比較手法では欠落していた2D-BBOXが、本手法では検出されることが分かる。また、図14の下段の図に示すように、比較手法では、本来の対象物である選手以外の人物を示す2D-BBOXが誤検出されているが、本手法では誤検出が解消されていることが分かる。

[0053] また、上記実施形態に係る情報処理装置は、例えば、体操競技の採点システムへの適用が可能である。ここで、図15を参照して、体操競技の採点システムの処理の概略について説明する。

[0054] 採点システムは、多視点画像が入力されると、多視点画像に含まれる各画像から人物の領域を検出する。次に、採点システムは、人物が存在する位置が競技エリアか否か等に基づいて、検出した領域が示す人物が選手か選手以外かを判定し、選手を示す領域を特定する。採点システムは、時系列の多視点画像において、同一の選手を示す領域を対応付けることにより、選手をトラッキングする。採点システムは、トラッキングされた一連の画像の各々から、認識モデル等を用いて選手の2次元の骨格情報を認識する。採点システムは、2次元の骨格情報から、カメラパラメータを用いて3次元の骨格情報を推定する。そして、採点システムは、時系列の3次元の骨格情報に対する平滑化等の後処理を行い、演技のフェーズ（切れ目）を推定した後、技を認識する。

[0055] 上記の採点システムにおいて、人物の領域を検出する処理に、上記実施形態に係る情報処理装置で生成された疑似ラベルを用いて機械学習が実行された検出器を適用可能である。

[0056] なお、上記実施形態では、推定した3次元位置情報である3次元の姿勢情報を、多視点画像に含まれる全ての画像に投影する場合について説明したが、これに限定されない。検出器により2D-BBOXが検出されていない画

像を対象にするなど、多視点画像の少なくとも1つの画像に投影するようにしてもよい。

[0057] また、開示の技術は、対象物を体操選手とする場合に限定されるものではなく、他の競技の選手や、一般の歩行者等、様々な人物を対象物として適用可能である。さらに、人物以外でも、動物や車両等を対象物として適用することも可能である。

[0058] また、上記実施形態では、情報処理プログラムが記憶装置に予め記憶（インストール）されているが、これに限定されない。開示の技術に係るプログラムは、CD-ROM、DVD-ROM、USBメモリ等の記憶媒体に記憶された形態で提供されてもよい。

符号の説明

- [0059] 10 情報処理装置
- 11 取得部
 - 12 推定部
 - 13 生成部
 - 14 選択部
 - 15 機械学習部
 - 22 検出器
 - 24 カメラパラメータDB
 - 30n カメラ
 - 40n 画像
 - 42n 2D-BBOX
 - 44n 疑似ラベル
 - 50 コンピュータ
 - 51 CPU
 - 52 メモリ
 - 53 記憶装置
 - 54 入出力装置

- 5 5 R/W装置
- 5 6 通信 I/F
- 5 7 バス
- 5 9 記憶媒体
- 6 0 情報処理プログラム
- 6 1 取得プロセス制御命令
- 6 2 推定プロセス制御命令
- 6 3 生成プロセス制御命令
- 6 4 選択プロセス制御命令
- 6 5 機械学習プロセス制御命令
- 7 0 情報記憶領域
- 9 0 対象物

請求の範囲

- [請求項1] 対象物を異なる複数の視点から撮影する複数のカメラの各々で撮影された複数の画像を取得し、
- 前記複数の画像の各々から検出された前記対象物の2次元位置情報と、前記複数のカメラの各々のカメラパラメータとに基づいて、前記対象物の3次元位置情報を推定し、
- 前記複数の画像の少なくとも1つの画像を撮影したカメラのカメラパラメータに基づいて、前記対象物の3次元位置情報を前記少なくとも1つの画像に投影し、前記少なくとも1つの画像における前記対象物の2次元位置情報を算出する、
- ことを含む処理をコンピュータが実行する情報処理方法。
- [請求項2] 前記対象物の2次元位置情報を算出する処理は、前記対象物の領域を示す情報を算出することを含む請求項1に記載の情報処理方法。
- [請求項3] 前記対象物が人物の場合、前記対象物の3次元位置情報を推定する処理は、前記対象物の2次元位置情報として、前記対象物の2次元の姿勢情報を前記複数の画像の各々から検出し、前記対象物の2次元の姿勢情報に基づいて、前記対象物の3次元の姿勢情報を推定することを含む請求項1又は請求項2に記載の情報処理方法。
- [請求項4] 算出された前記対象物の2次元位置情報を疑似ラベルとして付与した画像を訓練データとして、画像から前記対象物の2次元位置情報を検出するための機械学習モデルの機械学習を実行することを含む処理を前記コンピュータが実行する請求項1又は請求項2に記載の情報処理方法。
- [請求項5] 前記疑似ラベルのうち、投影元の前記対象物の3次元位置情報が、3次元空間における予め定めた範囲に含まれる前記疑似ラベルを付与した画像を、前記訓練データとする請求項4に記載の情報処理方法。
- [請求項6] 前記対象物が体操選手の場合、前記予め定めた範囲は、競技種目に応じた競技エリアである請求項5に記載の情報処理方法。

- [請求項7] 前記疑似ラベルのうち、対応する画像の撮影時刻が予め定めた時間範囲に含まれる前記疑似ラベルを付与した画像を、前記訓練データとする請求項4に記載の情報処理方法。
- [請求項8] 前記対象物が体操選手の場合、前記予め定めた時間範囲は、演技の開始から終了までに対応した時間範囲である請求項7に記載の情報処理方法。
- [請求項9] 前記疑似ラベルの生成と、前記疑似ラベルを付与した画像を訓練データとする前記機械学習モデルの機械学習とを繰り返し実行する請求項4に記載の情報処理方法。
- [請求項10] 検出された前記対象物の2次元位置情報が示す領域と、生成された前記疑似ラベルが示す領域との重複度が予め定めた閾値未満の前記疑似ラベルを付与した画像を前記訓練データから除外する請求項9に記載の情報処理方法。
- [請求項11] 対象物を異なる複数の視点から撮影する複数のカメラの各々で撮影された複数の画像を取得する取得部と、
前記複数の画像の各々から検出された前記対象物の2次元位置情報と、前記複数のカメラの各々のカメラパラメータとに基づいて、前記対象物の3次元位置情報を推定する推定部と、
前記複数の画像の少なくとも1つの画像を撮影したカメラのカメラパラメータに基づいて、前記対象物の3次元位置情報を前記少なくとも1つの画像に投影し、前記少なくとも1つの画像における前記対象物の2次元位置情報を算出する算出部と、
を含む情報処理装置。
- [請求項12] 前記算出部は、前記対象物の2次元位置情報として、前記対象物の領域を示す情報を算出する請求項11に記載の情報処理装置。
- [請求項13] 前記推定部は、前記対象物が人物の場合、前記対象物の3次元位置情報を推定する処理は、前記対象物の2次元位置情報として、前記対象物の2次元の姿勢情報を前記複数の画像の各々から検出し、前記対

象物の2次元の姿勢情報に基づいて、前記対象物の3次元の姿勢情報を推定する請求項11又は請求項12に記載の情報処理装置。

[請求項14] 算出された前記対象物の2次元位置情報を疑似ラベルとして付与した画像を訓練データとして、画像から前記対象物の2次元位置情報を検出するための機械学習モデルの機械学習を実行する機械学習部を含む請求項11又は請求項12に記載の情報処理装置。

[請求項15] 前記機械学習部は、前記疑似ラベルのうち、投影元の前記対象物の3次元位置情報が、3次元空間における予め定めた範囲に含まれる前記疑似ラベルを付与した画像を、前記訓練データとする請求項14に記載の情報処理装置。

[請求項16] 前記対象物が体操選手の場合、前記予め定めた範囲は、競技種目に応じた競技エリアである請求項15に記載の情報処理装置。

[請求項17] 前記機械学習部は、前記疑似ラベルのうち、対応する画像の撮影時刻が予め定めた時間範囲に含まれる前記疑似ラベルを付与した画像を、前記訓練データとする請求項14に記載の情報処理装置。

[請求項18] 前記対象物が体操選手の場合、前記予め定めた時間範囲は、演技の開始から終了までに対応した時間範囲である請求項17に記載の情報処理装置。

[請求項19] 前記疑似ラベルの生成と、前記疑似ラベルを付与した画像を訓練データとする前記機械学習モデルの機械学習とを繰り返し実行する請求項14に記載の情報処理装置。

[請求項20] 対象物を異なる複数の視点から撮影する複数のカメラの各々で撮影された複数の画像を取得し、

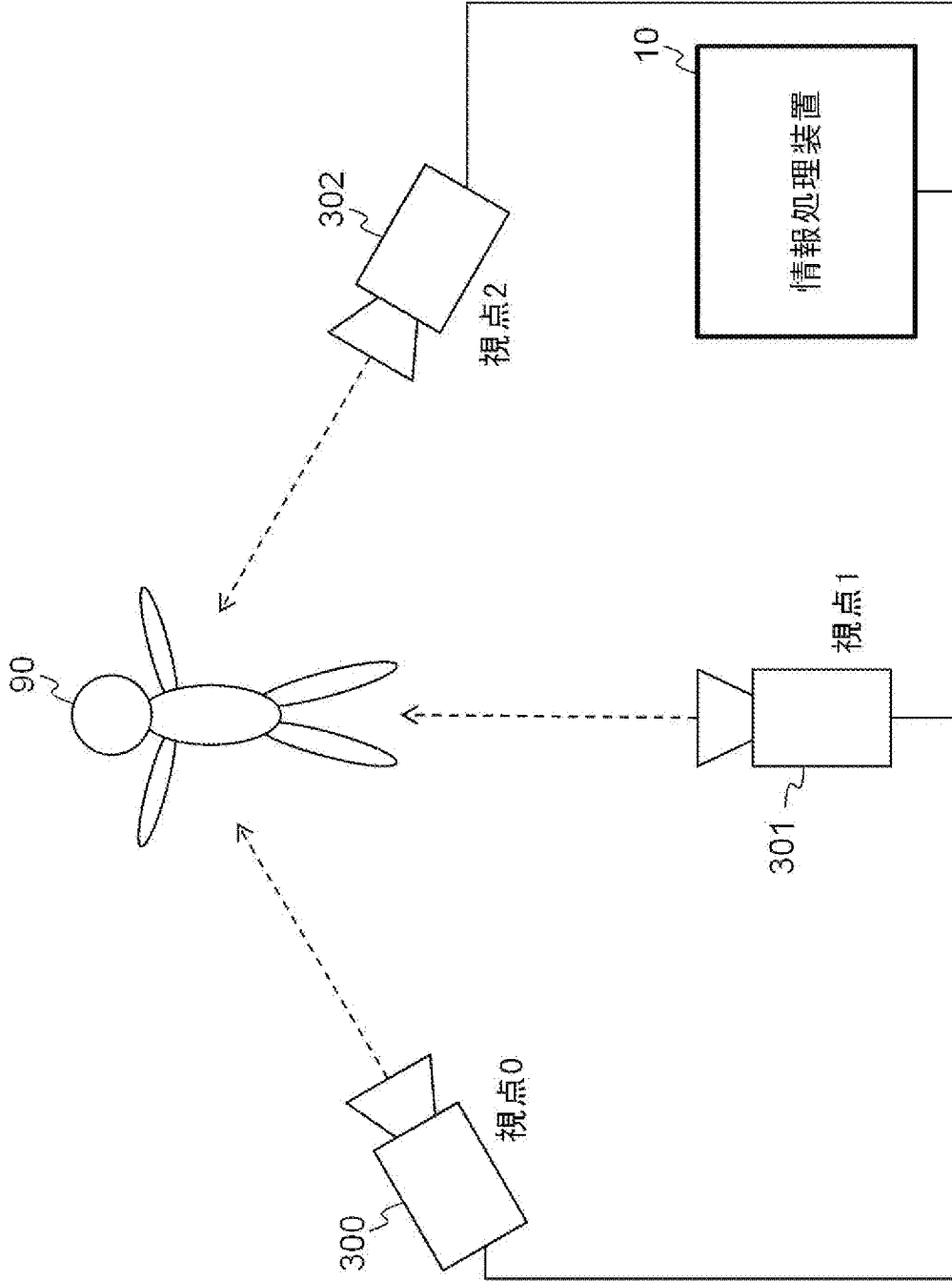
前記複数の画像の各々から検出された前記対象物の2次元位置情報と、前記複数のカメラの各々のカメラパラメータとに基づいて、前記対象物の3次元位置情報を推定し、

前記複数の画像の少なくとも1つの画像を撮影したカメラのカメラパラメータに基づいて、前記対象物の3次元位置情報を前記少なくとも

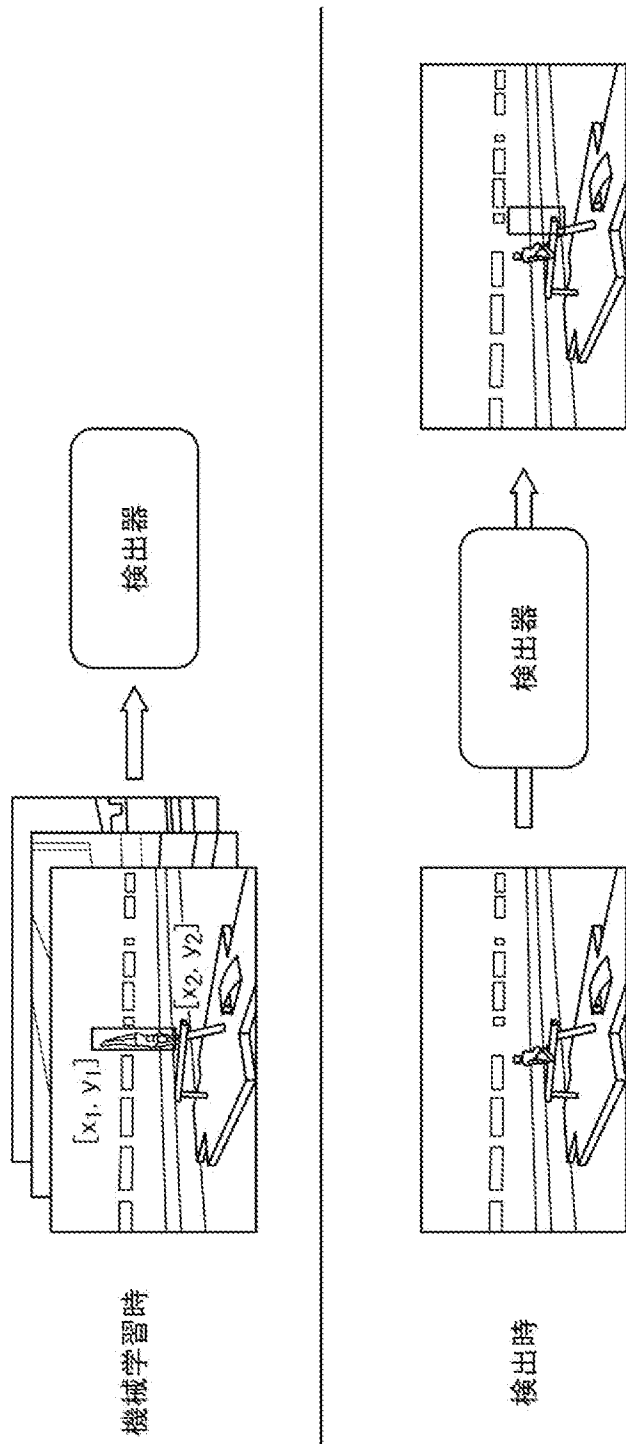
も1つの画像に投影し、前記少なくとも1つの画像における前記対象物の2次元位置情報を算出する、

ことを含む処理をコンピュータに実行させるための情報処理プログラム。

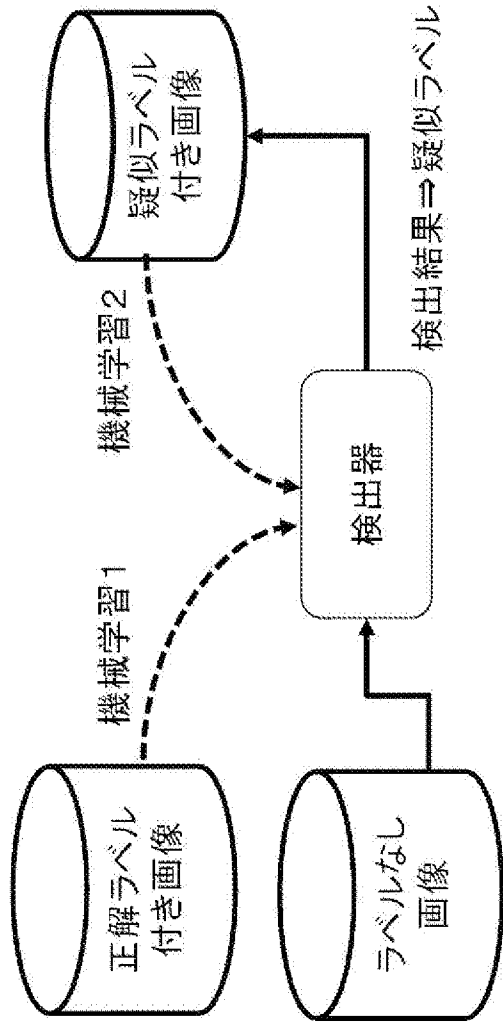
[図1]



[図2]

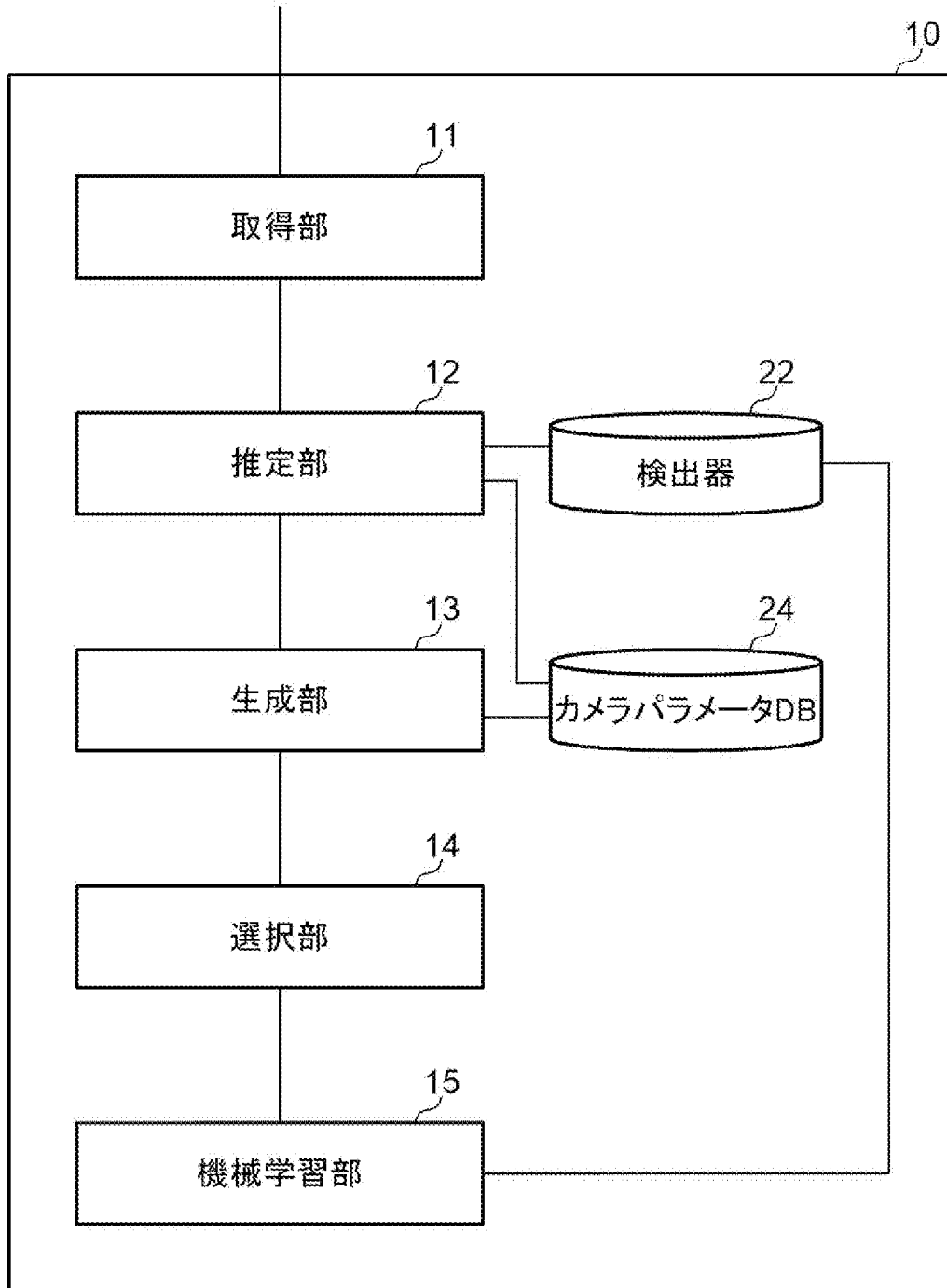


[図3]

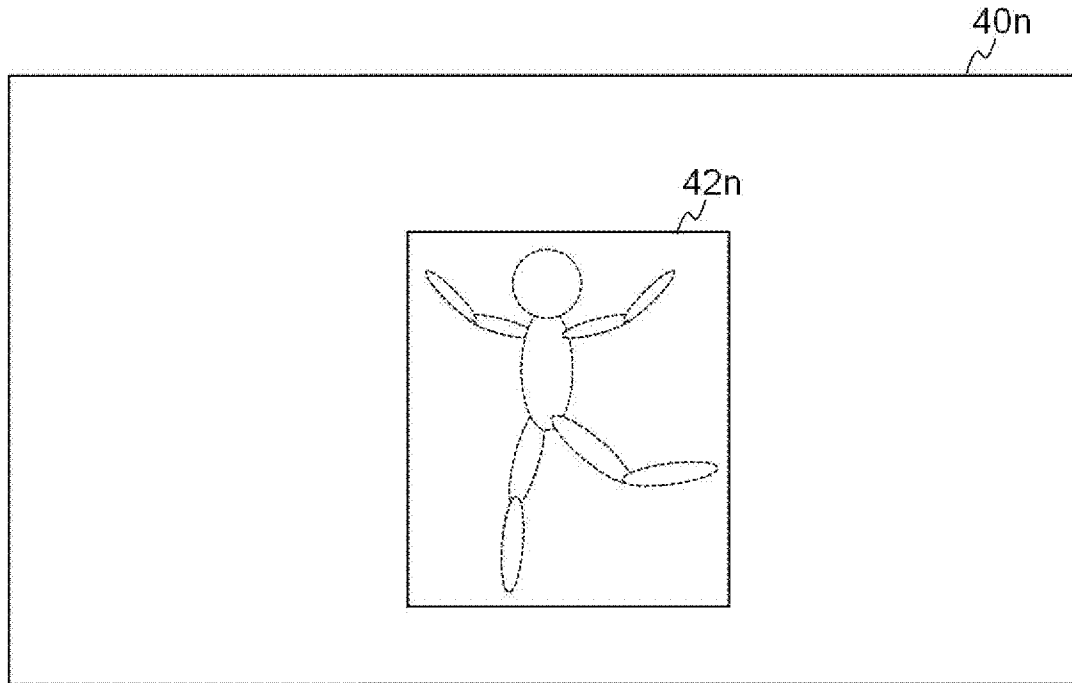


[図4]

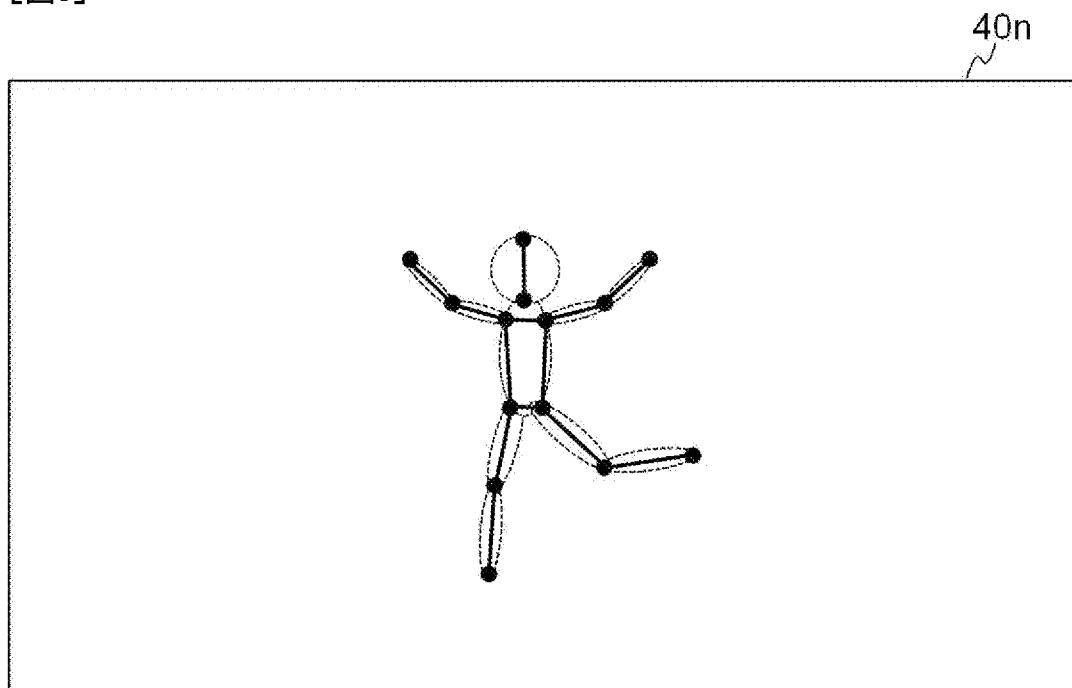
時系列の多視点画像



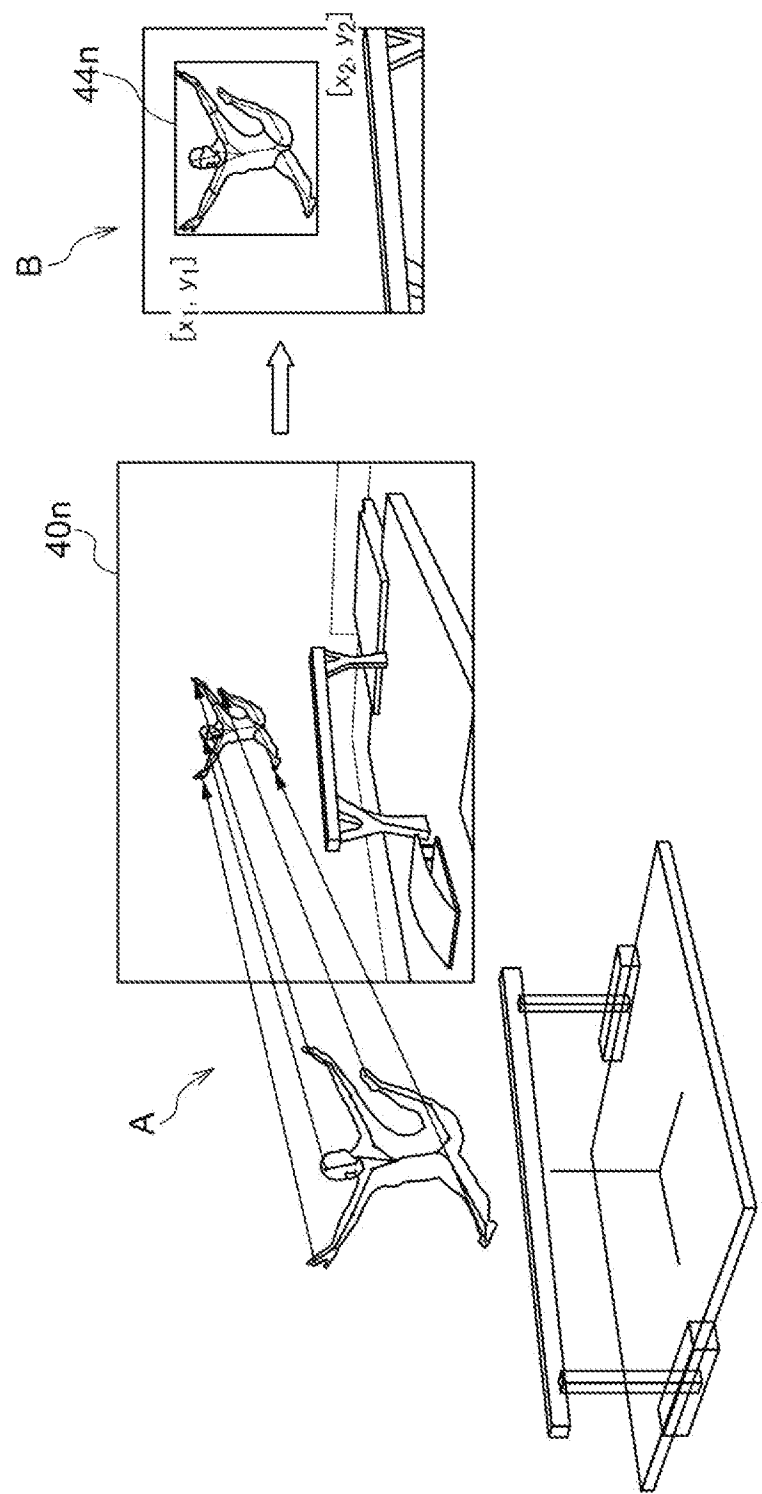
[図5]



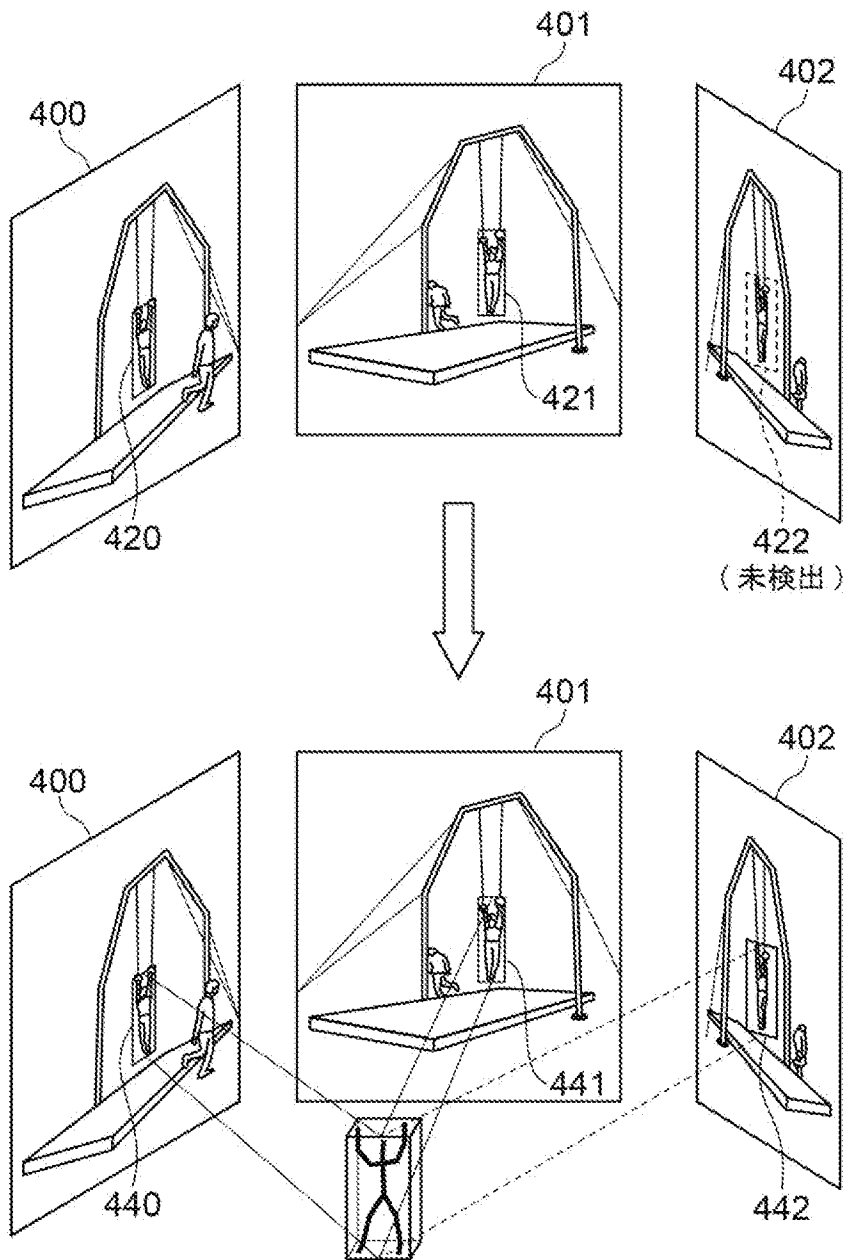
[図6]



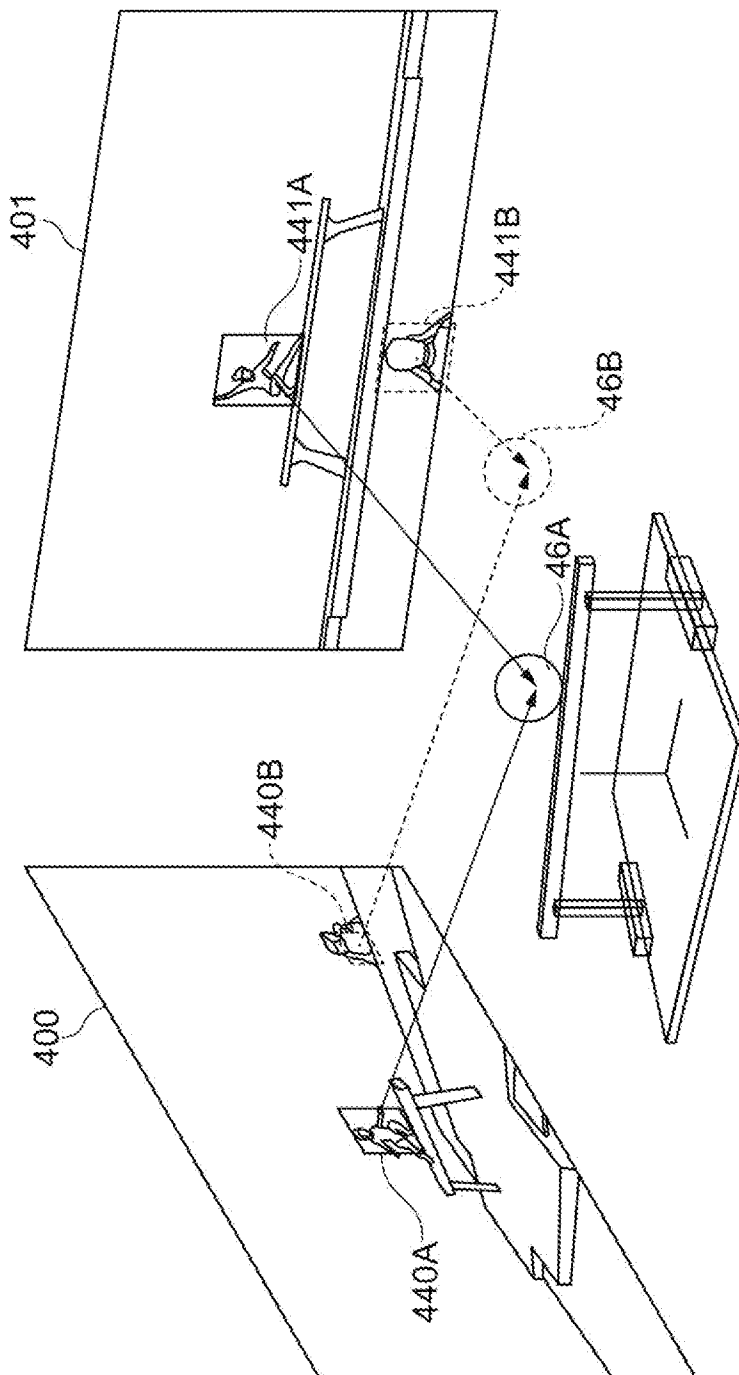
[図7]



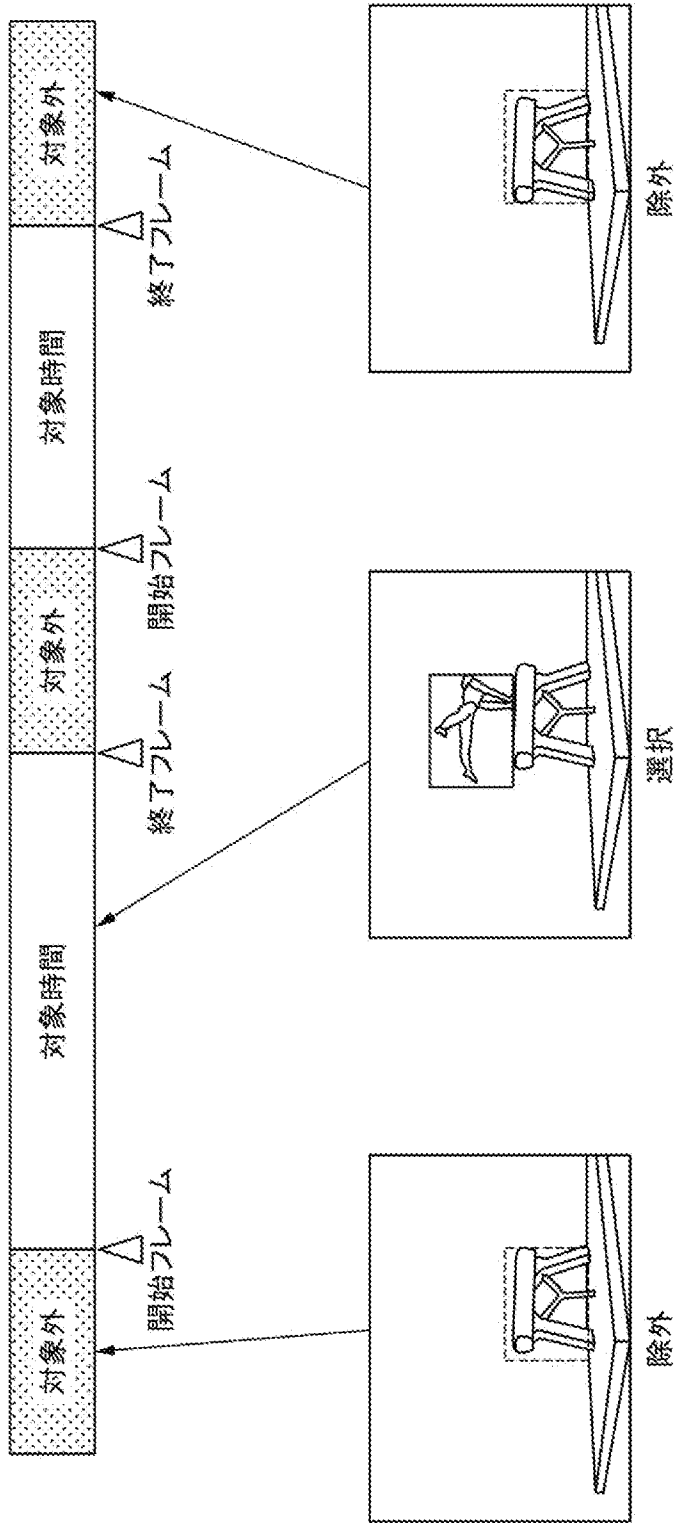
[図8]



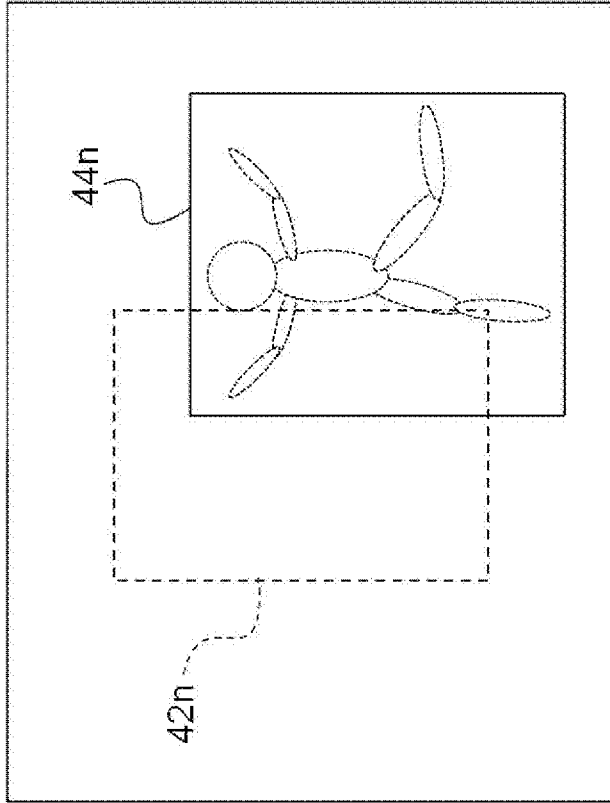
[図9]



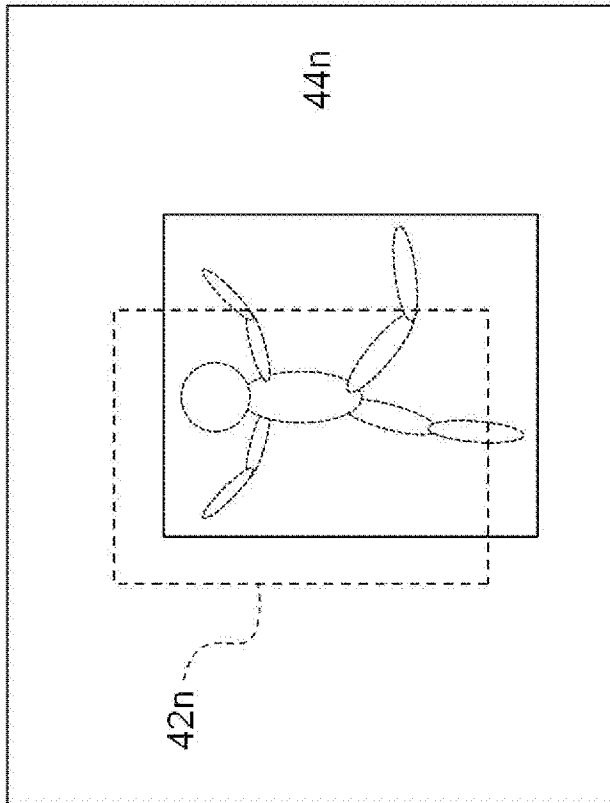
[図10]



[図11]

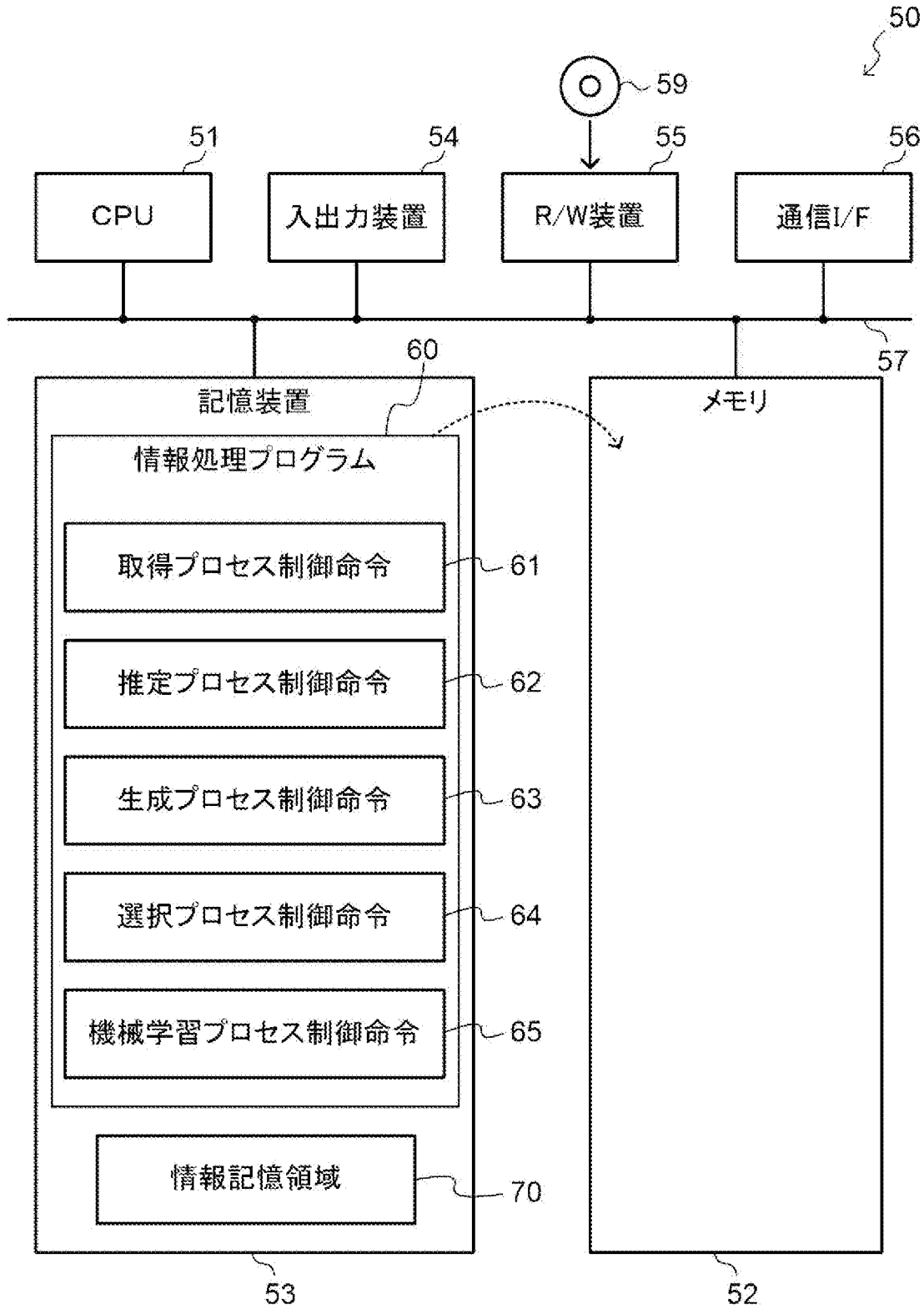


除外

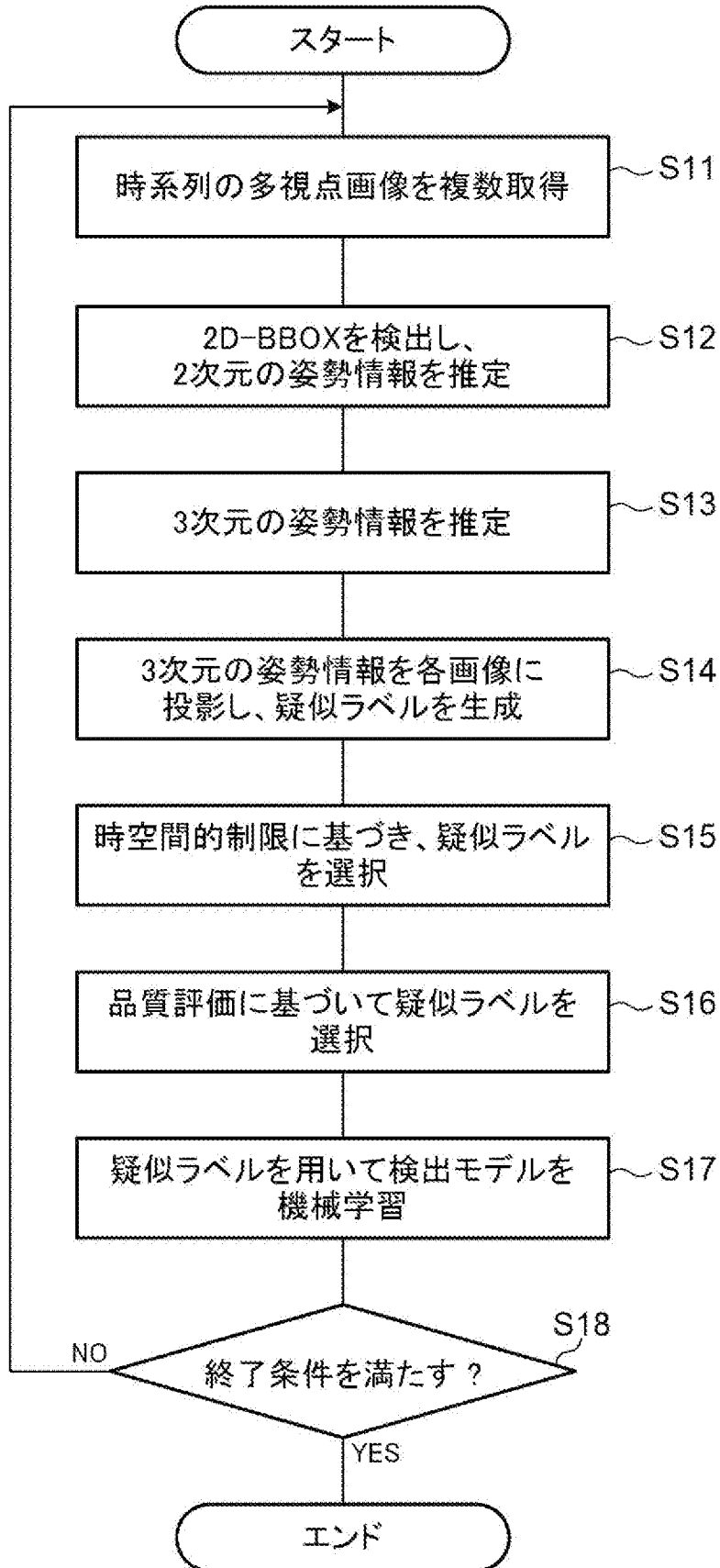


選択

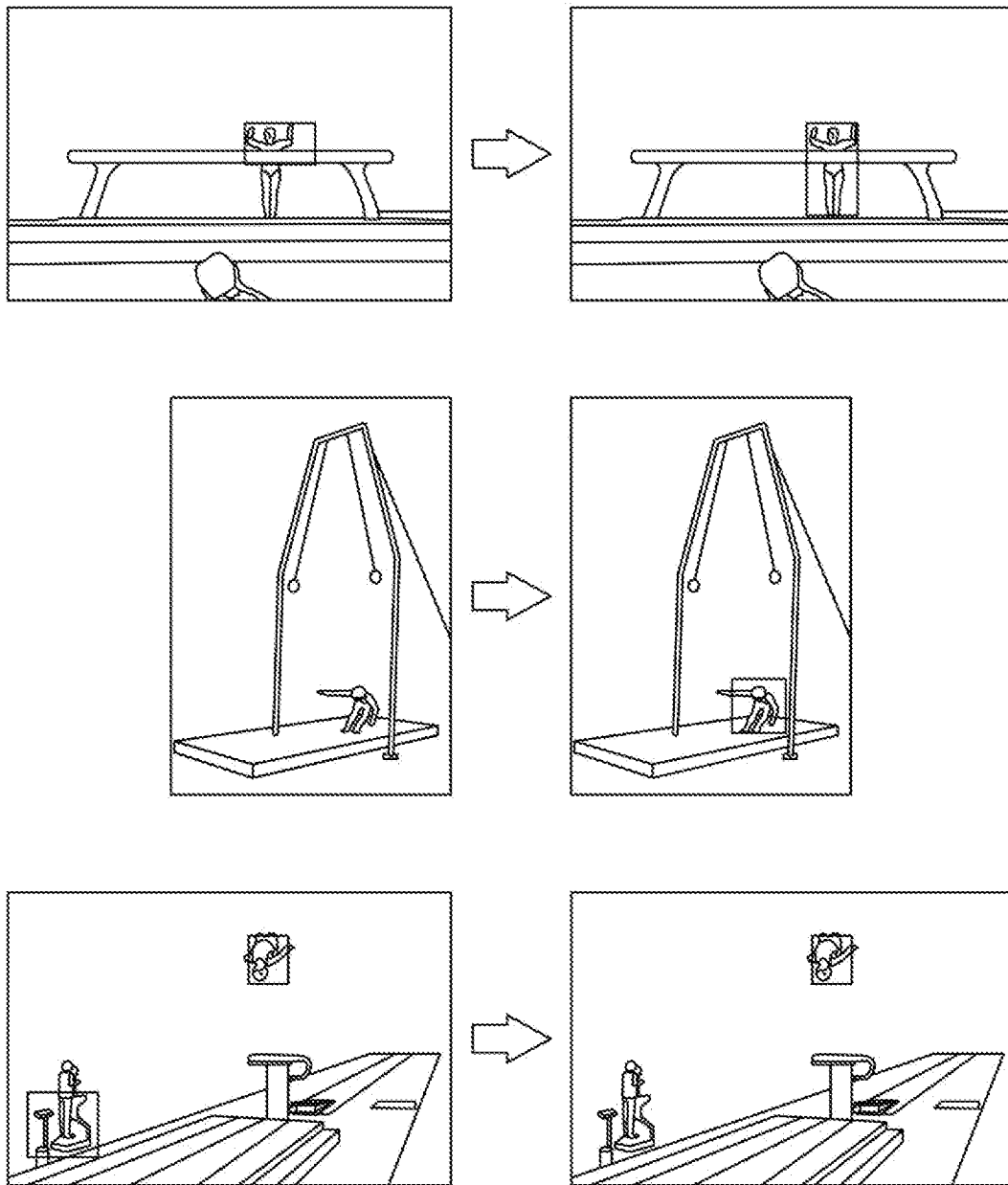
[図12]



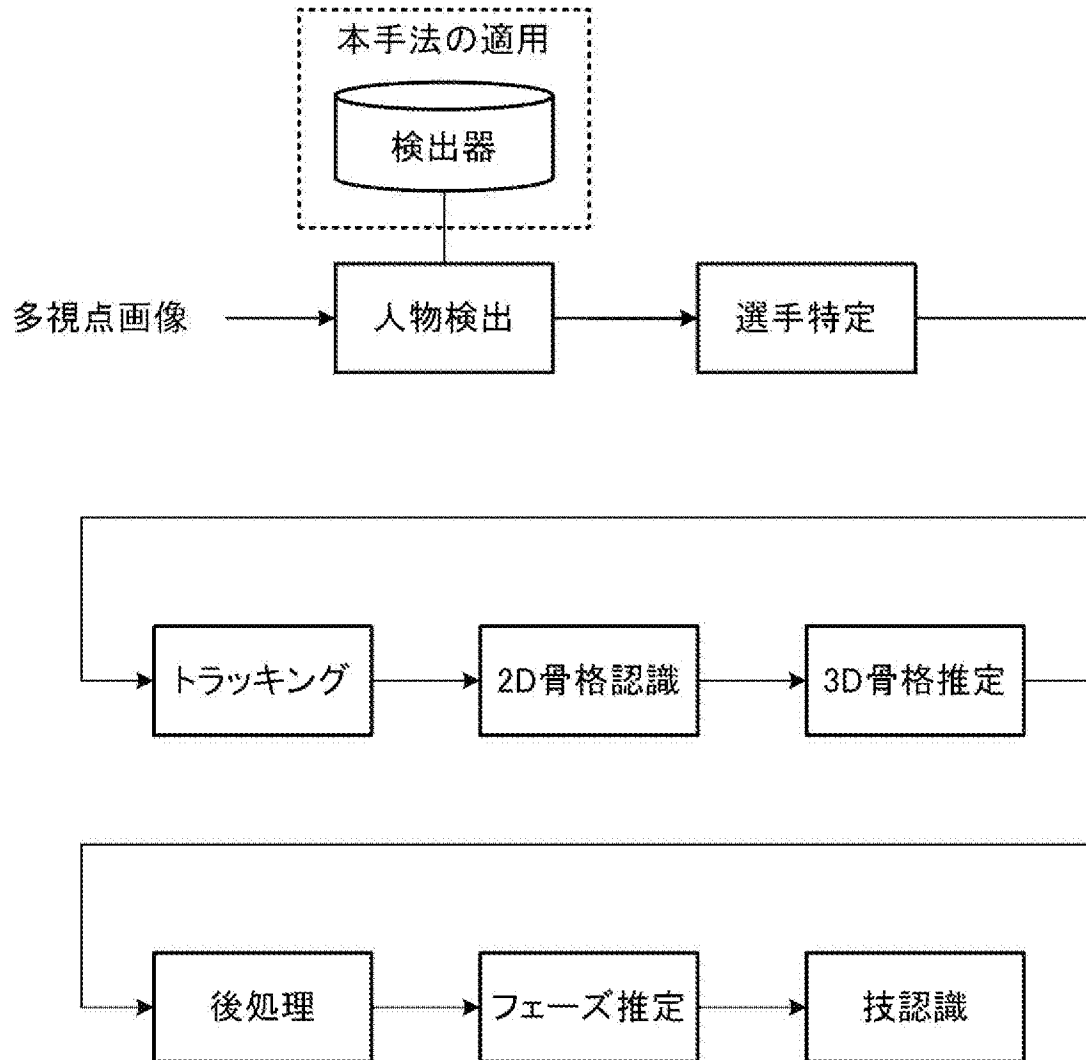
[図13]



[図14]



[図15]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/029178

| A. CLASSIFICATION OF SUBJECT MATTER | | |
|--|--|--|
| <i>H04N 7/18</i> (2006.01)i FI: H04N7/18 K | | |
| 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) H04N7/18; 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 | WO 2022/003963 A1 (FUJITSU LIMITED) 06 January 2022 (2022-01-06) paragraphs [0001]-[0006], [0011]-[0079], fig. 1-17 | 1-20 |
| A | JP 2014-240753 A (FUJITSU LIMITED) 25 December 2014 (2014-12-25) paragraphs [0032]-[0036], fig. 9-11 | 1-20 |
| A | JP 2022-64506 A (NEC SOLUTION INNOVATORS LTD) 26 April 2022 (2022-04-26) paragraphs [0002]-[0006], [0016]-[0078], fig. 1-11 | 1-20 |
| <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 06 October 2022 | | Date of mailing of the international search report 18 October 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/JP2022/029178

| Patent document cited in search report | Publication date (day/month/year) | Patent family member(s) | Publication date (day/month/year) |
|--|-----------------------------------|--|-----------------------------------|
| WO 2022/003963 A1 | 06 January 2022 | (Family: none) | |
| JP 2014-240753 A | 25 December 2014 | US 2014/0362193 A1 paragraphs [0068]-[0072], fig. 9-11 | |
| JP 2022-64506 A | 26 April 2022 | (Family: none) | |

| | | |
|--|---|--------------------------|
| A. 発明の属する分野の分類（国際特許分類（IPC）） H04N 7/18(2006.01)i FI: H04N7/18 K | | |
| 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）） H04N7/18; 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 | WO 2022/003963 A1（富士通株式会社）06.01.2022（2022 - 01 - 06） 段落[0001]-[0006], [0011]-[0079], 図1-17 | 1-20 |
| A | JP 2014-240753 A（富士通株式会社）25.12.2014（2014 - 12 - 25） 段落[0032]-[0036], 図9-11 | 1-20 |
| A | JP 2022-64506 A（NECソリューションイノベータ株式会社）26.04.2022（2022 - 04 - 26） 段落[0002]-[0006], [0016]-[0078], 図1-11 | 1-20 |
| <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 | 06.10.2022 | 国際調査報告の発送日 18.10.2022 |
| 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） 長谷川 素直 5P 2948 電話番号 03-3581-1101 内線 3581 | |

国際調査報告
特許ファミリーに関する情報

国際出願番号

PCT/JP2022/029178

| 引用文献 | 公表日 | 特許ファミリー文献 | 公表日 |
|-------------------|------------|--|-----|
| WO 2022/003963 A1 | 06.01.2022 | (ファミリーなし) | |
| JP 2014-240753 A | 25.12.2014 | US 2014/0362193 A1 段落[0068]-[0072], 図9-11 | |
| JP 2022-64506 A | 26.04.2022 | (ファミリーなし) | |