

(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
2025年3月27日(27.03.2025)

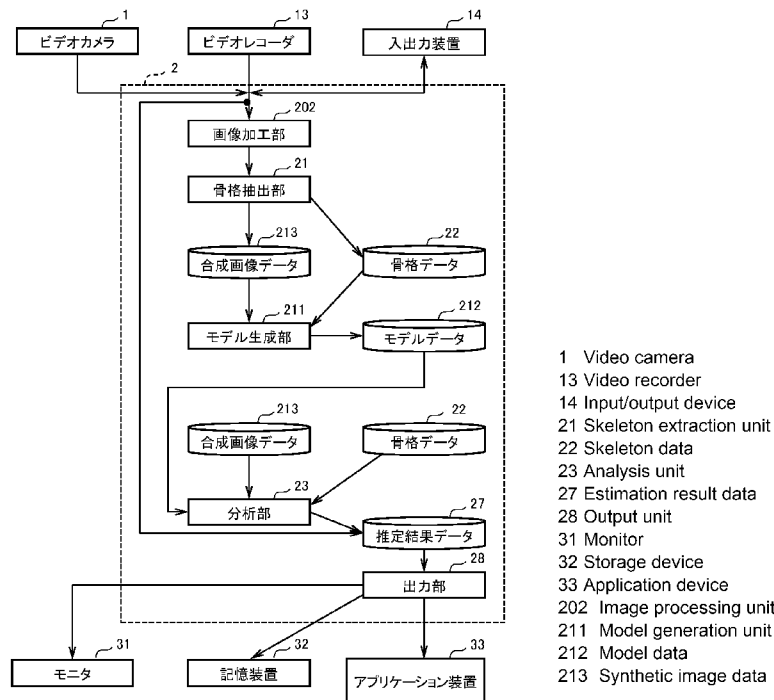


(10) 国際公開番号
WO 2025/062550 A1

- (51) 国際特許分類:
G06T 7/70 (2017.01)
- (21) 国際出願番号: PCT/JP2023/034205
- (22) 国際出願日: 2023年9月21日(21.09.2023)
- (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: 株式会社日立産業制御ソリューションズ(HITACHI INDUSTRY & CONTROL SOLUTIONS, LTD.) [JP/JP]; 〒1100006 東京都台東区秋葉原6番1号 Tokyo (JP).
- (72) 発明者: 木戸 眞一郎 (KIDO Shinichiro); 〒1100006 東京都台東区秋葉原6番1号 株式会社日立産業制御ソリューションズ内 Tokyo (JP).
川俣 昌之(KAWAMATA Yoshiyuki); 〒1100006
- (74) 代理人: 弁理士法人磯野国際特許商標事務所 (ISONO INTERNATIONAL PATENT OFFICE, P.C.); 〒1020082 東京都千代田区一番町2-1 一番町東急ビル 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, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL,

(54) Title: WORK MANAGEMENT DEVICE, WORK MANAGEMENT SYSTEM, AND WORK MANAGEMENT METHOD

(54) 発明の名称: 作業管理装置、作業管理システム、および、作業管理方法



(57) Abstract: A work management device (2) comprises: an image processing unit (202) that processes each frame of a moving image of a person into image data including a person region and an object region; a skeleton extraction unit (21) that extracts skeleton data of the person from the image data; a posture detection unit that detects the posture of the person from the skeleton data; and an analysis unit (23) that estimates the work of the person on the basis of the detection result of the posture of the person.

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, CV, 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, ME, 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))

(57) 要約 : 作業管理装置 (2) は、人物が撮影された動画の各フレームを、人物領域と物体領域を含む画像データに加工する画像加工部 (202) と、画像データからこの人物の骨格データを抽出する骨格抽出部 (21) と、骨格データから前記人物の姿勢を検出する姿勢検出部と、人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部 (23) と、を有する。

明 細 書

発明の名称：

作業管理装置、作業管理システム、および、作業管理方法

技術分野

[0001] 本発明は、作業管理装置、作業管理システム、および、作業管理方法に関する。

背景技術

[0002] 産業分野において、作業者が製品を組み立てる時間であるサイクルタイムの計測と、作業の抜け、または、定常的な作業ではない非定常作業の検知のための作業内容の分析といった処理に対するニーズがある。現在これらの処理は人手で行うことが主流となっている。

[0003] 作業内容の分析では、P C (Personal Computer) 操作、バーコード読込、ボタン押下などのように、人が作業の開始と終了を示す何らかの操作を作業時間の計測の契機としていた。または、ドリルのオン・オフ、スイッチのオン・オフ、装置の稼働を示す電流値などの作業に間接的に係っている装置からデータを取り出すことで、作業時間が計測される場合もある。

[0004] しかし、これらの計測手段は、新たな設備の追加や作業者への負担増となる。また、本来の業務ではない作業手順を追加して作業時間を計測する際には、実際にはそのような作業が行われず、正確にデータ化できないことが多い。

[0005] 作業者の作業状況をビデオカメラで記録し、人手によってその作業状況を解析することも行われている。しかし、長時間に亘って記録されたビデオカメラの映像から特定の作業者の状況を解析して、記録する作業は長時間に及ぶ。そのため多くの人的コストがかかると共に、限定的な範囲についてしかコンピュータの処理対象とすることができなかった。

[0006] 近年では、作業者の作業状況をビデオカメラで記録し、情報処理装置がその作業状況を解析することが行われるようになってきている。特許文献1には、

動作情報登録部が、骨格抽出部によって抽出された対象情報と類似する骨格情報である動作情報が示す動作内容を、対象者が行っている動作内容として特定する発明が記載されている。

先行技術文献

特許文献

[0007] 特許文献1：特許第6777819号公報

発明の概要

発明が解決しようとする課題

[0008] 上記のような情報処理装置では、入力画像サイズが大きく、高解像度であると演算処理の負担が大きくなり、リアルタイムで処理ができない場合がある。しかし、入力画像サイズを小さくすると、画素数が少なくなることで情報量が不足し、画像認識できない場合もある。

そこで、本発明は、入力画像から人物が行う作業を推定する装置にて、高い解像度の画像をリアルタイムに処理可能とすることを課題とする。

課題を解決するための手段

[0009] 前記した課題を解決するため、本発明の作業管理装置は、人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データに加工する画像加工部と、前記画像データから当該人物の骨格データを抽出する骨格抽出部と、前記骨格データから前記人物の姿勢を検出する姿勢検出部と、前記人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部と、を有することを特徴とする。

[0010] 本発明の作業管理システムは、人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データに加工する画像加工部と、前記画像データから当該人物の骨格データを抽出する骨格抽出部と、前記骨格データから前記人物の姿勢を検出する姿勢検出部と、前記人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部、を有することを特徴とする。

[0011] 本発明の作業管理方法は、画像加工部が、人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データに加工するステップと、骨格抽出部が、前記画像データから当該人物の骨格データを抽出するステップと、姿勢検出部が、前記骨格データから前記人物の姿勢を検出するステップと、分析部が、前記人物の姿勢の検出結果に基づき、前記人物の作業を推定するステップと、を有することを特徴とする。

その他の手段については、発明を実施するための形態のなかで説明する。

発明の効果

[0012] 本発明によれば、入力画像から人物が行う作業を推定する装置にて、高い解像度の画像をリアルタイムに処理可能となる。

図面の簡単な説明

[0013] [図1]第1の実施形態に関する作業管理システムの構成図である。

[図2]作業管理システムの運用を示すシーケンス図である。

[図3]画像データおよび骨格データの一例を示す図である。

[図4]合成後の画像データの一例を示す図である。

[図5]図3の骨格データを構成する特徴点データを示すテーブルである。

[図6]領域に関する処理部を示す構成図である。

[図7]姿勢に関する処理部を示す構成図である。

[図8]背景に関する処理部を示す構成図である。

[図9]手順に関する処理部と、その処理結果を出力する出力部とを示す構成図である。

[図10]背景定義部によるモデル定義を示すフローチャートである。

[図11]モデル定義の対象となる画像データを示す図である。

[図12]図11の画像データから生成される領域モデルの例を示す図である。

[図13]図11の画像データから生成される背景モデルの例を示す図である。

[図14]姿勢学習部によるモデル定義を示すフローチャートである。

[図15]図14の姿勢検出部の学習工程におけるGUI画面図である。

[図16]図14の姿勢検出部の学習工程の結果として生成される姿勢モデルを

示す図である。

[図17]手順学習部による学習結果である手順モデルを示す図である。

[図18]分析部のメイン処理を示すフローチャートである。

[図19]領域検出部のサブルーチン処理を示すフローチャートである。

[図20]図19の処理結果として「部品取り領域」で両手を認識したときの図である。

[図21]図19の処理結果として「完成品格納領域」で両手を認識したときの図である。

[図22]姿勢検出部のサブルーチン処理を示すフローチャートである。

[図23]図22の処理に用いられる画像データを示す図である。

[図24]図22の画像データに対する推論ラベル（姿勢ラベル）を示す姿勢データの図である。

[図25]背景検出部のサブルーチン処理を示すフローチャートである。

[図26]図25の処理結果として、ドライバが未使用の状態を認識したときの図である。

[図27]図25の処理結果として、ドライバが使用中の状態を認識したときの図である。

[図28]手順検出部が出力する手順データの例を示す図である。

[図29]図28の手順データをガントチャート形式で表示した画面図である。

[図30]第2の実施形態に関する作業管理装置の構成図である。

[図31]作業管理装置に入力する動画の一例を示す図である。

[図32]作業管理装置に入力する動画のフレームから抽出された骨格データの一例を示す図である。

[図33]作業管理装置に入力する画像データから認識したい部分を切り出した画像データの一例を示す図である。

[図34]骨格データの一例を示す図である。

[図35]領域モデルの一例を示す図である。

[図36]作業者の作業内容を示すタイムチャートである。

[図37]作業者の作業内容を示すタイムチャートである。

[図38]作業者の作業内容を示すタイムチャートである。

[図39]作業者の作業内容を示すタイムチャートである。

[図40]作業管理装置が判定する作業姿勢と、作業者の作業内容を示すタイムチャートである。

[図41A]作業管理処理を示すフローチャートである。

[図41B]作業管理処理を示すフローチャートである。

[図42]作業姿勢とドライバの関連付け処理を示すフローチャートである。

[図43]作業管理装置が表示するリアルタイム画面を示す図である。

[図44]作業管理装置が表示する履歴データ画面を示す図である。

[図45]作業管理装置が表示する認識レポート画面を示す図である。

発明を実施するための形態

[0014] 以降、本発明を実施するための形態を、各図を参照して詳細に説明する。

図1は、作業管理システムの構成図である。以下では、この作業管理システムを製造業の組み立て現場に導入し、作業者が小型のパソコンを組み立てる作業を分析する場合に適用した一例を説明する。

作業管理システムは、作業管理装置2を中心として、ビデオカメラ1と、ビデオレコーダ13と、入出力装置14と、モニタ31と、記憶装置32と、アプリケーション装置33とを含めて構成される。

[0015] これらの作業管理システムの各装置は、それぞれ、イーサネット（登録商標）などのネットワーク、USB（Universal Serial Bus）やその他、ハードウェア・インタフェースとして使用可能な適切なもので接続される。また、作業管理システムの各装置は、単独の装置として構成されていてもよいし、作業管理装置2などの計算機システム上のソフトウェアを実行することで実現してもよい。

[0016] ビデオカメラ1は、作業者を被写体として撮影する。ビデオレコーダ13には、ビデオカメラ1で撮影した動画が記録されている。入出力装置14は、グラフィックディスプレイ、マウスを備え、作業者などの利用者に情報表

示したり、利用者の指示を受け付けたりする。

モニタ31と、記憶装置32と、アプリケーション装置33とは、それぞれ作業管理装置2の分析結果の出力先である（詳細は図9）。

[0017] 作業管理装置2は、例えば、オンプレサーバ、もしくはクラウドサーバのような計算機システムである。作業管理装置2は、CPU（Central Processing Unit）と、メモリと、ハードディスクなどの記憶手段（記憶部）と、ネットワークインタフェースとを有するコンピュータとして構成される。このコンピュータは、CPUが、メモリ上に読み込んだプログラムを実行することにより、各処理部により構成される制御部（制御手段）を動作させる。

[0018] 作業管理装置2は、計算機システム上のプログラムを実行することで、画像加工部202と、骨格抽出部21と、モデル生成部211と、分析部23と、出力部28とを構成する。これらの構成された各処理部は、ハードディスクなどの不揮発メモリ上に蓄えられるデータ（モデルデータ212、推定結果データ27）にアクセスする。

[0019] 画像加工部202は、ビデオカメラ1またはビデオレコーダ13から入力された動画の各フレームである画像データ210をもとに、人物を認識したい領域と物体を認識したい領域を合成し、新たな画像データ213に加工する。この画像データ213の画素数は、画像データ210の画素数よりも少ない為、画像データ210が高解像度であっても演算処理の負担は少なく、リアルタイム処理性に優れる。

[0020] 画像加工部202は、動画の各フレームである画像データ210から、人物を認識したい領域と物体を認識したい領域を切り出し、人物を認識したい領域を一方に配置し、物体を認識したい領域を他方に配置して合成することで、画像データ213に加工する。認識したい対象物に応じて各領域を配置しているので、画像データ213の一方から人物を認識し、画像データ213の他方から物体を認識することができる。このとき演算処理の負担は少なく、リアルタイム処理性に優れる。

[0021] 画像加工部202は、動画の各フレームである画像データ210から、人

物を認識したい領域と物体を認識したい領域を切り出し、余白を入れて合成することで画像データ213に加工する。これにより、各領域の継ぎ目の部分の物体認識の誤検知や骨格抽出の誤検知を抑制できる。

[0022] 更に画像加工部202は、動画の各フレームである画像データ210から、人物を認識したい領域と物体を認識したい領域を切り出し、リサイズして合成することで画像データ213に加工する。認識したい対象物に応じて各領域をリサイズするので、画像データ210が高解像度であっても演算処理の負担は少なく、リアルタイム処理性に優れる。

[0023] 更に画像加工部202は、動画の各フレームである画像データ210から、人物を認識したい領域と物体を認識したい領域を切り出し、フィルタ処理して合成することで画像データ213に加工してもよい。これにより、物体認識の誤検知や骨格抽出の誤検知を抑制できる。

[0024] 画像加工部202は、動画の各フレームである画像データ210から人物を認識したい領域と物体を認識したい領域を切り出す位置を、パターンマッチングで調整するとよい。これにより、ビデオカメラ1の設置位置のズレや、作業者の作業位置のズレを調整可能である。

[0025] 画像加工部202は、動画の各フレームである画像データ210から人物を認識したい領域と物体を認識したい領域を切り出す位置を、マーカで調整するとよい。これにより、ビデオカメラ1の設置位置のズレや、作業者の作業位置のズレを調整可能である。

画像加工部202が動画の各フレームである画像データ210から人物を認識したい領域と物体を認識したい領域を切り出す位置を、手作業で調整可能な画面を表示させる調整部を備えるとよい。これにより、人物を認識したい領域と物体を認識したい領域を、利用者の所望の位置に調整可能である。

[0026] 骨格抽出部21は、画像加工部202から入力された画像データ213をもとに、骨格データ22を抽出する。骨格抽出部21は、画像データ213のうち人物を認識したい領域から、骨格データ22を抽出するとよい。画像データ210が高解像度であっても演算処理の負担は少なくなり、骨格抽出

部 2 1 は、リアルタイム処理性に優れる処理を提供可能である。

[0027] モデル生成部 2 1 1 は、学習用の画像データ 2 1 3 と、学習用の骨格データ 2 2 とを入力として、モデルデータ 2 1 2 を生成して不揮発メモリに保存する。モデルデータ 2 1 2 には、ユーザから明示的に定義された定義データと、ユーザから入力されたラベルデータを用いた学習結果である学習済データとが存在する。なお、モデル生成部 2 1 1 は、分析対象の作業に対してモデルデータ 2 1 2 を基本的には 1 度作成すればよいが、精度を向上させるために、すでに作成したモデルデータ 2 1 2 を更新（改良）してもよい。

分析部 2 3 は、分析用の画像データ 2 1 3 と、分析用の骨格データ 2 2 とを入力として、モデルデータ 2 1 2 を用いた推論処理により、推定結果データ 2 7 を求める。

出力部 2 8 は、画像データ 2 1 3 の座標で示された推定結果データ 2 7 を、画像データ 2 1 3 に加工する前のフレームである画像データ 2 1 0 の座標に変換する。出力部 2 8 は、変換した推定結果データ 2 7 を、画像データ 2 1 3 に加工する前のフレームである画像データ 2 1 0 に重畳して、外部装置（モニタ 3 1 と、記憶装置 3 2 と、アプリケーション装置 3 3）に出力する。

[0028] 図 2 は、作業管理システムの運用を示すシーケンス図である。

深層学習などの機械学習段階において、ビデオカメラ 1 から画像取得（ステップ S 1 0 1）された画像データ 2 1 0、または、ビデオカメラ 1 から画像取得（ステップ S 1 0 2）されてビデオレコーダ 1 3 が記録画像 3 2 D に画像記録（ステップ S 1 0 3）した画像データ 2 1 0 は、作業管理装置 2 に入力される。

作業管理装置 2 は、画像データ 2 1 0 から認識したい複数の画像部分を切り出し、1 枚の画像データ 2 1 3 としたのち、利用者から入出力装置 1 4 を介して受けた学習指示（ステップ S 1 1 1）により、学習処理（ステップ S 1 1 2）を実行し、その結果をモデルデータ 2 1 2 として出力する。

[0029] 分析段階において、ビデオカメラ 1 から画像取得（ステップ S 1 2 1）さ

れた画像データ210、または、ビデオカメラ1から画像取得（ステップS122）されてビデオレコーダ13が記録画像32Eに画像記録（ステップS123）した画像データ210は、作業管理装置2に入力される。

作業管理装置2は、画像データ210から認識したい複数の画像部分を切り出し、1枚の画像データ213としたのち、利用者から入出力装置14を介して受けた分析指示（ステップS131）により、モデルデータ212に基づく分析処理（ステップS132）を実行し、その結果を推定結果データ27として出力する。なお、作業管理装置2は、画像取得（ステップS121）された画像データ210から認識したい複数の画像部分を切り出し、1枚の画像データ213として、分析処理（ステップS132）をリアルタイムに実行してもよい。また、作業管理装置2は、利用者からの分析指示（ステップS131）の操作を介さずに、分析処理（ステップS132）を自動実行してもよい。

そして、作業管理装置2の出力部28は、出力処理（ステップS141）によって推定結果データ27をアプリケーション装置33などに出力する。

[0030] 図3は、画像データ210および骨格データ22の一例を示す図である。

画像データ210は、人物が写っている動画像において、フレームごとに1つ生成される。この画像データ210上には、部品取り領域901と完成品格納領域902と、工具領域903と、人物領域910とが予め定義されている。人物領域910は、人物が撮影されていることが前提となる人物領域である。部品取り領域901と完成品格納領域902と、工具領域903は、物体が撮影されていることが前提となる物体領域である。

骨格データ22は、骨格抽出部21が画像データ210から人物の骨格情報を抽出した結果である。骨格データ22は、人物の特徴点（関節点など）ごとに、1つずつ番号が割り当てられる（図では番号=0~9）。骨格抽出部21は、OpenPose（URL=https://github.com/CMU-Perceptual-Computing-Lab/openpose）などの公知の骨格情報取得技術を使用することができる。

[0031] 部品取り領域901は、作業者に対して右側に配置されており、作業者が

部品を取り出す領域である。分析部23は、この部品取り領域901に所定の骨格データ22を検出したときに、作業員の作業状態が「部品取り」であると推定する。

[0032] 工具領域903は、作業者に対して正面のやや右側に配置されており、作業者が工具を取り出し、使い終わった工具を載置する領域である。分析部23は、この工具領域903に工具が存在しているときに、作業員は未だこの工具を持っていないと推定する。

[0033] 完成品格納領域902は、作業者に対して左側に配置されており、作業者が完成品を格納する領域である。分析部23は、この完成品格納領域902と所定の骨格データ22とが重なったときに、作業員の一連の作業状態が「終了」であると推定する。

人物領域910は、作業者のイスを中心に配置されており、作業者を撮影するための領域である。

[0034] 図4は、合成後の画像データ213の一例を示す図である。

画像加工部202は、図3の画像データ210のうち不要な部分を取り除く。画像加工部202は、画像データ210のうち、部品取り領域901と、完成品格納領域902と、工具領域903と、人物領域910を合成して、新たな画像データ213に加工する。

大きなサイズの画像データ210から認識したい複数の領域を切り出し、1枚の画像とすることで、画像データ213は、元の画像データ210のサイズよりも小さくなる。

[0035] 図5は、図3の骨格データ22を構成する特徴点データを示すテーブルである。

このテーブルは、特徴点の番号ごとに、特徴点の名称（鼻など）と、特徴点の（ x , y ）座標とで構成される。特徴点の名称として、例えば、人物の首、左肩、左肘などの特徴点に対して別々の番号が割り当てられる。特徴点の名称や座標は、骨格抽出部21が画像データ213からそれぞれの関節点を画像認識で認識した結果である。

[0036] 以下、図6～図9を参照して、モデル生成部211および分析部23の詳細を説明する。作業管理装置2は、以下の(1)～(3)に示す中間的な分析結果をもとに、(4)の最終的な分析結果を求める。

(1) 「領域」の分析とは、画像データ213内にあらかじめ定義した領域内に、骨格データ22が示す作業者の身体が入っているか否かを分析することである(詳細は図6)。

[0037] (2) 「姿勢」の分析とは、骨格データ22が示す作業者の身体がどのような姿勢になっているかを分析することである(詳細は図7)。

(3) 「背景」の分析とは、画像データ213内にあらかじめ定義した背景領域内の状態を分析することである(詳細は図8)。

[0038] (4) 「手順」の分析とは、「領域、姿勢、背景」それぞれの分析結果の組み合わせを元に、画像データ213内の作業者が組み立て作業の中のどのような手順を行っているかを分析することである(詳細は図9)。

なお、出力部28は、(4)の最終的な分析結果を出力してもよいし、(1)～(3)に示す中間的な分析結果のうちの少なくとも1つを出力してもよい。

[0039] 図6は、領域に関する処理部を示す構成図である。

モデル生成部211の領域定義部231aは、画像データ210上の領域を、入出力装置14を介してユーザに多角形(四角形)座標データとして定義させ、その定義データをモデルデータ212の領域モデル24として保存する。

[0040] 分析部23の領域検出部231は、保存された領域モデル24と骨格データ22とを使用して、定義された領域内に、骨格データ22の特徴点が入っているか否かを分析し、その分析結果(人物作業状態)を推定結果データ27の領域データ24aとして出力する。

[0041] なお、領域定義部231aは、深層学習などの機械学習を用いてもよいし、用いなくてもよい。機械学習は精度が高く、汎用性も高い。しかし、学習に膨大な画像データ210を必要とし、学習に手間がかかる。また、深層学

習を使う技術が必要とされるため、製造現場において、生産管理を担当する担当者が使えるものではない。そこで、機械学習を用いる代わりに、領域モデル24を直接ユーザに定義させることで、製造現場の担当者の負担を軽減できる。

[0042] 図7は、姿勢に関する処理部を示す構成図である。

モデル生成部211の姿勢学習部232aは、骨格抽出部21が抽出した骨格データ22を表示し、その表示を見たユーザからの正解ラベル（姿勢ラベル）を受け付ける。姿勢学習部232aは、骨格データ22と姿勢ラベルとの組み合わせデータを学習し、その学習結果をモデルデータ212の姿勢モデル25として保存する。

分析部23の姿勢検出部232は、保存された姿勢モデル25と骨格データ22とを使用して、骨格データ22の人物の姿勢を分析し、その分析結果（人物作業状態）を推定結果データ27の姿勢データ25aとして出力する。

[0043] 図8は、背景に関する処理部を示す構成図である。

モデル生成部211の背景定義部233aは、画像データ210上の背景領域を、入出力装置14を介してユーザに多角形（四角形）座標データとして定義させ、その背景領域内に写っている画像内容を見たユーザからの正解ラベル（背景ラベル）を受け付ける。背景定義部233aは、背景領域と背景ラベルとの組み合わせデータを学習し、その学習結果をモデルデータ212の背景モデル291として保存する。

分析部23の背景検出部233は、保存された背景モデル291と画像データ210とを使用して、画像データ213の背景領域内の画像内容を分析し、その分析結果（機材作業状態）を推定結果データ27の背景データ291aとして出力する。

画像データ210をもとに、領域モデル24の領域と、背景モデル291の領域と、人物が主に撮影される領域を切り出して合成することで、画像データ213が生成される。

[0044] 図9は、手順に関する処理部と、その処理結果を出力する出力部28とを示す構成図である。

領域モデル24と、姿勢モデル25と、背景モデル291と、手順モデル297とを含むモデルデータ212、および、領域データ24aと、姿勢データ25aと、背景データ291aと、手順データ297aとを含む推定結果データ27は、それぞれ作業管理装置2の記憶部200に格納される。

手順検出部237は、「領域データ24a、姿勢データ25a、背景データ291a」それぞれの中間的な分析結果の推定結果データ27を組み合わせ、最終的な作業者の手順データ297a（作業状態）を決定する。中間的な分析結果の3種類のうちの1種類が誤った推定をしても、残りの2種類が正しく推定されることで、最終的な精度が向上する。

[0045] 手順検出部237による手順データ297aの決定処理には、「領域モデル24と、姿勢モデル25と、背景モデル291」それぞれの中間的なモデルデータ212の組み合わせから、手順データ297aを求めるためのモデルデータ212である手順モデル297が必要となる。

そこで、手順学習部237aは、「領域、姿勢、背景」それぞれの中間的な分析結果の組み合わせを表示し、その表示を見たユーザからの正解ラベル（手順ラベル）を受け付ける。手順学習部237aは、中間的な分析結果の組み合わせと、手順ラベルとを学習し、その学習結果をモデルデータ212の手順モデル297として保存する。このように、機械学習の手法を使った学習・推論を組み合わせることで、より短い時間で、効率的に分析できる。

[0046] 出力部28の出力演算部281は、推定結果データ27の通知を受け、出力先で要求されるデータとなるように、以下に例示する演算処理を実行させる。

HTML出力部282は、推定結果データ27をHTML形式（ブラウザ表示）に変換し、モニタ31に出力する。

CSV出力部283は、推定結果データ27をCSV形式のファイルに変換し、記憶装置32に出力する。

ソケット通信部 284 は、推定結果データ 27 をソケット通信でアプリケーション装置 33 に出力する。

[0047] 以下、図 10～図 17 を参照して、モデル生成部 211 の事例を説明する。

図 10 は、背景定義部 233 a によるモデル定義を示すフローチャートである。

ステップ S301 にて、背景定義部 233 a は、GUI (Graphical User Interface) を用いて選択されたフレームの画像データ 210 を取得する。

ステップ S302 にて、背景定義部 233 a は、選択されたフレームに対して背景ラベルのラベル付けの入力を受け付ける。

ステップ S303 にて、背景定義部 233 a は、選択されたフレームの画像データ 210 の一部である多角形 (四角形) 座標データで定義された背景領域の画像データを切り取る。

ステップ S304 にて、背景定義部 233 a は、ステップ S303 の画像データと、ステップ S302 の背景ラベルとの組み合わせを、選択されたフレームの学習データとして保持する。

[0048] ステップ S305 にて、背景定義部 233 a は、未処理のフレームが存在するときには、処理をステップ S301 に戻す。

ステップ S306 にて、背景定義部 233 a は、ステップ S306 の学習データを入力として機械学習を実行する。機械学習は、深層学習を含む、ニューラルネットワークやアンサンブル学習など、公知の技術を用いることができる。

ステップ S307 にて、背景定義部 233 a は、ステップ S306 の学習結果を、背景モデル 291 として保存すると、図 10 の処理は終了する。

以上、ステップ S301～S307 の処理により、背景定義部 233 a は、画像データ 213 から背景モデル 291 を定義した。

[0049] 図 11 は、モデル定義の対象となる画像データ 210 と 213 を示す図である。

領域定義部231aは、入出力装置14のGUIを使用して、画像データ210上の部品取り領域901と完成品格納領域902を領域モデル24用に定義させる。例えば、部品取り領域901を作業者のイスに対して右側に配置し、完成品格納領域902を左側に配置する。

背景定義部233aは、入出力装置14のGUIを使用して、画像データ210上の工具領域903を背景モデル291用に定義させる。例えば、ドライバ置き場であるドライバ領域を作業者のイスに対して右側に配置する（工具領域903）。

画像データ213は、部品取り領域901と、完成品格納領域902と、工具領域903と、人物領域910とを組み合わせたものである。このように画像加工部202は、モデルデータ212から分析対象の領域情報を取得して、設定可能であってもよい。

[0050] 図12は、図11の画像データ210から生成される領域モデル24の例を示す図である。

領域モデル24は、図10で入力された領域ごとに、領域ラベルと、特徴点番号と、判定論理と、多角形（四角形）座標データとを対応づけて構成される。例えば、領域モデル24の第1行は、「部品取り領域」として、作業者の骨格データ22の特徴点番号（4は右手首、7は左手首を示す）の両方（AND）が多角形（四角形）座標データ（4つの頂点座標、図11の部品取り領域901を示す）内に存在したときに、組み立てるパソコンの部品を作業者が取ったと認識される。

なお、判定論理の「AND」は特徴点番号のAND判定（例えば両手）を示し、「OR」は特徴点番号のOR判定（例えば片手）を示す。つまり、作業者の両手首が部品取り領域に入ったら「右側に両手が入る」という領域判定が行われる。

[0051] 図13は、図11の画像データ210から生成される背景モデル291の例を示す図である。

背景モデル291は、定義名と、ステップS303で入力される多角形（

四角形) 座標データと、ステップS 3 0 2で入力される背景ラベルとを対応づけて構成される。

例えば、利用者は、ドライバ置き場にドライバが置かれている状態の工具領域9 0 3の多角形(四角形)座標データに対して、背景ラベル「未使用(ドライバあり)」を対応づける。一方、図示は省略したが、利用者は、ドライバ置き場にドライバが置かれていない状態の工具領域9 0 3の多角形(四角形)座標データに対して、背景ラベル「使用中(ドライバなし)」を対応づける。つまり、多角形(四角形)座標データで示される画像データ2 1 3内の領域の位置が同じであっても、ドライバが置かれている画像データ2 1 0と、ドライバが置かれていない画像データ2 1 0とで、別々の背景ラベルが対応づけられる。

[0052] 図1 4は、姿勢学習部2 3 2 aによるモデル定義を示すフローチャートである。

ステップS 3 1 1にて、姿勢学習部2 3 2 aは、GUIを用いて選択されたフレームの画像データ2 1 0を取得する(図1 5で後記)。

ステップS 3 1 2にて、姿勢学習部2 3 2 aは、ステップS 3 1 1で選択されたフレームの骨格データ2 2を取得する。そして、姿勢学習部2 3 2 aは、画像データ2 1 0と骨格データ2 2とを表示して、その表示内容に対して姿勢モデルの入力を促す。

ステップS 3 1 3にて、姿勢学習部2 3 2 aは、姿勢ラベル(正解ラベル)がラベル付けされなかった場合、処理をステップS 3 1 1に戻して別のフレームを選択させる。

[0053] ステップS 3 1 4にて、姿勢学習部2 3 2 aは、ステップS 3 1 2の骨格データ2 2と、ステップS 3 1 3の姿勢ラベルとの組み合わせを、選択されたフレームの学習データとして保持する。

ステップS 3 1 5にて、姿勢学習部2 3 2 aは、未処理のフレームが存在するときには、処理をステップS 3 0 1に戻す。

ステップS 3 1 6にて、姿勢学習部2 3 2 aは、ステップS 3 0 6と同様

に、ステップS 3 1 4の学習データを入力として機械学習を実行する。

ステップS 3 1 7にて、姿勢学習部2 3 2 aは、ステップS 3 1 6の学習結果を、姿勢モデル2 5として保存する（図1 6で後記）。

[0054] 図1 5は、図1 4の姿勢検出部2 3 2の学習工程におけるG U I画面を示す図である。

利用者は、入出力装置1 4のG U I画面8 1 0を用いて、正解ラベル付けを行う。まず、利用者は、画像表示欄8 1 1から学習用の画像を見ながら、画像選択欄8 1 2からコマ送りボタンやスライダで画像を選択する。

利用者は、選択した画像が、「右から取り出す」「組み立てる」「左に置く」「その他」のいずれであるかを、正解ラベル入力欄8 1 3のボタンの押下によって、指示する。

[0055] 図1 6は、図1 4の姿勢検出部2 3 2の学習工程の結果として生成される姿勢モデル2 5を示す図である。

姿勢検出部2 3 2は、図1 5のG U Iから入力されたフレーム番号と、正解ラベルと、フレームで検出した人物の骨格データ2 2とを、機械学習の結果である姿勢モデル2 5として関連付ける。この姿勢モデル2 5は、例えば、ねじ回しの作業姿勢を学習し、ねじ回しを行っているのか否かを推定するために使用される。

なお、図1 6では、図1 5の画像表示欄8 1 1に表示された画像上の人物の姿勢から、「右から取り出す」「組み立てる」「左に置く」「その他」を判定するための姿勢モデル2 5が示される。

[0056] 図1 7は、手順学習部2 3 7 aによる学習結果である手順モデル2 9 7を示す図である。

手順モデル2 9 7は、領域モデル2 4、姿勢モデル2 5、背景モデル2 9 1の組み合わせを入力モデルとして、その入力モデルから推定される作業者の手順を出力するためのモデルである。例えば、組み立て作業は、以下の各手順などから構成される。

・部品取り手順は、作業者の右側にある組み立て対象の部品を取得する手

順である。

- ・組み立て手順は、ドライバを用いて、ねじ締を行う手順である。
- ・部品格納手順は、組み立て完了した部品を作業者の左側に置く手順である。

例えば、領域モデル24を単独に用いただけでは、人物の手が「どう」動いたのかは理解できても、人物の手が「何を」つかんだのかは不明である。

しかし、領域モデル24と背景モデル291とを併用し、背景領域内のドライバ置き場にドライバが存在しないことで、人物が「ドライバを」つかんだことが明確になる。さらに、姿勢モデル25も併用して、ドライバに手が伸びたことが分かった上で、肘の角度などにより、ドライバをとったのか、置いたのかを判定することができる。

[0057] 以下、図18～図29を参照して、分析部23の事例を説明する。

図18は、分析部23のメイン処理を示すフローチャートである。

ステップS411にて、分析部23は、モデルデータ212を取得する。

ステップS412にて、分析部23は、分析用の画像データ213を取得する。

ステップS413にて、分析部23は、ステップS412の画像データ213から、骨格抽出部21に骨格データ22を抽出させる。

領域モデル24が存在しているときには（ステップS421, Yes）、分析部23は、領域検出部231に領域データ24aの推定処理を実行させる（ステップS422, 詳細は図19）。

姿勢モデル25が存在しているときには（ステップS423, Yes）、分析部23は、姿勢検出部232に姿勢データ25aの推定処理を実行させる（ステップS424, 詳細は図22）。

[0058] ステップS425にて、分析部23は、未処理の人物がステップS412の画像データ210に存在するときには、処理をステップS21に戻す。

背景モデル291が存在しているときには（ステップS426, Yes）、分析部23は、背景検出部233に背景データ291aの推定処理を実行

させる（ステップS 4 2 7，詳細は図 2 5）。

ステップS 4 3 1にて、分析部 2 3は、未処理のフレームが存在するときには、処理をステップS 4 1 2に戻す。

ステップS 4 3 2にて、分析部 2 3は、ステップS 4 2 2，S 4 2 4，S 4 2 7の各分析結果から手順検出部 2 3 7に作業手順を推定させる。

[0059] 図 1 9は、領域検出部 2 3 1のサブルーチン処理を示すフローチャートである。

ステップS 2 2 1にて、領域検出部 2 3 1は、フレームごとに、そのフレームにて検出された人物の骨格データ 2 2を取得する。

ステップS 2 2 2にて、領域検出部 2 3 1は、領域モデル 2 4から1レコード分（1つの領域）を取得する。

領域検出部 2 3 1は、ステップS 2 2 1の骨格データ 2 2を構成する特徴点番号の座標が、ステップS 2 2 2で取得した領域内であるときには（ステップS 2 2 3，Y e s）、ステップS 2 2 2で取得したレコードの領域ラベルを保持する（ステップS 2 2 4）。

[0060] ステップS 2 2 5にて、領域検出部 2 3 1は、領域モデル 2 4内の未処理のレコードが存在するときには、処理をステップS 2 2 2に戻す。

ステップS 2 2 6にて、領域検出部 2 3 1は、ステップS 2 2 4で保持されたすべての結果を領域データ 2 4 aとして出力する。

ステップS 2 2 7にて、領域検出部 2 3 1は、未処理のフレームが存在するときには、処理をステップS 2 2 1に戻す。

[0061] 図 2 0は、図 1 9の処理結果として「部品取り領域」で両手を認識したときの図である。

領域検出部 2 3 1は、図 1 1の画像データ 2 1 3上の部品取り領域 9 0 1に対して、図 1 2の領域モデル 2 4の第1レコード（部品取り領域）の要件を満たす（つまり、右側に両手が入る）ことで、「右から部品をとる」という領域ラベルを含めた領域データ 2 4 aを推定する。

[0062] 図 2 1は、図 1 9の処理結果として「完成品格納領域」で両手を認識した

ときの図である。

領域検出部231は、図11の画像データ213上の完成品格納領域902に対して、図12の領域モデル24の第2レコード（完成品格納領域）の要件を満たす（つまり、左側に両手が入る）ことで、「左側に部品を格納する」という領域ラベルを含めた領域データ24aを推定する。

[0063] 図22は、姿勢検出部232のサブルーチン処理を示すフローチャートである。

ステップS241にて、姿勢検出部232は、動画の各フレームにて検出された人物の骨格データ22を取得する。

ステップS242にて、姿勢検出部232は、取得した骨格データ22を入力として、姿勢モデル25を用いて機械学習による推論を行う。これにより、骨格データ22に対応する姿勢ラベルが出力される。

なお、ユーザは、ステップS242の姿勢ラベルが実際と異なった場合（推論ミス）には（ステップS243, Yes）、姿勢学習部232aは、ユーザから正しい姿勢ラベルを受け付けてもよい。そして、姿勢学習部232aは、受け付けた姿勢ラベルと、取得した骨格データ22との組み合わせを新たな学習データとして、姿勢モデル25を修正（再学習）してもよい（ステップS244）。

[0064] ステップS245にて、姿勢検出部232は、出力された姿勢ラベルを推論結果として保持する。

ステップS246にて、姿勢検出部232は、未処理の人物がステップS241のフレームに存在するときには、処理をステップS243に戻す。

ステップS247にて、姿勢検出部232は、フレームに存在するすべての人物についてのステップS245で保持した推論結果を姿勢データ25aとして出力する。

ステップS248にて、姿勢検出部232は、未処理のフレームが存在するときには、処理をステップS241に戻す。

[0065] 図23は、図22の処理に用いられる画像データ210を示す図である。

出力部 28 は、時系列的に表示する画像データ 210 に対して、左側から右側にむかってフレーム番号 (f10=10番、f30=30番、…) を併記する。各フレームの画像データ 210 には、骨格抽出部 21 が認識した骨格データ 22 を示す線も人物の画像へ重畳表示されている。なお実際には、画像データ 210 の人物領域と物体領域を組み合わせた画像データ 213 にて検出処理が行われる。

[0066] 図 24 は、図 23 の画像データ 210 に対する推論ラベル (姿勢ラベル) を示す姿勢データ 25 a の図である。

姿勢検出部 232 は、ビデオカメラ 1 などから取得した画像データ 213 から、写っている人物の行動を構成する姿勢を分析し、その分析結果を姿勢データ 25 a として出力する。姿勢データ 25 a には、検出時刻を示すフレーム番号が付されている。

この出力される姿勢データ 25 a は、例えば、製造業の組み立て現場での組み立て作業の手順に係る作業姿勢や、製造業の製造現場での作業安全にかかわる身体的負担の大きい作業姿勢を検出するために活用できる。

[0067] 図 25 は、背景検出部 233 のサブルーチン処理を示すフローチャートである。

ステップ S271 にて、背景検出部 233 は、動画の各フレームである画像データ 210 を取得する。

ステップ S272 にて、背景検出部 233 は、背景モデル 291 から 1 レコード分 (1 つの背景領域) を取得する。

ステップ S273 にて、背景検出部 233 は、ステップ S271 の画像データ 213 から、ステップ S272 の背景領域の位置の画像を切り取る。

ステップ S274 にて、背景検出部 233 は、ステップ S273 で切り取った画像データ 210 を入力として、背景モデル 291 を用いて機械学習の推論を実行する。

ステップ S275 にて、背景検出部 233 は、ステップ S274 の推論結果として、背景ラベルを保持する。

ステップS 2 7 6にて、背景検出部 2 3 3は、背景モデル 2 9 1内の未処理のレコードが存在するときには、処理をS 2 7 2に戻す。

ステップS 2 7 7にて、背景検出部 2 3 3は、未処理のフレームが存在するときには、処理をステップS 2 7 1に戻す。

ステップS 2 7 8にて、背景検出部 2 3 3は、すべての背景ラベルの推論結果を出力する。

[0068] 図 2 6 は、図 2 5 の処理結果として、ドライバが未使用の状態を認識したときの図である。

出力部 2 8 は、人物行動と関連のある背景画像の情報として、背景ラベル 9 0 3 e で示した「未使用」を、人物行動画像へ重畳表示する。また、出力部 2 8 は、動画の各フレームである画像データ 2 1 0 について、骨格抽出部 2 1 が認識した骨格データ 2 2 を人物の画像へ重畳表示する。さらに、出力部 2 8 は、骨格データ 2 2 を構成する特徴点データ（関節点）をマーキングして表示する（図では丸印）。

[0069] 図 2 7 は、図 2 5 の処理結果として、ドライバが使用中の状態を認識したときの図である。出力部 2 8 は、人物行動と関連のある背景画像の情報として、背景ラベル 9 0 3 f で示した「使用中」を、人物行動画像へ重畳表示する。

[0070] 図 2 8 は、手順検出部 2 3 7 が出力する手順データ 2 9 7 a の例を示す図である。

図 9 でも説明したように、手順検出部 2 3 7 は、領域データ 2 4 a、姿勢データ 2 5 a、背景データ 2 9 1 a それぞれの分析結果を組み合わせ、作業者の手順データ 2 9 7 a（作業状態）をフレーム番号ごとに決定する。例えば、背景データ 2 9 1 a として、ドライバが未使用の状態（図 2 6）と使用中の状態（図 2 7）とを区別することで、出力手順が組み立て中か否かを決定することができる。

[0071] 図 2 9 は、図 2 8 の手順データ 2 9 7 a をガントチャート形式で表示した画面図である。

手順検出部 237 は、手順データ 297 a のフレーム番号は特定の時刻を示しているため、手順データ 297 a から時系列の作業手順（出力手順）を求めることができる。そこで、出力部 28 は、時系列の作業手順をガントチャート形式で表示することで、作業手順ごとの所要時間をユーザにわかりやすく示すことができる。

[0072] 以上説明した第 1 の実施形態では、作業管理装置 2 が深層学習などによる画像認識で画像データ 213 に写っている人物を検出し、その人物の骨格データ 22 を取得する。そして、作業管理装置 2 は、取得した骨格データ 22 と、事前に入力された正解ラベルとを姿勢モデル 25 として機械学習しておくことで、製造現場における作業者の姿勢を推定し、その姿勢から作業者の作業手順を特定する。

また、作業管理装置 2 は、骨格データ 22 との位置関係を判定するための領域モデル 24 と、人物が置かれている状況を示す背景モデル 291 とを併せて用いることで、作業手順の特定精度を向上させる。これにより、深層学習や画像認識の知識を問わず、簡便な方法で、作業者を撮影した画像データ 213 から、作業者の姿勢データ 25 a を分析し、作業者の手順データ 297 a を特定できる。

[0073] 図 30 は、第 2 の実施形態に関する作業管理装置 2 の構成図である。

以下では、この作業管理システムを製造業の組み立て現場に導入し、作業者が小型のパソコンを組み立てる作業を分析する場合に適用した一例を説明する。

[0074] 作業管理システムは、作業管理装置 2 を中心として、ビデオカメラ 1 と、ビデオレコーダ 13 と、マイク 12 と、作業管理装置 2 と、モニタ 31 と、記憶装置 32 とを含めて構成される。これらの作業管理システムの各装置は、それぞれ、イーサネット（登録商標）などのネットワーク、USB（Universal Serial Bus）やその他、ハードウェア・インタフェースとして使用可能な適切なもので接続される。

[0075] ビデオカメラ 1 は、作業者自体や、作業者の周辺にある物体や背景などを

被写体として撮影する。ビデオレコーダ13には、ビデオカメラ1で撮影した映像が記録されている。マイク12は、作業者の作業現場で発生した音を収録する。モニタ31と、記憶装置32とは、それぞれ作業管理装置2の分析結果の出力先である。つまり、作業管理装置2は、現在の作業中の作業者を撮影した画像データが入力されてもよく、また過去に作業を行った作業者を撮影した画像データが入力されてもよい。

[0076] 作業管理装置2は、例えば、オンプレミスのサーバ、もしくはクラウドサーバのような計算機システムである。作業管理装置2は、CPU (Central Processing Unit) と、メモリと、ハードディスクなどの記憶手段 (記憶部) と、ネットワークインタフェースとを有するコンピュータとして構成される。このコンピュータは、CPUが、メモリ上に読み込んだプログラム (アプリケーションや、その略のアプリとも呼ばれる) を実行することにより、各処理部により構成される制御部 (制御手段) を動作させる。なお、これに限られず、作業管理装置2がGPU (Graphics Processing Unit) を備え、このモデルの分析や表示にGPUを用いてもよい。

[0077] 作業管理装置2は、計算機システム上のプログラムを実行することで、画像加工部202と、前処理部201と、骨格抽出部21と、分析部23と、出力部28とを構成する。分析部23は更に、領域検出部231と、姿勢検出部232と、背景検出部233と、物体検出部234と、音検出部235と、移動検出部236を備える。

[0078] 画像加工部202は、後記する図31に示すように、ビデオカメラ1またはビデオレコーダ13から入力された動画像の各フレームを構成する画像データ210をもとに、認識したい複数の領域を新たな画像データ213に加工する。この画像データ213の画素数は、画像データ210の画素数よりも少ない為、画像データ210が高解像度であっても演算処理の負担は少なく、リアルタイム処理性に優れる。

[0079] 図30に戻り説明を続ける。マイク12で収録された音情報には、作業現場で計測される音には背景音 (例えば近傍の道路の車両通過音) などが含ま

れる。前処理部201は、音または音声の生データに対してFFT (Fast Fourier Transformation) 等を用いた前処理を行い、バックグラウンドノイズを除去し、作業現場で発生した音を音要素情報として抽出する。前処理部201が抽出した音要素情報は、音データ296である。

[0080] 骨格抽出部21は、画像加工部202から入力された画像データ213をもとに、骨格データ22を抽出する。

なお、骨格抽出部21は、画像データ213から骨格データ22を抽出するだけでなく、例えば人物領域910から作業者の骨格データ22を抽出するなどのように、予め定められた領域から人物を検出してもよい。

[0081] 分析部23は、領域モデル24と、姿勢モデル25と、物体モデル26と、背景モデル291と、組み合わせモデル292と、音モデル293と、移動分析モデル294を参照して、作業者が実行している作業内容を推定する。分析部23では、上記した学習段階において画像要素の情報のほかに、音要素の情報を加えて学習処理を実行する。この学習は例えば、各要素の学習や、同一時間帯に取得された画像と音からの6組の要素情報（領域、姿勢、背景、物体、音、移動）による6次元データのクラスタリング処理を行うものである。なお、分析部23は、領域、背景、物体、音、移動、人物の姿勢の検出結果から人物の作業を推定してもよく、推定に用いる検出結果は単独でも組み合わせてもよく、限定されない。

[0082] このクラスタリング処理により生成された複数のクラスタによれば、作業現場における類似作業（作業工程）毎にクラスタが生成されることになる。例えばある時間帯に画像中央部の台で作業者が座った姿勢で工具（ペンチ）を使用している状況についてクラスタXが生成されたのであれば、これは作業工程Aである可能性が高いと推定でき、画像右側で立ち作業をしている状況についてクラスタYが生成されたのであれば、これは作業工程Bであると推定できるということが判明する。

分析部23におけるクラスタリングの結果は、推定結果データ27として出力部28を介して出力され、推定結果や推定根拠などをモニタ31に表示

し、あるいは記憶装置 32 に記憶される。

[0083] 領域検出部 231 は、領域モデル 24 を参照して、予め定められた領域と、作業者の骨格との重なりを検出する。例えば、予め定められた部品領域と、作業者の骨格との重なりを検出することで、作業者が部品を取りだして作業を開始したことを推定可能である。

姿勢検出部 232 は、姿勢モデル 25 に格納された各姿勢と、映像から抽出された骨格データ 22 を参照して、画像データ中の作業者の姿勢を検出する。姿勢モデル 25 は、人の各姿勢と、その姿勢の場合における骨格データとの対応関係を記憶している。人の姿勢と作業との対応関係を参照することにより、作業者がどのような作業を行っているかを推定可能である。

[0084] 背景検出部 233 は、背景モデル 291 を参照して画像データから背景を検出する。背景モデル 291 は、背景の各構造物と、その構造物の外観データとの対応関係を記憶している。この背景モデル 291 により、背景の構造物の外観データを推定し、よって作業者がどのような作業を行っているかを推定可能である。

[0085] 物体検出部 234 は、物体モデル 26 を用いて、画像データから物体を検出する。物体モデル 26 は、工具などの物体と、その物体の外観データとの対応関係を記憶している。この物体モデル 26 により、工具などがどのように使われているかを知ることができ、よって作業者がどのような作業を行っているかを検知可能である。同様に、

これらの検出部にて検出した結果に加え、分析部 23 は、組み合わせモデル 292 を参照して、検出結果を組み合わせ、作業員が実施している作業内容を推定する。組み合わせモデル 292 は、作業員が実施する各作業内容と、その作業内容を示す各検出部の検出結果との組み合わせの対応関係を記憶している。

なお、物体検出部 234 は、画像データ 213 から物体を検出するだけでなく、例えば部品取り領域 901 から部品を検出し、完成品格納領域 902 から完成品を検出し、工具領域 903 から工具を検出するなどのように、予

め定められた領域から定められた物体を検出してもよい。

- [0086] 音検出部235は、音モデル293を参照して音データ296から作業内容を検出する。音モデル293は、音要素と作業内容との対応関係を記憶している。音データ296によって、作業者がどのような作業を行っているかを推定可能である。
- [0087] 移動検出部236は、骨格データ22を参照して移動分析データ295を生成し、この移動分析データ295から作業内容を検出する。移動分析モデル294は、移動分析データ295と作業内容との対応関係を記憶している。移動分析データ295によって、作業者がどのような作業を行っているかを推定可能である。
- [0088] これらの構成された各処理部は、ハードディスクなどの不揮発メモリ上に蓄えられる骨格データ22または推定結果データ27にアクセスする。
- [0089] 骨格抽出部21は、ビデオカメラ1またはビデオレコーダ13から入力された画像データをもとに、この画像データに撮影されている人物の骨格データ22を抽出する。分析部23は、分析用の画像データと、分析用の骨格データ22とを入力として、作業員が実行する作業を推定した推定結果データ27を求める。出力部28は、推定結果データ27を外部装置であるモニタ31と、記憶装置32に出力する。
- [0090] 物体検出部234は、機械学習モデルを利用して、動画を構成する各フレームから物体を検出する。ここで検出する物体は、説明上、ドライバなど作業に用いる工具としているが、画像上に映っている物体/模様などであればよく、限定されない。物体検出部234は、後記する図31の画像データ210などから物体の位置と種類を検出する。物体検出部234が物体を検出する範囲は、フレーム全面だけでなく、指定した骨格を基準にした領域内とする。例えば、作業者の手が届く範囲とし、作業者の把持できる範囲に制限して物体の検出を行うなどである。具体的に言うと、本実施形態の物体検出部234は、骨格データ22から決定した両手首近傍の検出範囲内で、所定の物体であるドライバを検出する。この範囲の制限により、物体検出部

- 234は、物体の検出を高速化できると共に、物体の誤検出を抑制できる。
- [0091] 物体検出部234は、物体モデル26の機械学習の結果を用いて、物体を検出した座標と確からしさを示すスコアを取得する。物体検出部234は、事前に定めた閾値により、明らかにスコアが低い検出結果を、物体モデル26の機械学習の結果から除外する。また、物体検出部234は、複数の連続するフレームで物体を検出した連続検出回数が閾値以上である条件を満たしたものを検出と判定したり、多数決などによる検出結果のフィルタリングを実施したりすることで、誤検知を抑制し、更に検出がフレーム毎にちらつくことを防ぐ。
- [0092] 物体検出部234は、作業者の把持する物体を検出するなど、指定した骨格を基準にした領域内で物体を検出する。しかし、作業者の身体や作業スペースにより物体の一部が隠れることが多い。一部が隠れる場合は、検出精度が低下し、作業時間の計測が困難になる。そこで、物体検出部234が、例えばプラスドライバまたはマイナスドライバを検出した際、同じ作業姿勢を取っている時間を、ネジ締結の作業と推定する。抽出したドライバの位置は、画像に重畳表示される。そして、作業時間は、グラフ化して出力される。
- [0093] 本実施形態では、事前の学習段階において、作業現場の作業状況をビデオカメラ1とマイク12から情報収集して、例えばクラスタリング処理により、作業状況を反映した複数のクラスタを生成しておく。なお生成したクラスタの中には、作業工程に特有の事象を色濃く反映したクラスタばかりではなく、作業工程を特定するには信頼度の低いものを含むことも考えられる。このことから、クラスタ生成の際に得られるスコアなどを参照して、発生したすべてのクラスタの中から、作業者自身によるクラスタの見直しを適宜行うのがよい。
- [0094] そのうえで、実利用段階では、実利用段階で入手した画像、音からの6組の要素情報（領域、姿勢、背景、物体、音、移動）による6次元データのクラスタリング処理を行い、既存のクラスタのどれに分類されるのか、あるいは新クラスタを生成するのかを判断する。この結果、クラスタXに分類され

るとしたら、この状況は作業工程Aである可能性が高いと推定できる。

[0095] なおここでは、学習がクラスタリングである事例について説明したが、要するに作業工程における画像と音の関係が、特定の作業工程を示すものであることが判明している場合に、この関係を利用して、特定の作業工程であることを判断することができるものであればよい。

[0096] 音の要素を加味した本発明によれば、推定精度の向上を期待できる。例えば画像のみの場合には、組立や検査を行う上で使われる工具や装置の状態が人物の陰等になり画像から認識できない場合がある。これは上記例では、画像による4次元の要素情報（領域、姿勢、背景、物体）のうち、物体の要素情報が失われた3次元の要素情報からの判断となるために精度が出ないということである。この点、本実施形態では、これに音の要素が加わるためにペンチを使用するときのパチン、パチンという音が確認できるのであれば、クラスタXに分類される確率が高くなるように評価されることになる。

[0097] まず音に関して、作業の上で使用中の音を発する工具音も学習しておくことで、人陰で画像では認識できない手持ち工具等も認識できるようになり、人物が製品の組立や検査を行っている状況を、より精度よく自動抽出できる。音または音声の生データまたはFFT等を用いた前処理を行った後のデータを機械学習にて推定するのがよく、音声による言葉認識からも判断できる。

[0098] また、作業者が把持／使用している物体の画像認識および、物体から発生する音を検出することで、作業者が行う作業内容の推定精度の向上を期待できる。画像認識と音認識は個別または組合せで作業内容を推定するのがよい。

[0099] また音または音声が発生する位置がおおよそ決まっている場合、指向性マイクを向けることで、周囲雑音を拾うことなく音または音声により精度よく認識が可能である。工具を取り出す際の音や、工具を置くときの音、さらに取り出すときは、重さが変化することでブザーを鳴らしての認識も可能であり、例えば音から電動ドライバを回した回数も認識可能である。

[0100] 学習機能に関して、画像、骨格、音（または音声）の誤検出について、機械学習の結果として確からしさを示すスコアを取得できる。スコアは事前に定めた閾値により、明らかにスコアが低い検出結果は結果から除外することに利用できる。また、連続検出回数や、多数決などによる検出結果のフィルタリングを実施することで、誤検知抑制や検出がちらつくことを防ぐのがよい。

[0101] 図31は、作業管理装置2に入力する映像の一例を示す図である。

画像データ210は、人物が写っている動画の各フレームに1つ生成される。この画像データ210に写っている人物は、電子装置を手元に持っており、未だドライバを手にしていない。部品取り領域901は、作業者が部品を取るための領域である。工具領域903は、作業者が作業で使う工具を載置する領域である。完成品格納領域902は、作業者が完成品を格納する領域である。

[0102] マーカ911, 912は、画像データ210に撮影されているマーカであり、画像加工部202が人物を認識するための人物領域910を決定するためのものである。画像加工部202は、作業管理処理を実行する前にマーカ911, 912を認識して、人物領域910を設定する。

[0103] このように、画像加工部202は、動画中に撮影されている人物を認識するための人物領域や物体を認識するための領域を、マーカに基づいて決定してもよい。これにより、作業者は、自身の作業領域を認識して、ビデオカメラ1に適切に撮影されることが可能である。また、作業管理装置2は、ビデオカメラ1で撮影された動画のフレームから、人物を認識するための人物領域を適切に設定可能である。

[0104] なお、画像加工部202は、人物を認識するための人物領域や物体を認識するための領域を、作業管理処理を実行する前に、パターンマッチングに基づいて決定してもよい。また、画像加工部202が人物を認識するための人物領域や物体を認識するための領域を調整可能な画面を表示して、手作業で調整可能としてもよい。

[0105] 図32は、作業管理装置2に入力する映像から抽出された骨格データ42と、予め設定された部品取り領域901、工具領域903、完成品格納領域902の一例を示す図である。

骨格データ42は、骨格抽出部21が画像データ210から人物の骨格情報を抽出した結果を、この画像データ210に重畳表示したものであり、各特徴点を線で結んで示している。骨格抽出部21は、OpenPose (URL=<https://github.com/CMU-Perceptual-Computing-Lab/openpose>) などの公知の骨格情報取得技術を使用することができる。

[0106] 部品取り領域901は、部品が置かれている領域をこの画像データ210に重畳表示したものである。分析部23は、この部品取り領域901と所定の骨格データ22とが重なったときに、作業員の作業状態が「部品取り」であると推定する。

工具領域903は、工具が置かれる領域をこの画像データ210に重畳表示したものである。分析部23は、この工具領域903に工具が存在しているときに、作業員は未だこの工具を持っていないと推定する。

[0107] 完成品格納領域902は、完成品が格納される領域をこの画像データ210に重畳表示したものである。分析部23は、この完成品格納領域902と所定の骨格データ22とが重なったときに、作業員の一連の作業状態が「終了」であると推定する。

[0108] 図33は、合成後の画像データ213の一例を示す図である。

画像加工部202は、図32の画像データ210のうち不要な部分を取り除く。画像加工部202は、画像データ210のうち、部品取り領域901と完成品格納領域902と、工具領域903と、人物領域910を合成して、新たな画像データ213に加工する。なお、これに限られず、画像加工部202は、合成時に各領域に対して必要に応じて変形、色の調整、フィルタ処理などを行ってもよい。

大きなサイズの画像データ210から認識したい複数の領域を切り出し、1枚の画像とすることで、画像データ213は、元の画像データ210のサ

イズよりも小さくなる。

[0109] 図34は、骨格データ22の一例を示す図である。

骨格データ22は、番号欄と、特徴点欄と、X軸座標欄と、Y軸座標欄と、スコア欄とを含んで構成される。各行には、人物の関節点などである特徴点が格納される。番号欄は、例えば、鼻は0、右肩は1、右肘は2などのように、各特徴点に1つずつ固有の番号が割り当てられている。特徴点欄は、この特徴点の名前を格納している。X軸座標欄は、この特徴点の画像データ210におけるX軸の座標値を格納している。Y軸座標欄は、この特徴点の画像データ210におけるY軸の座標値を格納している。スコア欄は、この特徴点の確からしさを表す数値を格納している。

[0110] 図35は、領域モデル24の一例を示す図である。

領域モデル24は、領域ラベル欄と、特徴点番号欄と、判定論理欄と、多角形座標欄とを含んで構成される。領域ラベル欄は、この領域の名前を格納する欄である。特徴点番号欄は、骨格データ22の特徴点の番号を1つ、または複数格納する欄である。特徴点番号欄で指定された特徴点が、多角形座標欄で指定された領域に入った場合、分析部23は、その特徴点が領域に入ったことを検知する。

[0111] 判定論理欄は、特徴点番号欄が複数であるとき、これら番号を有する特徴点がこの領域に進出したことを判定する論理を示している。多角形座標欄は、この領域を示す多角形の座標値を格納する欄である。

[0112] 例えば、領域モデル24の第1行は、図32の部品取り領域901として、作業者の骨格データ22の特徴点番号（#3は右手首、#6は左手首を示す）のうち何れか（論理和）が多角形（四角形）座標データ（4つの頂点座標）内に存在したときに、組み立てるパソコンの部品を作業者が取ったと認識される。

[0113] なお、判定論理の「論理積」は特徴点番号の論理積による判定（例えば両手）を示し、「論理和」は特徴点番号の論理和による判定（例えば片手）を示す。つまり、作業者の片手首が部品取り領域901に入ったら、「部品取

り領域に手が入る」という領域判定が行われる。

[0114] 領域モデル24の第2行は、図32の完成品格納領域902として、作業者の骨格データ22の特徴点番号（#3は右手首、#6は左手首を示す）のうち何れか（論理和）が多角形（四角形）座標データ（4つの頂点座標）内に存在したときに、組み立てた完成品を作業者が格納したと認識される。

[0115] 領域モデル24の第3行は、図32の工具領域903として、作業者の骨格データ22の特徴点番号（#3は右手首、#6は左手首を示す）の両方（論理積）が多角形（四角形）座標データ（4つの頂点座標）内に存在したときに、ドライバを作業者が取ったか、または作業者がドライバを戻したと認識される。

[0116] 図36、図37は、本発明の時の発生事象と推論結果の関係の例を示す図である。これらの図では、横軸に時間推移を示し、縦軸側に上から順に実態としての作業工程、画像情報の例として姿勢の情報、音情報、推論結果を示している。

図36は、作業者の作業内容を示すタイムチャートである。

図36は、画像から姿勢Aが認識できない区間を含む場合を示している。この例では、実態としての作業工程は時刻t0からt4まで作業工程Aを実施し、その後に作業工程Bに移行した。画像情報の中の姿勢は、時刻t0からt2間と、時刻t3からt4間で作業工程Aでの姿勢Aが計測されたが、時刻t2とt3の間はこの姿勢Aが計測されなかった。音情報DMは、ほぼこの期間全域で作業工程Aでの音Aが計測された。

[0117] この時の推論結果は、図39の左に示すように、この作業工程Aの実施期間T内に姿勢の不検出期間T0はあるものの、この期間T内で少なくとも一定時間以上の姿勢と音の双方が検知される時間帯T1が存在することから工程Aの可能性が高いと判断したものである。

[0118] 図37は、作業者の作業内容を示すタイムチャートである。

図37は、画像から姿勢Aか姿勢Cかが認識できない区間を含む場合を示している。この例では、実態としての作業工程は時刻t0からt4まで作業

工程 A を実施し、その後に工程 B に移行した。画像情報中の姿勢は、時刻 t_1 から t_4 間で作業工程 A での姿勢 A または姿勢 C が計測された。音情報は、ほぼこの期間全域で作業工程 A での音 A が計測された。

[0119] この時の推論結果は、図 39 の左に示すように、この作業工程 A の実施期間 T 内に姿勢 A と姿勢 C の混在する区間があるものの、この期間 T 内で少なくとも一定時間以上の姿勢 A と音の双方が検知される時間帯 T1 が存在することから工程 A の可能性が高いと判断したものである。

[0120] 図 38 は、作業者の作業内容を示すタイムチャートである。

図 38 は、作業時間を計測するために、作業の開始と終了を画像と音から判断する例を示している。画像からは作業の開始を意味する姿勢 A と音 A を検知したことで、作業開始と判断する。また作業の終了を意味する姿勢 B と音 C を検知したことで、作業終了と判断していることを示している。

[0121] 図 39 は、作業者の作業内容を示すタイムチャートである。

この作業工程 A の実施期間 T 内に姿勢 A と姿勢 C の混在する区間があるものの、この期間 T 内で少なくとも一定時間以上の姿勢 A と音の双方が検知される時間帯 T1 が存在することから工程 A の可能性が高いと判断したものである。

[0122] 図 40 は、作業管理装置 2 が判定する作業姿勢と、作業者の作業内容を示すタイムチャートである。

「姿勢」の右側の棒グラフは、時刻 t_1 から t_4 と、時刻 t_5 から t_6 に姿勢検出部 232 が検出した作業者の姿勢が、作業姿勢であることを示している。

[0123] 「ドライバ検出」の右側の棒グラフは、時刻 t_2 から t_3 に、物体検出部 234 が、作業者の右手首または左手首の近傍にドライバを検出したことを示している。

「作業」の右側の棒グラフは、分析部 23 が推定した作業員の作業内容を示している。時刻 t_1 から t_4 は、作業としてネジ締結が推定されている。これは時刻 t_1 から t_4 の一部にて、作業員の手首の近傍にドライバを検出

したためである。これに対し、時刻 t 5 から t 6 までは、ドライバを検出していないためネジ締結とは推定されていない。

[0124] 図 4 1 A と図 4 1 B は、作業管理処理を示すフローチャートである。

画像加工部 2 0 2 は、動画を構成する各フレームを画像データとして取得する（ステップ S 1 1）。そして、画像加工部 2 0 2 は、認識したい複数の領域を新たな画像データ 2 1 3 に加工する（ステップ S 1 2）。そして、骨格抽出部 2 1 は、加工した画像データ 2 1 3 中の作業者の骨格データ 2 2 を抽出する（ステップ S 1 3）。

[0125] その後、分析部 2 3 は、並行に検出処理を実施する。具体的にいうと、領域検出部 2 3 1 は、領域モデル 2 4 に基づき、フレーム中の領域を検出する（ステップ S 1 4）。

姿勢検出部 2 3 2 は、姿勢モデル 2 5 に基づき、骨格データ 2 2 から姿勢を検出する（ステップ S 1 5）。ここで姿勢検出部 2 3 2 は、検出した姿勢のスコア（確からしさ）を算出する。スコアが閾値以上であれば、その姿勢を検出したと見做し、検出を示す文字列（ラベル）を出力する。

[0126] 背景検出部 2 3 3 は、背景モデル 2 9 1 に基づき、フレームから背景を検出する（ステップ S 1 6）。ここで背景検出部 2 3 3 は、検出した背景のスコア（確からしさ）を算出する。スコアが閾値以上であれば、その背景を検出したと見做し、検出を示す文字列（ラベル）を出力する。

[0127] 物体検出部 2 3 4 は、物体モデル 2 6 と骨格データ 2 2 に基づき、フレームから物体を検出する（ステップ S 1 7）。ここで物体検出部 2 3 4 は、検出した物体のスコア（確からしさ）を算出する。スコアが閾値以上であれば、その物体を検出したと見做し、検出を示す文字列（ラベル）を出力する。物体検知については、物体モデル 2 6 に特徴点を定義しているため、予め抽出した骨格データ 2 2 の座標に応じて物体検出の範囲を設定して検出する。

音検出部 2 3 5 は、音モデル 2 9 3 と音データ 2 9 6 に基づき、対応する作業内容を検出する（ステップ S 1 8）。移動検出部 2 3 6 は、移動分析モデル 2 9 4 と移動分析データ 2 9 5 に基づき、移動を検出する（ステップ S

19)。

[0128] これらの検出処理が終了すると、ステップS20に進む。

[0129] ステップS20にて、分析部23は、過去の検出回数によるフィルタ処理を実施して、検出した領域と姿勢と背景と物体のちらつきを抑制する。そして分析部23は、領域と姿勢と背景と物体の検出時間を算出する（ステップS21）。

そして、分析部23は、これらの算出した検出時間を基に、作業を推定する。具体的にいうと、分析部23は、これら検出結果の組み合わせ処理を実施する（ステップS22）。

[0130] 分析部23は、検出結果の組み合わせ処理において、組み合わせモデル292として定義したデータに基づき、算出した検出結果を組み合わせで演算し、作業員が実施している作業内容を推定する。

[0131] 組み合わせモデル292には、例えば部品取り領域に手が入ったことの検出と、部品を取る姿勢をしていることの検出との論理積と、開始作業との組み合わせが格納されている。更に組み合わせモデル292には、作業姿勢の検出と、ドライバを手首近傍における検出との論理積と、ねじ締結作業との組み合わせが格納されている。

[0132] 分析部23は、例えば、領域検出で部品が置かれている場所である部品取り領域に手が入ったことの検出と、姿勢検出で部品を取る姿勢をしていることの検出との論理積を行った結果、真であれば開始作業を行っていると思わず。

[0133] 図42は、作業姿勢とドライバとの関連付け処理を示すフローチャートである。

分析部23は、物体検出部234がドライバを検出したか否かを判定する（ステップS40）。分析部23は、物体検出部234がドライバを検出なかったならば（No）、図42の処理を終了し、物体検出部234がドライバを検出していたならば（Yes）、ステップS41に進む。

[0134] 次に分析部23は、姿勢検出部232が作業姿勢を検出したか否かを判定

する（ステップS 4 1）。分析部 2 3は、作業姿勢を検出しなかったならば（N o）、図 4 2の処理を終了し、作業姿勢を検出したならば（Y e s）、ステップS 4 2に進む。

[0135] ステップS 4 2にて、分析部 2 3は、作業姿勢の検出時間と同じ作業時間を、ねじ締め作業時間とすると、図 4 2の処理を終了する。分析部 2 3は、検出した結果について、論理演算での検出有無を算出することに加えて、作業時間についても、論理和または論理積での演算を行うことができる。

[0136] なお、物体検出部 2 3 4が検出するのはドライバに限られず、筆またはエアガン、半田ごて、スパナなどであってもよい。分析部 2 3が推定する作業内容は、ねじ締結に限られず、筆またはエアガンを検出したときの塗装作業、半田ごてを検出したときの半田づけ作業、スパナを検出したときのボルト締結作業などであってもよい。

[0137] 図 4 1 Bに戻り説明を続ける。分析部 2 3は、最新のフレームから推定した作業状態が開始作業で有るか否かを判定する（ステップS 2 3）。分析部 2 3は、開始作業ならば（Y e s）、作業時間の計測を開始して（ステップS 2 4）、ステップS 2 7に進む。分析部 2 3は、開始作業では無いならば（N o）、ステップS 2 5に進む。

[0138] ステップS 2 5にて、分析部 2 3は、最新のフレームから推定した作業状態が終了作業で有るか否かを判定する。分析部 2 3は、終了作業ならば（Y e s）、作業時間の計測を終了して（ステップS 2 6）、ステップS 2 7に進む。分析部 2 3は、終了作業では無いならば（N o）、ステップS 2 7に進む。

[0139] ステップS 2 7にて、分析部 2 3は、作業順序を確認する。例えば分析部 2 3は、予め定められた作業順序から逸脱した作業について、その旨を作業属性として記録するとよい。これにより、分析部 2 3は、予め定められた作業順序に沿った作業と、予め定められた作業順序から逸脱した作業を区別して表示することができる。

[0140] そして分析部 2 3は、途中結果を出力する（ステップS 2 8）。そして、

分析部 23 は、未処理のフレームが有るか否かを判定する（ステップ S 29）。分析部 23 は、未処理のフレームが有るならば（Yes）、ステップ S 11 の処理に戻り、全てのフレームを処理したならば（No）、ステップ S 30 に進む。

ステップ S 30 にて、分析部 23 は、推定結果データ 27 を出力すると、図 41A と図 41B の処理を終了する。

図 41A と図 41B のフローチャートで説明したように、分析部 23 は、ドライバの検出結果と作業者の姿勢の検出結果とに基づき、この作業者が行っている作業を推定する。これにより、作業員が行う複数の異なる作業が同様な姿勢であっても、各作業内容を好適に抽出できる。

[0141] 図 43 は、作業管理装置 2 が表示するリアルタイム画面 5 を示す図である。

図 43 に示したリアルタイム画面 5 は、分析の実行、中断、停止と、分析した結果を示す画面である。出力部 28 が、このリアルタイム画面 5 をモニタ 31 に表示する。リアルタイム画面 5 は、実行ボタン 51 と、ポーズボタン 52 と、停止ボタン 53 と、映像表示領域 56 と、状態表示欄 54 と、状態累計時間グラフ 55 と、状態タイムチャート 57 とを含んで表示される。

[0142] 映像表示領域 56 は、入力した画像データに分析結果を重畳して表示する領域である。実行ボタン 51 は、分析の実行を行うためのものである。ポーズボタン 52 は、分析の実行にポーズを掛けるものである。停止ボタン 53 は、分析の実行を停止させるものである。

[0143] 状態表示欄 54 は、現在の分析状況を示す欄であり、作業の開始、終了や主な作業の実施状況を表示する。状態累計時間グラフ 55 は、分析中の映像における作業の累計時間を示す棒グラフである。状態タイムチャート 57 は、直近で分析した結果をタイムチャートで示す。

[0144] 図 44 は、作業管理装置 2 が表示する履歴データ画面 6 を示す図である。履歴データ画面 6 は、分析結果を示す各タイムチャートを含んでいる。状態タイムチャート 60 は、各時間における分析結果のスコア（確からし

さ)に基づいて、総合的に判定した状態を示すタイムチャートである。

[0145] 「開始」タイムチャート61は、各時間における映像の被写体の「開始」状態のスコアに基づいて判定した状態を示すタイムチャートである。ここで分析部23は、被写体の骨格が部品取り領域901と重なっており、かつ未だ作業が開始されていないときに、「開始」状態のスコアが高いと判定する。

[0146] 「部品を取る」タイムチャート62は、各時間における映像の被写体の「部品を取る」状態のスコアに基づいて判定した状態を示すタイムチャートである。ここで分析部23は、被写体の骨格が部品取り領域901と重なっており、かつ作業が開始されているときに、「開始」状態のスコアが高いと判定する。

[0147] 「ドライバ使用中」タイムチャート63は、各時間における映像の被写体の「ドライバ使用中」状態のスコアに基づいて判定した状態を示すタイムチャートである。

[0148] 「ドライバ未使用」タイムチャート64は、各時間における映像の被写体の「ドライバ未使用」状態のスコアに基づいて判定した状態を示すタイムチャートである。

「ネジ締め」タイムチャート65は、各時間における映像の被写体の「ネジ締め」状態のスコアに基づいて判定した状態を示すタイムチャートである。

[0149] 「右向き」タイムチャート66は、各時間における映像の被写体の「右向き」状態のスコアに基づいて判定した状態を示すタイムチャートである。

「左向き」タイムチャート67は、各時間における映像の被写体の「左向き」状態のスコアに基づいて判定した状態を示すタイムチャートである。

[0150] 「終了」タイムチャート68は、各時間における映像の被写体の「終了」状態のスコアに基づいて判定した状態を示すタイムチャートである。

「部品を置く」タイムチャート69は、各時間における映像の被写体の「部品を置く」状態のスコアに基づいて判定した状態を示すタイムチャートで

ある。

[0151] 図45は、作業管理装置2が表示する認識レポート画面7を示す図である。

認識レポート画面7には、凡例70と、ラベル71と、積み上げ棒グラフ72と、場所欄73と、日時欄74とを含んで表示している。

[0152] 積み上げ棒グラフ72は、分析部23が認識した姿勢に基づく作業の累計時間を示すものである。ここで4本の積み上げ棒グラフ72は、9時から10時30分、10時50分から12時20分、13時20分から14時50分、15時20分から16時50分の各時間帯における各作業の累計時間を示している。

[0153] ラベル71は、積み上げ棒グラフ72に表示する作業を選択するものである。凡例70は、積み上げ棒グラフ72に表示された表示態様と、その表示態様が示す作業との関係を示すものである。場所欄73は、認識レポート画面7の対象となる場所を示すものである。日時欄74は、認識レポート画面7の対象となる日時を示すものである。ここでは4本の積み上げ棒グラフにより、それぞれの日における作業の累計時間を示している。棒グラフは場所や年月日ごとなど、指定する単位で表示できる。

[0154] この認識レポート画面7によれば、作業の管理者は、各作業者の作業の累計時間を容易に把握できる。

[0155] 以下に本発明の構成と作用効果について説明する。

[0156] [1] 人物が撮影された動画の各フレームを、人物領域と物体領域を含む画像データ(213)に加工する画像加工部(202)と、

前記画像データ(213)から当該人物の骨格データ(22)を抽出する骨格抽出部(21)と、

前記骨格データ(22)から前記人物の姿勢を検出する姿勢検出部(232)と、

前記人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部(23)と、

を有することを特徴とする作業管理装置（２）。

[0157] これにより、処理対象の画像データの情報量が少なくなるので、骨格抽出部（２１）は、高い解像度の画像をリアルタイムに処理可能となる。

[0158] [２] 前記骨格抽出部（２１）は、前記画像データ（２１３）のうち人物を認識したい領域から当該人物の骨格データ（２２）を抽出する、
ことを特徴とする[１]に記載の作業管理装置。

[0159] 動画の各フレームである画像データ（２１０）が高解像度であっても演算処理の負担は少なくなり、骨格抽出部（２１）は、リアルタイム処理性に優れた処理を提供可能である。

[0160] [３] 前記画像加工部（２０２）は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、人物を認識したい領域を一方に配置し、物体を認識したい領域を他方に配置して合成することで、前記画像データ（２１３）に加工する、
ことを特徴とする[１]に記載の作業管理装置。

[0161] 画像データ（２１３）の一方から人物を認識し、画像データ（２１３）の他方から物体を認識することができる。このとき演算処理の負担は少なく、リアルタイム処理性に優れる。

[0162] [４] 前記画像データ（２１３）のうち、物体を認識したい領域が配置された部分から所定の物体を検出する物体検出部（２３４）、
を更に備えることを特徴とする[３]に記載の作業管理装置。

[0163] 物体検出部（２３４）は、予め定められた物体を認識したい領域から定められた物体を検出しているので、演算処理の負担は少なく、リアルタイム処理性に優れる。

[0164] [５] 前記画像加工部（２０２）は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、余白を入れて合成することで前記画像データ（２１３）に加工する、
ことを特徴とする[１]に記載の作業管理装置。

[0165] これにより、物体検出部（２３４）は、各領域の継ぎ目の部分の物体認識

の誤検知を抑制できる。更に骨格抽出部（21）は、骨格抽出の誤検知を抑制できる。

[0166] [6] 前記画像加工部（202）は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、リサイズして合成することで前記画像データ（213）に加工する、

ことを特徴とする[1]に記載の作業管理装置。

[0167] 認識したい対象物に応じて各領域をリサイズするので、画像データ（210）が高解像度であっても、物体検出部（234）や骨格抽出部（21）の演算処理の負担は少なく、リアルタイム処理性に優れる。

[0168] [7] 前記画像加工部（202）は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、フィルタ処理して合成することで前記画像データ（213）に加工する、

ことを特徴とする[1]に記載の作業管理装置。

[0169] これにより、物体認識の誤検知や骨格抽出の誤検知を抑制できる。

[0170] [8] 前記画像加工部（202）は、前記動画の各フレームから人物を認識したい領域と物体を認識したい領域を切り出す位置を、パターンマッチングで調整する、

ことを特徴とする[1]に記載の作業管理装置。

[0171] これにより、ビデオカメラ（1）の設置位置のズレや、作業者の作業位置のズレを調整可能である。

[0172] [9] 前記画像加工部（202）は、前記動画の各フレームから人物を認識したい領域と物体を認識したい領域を切り出す位置を、マーカで調整する、

ことを特徴とする[1]に記載の作業管理装置。

[0173] これにより、ビデオカメラ（1）の設置位置のズレや、作業者の作業位置のズレを調整可能である。

[0174] [10] 前記画像加工部（202）が前記動画の各フレームから人物を認識したい領域と物体を認識したい領域を切り出す位置を、手作業で調整可能な画面を表示させる調整部、

を備えることを特徴とする [1] に記載の作業管理装置。

[0175] これにより、ビデオカメラ (1) の設置位置のズレや、作業者の作業位置のズレを調整可能である。

[0176] [1 1] 前記姿勢検出部 (2 3 2) による検出の結果に基づき、前記骨格抽出部 (2 1) が抽出した前記骨格データ (2 2) の特徴点データを、前記画像データ (2 1 3) に加工する前のフレームにマーキングして表示部に表示させる出力部 (2 8)、

を備えることを特徴とする [1] に記載の作業管理装置。

[0177] これにより、ビデオカメラ (1) で撮影されている作業場所の画像に、骨格データ (2 2) の特徴点データを表示させることができる。

[0178] [1 2] 各骨格データに姿勢ラベルが対応づけられている姿勢モデルの記憶部を更にそなえ、

前記姿勢検出部 (2 3 2) は、前記姿勢モデルに基づき、前記人物の骨格データから前記人物の姿勢を検出する、

ことを特徴とする [1] に記載の作業管理装置。

[0179] これにより、姿勢検出部 (2 3 2) は、作業者の作業に対応した固有の姿勢を適切に検出可能である。

[0180] [1 3] 前記分析部 (2 3) は更に、領域、背景、物体、音、移動、前記人物の姿勢の検出結果を単独、もしくは組み合わせて人物の作業を推定する、

ことを特徴とする請求項 1 に記載の作業管理装置。

[0181] これにより、姿勢の検出結果と他の要素との組み合わせに基づき、人物の作業をより正確に推定可能である。

[0182] [1 4] 人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データ (2 1 3) に加工する画像加工部 (2 0 2) と、

前記画像データ (2 1 3) から当該人物の骨格データ (2 2) を抽出する骨格抽出部 (2 1) と、

前記骨格データ (2 2) から前記人物の姿勢を検出する姿勢検出部 (2 3

2) と、

前記人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部 (23) と、

を有することを特徴とする作業管理システム。

[0183] これにより、処理対象の画像データの情報量が少なくなるので、骨格抽出部が高い解像度の画像をリアルタイムに処理可能となる。

[0184] [15] 画像加工部 (202) が、人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データ (213) に加工するステップと、

骨格抽出部 (21) が、前記画像データ (213) から当該人物の骨格データ (22) を抽出するステップと、

姿勢検出部 (232) が、前記骨格データ (22) から前記人物の姿勢を検出するステップと、

分析部 (23) が、前記人物の姿勢の検出結果に基づき、前記人物の作業を推定するステップと、

を有することを特徴とする作業管理方法。

[0185] これにより、処理対象の画像データの情報量が少なくなるので、骨格抽出部が高い解像度の画像をリアルタイムに処理可能となる。

[0186] 《変形例》

本発明は上記した実施形態に限定されるものではなく、様々な変形例が含まれる。例えば上記した実施形態は、本発明を分かりやすく説明するために詳細に説明したものであり、必ずしも説明した全ての構成を備えるものに限定されるものではない。ある実施形態の構成の一部を他の実施形態の構成に置き換えることが可能であり、ある実施形態の構成に他の実施形態の構成を加えることも可能である。また、各実施形態の構成の一部について、他の構成の追加・削除・置換をすることも可能である。

[0187] 上記の各構成、機能、処理部、処理手段などは、それらの一部または全部を、例えば集積回路などのハードウェアで実現してもよい。上記の各構成、

機能などは、プロセッサがそれぞれの機能を実現するプログラムを解釈して実行することにより、ソフトウェアで実現してもよい。各機能を実現するプログラム、テーブル、ファイルなどの情報は、メモリ、ハードディスク、SSD (Solid State Drive) などの記録装置、または、フラッシュメモリカード、DVD (Digital Versatile Disk) などの記録媒体に置くことができる。

[0188] 各実施形態に於いて、制御線や情報線は、説明上必要と考えられるものを示しており、製品上必ずしも全ての制御線や情報線を示しているとは限らない。実際には、殆ど全ての構成が相互に接続されていると考えてもよい。

符号の説明

- [0189] 1 ビデオカメラ
- 1 3 ビデオレコーダ
 - 1 4 入出力装置
 - 2 作業管理装置
 - 2 1 骨格抽出部
 - 2 2 骨格データ
 - 2 1 1 モデル生成部
 - 2 0 2 画像加工部
 - 2 1 3 画像データ
 - 2 3 1 a 領域定義部
 - 2 3 2 a 姿勢学習部
 - 2 3 3 a 背景定義部
 - 2 3 7 a 手順学習部
 - 2 3 分析部
 - 2 3 1 領域検出部
 - 2 3 2 姿勢検出部
 - 2 3 3 背景検出部
 - 2 3 4 物体検出部

- 2 3 5 音検出部
- 2 3 6 移動検出部
- 2 3 7 手順検出部
- 2 4 領域モデル
- 2 5 姿勢モデル
- 2 6 物体モデル
- 2 2 骨格データ
- 2 7 推定結果データ
- 2 8 出力部
- 3 1 モニタ
- 3 2 記憶装置
- 3 3 アプリケーション装置
- 9 0 1 部品取り領域
- 9 0 3 工具領域
- 9 0 2 完成品格納領域
- 9 1 0 人物領域

請求の範囲

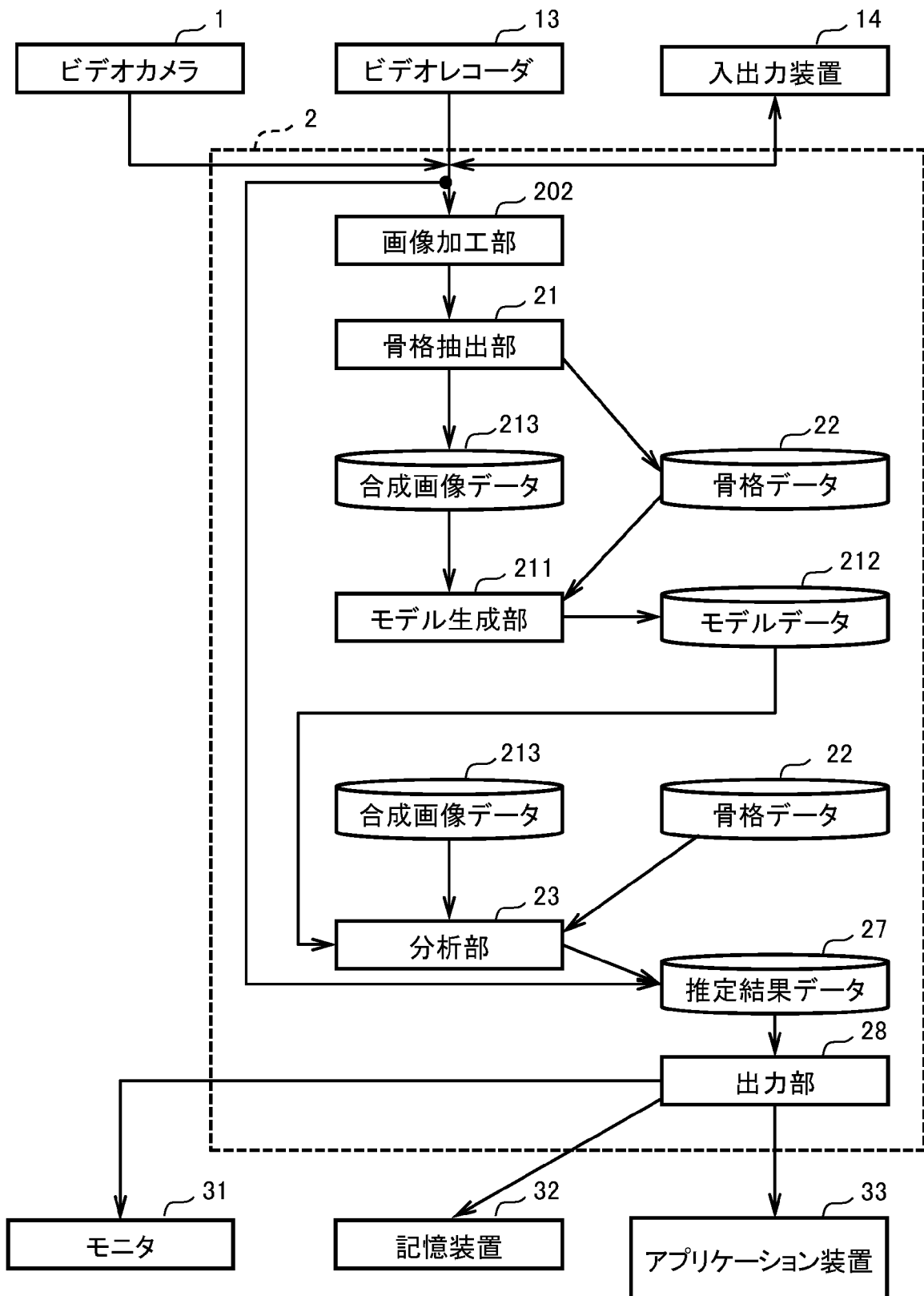
- [請求項1] 人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データに加工する画像加工部と、
前記画像データから当該人物の骨格データを抽出する骨格抽出部と、
、
前記骨格データから前記人物の姿勢を検出する姿勢検出部と、
前記人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部と、
を有することを特徴とする作業管理装置。
- [請求項2] 前記骨格抽出部は、前記画像データのうち人物を認識したい領域から当該人物の骨格データを抽出する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項3] 前記画像加工部は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、人物を認識したい領域を一方に配置し、物体を認識したい領域を他方に配置して合成することで、前記画像データに加工する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項4] 前記画像データのうち、物体を認識したい領域が配置された部分から所定の物体を検出する物体検出部、
を更に備えることを特徴とする請求項3に記載の作業管理装置。
- [請求項5] 前記画像加工部は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、余白を入れて合成することで前記画像データに加工する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項6] 前記画像加工部は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、リサイズして合成することで前記画像データに加工する、
ことを特徴とする請求項1に記載の作業管理装置。

- [請求項7] 前記画像加工部は、前記動画の各フレームから、人物を認識したい領域と物体を認識したい領域を切り出し、フィルタ処理して合成することで前記画像データに加工する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項8] 前記画像加工部は、前記動画の各フレームから人物を認識したい領域と物体を認識したい領域を切り出す位置を、パターンマッチングで調整する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項9] 前記画像加工部は、前記動画の各フレームから人物を認識したい領域と物体を認識したい領域を切り出す位置を、マーカで調整する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項10] 前記画像加工部が前記動画の各フレームから人物を認識したい領域と物体を認識したい領域を切り出す位置を、手作業で調整可能な画面を表示させる調整部、
を備えることを特徴とする請求項1に記載の作業管理装置。
- [請求項11] 前記姿勢検出部による検出の結果に基づき、前記骨格抽出部が抽出した前記骨格データの特徴点データを、前記画像データに加工する前のフレームにマーキングして表示部に表示させる出力部、
を備えることを特徴とする請求項1に記載の作業管理装置。
- [請求項12] 各骨格データに姿勢ラベルが対応づけられている姿勢モデルの記憶部を更にそなえ、
前記姿勢検出部は、前記姿勢モデルに基づき、前記人物の骨格データから前記人物の姿勢を検出する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項13] 前記分析部は更に、領域、背景、物体、音、移動、前記人物の姿勢の検出結果を単独、もしくは組み合わせて人物の作業を推定する、
ことを特徴とする請求項1に記載の作業管理装置。
- [請求項14] 人物が撮影された動画の各フレームを、人物を認識したい領域と物

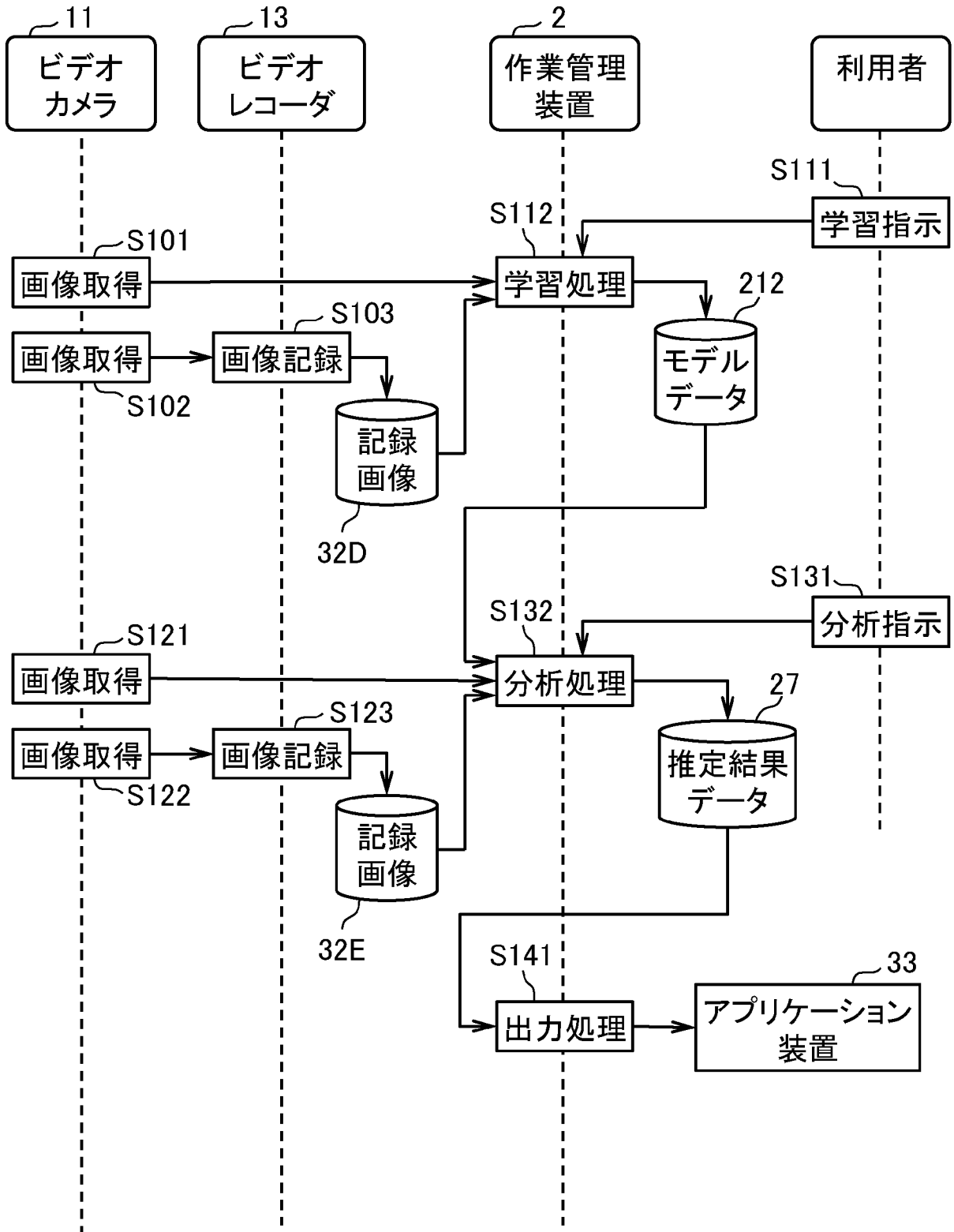
体を認識したい領域を含む画像データに加工する画像加工部と、
前記画像データから当該人物の骨格データを抽出する骨格抽出部と、
、
前記骨格データから前記人物の姿勢を検出する姿勢検出部と、
前記人物の姿勢の検出結果に基づき、前記人物の作業を推定する分析部と、
を有することを特徴とする作業管理システム。

[請求項15] 画像加工部が、人物が撮影された動画の各フレームを、人物を認識したい領域と物体を認識したい領域を含む画像データに加工するステップと、
骨格抽出部が、前記画像データから当該人物の骨格データを抽出するステップと、
姿勢検出部が、前記骨格データから前記人物の姿勢を検出するステップと、
分析部が、前記人物の姿勢の検出結果に基づき、前記人物の作業を推定するステップと、
を有することを特徴とする作業管理方法。

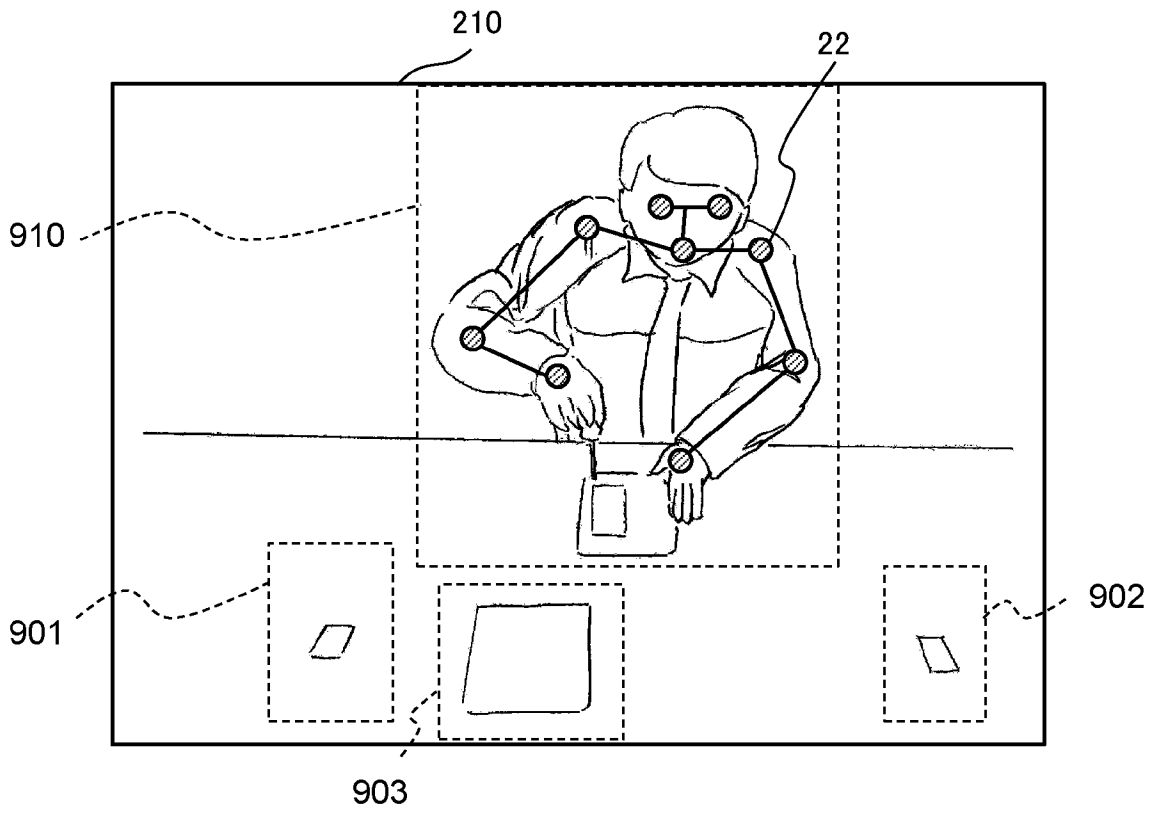
[図1]



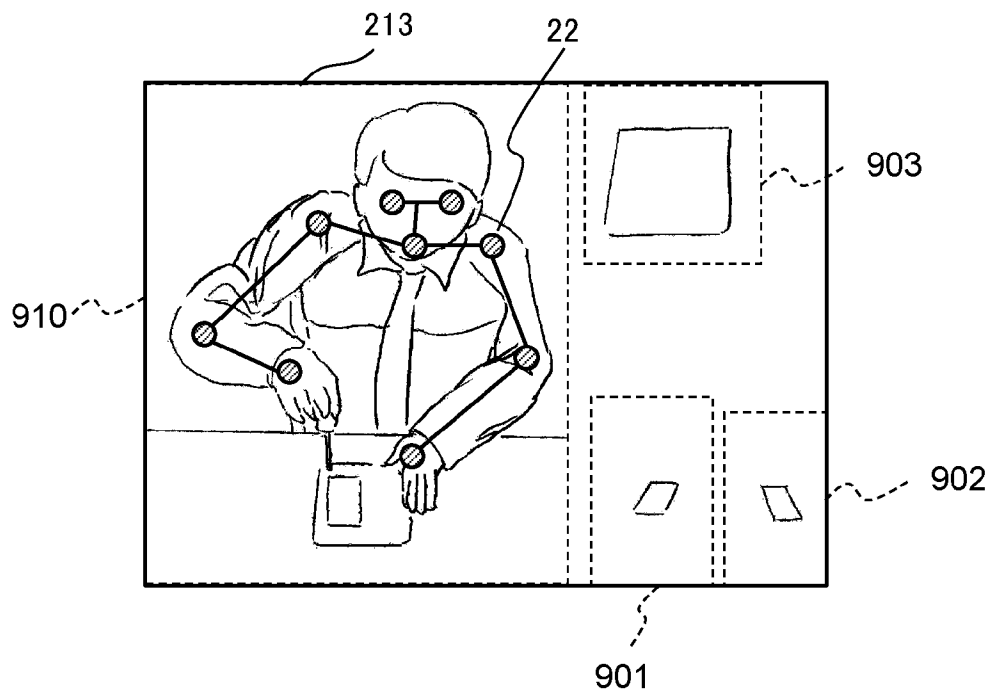
[図2]



[図3]



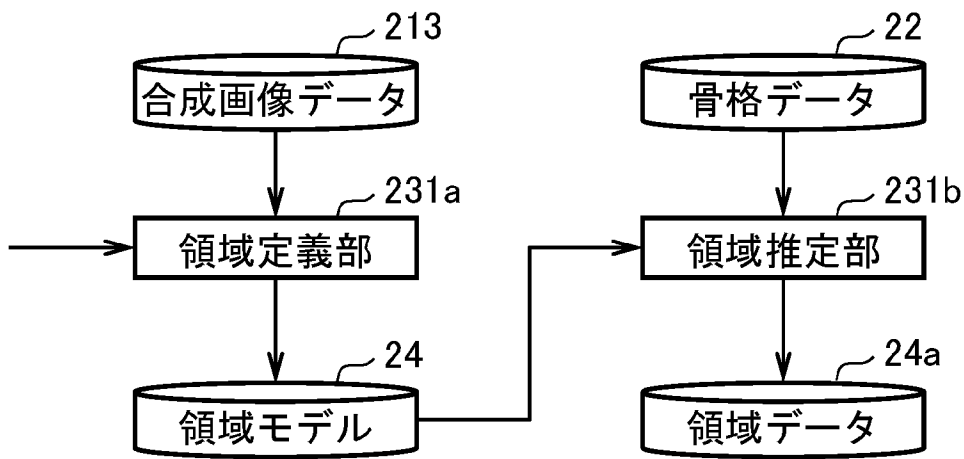
[図4]



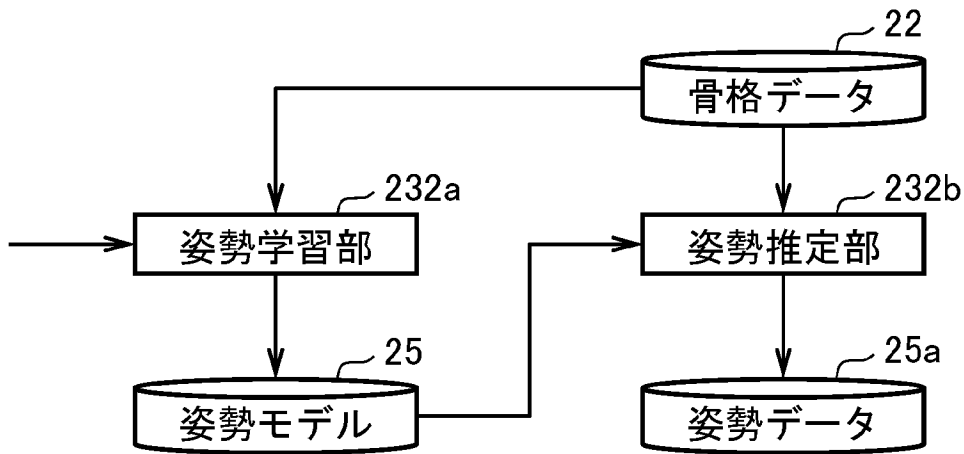
[図5]

番号	特徴点	座標(x軸)	座標(y軸)
0	鼻	240	161
1	首	241	189
2	右肩	150	191
3	右肘	141	251
4	右手首	142	300
5	左肩	350	190
6	左肘	255	253
7	左手首	249	315
8	右目	225	100
9	左目	255	140
:	:	:	:

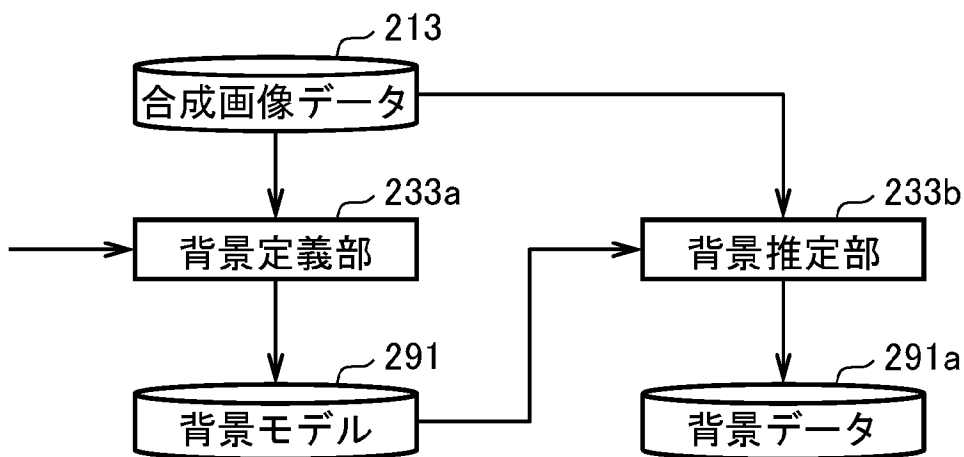
[図6]



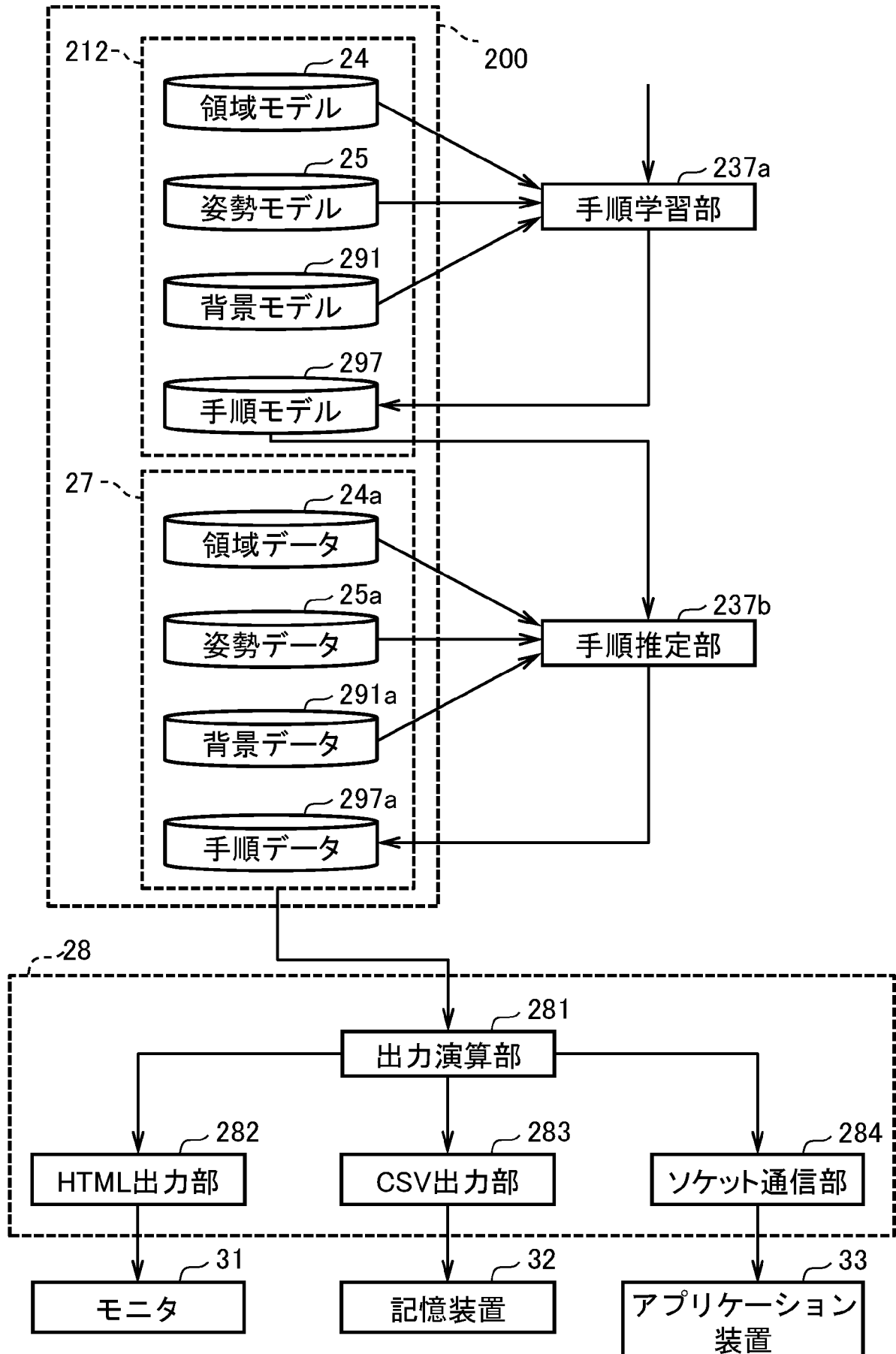
[図7]



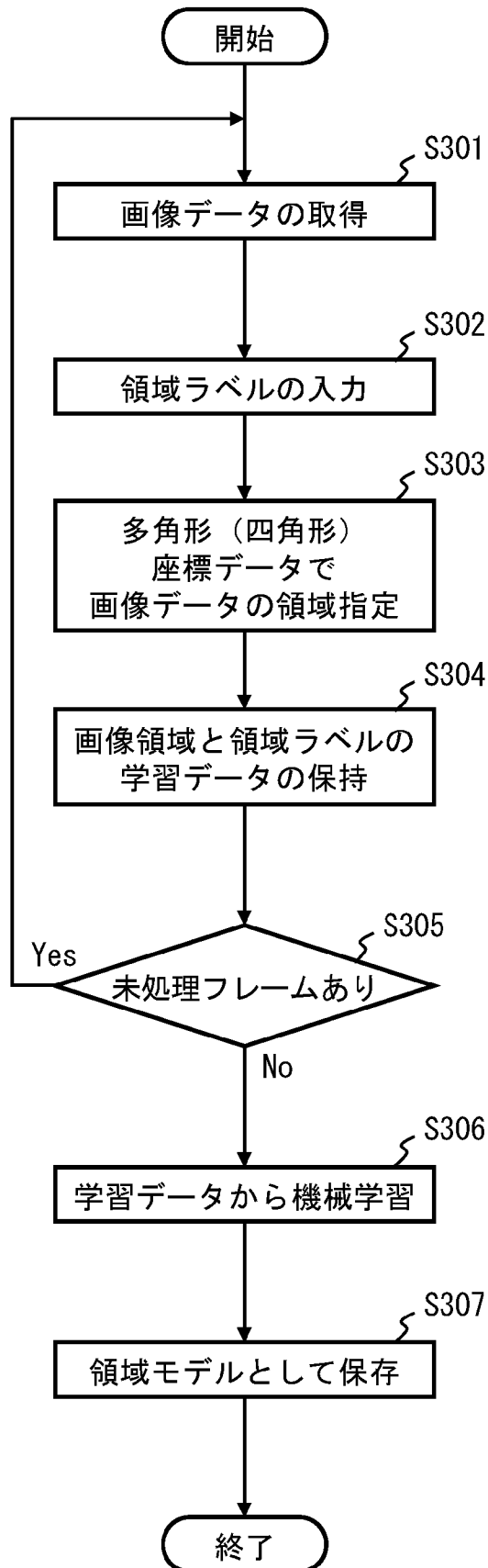
[図8]



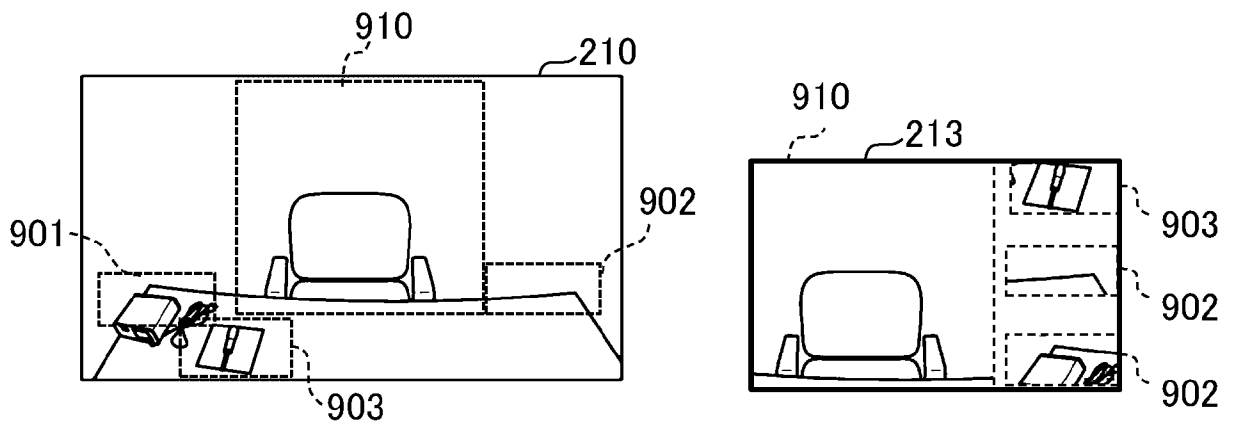
[図9]



[図10]



[図11]

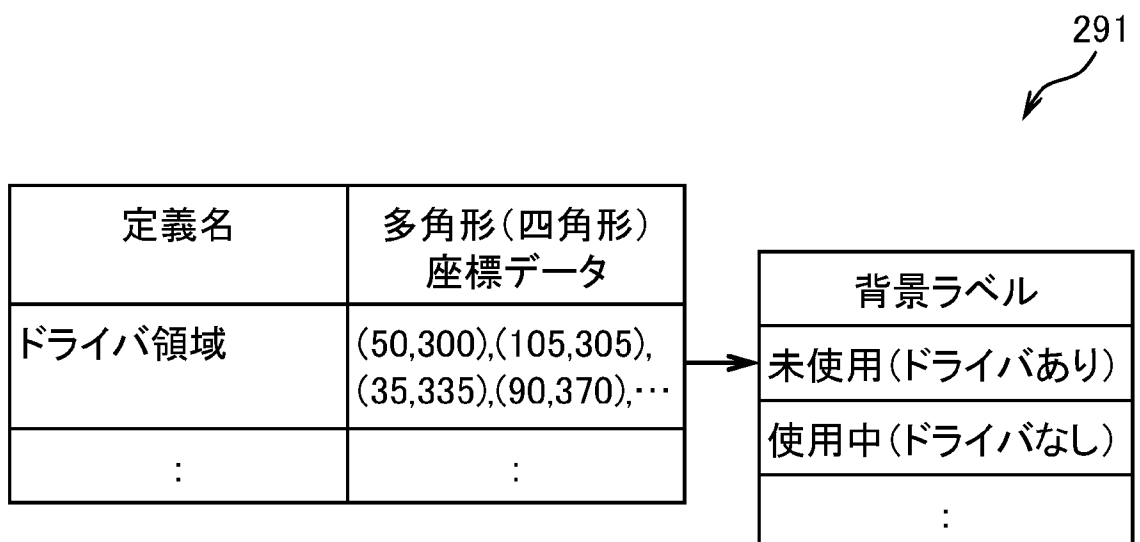


[図12]

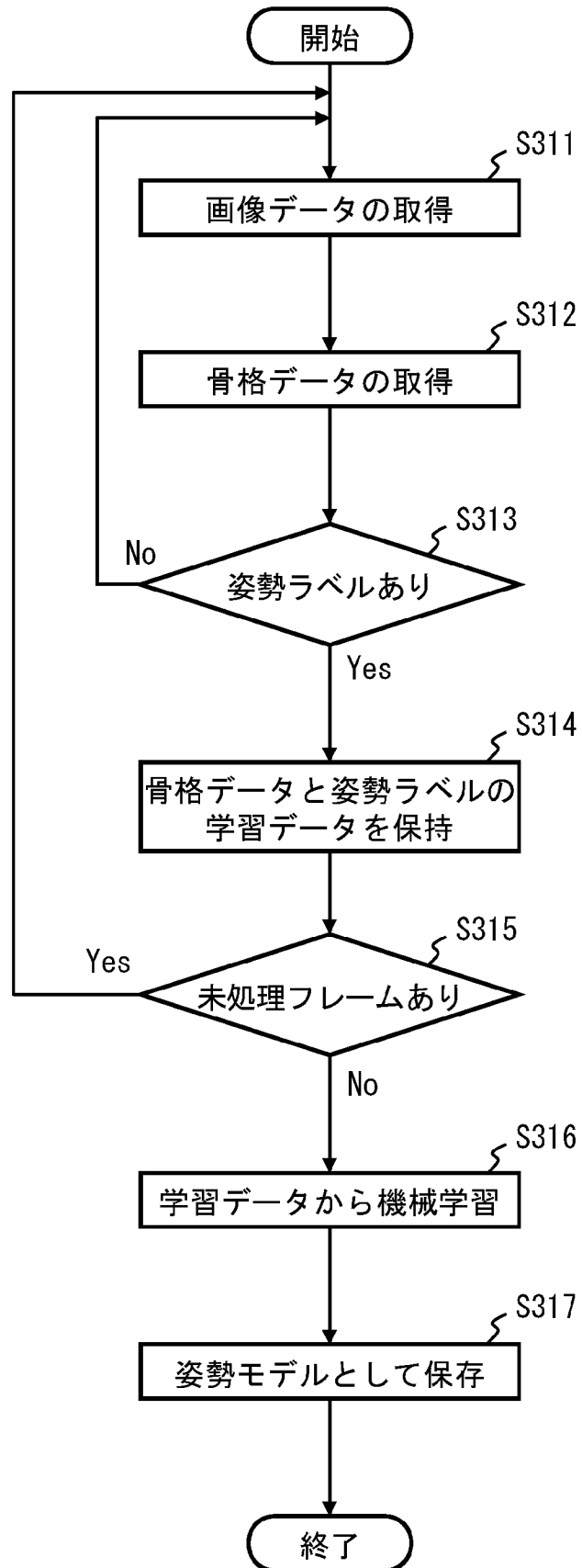
24

領域ラベル	特徴点番号	判定論理	多角形(四角形) 座標データ
部品取り領域	4,7	AND	(50,300),(105,305), (35,335),(90,370),...
完成品格納領域	4,7	AND	(850,307),(905,322), (835,356),(990,380),...
:	:	:	:

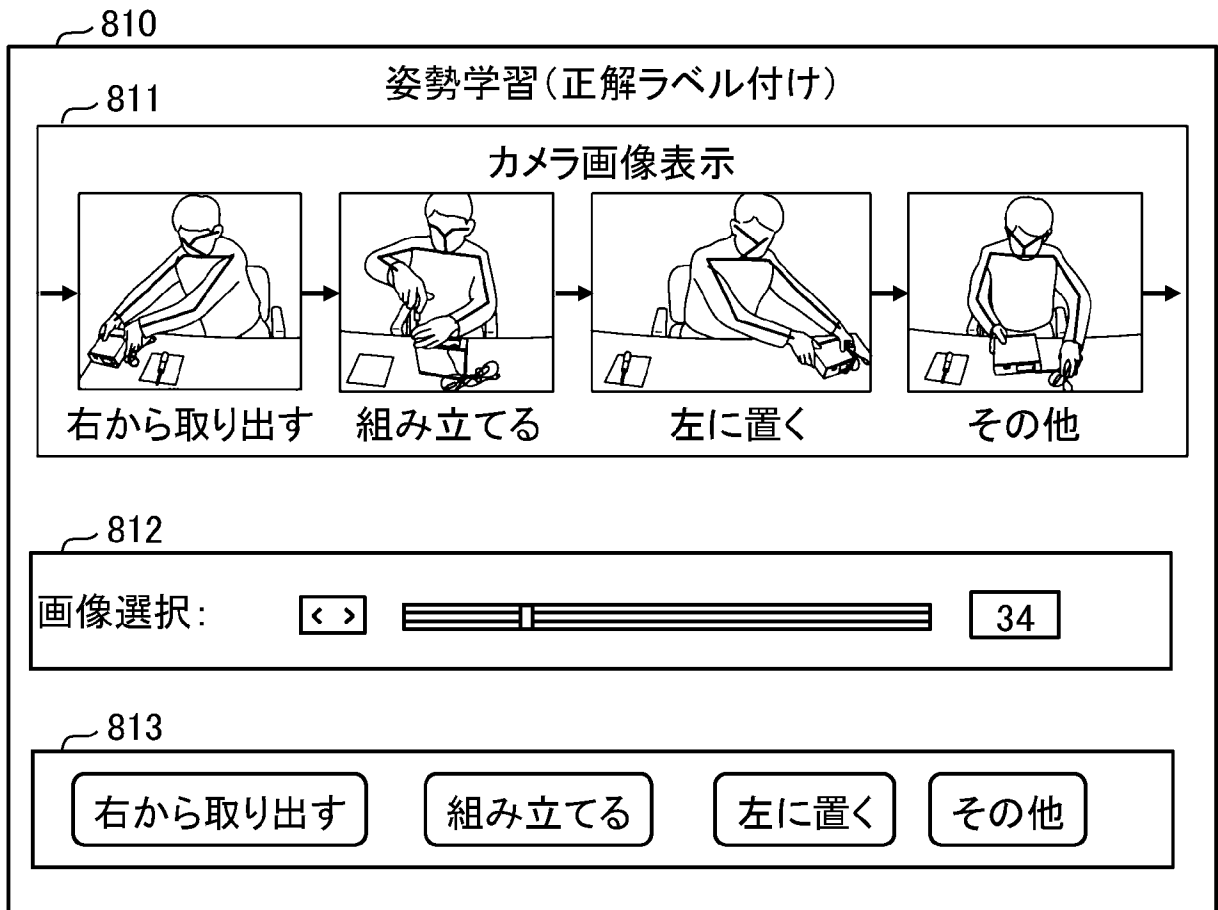
[図13]



[図14]



[図15]



[図16]

25

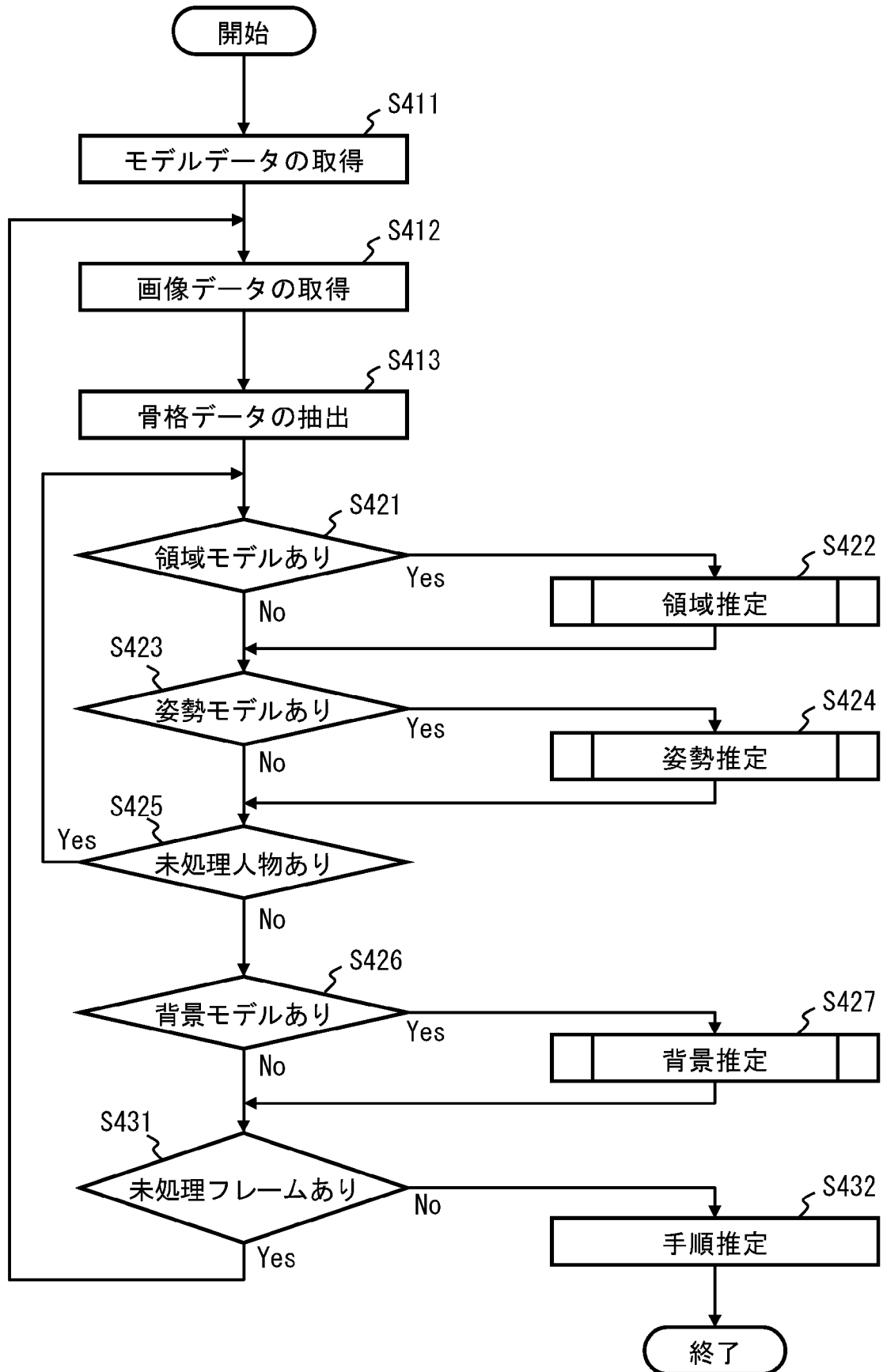
フレーム番号	正解ラベル	骨格データ
1	その他	150, 200, 199, 203, 647...
:	:	:
:	右から取り出す	:
:	:	:
:	組み立てる	:
:	:	:
:	左に置く	:
:	:	:

[図17]

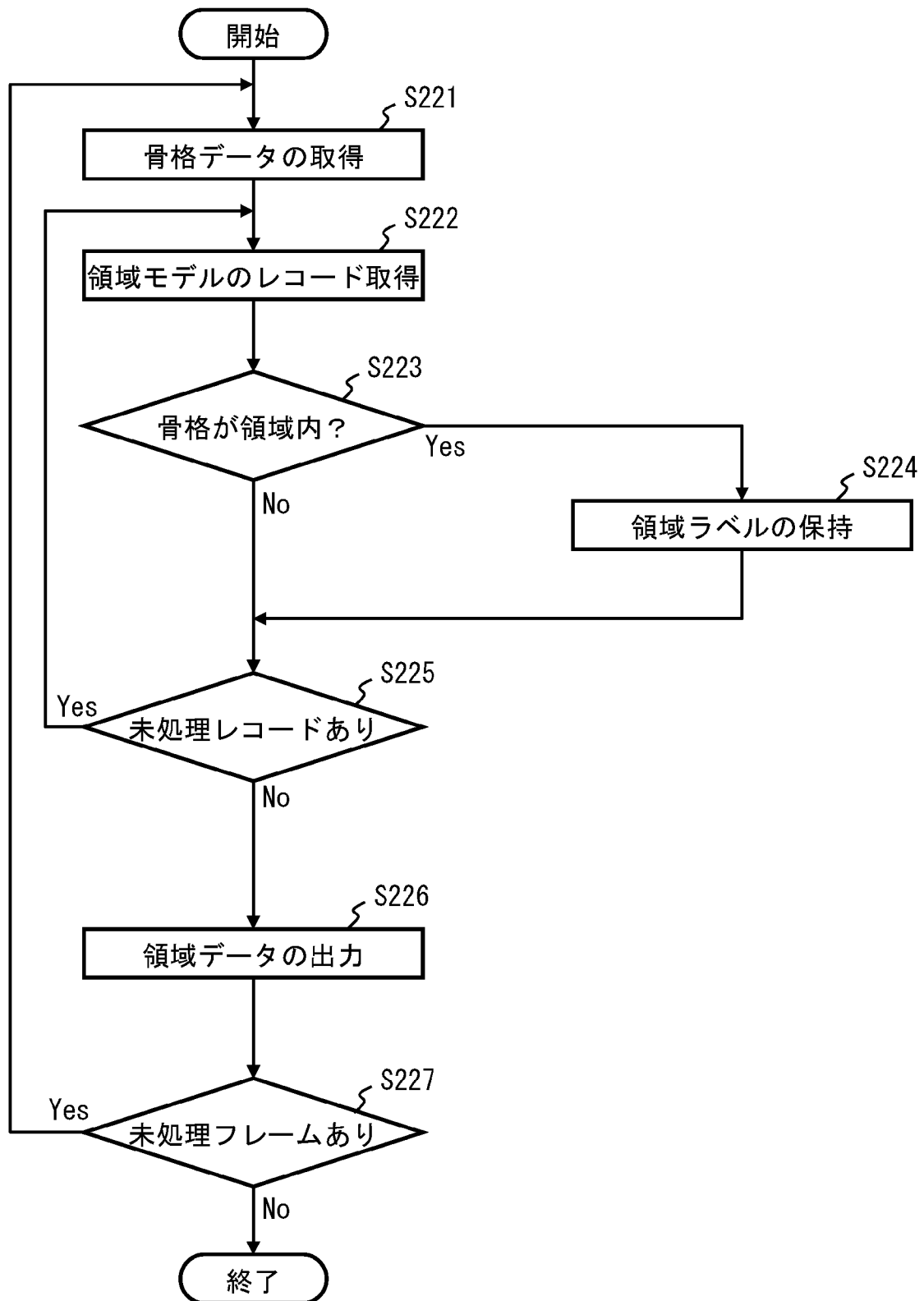
297

入力モデル			出力手順
領域モデル	姿勢モデル	背景モデル	
未検出	組み立てる	使用中 (ドライバなし)	組み立て (ねじ締め付け)
同上	同上	未使用 (ドライバあり)	組み立て (その他)
同上	組み立てる、以外	N/A	部品待ち・ 間作業
右側に両手首入る	右から取り出す	N/A	部品取り
同上	右から取り出す、以外	N/A	部品待ち・ 間作業
左側に両手首入る	左に置く	N/A	部品格納
同上	左に置く、以外	N/A	部品待ち・ 間作業

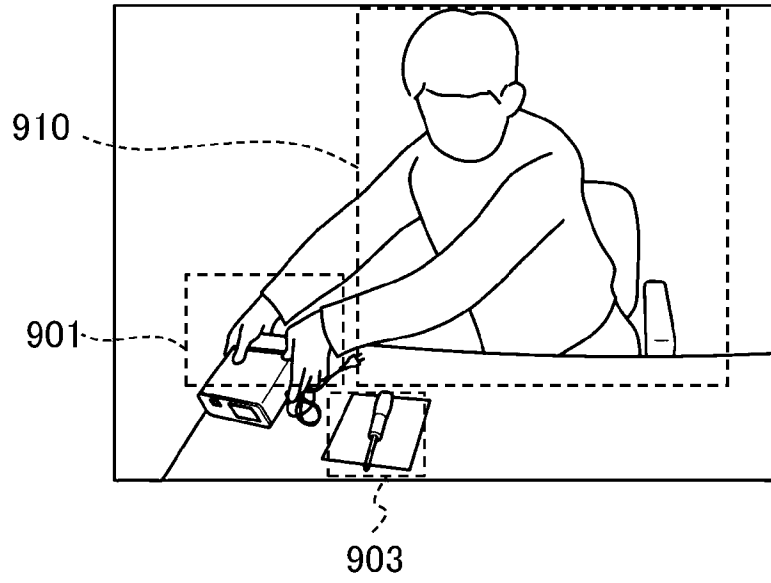
[図18]



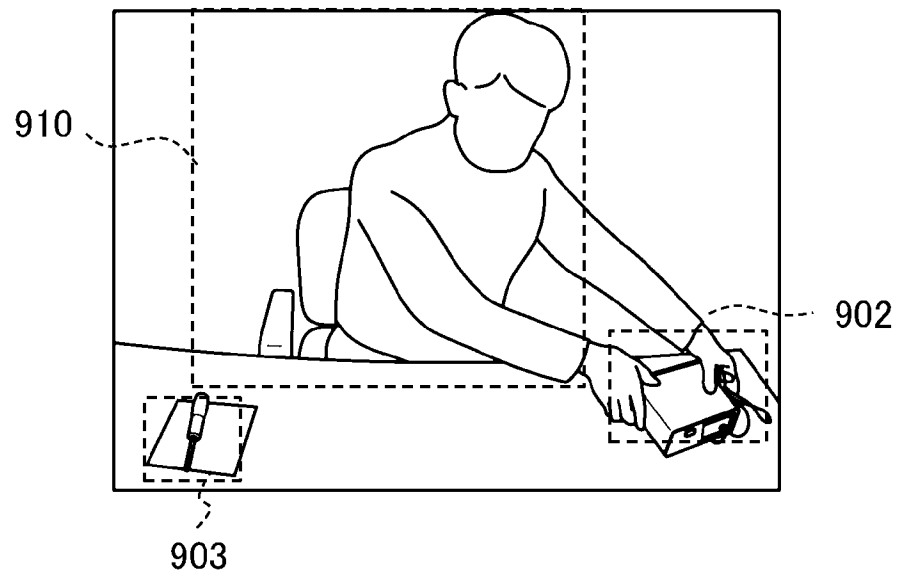
[図19]



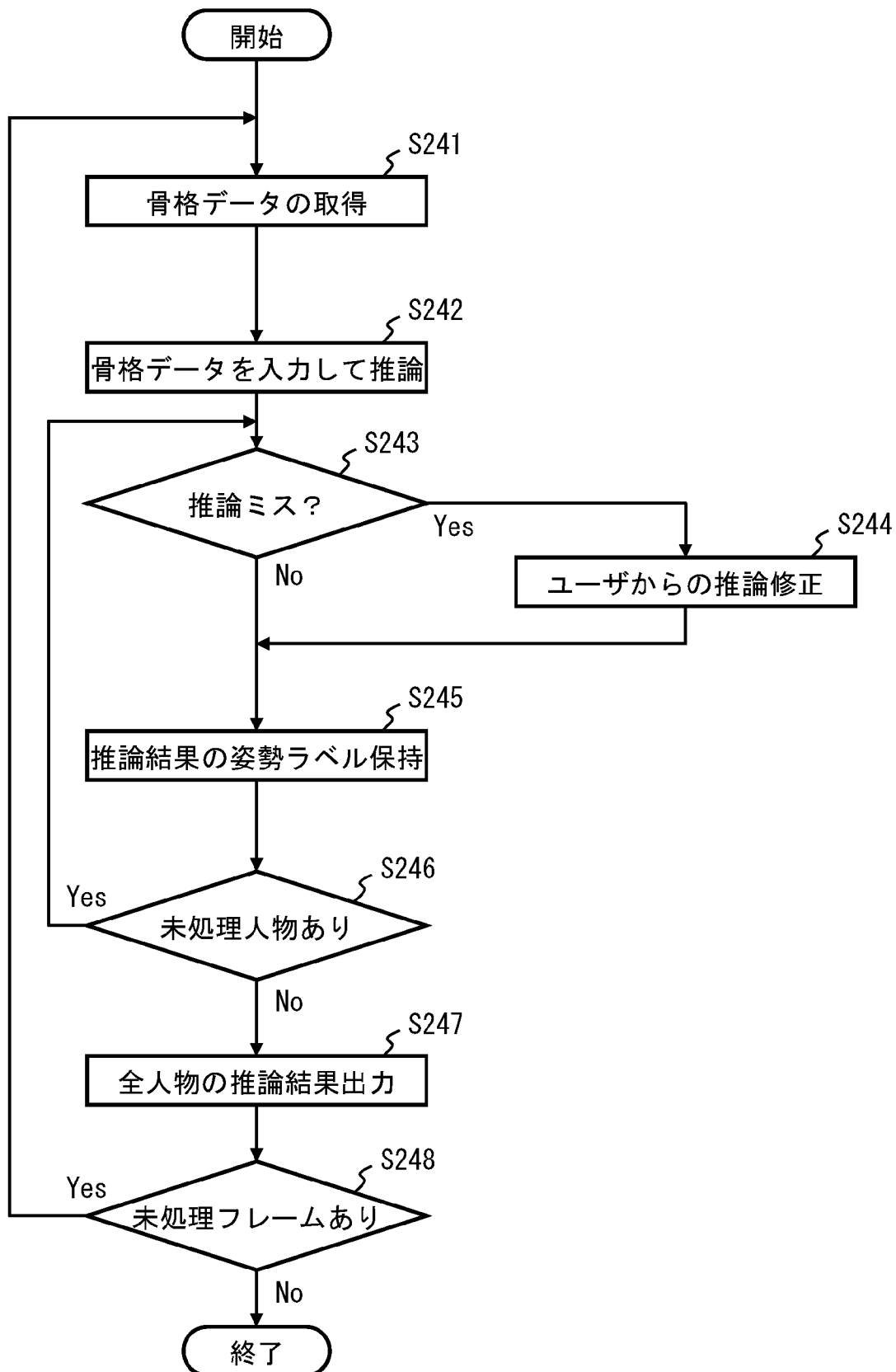
[図20]



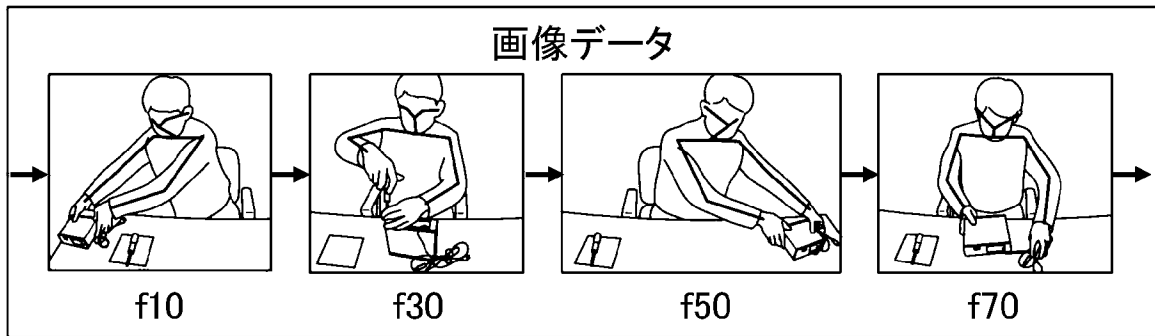
[図21]



[図22]



[図23]

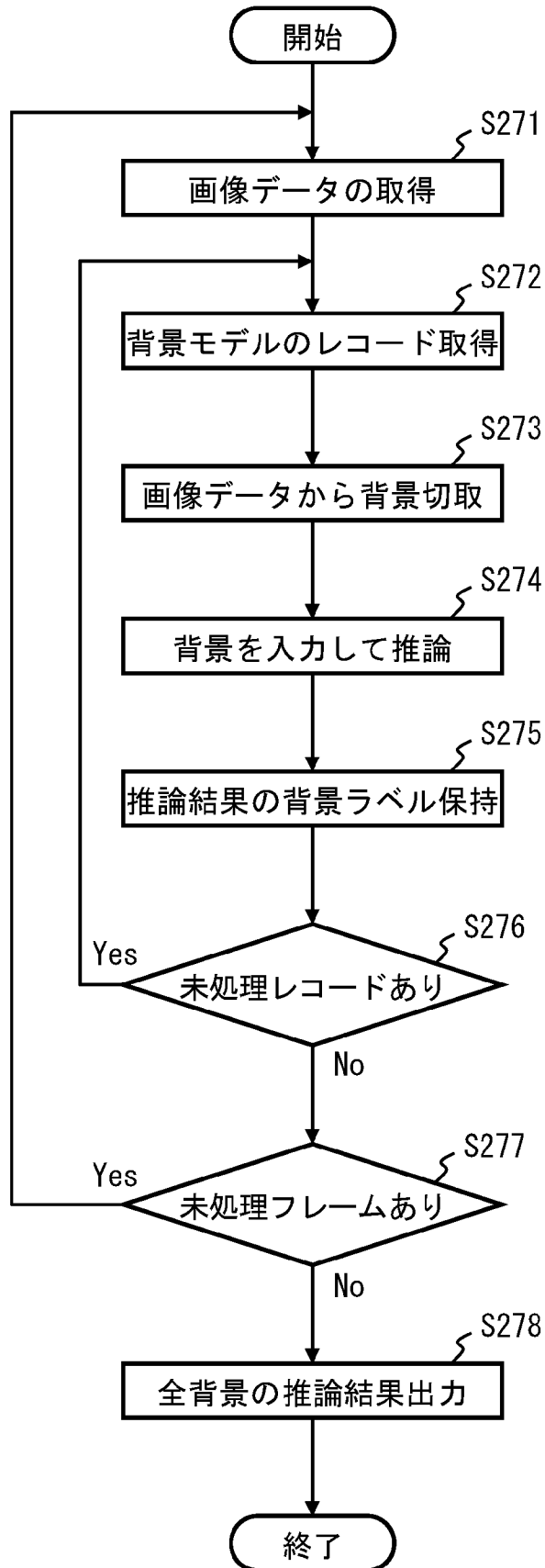


[図24]

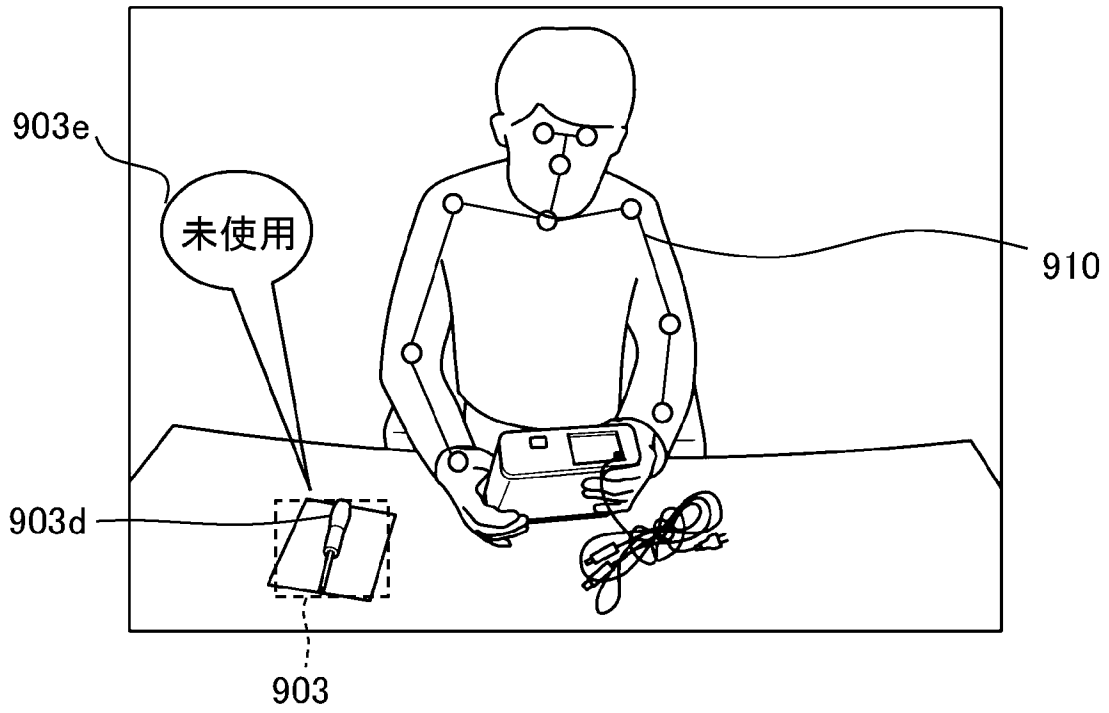
25a

フレーム番号	推論ラベル
1	その他
:	:
10	右から取り出す
:	:
:	:
:	:
30	組み立てる
:	:
:	その他
:	:
50	左に置く
:	:
70	その他
:	:

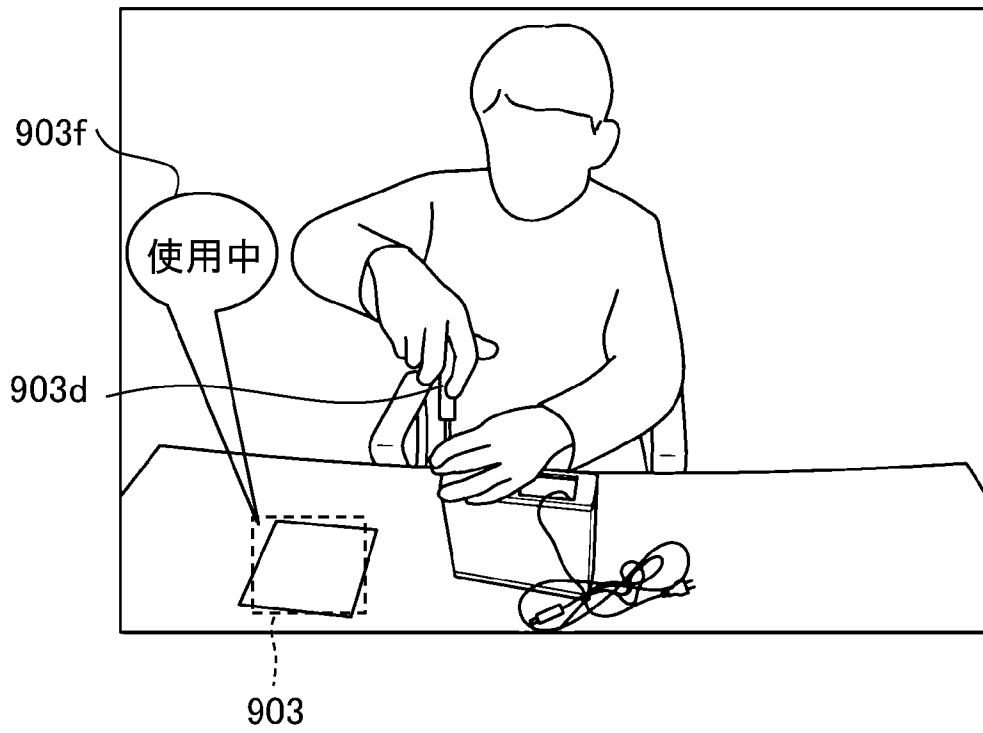
[図25]



[図26]



[図27]



[図28]

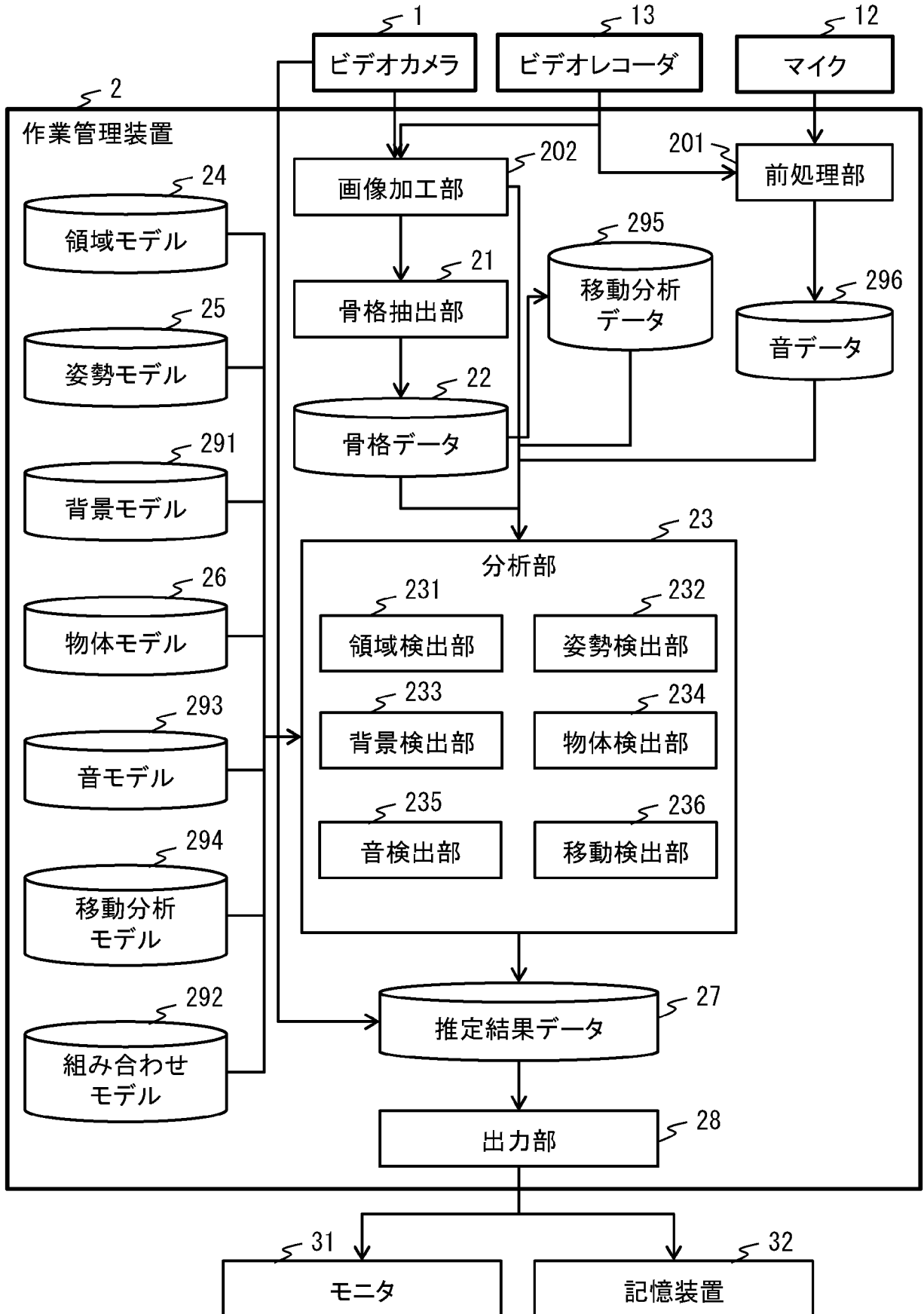
29

フレーム番号	領域データ	姿勢データ	背景データ	出力手順
1		その他	未使用 (ドライバあり)	部品待ち
:	:	:	:	:
:	右側に 両手首入る	右から 取り出す	未使用 (ドライバあり)	部品取り
:	:	:	:	:
:	なし	組み立てる	使用中 (ドライバなし)	組み立て
:	:	:	:	:
:		左に置く	未使用 (ドライバあり)	部品格納
:	:	:	:	:

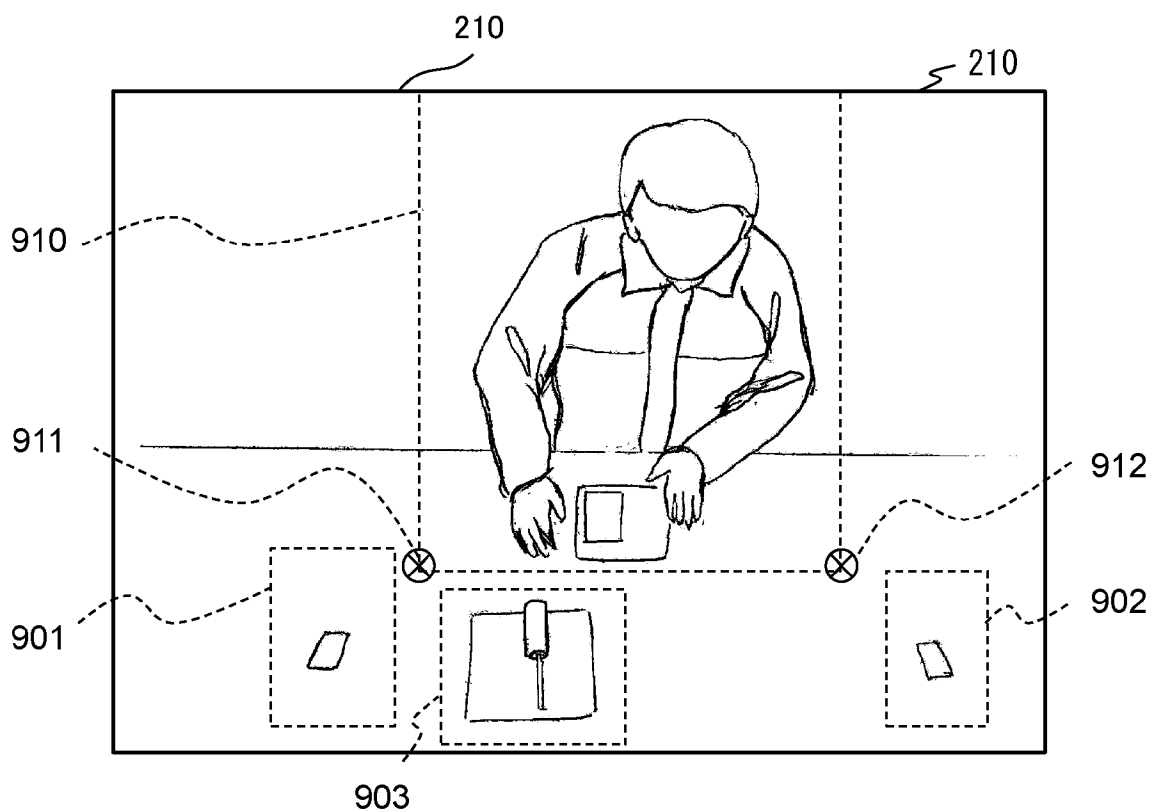
[図29]

作業手順	作業時間
部品待ち	
部品取り	
間作業	
組み立て	
間作業	
部品格納	

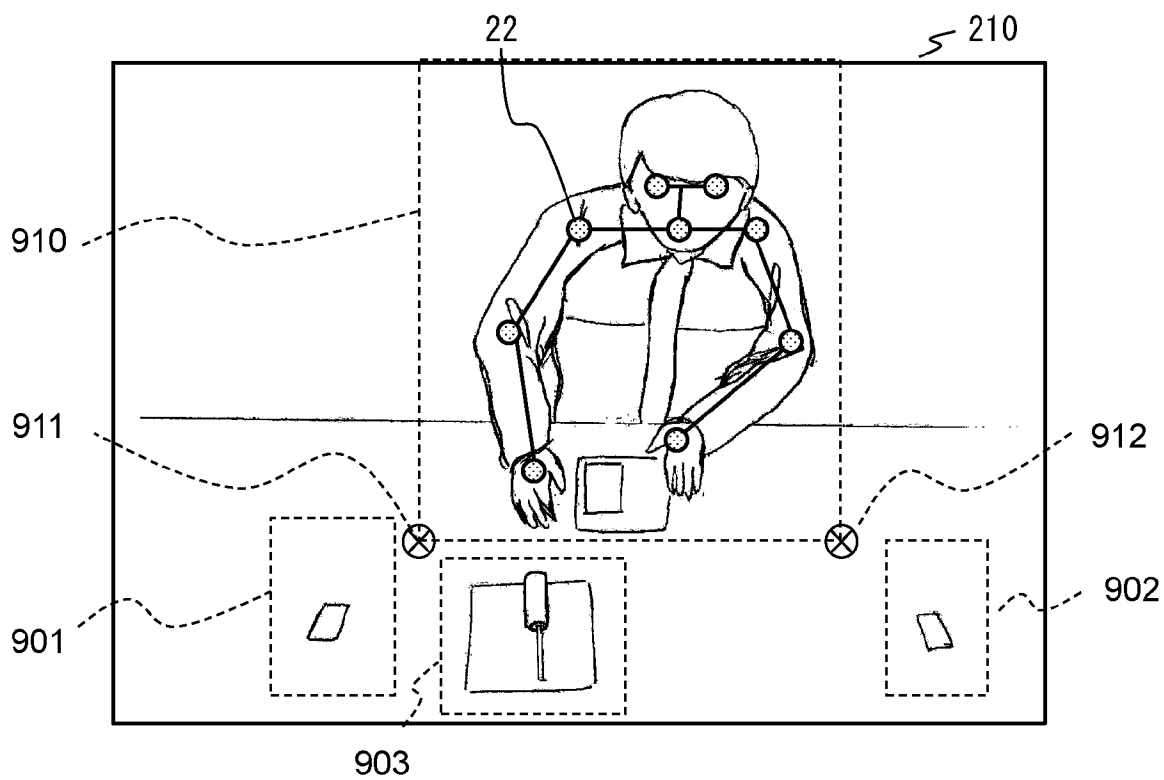
[図30]



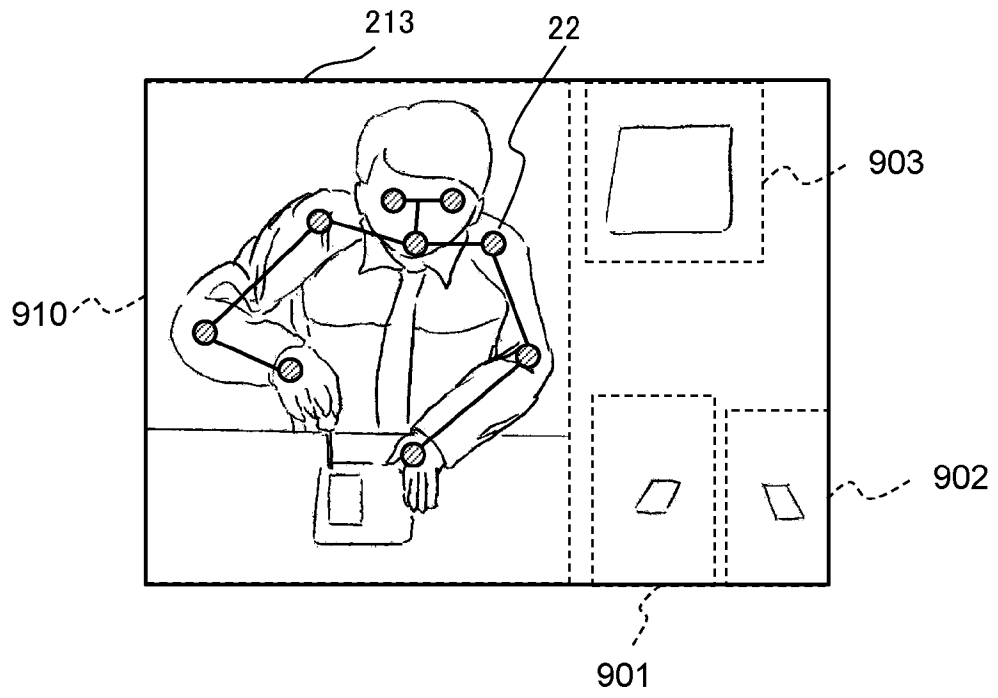
[図31]



[図32]



[図33]



[図34]

22

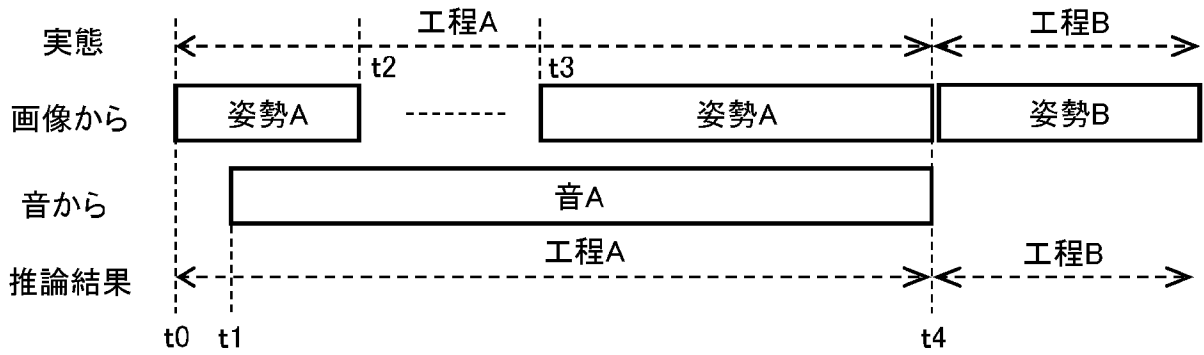
番号	特徴点	X軸座標	Y軸座標	スコア
0	鼻
1	右肩
2	右肘
3	右手首
4	左肩
5	左肘
6	左手首
7	右目
8	左目

[図35]

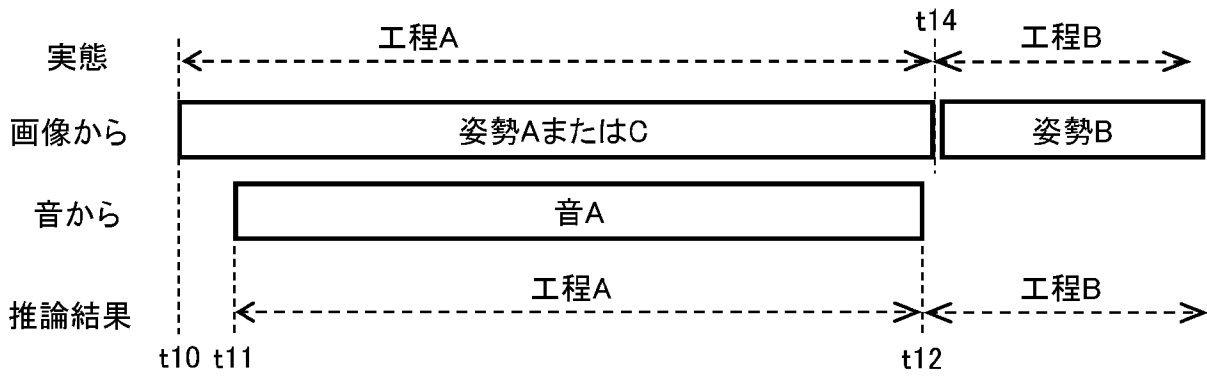
24

領域ラベル	特徴点番号	判定論理	多角形座標
部品取り領域	3,6	論理和	...
完成品格納領域	3,6	論理和	...
工具領域	3,6	論理積	...

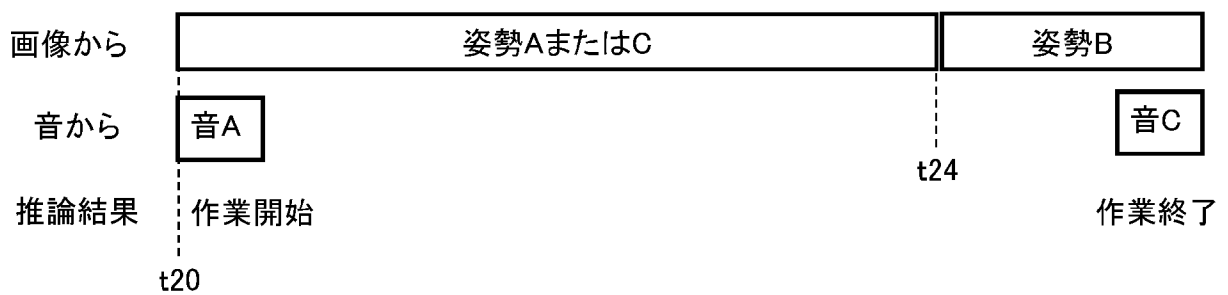
[図36]



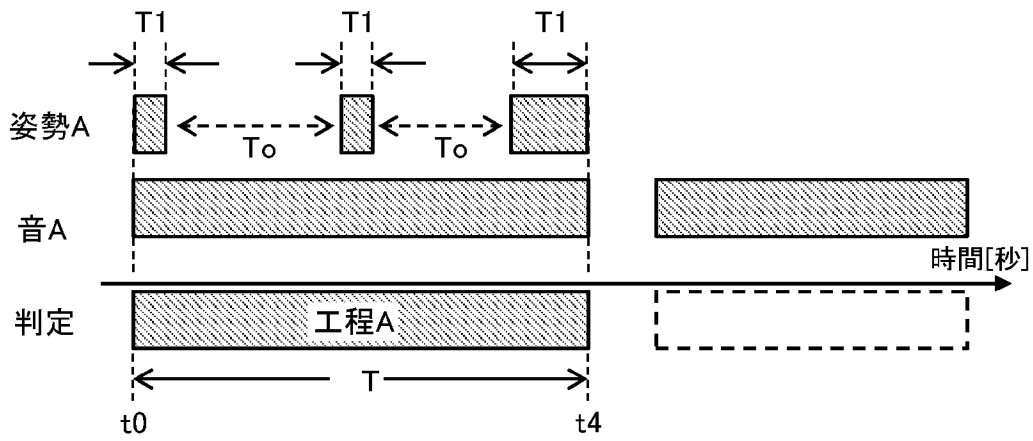
[図37]



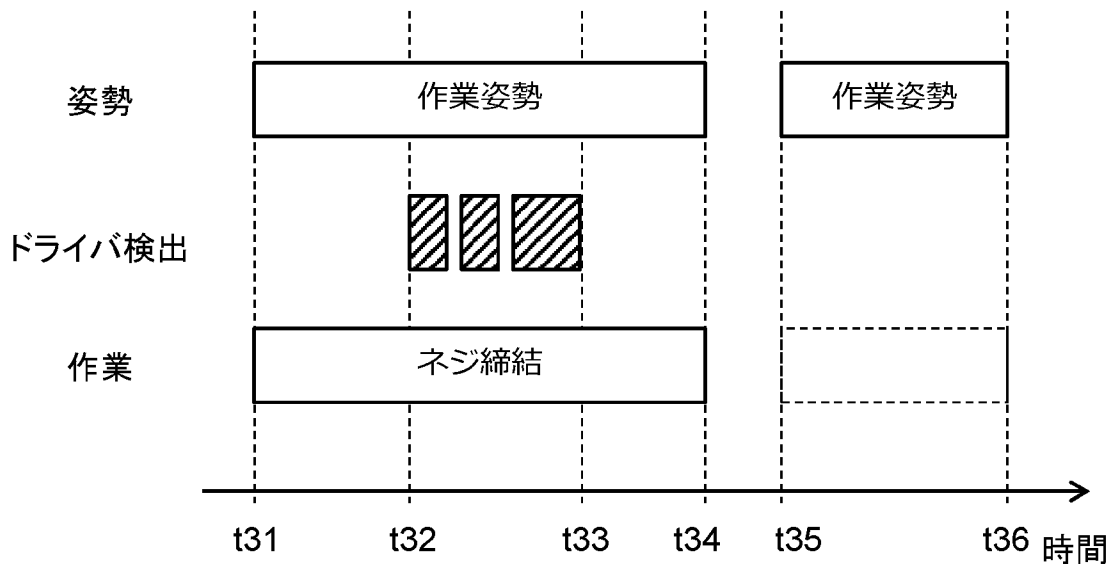
[図38]



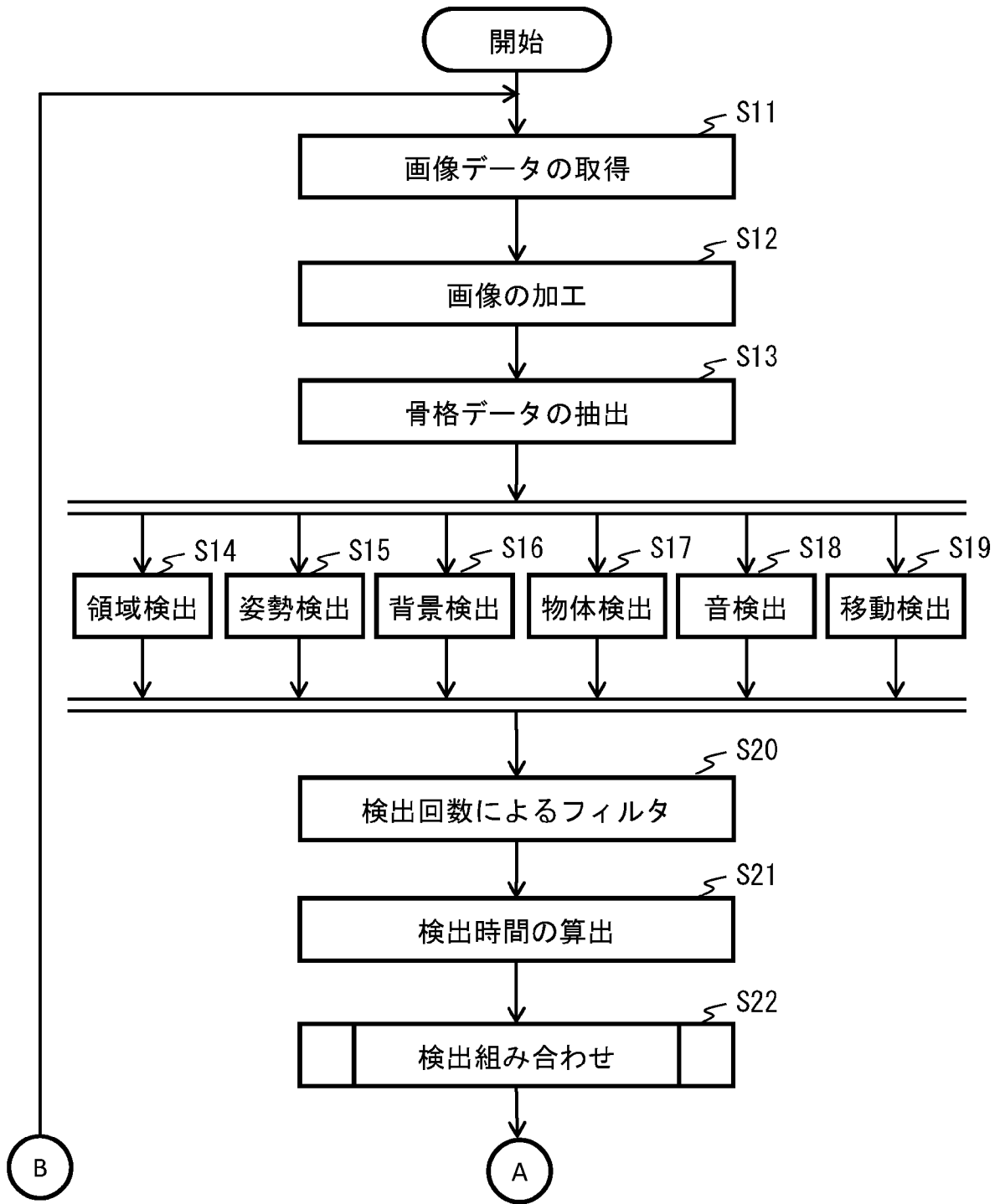
[図39]



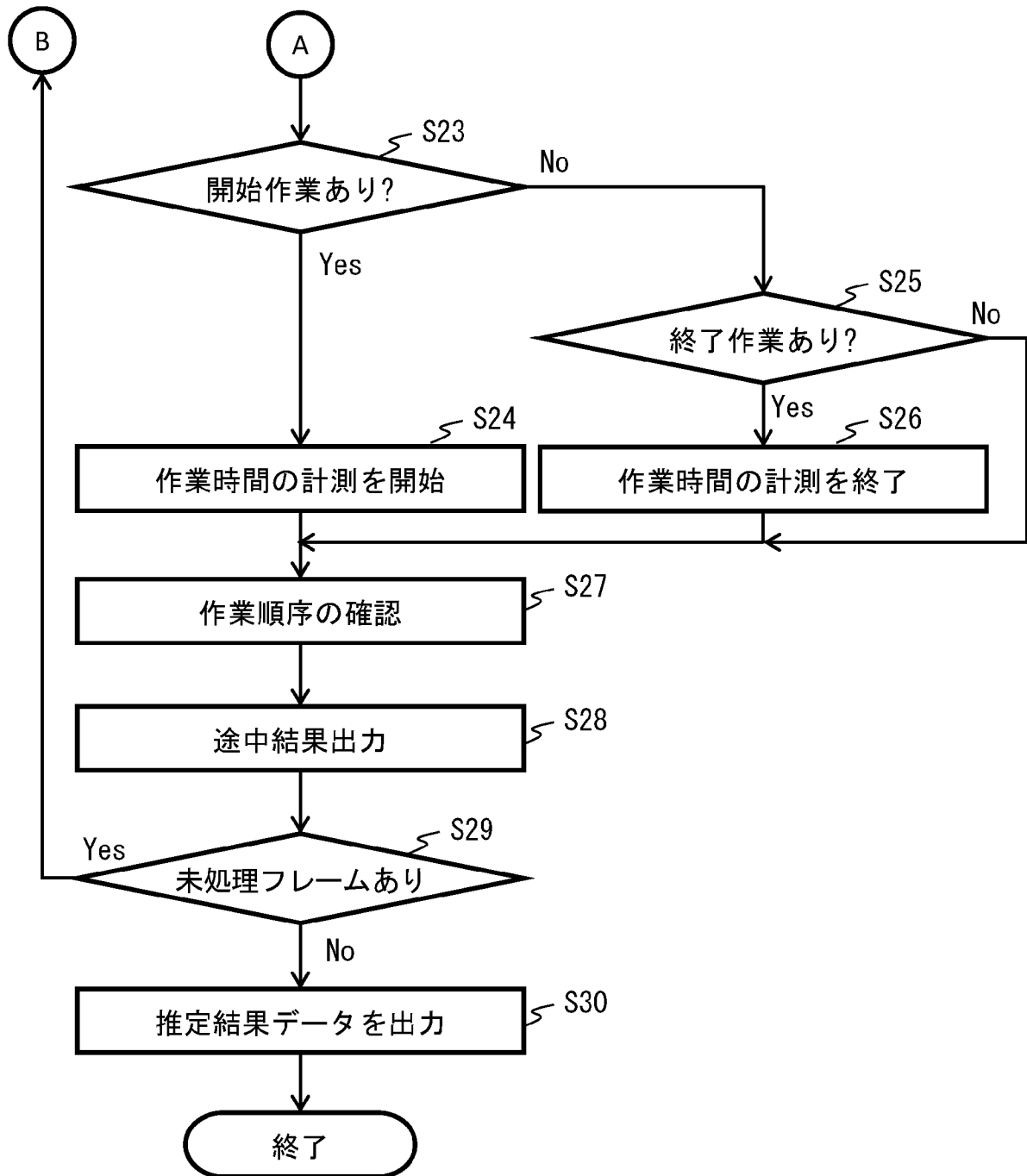
[図40]



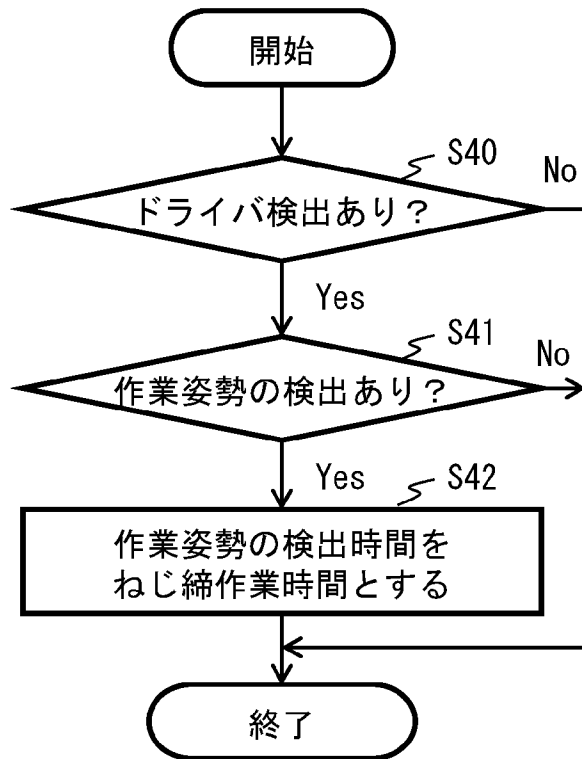
[図41A]



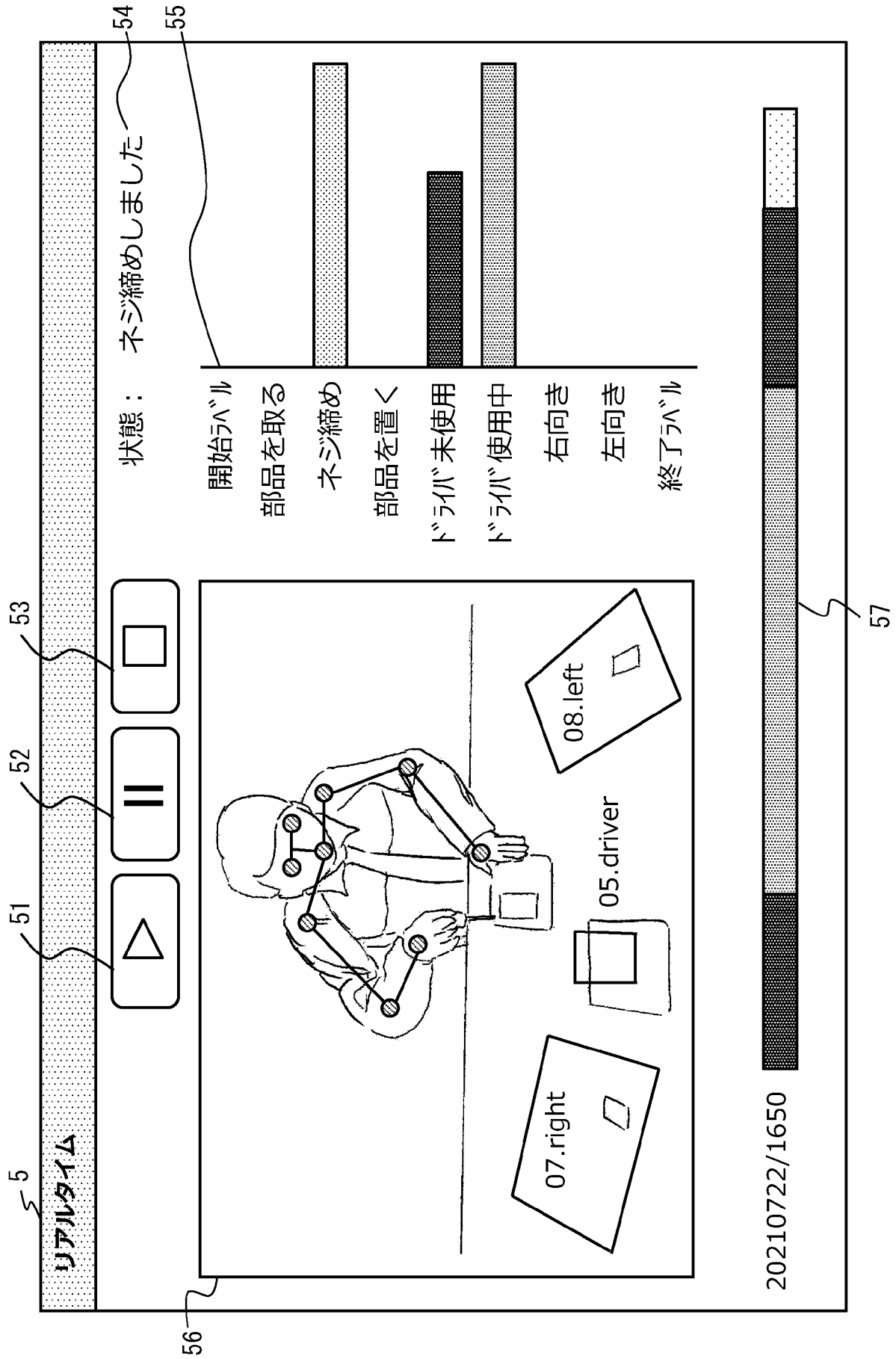
[図41B]



[図42]



[図43]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2023/034205

A. CLASSIFICATION OF SUBJECT MATTER		
G06T 7/70(2017.01)i FI: G06T7/70 Z		
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) G06T 7/00 - 7/90, G06V10/00 - 20/90, G06V30/418, G06V40/16, G06V40/20		
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 1971-2023 Registered utility model specifications of Japan 1996-2023 Published registered utility model applications of Japan 1994-2023		
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 2020-201772 A (HITACHI INDUSTRY & CONTROL SOLUTIONS, LTD.) 17 December 2020 (2020-12-17) paragraphs [0013]-[0016], [0022], [0024]-[0026], [0031], [0050], fig. 1, 5-8, 10, 25	1-2, 11-15
Y	JP 2021-72475 A (PANASONIC INTELLECTUAL PROPERTY MANAGEMENT CO., LTD.) 06 May 2021 (2021-05-06) paragraphs [0029]-[0035], [0063], [0066]-[0067], [0145]-[0146], [0148], fig. 1, 6, 15	1-2, 11-15
A	JP 2021-163188 A (KOBELCO CONSTRUCTION MACHINERY CO., LTD.) 11 October 2021 (2021-10-11) paragraphs [0009], [0014], [0018], [0028], fig. 1	3-10
<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 13 October 2023		Date of mailing of the international search report 31 October 2023
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/JP2023/034205

Patent document cited in search report	Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
JP 2020-201772	A 17 December 2020	(Family: none)	
JP 2021-72475	A 06 May 2021	(Family: none)	
JP 2021-163188	A 11 October 2021	US 2023/0105895 A1 paragraphs [0016], [0021], [0025], [0035], fig. 1 EP 4105391 A1 CN 115362297 A	

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）） G06T 7/00 - 7/90, G06V10/00 - 20/90, G06V30/418, G06V40/16, G06V40/20 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 - 2023年 日本国実用新案登録公報 1996 - 2023年 日本国登録実用新案公報 1994 - 2023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 2020-201772 A（株式会社 日立産業制御ソリューションズ）17.12.2020（2020 - 12 - 17） 段落[0013]～[0016], [0022], [0024]～[0026], [0031], [0050], 図1, 5～8, 10, 25	1-2, 11-15
Y	JP 2021-72475 A（パナソニックIPマネジメント株式会社）06.05.2021（2021 - 05 - 06） 段落[0029]～[0035], [0063], [0066]～[0067], [0145]～[0146], [0148], 図1, 6, 15	1-2, 11-15
A	JP 2021-163188 A（コベルコ建機株式会社）11.10.2021（2021 - 10 - 11） 段落[0009], [0014], [0018], [0028], 図1	3-10
<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、一般的な技術水準を示すもの “E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの “L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す） “O” 口頭による開示、使用、展示等に言及する文献 “P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献 “T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの “X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの “Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの “&” 同一パテントファミリー文献		
国際調査を完了した日	13.10.2023	国際調査報告の発送日 31.10.2023
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） 伊知地 和之 5H 9291 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/JP2023/034205

引用文献	公表日	パテントファミリー文献	公表日
JP 2020-201772 A	17.12.2020	(ファミリーなし)	
JP 2021-72475 A	06.05.2021	(ファミリーなし)	
JP 2021-163188 A	11.10.2021	US 2023/0105895 A1 段落[0016],[0021],[0025], [0035], 図1 EP 4105391 A1 CN 115362297 A	