

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-29021
(P2019-29021A)

(43) 公開日 平成31年2月21日(2019.2.21)

(51) Int.Cl.		F I		テーマコード (参考)
G06T 7/00	(2017.01)	G06T 7/00	350B	3C707
B25J 13/08	(2006.01)	B25J 13/08	A	5L096

審査請求 有 請求項の数 19 O L (全 26 頁)

(21) 出願番号	特願2018-141910 (P2018-141910)	(71) 出願人	504143441
(22) 出願日	平成30年7月27日 (2018.7.27)		国立大学法人 奈良先端科学技術大学院大学
(31) 優先権主張番号	特願2017-147305 (P2017-147305)		奈良県生駒市高山町8916-5
(32) 優先日	平成29年7月30日 (2017.7.30)	(74) 代理人	110000822
(33) 優先権主張国	日本国 (JP)		特許業務法人グローバル知財
		(72) 発明者	友近 圭汰
			奈良県生駒市高山町8916-5 国立大 学法人奈良先端科学技術大学院大学内
		(72) 発明者	清川 拓哉
			奈良県生駒市高山町8916-5 国立大 学法人奈良先端科学技術大学院大学内
		(72) 発明者	小笠原 司
			奈良県生駒市高山町8916-5 国立大 学法人奈良先端科学技術大学院大学内

最終頁に続く

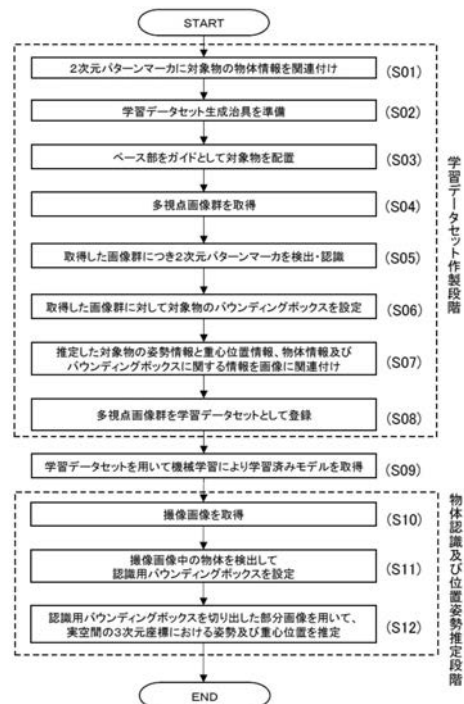
(54) 【発明の名称】 学習データセット作製方法及び、物体認識及び位置姿勢推定方法

(57) 【要約】

【課題】 工場の自動化システムに用いるロボット制御に向け、複数の物体を対象として、機械学習に用いる学習データセットを効率的に作製できる方法及び装置、学習データセットを用いて物体認識及び位置姿勢を同時に高精度で推定できる方法及び装置を提供する。

【解決手段】 位置姿勢検出用マーカに対象物の物体情報を関連付けし、対象物の配置位置のガイドとなるベース部と、ベース部の上方に固定されたマーカとから構成される学習データセット生成治具を用いて、ベース部をガイドとして対象物を配置した状態で、マーカを含む物体全体の多視点画像群を取得する。そして、取得した画像群に対して対象物のバウンディングボックスを設定し、撮像画像から推定した対象物の姿勢情報と重心位置情報、物体情報及びバウンディングボックスに関する情報を撮像画像に関連付けして、対象物の物体認識及び位置姿勢推定を行うための学習データセットを生成する。

【選択図】 図 1



【特許請求の範囲】**【請求項 1】**

位置姿勢検出用マーカに対象物の物体情報を関連付けし、
対象物の配置位置のガイドとなるベース部と、該ベース部の上方に固定された前記マーカとから構成される学習データセット生成治具を用いて、前記ベース部をガイドとして対象物を配置した状態で、前記マーカを含む物体全体の多視点画像群を取得し、
取得した画像群に対して対象物のバウンディングボックスを設定し、
撮像画像から推定した対象物の姿勢情報と位置情報、前記物体情報及び前記バウンディングボックスに関する情報を、前記撮像画像に関連付けして、対象物の物体認識及び位置姿勢推定を行うための学習データセットを生成することを特徴とする学習データセット作製方法。

10

【請求項 2】

前記位置情報は、前記ベース部をガイドとして配置された対象物の位置と、予め定義した基準座標系における基準姿勢の位置との差分情報であることを特徴とする請求項 1 に記載の学習データセット作製方法。

【請求項 3】

前記姿勢情報は、前記マーカの撮像画像と、前記マーカと対象物との相対位置関係を用いて算出した情報であり、前記ベース部をガイドとして配置された対象物の姿勢と、予め定義した基準座標系における基準姿勢との差分情報であることを特徴とする請求項 1 又は 2 に記載の学習データセット作製方法。

20

【請求項 4】

前記学習データセット生成治具において、前記ベース部は対象物を載せる台座部であり、
前記マーカは、該台座部に上に固着された支柱の上方に脱着自在に固定されたことを特徴とする請求項 1 ~ 3 の何れかに記載の学習データセット作製方法。

【請求項 5】

前記マーカは、ARマーカを含む2次元パターンマーカ、又は、3次元マーカであることを特徴とする請求項 1 ~ 4 の何れかに記載の学習データセット作製方法。

【請求項 6】

前記多視点画像群は、対象物を前記学習データセット生成治具に配置して搬送手段により搬送しながら撮像、対象物を前記学習データセット生成治具に配置して回転手段により回転させながら撮像、及び、対象物を前記学習データセット生成治具に配置して移動手段により移動させながら撮像の少なくとも何れかにより取得されたことを特徴とする請求項 1 ~ 5 の何れかに記載の学習データセット作製方法。

30

【請求項 7】

請求項 1 ~ 6 の何れかの作製方法により作製した前記学習データセット。

【請求項 8】

請求項 1 ~ 6 の何れかの作製方法に用いる前記学習データセット生成治具。

【請求項 9】

請求項 7 の学習データセットを用いて機械学習を行い、学習済みモデルを取得し、
新たに撮像された撮像画像から、画像認識によって物体の名称を検出して認識用バウンディングボックスを作成し、
前記認識用バウンディングボックスを切り出した部分画像を生成し、前記学習済みモデルを用いて、生成した部分画像に基づき検出した物体の姿勢を推定し、
前記認識用バウンディングボックスの位置、幅及び高さを取得し、前記学習済みモデルを用いて、検出した物体に関する実空間の3次元座標における位置を推定することを特徴とする物体認識及び位置姿勢推定方法。

40

【請求項 10】

前記学習済みモデルは、物体の姿勢もしくは位置を推定するための分類器又は回帰器の少なくとも何れかが含まれることを特徴とする請求項 9 に記載の物体認識及び位置姿勢推

50

定方法。

【請求項 1 1】

前記撮像画像は、複数の物体が撮像された画像であり、
検出した各物体に対して、各々認識用バウンディングボックスを作成し、
検出した全ての物体の名称、並びに、推定した姿勢及び位置を実空間の 3 次元座標として算出することを特徴とする請求項 9 又は 10 に記載の物体認識及び位置姿勢推定方法。

【請求項 1 2】

前記撮像画像は、前記学習データセットの作製環境と同一又は近似した環境下で撮像された画像であることを特徴とする請求項 9 ~ 1 1 の何れかに記載の物体認識及び位置姿勢推定方法。

10

【請求項 1 3】

請求項 1 ~ 6 の何れかの学習データセット作製方法の一部の処理をコンピュータに実行させるためのプログラムであって、

前記多視点画像群に対して対象物の前記バウンディングボックスを設定するステップ、
前記撮像画像から推定した対象物の姿勢情報と位置情報、前記物体情報及び前記バウンディングボックスに関する情報を、前記撮像画像に関連付けして、前記学習データセットを生成するステップ、

をコンピュータに実行させるための学習データセット作製プログラム。

【請求項 1 4】

請求項 9 ~ 1 2 の何れかの物体認識及び位置姿勢推定方法の一部の処理をコンピュータに実行させるためのプログラムであって、

新たに撮像された撮像画像から、画像認識によって物体の名称を検出して認識用バウンディングボックスを作成するステップ、

前記認識用バウンディングボックスを切り出した部分画像を生成し、生成した部分画像に基づき前記学習済みモデルを用いて、検出した物体の姿勢を推定するステップ、

前記認識用バウンディングボックスの位置、幅及び高さを取得し、前記学習済みモデルを用いて、検出した物体に関して実空間の 3 次元座標における位置を推定するステップ、
をコンピュータに実行させるための物体認識及び位置姿勢推定プログラム。

20

【請求項 1 5】

対象物の配置位置のガイドとなるベース部と、該ベース部の上方に固定され、対象物の物体情報が関連付けされる位置姿勢検出用マーカとから構成される学習データセット生成治具、

前記ベース部をガイドとして対象物を配置した状態で、前記マーカを含む物体全体の多視点画像群を取得する撮像手段、

取得した画像群に対して対象物のバウンディングボックスを設定する手段、

撮像画像から推定した対象物の姿勢情報と位置情報、前記物体情報及び前記バウンディングボックスに関する情報を、前記撮像画像に関連付けして、対象物の物体認識及び位置姿勢推定を行うための学習データセットを生成する手段、

を備えたことを特徴とする学習データセット作製装置。

30

【請求項 1 6】

前記学習データセット生成治具において、前記ベース部は対象物を載せる台座部であり、

前記マーカは、該台座部に上に固着された支柱の上方に脱着自在に固定されたことを特徴とする請求項 1 5 に記載の学習データセット作製装置。

40

【請求項 1 7】

請求項 1 ~ 6 の何れかの作製方法により作製した前記学習データセットを用いて機械学習を行った産業用ロボットによる工場の自動化システム。

【請求項 1 8】

請求項 9 ~ 1 2 の何れかの物体認識及び位置姿勢推定方法を用いた産業用ロボットによる工場の自動化システム。

50

【請求項 19】

請求項 14 に記載の物体認識及び位置姿勢推定プログラムが搭載された産業用ロボットによる工場の自動化システム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、工場自動化に用いるロボット制御システムに向けた、学習データセットの作製方法と、その学習データセットを用いた物体認識と位置姿勢の同時推定方法に関するものである。

10

【背景技術】**【0002】**

従来から、工場等での作業を自動化するものとして、人工知能 (Artificial Intelligence、以下“AI”とする。) を搭載したロボットが利用されている。近年、機械学習やディープラーニング (深層学習) の発展により、工場等における生産システムにおいても、工場等の完全自動化に向けた取り組みとして、機械学習等を利用した AI の開発が急速に行われている。

ロボットによる作業の自動化は、あらゆる業界にニーズが存在するが、中でも、食品業界や物流業界は今後の成長が期待される分野であり、ロボットによる作業の自動化のニーズは高く存在する。

20

しかしながら、食品業界や物流業界において扱う製品は、柔軟なものが多く、取り扱いの際に複雑に形状が変化してしまうため、ロボットに備えられたロボットハンドでのハンドリングが難しいという問題がある。また、食品業界や物流業界において扱う商品は、多品種であることから、商品の認識を正確かつ迅速に行うことが難しいという問題もある。すなわち、多品種な商品を短時間で認識して、詰め合わせギフトや不良品の除去作業などのような仕分け作業を正確に行うことが求められるのである。

これらの事情から、食品業界や物流業界においては、未だにロボットによる工場の自動化が十分に進んでいないのが現実である。

【0003】

例えば、ロボットを用いて工場での仕分け作業を行う場合、対象となる物体をカメラ等で撮影し、物体を認識することが必要なところ、カメラから入力された画像の認識精度を高める表示データ解析装置が知られている (特許文献 1 を参照)。これは、マーカを検出して物体の情報を表示するセグメント領域を決定するための技術であり、物体の情報と位置のみを推定するものである。

30

しかしながら、特許文献 1 に開示された表示データ解析装置では、マーカの位置情報と姿勢情報を表示するのみであり、物体を囲むセグメント領域の情報を取得できていないため、機械学習に用いる学習データセットとしては活用できないという問題がある。

【0004】

また、3次元的に任意の位置姿勢をとりうる物体の位置姿勢を推定する装置が知られている (特許文献 2 を参照)。これは、同一の物体に関して、画像から物体の空間的な位置および姿勢を推定するものである。

40

しかしながら、特許文献 2 に開示された推定装置では、複数の物体を用いる場合には、それぞれの位置や姿勢を推定できない場合があるという問題がある。

【0005】

また、画像から物体を検出するものとしては、顔検出のための学習画像を収集するために、取得した画像自体に変形を加えて、学習画像を生成する手法が提案されている (非特許文献 1 を参照)。

しかしながら、非特許文献 1 に開示されている学習画像の生成手法では、取得した画像を加工する必要があるところ、変形を加えるという加工により作製された学習データセットは、実際の環境で作製されたものとは異なるため、実環境での学習データセットが必要

50

とされるという要求には適用が困難であり、また、実環境での学習データセットでは無いことから十分に学習の効果を発揮しないという問題がある。さらに、取得した画像に変形を加えるため、学習データセットの準備に時間がかかるという問題もある。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2002-24800号公報

【特許文献2】特開2011-133273号公報

【非特許文献】

【0007】

【非特許文献1】“ディープラーニングのための訓練データ自動生成”，大西一徳ら，第14回情報科学技術フォーラム予稿集，第3分冊，第443～446頁

【発明の概要】

【発明が解決しようとする課題】

【0008】

かかる状況に鑑みて、本発明は、工場の自動化システムに用いるロボット制御に向け、複数の物体を対象として、機械学習に用いる学習データセットを効率的に作製できる方法及び装置、並びに、作製した学習データセットを用いて物体認識及び位置姿勢を同時に高精度で推定できる方法及び装置を提供することを目的とする。

【課題を解決するための手段】

【0009】

上記課題を解決すべく、本発明の学習データセット作製方法は、位置姿勢検出用マーカに対象物の物体情報を関連付けし、対象物の配置位置のガイドとなるベース部と、ベース部の上方に固定されたマーカとから構成される学習データセット生成治具を用いて、ベース部をガイドとして対象物を配置した状態で、マーカを含む物体全体の多視点画像群を取得する。そして、取得した画像群に対して対象物のバウンディングボックスを設定し、撮像画像から推定した対象物の姿勢情報と位置情報、物体情報及びバウンディングボックスに関する情報を、撮像画像に関連付けして、対象物の物体認識及び位置姿勢推定を行うための学習データセットを生成する。

【0010】

位置姿勢検出用マーカを用いることにより、撮像画像から、対象物の物体情報だけではなく、姿勢情報と位置情報を同時に取得することが可能となり、学習データセットの生成を効率良く行うことができる。多視点画像群の取得方法としては、対象物を移動等させながら1台のカメラにより撮影し取得したものでよいし、複数台のカメラを用いて取得したものでよい。いずれにしても、対象物を異なる角度から撮影した大量の撮像画像を収集することで、機械学習のための学習データとして用いることが可能である。

学習データセットの生成に用いる撮像画像自体には、加工は行われないため、学習済みモデルを利用する実際の環境で得られる撮像画像に近い撮像画像を用いて、機械学習を行うことができる。そのため、物体認識と位置・姿勢推定の精度を向上させることができる。

なお、対象物の物体情報とは、対象物の名称、サイズといった基本情報のことであり、学習データセットの生成に当たって、予めコンピュータにこれらの情報が記憶されている。

また、バウンディングボックスとは、画像等において、対象物が収まるように取り囲まれた表示領域であり、矩形などの領域をいう。取得した画像群に対して対象物のバウンディングボックスを設定する際は、バウンディングボックスが対象物全体を可能な限り確実に囲うようにするために、物体の位置情報として、物体の重心位置を取得して設定する。

【0011】

本発明の学習データセット作製方法における姿勢情報は、具体的には、マーカの撮像画像と、マーカと対象物との相対位置関係を用いて算出した情報であり、ベース部をガイド

10

20

30

40

50

として配置された対象物の姿勢と、予め定義した基準座標系における基準姿勢との差分情報である。

マーカの撮像画像と、マーカと対象物との相対位置関係を用いて、対象物の姿勢を推定することにより、精度の高い姿勢推定が可能となる。

基準座標系は、例えば、予め学習データセット生成治具に取り付けた位置姿勢検出用マーカを基準マーカとして利用して、基準としたい位置及び姿勢に配置することにより定義することができる。そのようにすることで、基準マーカの基準位置及び基準姿勢に関するデータをデータベースとして記憶し、学習データセット作製段階に、実空間の基準座標系を用いて、その基準座標系における位置と姿勢の差分を、位置と姿勢の情報として画像に紐づけることができる。

10

基準座標系を示す基準マーカが撮像画像内に無い場合であっても、撮像画像から基準座標系を認識することは可能である。例えば、カメラが単一で固定されている場合には、ズームする前の撮像画像内で基準マーカにより基準座標系が設定できていれば、ズームを行い見えなくなった場合でもズーム倍率がわかる限り、撮像画像外の基準マーカの位置と姿勢を計算することが可能である。

一方、複数台のカメラが固定されている場合には、それぞれのカメラ間の位置関係が予めわかっているならば、基準座標系を示す基準マーカが撮像画像内に無い場合であっても、撮像画像外の基準マーカの位置と姿勢を計算することが可能である。また、1台のカメラでしか見えない位置に基準マーカを配置した場合でも、見えているカメラと見えていないカメラで同時に撮像可能な場所に、仮の基準マーカを配置した状態で撮像した映像を用いて、基準座標系を算出することが可能である。さらに、カメラが動的に動く場合についても、最初に基準座標系を撮影したときのカメラの位置と移動先のカメラとの間の位置関係が分かれば、計算により基準座標系を算出できる。

20

【0012】

本発明の学習データセット作製方法における位置情報は、対象物の任意のポイントの位置情報を用いることができる。任意のポイントに関して、対象物の基準姿勢における当該ポイントの位置情報を予め登録しておく。その場合、位置情報は、ベース部をガイドとして配置された対象物の該ポイントの位置と、基準座標系における基準姿勢の該ポイント位置との差分情報である。

また、本発明の学習データセット作製方法における位置情報が重心位置情報である場合、重心位置情報は、ベース部をガイドとして配置された対象物の重心位置と、基準座標系における基準姿勢の重心位置との差分情報である。

30

ここで、基準座標系の代わりに、カメラ座標系を用いて、対象物の姿勢情報及び位置情報を表すことも可能である。例えば、1台の固定カメラで撮像する場合、そのカメラのカメラ座標系での3次元座標で、対象物の姿勢情報と位置情報を表す。そして、カメラ座標系の3次元座標から、実空間の3次元座標に変換し、ロボットに実空間の3次元座標を伝達することにより、ロボットが物体を正確に捉えさせることができる。

【0013】

本発明の学習データセット作製方法に用いる学習データセット生成治具において、ベース部は対象物を載せる台座部であり、マーカは、台座部に上に固着された支柱の上方に脱着自在に固定されたことでもよい。マーカが、対象物より上方に設けられることにより、上方からカメラで撮影した場合に、マーカが対象物に隠れてしまうことを防止でき、物体認識の精度を向上させることができる。また、マーカが脱着自在に固定されることで、対象物を変えて、容易にベース部を再利用することが可能である。

40

ベース部が、対象物を載せる台座部とされることで、学習データセット生成治具に対象物が配置された際に、対象物が、支柱の上方に設けられたマーカを安定させる重石としての役割を果たすことにもなる。

【0014】

本発明の学習データセット作製方法に用いる位置姿勢検出用マーカは、AR (Augmented Reality: 拡張現実) マーカを含む2次元パターンマーカや3次元マーカを用いること

50

ができる。ARマーカとは、拡張現実システムにおいて、付加情報を表示する位置を指定するための標識となるパターン画像をいう。また、3次元マーカとは、例えば、光学式モーションキャプチャシステムで用いられる球形状のマーカを用いることができる。これらの球形状のマーカで位置及び姿勢を検出する場合には、学習データセット生成治具のベース部の上方に、3個以上のマーカを固定する。

【0015】

本発明の学習データセット作製方法において取得される多視点画像群は、対象物を学習データセット生成治具に配置して搬送手段により搬送しながら撮像して取得されても良く、また、対象物を学習データセット生成治具に配置して回転手段により回転させながら撮像取得されてもよく、さらに、対象物を前記学習データセット生成治具に配置して移動手段により移動させながら撮像取得されてもよい。これらの撮像取得のやり方は、何れかを組合せて用いることも可能である。搬送手段や回転手段や移動手段が用いられるにより、撮像手段として用いるカメラが1台であっても、多視点画像群を簡便に取得でき、学習データセット生成作業の利便性、時間短縮を行うことができる。

10

ここで、搬送手段とは、搬送路をレーンとして対象物を搬送するものであり、具体的には、対象物を学習データセット生成治具に配置した状態で、搬送ベルトで搬送するベルトコンベアが好適に用いられる。また、回転手段として、例えば、対象物を学習データセット生成治具に配置した状態で、水平面で回動可能な回転機構が用いられる。また、移動手段として、例えば、対象物を学習データセット生成治具に配置した状態で、学習データセット生成治具が水平の直線レール上の移動可能な直動機構が用いられる。さらには、回転機構と直動機構を組合せ、回転しながら移動できる回転移動機構も使用することができる。これらの回転機構、直動機構や回転移動機構を、ベルトコンベアに載せて、カメラ1台で多視点画像群を取得することができる。

20

【0016】

学習データセット生成治具は、マーカを除き、透明性を有することでも構わない。学習データセット生成治具のマーカ以外の部分が透明性を有することにより、撮像画像において、学習データセット生成治具が対象物を遮ることを防止でき、物体認識の精度を向上させることができる。また、学習データセット生成治具は、画像解析処理によって、取得した撮像画像中から消し去る加工処理を施してもよい。

【0017】

本発明の学習データセットは、本発明の学習データセット作製方法により作製したものである。また、本発明の学習データセット生成治具は、本発明の学習データセット作製方法に用いられるものであり、上述の如く、対象物の配置位置のガイドとなるベース部と、ベース部の上方に固定されるマーカとから構成され、マーカには、対象物の物体名称などの物体情報が関連付けされる。

30

【0018】

本発明の物体認識及び位置姿勢推定方法では、本発明の学習データセット作製方法により作製した学習データセットを用いて機械学習を行い、学習済みモデルを取得し、新たに撮像された撮像画像から、画像認識によって物体を検出して認識用バウンディングボックスを作成する。そして、認識用バウンディングボックスを切り出した部分画像を生成し、学習済みモデルを用いて、生成した部分画像に基づき検出した物体の情報（例えば、物体名称等）及び姿勢を推定する。また、認識用バウンディングボックスの位置、幅及び高さを取得し、学習済みモデルを用いて、検出した物体に関して実空間の3次元座標における位置を推定する。

40

なお、検出した物体全体を取り囲む認識用バウンディングボックスの位置は、取り囲んだバウンディングボックスの重心位置でもよい。また、検出した物体に関する実空間の3次元座標における位置は、物体の重心位置でもよい。

【0019】

本発明の物体認識及び位置姿勢推定方法において、学習済みモデルは、具体的には、物体の姿勢もしくは位置を推定するための分類器又は回帰器の少なくとも何れかが含まれる

50

。推定する位置は、物体の重心位置であってもよい。

【0020】

本発明の物体認識及び位置姿勢推定方法では、撮像画像は、複数の物体が撮像された画像であり、検出した各物体に対して、各々認識用バウンディングボックスを作成し、検出した全ての物体の名称、並びに、推定した姿勢及び位置を実空間の3次元座標として算出する。

これにより、検出された各物体に関して、推定した実空間の3次元座標における姿勢及び位置を、工場の自動化システムのロボットに対して伝達できる。また、検出された各物体に関して、実空間の3次元座標における姿勢及び重心位置を推定し算出することにより、ロボットが正確に物体を捉えることができる。

10

【0021】

本発明の物体認識及び位置姿勢推定方法における撮像画像は、学習データセットの作製環境と同一又は近似した環境下で撮像された撮像画像であることが好ましい。

学習データセットの作製環境と同一又は近似した環境下で撮像された撮像画像が用いられることにより、物体認識と位置姿勢の推定精度を向上させることができる。

【0022】

本発明の学習データセット作製プログラムは、本発明の学習データセット作製方法の一部の処理である下記a)、b)のステップをコンピュータに実行させるためのプログラムである。

a) 多視点画像群に対して対象物のバウンディングボックスを設定するステップ、
b) 撮像画像から推定した対象物の姿勢情報と位置情報、物体情報及びバウンディングボックスに関する情報を、撮像画像に関連付けして、学習データセットを生成するステップ。

20

【0023】

本発明の物体認識及び位置姿勢推定プログラムは、本発明の物体認識及び位置姿勢推定方法の一部の処理である下記c)~e)をコンピュータに実行させるためのプログラムである。

c) 新たに撮像された撮像画像から、画像認識によって物体の名称を検出して認識用バウンディングボックスを作成するステップ、
d) 認識用バウンディングボックスを切り出した部分画像を生成し、生成した部分画像に基づき学習済みモデルを用いて、検出した物体の姿勢を推定するステップ、
e) 認識用バウンディングボックスの位置、幅及び高さを取得し、学習済みモデルを用いて、検出した物体に関して実空間の3次元座標における位置を推定するステップ。

30

なお、検出した物体全体を取り囲む認識用バウンディングボックスの位置は、取り囲んだバウンディングボックスの重心位置でもよい。また、検出した物体に関する実空間の3次元座標における位置は、物体の重心位置でもよい。

【0024】

本発明の学習データセット作製装置は、下記1)~4)を備える。

1) 対象物の配置位置のガイドとなるベース部と、該ベース部の上方に固定され、対象物の物体情報が関連付けされる位置姿勢検出用マーカとから構成される学習データセット生成治具、
2) 学習データセット生成治具のベース部をガイドとして対象物を配置した状態で、マーカを含む物体全体の多視点画像群を取得する撮像手段、
3) 取得した画像群に対して対象物のバウンディングボックスを設定する手段、
4) 撮像画像から推定した対象物の姿勢情報と位置情報、物体情報及びバウンディングボックスに関する情報を、撮像画像に関連付けして、対象物の物体認識及び位置姿勢推定を行うための学習データセットを生成する手段。

40

【0025】

本発明の学習データセット作製装置の学習データセット生成治具において、具体的には、ベース部は対象物を載せる台座部であり、マーカは、該台座部に上に固着された支柱の

50

上方に脱着自在に固定される。また、マーカは、ARマーカを含む2次元パターンマーカ、又は、3次元マーカである。

本発明の学習データセット作製方法により作製した学習データセットを用いて機械学習を行った産業用ロボット、本発明の物体認識及び位置姿勢推定方法を用いた産業用ロボット、そして、本発明の物体認識及び位置姿勢推定プログラムが搭載された産業用ロボットは、食品業界や物流業界における仕分け作業や加工作業などの工場の自動化システムに利用できる。

【発明の効果】

【0026】

本発明によれば、複数の物体を対象として、機械学習に用いる学習データセットを効率的に作製し、それを用いて物体認識及び位置姿勢を高精度に推定でき、工場の自動化システムに用いるロボット制御に活用できるといった効果がある。

10

【図面の簡単な説明】

【0027】

【図1】学習データセットの作製処理、並びに、物体認識及び位置姿勢推定処理の概略フロー図

【図2】学習データセットの作製装置、並びに、物体認識及び位置姿勢推定装置の機能ブロック図

【図3】学習データセット作製装置のイメージ図

【図4】学習データセット作製処理の詳細フロー図

20

【図5】物体認識及び位置姿勢推定の説明図であり、(1)は台座に支柱が設けられたもの、(2)は台座に支柱が設けられていないものを示している。

【図6】学習データセット作製に用いられる画像イメージ図

【図7】対象物の撮影イメージ図であり、画像中の対象物に関して、(1)は左方、(2)は中央、(3)は右方に位置する場合を示している。

【図8】バウンディングボックス設定の説明図

【図9】学習済みモデル生成のフロー図

【図10】物体認識及び位置姿勢推定時のロボット制御システムのシステム構成図

【図11】物体認識及び位置姿勢推定処理フロー図

【図12】物体認識及び位置姿勢推定処理に用いられる画像の一例

30

【図13】搬送装置と回転移動機構の説明図

【図14】対象物の把持位置計測用マーカの説明図

【図15】比較実験における撮影方法の説明図

【図16】位置推定に関する比較実験結果を表すグラフ

【発明を実施するための最良の形態】

【0028】

以下、本発明の実施形態の一例を、図面を参照しながら詳細に説明していく。なお、本発明の範囲は、以下の実施例や図示例に限定されるものではなく、幾多の変更及び変形が可能である。

【実施例1】

40

【0029】

図1は、学習データセットの作製処理の概略フロー及び物体認識及び位置姿勢推定処理の概略フローを示している。図1に示すように、学習データセット生成段階では、まず、位置姿勢検出用マーカであるARマーカ(2次元パターンマーカ)に対象物の物体情報を関連付けする(ステップS01)。次に、関連付けした2次元パターンマーカを用いた学習データセット生成治具を準備する(ステップS02)。学習データセット生成治具のベース部をガイドとして対象物を配置する(ステップS03)。対象物を配置した状態で、対象物の多視点画像群を取得する(ステップS04)。取得した画像群につき2次元パターンマーカを検出し認識する(ステップS05)。取得した画像群に対して対象物全体を取り囲むバウンディングボックスを設定する(ステップS06)。推定した対象物の姿勢

50

情報と重心位置情報、物体情報及びバウンディングボックスに関する情報を画像に関連付けする（ステップS07）。そして、多視点画像群を学習データセットとしてデータベースに登録する（ステップS08）。

なお、上記ステップS06のバウンディングボックスの設定において、可能な限り対象物全体を取り囲めればよく、多少、対象物を囲めていない場合であっても、学習データセットとして使用することができる。仮に対象物が他の物体で隠れていた場合であっても、隠れずに見えている範囲で囲めていけば構わない。

【0030】

次に、登録された学習データセットを用いて、深層学習などの機械学習により学習済みモデルを取得する（ステップS09）。

そして、物体認識及び位置姿勢推定段階では、まず撮像画像を取得し（ステップS10）、画像認識により撮像画像中の物体を検出して認識用バウンディングボックスを設定する（ステップS11）。なお、物体の名称は、物体を検出し、物体を認識した時点で取得できる。物体の認識用バウンディングボックスを切り出した部分画像を用いて、予め定義された基準座標系（実空間の3次元座標）における姿勢及び重心位置を推定する（ステップS12）。

基準座標系は、予め学習データセット生成治具に取り付けた位置姿勢検出用マーカを、基準としたい位置姿勢に配置することにより定義する。

【0031】

図2は、学習データセットの作製装置、物体認識及び位置姿勢推定装置の機能ブロック図を示している。図2に示すように、学習データセット作製装置10では、撮像手段20及び学習データセット生成治具30を用いて対象物の画像を撮影する。撮像された画像は、学習データセット作製装置10のコンピュータ40がARマーカ認識手段40aとして動作することにより、ARマーカが認識される。予めデータベース80には、ARマーカと認識対象となる対象物の物体情報とが関連付けされており、コンピュータ40が対象物認識手段40bとして動作することにより、対象物の物体名称など物体情報が認識される。ARマーカ認識手段40a及び対象物認識手段40bから得られた情報に基づいて、コンピュータ40がバウンディングボックス設定手段40cとして動作することにより、バウンディングボックスが設定される。また、コンピュータ40が対象物の姿勢推定手段40dとして動作することにより、対象物の姿勢が推定される。また、コンピュータ40が重心位置推定手段40eとして動作することにより、対象物の重心位置が推定される。取得された対象物に関する情報や、姿勢・重心位置に関する情報は、コンピュータ40がバウンディングボックス割付手段40fとして動作することにより、設定されたバウンディングボックスに割付けされ、学習データセットとしてデータベース80に保存される。

【0032】

学習データセットは、モデル生成コンピュータ9による機械学習により、分類器9a及び回帰器9bを作成し、学習済みモデルを取得する。

生成された学習済みモデルは、物体認識及び位置姿勢推定装置11において利用される。物体認識及び位置姿勢推定装置11では、撮像手段21により、物体が撮影される。得られた撮像画像は、コンピュータ41が物体認識手段41aとして動作することにより、画像中に含まれる物体が検出され、物体名称などの物体情報が取得される。検出された物体は、コンピュータ41が認識用バウンディングボックス設定手段41bとして動作することにより、認識用バウンディングボックスが設定される。認識用バウンディングボックスが設定されると、かかる認識用バウンディングボックス毎に、コンピュータ41が学習済みモデル41cとして動作することにより、姿勢及び重心位置が推定される。

なお、撮像手段20と撮像手段21、コンピュータ40とコンピュータ41、又は、データベース80の記憶手段とデータベース81の記憶手段は、いずれも、同一の機器を用いた構成であってもよい。

【0033】

（学習データセット作製について）

10

20

30

40

50

図3は、学習データセット作製の際のシステム構成図を示している。図3に示すように、学習データセット作製の際は、1台のカメラ2、学習データセット生成治具である台座14、基準マーカ33、台座14とその上の対象物5の位置を動かすベルトコンベア12及びコンピュータ4から成る。

学習データセット生成治具である台座14は、ARマーカ3、支柱18及び台座部17から成り、ARマーカ3と台座部17は、支柱18を介して接続されている。台座部17上には、認識対象となる対象物5が配置されている。対象物5の撮影は上方に設けられたカメラ2によって行う。撮影時にARマーカ3が対象物5に隠れることを防止するため、ARマーカ3は対象物5よりも高い位置に設けられている。

基準マーカ33は、ベルトコンベア12の外側に配置され固定された位置に置かれる。基準マーカ33は、学習データセット生成治具である台座14をそのまま利用可能であり、台座14のARマーカ3が、基準マーカ33として用いることができる。基準マーカ33を、基準としたい位置及び姿勢に配置することにより、基準座標系を定義する。

コンピュータ4には、データベース8が設けられており、カメラ2において撮像した画像群、ARマーカ3、対象物5に関する位置姿勢の情報が保存される。

【0034】

図4は、学習データセット生成のフロー図を示している。図4に示すように、まず、カメラ2によって、台座部17上に配置された対象物5を撮影し、1枚の撮像画像をコンピュータ4に入力する(ステップS101)。コンピュータ4では、取得した画像に含まれるARマーカを画像解析により検出する(ステップS102)。データベース8に登録されているARマーカに関連付けされた物体情報には、対象物に関する一般名称などの物体を識別するための情報が登録されている。次に、データベース8から、ARマーカに関連付けされた物体情報を取得する(ステップS103)。データベース8に登録されているARマーカに関連付けされた情報には、対象物の大きさ(サイズ)に関する情報が登録されている。ARマーカの姿勢情報に基づいて、対象物の姿勢を推定する(ステップS104)。さらに、対象物の重心位置を算出する(ステップS105)。ここで、対象物の重心位置とは、対象物の形状および学習データセット生成治具への配置から算出する重心位置のことであり、基準座標系における対象物重心の位置(x, y, z)情報を基に算出する。なお、重心位置の算出については、予め物体情報として、物体の高さと学習データセット生成治具への配置位置及び配置姿勢に加えて、学習データセット生成治具の情報として、治具の高さと治具へのARマーカの配置位置及び姿勢を所持しており、それらを用いて物体の重心位置を算出するものでも構わない。

【0035】

そして、対象物を取り囲むバウンディングボックスを設定し(ステップS106)、物体情報、姿勢、重心位置及び設定したバウンディングボックス情報(画像中の位置、幅、高さ)を画像に関連付けする(ステップS107)。画像内に未だ処理していないARマーカが存在する場合には、未処理のARマーカを同様に処理する。画像内の全てのARマーカが処理済みの場合(ステップS108)は、対象物の物体情報、姿勢・重心位置情報及びバウンディングボックス情報が紐付いた1枚の画像を出力する(ステップS109)。画像は複数回撮影され、全ての画像につき処理が完了することで、学習データセットが作製される(ステップS110)。

【0036】

ここで、学習データセット作製の際、ARマーカと対象物の関係について、図5を参照しながら説明する。図5は、物体認識及び位置姿勢推定の説明図であり、(1)は台座に支柱が設けられたもの、(2)は台座に支柱が設けられていないものを示している。図5(1)に示すように、ARマーカ3には、支柱18を介して台座部17が設けられており、台座部17上には、認識対象となる対象物5が配置されている。台座部17と支柱18は固着されているが、ARマーカ3は、支柱18から脱着自在に固定されているため、異なるARマーカを取り付けて使用することが可能である。

ARマーカ3が支柱18に取り付けられると、台座部17とARマーカ3の相対的位置

及び姿勢が固定されることになる。また、対象物 5 を、一定の基準を設定した後に、台座部 17 上に配置すると、対象物 5 と台座部 17 の相対的位置及び姿勢が固定されることになる。これにより、対象物 5 を台座部 17 上に配置するだけで、容易に A R マーカ 3 と対象物 5 の間の相対的位置及び姿勢を固定することができる。

【 0 0 3 7 】

ここで、対象物 5 を台座部 17 に配置する際に設定する一定の基準とは、例えば、台座の長辺と対象物の長辺を略並行にして配置するといったものである。ここで設定する基準については、深層学習などの機械学習をより効率的に行うために設定するものであるから、配置の際に必ずしも厳密に守られる必要は無く、例えば、やや斜めにずれて配置されたといった場合でも、かかるデータも後の機械学習の際に用いることはできる。

10

【 0 0 3 8 】

また、台座に支柱が設けられることは必須ではない。支柱 18 は、A R マーカ 3 の撮像を容易にするために設けられたものであるため、A R マーカ 3 の撮像に差し支えない場合には、支柱 18 は必須ではないからである。例えば、図 5 (2) に示すように、台座部 17 に直接 A R マーカ 3 が取り付けられた場合でも、対象物 50 が偏平形状であるような場合には、矢印 (19 a , 19 b , 19 c) の何れかの方向から撮影されても、A R マーカ 3 の撮像には差し支えないため、支柱は設けられなくてもよい。

【 0 0 3 9 】

データベース 8 には、予め、A R マーカの種類・形状、A R マーカに対応する対象物の名称、サイズとして、長辺 (mm)、短辺 (mm) 及び高さ (mm)、並びに、基準座標系における対象物の姿勢 (deg) が登録されている。下記表 1 はデータベースのテーブルの例を示している。表 1 において、対象物コードは対象物の識別コードとして用いている。

20

【 0 0 4 0 】

【表 1】

対象物 コード	マーカ コード	対象物 長辺 [mm]	対象物 短辺 [mm]	対象物 高さ [mm]	基準座標系に おける対象物 の重心位置 (x , y , z)	基準座標系 における対象 物の姿勢 [deg]
1	3	100	50	100	(20 , 30 , -150)	0
2	2	200	150	10	(10 , 20 , -30)	0
3	1	300	250	50	(12 , 28 , -80)	10
...

30

【 0 0 4 1 】

上記表 1 に示すように、対象物コード 1 については、マーカコードは 3、対象物長辺は 100 mm、対象物短辺は 50 mm、対象物高さは 100 mm、基準座標系における対象物重心の位置は、x が 20、y が 30、z が -150、基準座標系における対象物の姿勢は 0 (deg) とデータベース上に登録されている。また、対象物コード 2 については、マーカコードは 2、対象物長辺は 200 mm、対象物短辺は 150 mm、対象物高さは 10 mm、基準座標系における対象物重心の位置は、x が 10、y が 20、z が -30、基準座標系における対象物の姿勢は 0 (deg) と登録されている。対象物コード 3 については、マーカコードは 1、対象物長辺は 300 mm、対象物短辺は 250 mm、対象物高さは 50 mm、基準座標系における対象物重心の位置は、x が 12、y が 28、z が -80、基準座標系における対象物の姿勢は 10 (deg) と登録されている。以下同様に、対象物毎に各データが登録される。

40

【 0 0 4 2 】

50

そのため、ARマーカ3と対象物5の間の相対的な位置姿勢関係が固定された状態で、カメラ2を用いてARマーカ3及び対象物5を撮影すると、コンピュータ4は、ARマーカ3を検出することで、ARマーカの種類がARマーカ3であること、ARマーカ3の位置姿勢及び対応する対象物5の名称などの情報が取得される。

また、ARマーカ3と台座部17と対象物5の間の相対的な位置姿勢関係から、基準座標系における対象物5の姿勢を推定する。また、ARマーカ3と台座部17と対象物5の間の相対的位置の関係と、対象物5および台座部17のサイズから、基準座標系における対象物5の重心位置を推定する。

データベース8に予め登録されている対象物の大きさに関するデータを用いて、上記の推定された位置や姿勢のデータに基づいて、対象物5を囲むバウンディングボックスを設定する。

10

【0043】

1つの対象物5につき、アングルを変えて、複数回の撮影を行うことで、多数のパターンの出力画像が得られることになる。例えば、図5に示すように、ARマーカ3と対象物5の間の相対的な位置姿勢関係が固定された状態でも、矢印19aから撮影する場合と、矢印19b又は矢印19cから撮影する場合とでは、取得される画像が異なるからである。撮影方法としては、本実施例では、1台のカメラ2を固定した上で、台座部17及び対象物5を移動させながら、複数回の撮影を行っている。この他、台座部17及び対象物5を固定したままで、1台のカメラ2を移動して撮影してもよく、或は、複数のカメラを用いて異なるアングルから撮影してもよい。

20

【0044】

図6は、学習データセット作成に用いられる画像イメージ図を示している。図6に示す画像7aは、図2で示すカメラ2によって撮影された1枚の画像である。図6に示すように、画像7aには、ベルトコンベア12、対象物(5a~5c)、台座(13a~13c)及びバウンディングボックス(6a~6c)が表示されている。

対象物5aは、台座14aの台座部17a上に配置されている。同様に、対象物5bは、台座14bの台座部17b上に、対象物5cは、台座14cの台座部17c上にそれぞれ配置されている。対象物5aは紙製容器入り食品であり、容器を立てた状態で配置しているが、寝かせた状態で配置してもよい。なお、対象物5bはビニール袋に入った軟質の食品であり、対象物5cは筒状容器入り食品である。ここでは、対象物としては食品を用いているが、食品以外の対象物でもよい。

30

矢印19dは、ベルトコンベア12が流れる方向を示したものであり、画像7aは、ベルトコンベア12が矢印19dの方向に流れる中で撮影したものである。なお、画像7aでは、上記以外の物体等は表示されていないが、背景に別の装置等の物体が配置され、映り込んでいても構わない。

【0045】

図6に示すように、例えば、対象物5aは、ベルトコンベア12の移動方向に並行に配置されているが、対象物(5b, 5c)はベルトコンベア12の移動方向に対して斜めに配置されている。このようにあらゆる向きに対象物が配置されていても、対象物とARマーカの相対的な位置姿勢関係が固定されていればよい。また、同一の対象物についてもあらゆる向きに設置してカメラ2による撮影を行うことで、多様なパターンの画像が得られ、後の深層学習などの機械学習により、より精度の高いモデルの生成が可能となる。

40

ここで、対象物(5a~5c)の撮影は、ベルトコンベア12上を一度流される際に、複数回行われる。

【0046】

図7は、対象物の撮影イメージ図であり、画像中の対象物につき(1)は左方、(2)は中央、(3)は右方に位置する場合を示している。図7(1)に示すように、対象物15は、略立方体形状の物体であるが、画像中の左方に位置する場合には、面(15a~15c)が画像中に写り込んでいる。これに対して、図7(2)に示すように、対象物15が画像中の中央に位置する場合には、面(15a, 15b)は画像中に写り込んでいるが

50

、面 15c は画像中に写り込んでいない。また、図 7(3) に示すように、対象物 15 が画像中の右方に位置する場合には、面 (15a ~ 15d) が画像中に写り込んでいる。このように、同じ対象物 15 であっても、撮影する角度によって異なる画像が得られるため、ベルトコンベア 12 上を一度流すだけでも、多様なパターンの物体画像が得られることになる。さらに、対象物 (5a ~ 5c) の設置の向きを変えて、さらに複数回ベルトコンベア 12 上を流され、撮影が行われる。

なお、図 7(1) ~ (3) に示すように、対象物 15 と共に AR マーカ 3 も、対象物 15 との相対的位置及び姿勢が固定された状態で移動するため、AR マーカ 3 の位置や姿勢も変化することとなる。

【0047】

以上のような方法で撮影が行われると、図 6 に示す画像 7a においては、画像データが取得され、各画像からまず、AR マーカ (3a ~ 3c) の種類と位置姿勢が検出される。AR マーカ (3a ~ 3c) については、予め、AR マーカの種類毎に、対象物の名称、長辺 (mm)、短辺 (mm)、高さ (mm)、及び、基準座標系における対象物の姿勢 (deg) が登録されているため、AR マーカ (3a ~ 3c) の検出により、対象物の名称、長辺 (mm)、短辺 (mm)、高さ (mm)、及び、基準座標系における対象物の姿勢 (deg) が取得される。ここでは、AR マーカ 3a の検出により、対象物 5a の名称等が取得される。同様に、AR マーカ 3b の検出により、対象物 5b の名称等が取得され、AR マーカ 3c の検出により、対象物 5c の名称等が取得される。

AR マーカ (3a ~ 3c) に定義された位置姿勢情報に基づいて、対象物 (5a ~ 5c) の位置姿勢情報を推定する。

【0048】

具体的には、まず、AR マーカ 3a に定義された姿勢情報に基づいて対象物 5a の姿勢を算出する。同様に、AR マーカ 3b に定義された姿勢情報に基づいて対象物 5b の姿勢を算出し、AR マーカ 3c に定義された姿勢情報に基づいて対象物 5c の姿勢を算出する。

【0049】

次に、AR マーカ 3a に定義された位置情報と、AR マーカ 3a と対象物 5a の相対的な位置関係の情報に基づいて対象物 5a の重心位置を算出する。同様に、AR マーカ 3b に定義された位置情報と、AR マーカ 3b と対象物 5b の相対的な位置関係の情報に基づいて対象物 5b の重心位置を算出し、AR マーカ 3c に定義された位置情報と、AR マーカ 3c と対象物 5c の相対的な位置関係の情報に基づいて対象物 5c の重心位置を算出する。

【0050】

上記の算出されたデータと、対象物 (5a ~ 5c) の大きさに関するデータを用いて、対象物 (5a ~ 5c) を囲むバウンディングボックス (6a ~ 6c) を生成する。ここで、例えば、AR マーカ (3a, 3c) は、それぞれ対象物 (5a, 5c) を囲むバウンディングボックス (6a, 6c) からはみ出ているが、バウンディングボックスは対象物を囲むことができればよく、問題とはならない。

【0051】

図 8 は、バウンディングボックス設定の説明図を示している。図 8 は、図 6 における画像 7a におけるバウンディングボックス 6b を拡大したものである。

図 8 に示すように、対象物 5b を長方形として認識し、例えば、下記数 1 をように、矩形 (四角形) の幅 (x') と高さ (y') にそれぞれマージンとなるオフセットを加算して、バウンディングボックスの幅と高さを決定し、バウンディングボックスを設定する。ここで、 h は物体の高さ、 θ は鉛直方向に対するカメラの設置角度である。

【0052】

10

20

30

40

【数 1】

$$x' = x \cos \theta + y \sin \theta$$

$$y' = y \cos \theta + x \sin \theta + h \sin \phi$$

【0053】

なお、図 8 に示すように、バウンディングボックス 6 b を設定する際には、上記のオフセットを設けるため、上記数 1 により算出されたバウンディングボックスよりも、面積は広く設けられる。

【0054】

図 6 では、画像 7 a を例に説明をしているが、ベルトコンベア 1 2 を稼動しながら、同様に、撮影を重ねることで、画像 (7 b ~ 7 f) のように、複数の画像を取得することができる。取得する画像は、画像 (7 a ~ 7 f) に限られず、さらに多数の画像が取得される。

このように、バウンディングボックス (6 a ~ 6 c) が挿入された画像 7 a のような画像が集積することで、学習データセットが作製される。

【0055】

(機械学習について)

学習データセットが生成された後、学習データセットを具体的なロボット制御に利用する場合には、先立って学習データセットを使用した深層学習などの機械学習が必要となる。そこで、コンピュータ 4 は、データベース 8 に保存された学習データセットを用いて、推定を行う対象物について、深層学習を行い、学習済みモデルを取得する。

図 9 は、学習済みモデル生成のフロー図を示している。図 9 に示すように、まず学習データセットを入力する (ステップ S 2 1)。入力された学習データセットを基に、深層学習を行う (ステップ S 2 2)。本実施例では、Google (登録商標) が開発しオープンソースとして公開した人工知能のソフトウェアライブラリである TensorFlow (登録商標) を使用して深層学習を行う。深層学習により得られた学習済みモデルを出力する (ステップ S 2 3)。

【0056】

(物体認識及び位置姿勢推定について)

図 10 は、物体認識及び位置姿勢推定時のロボット制御システムのシステム構成図を示している。図 10 に示すように、物体認識及び位置姿勢推定時においては、ロボット制御システム 1 は、カメラ 2、コンピュータ 4 及びロボット 1 3 から成る。

ベルトコンベア 1 2 上には、認識対象となる対象物 5 が配置されている。コンピュータ 4 には、データベース 8 が設けられており、カメラ 2 において撮影した画像等が保存される。ロボット 1 3 には、ロボットアーム 1 3 a が設けられており、ロボットアーム 1 3 a によって、対象物を把持できる構造となっている。

【0057】

図 11 は、物体認識及び位置姿勢推定フロー図を示している。図 11 に示すように、カメラ 2 を用いて、ベルトコンベア 1 2 上を撮影し、1 枚の撮像画像を入力する (ステップ S 3 0 1)。画像認識により、物体を検出し、画像座標系における位置を検出する (ステップ S 3 0 2)。ここで、物体情報 (名称など) が取得される。また、認識用バウンディングボックスを設定する (ステップ S 3 0 3)。設定された認識用バウンディングボックスの内、いずれか 1 つを選択する (ステップ S 3 0 4)。撮像画像から、認識用バウンディングボックス内を切り抜く (ステップ S 3 0 5)。そして、深層学習など機械学習によって得られた学習済みモデルの分類器を用いて、物体の姿勢を推定する (ステップ S 3 0 6)。次に、認識用バウンディングボックスの幅、高さを取得し (ステップ S 3 0 7)、学習済みモデルの回帰器により、実空間の 3 次元座標における物体の重心位置を推定する (ステップ S 3 0 8)。認識用バウンディングボックスの内、選択していないものが存在するか否かを判断し (ステップ S 3 0 9)、未選択の認識用バウンディングボックスがあ

10

20

30

40

50

れば、いずれか1つを選択する。認識用バウンディングボックスを全て選択し終えた場合は、検出した全ての物体の名称、姿勢、実空間の3次元座標における重心位置を出力する(ステップS310)。

【0058】

図12は、物体認識及び位置姿勢推定に用いられる画像の一例を示している。図12に示すように、画像70上には、ベルトコンベア12、対象物(5a~5c)及び認識用バウンディングボックス(16a~16c)が表示されている。画像70は、ベルトコンベア12上に対象物(5a~5c)を並べて、矢印19dに示す方向に対象物(5a~5c)を流し、かかる状態で、カメラ2を用いて撮影したものである。

学習済みモデルの利用に先立って、学習データセットを用いた深層学習が行われたことにより、対象物を認識し、位置及び姿勢を推定することが可能となっている。

10

【0059】

図12においては、対象物(5a~5c)を囲むように認識用バウンディングボックス(16a~16c)が表示されているのみであるが、コンピュータ4においては、対象物(5a~5c)の名称や、姿勢・重心位置に関する情報が推定されており、ロボット13の把持動作を予め計画しておくことで、コンピュータ4において推定された情報を基に、ロボット13に対象物(5a~5c)の仕分け作業を行わせるといったことが可能である。具体的には、例えば、詰め合わせギフトのように、所定の箱に複数種類の商品を1つずつ配置したり、レーンを通る多数の物品の中から不良品を除去したりといった作業が可能である。

20

【実施例2】

【0060】

搬送装置と回転移動機構について図13を参照して説明する。図13に示すコンベア120は、コンベアベルト120aを水平に移動させるものである。コンベア120は、搬送装置であり、コンベアベルト120aの上に載せた物体を水平に移動する。

コンベアベルト120aの上に、学習データセット生成治具の台座14を載せ、台座部17をガイドとして対象物5を配置した状態で、コンベアベルト120aを水平(矢印19fに示す方向)に移動させ、ARマーカ3を含む物体5全体の多視点画像群を効率よく取得することができる。

さらに、図13に示すように、回転機構120bと直動機構120cを組合せた装置に、学習データセット生成治具の台座14を取付け、当該装置をコンベアベルト120aの上に載せて、コンベアベルト120aを水平に移動させることにより、更に効率よく多視点画像群を取得することができる。すなわち、コンベア120aが矢印19fに示す方向に移動しながら、回転機構120bが矢印19gに示す方向に水平に回転し、さらに、直動機構120cが矢印19hに示す方向に水平に移動することで、マーカを含む物体全体の多視点画像群を効率よく取得するのである。これにより、対象物5及び台座14をコンベアベルト120a上に一度流すだけで、あらゆる角度から撮像することが可能であり、短時間で、効率的に学習データセットを作製することができる。

30

なお、矢印19fに示す移動方向や矢印19gに示す回転方向は、逆方向であってもよく、更には、正方向と逆方向に交互に動くものであってもよい。また、矢印19hに示す移動方向は往復移動であるが、片方向のみに移動するものでもよい。

40

【実施例3】

【0061】

学習データセット作製方法の一部の処理をコンピュータに実行させるための学習データセット作製プログラムについて図4を参照して説明する。学習データセット作製プログラムは、1)多視点画像群に対して対象物のバウンディングボックスを設定するステップと、2)撮像画像から推定した対象物の姿勢情報と重心位置情報、物体情報及びバウンディングボックスに関する情報を、撮像画像に関連付けして、学習データセットを生成するステップをコンピュータに実行させる。

上記1)のステップでは、取得した画像に含まれるARマーカを画像解析により検出し

50

(ステップS102)、ARマーカに関連付けされた物体情報を取得し(ステップS103)、ARマーカの姿勢情報に基づいて、対象物の姿勢を推定し(ステップS104)、対象物の重心位置を算出し(ステップS105)、対象物を取り囲むバウンディングボックスを設定する(ステップS106)。

上記2)のステップでは、物体情報、姿勢、重心位置及び設定したバウンディングボックス情報(画像中の位置、幅、高さ)を画像に関連付けし(ステップS107)、画像内に未だ処理していないARマーカが存在する場合には、未処理のARマーカを同様に処理し、画像内の全てのARマーカが処理済みの場合(ステップS108)には、対象物の物体情報、姿勢・重心位置情報及びバウンディングボックス情報が紐付いた1枚の画像を出力し(ステップS109)、全ての撮像画像につき処理を行い、学習データセットを作製する(ステップS110)。

10

【0062】

また、物体認識及び位置姿勢推定方法の一部の処理をコンピュータに実行させるための物体認識及び位置姿勢推定プログラムについて図11を参照して説明する。物体認識及び位置姿勢推定プログラムは、a)新たに撮像された撮像画像から、画像認識によって物体の名称を検出して認識用バウンディングボックスを作成するステップと、b)認識用バウンディングボックスを切り出した部分画像を生成し、生成した部分画像に基づき学習済みモデルを用いて、検出した物体の姿勢を推定するステップと、c)認識用バウンディングボックスの重心位置、幅及び高さを取得し、学習済みモデルを用いて、検出した物体の重心位置を推定するステップをコンピュータに実行させる。

20

上記a)のステップでは、画像認識により、物体を検出し、画像座標系における位置を検出し(ステップS302)、物体情報(名称など)を取得して、認識用バウンディングボックスを設定する(ステップS303)。

上記b)のステップでは、設定された認識用バウンディングボックスの内、いずれか1つを選択し(ステップS304)、撮像画像から、認識用バウンディングボックス内を切り抜き(ステップS305)、深層学習など機械学習によって得られた学習済みモデルの分類器を用いて、物体の姿勢を推定する(ステップS306)。

上記c)のステップでは、認識用バウンディングボックスの幅、高さを取得し(ステップS307)、学習済みモデルの回帰器により、実空間の3次元座標における物体の重心位置を推定する(ステップS308)。

30

【実施例4】

【0063】

学習データセット生成治具のARマーカと、基本座標系における対象物の重心位置との相対位置関係について、実施例1では手入力で予めデータベースのテーブルに設定している(実施例1の表1を参照)。しかしながら、実際にロボットを用いて対象物を把持させる場合には、対象物の外周面において、把持できる箇所を推定する必要がある。

そこで、例えば、親指及び人差し指の指先と手の第一指間部の3ヶ所に装着し、学習データセット生成治具に配置しているARマーカと、人が把持する場合の親指及び人差し指の指先と手の第一指間部に装着するARマーカとの相対的な位置姿勢の関係を予め計測する。なお、人の手に取り付けるマーカの位置と数は任意に決定することができる。

40

そして、対象物の重心位置は、例えば、3ヶ所に装着したARマーカの位置情報を円周上にあると仮定し、近似したときの円の中心位置を、把持する重心位置とし、データベースに、学習データセット生成治具のARマーカと、基本座標系における対象物の重心位置との相対位置関係として登録する。

【0064】

図14に示すように、例えば、飲料容器のペットボトル36を対象物として想定すると、円柱状のペットボトル36の中央辺りを、人の手34の親指の指先と人差し指の指先とで挟んで掴むことになる。親指及び人差し指の指先と手の第一指間部の3ヶ所に装着したARマーカ(把持位置計測用マーカ:35a~35c)の位置情報は、学習データセット生成治具14に配置しているARマーカ3と、人が把持する場合の親指及び人差し指の指

50

先と手の第一指間部に装着するARマーカ(35a~35c)との相対的な位置姿勢関係により計測する。そして、計測した3ヶ所の位置が円周上にあるとして近似したときの円の中心位置を、ペットボトル36の把持する重心位置とする。

また、基準座標系における対象物の姿勢もデータベースに登録はする。これはロボットを使用して実際に把持を行う際に活用する。

学習データセット生成治具14のARマーカ3と、親指及び人差し指の指先と手の第一指間部に装着するARマーカ(35a~35c)との相対的な位置姿勢の計測方法としては、把持位置、すなわち、把持位置計測用マーカ(35a~35c)の位置が固定された状態で、1台のカメラで対象物の把持の姿勢を複数回撮像もしくは複数台のカメラで同時に撮像する。その際、撮像を行うカメラ位置姿勢は、学習データセット生成治具14のARマーカ3との相対位置姿勢のみ使用するため、任意のカメラ位置姿勢で構わない。

【0065】

上述の計測を実際に使用する工場の自動化システムにおけるロボットハンドを用いて、同様に行うことができる。例えば、ロボットハンドで把持する場合、ロボットハンドの第1指先と第2指先と第一指間部の3ヶ所に装着して計測する。

すなわち、人間もしくはロボットハンドを使用することで、学習データセット生成治具のARマーカに対する相対位置の設定と、ロボットによる把持の際に使用する情報をデータベースに登録することができる。

【実施例5】

【0066】

ARマーカによる学習データ収集と人手による学習データ収集の2つの場合について、学習データセットの生成時間、物体認識精度、及び、位置推定精度の比較を行った結果について以下に説明する。

まず、撮像画像を得るためのカメラとしては、RGBカメラ(ポイントグレイリサーチ株式会社製:Flea3 FL3-U3-88S2C)を使用し、工場環境を再現するために、オークラ輸送機株式会社(登録商標)製のコンベアベルトを使用した。

画像処理ライブラリは、OpenCV 3.2.0を使用した。また、物体認識モデルのニューラルネットワークは、Tensor Flow(登録商標)(tensorflow-gpu 1.3.0)上で動作するKeras 1.2.2を使用した。回帰モデルは、位置を推定するためにscikit-learn 0.19.0を使用した。

実験に用いた対象物は、背の高い物体としてボトルA及びボトルBを選択し、変形可能な物体としてパウチA及びパウチBを選択し、背の低い物体として缶A及び缶Bを選択した。なお、ボトルA及びボトルBとは、ペットボトルのことである。

実験に使用したARマーカは、一辺が40mmの正方形のARマーカであり、対象物とARマーカが重なることを防ぐべく、台座部から高さ200mmの位置に設けた。ARマーカの検出には、ArUcoと呼ばれるARライブラリを用いて行った。

【0067】

(学習データセットの生成時間の評価結果)

比較実験における撮影方法の説明図を図15に示す。図に示すように、撮影の際は、データの偏りが起こらないように、3×7のマトリックスのポイント38を画像71中に設定し、それぞれのポイント毎に配置して、撮影を行った。縦横のポイント同士の間隔はいずれも100mmである。

撮影は2人で行い、いずれの場合も、1人が物体の姿勢を変化させ、もう1人がカメラで撮影した。モデルをトレーニングするために400画像、モデルの検証に100画像、テストに100画像を使用した。

人手によるアノテーションは2人で行った。具体的には、グラフィカルにアノテーション可能なツールである“LabelImg”を使用した。各対象物の周囲にバウンディングボックスを作成し、そのオブジェクト名を同時に入力を行った。また、全ての対象物の位置を手動で割り当てた。

ここで、アノテーションとは、撮像画像に対するバウンディングボックスの設定だけで

10

20

30

40

50

はなく、物体情報の関連付けも含めたものを意味する用語である。

【0068】

下記表2は、ARマーカ及び人手による場合の双方について、学習データセット生成時間の測定結果を表している。比較対象とする撮像画像とすべく、ARマーカ及び人手による場合のそれぞれにつき500枚撮影した。

【0069】

【表2】

	ARマーカ	人手
撮影時間(分)	75	80
アノテーション時間(分)	37.23	886
合計時間(分)	112.23 (1.87時間)	966 (16.1時間)

10

20

【0070】

上記表2に示すように、ARマーカによる場合の撮影時間は75分、アノテーションにかかる時間は37.23分、合計時間は112.23分であった。また、人手による場合の撮影時間は80分、アノテーションにかかる時間は886分、合計時間は966分であった。

以上より、ARマーカによる学習データセットの生成は、人手による学習データセットの生成と比較して、88.4%の効率化に成功したといえる。また、アノテーション時間についても、ARマーカによる場合では、人手による場合に比べて5%未満に時間が短縮されている。

【0071】

30

(物体認識に関する比較実験結果)

下記表3は、ARマーカによる場合についての物体認識精度に関する測定結果を表し、下記表4は、人手による場合についての物体認識精度に関する測定結果を表している。対象物であるボトルA、ボトルB、パウチA、パウチB、缶A及び缶Bにつき、“F値”、“Precision”及び“Recall”を測定した。

ここで、“Recall”は、再現率であり、実際に正であるもののうち、正であると予測されたものの割合を意味し、見つけなければいけないものをどれだけ見つけたかを数値化したものである。“Recall”は、最小値が0%であり100%あるいは100%に近いほど優れている。また、“Precision”は、判定したものの正確さを表す指標であり、 $TP / (TP + FP)$ から算術する(TP : True Positive, FP : False Positive)。“Precision”も“Recall”と同じく、最小値が0%であり100%に近いほど優れている。また、“F値”は、機械学習における予測結果の評価尺度の一つであり、精度と再現率の調和平均を示している。具体的には、F値は、Recallの値を“A”、Precisionの値を“B”とすると、 $2 \times B \times A / (A + B)$ から算術する。F値も、最小値が0%であり100%に近いほど優れている。

40

【0072】

【表 3】

	F値(%)	Precision(%)	Recall(%)
ボトルA	100	100	100
ボトルB	98	99	98
パウチA	79	83	75
パウチB	83	84	82
缶A	82	91	75
缶B	81	94	71
平均	87	92	84

10

【0073】

【表 4】

	F値(%)	Precision(%)	Recall(%)
ボトルA	99	98	100
ボトルB	99	98	99
パウチA	72	88	61
パウチB	80	71	92
缶A	82	79	85
缶B	77	93	66
平均	85	88	84

20

30

【0074】

上記表 3 及び 4 に示すように、AR マーカによる場合の F 値の平均は 87%、人手による場合の F 値の平均は 85% であり、AR マーカによる場合は人手による場合以上の認識性能となることを確認できる。

【0075】

(位置推定に関する比較実験結果)

図 16 は、位置推定に関する比較実験結果を表すグラフであり、AR マーカによる場合及び人手による場合についての位置推定誤差を表すグラフである。なお、図 16 では、ボトル A、ボトル B、パウチ A、パウチ B、缶 A 及び缶 B につき比較しており、グラフの数値は小数第二位を四捨五入して説明する。

40

【0076】

図 16 に示すように、人手による場合の位置推定誤差は、ボトル A が 28.0 mm、ボトル B が 27.0 mm、パウチ A が 23.6 mm、パウチ B が 29.4 mm、缶 A が 13.3 mm、缶 B が 15.2 mm であり、これに対して、AR マーカによる場合は、ボトル A が 28.3 mm、ボトル B が 28.4 mm、パウチ A が 46.8 mm、パウチ B が 51.1 mm、缶 A が 27.7 mm、缶 B が 32.1 mm であった。

したがって、いずれの対象物についても、人手による場合の方が AR マーカによる場合よりも誤差は小さかったが、ボトル A 及びボトル B については、AR マーカによる場合でも人手による場合と殆ど変わらない誤差であることが判った。

50

【 0 0 7 7 】

(実験結果のまとめ)

以上より、ARマーカを用いた場合は、88.4%のデータ収集時間の効率化が図られ、物体認識においては、十分な性能が得られることが判った。また、位置推定についても物体の形状によっては、十分な性能が得られることが判った。

【 産業上の利用可能性 】

【 0 0 7 8 】

本発明は、産業用ロボットによる工場の自動化システムとして有用である。すなわち、本発明の学習データセット作製方法により作製した学習データセットを用いて機械学習を行った産業用ロボット、本発明の物体認識及び位置姿勢推定方法を用いた産業用ロボット、或は、本発明の物体認識及び位置姿勢推定プログラムが搭載された産業用ロボットは、食品業界や物流業界における仕分け作業や加工作業などの工場の自動化システムに利用可能である。また、建物等における防犯システムにも利用可能である。

10

【 符号の説明 】

【 0 0 7 9 】

- 1 ロボット制御システム
- 2 カメラ
- 3 , 3 a ~ 3 c ARマーカ
- 4 , 4 0 , 4 1 コンピュータ
- 5 , 5 a ~ 5 c , 1 5 , 5 0 対象物
- 6 a ~ 6 c バウンディングボックス
- 7 a ~ 7 f , 7 0 , 7 1 画像
- 8 , 8 0 , 8 1 データベース
- 9 モデル生成コンピュータ
- 9 a 分類器
- 9 b 回帰器
- 1 0 学習データセット生成手段
- 1 1 物体認識及び位置姿勢推定手段
- 1 2 ベルトコンベア
- 1 3 ロボット
- 1 3 a ロボットアーム
- 1 4 , 1 4 a ~ 1 4 c 学習データセット生成治具 (台座)
- 1 5 a ~ 1 5 d 面
- 1 6 a ~ 1 6 c 認識用バウンディングボックス
- 1 7 , 1 7 a ~ 1 7 c 台座部
- 1 8 支柱
- 1 9 a ~ 1 9 h 矢印
- 2 0 , 2 1 撮像手段
- 3 0 学習データセット生成治具
- 3 3 基準マーカ (ARマーカ)
- 3 4 人の手
- 3 5 a ~ 3 5 c 把持位置計測用マーカ (ARマーカ)
- 3 6 ペットボトル
- 3 8 ポイント
- 4 0 a ARマーカ認識手段
- 4 0 b 対象物認識手段
- 4 0 c バウンディングボックス設定手段
- 4 0 d 対象物の姿勢推定手段
- 4 0 e 重心位置推定手段
- 4 0 f バウンディングボックス割付手段

20

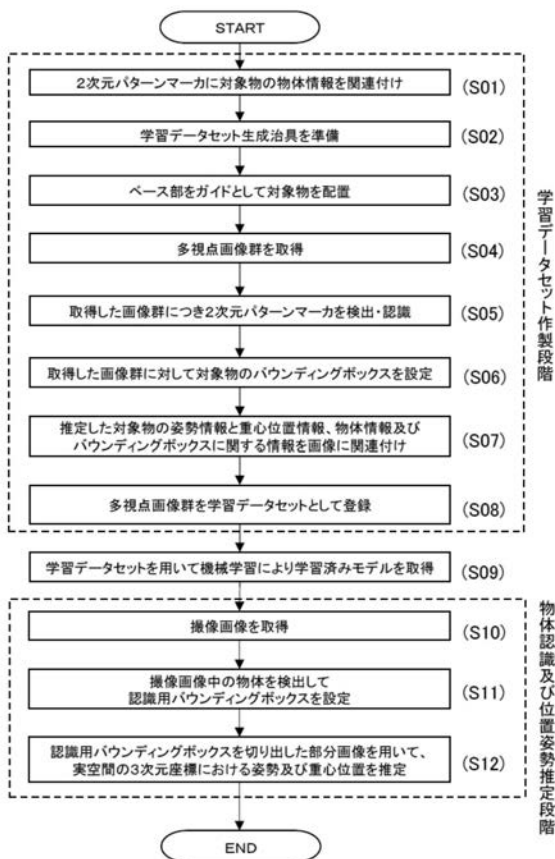
30

40

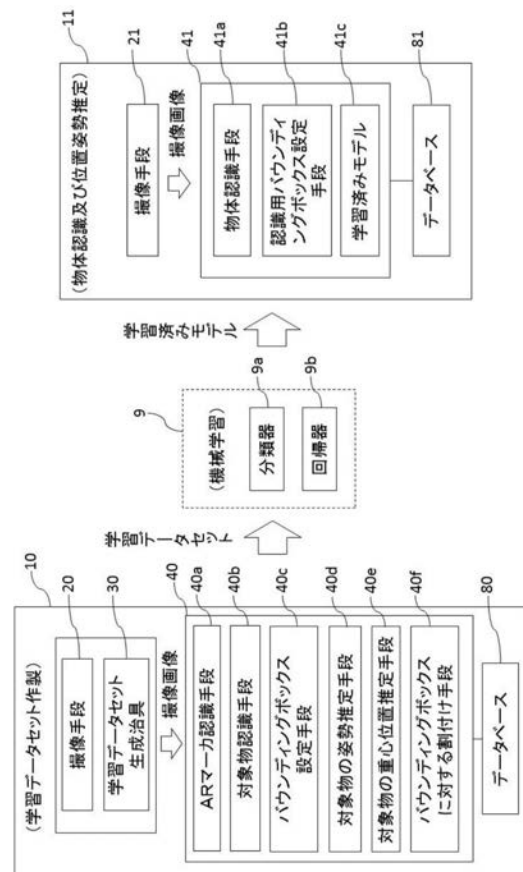
50

- 4 1 a 物体認識手段
- 4 1 b 認識用バウンディングボックス設定手段
- 4 1 c 学習済みモデル
- 1 2 0 コンペア
- 1 2 0 a コンペアベルト
- 1 2 0 b 回転機構
- 1 2 0 c 直動機構

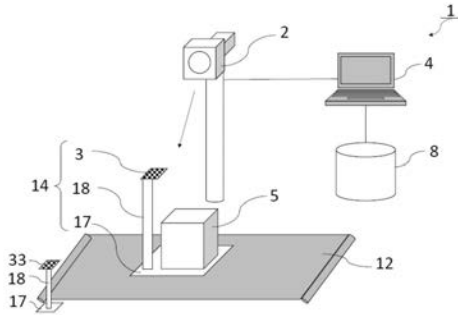
【 図 1 】



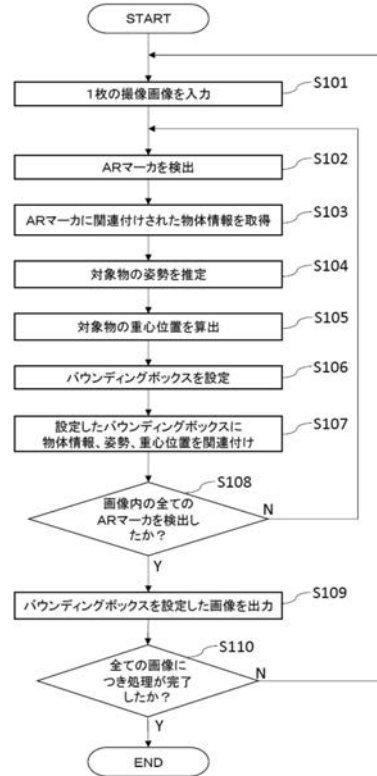
【 図 2 】



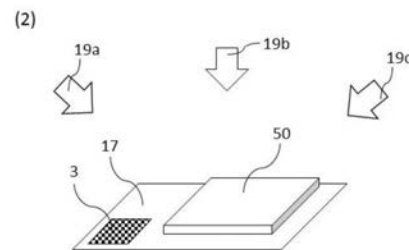
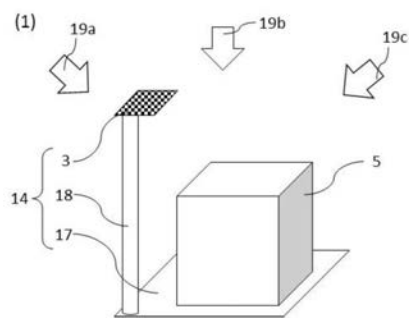
【 図 3 】



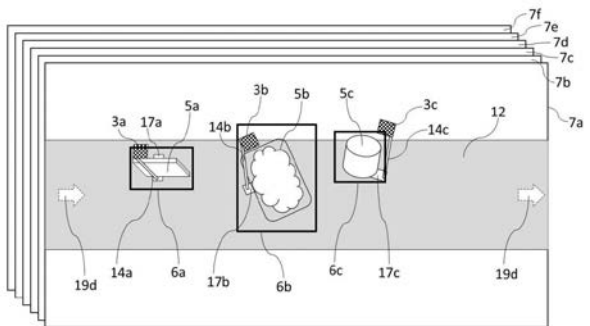
【 図 4 】



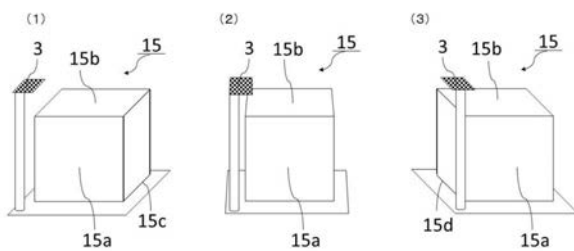
【 図 5 】



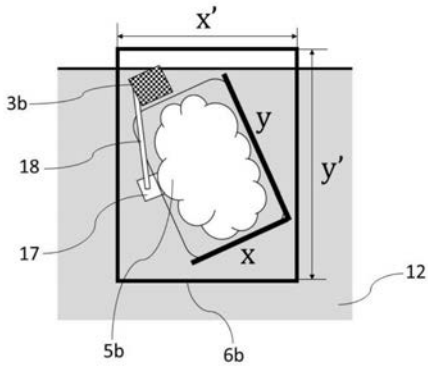
【 図 6 】



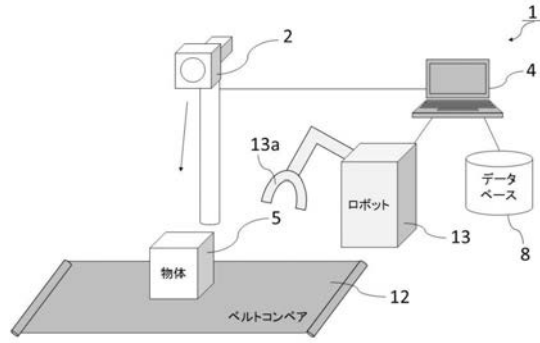
【 図 7 】



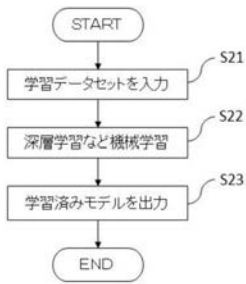
【 図 8 】



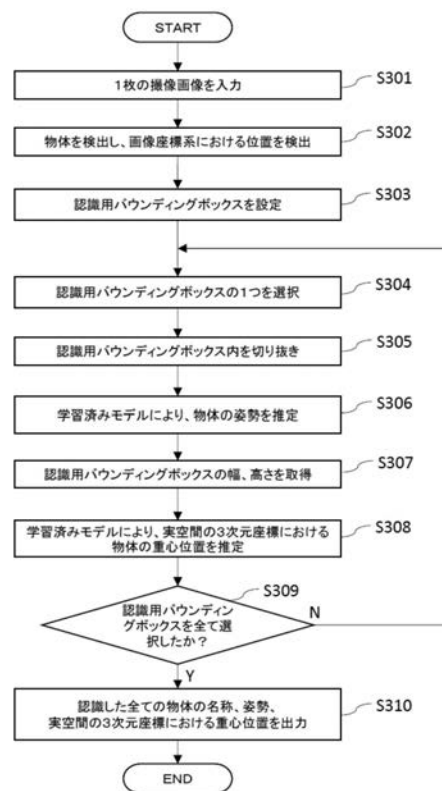
【 図 1 0 】



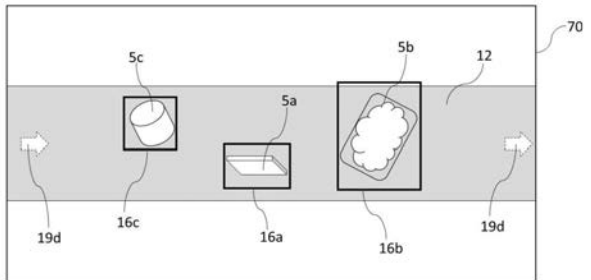
【 図 9 】



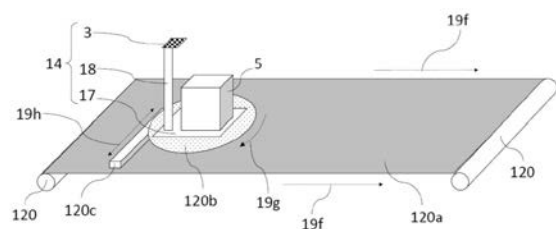
【 図 1 1 】



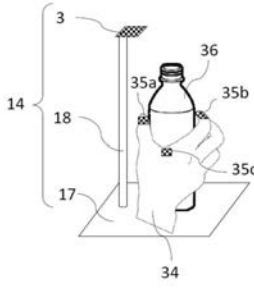
【 図 1 2 】



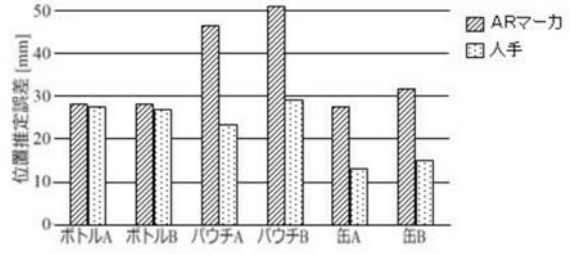
【 図 1 3 】



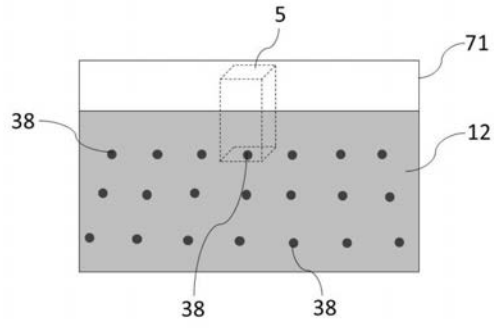
【図14】



【図16】



【図15】



フロントページの続き

(72)発明者 高松 淳

奈良県生駒市高山町 8 9 1 6 - 5 国立大学法人奈良先端科学技術大学院大学内

(72)発明者 丁 明

奈良県生駒市高山町 8 9 1 6 - 5 国立大学法人奈良先端科学技術大学院大学内

Fターム(参考) 3C707 KS03 KS04 KT01 KT06 LW12 NS02

5L096 AA06 BA05 CA04 DA02 FA60 FA64 FA67 FA69 HA08 KA04
KA15