

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7200610号
(P7200610)

(45)発行日 令和5年1月10日(2023.1.10)

(24)登録日 令和4年12月26日(2022.12.26)

(51)国際特許分類		F I	
G 0 6 T	7/70 (2017.01)	G 0 6 T	7/70 A
G 0 6 T	7/00 (2017.01)	G 0 6 T	7/00 3 5 0 C
G 0 6 N	3/045(2023.01)	G 0 6 N	3/04 1 5 4
B 2 5 J	13/08 (2006.01)	B 2 5 J	13/08 A

請求項の数 8 (全15頁)

(21)出願番号	特願2018-210460(P2018-210460)	(73)特許権者	000005223 富士通株式会社 神奈川県川崎市中原区上小田中4丁目1番1号
(22)出願日	平成30年11月8日(2018.11.8)	(74)代理人	110002147 弁理士法人酒井国際特許事務所
(65)公開番号	特開2020-77231(P2020-77231A)	(72)発明者	横田 泰斗 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
(43)公開日	令和2年5月21日(2020.5.21)	(72)発明者	鈴木 彼方 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内
審査請求日	令和3年8月10日(2021.8.10)	審査官	佐藤 実

最終頁に続く

(54)【発明の名称】 位置検出プログラム、位置検出方法及び位置検出装置

(57)【特許請求の範囲】

【請求項1】

コンピュータに、

それぞれが相似形でない各形状を示す単純形状データを入力して前記形状の位置情報を検出する、各形状に対応した各学習モデルに、撮影手段により物体が撮影された画像データを入力し、

前記各学習モデルから位置情報および検出の確からしさを示す信頼度を取得し、

前記各学習モデルから得られた各信頼度の合計値が閾値以上である場合、前記各学習モデルから得られた位置情報を用いて、前記物体の位置を検出する、

処理を実行させることを特徴とする位置検出プログラム。

10

【請求項2】

前記コンピュータに、

前記物体の位置から前記物体の把持位置を検出する

処理をさらに実行させることを特徴とする請求項1に記載の位置検出プログラム。

【請求項3】

前記各学習モデルそれぞれは、前記各形状の位置情報および前記各形状の位置情報の信頼度を出力するモデルであり、

前記検出する処理は、アンサンブル学習モデルを用いて前記物体の位置を特定し、

前記アンサンブル学習モデルは、前記各学習モデルそれぞれの検出結果のうち、前記信頼度から高い順に所定個の位置情報を特定し、前記各学習モデルそれぞれから所定個ずつ

20

取得された複数の位置情報を用いて前記物体の位置を検出する、ことを特徴とする請求項 1 又は 2 に記載の位置検出プログラム。

【請求項 4】

前記学習モデルは、深層学習を用いる畳み込み型のニューラルネットワークであって、検出結果として、前記画像データから前記物体の検出位置を示すバウンディングボックスと前記信頼度を対応付けて複数出力し、

前記アンサンブル学習モデルは、第 1 の数の前記学習モデルからそれぞれ第 2 の数の位置候補を取得し、前記第 1 の数に前記第 2 の数を乗じた数の位置候補のバウンディングボックスの中に重なる領域がある場合には、重なる領域を有する位置候補をグループ化し、グループ化した位置候補の信頼度の合計をグループの信頼度とし、信頼度が最大であるグループ又は位置候補のバウンディングボックスを前記物体の位置の情報として出力することを特徴とする請求項 3 に記載の位置検出プログラム。

10

【請求項 5】

前記単純形状データは、シミュレータにより前記単純形状の立体のモデルから作成されたデータであることを特徴とする請求項 1 ~ 4 のいずれか 1 つに記載の位置検出プログラム。

【請求項 6】

前記単純形状データにおける前記単純形状は、円柱、直方体、立方体、球及びバネを含むことを特徴とする請求項 1 ~ 5 のいずれか 1 つに記載の位置検出プログラム。

【請求項 7】

コンピュータが、
それぞれが相似形でない各形状を示す単純形状データを入力して前記形状の位置情報を検出する、各形状に対応した各学習モデルに、撮影手段により物体が撮影された画像データを

20

前記各学習モデルから位置情報および検出の確からしさを示す信頼度を取得し、
前記各学習モデルから得られた各信頼度の合計値が閾値以上である場合、前記各学習モデルから得られた位置情報を用いて、前記物体の位置を検出する、
処理を実行することを特徴とする位置検出方法。

【請求項 8】

それぞれが相似形でない各形状を示す単純形状データを入力して前記形状の位置情報を検出する、各形状に対応した各学習モデルに、撮影手段により物体が撮影された画像データを

30

前記各学習モデルから位置情報および検出の確からしさを示す信頼度を取得し、
前記各学習モデルから得られた各信頼度の合計値が閾値以上である場合、前記各学習モデルから得られた位置情報を用いて、前記物体の位置を検出する検出部と
を有することを特徴とする位置検出装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、位置検出プログラム、位置検出方法及び位置検出装置に関する。

40

【背景技術】

【0002】

近年、画像内の人や物体を検出する技術が広く使われている。例えば、工場や倉庫におけるピッキングロボットでは、把持する物体の位置を画像から検出する技術が利用されている。

【0003】

画像内の物体を検出する技術として、テンプレートマッチングがある。図 1 4 は、テンプレートマッチングを説明するための図である。テンプレートマッチングでは、検出装置は、検出する物体の画像をテンプレート画像 9 1 として予め記憶し、テンプレート画像 9 1 と類似度が高い範囲を入力画像 9 2 から探索する。検出装置は、入力画像上で探索窓 9

50

3をスライドさせながら、探索窓内の画像とテンプレート画像91の類似度を計算し、類似度が最も高い位置に物体が存在すると判断する。

【0004】

また、深層学習(Deep Learning)を行うニューラルネットワーク(NN: Neural Network)を用いて画像内の物体を検出する技術としてSSD(Single Shot Multibox Detector)がある。SSDでは、利用者は、事前に正解ラベル付きの入力画像のデータセットを用意し、用意したデータセットを画像の学習に特化したNNである畳み込みNNに学習させる。学習後のNNは、入力画像中の物体を検出し、検出した物体の位置と検出対象の確からしさを示す信頼度を出力する。

【0005】

なお、ロボットアームの先端に取り付けたカメラからの画像情報をコントローラに入力し、深層学習を用いてロボット自らリアルタイムに動作を学習・生成することで教示を不要とする技術がある。

【先行技術文献】

【非特許文献】

【0006】

【文献】「ロボットによる対象物の多様なつかみ方を実現するAIピッキング機能を開発」、[平成30年10月22日検索]、インターネット<URL:https://www.yaskawa.co.jp/newsrelease/technology/35697>

【発明の概要】

【発明が解決しようとする課題】

【0007】

テンプレートマッチングやSSDには、事前に想定した形状と異なる形状の物体を検出できないという問題がある。

【0008】

本発明は、1つの側面では、事前に想定した形状と異なる形状の物体の検出可能性を高めることを目的とする。

【課題を解決するための手段】

【0009】

1つの態様では、位置検出プログラムは、コンピュータに、それぞれが相似形でない、単純形状の立体のデータを複数取得し、取得した複数の前記データを用いて学習を行う処理を実行させる。そして、前記位置検出プログラムは、前記コンピュータに、撮像手段により撮像された画像を取得し、前記学習に基づいて生成された第1学習モデルを用いて、取得した前記画像から物体の位置を検出する処理を実行させる。

【発明の効果】

【0010】

1つの側面では、本発明は、事前に想定した形状と異なる形状の物体の検出可能性を高めることができる。

【図面の簡単な説明】

【0011】

【図1】図1は、ピッキングロボットの概略を示す図である。

【図2】図2は、ピッキングの流れの一例を示す図である。

【図3】図3は、実施例に係るピッキングロボットシステムの構成を示す図である。

【図4】図4は、アンサンブル学習モデルの作成の流れを示す図である。

【図5】図5は、単純形状モデルによる検出例を示す図である。

【図6】図6は、アンサンブル学習モデルによる検出例を示す図である。

【図7】図7は、位置検出装置の機能構成を示す図である。

【図8】図8は、入力画像から対象物体の位置を推定する処理を説明するための図である。

【図9】図9は、検出位置グループを説明するための図である。

【図10】図10は、位置検出装置による処理のフローを示すフローチャートである。

10

20

30

40

50

【図 1 1】図 1 1 は、アンサンブル学習モデルによる処理のフローを示すフローチャートである。

【図 1 2】図 1 2 は、検出位置グループの効果を説明するための図である。

【図 1 3】図 1 3 は、実施例に係る位置検出プログラムを実行するコンピュータのハードウェア構成を示す図である。

【図 1 4】図 1 4 は、テンプレートマッチングを説明するための図である。

【発明を実施するための形態】

【0012】

以下に、本願の開示する位置検出プログラム、位置検出方法及び位置検出装置の実施例を図面に基づいて詳細に説明する。なお、この実施例は開示の技術を限定するものではない。

10

【実施例】

【0013】

まず、ピッキングロボットについて説明する。図 1 は、ピッキングロボットの概略を示す図である。図 1 に示すように、ピッキングロボット 1 は、アーム 2 の先端に取り付けたグリッパ 3 を用いて対象物体 4 をつかんで移動させるロボットである。ピッキングロボット 1 は、工場で部品や製品をつかんで移動させる場合、あるいは、倉庫で荷物をつかんで移動させる場合などに用いられる。

【0014】

図 2 は、ピッキングの流れの一例を示す図である。図 2 に示すように、事前作業として、利用者は、ピッキングの対象物体 4 とアーム動作を決め (t_1)、ティーチングを実施する (t_2)。ここで、「ティーチング」とは、事前に対象物体 4 の形状などの情報やアーム動作をロボット制御プログラムに設定することである。

20

【0015】

そして、ティーチングが行われたピッキングロボット 1 は、ピッキング動作を行う。すなわち、ピッキングロボット 1 は、カメラで画像を取得し (t_3)、画像から対象物体 4 を検出し、位置を推定する (t_4)。そして、ピッキングロボット 1 は、対象物体 4 をつかむ動作を実施し (t_5)、指定場所へ移動して対象物体 4 を放す (t_6)。

【0016】

このように、ピッキングロボット 1 を動作させるためには、事前にティーチングを行う必要がある。対象物体 4 の形状や周囲の環境が変わると、ティーチングをし直す必要があり、ティーチングの工数を如何に減らすかが重要となっている。

30

【0017】

また、深層学習による画像認識では、大量の学習データを用いた学習を行うことで、認識精度の向上を図っており、大量の学習データが必要となる。利用可能なオープンデータが多数ある場合には、学習データの数を確保することに困難はないが、ピッキング作業の分野においては、オープンデータはほとんどなく、学習データとして利用できるデータは多くない。したがって、ピッキング作業の分野では、学習データをいかに用意するかが重要となっている。

【0018】

また、ピッキングという観点では、対象物体 4 の全体を認識する必要はなく、把持可能な部分を認識できることが重要となる。把持可能な形状は、円柱、直方体、立方体、正多面体、円錐、角錐、球、トーラス(ドーナツ型)、螺旋などの単純な形状であり、ピッキングの対象物体 4 は、これらの単純な形状な部分を備える。このような単純な形状は、一般にプリミティブと呼ばれることがある。

40

【0019】

次に、実施例に係るピッキングロボットシステムの構成について説明する。図 3 は、実施例に係るピッキングロボットシステムの構成を示す図である。図 3 に示すように、実施例に係るピッキングロボットシステム 5 は、ピッキングロボット 1 と、カメラ 6 と、ロボット制御装置 7 と、位置検出装置 10 とを有する。

50

【 0 0 2 0 】

カメラ 6 は、対象物体 4 を撮像し、撮像した画像を位置検出装置 1 0 に出力する。位置検出装置 1 0 は、カメラ 6 により撮像された画像を入力して把持位置情報をロボット制御装置 7 に出力する。ロボット制御装置 7 は、把持位置情報に基づいてピッキングロボット 1 に動作指示を行う。ピッキングロボット 1 は、対象物体 4 に対してピッキングを実施する。

【 0 0 2 1 】

位置検出装置 1 0 は、アンサンブル学習モデル 1 0 a と、変換部 3 2 とを有する。アンサンブル学習モデル 1 0 a は、学習モデル A ~ 学習モデル E で表される 5 つの学習モデル 1 1 を有する。学習モデル 1 1 は、学習用データセットを深層学習した NN である。学習モデル 1 1 は、SSD に基づく NN である。学習モデル 1 1 は、カメラ 6 が撮影した画像を入力して対象物体 4 を検出し、位置情報と信頼度を出力する。アンサンブル学習モデル 1 0 a は、各学習モデル 1 1 が出力した位置情報と信頼度に基づいて対象物体 4 の位置情報を出力する。変換部 3 2 は、対象物体 4 の位置情報を入力して把持位置情報を出力する。

【 0 0 2 2 】

図 4 は、アンサンブル学習モデル 1 0 a の作成の流れを示す図である。利用者は、単純形状モデル A ~ 単純形状モデル E で表される 5 種類の単純形状モデル 1 2 を CG (Computer Graphics) を用いて用意する。ここで、単純形状モデル 1 2 とは、単純な形状の 3 次元モデルである。各単純形状モデル 1 2 は他の単純形状モデル 1 2 と相似形でない。図 4 では、単純形状モデル A は円柱の 3 次元モデルであり、単純形状モデル B は直方体の 3 次元モデルであり、単純形状モデル C は立方体の 3 次元モデルである。また、単純形状モデル D は球の 3 次元モデルであり、単純形状モデル E はバネの 3 次元モデルである。

【 0 0 2 3 】

そして、利用者は、ピッキングロボット 1 のシミュレータを使い、単純形状モデル 1 2 ごとに学習用データセット 1 3 を作成する。そして、位置検出装置 1 0 が、1 つの学習用データセット 1 3 に対して 1 つの学習モデル 1 1 を作成する。学習データセット A から学習モデル A が作成され、学習データセット B から学習モデル B が作成され、学習データセット C から学習モデル C が作成され、学習データセット D から学習モデル D が作成され、学習データセット E から学習モデル E が作成される。

【 0 0 2 4 】

そして、位置検出装置 1 0 は、5 つの学習モデル 1 1 を用いて対象物体 4 の位置を検出するアンサンブル学習モデル 1 0 a を作成する。アンサンブル学習モデル 1 0 a は、例えば、各位置情報について 5 つの学習モデル 1 1 の信頼度の合計をアンサンブル学習モデル 1 0 a の信頼度とし、信頼度が基準値以上であれば、対象物体 4 を検出したと判定する。基準値は、例えば、1 . 0 である。

【 0 0 2 5 】

なお、学習に用いられる単純形状モデル 1 2 の数は、5 より多くても少なくともよいが、単純形状モデル 1 2 の数が多いほど位置検出装置 1 0 の汎用性が増すが、処理に時間がかかるため、5 ~ 1 0 がよい。ピッキング対象の形状の種類が事前に決まっている場合は、位置検出装置 1 0 は、ピッキング対象の形状を抽象化したモデルを用いてもよい。あるいは、ピッキング対象の形状に含まれる部分的な単純形状が単純形状モデルとして選択されることとしてもよい。

【 0 0 2 6 】

図 5 は、単純形状モデル 1 2 による検出例を示す図である。ケース # 1 は、検出対象の物体が「ボルト」である場合を示し、ケース # 2 は、検出対象の物体が「ICチップ」である場合を示す。また、検出判定の基準の信頼度を 0 . 5 以上であるとする。

【 0 0 2 7 】

ケース # 1 では、学習用データセット 1 3 の対象が実写のボルトである場合には、信頼度が 0 . 9 5 であるのでボルトは検出されるが、学習用データセット 1 3 の対象が実写の ICチップである場合には、信頼度が 0 . 0 9 であるのでボルトは検出されない。また、

10

20

30

40

50

学習用データセット 13 の対象が単純形状モデル 12 の 1 つである円柱の 3 次元モデルである場合には、信頼度が 0.83 であるのでボルトが検出される。

【0028】

ケース # 2 では、学習用データセット 13 の対象が実写のボルトである場合には、信頼度が 0.13 であるので IC チップは検出されず、学習用データセット 13 の対象が実写の IC チップである場合には、信頼度が 0.92 であるので IC チップは検出される。また、学習用データセット 13 の対象が単純形状モデル 12 のうちの 1 つである円柱の 3 次元モデルである場合には、信頼度が 0.61 であるので IC チップが検出される。

【0029】

このように、位置検出装置 10 は、単純形状モデル 12 を学習させることによって、対象物体 4 の形状を過学習することを避け、検出可能な形状の類似度の幅を広げることができる。

10

【0030】

図 6 は、アンサンプル学習モデル 10 a による検出例を示す図である。図 6 では、検出対象の物体形状は未知である。図 6 において、単体の学習モデル 11 では、検出判定の基準の信頼度を 0.5 以上であるとし、アンサンプル学習モデル 10 a では、検出判定の基準の信頼度を 1.0 以上であるとする。

【0031】

図 6 に示すように、各単純形状モデル 12 からそれぞれ作成された各学習モデル 11 は、信頼度が 0.5 未満のため、対象物体 4 を検出することはできない。一方、アンサンプル学習モデル 10 a は、各学習モデル 11 の結果を利用することで、対象物体 4 を検出することができる。図 6 では、5 つの学習モデル 11 の信頼度の合計が $0.39 + 0.37 + 0.25 + 0.11 + 0.09 = 1.21$ で 1.0 以上となるので、アンサンプル学習モデル 10 a は、対象物体 4 を検出することができる。

20

【0032】

このように、位置検出装置 10 は、アンサンプル学習モデル 10 a を用いることによって、検出可能な形状の範囲を飛躍的に広げることができる。

【0033】

次に、位置検出装置 10 の機能構成について説明する。図 7 は、位置検出装置 10 の機能構成を示す図である。図 7 に示すように、位置検出装置 10 は、モデル作成部 20 と把持位置検出部 30 とを有する。モデル作成部 20 は、5 つの単純形状モデル 12 の学習用データセット 13 を取得し、5 つの単純形状モデル 12 に基づくアンサンプル学習モデル 10 a を作成する。把持位置検出部 30 は、対象物体 4 を撮像した画像を取得し、アンサンプル学習モデル 10 a を用いて画像内の対象物体 4 の位置を検出し、検出した位置に基づいて把持位置を特定して把持位置情報を出力する。

30

【0034】

モデル作成部 20 は、個別モデル作成部 21 とアンサンプルモデル作成部 22 を有する。把持位置検出部 30 は、物体位置検出部 31 と、変換部 32 とを有する。

【0035】

個別モデル作成部 21 は、5 つの単純形状モデル 12 の学習用データセット 13 を取得し、取得した 5 つの学習用データセット 13 をそれぞれ学習した 5 つの学習モデル 11 を作成する。各学習用データセット 13 には、例えば、検出位置の正解ラベル付きデータが 10000 個含まれる。また、個別モデル作成部 21 は、学習用データセット 13 ごとに学習モデル 11 を 100 エポック訓練させる。

40

【0036】

アンサンプルモデル作成部 22 は、個別モデル作成部 21 により作成された 5 つの学習モデル 11 を用いてアンサンプル学習モデル 10 a を作成する。アンサンプルモデル作成部 22 は、各位置情報について 5 つの学習モデル 11 の信頼度の合計を信頼度とし、信頼度が基準値以上であれば対象物体 4 を検出したと判定するアンサンプル学習モデル 10 a を作成する。

50

【 0 0 3 7 】

物体位置検出部 3 1 は、対象物体 4 を撮像した画像を取得し、アンサンブル学習モデル 1 0 a を用いて画像内の対象物体 4 の位置を検出し、位置情報を出力する。図 8 は、入力画像から対象物体 4 の位置を推定する処理を説明するための図である。図 8 に示すように、物体位置検出部 3 1 は、入力画像から対象物体 4 を検出し、位置情報としてバウンディングボックスの左上の点 (x u , y u)、右下の点 (x v , y v) を出力する。

【 0 0 3 8 】

ここで、バウンディングボックスは、対象物体 4 の位置を定義する長方形の枠である。バウンディングボックスは、例えば、左上と右下の 2 点の x 座標及び y 座標で定義される。座標軸の原点は、例えば、入力画像の左下である。

10

【 0 0 3 9 】

アンサンブル学習モデル 1 0 a は、各学習モデル 1 1 が出力した位置情報の中から信頼度の高い順に 5 個の位置情報を検出位置候補として選択し、5 個の学習モデル 1 1 で合計 2 5 個の検出位置候補に基づいて、対象物体 4 の位置情報を作成する。

【 0 0 4 0 】

具体的には、アンサンブル学習モデル 1 0 a は、検出位置候補のバウンディングボックスに重なりがある場合には、検出位置候補をグルーピングして検出位置グループとする。図 9 は、検出位置グループを説明するための図である。図 9 では、2 5 個の検出位置候補のうち 5 個の検出位置候補だけを示す。

【 0 0 4 1 】

図 9 では、5 個のバウンディングボックスのうち 3 個のバウンディングボックスに重なりがあるため、3 個のバウンディングボックスがグループ化され、太線で示す枠が検出位置グループを示す。検出位置グループの信頼度は、重なりのある検出位置候補の信頼度の合計である。

20

【 0 0 4 2 】

そして、アンサンブル学習モデル 1 0 a は、信頼度がもっとも高い検出位置候補又は検出位置グループを対象物体 4 の位置情報として選択する。アンサンブル学習モデル 1 0 a は、対象物体 4 の位置情報として検出位置グループを選択した場合には、バウンディングボックスの左上の点の座標 (X , Y) を、

$$(X , Y) = (s 1 (x 1 , y 1) + s 2 (x 2 , y 2) + \dots + s n (x n , y n)) / (s 1 + s 2 + \dots + s n)$$

30

により計算する。ここで、n は検出位置グループに含まれる検出位置候補の数であり、s i (i = 1 , 2 , \dots , n) は検出位置候補 i の信頼度であり、(x i , y i) は検出位置候補 i のバウンディングボックスの左上の点の座標である。バウンディングボックスの右下の点の座標も同様に計算される。

【 0 0 4 3 】

そして、アンサンブル学習モデル 1 0 a は、対象物体 4 の位置情報としてバウンディングボックスの左上と右下の座標を出力する。物体位置検出部 3 1 は、アンサンブル学習モデル 1 0 a が出力した位置情報を変換部 3 2 に渡す。

【 0 0 4 4 】

変換部 3 2 は、物体位置検出部 3 1 から位置情報を取得し、取得した位置情報を把持位置情報に変換する。変換部 3 2 は、位置情報と対象物体 4 の把持可能な形状の個所とに基づいて位置情報を把持位置情報に変換する。

40

【 0 0 4 5 】

次に、位置検出装置 1 0 による処理のフローについて説明する。図 1 0 は、位置検出装置 1 0 による処理のフローを示すフローチャートである。図 1 0 に示すように、位置検出装置 1 0 は、単純形状モデル 1 2 の学習用データセット 1 3 を 5 個取得し (ステップ S 1)、5 個の学習用データセット 1 3 を用いて 5 個の学習モデル 1 1 を作成する (ステップ S 2)。そして、位置検出装置 1 0 は、5 個の学習モデル 1 1 からアンサンブル学習モデル 1 0 a を作成する (ステップ S 3)。

50

【 0 0 4 6 】

そして、位置検出装置 1 0 は、対象物体 4 が撮像された画像を取得し（ステップ S 4 ）、アンサンブル学習モデル 1 0 a を用いて、取得した画像から対象物体 4 を検出し、位置情報を作成する（ステップ S 5 ）。そして、位置検出装置 1 0 は、位置情報を把持位置情報に変換し（ステップ S 6 ）、把持位置情報を出力する（ステップ S 7 ）。

【 0 0 4 7 】

このように、位置検出装置 1 0 は、5 つの単純形状モデル 1 2 の学習用データセット 1 3 をそれぞれ学習した 5 つの学習モデル 1 1 を用いてアンサンブル学習モデル 1 0 a を作成し、アンサンブル学習モデル 1 0 a を用いて対象物体 4 を検出する。したがって、位置検出装置 1 0 は、事前に想定した形状と異なる形状の物体を検出することができる。

10

【 0 0 4 8 】

次に、アンサンブル学習モデル 1 0 a による処理のフローについて説明する。図 1 1 は、アンサンブル学習モデル 1 0 a による処理のフローを示すフローチャートである。図 1 1 に示すように、アンサンブル学習モデル 1 0 a は、各学習モデル 1 1 から信頼度が高い順に 5 個のバウンディングボックス及び信頼度を取得する（ステップ S 1 1 ）。

【 0 0 4 9 】

そして、アンサンブル学習モデル 1 0 a は、バウンディングボックスが重なっている領域があるか否かを判定する（ステップ S 1 2 ）。そして、ある場合には、アンサンブル学習モデル 1 0 a は、領域が重なっていた複数の検出位置候補を、1 つの検出位置グループと定義し、領域が重なっていた複数の検出位置候補の信頼度を合算して検出位置グループの信頼度を計算する（ステップ S 1 3 ）。

20

【 0 0 5 0 】

そして、アンサンブル学習モデル 1 0 a は、信頼度がもっとも高い検出位置候補又は検出位置グループを対象物体 4 の位置情報として選択し（ステップ S 1 4 ）、最も信頼度が高いのは検出位置グループか否かを判定する（ステップ S 1 5 ）。そして、最も信頼度が高いのは検出位置グループである場合には、アンサンブル学習モデル 1 0 a は、元の検出位置候補の信頼度で重み付けした座標を検出位置グループの座標として計算する（ステップ S 1 6 ）。

【 0 0 5 1 】

そして、アンサンブル学習モデル 1 0 a は、対象物体 4 の位置情報として、計算した座標を出力する（ステップ S 1 7 ）。

30

【 0 0 5 2 】

このように、アンサンブル学習モデル 1 0 a は、領域が重なっていた検出位置候補を、1 つの検出位置グループと定義し、重なっていた検出位置候補の信頼度を合算して検出位置グループの信頼度とする。したがって、アンサンブル学習モデル 1 0 a は、単体の学習モデル 1 1 のみでは検出ができない対象物体 4 を検出することができる。

【 0 0 5 3 】

図 1 2 は、検出位置グループの効果を説明するための図である。図 1 2 では、単体の学習モデル 1 1 の基準値は 0 . 5 であり、アンサンブル学習モデル 1 0 a の基準値は 1 . 0 であるとする。

40

【 0 0 5 4 】

図 1 2 に示すように、検出対象の物体形状がバネのケースでは、検出位置グループは定義されない。また、単体の学習モデル 1 1 では、バネを学習した学習モデル 1 1 の信頼度が 0 . 5 以上となり、バネを学習した学習モデル 1 1 のみで検出が可能である。

【 0 0 5 5 】

一方、検出対象の物体形状が未知のケースでは、円柱を学習した学習モデル 1 1 が出力した検出位置候補と直方体を学習した学習モデル 1 1 が出力した検出位置候補と立方体を学習した学習モデル 1 1 が出力した検出位置候補がグループ化される。そして、検出位置グループの信頼度が $0 . 3 9 + 0 . 3 7 + 0 . 2 5 = 1 . 0 1$ となり、検出位置グループを用いるアンサンブル学習モデル 1 0 a により物体形状が未知の物体が検出される。

50

【 0 0 5 6 】

上述してきたように、実施例では、個別モデル作成部 2 1 が、複数の単純形状モデル 1 2 からそれぞれ作成された複数の学習用データセット 1 3 をそれぞれ用いて複数の学習モデル 1 1 を作成する。そして、アンサンブルモデル作成部 2 2 が、複数の学習モデル 1 1 に基づいてアンサンブル学習モデル 1 0 a を作成する。そして、物体位置検出部 3 1 が、アンサンブル学習モデル 1 0 a を用いて、対象物体 4 を撮像した画像から対象物体 4 を検出して位置情報を作成する。したがって、位置検出装置 1 0 は、対象物体 4 の形状を過学習することを避け、事前に想定した形状と異なる形状の物体の検出可能性を高めることができる。

【 0 0 5 7 】

また、実施例では、位置検出装置 1 0 は、位置情報を把持位置情報に変換してロボット制御装置 7 に出力するので、対象物体 4 の形状が変更になった場合でもピッキングロボット 1 のティーチングを不要とすることができ、ティーチングの工数を減らすことができる。

【 0 0 5 8 】

また、実施例では、アンサンブル学習モデル 1 0 a は、検出位置候補のバウンディングボックスに重なりがある場合には、検出位置候補をグルーピングして検出位置グループとし、検出位置グループの信頼度を、重なりのある検出位置候補の信頼度の合計とする。したがって、位置検出装置 1 0 は、1 つの学習モデル 1 1 だけでは検出できない未知の物体を検出することができる。

【 0 0 5 9 】

また、実施例では、単純形状モデル 1 2 の学習用データセット 1 3 は、ロボットシミュレータにより作成されるので、学習に十分な量の学習用データセット 1 3 を簡単に作成することができる。

【 0 0 6 0 】

また、実施例では、位置検出装置 1 0 は、単純形状として円柱、直方体、立方体、球及びバネを用いるので、検出できる物体形状の汎用性を高めることができる。

【 0 0 6 1 】

なお、実施例では、位置検出装置 1 0 について説明したが、位置検出装置 1 0 が有する構成をソフトウェアによって実現することで、同様の機能を有する位置検出プログラムを得ることができる。そこで、位置検出プログラムを実行するコンピュータについて説明する。

【 0 0 6 2 】

図 1 3 は、実施例に係る位置検出プログラムを実行するコンピュータのハードウェア構成を示す図である。図 1 3 に示すように、コンピュータ 5 0 は、メインメモリ 5 1 と、CPU (Central Processing Unit) 5 2 と、LAN (Local Area Network) インタフェース 5 3 と、HDD (Hard Disk Drive) 5 4 とを有する。また、コンピュータ 5 0 は、スーパー I/O (Input Output) 5 5 と、DVI (Digital Visual Interface) 5 6 と、ODD (Optical Disk Drive) 5 7 とを有する。

【 0 0 6 3 】

メインメモリ 5 1 は、プログラムやプログラムの実行途中結果等を記憶するメモリである。CPU 5 2 は、メインメモリ 5 1 からプログラムを読み出して実行する中央処理装置である。CPU 5 2 は、メモリコントローラを有するチップセットを含む。

【 0 0 6 4 】

LAN インタフェース 5 3 は、コンピュータ 5 0 を LAN 経由で他のコンピュータに接続するためのインタフェースである。HDD 5 4 は、プログラムやデータを格納するディスク装置であり、スーパー I/O 5 5 は、マウスやキーボード等の入力装置を接続するためのインタフェースである。DVI 5 6 は、液晶表示装置を接続するインタフェースであり、ODD 5 7 は、DVD の読み書きを行う装置である。

【 0 0 6 5 】

LAN インタフェース 5 3 は、PCI エクスプレス (PCIe) により CPU 5 2 に接

10

20

30

40

50

続され、HDD 5 4 及びODD 5 7 は、SATA (Serial Advanced Technology Attachment) によりCPU 5 2 に接続される。スーパーIO 5 5 は、LPC (Low Pin Count) によりCPU 5 2 に接続される。

【0066】

そして、コンピュータ50において実行される位置検出プログラムは、コンピュータ50により読み出し可能な記録媒体の一例であるDVDに記憶され、ODD 5 7 によってDVDから読み出されてコンピュータ50にインストールされる。あるいは、位置検出プログラムは、LANインタフェース53を介して接続された他のコンピュータシステムのデータベース等に記憶され、これらのデータベースから読み出されてコンピュータ50にインストールされる。そして、インストールされた位置検出プログラムは、HDD 5 4 に記憶され、メインメモリ51に読み出されてCPU 5 2 によって実行される。

10

【0067】

また、実施例では、アンサンブル学習モデルが位置情報を作成し、変換部が位置情報を把持位置情報に変換する場合について説明したが、アンサンブル学習モデルは、把持位置情報を直接作成してもよい。この場合、アンサンブル学習モデルは、物体検出範囲を把持位置範囲に置き換えて、信頼度及び把持角度とともに出力する。

【0068】

また、実施例では、ピッキングロボット1で使われる場合について説明したが、位置検出装置10は、医療用ロボット、調理ロボット、土木及び建設用機械、自動走行車、ドローン、家庭用電化製品などの他の装置でも使われてもよい。

20

【符号の説明】

【0069】

- 1 ピッキングロボット
- 2 アーム
- 3 グリッパ
- 4 対象物体
- 5 ピッキングロボットシステム
- 6 カメラ
- 7 ロボット制御装置
- 10 位置検出装置
- 10 a アンサンブル学習モデル
- 11 学習モデル
- 12 単純形状モデル
- 13 学習用データセット
- 32 変換部
- 50 コンピュータ
- 51 メインメモリ
- 52 CPU
- 53 LANインタフェース
- 54 HDD
- 55 スーパーIO
- 56 DVI
- 57 ODD
- 91 テンプレート画像
- 92 入力画像
- 93 探索窓

30

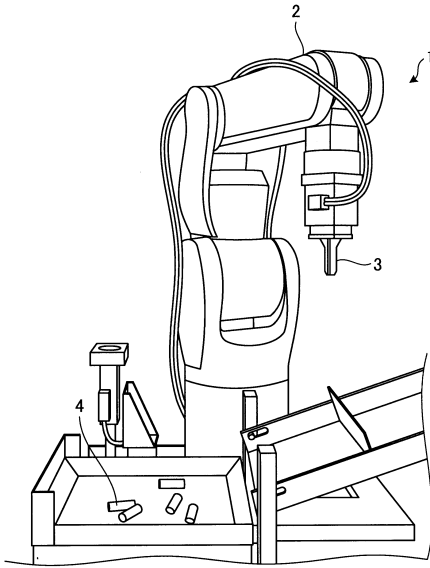
40

50

【図面】

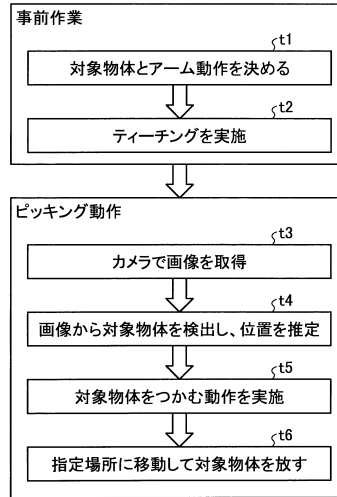
【図 1】

ピッキングロボットの概略を示す図



【図 2】

ピッキングの流れの一例を示す図

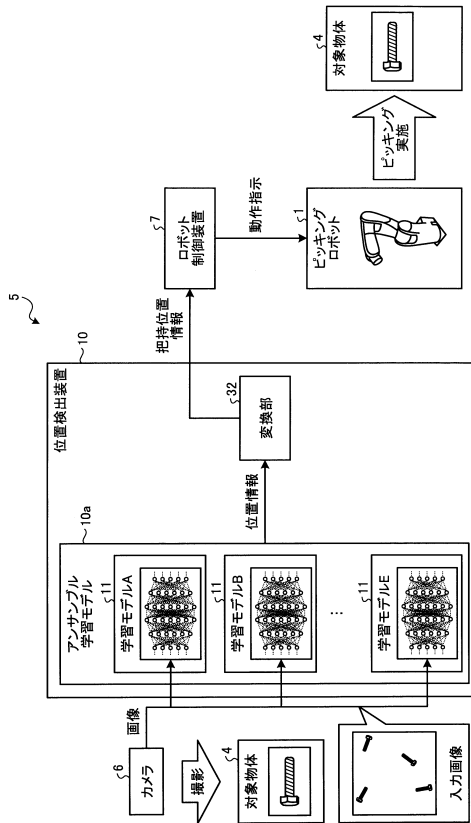


10

20

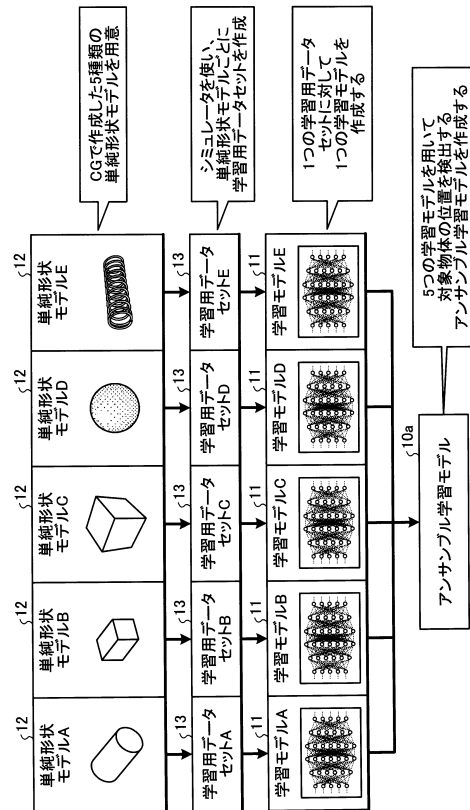
【図 3】

実施例に係るピッキングロボットの構成を示す図



【図 4】

アンサンブル学習モデルの作成の流れを示す図



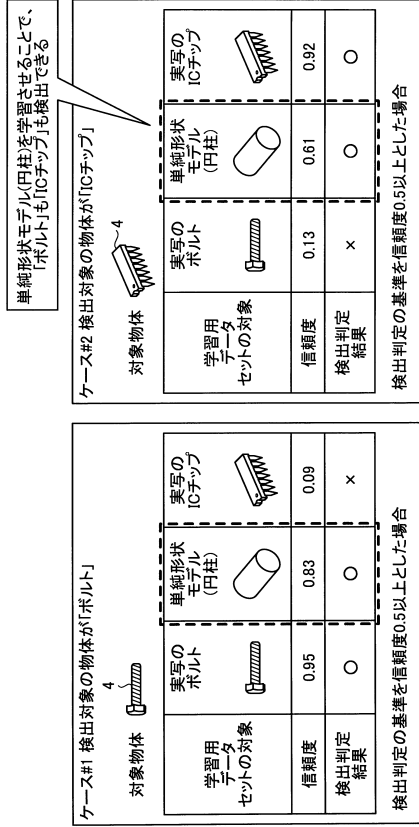
30

40

50

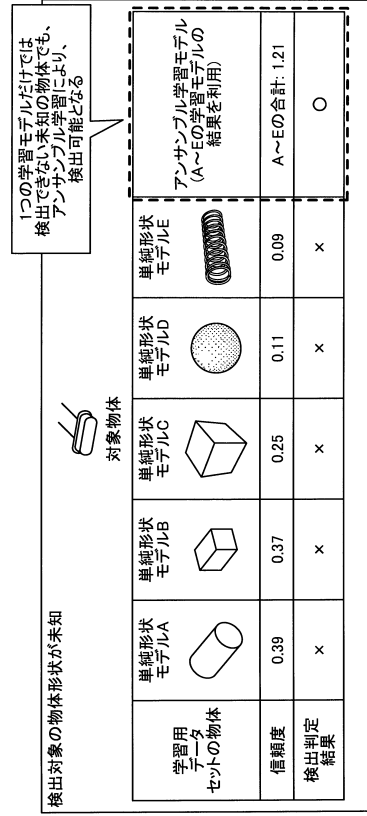
【図 5】

単純形状モデルによる検出例を示す図



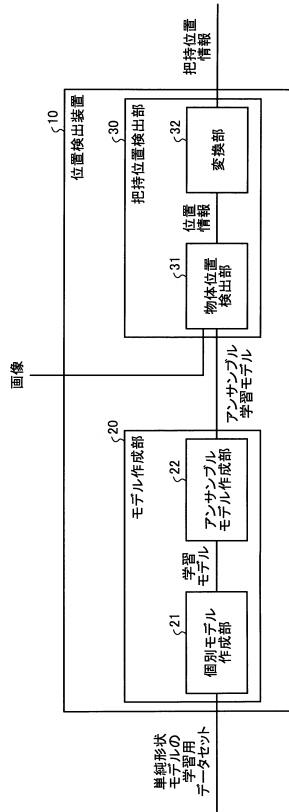
【図 6】

アンサンブル学習モデルによる検出例を示す図



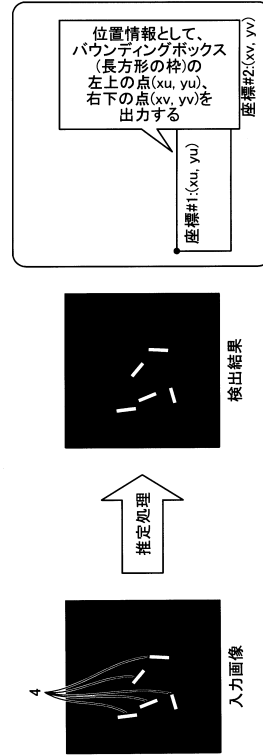
【図 7】

位置検出装置の機能構成を示す図



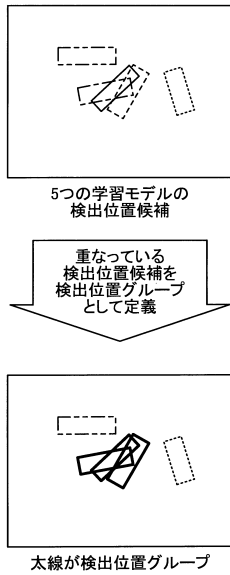
【図 8】

入力画像から対象物体の位置を推定するための図



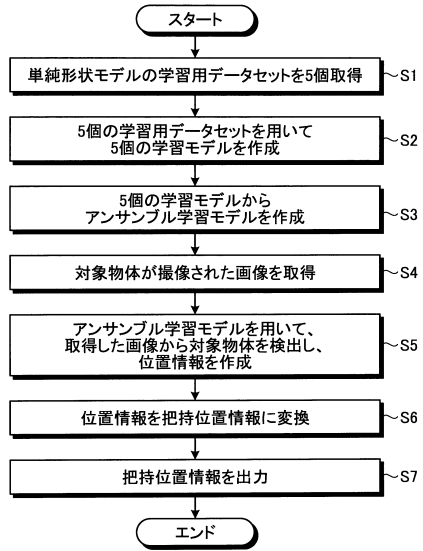
【 図 9 】

検出位置グループを説明するための図



【 図 1 0 】

位置検出装置による処理のフローを示すフローチャート

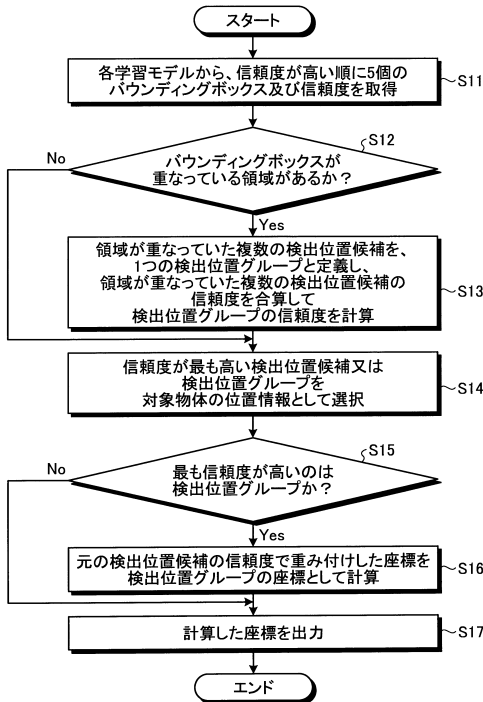


10

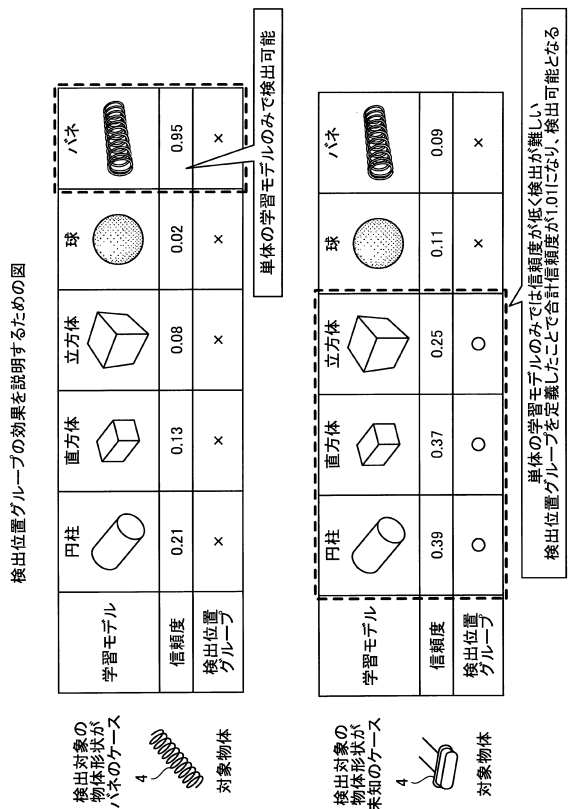
20

【 図 1 1 】

アンサンブル学習モデルによる処理のフローを示すフローチャート



【 図 1 2 】



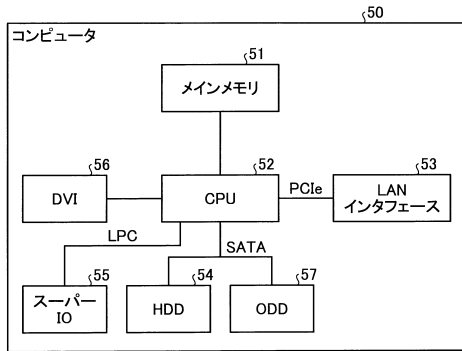
30

40

50

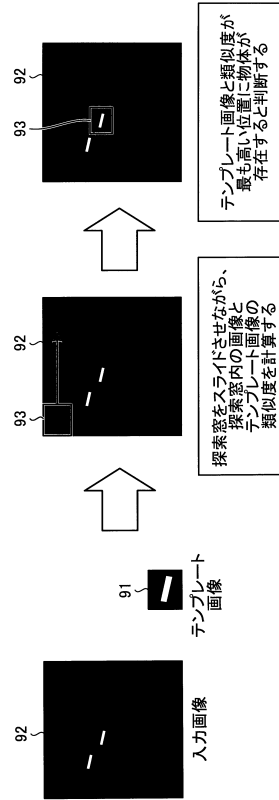
【 図 1 3 】

実施例に係る位置検出プログラムを実行するコンピュータのハードウェア構成を示す図



【 図 1 4 】

テンプレートマッチングを説明するための図



10

20

30

40

50

フロントページの続き

(56)参考文献 特開2018-116599(JP,A)

(58)調査した分野 (Int.Cl., DB名)

G06T 7/00 - 7/90

G06N 3/04

B25J 13/08