

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7345680号  
(P7345680)

(45)発行日 令和5年9月15日(2023.9.15)

(24)登録日 令和5年9月7日(2023.9.7)

(51)国際特許分類

F I

G 0 6 T 7/00 (2017.01)

G 0 6 T 7/00 3 5 0 C

G 0 6 V 10/82 (2022.01)

G 0 6 V 10/82

G 0 6 V 10/762 (2022.01)

G 0 6 V 10/762

請求項の数 6 (全19頁)

|             |                             |          |                   |
|-------------|-----------------------------|----------|-------------------|
| (21)出願番号    | 特願2022-562623(P2022-562623) | (73)特許権者 | 000006013         |
| (86)(22)出願日 | 令和4年8月2日(2022.8.2)          |          | 三菱電機株式会社          |
| (86)国際出願番号  | PCT/JP2022/029598           |          | 東京都千代田区丸の内二丁目7番3号 |
| (87)国際公開番号  | WO2023/277201               | (74)代理人  | 110003166         |
| (87)国際公開日   | 令和5年1月5日(2023.1.5)          |          | 弁理士法人山王内外特許事務所    |
| 審査請求日       | 令和4年10月14日(2022.10.14)      | (72)発明者  | 澤田 友哉             |
|             |                             |          | 東京都千代田区丸の内二丁目7番3号 |
|             |                             |          | 三菱電機株式会社内         |
|             |                             | 審査官      | 藤原 敬利             |

最終頁に続く

(54)【発明の名称】 推論装置、推論方法及び推論プログラム

(57)【特許請求の範囲】

【請求項1】

学習用画像と画像のドメインが異なる場合と、事前学習したタスクと認識タスクの異なる場合とのうち、1つ以上が異なる場合において、検知対象物体が映っている画像である推論対象画像を示す画像信号を取得する画像信号取得部と、

前記画像信号取得部により取得された画像信号を、前記学習用画像の学習が済んでいる第1の学習モデルに与えて、前記第1の学習モデルから、前記推論対象画像に映っている検知対象物体の複数の特徴量のそれぞれがばかされてから前記複数の特徴量が結合された特徴量であって、前記検知対象物体の3次元位置の推論に用いられる特徴量である推論時特徴量を取得する特徴量抽出部と、

前記推論対象画像と画像のドメイン及び認識タスクのそれぞれが同じ対象である変換用画像に映っている検知対象物体の登録済みの特徴量である代表特徴量と前記特徴量抽出部により取得された推論時特徴量とに基づいて、前記推論対象画像に映っている検知対象物体の3次元位置を推定する3次元位置推定部と、

前記3次元位置推定部による3次元位置の推定結果に基づいて、前記推論対象画像に映っている検知対象物体の3次元位置の時間的な変化を解析する変化解析部とを備えた推論装置。

【請求項2】

前記画像信号取得部は、  
前記変換用画像を示す画像信号を取得し、

前記特徴量抽出部は、

前記変換用画像を示す画像信号を前記第 1 の学習モデルに与えて、前記第 1 の学習モデルから、前記変換用画像に映っている検知対象物体の複数の特徴量のそれぞれがぼかされてから当該複数の特徴量が結合された特徴量であって、前記検知対象物体の 3 次元位置の推論に用いられる特徴量である代表特徴量を取得し、

前記特徴量抽出部により取得された代表特徴量を登録する代表特徴量登録部を備えたことを特徴とする請求項 1 記載の推論装置。

【請求項 3】

前記 3 次元位置推定部は、

種類及び存在している領域のそれぞれが互いに異なる複数の検知対象物体の代表特徴量と前記特徴量抽出部により取得された推論時特徴量とを比較して、前記複数の検知対象物体の代表特徴量の中で、前記特徴量抽出部により取得された推論時特徴量に対応している代表特徴量を特定し、前記代表特徴量の特定結果に基づいて、前記推論対象画像に映っている検知対象物体の認識として、前記検知対象物体の種類及び存在している 3 次元領域のそれぞれを推定することを特徴とする請求項 1 記載の推論装置。

10

【請求項 4】

前記変化解析部は、

3 次元位置を学習して 3 次元位置の時間的な変化を回帰する第 2 の学習モデルに対して、前記 3 次元位置推定部による 3 次元位置の推定結果を与えて、前記第 2 の学習モデルから、当該検知対象物体の 3 次元位置の時間的な変化を示す位置データを取得することを特徴とする請求項 1 記載の推論装置。

20

【請求項 5】

画像信号取得部が、学習用画像と画像のドメインが異なる場合と、事前学習したタスクと認識タスクの異なる場合とのうち、1 つ以上が異なる場合において、検知対象物体が映っている画像である推論対象画像を示す画像信号を取得し、

特徴量抽出部が、前記画像信号取得部により取得された画像信号を、前記学習用画像の学習が済んでいる第 1 の学習モデルに与えて、前記第 1 の学習モデルから、前記推論対象画像に映っている検知対象物体の複数の特徴量のそれぞれがぼかされてから前記複数の特徴量が結合された特徴量であって、前記検知対象物体の 3 次元位置の推論に用いられる特徴量である推論時特徴量を取得し、

30

3 次元位置推定部が、前記推論対象画像と画像のドメイン及び認識タスクのそれぞれが同じ対象である変換用画像に映っている検知対象物体の登録済みの特徴量である代表特徴量と前記特徴量抽出部により取得された推論時特徴量とに基づいて、前記推論対象画像に映っている検知対象物体の 3 次元位置を推定し、

変化解析部が、前記 3 次元位置推定部による 3 次元位置の推定結果に基づいて、前記推論対象画像に映っている検知対象物体の 3 次元位置の時間的な変化を解析する

推論方法。

【請求項 6】

画像信号取得部が、学習用画像と画像のドメインが異なる場合と、事前学習したタスクと認識タスクの異なる場合とのうち、1 つ以上が異なる場合において、検知対象物体が映っている画像である推論対象画像を示す画像信号を取得する画像信号取得手順と、

40

特徴量抽出部が、前記画像信号取得部により取得された画像信号を、前記学習用画像の学習が済んでいる第 1 の学習モデルに与えて、前記第 1 の学習モデルから、前記推論対象画像に映っている検知対象物体の複数の特徴量のそれぞれがぼかされてから前記複数の特徴量が結合された特徴量であって、前記検知対象物体の 3 次元位置の推論に用いられる特徴量である推論時特徴量を取得する特徴量取得手順と、

3 次元位置推定部が、前記推論対象画像と画像のドメイン及び認識タスクのそれぞれが同じ対象である変換用画像に映っている検知対象物体の登録済みの特徴量である代表特徴量と前記特徴量抽出部により取得された推論時特徴量とに基づいて、前記推論対象画像に映っている検知対象物体の 3 次元位置を推定する 3 次元位置推定手順と、

50

変化解析部が、前記３次元位置推定部による３次元位置の推定結果に基づいて、前記推論対象画像に映っている検知対象物体の３次元位置の時間的な変化を解析する変化解析手順とをコンピュータに実行させるための推論プログラム。

【発明の詳細な説明】

【技術分野】

【０００１】

本開示は、推論装置、推論方法及び推論プログラムに関するものである。

【背景技術】

【０００２】

例えば、特許文献１には、自律移動装置の異常発生を検知する異常検知方法が開示されている。当該異常検知方法では、センサ群及び制御部のそれぞれを用いて、自律移動装置の異常発生を検知する。

10

センサ群は、自律移動装置における現在の状態を検知する。制御部は、センサ群から、検知開始時点から現時点までのセンサデータである時系列データを取得する。制御部は、時系列データを第１所定間隔毎に分割することで、複数の分割データを生成し、複数の分割データと時系列データとから複数のグラフを生成する。また、制御部は、複数のグラフを学習モデルに与えて、学習モデルから、異常発生の検知結果を取得する。異常発生の検知結果の中には、自律移動装置の異常な位置が含まれている。

【先行技術文献】

【特許文献】

20

【０００３】

【文献】特開２０２１－１１０９７３号公報

【発明の概要】

【発明が解決しようとする課題】

【０００４】

自律移動装置が飛翔体であれば、自律移動装置の航行は、自然環境の状態変化に影響される。自然環境の状態としては、例えば、風の強さ、風の向き、雨の有無、降雨量、雪の有無、又は、積雪量がある。

特許文献１に開示されている異常検知方法では、自然環境の状態がどのように変化しても、自律移動装置の異常発生を検知できるようにするには、自律移動装置の航行時に想定される、全ての自然環境の状態を考慮して、学習モデルを学習させる必要がある。しかしながら、全ての自然環境の状態を考慮して、学習モデルを学習させるには、極めて膨大な学習データを用意する必要があるが、現実的には、十分な学習データを用意することができないことがある。このため、当該異常検知方法では、自然環境の状態によっては、自律移動装置の異常発生を検知できないことがあるという課題があった。

30

【０００５】

本開示は、上記のような課題を解決するためになされたもので、自然環境の状態を考慮して、学習モデルを学習させることなく、検知対象物体の３次元位置の時間的な変化を解析することができる推論装置を得ることを目的とする。

【課題を解決するための手段】

40

【０００６】

本開示に係る推論装置は、学習用画像と画像のドメインが異なる場合と、事前学習したタスクと認識タスクの異なる場合とのうち、１つ以上が異なる場合において、検知対象物体が映っている画像である推論対象画像を示す画像信号を取得する画像信号取得部と、画像信号取得部により取得された画像信号を、学習用画像の学習が済んでいる第１の学習モデルに与えて、第１の学習モデルから、推論対象画像に映っている検知対象物体の複数の特徴量のそれぞれがばかされてから複数の特徴量が結合された特徴量であって、検知対象物体の３次元位置の推論に用いられる特徴量である推論時特徴量を取得する特徴量抽出部とを備えている。また、推論装置は、推論対象画像と画像のドメイン及び認識タスクのそれぞれが同じ対象である変換用画像に映っている検知対象物体の登録済みの特徴量である

50

代表特徴量と特徴量抽出部により取得された推論時特徴量とに基づいて、推論対象画像に映っている検知対象物体の３次元位置を推定する３次元位置推定部と、３次元位置推定部による３次元位置の推定結果に基づいて、推論対象画像に映っている検知対象物体の３次元位置の時間的な変化を解析する変化解析部とを備えている。

【発明の効果】

【０００７】

本開示によれば、自然環境の状態を考慮して、学習モデルを学習させることなく、検知対象物体の３次元位置の時間的な変化を解析することができる。

【図面の簡単な説明】

【０００８】

10

【図１】実施の形態１に係る推論装置４を示す構成図である。

【図２】実施の形態１に係る推論装置４のハードウェアを示すハードウェア構成図である。

【図３】推論装置４が、ソフトウェア又はファームウェア等によって実現される場合のコンピュータのハードウェア構成図である。

【図４】学習装置７を示す構成図である。

【図５】学習装置７のハードウェアを示すハードウェア構成図である。

【図６】学習装置７が、ソフトウェア又はファームウェア等によって実現される場合のコンピュータのハードウェア構成図である。

【図７】ドメイン変換時の推論装置４の処理手順を示すフローチャートである。

【図８】位置推定時の推論装置４の処理手順である推論方法を示すフローチャートである。

20

【図９】検知対象物体の３次元位置の時間的な変化を示す説明図である。

【発明を実施するための形態】

【０００９】

以下、本開示をより詳細に説明するために、本開示を実施するための形態について、添付の図面に従って説明する。

【００１０】

実施の形態１．

図１は、実施の形態１に係る推論装置４を示す構成図である。

図２は、実施の形態１に係る推論装置４のハードウェアを示すハードウェア構成図である。

30

図１において、モデル記憶部１は、例えば、ハードディスク、又は、ＲＡＭ（Ｒａｎｄｏｍ Ａｃｃｅｓｓ Ｍｅｍｏｒｙ）によって実現される。

モデル記憶部１は、第１の学習モデル１ａを記憶している。

【００１１】

第１の学習モデル１ａは、例えば、多層ニューラルネットワーク（Ｄｅｅｐ Ｎｅｕｒａｌ Ｎｅｔｗｏｒｋｓ：ＤＮＮｓ）によって実現される。ＤＮＮｓの中には、畳み込みニューラルネットワークであるＣＮＮｓ（Ｃｏｎｖｏｌｕｔｉｏｎａｌ Ｎｅｕｒａｌ Ｎｅｔｗｏｒｋｓ）が含まれる。

学習モデル１ａは、学習時に、学習データとして、学習用画像を示す画像信号が与えられて、学習用画像の学習が済んでいる。学習用画像は、例えば、画像認識タスクに用いられる画像である。

40

学習用画像のドメインである画像の種類は、どのようなものでもよく、学習用画像は、例えば、ＲＧＢ画像、ＴＩＲ画像、又は、ＣＧシミュレータによって生成された画像のいずれかである。

図１に示す推論装置４では、説明の便宜上、学習用画像がＲＧＢ画像であるものとして説明する。学習モデル１ａは、大量のＲＧＢ画像が与えられて、ＲＧＢ画像を学習しているものである。

学習モデル１ａは、後述する特徴量抽出部１２から、学習用画像と画像のドメイン及び認識タスクのそれぞれが異なる場合において、検知対象物体が映っている画像である変換用画像を示す画像信号が与えられたとき、変換用画像に映っている検知対象物体の複数の

50

特徴量のそれぞれがばかされてから複数の特徴量が結合された特徴量であって、検知対象物体の3次元位置の推論に用いられる特徴量を示す特徴ベクトルを特徴量抽出部12に出力する。

変換用画像は、学習用画像と画像のドメイン及び認識タスクのうち、1つ以上が異なる画像であればよい。図1に示す推論装置4では、説明の便宜上、変換用画像がTIR画像であるものとして説明する。

#### 【0012】

第1の学習モデル1aを実現するCNNsは、非常にディープなCNNsである。非常にディープなCNNsとしては、例えば、101層あるResNetがある。このため、例えば3次元位置の推定時において、第1の学習モデル1aの入力層に画像信号が与えられたときに、第1の学習モデル1aの出力層から出力される特徴ベクトルが示す特徴量は、高次元特徴量である。高次元特徴量は、複数の次元の特徴量を含むものであり、高次元特徴量を示す特徴ベクトルとしては、例えば、Tensorが用いられる。

10

学習モデル1aに含まれている複数段の隠れ層のうち、浅い層から出力される低次元特徴量は、例えば、色、輝度、又は、方向を示すものである。このため、低次元特徴量は、入力層に与えられる画像信号が示す画像のドメインに依存する。つまり、RGB画像の画像信号が学習モデル1aの入力層に与えられたときに、学習モデル1aの浅い層から出力される特徴ベクトルが示す特徴量と、TIR画像の画像信号が学習モデル1aの入力層に与えられたときに、学習モデル1aの浅い層から出力される特徴ベクトルが示す特徴量とは、大きく異なることがある。

20

一方、学習モデル1aの十分に深い中間層から出力される高次元特徴量は、検知対象物体の良否等を表現する概念的な特徴を示すものである。このため、高次元特徴量は、入力層に与えられる画像信号が示す画像のドメインへの依存度が極めて低い概念的な情報となる。また、より深い層の高次元特徴を採用することで、タスクへの依存度の低い汎用的な情報を取得することが可能となる。概念的な物体の特徴は、例えば、“Objectness”、又は、“Informativeness”が知られている。

つまり、RGB画像の画像信号が学習モデル1aの入力層に与えられたときに、学習モデル1aの出力層から出力される特徴ベクトルが示す高次元特徴量と、TIR画像の画像信号が学習モデル1aの入力層に与えられたときに、学習モデル1aの出力層から出力される特徴ベクトルが示す高次元特徴量との差異は小さい。

30

したがって、学習モデル1aがCNNsによって実現されて、推論装置4が、CNNsの十分に深い中間層から出力される特徴ベクトルが示す高次元特徴量を用いる場合、入力層に与えられる画像信号が示す画像のドメインへの依存性と認識タスクへの依存性が軽減される。

#### 【0013】

モデル記憶部2は、例えば、ハードディスク、又は、RAMによって実現される。

モデル記憶部2は、第2の学習モデル2aを記憶している。

第2の学習モデル2aは、例えば、回帰型ニューラルネットワークであるRNNs(Recurrent Neural Networks)によって実現される。

第2の学習モデル2aは、自学習の学習モデルであって、3次元位置を学習して3次元位置の時間的な変化を回帰する学習モデルである。

40

第2の学習モデル2aは、後述する変化解析部16から、後述する3次元位置推定部15による3次元位置の推定結果が与えられたとき、検知対象物体の3次元位置の時間的な変化を示す位置データとして、検知対象物体の将来の時刻の3次元位置を示す信号を変化解析部16に出力する。

#### 【0014】

図1に示す推論装置4では、第1の学習モデル1a及び第2の学習モデル2aのそれぞれが推論装置4の外部に設けられている。しかし、これは一例に過ぎず、例えば、第1の学習モデル1aが特徴量抽出部12に内蔵され、第2の学習モデル2aが変化解析部16に内蔵されていてもよい。

50

## 【 0 0 1 5 】

検知対象物体は、例えば、ドローン、空飛ぶクルマ、ヘリコプター、自動車、又は、船舶である。図 1 に示す推論装置 4 では、説明の便宜上、検知対象物体が、ドローン、空飛ぶクルマ、又は、ヘリコプターのいずれかに認識される例を説明する。

また、図 1 に示す推論装置 4 では、3 次元位置推定部 1 5 により認識される検知対象物体の 3 次元位置が、例えば、高次元特徴における物体の存在範囲 ( O b j e c t n e s s ) を含む特徴ベクトルが登録された代表特徴量記憶部 1 4 より、検知対象物体の存在範囲 ( O b j e c t n e s s ) を含む代表特徴量を取得し、代表特徴量と特徴抽出部 1 2 より抽出された推論時特徴量とを比較して、複数の検知対象物体の代表特徴量の中で、特徴量抽出部 1 2 により抽出された推論時特徴量に対応している代表特徴量を特定する。最も類似した代表特徴量を持つクラスが検出対象物体の所属するクラスとなる。さらに、推論時特徴量には、推論時に入力された画像に含まれる検出対象物体の存在範囲 ( O b j e c t n e s s ) を含むため、推論時特徴量の高次元特徴の記述方法を T e n s o r 等から二次元空間へ変換することで空間的な物体の存在位置を表すことが可能となる。

10

## 【 0 0 1 6 】

検知対象物体が、ドローン、空飛ぶクルマ、又は、ヘリコプターのいずれかに認識される場合には、第 1 の学習モデル 1 a に与えられる学習データは、学習用画像を示す画像信号を含む学習データである。学習用画像は、説明の便宜上、R G B 画像であるものとする。

## 【 0 0 1 7 】

第 2 の学習モデル 2 a に与えられる学習データは、3 次元位置推定部 1 5 による 3 次元位置の推定結果である。

20

第 2 の学習モデル 2 a は、3 次元位置の推定結果が与えられると、3 次元位置を学習して 3 次元位置の時間的な変化を回帰する。

## 【 0 0 1 8 】

カメラ 3 は、例えば、赤外線カメラによって実現される。

カメラ 3 は、検知対象物体を撮影する。

推論装置 4 が学習時とは異なるドメインの画像を登録する際 ( 以下「ドメイン変換時」という ) には、カメラ 3 は、検知対象物体が映っている変換用画像を示す画像信号として、例えば、検知対象物体が映っている T I R 画像を示す画像信号を推論装置 4 に出力する。

推論装置 4 が検知対象物体の 3 次元位置を推定する位置推定時には、カメラ 3 は、検知対象物体が映っている推論対象画像を示す画像信号として、例えば、検知対象物体が映っている T I R 画像を示す画像信号を推論装置 4 に出力する。

30

図 1 では、カメラ 3 が、変換用画像を示す画像信号を推論装置 4 に出力している。しかし、これは一例に過ぎず、検知対象物体が映っている変換用画像を示す画像信号を記憶している図示せぬ記憶部から、当該画像信号が推論装置 4 に出力されるものであってもよい。

## 【 0 0 1 9 】

画像信号取得部 1 1 は、例えば、図 2 に示す画像信号取得回路 2 1 によって実現される。

ドメイン変換時には、画像信号取得部 1 1 は、カメラ 3 から、検知対象物体が映っている変換用画像を示す画像信号を取得する。

そして、画像信号取得部 1 1 は、変換用画像を示す画像信号を特徴量抽出部 1 2 に出力する。

40

位置推定時には、画像信号取得部 1 1 は、カメラ 3 から、検知対象物体が映っている推論対象画像を示す画像信号を取得する。

そして、画像信号取得部 1 1 は、推論対象画像を示す画像信号を特徴量抽出部 1 2 に出力する。

変換用画像と推論対象画像とは、画像の種類が同じであり、例えば、共に T I R 画像である。

## 【 0 0 2 0 】

特徴量抽出部 1 2 は、例えば、図 2 に示す特徴量抽出回路 2 2 によって実現される。

ドメイン変換時には、特徴量抽出部 1 2 は、画像信号取得部 1 1 により取得された画像

50

信号を第1の学習モデル1aに与えて、第1の学習モデル1aから、変換用画像に映っている検知対象物体の複数の特徴量のそれぞれがぼかされてから複数の特徴量が結合された特徴量である代表特徴量を示す特徴ベクトルを取得する。代表特徴量は、検知対象物体の3次元位置の推論に用いられる特徴量である。

特徴量抽出部12は、特徴ベクトルを代表特徴量登録部13に出力する。

位置推定時には、特徴量抽出部12は、画像信号取得部11により取得された画像信号を第1の学習モデル1aに与えて、第1の学習モデル1aから、推論対象画像に映っている検知対象物体の特徴量のそれぞれがぼかされてから複数の特徴量が結合された特徴量である推論時特徴量を示す特徴ベクトルを取得する。推論時特徴量は、検知対象物体の3次元位置の推論に用いられる特徴量である。

10

特徴量抽出部12は、特徴ベクトルを3次元位置推定部15に出力する。

複数の特徴量のそれぞれをぼかす処理としては、“Pooling Operation”が知られている。

【0021】

代表特徴量登録部13は、例えば、図2に示す代表特徴量登録回路23によって実現される。

代表特徴量登録部13は、特徴量抽出部12により取得された代表特徴量を登録する。

即ち、代表特徴量登録部13は、いずれかの領域に存在しているドローンの特徴量を示す特徴ベクトルを取得し、特徴ベクトルを代表特徴量記憶部14に記憶させる。

また、代表特徴量登録部13は、いずれかの領域に存在している空飛ぶクルマの特徴量を示す特徴ベクトルを取得し、特徴ベクトルを代表特徴量記憶部14に記憶させる。

20

また、代表特徴量登録部13は、いずれかの領域に存在しているヘリコプターの特徴量を示す特徴ベクトルを取得し、特徴ベクトルを代表特徴量記憶部14に記憶させる。

【0022】

代表特徴量記憶部14は、例えば、図2に示す代表特徴量記憶回路24によって実現される。

代表特徴量記憶部14は、代表特徴量を示す特徴ベクトルを記憶する。

【0023】

3次元位置推定部15は、例えば、図2に示す3次元位置推定回路25によって実現される。

30

3次元位置推定部15は、特徴量抽出部12から、推論対象画像に映っている検知対象物体の特徴量を示す特徴ベクトルを取得し、代表特徴量記憶部14から、代表特徴量を示す特徴ベクトルを取得する。

3次元位置推定部15は、代表特徴量と推論対象画像に映っている検知対象物体の特徴量とに基づいて、推論対象画像に映っている検知対象物体の3次元位置を推定する。

具体的には、3次元位置推定部15は、代表特徴量と推論時特徴量とに基づいて、推論対象画像に映っている検知対象物体の種類及び存在している領域のそれぞれを認識する。

具体的には、3次元位置推定部15は、高次元特徴における物体の存在範囲(Objectness)及び物体の種類を双方を含む特徴ベクトルが登録された代表特徴量記憶部14より、検知対象物体の存在範囲(Objectness)及び物体の種類を双方を含む代表特徴量を取得し、代表特徴量と特徴抽出部12より抽出された推論時特徴量とを比較して、複数の検知対象物体の代表特徴量の中で、特徴量抽出部12により抽出された推論時特徴量に対応している代表特徴量を特定する。最も類似した代表特徴量を持つクラスが検出対象物体の所属するクラスとなる。さらに推論時特徴量には、推論時に入力された画像に含まれる検出対象物体の存在範囲(Objectness)を含むため、推論時特徴量の高次元特徴の記述方法をTensor等から二次元空間へ変換することで空間的な物体の存在位置を表すことが可能となる。

40

3次元位置推定部15は、検知対象物体の認識結果を示す表示データを生成し、表示データを表示装置5に出力する。ここで、二次元空間上に矩形として表現した場合は、物体検出(Object Detection)タスクとなり、二次元空間上に領域として表

50

現した場合は、セグメンテーション ( S e m a t i c   S e g m e n t a t i o n ) タスクとなる。

【 0 0 2 4 】

変化解析部 1 6 は、例えば、図 2 に示す変化解析回路 2 6 によって実現される。

変化解析部 1 6 は、3 次元位置推定部 1 5 による 3 次元位置の推定結果に基づいて、推論対象画像に映っている検知対象物体の 3 次元位置の時間的な変化を解析する。

具体的には、変化解析部 1 6 は、3 次元位置推定部 1 5 による次元位置の推定結果を第 2 の学習モデル 2 a に与えて、第 2 の学習モデル 2 a から、検知対象物体の 3 次元位置の時間的な変化を示す位置データを取得する。

また、変化解析部 1 6 は、位置データに基づいて、検知対象物体が落下するか否かを判定する処理のほか、検知対象物体を多クラスに分類する処理として、検知対象物体が、例えば、ホバリングしているのか、上昇しているのか、下降しているのか、前進しているのか、又は、後退しているのかを判定する。

10

変化解析部 1 6 は、検知対象物体の 3 次元位置の時間的な変化の解析結果を示す表示データを生成し、表示データを表示装置 5 に出力する。

また、変化解析部 1 6 は、上記の判定処理の判定結果を示す表示データを生成し、表示データを表示装置 5 に出力する。

【 0 0 2 5 】

表示装置 5 は、3 次元位置推定部 1 5 から出力された表示データに従って、3 次元位置の推定結果を図示せぬディスプレイに表示させる。

20

また、表示装置 5 は、変化解析部 1 6 から出力された表示データに従って、検知対象物体の 3 次元位置の時間的な変化の解析結果と、判定処理の判定結果とを図示せぬディスプレイに表示させる。

【 0 0 2 6 】

図 1 では、推論装置 4 の構成要素である画像信号取得部 1 1、特徴量抽出部 1 2、代表特徴量登録部 1 3、代表特徴量記憶部 1 4、3 次元位置推定部 1 5 及び変化解析部 1 6 のそれぞれが、図 2 に示すような専用のハードウェアによって実現されるものを想定している。即ち、推論装置 4 が、画像信号取得回路 2 1、特徴量抽出回路 2 2、代表特徴量登録回路 2 3、代表特徴量記憶回路 2 4、3 次元位置推定回路 2 5 及び変化解析回路 2 6 によって実現されるものを想定している。

30

代表特徴量記憶回路 2 4 は、例えば、RAM ( R a n d o m   A c c e s s   M e m o r y )、ROM ( R e a d   O n l y   M e m o r y )、フラッシュメモリ、EPROM ( E r a s a b l e   P r o g r a m m a b l e   R e a d   O n l y   M e m o r y )、EEPROM ( E l e c t r i c a l l y   E r a s a b l e   P r o g r a m m a b l e   R e a d   O n l y   M e m o r y ) 等の不揮発性又は揮発性の半導体メモリ、磁気ディスク、フレキシブルディスク、光ディスク、コンパクトディスク、ミニディスク、あるいは、DVD ( D i g i t a l   V e r s a t i l e   D i s c ) が該当する。

画像信号取得回路 2 1、特徴量抽出回路 2 2、代表特徴量登録回路 2 3、3 次元位置推定回路 2 5 及び変化解析回路 2 6 のそれぞれは、例えば、単回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ASIC ( A p p l i c a t i o n   S p e c i f i c   I n t e g r a t e d   C i r c u i t )、FPGA ( F i e l d - P r o g r a m m a b l e   G a t e   A r r a y )、又は、これらを組み合わせたものが該当する。

40

【 0 0 2 7 】

推論装置 4 の構成要素は、専用のハードウェアによって実現されるものに限るものではなく、推論装置 4 が、ソフトウェア、ファームウェア、又は、ソフトウェアとファームウェアとの組み合わせによって実現されるものであってもよい。

ソフトウェア又はファームウェアは、プログラムとして、コンピュータのメモリに格納される。コンピュータは、プログラムを実行するハードウェアを意味し、例えば、CPU ( C e n t r a l   P r o c e s s i n g   U n i t )、GPU ( G r a p h i c a l   P

50

rocessing Unit)、中央処理装置、処理装置、演算装置、マイクロプロセッサ、マイクロコンピュータ、プロセッサ、あるいは、DSP(Digital Signal Processor)が該当する。

【0028】

図3は、推論装置4が、ソフトウェア又はファームウェア等によって実現される場合のコンピュータのハードウェア構成図である。

推論装置4が、ソフトウェア又はファームウェア等によって実現される場合、代表特徴量記憶部14がコンピュータのメモリ31上に構成される。画像信号取得部11、特徴量抽出部12、代表特徴量登録部13、3次元位置推定部15及び変化解析部16におけるそれぞれの処理手順をコンピュータに実行させるためのプログラムがメモリ31に格納される。そして、コンピュータのプロセッサ32がメモリ31に格納されているプログラムを実行する。

10

【0029】

また、図2では、推論装置4の構成要素のそれぞれが専用のハードウェアによって実現される例を示し、図3では、推論装置4がソフトウェア又はファームウェア等によって実現される例を示している。しかし、これは一例に過ぎず、推論装置4における一部の構成要素が専用のハードウェアによって実現され、残りの構成要素がソフトウェア又はファームウェア等によって実現されるものであってもよい。

【0030】

図4は、学習装置7を示す構成図である。

20

図5は、学習装置7のハードウェアを示すハードウェア構成図である。

学習データ記憶部6は、例えば、ハードディスク、又は、RAMによって実現される。

学習データ記憶部6は、学習データとして、学習用画像を示す画像信号を記憶している。

【0031】

学習装置7は、学習データ取得部41及び学習処理部42を備えている。

学習データ取得部41は、例えば、図5に示す学習データ取得回路51によって実現される。

学習データ取得部41は、学習データ記憶部6から、学習データを取得する。

学習データ取得部41は、学習データを学習処理部42に出力する。

【0032】

30

学習処理部42は、例えば、図5に示す学習処理回路52によって実現される。

学習処理部42は、学習データ取得部41から、大量の学習データを取得する。

学習処理部42は、それぞれの学習データを学習モデル1aに与えて、それぞれの学習データに含まれている画像信号が示す学習用画像を学習モデル1aに学習させる。

学習済みの学習モデル1aは、ドメイン変換時、又は、位置推定時において、画像信号が与えられたとき、当該画像信号に対応する特徴ベクトルを出力する。

【0033】

図4では、学習装置7の構成要素である学習データ取得部41及び学習処理部42のそれぞれが、図5に示すような専用のハードウェアによって実現されるものを想定している。即ち、学習装置7が、学習データ取得回路51及び学習処理回路52によって実現されるものを想定している。

40

学習データ取得部41及び学習処理部42のそれぞれは、例えば、単回路、複合回路、プログラム化したプロセッサ、並列プログラム化したプロセッサ、ASIC、FPGA、又は、これらを組み合わせたものが該当する。

【0034】

学習装置7の構成要素は、専用のハードウェアによって実現されるものに限るものではなく、学習装置7が、ソフトウェア、ファームウェア、又は、ソフトウェアとファームウェアとの組み合わせによって実現されるものであってもよい。

図6は、学習装置7が、ソフトウェア又はファームウェア等によって実現される場合のコンピュータのハードウェア構成図である。

50

学習装置 7 が、ソフトウェア又はファームウェア等によって実現される場合、学習データ取得部 4 1 及び学習処理部 4 2 におけるそれぞれの処理手順をコンピュータに実行させるためのプログラムがメモリ 6 1 に格納される。そして、コンピュータのプロセッサ 6 2 がメモリ 6 1 に格納されているプログラムを実行する。

【 0 0 3 5 】

また、図 5 では、学習装置 7 の構成要素のそれぞれが専用のハードウェアによって実現される例を示し、図 6 では、学習装置 7 がソフトウェア又はファームウェア等によって実現される例を示している。しかし、これは一例に過ぎず、学習装置 7 における一部の構成要素が専用のハードウェアによって実現され、残りの構成要素がソフトウェア又はファームウェア等によって実現されるものであってもよい。

10

【 0 0 3 6 】

最初に、図 4 に示す学習装置 7 の動作について説明する。

学習データ記憶部 6 には、大量の学習データが記憶されており、それぞれの学習データには、学習用画像を示す画像信号が含まれている。

【 0 0 3 7 】

学習装置 7 の学習データ取得部 4 1 は、学習データ記憶部 6 から、大量の学習データを取得する。

学習データ取得部 4 1 は、それぞれの学習データを学習処理部 4 2 に出力する。

【 0 0 3 8 】

学習処理部 4 2 は、学習データ取得部 4 1 から、それぞれの学習データを取得する。

20

学習処理部 4 2 は、それぞれの学習データを第 1 の学習モデル 1 a に与えて、それぞれの学習データに含まれている画像信号が示す学習用画像を学習モデル 1 a に学習させる。

学習済みの学習モデル 1 a は、例えば、RGB 画像を示す画像信号が入力層に与えられると、出力層から、当該画像信号に対応する特徴ベクトルとして、RGB 画像に映っている検知対象物体の高次元特徴量を示す特徴ベクトルを出力する。

学習モデル 1 a に学習に用いられている学習用画像が、例えば、RGB 画像であって、TIR 画像が、学習用画像として用いられていない場合、RGB 画像に映っている検知対象物体と TIR 画像に映っている検知対象物体とが、共に正常な同一物体であったとしても、RGB 画像を示す画像信号が入力層に与えられたときに、出力層から出力される特徴ベクトルと、TIR 画像を示す画像信号が入力層に与えられたときに、出力層から出力される特徴ベクトルとが相違することがある。

30

ただし、学習モデル 1 a を実現する CNN s は、非常にディープな CNN s であり、学習モデル 1 a の十分に深い中間層から出力される特徴ベクトルは、高次元特徴量を示すものである。このため、上記の相違は僅かなものである。

また、学習モデル 1 a の出力層から出力される特徴ベクトルが示す特徴量は、上述したように、検知対象物体の、十分に深い中間層の複数の層の複数の特徴量のそれぞれがぼかされてから複数の層の特徴量が結合された特徴量である。このため、特徴ベクトルが示す特徴量は、画像のドメインの依存性及び認識タスクの依存性のそれぞれが排除された頑健な特徴を示すものである。

【 0 0 3 9 】

40

次に、ドメイン変換時における推論装置 4 の動作について説明する。

図 7 は、ドメイン変換時の推論装置 4 の処理手順を示すフローチャートである。

カメラ 3 は、検知対象物体を撮影する。カメラ 3 により撮影される検知対象物体は、ドローン、空飛ぶクルマ、又は、ヘリコプターのいずれかである。ただし、推論装置 4 は、検知対象物体を、例えば 1 0 0 0 に分類することも可能である。このため、検知対象物体を、ドローン、空飛ぶクルマ、又は、ヘリコプターの 3 つに分類することは、あくまでも、一例である。また、カメラ 3 により撮影される検知対象物体は、いずれかの領域に存在している。

カメラ 3 は、いずれかの領域に存在している検知対象物体が映っている変換用画像を示す画像信号として、例えば、検知対象物体が映っている TIR 画像を示す画像信号を推論

50

装置 4 に出力する。

【 0 0 4 0 】

画像信号取得部 1 1 は、カメラ 3 から、検知対象物体が映っている変換用画像を示す画像信号を取得する（図 7 のステップ S T 1 ）。

具体的には、画像信号取得部 1 1 は、カメラ 3 から、いずれかの領域に存在しているドローンが映っている変換用画像を示す画像信号を取得し、変換用画像を示す画像信号を特徴量抽出部 1 2 に出力する。

また、画像信号取得部 1 1 は、カメラ 3 から、いずれかの領域に存在している空飛ぶクルマが映っている変換用画像を示す画像信号を取得し、変換用画像を示す画像信号を特徴量抽出部 1 2 に出力する。

10

また、画像信号取得部 1 1 は、カメラ 3 から、いずれかの領域に存在しているヘリコプターが映っている変換用画像を示す画像信号を取得し、変換用画像を示す画像信号を特徴量抽出部 1 2 に出力する。

【 0 0 4 1 】

特徴量抽出部 1 2 は、画像信号取得部 1 1 から、いずれかの領域に存在している検知対象物体が映っている変換用画像を示す画像信号を取得する。

特徴量抽出部 1 2 は、それぞれの画像信号から、それぞれの変換用画像に映っている検知対象物体の特徴量を抽出する（図 7 のステップ S T 2 ）。

具体的には、特徴量抽出部 1 2 は、画像信号取得部 1 1 から、いずれかの領域に存在しているドローンが映っている変換用画像を示す画像信号を取得する。

20

特徴量抽出部 1 2 は、それぞれの画像信号を第 1 の学習モデル 1 a に与えて、第 1 の学習モデル 1 a から、いずれかの領域に存在しているドローンの複数の特徴量のそれぞれがばかされてから複数の特徴量が結合された代表特徴量を示す特徴ベクトルを取得する。

特徴量抽出部 1 2 は、特徴ベクトルを代表特徴量登録部 1 3 に出力する。

【 0 0 4 2 】

また、特徴量抽出部 1 2 は、画像信号取得部 1 1 から、いずれかの領域に存在している空飛ぶクルマが映っている変換用画像を示す画像信号を取得する。

特徴量抽出部 1 2 は、それぞれの画像信号を第 1 の学習モデル 1 a に与えて、第 1 の学習モデル 1 a から、いずれかの領域に存在している空飛ぶクルマの複数の特徴量のそれぞれがばかされてから複数の特徴量が結合された代表特徴量を示す特徴ベクトルを取得する。

30

特徴量抽出部 1 2 は、特徴ベクトルを代表特徴量登録部 1 3 に出力する。

【 0 0 4 3 】

また、特徴量抽出部 1 2 は、画像信号取得部 1 1 から、いずれかの領域に存在しているヘリコプターが映っている変換用画像を示す画像信号を取得する。

特徴量抽出部 1 2 は、それぞれの画像信号を第 1 の学習モデル 1 a に与えて、第 1 の学習モデル 1 a から、いずれかの領域に存在しているヘリコプターの複数の特徴量のそれぞれがばかされてから複数の特徴量が結合された代表特徴量を示す特徴ベクトルを取得する。

特徴量抽出部 1 2 は、特徴ベクトルを代表特徴量登録部 1 3 に出力する。

【 0 0 4 4 】

代表特徴量登録部 1 3 は、特徴量抽出部 1 2 から、それぞれの特徴ベクトルを取得する。

40

代表特徴量登録部 1 3 は、それぞれの特徴ベクトルを代表特徴量記憶部 1 4 に記憶させることで、代表特徴量を登録する（図 7 のステップ S T 3 ）。

【 0 0 4 5 】

ここで、代表特徴量を示す特徴ベクトルは、Tensor で表現されている。Tensor は、Vector よりも高次元な情報を表現できるものであり、フィーチャマップと呼ばれることがある。

Tensor は、高次元な情報を表現できるものであるため、代表特徴量登録部 1 3 が、Tensor で表現されている特徴ベクトルをそのまま代表特徴量記憶部 1 4 に記憶させた場合、3 次元位置推定部 1 5 が特徴ベクトルを照合する際に、多くの処理時間を要することがある。

50

３次元位置推定部１５が特徴ベクトルを照合する際に要する処理時間を短縮するために、代表特徴量登録部１３が、代表特徴量を示す特徴ベクトルを、Tensorよりも次元数が少ないワンホットベクトル(One-hot-vector)に変換し、ワンホットベクトルを代表特徴量記憶部１４に記憶させるようにしてもよい。

代表特徴量登録部１３によって登録される特徴ベクトルが、Tensor、又は、ワンホットベクトルのいずれであっても、数百次元等の高次元の情報が表現されるものである。このため、種類が同一である複数の検知対象物体の間に多少のばらつきがあったとしても、特徴ベクトルは、当該検知対象物体の代表的な特徴が高次元で記述されたものになっている。

#### 【００４６】

次に、位置推定時の推論装置４の動作について説明する。

図８は、位置推定時の推論装置４の処理手順である推論方法を示すフローチャートである。

カメラ３は、検知対象物体を撮影する。カメラ３により撮影される検知対象物体は、ドローン、空飛ぶクルマ、又は、ヘリコプターのいずれであるのかが不明である。また、カメラ３により撮影される検知対象物体が存在している領域は、不明である。

カメラ３は、検知対象物体が映っている推論対象画像を示す画像信号として、例えば、検知対象物体が映っているTIR画像を示す画像信号を推論装置４に出力する。

#### 【００４７】

画像信号取得部１１は、カメラ３から、検知対象物体が映っている推論対象画像を示す画像信号を取得する(図８のステップＳＴ１１)。

画像信号取得部１１は、推論対象画像を示す画像信号を特徴量抽出部１２に出力する。

#### 【００４８】

特徴量抽出部１２は、画像信号取得部１１から、推論対象画像を示す画像信号を取得する。

特徴量抽出部１２は、画像信号から、推論対象画像に映っている検知対象物体の特徴量を抽出する(図８のステップＳＴ１２)。

具体的には、特徴量抽出部１２は、画像信号を第１の学習モデル１ａに与えて、第１の学習モデル１ａから、推論対象画像に映っている検知対象物体の高次元特徴量を示す特徴ベクトルを取得する。

特徴量抽出部１２は、特徴ベクトルを３次元位置推定部１５に出力する。

#### 【００４９】

３次元位置推定部１５は、特徴量抽出部１２から、特徴ベクトルを取得する。

３次元位置推定部１５は、代表特徴量記憶部１４から、代表特徴量を示す特徴ベクトルを複数取得する。

３次元位置推定部１５は、複数の代表特徴量を示す特徴ベクトルと、特徴量抽出部１２から取得した特徴ベクトルとに基づいて、推論対象画像に映っている検知対象物体の３次元位置を推定する(図８のステップＳＴ１３)。

#### 【００５０】

具体的には、３次元位置推定部１５は、複数の代表特徴量を示す特徴ベクトルと、特徴量抽出部１２から取得した特徴ベクトルとの類似度をそれぞれ算出する。

#### 【００５１】

３次元位置推定部１５は、認識したい対象の数だけ登録してある代表特徴量と推論時特徴量との類似度の中で、最も高い類似度を特定し、最も高い類似度に対応する代表特徴量を特定する。３次元位置推定部１５は、特定することで、どのクラスに属するのかを判別することが可能となる。

#### 【００５２】

３次元位置推定部１５は、最も類似した代表特徴量を持つクラスが検出対象物体の所属するクラスとなる。例えば、代表特徴量と推論時特徴量のいずれもがTIR画像を入力とした場合、学習時のドメインを超えたクラス識別が可能となる。

10

20

30

40

50

さらに、推論時特徴量には推論時に入力された画像に含まれる検出対象物体の存在範囲 (Objectness) を含むため、推論時特徴量の高次元特徴の記述方法をTensor等から二次元空間へ変換することで空間的な物体の存在位置を表すことが可能となる。これによって、学習時のタスクが画像識別 (Image Classification) であった場合、タスクを超えた認識が可能となる。

【0053】

3次元位置推定部15は、最も類似度の高い代表特徴量が、例えば、ドローンの代表特徴量であれば、推論対象画像に映っている検知対象物体の種類がドローンであることを認識し、かつ、検知対象物体の存在している領域を認識する。

3次元位置推定部15は、最も類似度の高い代表特徴量が、例えば、空飛ぶクルマの代表特徴量であれば、推論対象画像に映っている検知対象物体の種類が空飛ぶクルマであることを認識し、かつ、検知対象物体の存在している領域を認識する。

10

3次元位置推定部15は、最も類似度の高い代表特徴量が、例えば、ヘリコプターの代表特徴量であれば、推論対象画像に映っている検知対象物体の種類がヘリコプターであることを認識し、かつ、検知対象物体の存在している領域を認識する。

【0054】

3次元位置推定部15は、3次元位置の推定結果を変化解析部16に出力する。

また、3次元位置推定部15は、3次元位置の推定結果を示す表示データを生成し、表示データを表示装置5に出力する。

なお、3次元位置推定部15は、画像信号取得部11が画像信号を取得する毎に、3次元位置の推定結果を変化解析部16に出力する。画像信号取得部11が画像信号を取得するサンプリング時刻が $t_n$ であれば、3次元位置推定部15は、サンプリング時刻 $t_n$ における3次元位置の推定結果を変化解析部16に出力する。

20

【0055】

変化解析部16は、3次元位置推定部15から、サンプリング時刻 $t_n$ における3次元位置の推定結果を取得し、3次元位置の推定結果を第2の学習モデル2aに与える。

なお、変化解析部16は、ドローンの3次元位置の時間的な変化を示す位置データを取得する必要がある場合、ドローンの3次元位置の推定結果を第2の学習モデル2aに与える。

変化解析部16は、空飛ぶクルマの3次元位置の時間的な変化を示す位置データを取得する必要がある場合、空飛ぶクルマの3次元位置の推定結果を第2の学習モデル2aに与える。

30

変化解析部16は、ヘリコプターの3次元位置の時間的な変化を示す位置データを取得する必要がある場合、ヘリコプターの3次元位置の推定結果を第2の学習モデル2aに与える。

【0056】

第2の学習モデル2aは、3次元位置を学習して3次元位置の時間的な変化を回帰する学習モデルである。このため、第2の学習モデル2aは、変化解析部16から、サンプリング時刻 $t_n$ における3次元位置の推定結果が与えられると、当該推定結果に対応する将来のサンプリング時刻における検知対象物体の3次元位置の時間的な変化を示す位置データを変化解析部16に出力する。

40

変化解析部16は、第2の学習モデル2aから、将来のサンプリング時刻における検知対象物体の3次元位置の時間的な変化を示す位置データとして、例えば、将来のサンプリング時刻 $t_{n+1} \sim t_{n+3}$ における検知対象物体の3次元位置の時間的な変化を示す位置データを取得する。

【0057】

図1に示す推論装置4では、変化解析部16が、第2の学習モデル2aを用いて、検知対象物体の3次元位置の時間的な変化を示す位置データを取得している。しかし、これは一例に過ぎず、変化解析部16は、検知対象物体の3次元位置を予測するための予測関数に対して、サンプリング時刻 $t_n$ における検知対象物体の3次元位置の推定結果を与える

50

ことで、検知対象物体の３次元位置の時間的な変化を示す位置データを取得するようにしてもよい。

【００５８】

変化解析部１６は、位置データに基づいて、検知対象物体が落下するか否かを判定する。

例えば、検知対象物体の将来の位置が、地上の位置を示していれば、変化解析部１６は、検知対象物体が落下すると判定する。検知対象物体の将来の位置が、地上の位置を示していなければ、変化解析部１６は、検知対象物体が落下しないと判定する。

また、変化解析部１６は、位置データに基づいて、検知対象物体を多クラスに分類する処理として、検知対象物体が、例えば、ホバリングしているのか、上昇しているのか、下降しているのか、前進しているのか、又は、後退しているのかを判定する。

10

【００５９】

変化解析部１６は、検知対象物体の３次元位置の時間的な変化の解析結果を示す表示データを生成し、表示データを表示装置５に出力する。

また、変化解析部１６は、上記の判定処理の判定結果を示す表示データを生成し、表示データを表示装置５に出力する。

【００６０】

表示装置５は、３次元位置推定部１５から出力された表示データに従って、３次元位置の推定結果を図示せぬディスプレイに表示させる。３次元位置の推定結果は、検知対象物体の種類と検知対象物体の３次元位置とを示すものである。

また、表示装置５は、図９に示すように、変化解析部１６から出力された表示データに従って、検知対象物体の３次元位置の時間的な変化の解析結果と、検知対象物体が落下するか否か等の判定結果とを図示せぬディスプレイに表示させる。

20

図９は、検知対象物体の３次元位置の時間的な変化を示す説明図である。

図９は、時刻 $t$ における検知対象物体の $x$ 方向の位置と、時刻 $t$ における検知対象物体の $y$ 方向の位置とを示している。

図９において、 $x$ 方向は、カメラ３の撮影方向と直交する方向であって、例えば、地面と水平な方向である。

$y$ 方向は、カメラ３の撮影方向と平行な方向である。 $z$ 方向は、カメラ３の撮影方向と直交する方向であって、例えば、地面と垂直な方向である。

図９の例では、検知対象物体の $z$ 方向の位置が、或る時刻において、地上の位置になっているため、検知対象物体が落下する旨の判定結果を明示している。図９において、検知対象物体が落下する旨の判定結果が“メッセージ”として表示されるようにしてもよい。

30

【００６１】

以上の実施の形態１では、学習用画像と画像のドメインが異なる場合と、事前学習したタスクと認識タスクの異なる場合とのうち、１つ以上が異なる場合において、検知対象物体が映っている画像である推論対象画像を示す画像信号を取得する画像信号取得部１１と、画像信号取得部１１により取得された画像信号を、学習用画像の学習が済んでいる第１の学習モデル１ａに与えて、第１の学習モデル１ａから、推論対象画像に映っている検知対象物体の複数の特徴量のそれぞれがほかされてから複数の特徴量が結合された特徴量であって、検知対象物体の３次元位置の推論に用いられる特徴量である推論時特徴量を取得する特徴量抽出部１２とを備えるように、推論装置４を構成した。また、推論装置４は、推論対象画像と画像のドメイン及び認識タスクのそれぞれが同じ対象である変換用画像に映っている検知対象物体の登録済みの特徴量である代表特徴量と特徴量抽出部１２により取得された推論時特徴量とに基づいて、推論対象画像に映っている検知対象物体の３次元位置を推定する３次元位置推定部１５と、３次元位置推定部１５による３次元位置の推定結果に基づいて、推論対象画像に映っている検知対象物体の３次元位置の時間的な変化を解析する変化解析部１６とを備えている。したがって、推論装置４は、自然環境の状態を考慮して、第１の学習モデルを学習させることなく、検知対象物体の３次元位置の時間的な変化を解析することができる。

40

【００６２】

50

また、実施の形態 1 では、画像信号取得部 1 1 が、変換用画像を示す画像信号を取得し、特徴量抽出部 1 2 が、変換用画像を示す画像信号を第 1 の学習モデル 1 a に与えて、第 1 の学習モデル 1 a から、変換用画像に映っている検知対象物体の複数の特徴量のそれぞれがばかされてから複数の特徴量が結合された特徴量である代表特徴量を取得するように、推論装置 4 を構成した。また、推論装置 4 は、特徴量抽出部 1 2 により取得された代表特徴量を登録する代表特徴量登録部 1 3 を備えている。したがって、推論装置 4 は、検知対象物体の 3 次元位置の推定に用いることが可能な代表特徴量を登録することができる。

#### 【 0 0 6 3 】

図 1 に示す推論装置 4 では、特徴量抽出部 1 2 が、非常にディープな CNNs によって実現されている第 1 の学習モデル 1 a に対して画像信号を与えて、第 1 の学習モデル 1 a から、検知対象物体の複数の特徴量のそれぞれがばかされてから複数の特徴量が結合された特徴量を取得している。

10

第 1 の学習モデル 1 a が非常にディープな CNNs によって実現されている場合には、上述したように、第 1 の学習モデル 1 a の入力層に与えられる画像信号が示す推論対象画像のドメインが学習用画像と異なっている場合、また、検知対象物体が異なっている場合、出力層から出力される特徴ベクトルの相違は僅かなものとなる。

これに対して、第 1 の学習モデル 1 a が一般的なニューラルネットワーク等によって実現されている場合、第 1 の学習モデル 1 a の入力層に与えられる画像信号が示す推論対象画像のドメイン、又は、推論対象画像に映っている検知対象物体が、学習用画像と異なっていれば、出力層から出力される特徴ベクトルの相違が大きくなることがある。

20

しかしながら、変換用画像のドメインと推論対象画像のドメインとが同じである。このため、第 1 の学習モデル 1 a が一般的なニューラルネットワーク等によって実現されている場合でも、推論対象画像に映っている検知対象物体が正常な物体であれば、ドメイン変換時に特徴量抽出部 1 2 により取得される代表特徴量と、位置推定時に特徴量抽出部 1 2 により取得される推論時特徴量とは、概ね同じような値となる。

一方、推論対象画像に映っている検知対象物体が異常な物体であれば、ドメイン変換時に特徴量抽出部 1 2 により取得される代表特徴量と、位置推定時に特徴量抽出部 1 2 により取得される推論時特徴量とは、大きく異なる値となる。

したがって、第 1 の学習モデル 1 a が一般的なニューラルネットワーク等によって実現されている場合でも、3 次元位置推定部 1 5 が、高精度に検知対象物体の 3 次元位置を推定することが可能である。

30

#### 【 0 0 6 4 】

なお、本開示は、実施の形態の任意の構成要素の変形、もしくは実施の形態の任意の構成要素の省略が可能である。

#### 【産業上の利用可能性】

#### 【 0 0 6 5 】

本開示は、推論装置、推論方法及び推論プログラムに適している。

#### 【符号の説明】

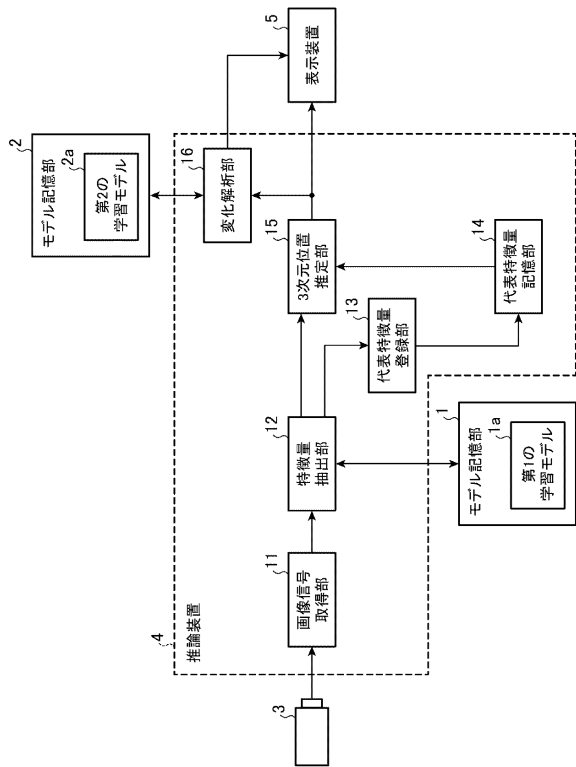
#### 【 0 0 6 6 】

1 モデル記憶部、1 a 第 1 の学習モデル、2 モデル記憶部、2 a 第 2 の学習モデル、3 カメラ、4 推論装置、5 表示装置、6 学習データ記憶部、7 学習装置、1 1 画像信号取得部、1 2 特徴量抽出部、1 3 代表特徴量登録部、1 4 代表特徴量記憶部、1 5 3 次元位置推定部、1 6 変化解析部、2 1 画像信号取得回路、2 2 特徴量抽出回路、2 3 代表特徴量登録回路、2 4 代表特徴量記憶回路、2 5 3 次元位置推定回路、2 6 変化解析回路、3 1 メモリ、3 2 プロセッサ、4 1 学習データ取得部、4 2 学習処理部、5 1 学習データ取得回路、5 2 学習処理回路、6 1 メモリ、6 2 プロセッサ。

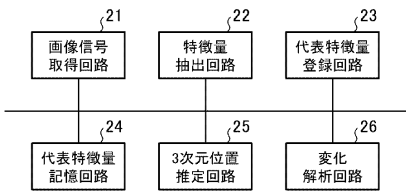
40

【図面】

【図 1】



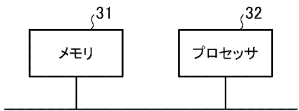
【図 2】



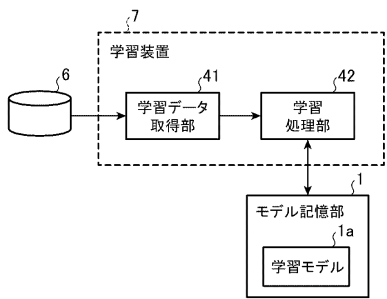
10

20

【図 3】



【図 4】

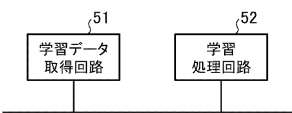


30

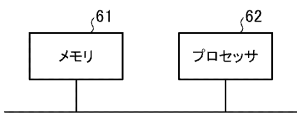
40

50

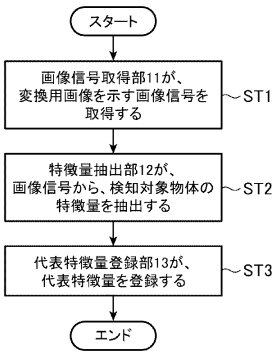
【図 5】



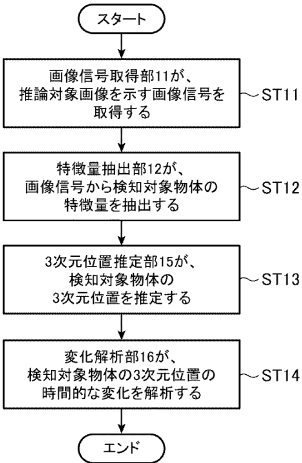
【図 6】



【図 7】



【図 8】



10

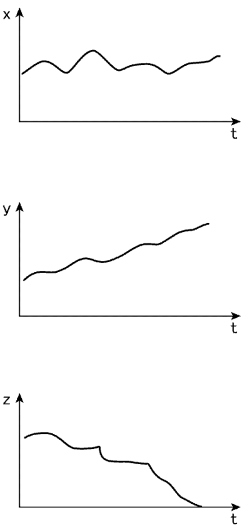
20

30

40

50

【 図 9 】



10

20

30

40

50

---

フロントページの続き

- (56)参考文献 特開 2 0 1 9 - 1 8 5 1 2 7 ( J P , A )  
特開 2 0 2 0 - 1 0 1 9 4 8 ( J P , A )  
特開 2 0 1 9 - 2 1 2 2 9 6 ( J P , A )  
特開 2 0 2 2 - 0 3 7 9 5 5 ( J P , A )
- (58)調査した分野 (Int.Cl. , D B 名)
- |         |                       |
|---------|-----------------------|
| G 0 6 N | 3 / 0 0 - 9 9 / 0 0   |
| G 0 6 T | 1 / 0 0 - 1 / 4 0     |
| G 0 6 T | 3 / 0 0 - 7 / 9 0     |
| G 0 6 V | 1 0 / 0 0 - 2 0 / 9 0 |