

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7707332号
(P7707332)

(45)発行日 令和7年7月14日(2025.7.14)

(24)登録日 令和7年7月4日(2025.7.4)

(51)国際特許分類 F I
G 0 6 T 7/73 (2017.01) G 0 6 T 7/73
G 0 6 T 7/00 (2017.01) G 0 6 T 7/00 3 0 0 D

請求項の数 6 (全14頁)

| | | | |
|-------------------|----------------------------------|----------|------------------------------------|
| (21)出願番号 | 特願2024-14382(P2024-14382) | (73)特許権者 | 520461440 |
| (22)出願日 | 令和6年2月1日(2024.2.1) | | 所羅門股 分 有限公司 |
| (65)公開番号 | 特開2024-114637(P2024-114637 A) | (74)代理人 | 100118256 台湾臺北市 内 湖區行忠路4 2 號6 樓 |
| (43)公開日 | 令和6年8月23日(2024.8.23) | | 弁理士 小野寺 隆 |
| 審査請求日 | 令和6年2月1日(2024.2.1) | (74)代理人 | 100166338 弁理士 関口 正夫 |
| (31)優先権主張番号 | 112104766 | (72)発明者 | 陳 政隆 |
| (32)優先日 | 令和5年2月10日(2023.2.10) | | 台湾臺北市 内 湖區行忠路4 2 號6 樓 |
| (33)優先権主張国・地域又は機関 | 台湾(TW) | (72)発明者 | スアン ロック グエン |
| | | | 台湾臺北市 内 湖區行忠路4 2 號6 樓 |
| | | (72)発明者 | タフジラ ヌグラハ ブリリアン |
| | | | 台湾臺北市 内 湖區行忠路4 2 號6 樓 |
| | | (72)発明者 | マン クアン グエン |
| | | | 台湾臺北市 内 湖區行忠路4 2 號6 樓 |

最終頁に続く

(54)【発明の名称】 目標物体の姿勢を識別する方法およびそれを実行する計算装置

(57)【特許請求の範囲】

【請求項1】

目標物体の姿勢を識別する方法であって、該方法は、データベースを記憶する計算装置により実行され、前記データベースは、前記目標物体が属する特定の種類に関連し、複数のテンプレート画像を含み、前記テンプレート画像は、前記特定の種類に属する参照物体をそれぞれ含み、前記参照物体が捕捉される参照角度に対する異なる偏向角度にそれぞれ対応し、前記データベースは、前記テンプレート画像にそれぞれ対応する複数の参照特徴データセットをさらに含み、前記参照特徴データセットのそれぞれは、前記テンプレート画像の対応する1つに示される角度での前記参照物体の外観特徴を示し、前記テンプレート画像は、最初のテンプレート画像と、複数の加工されたテンプレート画像と、を含み、前記方法は、

前記最初のテンプレート画像を得るステップと、

前記最初のテンプレート画像に基づいて、前記最初のテンプレート画像に対応する前記参照特徴データセットを生成するステップと、

前記最初のテンプレート画像を複数回回転することにより、前記複数の加工されたテンプレート画像を生成するステップであって、毎回の回転で1つの前記加工されたテンプレート画像を生成するステップと、

前記加工されたテンプレート画像に基づいて、前記加工されたテンプレート画像にそれぞれ対応する前記参照特徴データセットを生成するステップと、

前記特定の種類に属する前記目標物体を含む入力画像を得るステップと、

前記入力画像における前記目標物体の外観に基づいて、前記データベースの前記テンプレート画像から前記入力画像と最もマッチするマッチング画像を選択するステップであって、前記マッチング画像は、前記テンプレート画像に示される前記参照物体の前記角度が前記入力画像に示される前記目標物体の角度と最も近い1つであるステップと、

前記入力画像と前記マッチング画像とに基づいて、キーポイントマッチング工程を実行することによって、前記入力画像に示されかつ前記目標物体の外観に関連する複数の第1の特徴点と、前記マッチング画像に示されかつ前記第1の特徴点とそれぞれマッチする複数の第2の特徴点と、を認識するステップと、

前記第1の特徴点と前記第2の特徴点との間の相対姿勢関係に基づいて、前記目標物体の前記姿勢を示す姿勢認識結果を生成するステップと、を含み、

マッチング画像を選択するステップは、

前記入力画像に示される前記目標物体の複数の特徴部分に基づいて、前記入力画像に対応する目標特徴データセットを生成するステップであって、前記目標特徴データセットは、前記入力画像に示される前記角度での前記目標物体の外観特徴を示すステップと、

前記参照特徴データセットのそれぞれについて、前記目標特徴データセットと該参照特徴データセットとの間のマッチング程度を計算するステップと、

前記テンプレート画像の前記目標特徴データセットとの前記マッチング程度が最も高い前記参照特徴データセットに対応する1つを、前記マッチング画像として選択するステップと、を含み、

前記複数の加工されたテンプレート画像は、前記最初のテンプレート画像を所定の角度で複数回回転することにより生成され、前記所定の角度は、25度未満の任意の角度であり、前記テンプレート画像の数と前記所定の角度との関係は、前記テンプレート画像の数に前記所定の角度を乗じた値が360度に等しいという条件を満たす、

方法。

【請求項2】

前記目標特徴データセットと前記参照特徴データセットとは、それぞれベクトルで表現され、

マッチング程度を計算するステップは、前記参照特徴データセットのそれぞれについて、前記目標特徴データセットと該参照特徴データセットとの間のミンコフスキー距離を計算するステップを含む、

請求項1に記載の方法。

【請求項3】

前記データベースは、前記テンプレート画像にそれぞれ対応する複数の参照姿勢データセットをさらに含み、前記参照姿勢データセットのそれぞれは、前記テンプレート画像の対応する1つに示される前記参照物体の姿勢を示し、

前記方法は、前記第1の特徴点と前記第2の特徴点とに基づいて、前記第1の特徴点と前記第2の特徴点との間の前記相対姿勢関係を示す校正データセットを生成することをさらに含み、

前記姿勢認識結果は、前記マッチング画像に対応する前記参照姿勢データセットと前記校正データセットとに基づいて生成される、

請求項1に記載の方法。

【請求項4】

目標物体の姿勢を認識する計算装置であって、

処理ユニットと、

前記処理ユニットに電氣的に接続し、データベースを記憶する記憶ユニットと、を含み、

前記データベースは、前記目標物体が属する特定の種類に関連し、複数のテンプレート画像を含み、前記テンプレート画像は、前記特定の種類に属する参照物体をそれぞれ含み、前記参照物体が捕捉される参照角度に対する異なる偏向角度にそれぞれ対応し、前記データベースは、前記テンプレート画像にそれぞれ対応する複数の参照特徴データセットをさらに含み、前記参照特徴データセットのそれぞれは、前記テンプレート画像の対応する

10

20

30

40

50

1つに示される角度での前記参照物体の外観特徴を示し、前記テンプレート画像は、最初のテンプレート画像と、複数の加工されたテンプレート画像と、を含み、

前記処理ユニットは、

前記特定の種類の属する前記目標物体を含む入力画像を得て、

前記入力画像における前記目標物体の外観に基づいて、前記データベースの前記テンプレート画像から前記入力画像と最もマッチするマッチング画像を選択し、前記マッチング画像は、前記テンプレート画像に示される前記参照物体の前記角度が前記入力画像に示される前記目標物体の角度と最も近い1つであり、

前記入力画像と前記マッチング画像とに基づいて、キーポイントマッチング工程を実行することによって、前記入力画像に示されかつ前記目標物体の外観に関連する複数の第1の特徴点と、前記マッチング画像に示されかつ前記第1の特徴点とそれぞれマッチする複数の第2の特徴点と、を認識し、

前記第1の特徴点と前記第2の特徴点との間の相対姿勢関係に基づいて、前記目標物体の前記姿勢を示す姿勢認識結果を生成する、ように構成され、

前記処理ユニットは、

前記入力画像に示される前記目標物体の複数の特徴部分に基づいて、前記入力画像に対応する目標特徴データセットを生成し、前記目標特徴データセットは、前記入力画像に示される前記角度での前記目標物体の外観特徴を示し、

前記参照特徴データセットのそれぞれについて、前記目標特徴データセットと該参照特徴データセットとの間のマッチング程度を計算し、

前記テンプレート画像の前記目標特徴データセットとの前記マッチング程度が最も高い前記参照特徴データセットに対応する1つを、前記マッチング画像として選択することによって、前記マッチング画像を選択するように構成され、

前記処理ユニットは、前記入力画像を得ることと前記マッチング画像を選択することとの前に、

前記最初のテンプレート画像を得て、

前記最初のテンプレート画像に基づいて、前記最初のテンプレート画像に対応する前記参照特徴データセットを生成し、

前記最初のテンプレート画像を複数回回転することにより、前記複数の加工されたテンプレート画像を生成し、毎回の回転で1つの前記加工されたテンプレート画像を生成し、

前記加工されたテンプレート画像に基づいて、前記加工されたテンプレート画像にそれぞれ対応する前記参照特徴データセットを生成するように、さらに構成され、

前記複数の加工されたテンプレート画像は、前記処理ユニットが前記最初のテンプレート画像を所定の角度で複数回回転することにより生成され、前記所定の角度は、25度未満の任意の角度であり、前記テンプレート画像の数と前記所定の角度との関係は、前記テンプレート画像の数に前記所定の角度を乗じた値が360度に等しいという条件を満たす計算装置。

【請求項5】

前記目標特徴データセットと前記参照特徴データセットとは、それぞれベクトルで表現され、

前記処理ユニットは、前記参照特徴データセットのそれぞれについて、前記目標特徴データセットと該参照特徴データセットとの間のミンコフスキー距離を計算することによって、前記マッチング程度を計算するように構成される、

請求項4に記載の計算装置。

【請求項6】

前記データベースは、前記テンプレート画像にそれぞれ対応する複数の参照姿勢データセットをさらに含み、前記参照姿勢データセットのそれぞれは、前記テンプレート画像の対応する1つに示される前記参照物体の姿勢を示し、

前記処理ユニットは、前記第1の特徴点と前記第2の特徴点とに基づいて、前記第1の特徴点と前記第2の特徴点との間の前記相対姿勢関係を示す校正データセットを生成する

10

20

30

40

50

ようにさらに構成され、

前記処理ユニットは、前記マッチング画像に対応する前記参照姿勢データセットと前記校正データセットとに基づいて、前記姿勢認識結果を生成するように構成される、

請求項4に記載の計算装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、目標物体の姿勢を認識する方法に関し、より具体的には、画像認識に關与する目標物体の姿勢を認識する方法に関する。本発明は、該方法を実行する計算装置にさら

10

【背景技術】

【0002】

さまざまな自動化操作において、物体の姿勢（位置と方向付け）を識別することは、重要な工程である。例えば、ロボットアームを用いて物体を掴む操作において、物体の姿勢に基づいてロボットアームの動作を決めることによって、ロボットアームは正しい角度から物体を安定的に掴むことができる。また、物体を自動化加工する操作において、物体の姿勢に基づいて加工する位置または経路を調整することによって、加工の精確性を向上

【先行技術文献】

【特許文献】

20

【0003】

【文献】中国特許出願公開第113939852号明細書

【発明の概要】

【発明が解決しようとする課題】

【0004】

したがって、本発明の目的は、目標物体の姿勢を識別する方法およびそれを実行する計算装置を提供することにある。

【課題を解決するための手段】

【0005】

該方法は、データベースを記憶する計算装置により実行される。データベースは、目標物体が属する特定の種類に関連し、複数のテンプレート画像を含む。テンプレート画像は、特定の種類に属する参照物体をそれぞれ含み、参照物体が捕捉される参照角度に対する異なる偏向角度にそれぞれ対応する。

30

【0006】

該方法は、特定の種類に属する目標物体を含む入力画像を得るステップと、入力画像における目標物体の外観に基づいて、データベースのテンプレート画像から入力画像と最もマッチするマッチング画像を選択するステップであって、マッチング画像は、テンプレート画像に示される参照物体の角度が入力画像に示される目標物体の角度と最も近い1つであるステップと、入力画像とマッチング画像とに基づいて、キーポイントマッチング工程を実行することによって、入力画像に示されかつ目標物体の外観に関連する複数の第1の特徴点と、マッチング画像に示されかつ第1の特徴点とそれぞれマッチする複数の第2の特徴点と、を認識するステップと、第1の特徴点と第2の特徴点との間の相対姿勢関係に基づいて、目標物体の姿勢を示す姿勢認識結果を生成するステップと、を含む。

40

【0007】

計算装置は、処理ユニットと、処理ユニットに電氣的に接続し、データベースを記憶する記憶ユニットと、を含む。

【0008】

データベースは、目標物体が属する特定の種類に関連し、複数のテンプレート画像を含む。テンプレート画像は、特定の種類に属する参照物体をそれぞれ含み、参照物体が捕捉される参照角度に対する異なる偏向角度にそれぞれ対応する。

50

【 0 0 0 9 】

処理ユニットは、特定の種類に属する目標物体を含む入力画像を得て、入力画像における目標物体の外観に基づいて、データベースのテンプレート画像から入力画像と最もマッチするマッチング画像を選択し、マッチング画像は、テンプレート画像に示される参照物体の角度が入力画像に示される目標物体の角度と最も近い1つであり、入力画像とマッチング画像とに基づいて、キーポイントマッチング工程を実行することによって、入力画像に示されかつ目標物体の外観に関連する複数の第1の特徴点と、マッチング画像に示されかつ第1の特徴点とそれぞれマッチする複数の第2の特徴点と、を認識し、第1の特徴点と第2の特徴点との間の相対姿勢関係に基づいて、目標物体の姿勢を示す姿勢認識結果を生成する、ように構成される。

10

【 発明の効果 】

【 0 0 1 0 】

目標物体を認識する方法を実行することによって、計算装置は、目標物体を含む入力画像を得ると、データベースから入力画像と最もマッチするマッチング画像を選択し、入力画像とマッチング画像とに基づいてキーポイントマッチング工程を実行して、第1の特徴点と第2の特徴点とを認識し、第1の特徴点と第2の特徴点とに基づいて、目標物体の姿勢を認識する。

【 0 0 1 1 】

本発明の他の特徴及び利点は、添付の図面を参照する以下の実施形態の詳細な説明において明白になる。

20

【 図面の簡単な説明 】

【 0 0 1 2 】

【 図 1 】 本発明に係る一実施形態の目標物体の姿勢を認識する計算装置、データベース、および撮影装置を示す模式図である。

【 図 2 】 本発明に係る一実施形態の目標物体の姿勢を認識する方法を示すフローチャートである。

【 発明を実施するための形態 】

【 0 0 1 3 】

本発明をより詳細に説明する前に、適切と考えられる場合において、参照符号または参照符号の末端部は、同様の特性を有し得る対応のまたは類似の要素を示すために各図面間で繰り返し用いられることに留意されたい。

30

【 0 0 1 4 】

本明細書において、「結合」又は「接続」という用語は、複数の電気機器/装置/設備の間が導電材料(例えば、電線)により直接的に接続されること、あるいは、2つの電気機器/装置/設備の間が他の一つ以上の機器/装置/設備または無線通信により間接的に接続されることを意味する。

【 0 0 1 5 】

図1を参照すると、本発明に係る目標物体の姿勢を認識する計算装置1は、パーソナルコンピューター、ラップトップコンピューター、サーバーなどとして実施され得る。計算装置1は、処理ユニット11と、処理ユニット11に電氣的に接続する記憶ユニット12と、を含む。なお、「姿勢」という用語は、物体の位置と方向付け(即ち物体の向き)とを表している。

40

【 0 0 1 6 】

本実施形態において、処理ユニット11は、データ演算および処理機能を有するチップ(すなわち、集積回路)であり、中央処理装置(CPU)を含んでもよいが、これに限定されない。記憶ユニット12は、デジタルデータを記憶するデータ記憶装置(例えば、ハードディスク)である。いくつかの実施形態において、処理ユニット11は、チップの集合体であってもよく、記憶ユニット12は、同じタイプまたは異なるタイプのコンピュータ読み取り可能な記憶媒体の集合体であってもよいが、これに限定されない。いくつかの実施形態において、計算装置1は、スマートフォン、タブレットコンピューターなどの電

50

子装置であってもよい。したがって、計算装置 1 のハードウェアの実施は、本実施形態に限定されない。

【0017】

本実施形態において、処理ユニット 11 は、図 1 に示されるデータベース DB 1 を生成し、データベース DB 1 を記憶ユニット 12 に記憶するように構成される。具体的には、データベース DB 1 は、特定の種類の物体（例えば、特定の形式の物体、特定の仕様に準拠する物体など）に関連し、本発明のいくつかの実施形態に係る該特定の種類に属する目標物体の姿勢を認識する方法の第 1 の段階を実行する処理ユニット 11 によって生成される。すなわち、処理ユニット 11 が該方法の第 1 の段階を実行する前に、データベース DB 1 はまだ生成されておらず、記憶ユニット 12 に記憶されていない。なお、図 1 では、計算装置 1 の動作を容易に説明するために、データベース DB 1 の一例を予め示している。さらに、処理ユニット 11 は、動画を撮るように構成される撮影装置 2（例えば、カメラ）に電氣的に接続できるように構成され、処理ユニット 11 は、撮影装置 2 から動画を受信し、目標物体の姿勢を認識する方法の第 2 の段階を実行することによって、データベース DB 1 に基づいて、動画における物体の姿勢を認識するように構成される。

10

【0018】

図 2 をさらに参照すると、計算装置 1 によって実行される目標物体の姿勢を認識する方法が示される。該方法は、ステップ S 1 からステップ S 9 を含み、ステップ S 1 からステップ S 5 は方法の第 1 の段階に関連し、ステップ S 6 からステップ S 9 は方法の第 2 の段階に関連する。

20

【0019】

ステップ S 1 において、処理ユニット 11 は、テンプレート画像 D 1 を得る。ここで、該テンプレート画像 D 1 は、本実施形態の最初のテンプレート画像 D 1' とみなされ、最初のテンプレート画像 D 1' は、参照物体を含む。

【0020】

具体的には、最初のテンプレート画像 D 1' に含まれる参照物体は、後で計算装置 1 によって該方法を用いて認識される目標物体が属する特定の種類に属する。参照物体は、例えば、特定の種類のワークピースまたはコンポーネントである。例えば、特定の種類は、UTS (UNIFIED THREAD STANDARD) 規格番号「# 6 - 3 2 UNC 2 B」であり、参照物体および目標物体は、「# 6 - 3 2 UNC 2 B」に準拠するねじである。最初のテンプレート画像 D 1' に示される参照物体の角度は、最初のテンプレート画像 D 1' が撮影される際の撮影視点からの特定の種類の物体の参照角度とみなされる。より詳細に説明すると、撮影視点は、最初のテンプレート画像 D 1' が撮影される際の参照物体に対するカメラ装置（例えば、図 1 の撮影装置 2 であるがこれに限定されない）の視点である。なお、特定の種類に属する他の物体（すなわち、参照物体と実質的に同じ外観を有する他の物体）の姿勢を認識するために、最初のテンプレート画像 D 1' は、特定の種類に関連するデータベース DB 1 を生成するために処理ユニット 11 によって利用される。したがって、参照物体自体が、方法を用いて計算装置 1 によって認識される実際のターゲットであるとは限らない。

30

【0021】

さらに、最初のテンプレート画像 D 1' は、複数の特徴マーカを有し、特徴マーカは、最初のテンプレート画像 D 1' に示される参照物体の複数の特徴部分（以下、「参照特徴部分」と称する）をそれぞれ定義する。参照特徴部分のそれぞれは、最初のテンプレート画像 D 1' に示される参照物体の外観の部分であり、参照物体の外観の部分は、コンピュータビジョンにおける認識されるターゲットとして適合するものであり、例えば、最初のテンプレート画像 D 1' に示される参照物体のパターン、輪郭、またはそれらの組み合わせであるが、これらに限定されない。さらに、特徴マーカは、手動操作によって、最初のテンプレート画像 D 1' に追加されるバーチャルマーカであってもよいが、これに限定されない。

40

【0022】

本実施形態において、処理ユニット 11 は、ユーザにより制御されて記憶ユニット 12

50

から、記憶ユニット12に予め記憶されている最初のテンプレート画像D1'を読み取ることによって、最初のテンプレート画像D1'を得る。いくつかの実施形態において、処理ユニット11は、電子装置（図示せず。例えば、カメラ装置やコンピュータ）に電氣的に接続し、電子装置から最初のテンプレート画像D1'を得てもよい。いくつかの実施形態において、処理ユニット11は、最初のテンプレート画像D1'を記憶する外部記憶装置（図示せず。例えば、メモリカードやフラッシュドライブ）に電氣的に接続し、外部電子装置から最初のテンプレート画像D1'を得てもよい。

【0023】

最初のテンプレート画像D1'を得た後、フローはステップS2に進む。

【0024】

ステップS2において、処理ユニット11は、最初のテンプレート画像D1'に基づいて、最初のテンプレート画像D1'に対応する参照姿勢データセットD2と、最初のテンプレート画像D1'に対応する参照特徴データセットD3とを生成する。

【0025】

本実施形態において、最初のテンプレート画像D1'に対応する参照姿勢データセットD2は、マトリックスで表現され、最初のテンプレート画像D1'に示される参照物体の姿勢（すなわち、撮影視点から見た特定の種類の物体の参照角度での姿勢）を示す。具体的には、参照姿勢データセットD2は、3次元空間における6自由度を用いて、最初のテンプレート画像D1'に示される参照物体の姿勢を示す。参照姿勢データセットD2によって示される姿勢は、撮影視点に基づいて定義されるバーチャル3次元座標系の座標セットであってもよく、撮影視点から見た特定の種類に属する物体の参照姿勢と見なされてもよい。例えば、参照姿勢での特定の種類に属する物体は、3次元座標系の原点に位置し、3次元のそれぞれにある方向付けはゼロである（すなわち、ヨー、ピッチ、ロールはそれぞれゼロである）。

【0026】

本実施形態において、最初のテンプレート画像D1'に対応する参照特徴データセットD3は、複数の成分（例えば、2304個の成分）を含むベクトルで表現され、成分を用いて最初のテンプレート画像D1'に示される参照角度での参照物体（特定の種類の物体）の外観特徴を示す。外観特徴は、撮影視点から見て、特徴の種類に属する物体が該参照角度にある場合のみに存在する。例えば、外観特徴は、最初のテンプレート画像D1'に示される参照特徴部分それぞれの座標セットおよび/または方向付けであり、外観特徴は、撮影視点から見た参照物体が参照角度にある際の参照特徴部分の間の相対位置関係を示すことができるが、これに限定されない。

【0027】

本実施形態において、処理ユニット11は、最初のテンプレート画像D1'を、予めトレーニングされた畳み込みニューラルネットワーク（CNN）に入力して、出力となる参照特徴データセットD3を得る。すなわち、参照特徴データセットD3は、処理ユニット11が畳み込みニューラルネットワーク（CNN）を用いて取得した最初のテンプレート画像D1'の分析結果である。なお、畳み込みニューラルネットワーク（CNN）は、本発明のポイントではなく、当業者にとって周知であるため、簡潔性を考慮して、その詳細は省略される。

【0028】

処理ユニット11は、最初のテンプレート画像D1'に対応する参照姿勢データセットD2と参照特徴データセットD3とを生成した後、フローはステップS3に進む。

【0029】

ステップS3において、処理ユニット11は、最初のテンプレート画像D1'をN回回転することによって、最初のテンプレート画像D1'と異なるN個のテンプレート画像D1'を生成する。毎回の回転で1つのテンプレート画像D1'を生成する。テンプレート画像D1'のそれぞれは、参照物体を含む。本実施形態において、最初のテンプレート画像D1'をN回回転することにより生成されたN個のテンプレート画像D1'は、N個の加工されたテン

10

20

30

40

50

プレート画像 $D1'$ とみなされる。具体的には、 N は 1 以上の整数であり、回転は時計回りまたは反時計回りであってもよい。

【0030】

より詳細に説明すると、処理ユニット 11 は、所定の角度に基づいて、最初のテンプレート画像 $D1'$ を複数回（すなわち、 N 回）回転し、最初のテンプレート画像 $D1'$ が回転される度に、1 つの加工されたテンプレート画像 $D1''$ が生成される。本実施形態において、 N は 35 であり、所定の角度は 10 度である。したがって、処理ユニット 11 によって生成された加工されたテンプレート画像 $D1''$ の 1 つ目（以下、「第 1 の加工されたテンプレート画像」と称する）は、最初のテンプレート画像 $D1'$ を（例えば、時計回りに）10 度回転した結果を示し、処理ユニット 11 によって生成された加工されたテンプレート画像 $D1''$ の 2 つ目（以下、「第 2 の加工されたテンプレート画像」と称する）は、最初のテンプレート画像 $D1'$ をさらに 10 度回転（すなわち、時計回りに合計 20 度回転）した結果を示し、その他の加工されたテンプレート画像 $D1''$ も同様である。このように、第 1 の加工されたテンプレート画像 $D1''$ に示される参照物体の角度（以下、「参照物体の表示角度」と称する）は、最初のテンプレート画像 $D1'$ に示される参照物体の表示角度に対して時計回りに 10 度ずれている。同様に、第 2 の加工されたテンプレート画像 $D1''$ に示される参照物体の表示角度は、第 1 の加工されたテンプレート画像 $D1''$ に示される参照物体の表示角度に対して時計回りに 10 度ずれており、最初のテンプレート画像 $D1'$ に示される参照物体の表示角度に対して時計回りに 20 度ずれている。最終的に、本実施形態において、処理ユニット 11 は、最初のテンプレート画像 $D1'$ を 35 回回転して 35 個の加工されたテンプレート画像 $D1''$ を生成することによって、合計 36 個のテンプレート画像 $D1$ を得る。すなわち、36 個のテンプレート画像 $D1$ は、最初のテンプレート画像 $D1'$ と 35 個の加工されたテンプレート画像 $D1''$ とを含み、テンプレート画像 $D1$ のそれぞれにおいて、各テンプレート画像 $D1$ の参照物体の表示角度は、テンプレート画像 $D1$ の前の 1 つまたは後の 1 つの参照物体の表示角度に対して 10 度ずれている。

【0031】

なお、処理ユニット 11 によって最初のテンプレート画像 $D1'$ を回転するための所定の角度と回転の回数（すなわち、 N 回）は、上述の例に限定されない。他の実施形態において、所定の角度は、テンプレート画像 $D1$ の数と所定の角度との関係が、テンプレート画像 $D1$ の数に所定の角度を乗じた値が 360 度に等しいという条件を満たす限り、25 度より小さい任意の角度であってもよい。

【0032】

より詳細に説明すると、本実施形態において、テンプレート画像 $D1$ は、参照物体が捕捉される参照角度に対する異なる偏向角度にそれぞれ対応する。すなわち、偏向角度のそれぞれは、対応のテンプレート画像 $D1$ の参照物体の表示角度と最初のテンプレート画像 $D1'$ の参照物体の表示角度との間の差である。本実施形態の一例において、最初のテンプレート画像 $D1'$ に対応する偏向角度は 0 度であり、加工されたテンプレート画像 $D1''$ にそれぞれ対応する残りの偏向角度は、次から次へと 10 度ずつずれている（すなわち、10 度、20 度、30 度、……、350 度）が、本発明はこれに限定されない。

【0033】

処理ユニット 11 が加工されたテンプレート画像 $D1''$ を生成した後、フローはステップ S4 に進む。

【0034】

ステップ S4 において、処理ユニット 11 は、加工されたテンプレート画像 $D1''$ に基づいて、加工されたテンプレート画像 $D1''$ にそれぞれ対応する複数の参照姿勢データセット D2 と、加工されたテンプレート画像 $D1''$ にそれぞれ対応する複数の参照特徴データセット D3 と、をさらに生成する。

【0035】

ステップ S2 と同様に、本実施形態において、参照姿勢データセット D2 は、マトリックスで表現され、対応の加工されたテンプレート画像 $D1''$ に示される参照物体の姿勢（

すなわち、撮影視点から見た対応の加工されたテンプレート画像 D 1 ' ' に対応する偏向角度での特定の種類の物体の姿勢)を示す。具体的には、参照姿勢データセット D 2 のそれぞれは、3次元空間における6自由度を用いて、対応の加工されたテンプレート画像 D 1 ' ' に示される参照物体の姿勢を示す。加工されたテンプレート画像 D 1 ' ' が最初のテンプレート画像 D 1 ' を所定の角度で複数回回転することによって得られるため、加工されたテンプレート画像 D 1 ' ' にそれぞれ対応する参照姿勢データセット D 2 は、最初のテンプレート画像 D 1 ' に対応する参照姿勢データセット D 2 に基づいて計算して得ることができる。

【0036】

ステップ S 2 と同様に、本実施形態において、参照特徴データセット D 3 は、複数の成分(例えば、2304の個成分)を含むベクトルで表現され、成分を用いて対応の加工されたテンプレート画像 D 1 ' ' に示される参照角度での参照物体(特定の種類の物体)の外観特徴を示す。

10

【0037】

外観特徴は、撮影視点から見て、特徴の種類に属する物体が該参照角度から対応の加工されたテンプレート画像 D 1 ' ' (例えば、第1の加工されたテンプレート画像 D 1 ' ') に対応する偏向角度(例えば、10度)ずれている場合のみに存在する。

【0038】

例えば、外観特徴は、対応の加工されたテンプレート画像 D 1 ' ' に示される参照特徴部分それぞれの座標セットおよび/または方向付けであり、外観特徴は、撮影視点から見た参照物体が参照角度から偏向角度(例えば、10度)ずれている際の参照特徴部分の間の相対位置関係を示すことができるが、これに限定されない。

20

【0039】

さらに、本実施形態において、処理ユニット 1 1 は、加工されたテンプレート画像 D 1 ' ' のそれぞれを予めトレーニングされた畳み込みニューラルネットワーク(CNN)に入力して、出力となる対応の参照特徴データセット D 3 を得る。すなわち、参照特徴データセット D 3 のそれぞれは、処理ユニット 1 1 が畳み込みニューラルネットワーク(CNN)を用いて取得した対応の加工されたテンプレート画像 D 1 ' ' の分析結果である。

【0040】

処理ユニット 1 1 が加工されたテンプレート画像 D 1 ' ' にそれぞれ対応する複数の参照姿勢データセット D 2 と加工されたテンプレート画像 D 1 ' ' にそれぞれ対応する複数の参照特徴データセット D 3 とを得た後、フローはステップ S 5 に進む。

30

【0041】

ステップ S 5 において、処理ユニット 1 1 は、特定の種類に関連するデータベース D B 1 を確立して、データベース D B 1 を記憶ユニット 1 2 に記憶する。具体的には、データベース D B 1 は、異なる偏向角度にそれぞれ対応するテンプレート画像 D 1 (すなわち、最初のテンプレート画像 D 1 ' および加工されたテンプレート画像 D 1 ' ') と、テンプレート画像 D 1 にそれぞれ対応する参照姿勢データセット D 2 と、テンプレート画像 D 1 にそれぞれ対応する参照特徴データセット D 3 と、を含む。

【0042】

処理ユニット 1 1 がデータベース D B 1 を記憶ユニット 1 2 に記憶した後、フローはステップ S 6 に進む。

40

【0043】

ステップ S 6 において、処理ユニット 1 1 は、特定の種類に属する目標物体を含む入力画像を得ると、入力画像における目標物体の外観に基づいて、データベース D B 1 のテンプレート画像 D 1 から入力画像と最もマッチするマッチング画像 D 1 * を選択する。具体的には、マッチング画像 D 1 * は、テンプレート画像 D 1 のうちの示される参照物体の表示角度が入力画像に示される目標物体の角度(以下、「目標角度」と称する)と最も近い1つである。

【0044】

より具体的に説明すると、入力画像は、例えば、リアルタイム動画のフレームであり、

50

処理ユニット 11 は、例えば、処理ユニット 11 に電氣的に接続する撮影装置 2 からリアルタイムで動画を受信することによって、入力画像を得てもよい。目標物体は、本発明に係る方法を用いて計算装置 1 によって認識される実際のターゲットであり、特定の種類に属して参照物体と實質的に同様の外観を有する物体である。目標物体は、参照物体自体であってよい。

【0045】

本実施形態において、処理ユニット 11 は、以下の手段により、マッチング画像 D1* を選択する。まず、処理ユニット 11 は、例えば、入力画像に対して画像認識を実行することによって、入力画像に示される目標物体の複数の特徴部分（以下、「目標特徴部分」と称する）を認識する。具体的には、処理ユニット 11 は、最初のテンプレート画像 D1 における特徴マーカによって参照特徴部分に基づいて、入力画像に示される目標特徴部分を認識してもよい。したがって、テンプレート画像 D1 のそれぞれに対して、入力画像に示される目標特徴部分は、該テンプレート画像 D1 に示され、該テンプレート画像 D1 の特徴マーカによって定義される参照特徴部分にそれぞれ対応するべきである。

10

【0046】

入力画像に示される目標特徴部分を認識した後、処理ユニット 11 は、入力画像に示される目標特徴部分に基づいて、入力画像に対応する目標特徴データセットを生成する。参照特徴データセット D3 と同様に、本実施形態の目標特徴データセットは、複数の成分（例えば、2304 個の成分）を含むベクトルで表現され、成分を用いて入力画像に示される目標角度での目標物体の外観特徴を示す。外観特徴は、撮影視点から見て、目標物体が目標角度にある場合のみに存在する。

20

【0047】

目標特徴データセットが生成された後、処理ユニット 11 は、参照特徴データセット D3 のそれぞれについて、目標特徴データセットと該参照特徴データセット D3 との間のマッチング度を計算する。具体的には、本実施形態の目標特徴データセットおよび参照特徴データセット D3 がそれぞれベクトルで表現されるため、処理ユニット 11 は、参照特徴データセット D3 のそれぞれについて、目標特徴データセットと該参照特徴データセット D3 との間のミンコフスキー距離を計算し、計算されたミンコフスキー距離をマッチング程度とするが、本発明はこれに限定されない。なお、目標特徴データセットを表すベクトルと参照特徴データセット D3 のいずれかが 1 つを表すベクトルとの間のミンコフスキー距離が小さいほど、該 2 つのベクトルが互いに近く、入力画像に示される目標角度と対応のテンプレート画像 D1 に示される参照物体の表示角度とが近いことを意味する。すなわち、ミンコフスキー距離が小さいほど、マッチング程度が高いことを意味する。

30

【0048】

目標特徴データセットと各参照特徴データセット D3 との間のマッチング度を計算した後、処理ユニット 11 は、テンプレート画像 D1 のうちの目標特徴データセットとのマッチング度が最も高い参照特徴データセット D3 に対応する 1 つを、マッチング画像 D1* として選択する。このように、入力画像に示される目標角度が参照角度から 21.5 度ずれていると仮定すると、処理ユニット 11 は、20 度の偏向角度に対応する 1 つのテンプレート画像 D1 をマッチング画像 D1* として選択する（すなわち、選択された該 1 つのテンプレート画像 D1 は、入力画像の目標角度と最も近い参照物体の表示角度を有する）。

40

【0049】

処理ユニット 11 がテンプレート画像 D1 からマッチング画像 D1* を選択した後、フローはステップ S7 に進む。

【0050】

ステップ S7 において、処理ユニット 11 は、入力画像とマッチング画像 D1* とに基づいて、キーポイントマッチング工程を実行することによって、入力画像に示されかつ目標物体の外観に関連する複数の第 1 の特徴点と、マッチング画像 D1* に示されかつ第 1 の特徴点とそれぞれマッチする複数の第 2 の特徴点と、を認識する。具体的には、第 1 の

50

特徴点のそれぞれは、第2の特徴点の1つとは一対一の関係で同質(homogeneous)である。すなわち、第1の特徴点のそれぞれは第2の特徴点の1つと同質であり、その逆も同様である。本実施形態において、処理ユニット11は、従来技術のSe2-LoFTRを用いてキーポイントマッチング工程を実行する。いくつかの実施形態において、処理ユニット11は、Coarse-LoFTR、LoFTR、SuperGlueなどの従来技術を用いて、キーポイントマッチング工程を実行してもよく、本実施形態に限定されない。

【0051】

処理ユニット11が第1の特徴点と第2の特徴点とを認識した後、フローはステップS8へ進む。

10

【0052】

ステップS8において、処理ユニット11は、第1の特徴点と第2の特徴点とに基づいて、校正データセットを生成する。校正データセットは、マトリックスで表現され、第1の特徴点と第2の特徴点との間の相対姿勢関係を示す。具体的には、校正データセットは、第1の特徴点のそれぞれについて、3次元空間における6自由度を用いて、該第1の特徴点と対応の第2の特徴点との間の相対姿勢関係を示すことによって、マッチング画像D1*に示される参照物体の姿勢に対する入力画像に示される目標物体の姿勢を6自由度(すなわち、位置および方向付けの違い)で示す。

【0053】

処理ユニット11が校正データセットを生成した後、フローはステップS9に進む。

20

【0054】

ステップS9において、マッチング画像D1*に対応する参照姿勢データセットD2と校正データセット(該第1の特徴点と対応の第2の特徴点との間の相対姿勢関係を示す)とに基づいて、目標物体の姿勢を示す姿勢認識結果を生成する。具体的には、姿勢認識結果は、6自由度を用いて、入力画像の目標物体の姿勢と参照姿勢との間の違いを示すことによって、最初のテンプレート画像D1'に示される参照物体の姿勢に対する入力画像に示される目標物体の姿勢を6自由度で示す。より具体的には、本実施形態のマッチング画像D1*に対応する参照姿勢データセットD2と校正データセットとは、それぞれマトリックスで表現される。マッチング画像D1*に対応する参照姿勢データセットD2が第1のマトリックス(M_1 で表記)であり、校正データセットが第2のマトリックス(M_2 で表記)であると仮定すると、本実施形態の姿勢認識結果は、第2のマトリックスに第1のマトリックスを乗算することによって生成される(すなわち、 $M_2 M_1$)。以上、フローは完了する。

30

【0055】

本実施形態において、姿勢認識結果は、従来技術の座標系変換方法により、6自由度で他の座標系(例えば、ロボットアームの座標系)における位置および方向付けに変換されてもよい。さらに、姿勢認識結果が生成された後、処理ユニット11は、姿勢認識結果に基づいて目標物体に対して自動化操作をさらに実行してもよい。例えば、自動化操作は、ロボットアーム(図示せず)で目標物体を掴んで移動すること、または目標物体に対して加工する(例えば、溶接、接着、切削など)ことなどを含んでもよい。いくつかの実施形態において、姿勢認識結果が生成された後、処理ユニット11は、目標物体に対して自動化操作を実行する自動化システムに、姿勢認識結果を送信してもよく、これによって、自動化システムは、姿勢認識結果に基づいて、目標物体に対して動作する。なお、姿勢認識結果は様々な用途があるため、本実施形態は姿勢認識結果が生成された後の動作について制限しない。

40

【0056】

なお、ステップS1からステップS9および図2に示されるフローチャートは、本発明の一例を示すものに過ぎなく、本発明の方法と実質的に同じ方法で実質的に同じ機能を実現することができれば、ステップS1からステップS9を組み合わせたか、分割したり、順序を入れ替えたりしてもよい。例えば、処理ユニット11は、最初のテンプレート画像

50

D 1 ' を得た（すなわち、ステップ S 1 ）後、まず最初のテンプレート画像 D 1 ' に基づいて加工されたテンプレート画像 D 1 ' ' を生成（すなわち、ステップ S 3 ）してから、テンプレート画像 D 1 にそれぞれ対応する参照姿勢データセット D 2 およびテンプレート画像 D 1 にそれぞれ対応する参照特徴データセット D 3 を生成し（すなわち、ステップ S 2 およびステップ S 4 ）てもよい。したがって、本発明は、図 2 のフローチャートに限定されない。

【 0 0 5 7 】

要約すると、目標物体の姿勢を認識する方法を実行することによって、計算装置 1 は、目標物体が属する特定の種類に属する参照物体を含む 1 つのテンプレート画像 D 1 （すなわち、最初のテンプレート画像 D 1 ' ）のみを用いて、最初のテンプレート画像 D 1 ' を回転することによって加工されたテンプレート画像 D 1 ' ' を生成し、テンプレート画像 D 1 に対応する参照姿勢データセット D 2 および参照特徴データセット D 3 を生成することによって、特定の種類に関連するデータベース D B 1 を確立する。さらに、データベース D B 1 が確立された後、計算装置 1 が目標物体を含む入力画像を得ると、計算装置 1 は、データベース D B 1 の参照特徴データセット D 3 に基づいて入力画像と最もマッチするマッチング画像 D 1 * を選択し、そして入力画像とマッチング画像 D 1 * とに基づいてキーポイントマッチング工程を実行して、校正データセットを生成し、最後に、校正データセットとマッチング画像 D 1 * に対応する参照姿勢データセット D 2 とに基づいて、目標物体の姿勢を認識する。

【 0 0 5 8 】

なお、人工ニューラルネットワークを用いても目標物体の姿勢を認識することができるが、人工ニューラルネットワークは、目標物体の外観について予め学習する必要があるため、人工ニューラルネットワークの学習に異なる角度から撮られた複数の目標物体の画像を必要とする。人工ニューラルネットワークを用いる場合と比較して、本発明の計算装置 1 は、異なる角度から撮られた目標物体の画像を必要することなく、目標物体が属する特定の種類に属する参照物体を示す 1 つのテンプレート画像 D 1 （すなわち、最初のテンプレート画像 D 1 ' ）のみで、特定の種類に関連するデータベース D B 1 を確立することができる。さらに、処理ユニット 1 1 が最初のテンプレート画像 D 1 ' に基づいてデータベース D B 1 の生成に要する時間は、人工ニューラルネットワークの学習に要する時間よりはるかに短い。したがって、人工ニューラルネットワークを使用する場合と比較して、本発明は、人力と時間とを節約することができ、特に、認識する対象が変更される（例えば、目標物体が他の特定の種類に属する物体に変更される）場合において、人力と時間の節約はより顕著になる。

【 0 0 5 9 】

上記の説明では、説明の目的のために、実施形態の完全な理解を提供するために多数の特定の詳細が述べられた。しかしながら、当業者であれば、一又はそれ以上の他の実施形態が具体的な詳細を示さなくとも実施され得ることが明らかである。また、本明細書における「一実施形態」「一つの実施形態」を示す説明において、序数などの表示を伴う説明は全て、特定の態様、構造、特徴を有する本発明の具体的な実施に含まれ得るものであることと理解されたい。更に、本明細書において、時には複数の変化例が一つの実施形態、図面、又はこれらの説明に組み込まれているが、これは本明細書を合理化させるためのもので、本発明の多面性が理解されることを目的としたものであり、また、一実施形態における一又はそれ以上の特徴あるいは特定の具体例は、適切な場合には、本発明の実施において、他の実施形態における一またはそれ以上の特徴あるいは特定の具体例と共に実施され得る。

【 0 0 6 0 】

以上、本発明の実施形態および変化例を説明したが、本発明はこれらに限定されるものではなく、最も広い解釈の精神および範囲内に含まれる様々な構成として、全ての修飾および均等な構成を包含するものとする。

【 符号の説明 】

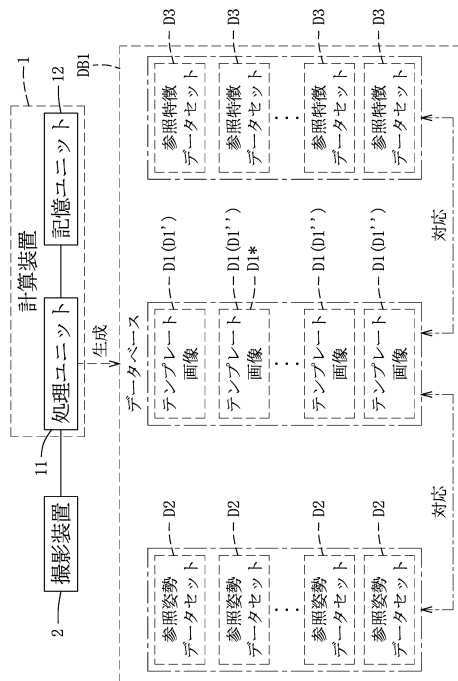
【 0 0 6 1 】

- 1 計算装置
- 1 1 処理ユニット
- 1 2 記憶ユニット
- 2 撮影装置
- D B 1 データベース
- D 1 テンプレート画像
- D 1 ' 最初のテンプレート画像
- D 1 ' ' 加工されたテンプレート画像
- D 1 * マッチング画像
- D 2 参照姿勢データセット
- D 3 参照特徴データセット
- S 1 - S 9 ステップ

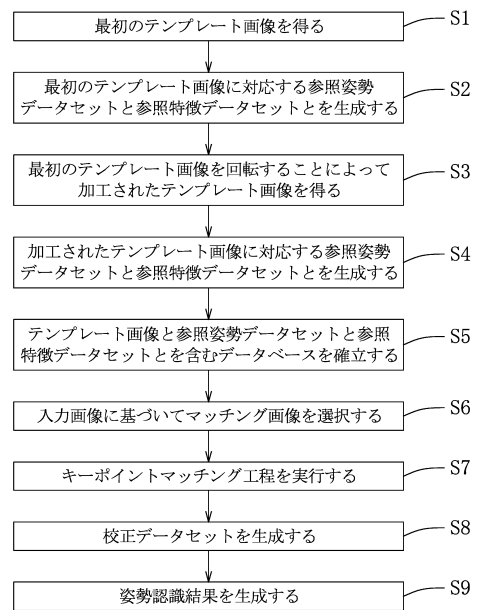
10

【 図 面 】

【 図 1 】



【 図 2 】



20

30

40

50

フロントページの続き

(72)発明者 ツァイ チエ

台湾臺北市 内 湖區行忠路42號6樓

審査官 鈴木 圭一郎

(56)参考文献 特開2007-219765(JP, A)

Yuan Liu, Gen6D: Generalizable Model-Free 6-DoF Object Pose Estimation from RGB Images, Computer Vision and Pattern Recognition (cs.CV), 2023年01月27日, p1-30, <https://doi.org/10.48550/arXiv.2204.10776>

石橋 研二 Kenji Ishibashi, データエンジニアリングIII, 技報 UNISYS TECHNOLOGY REVIEW Vol. 32 No. 4, 日本ユニシス株式会社, 第32巻
Sunyawit Sassanapitak, An efficient translation-rotation template matching using pre-computed scores of rotated templates, 2009 6th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology, IEEE, 2009年03月06日

(58)調査した分野 (Int.Cl., DB名)

G 0 6 T 7 / 7 3

G 0 6 T 7 / 0 0