

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5380789号
(P5380789)

(45) 発行日 平成26年1月8日(2014.1.8)

(24) 登録日 平成25年10月11日(2013.10.11)

(51) Int. Cl. F I
GO 1 B 11/00 (2006.01) GO 1 B 11/00 H
GO 6 T 1/00 (2006.01) GO 6 T 1/00 3 1 5

請求項の数 19 (全 37 頁)

(21) 出願番号	特願2007-150765 (P2007-150765)	(73) 特許権者	000002185
(22) 出願日	平成19年6月6日(2007.6.6)		ソニー株式会社
(65) 公開番号	特開2008-304268 (P2008-304268A)		東京都港区港南1丁目7番1号
(43) 公開日	平成20年12月18日(2008.12.18)	(74) 代理人	100093241
審査請求日	平成22年2月26日(2010.2.26)		弁理士 官田 正昭
		(74) 代理人	100101801
			弁理士 山田 英治
		(74) 代理人	100086531
			弁理士 澤田 俊夫
		(74) 代理人	100095496
			弁理士 佐々木 榮二
		(72) 発明者	多井 堅一郎
			東京都港区港南1丁目7番1号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 情報処理装置、および情報処理方法、並びにコンピュータ・プログラム

(57) 【特許請求の範囲】

【請求項1】

環境マップの生成処理を実行する情報処理装置であり、
 画像を撮影するカメラと、
 前記カメラの取得した画像に基づいてカメラの位置および姿勢を検出する自己位置検出部と、
 前記カメラの取得した画像からオブジェクトを検出する画像認識処理部と、
 前記自己位置検出部の検出したカメラの位置姿勢情報と、前記画像認識処理部の検出したオブジェクト情報を入力して、環境マップの生成または更新処理を実行するデータ構築部と、
 オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを格納した辞書データ記憶部を有し、
 前記画像認識処理部は、
 前記辞書データを参照して前記カメラの取得した画像からオブジェクトを検出する処理を実行し、
 前記データ構築部は、
 前記辞書データに登録された三次元形状データを適用して環境マップへのオブジェクト配置を実行する構成であることを特徴とする情報処理装置。

【請求項2】

前記画像認識処理部は、

前記カメラの取得した画像に含まれるオブジェクトの特徴点の位置を識別し、前記データ構築部へ出力する処理を実行する構成であり、

前記データ構築部は、

前記画像認識処理部から入力する特徴点情報と、前記自己位置検出部から入力するカメラの位置姿勢情報とに基づいて、オブジェクトの世界座標系における位置および姿勢を算出して環境マップに登録する処理を実行する構成であることを特徴とする請求項1に記載の情報処理装置。

【請求項3】

前記自己位置検出部は、

前記カメラから入力する画像内の特徴点を利用して世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラ位置(Cw)、カメラの回転行列(Rw)を算出して前記データ構築部へ出力する処理を実行する構成であることを特徴とする請求項1に記載の情報処理装置。

【請求項4】

前記自己位置検出部は、

前記カメラから入力する画像内の特徴点の位置と、カメラの位置姿勢を併せて検出するSLAM(simultaneous localization and mapping)を実行する構成であることを特徴とする請求項1に記載の情報処理装置。

【請求項5】

前記情報処理装置は、さらに、

前記画像認識処理部におけるオブジェクト検出対象フレームより以前の画像フレーム中のオブジェクト中の特徴点位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換する座標変換部を有し、

前記画像認識処理部は、

前記座標変換部の提供する特徴点情報を前記データ構築部へ出力する処理を実行する構成であることを特徴とする請求項1に記載の情報処理装置。

【請求項6】

前記自己位置検出部は、

前記カメラから入力する画像内の特徴点の位置と、カメラの位置姿勢を併せて検出するSLAM(simultaneous localization and mapping)を実行する構成であり、

前記情報処理装置は、さらに、

前記画像認識処理部において検出した第1の特徴点と、前記自己位置検出部が前記SLAMの実行により検出した第2の特徴点の位置を比較し、各特徴点の位置が予め定めた閾値以内の距離にある場合に、前記第2の特徴点を特徴点データベースに登録する特徴点照合部を有し、

前記座標変換部は、前記特徴点データベースに登録された特徴点の位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換して画像認識処理部において利用可能とする処理を行なう構成であることを特徴とする請求項5に記載の情報処理装置。

【請求項7】

前記データ構築部は、

生成済みの環境マップを格納した環境マップデータベースと、

前記環境マップデータベースから環境マップを取得する環境情報取得部と、

前記環境情報取得部の取得した環境マップと、前記画像認識処理部から入力するオブジェクト検出情報とを比較し、比較結果を環境マップ更新部へ出力する認識結果比較部と、

前記認識結果比較部から入力する比較結果に基づいて、前記環境マップデータベースに格納された環境マップの更新処理を実行する環境マップ更新部を有する構成であることを特徴とする請求項1に記載の情報処理装置。

【請求項8】

10

20

30

40

50

前記環境情報取得部は、
 前記環境マップデータベースから環境マップを取得し、取得した環境マップに含まれる特徴点の位置を含む特徴点情報を生成する特徴点情報生成部を有し、
 前記認識結果比較部は、
 前記画像認識処理部から入力するオブジェクト対応の特徴点情報と、前記特徴点情報生成部の生成した特徴点情報を比較して、比較情報を前記環境マップ更新部内の画像内変更領域抽出部に出力する特徴点情報比較部を有し、
 前記環境マップ更新部は、
 前記特徴点情報比較部から前記比較情報を入力し、マッチした特徴点の距離が予め設定した閾値より近い領域以外を環境マップの更新領域として抽出する画像内変更領域抽出部と、
 前記環境マップデータベースに格納された環境マップに対して、前記画像内変更領域抽出部の抽出した更新領域のみを更新対象として、前記画像認識処理部から入力する特徴点情報を用いた更新処理を実行する環境マップ更新部と、
 を有する構成であることを特徴とする請求項 7 に記載の情報処理装置。

10

【請求項 9】

前記環境情報取得部は、
 前記環境マップデータベースから環境マップを取得し、取得した環境マップに含まれるオブジェクトの重心・姿勢の存在確率分布を抽出する存在確率分布抽出部を有し、
 前記認識結果比較部は、
 前記画像認識処理部から入力するオブジェクト認識結果を入力して、認識オブジェクトの存在確率分布を生成する存在確率分布生成部と、
 前記存在確率分布抽出部が環境マップに基づいて生成した認識オブジェクトの存在確率分布と、前記存在確率分布生成部が前記画像認識処理部の認識結果に基づいて生成したオブジェクトの存在確率分布との比較を行い、比較情報を前記環境マップ更新部に出力する確率分布比較部を有し、
 前記環境マップ更新部は、
 前記存在確率分布比較部から入力する比較情報に基づいて環境マップの更新領域を決定し、前記環境マップデータベースに格納された環境マップに対して、前記更新領域のみを更新対象として、前記画像認識処理部から入力する特徴点情報を用いた更新処理を実行する構成であることを特徴とする請求項 7 に記載の情報処理装置。

20

30

【請求項 10】

前記確率分布比較部は、
 前記存在確率分布抽出部が環境マップに基づいて生成した認識オブジェクトの存在確率分布と、前記存在確率分布生成部が前記画像認識処理部の認識結果に基づいて生成したオブジェクトの存在確率分布との差分を示すマハラノビス距離 (s) を算出して前記環境マップ更新部に出力する構成であり、
 前記環境マップ更新部は、
 前記マハラノビス距離 (s) が、予め設定した閾値より大きい場合に、前記環境マップデータベースに登録された環境マップを更新する処理を実行する構成であることを特徴とする請求項 9 に記載の情報処理装置。

40

【請求項 11】

情報処理装置において、環境マップの生成処理を実行する情報処理方法であり、
 カメラが画像を撮影する画像撮影ステップと、
 自己位置検出部が、前記カメラの取得した画像に基づいてカメラの位置および姿勢を検出する自己位置検出ステップと、
 画像認識処理部が、前記カメラの取得した画像からオブジェクトを検出する画像認識処理ステップと、
 データ構築部が、前記自己位置検出部の検出したカメラの位置姿勢情報と、前記画像認識処理部の検出したオブジェクト情報を入力して、環境マップの生成または更新処理を実

50

行するデータ構築ステップとを有し、

前記画像認識処理ステップは、

オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを参照して前記カメラの取得した画像からオブジェクトを検出する処理を実行し、

前記データ構築ステップは、

前記辞書データに登録された三次元形状データを適用して環境マップへのオブジェクト配置を実行することを特徴とする情報処理方法。

【請求項 1 2】

前記画像認識処理ステップは、

前記カメラの取得した画像に含まれるオブジェクトの特徴点の位置を識別し、前記データ構築部へ出力する処理を実行するステップであり、

前記データ構築ステップは、

前記画像認識処理部から入力する特徴点情報と、前記自己位置検出部から入力するカメラの位置姿勢情報とに基づいて、オブジェクトの世界座標系における位置および姿勢を算出して環境マップに登録する処理を実行するステップであることを特徴とする請求項 1 1 に記載の情報処理方法。

【請求項 1 3】

前記自己位置検出ステップは、

前記カメラから入力する画像内の特徴点を利用して世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラ位置 (Cw)、カメラの回転行列 (Rw) を算出して前記データ構築部へ出力する処理を実行するステップであることを特徴とする請求項 1 1 に記載の情報処理方法。

【請求項 1 4】

前記自己位置検出ステップは、

前記カメラから入力する画像内の特徴点の位置と、カメラの位置姿勢を併せて検出する SLAM (simultaneous localization and mapping) を実行するステップであることを特徴とする請求項 1 1 に記載の情報処理方法。

【請求項 1 5】

前記情報処理方法は、さらに、

座標変換部が、前記画像認識処理部におけるオブジェクト検出対象フレームより以前の画像フレーム中のオブジェクト中の特徴点位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換する座標変換ステップを有し、

前記画像認識処理ステップは、

前記座標変換部の提供する特徴点情報を前記データ構築部へ出力する処理を実行するステップであることを特徴とする請求項 1 1 に記載の情報処理方法。

【請求項 1 6】

前記自己位置検出ステップは、

前記カメラから入力する画像内の特徴点の位置と、カメラの位置姿勢を併せて検出する SLAM (simultaneous localization and mapping) を実行するステップであり、

前記情報処理方法は、さらに、

特徴点照合部が、前記画像認識処理部において検出した第 1 の特徴点と、前記自己位置検出部が前記 SLAM の実行により検出した第 2 の特徴点の位置を比較し、各特徴点の位置が予め定めた閾値以内の距離にある場合に、前記第 2 の特徴点を特徴点データベースに記録する特徴点照合ステップを有し、

前記座標変換ステップは、前記特徴点データベースに記録された特徴点の位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換して画像認識処理部において利用可能とする処理を行なうステップであることを特徴とする請求項 1 5 に記載の情報処理方法。

10

20

30

40

50

【請求項 17】

前記データ構築ステップは、

環境情報取得部が、生成済みの環境マップを環境マップデータベースから取得する環境情報取得ステップと、

認識結果比較部が、環境マップデータベースから取得した環境マップと、前記画像認識処理部から入力するオブジェクト検出情報とを比較し、比較結果を環境マップ更新部に出力する認識結果比較ステップと、

環境マップ更新部が、前記認識結果比較部から入力する比較結果に基づいて、前記環境マップデータベースに格納された環境マップの更新処理を実行する環境マップ更新ステップを有することを特徴とする請求項 11 に記載の情報処理方法。

10

【請求項 18】

前記環境情報取得ステップは、

特徴点情報生成部が、前記環境マップデータベースから環境マップを取得し、取得した環境マップに含まれる特徴点の位置を含む特徴点情報を生成する特徴点情報生成ステップを有し、

前記認識結果比較ステップは、

特徴点情報比較部が、前記画像認識処理部から入力するオブジェクト対応の特徴点情報と、前記特徴点情報生成部の生成した特徴点情報を比較して、比較情報を前記環境マップ更新部内の画像内変更領域抽出部に出力する特徴点情報比較ステップを有し、

前記環境マップ更新ステップは、

画像内変更領域抽出部が、前記特徴点情報比較部から前記比較情報を入力し、マッチした特徴点の距離が予め設定した閾値より近い領域以外を環境マップの更新領域として抽出する画像内変更領域抽出ステップと、

20

環境マップ更新部が、前記環境マップデータベースに格納された環境マップに対して、前記画像内変更領域抽出部の抽出した更新領域のみを更新対象として、前記画像認識処理部から入力する特徴点情報を用いた更新処理を実行する環境マップ更新ステップと、

を有することを特徴とする請求項 17 に記載の情報処理方法。

【請求項 19】

情報処理装置において、環境マップの生成処理を実行させるコンピュータ・プログラムであり、

30

カメラに画像を撮影させる画像撮影ステップと、

自己位置検出部に、前記カメラの取得した画像に基づいてカメラの位置および姿勢を検出させる自己位置検出ステップと、

画像認識処理部に、前記カメラの取得した画像からオブジェクトを検出させる画像認識処理ステップと、

データ構築部に、前記自己位置検出部の検出したカメラの位置姿勢情報と、前記画像認識処理部の検出したオブジェクト情報を入力して、環境マップの生成または更新処理を実行させるデータ構築ステップとを有し、

前記画像認識処理ステップは、

オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを参照して前記カメラの取得した画像からオブジェクトを検出するステップであり、

40

前記データ構築ステップは、

前記辞書データに登録された三次元形状データを適用して環境マップへのオブジェクト配置を実行するステップであることを特徴とするコンピュータ・プログラム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、情報処理装置、および情報処理方法、並びにコンピュータ・プログラムに関する。さらに詳細には、カメラによって撮影された画像に基づいてカメラ周囲の地図（環

50

境地図)の作成(mapping)、すなわち環境マップ作成処理を実行する情報処理装置、および情報処理方法、並びにコンピュータ・プログラムに関する。

【0002】

本発明は、例えばカメラを備えたロボットなどのエージェント(移動体)において移動環境を観測し、観測状況に応じてエージェント周囲の地図(環境地図)の作成(mapping)、すなわち環境マップ作成処理を実行し、さらに環境マップ作成処理に併せてエージェントの位置や姿勢の推定、すなわち自己位置もしくは自己姿勢の同定処理(ローカリゼーション)を実行する情報処理装置、および情報処理方法、並びにコンピュータ・プログラムに関する。

【背景技術】

【0003】

例えばカメラを備えたロボットなどのエージェント(移動体)において移動環境を観測し、観測状況に応じてエージェント周囲の地図(環境地図)を作成する環境マップ構築処理は、車やロボットといった移動体の経路探索のために行われているものが多い。特許文献1(特開2003-269937)には、複数のカメラによって撮影された画像から生成される距離画像(ステレオ画像)から平面を検出し、検出した平面群と撮像デバイスの位置・姿勢から床面を求め、最後に床面から障害物を認識する技術が開示されている。しかし、この特許文献1において実行する周囲環境の解析では、環境は「床」か「障害物」としか区別できない。

【0004】

また上記手法では、複数のカメラの撮影画像を用いて距離画像を生成することが必要となる。このような距離画像を生成することなく1つのカメラ撮影画像の画像解析処理に基づいて環境マップを作成する技術についても盛んに開発されている。多くは単眼で取得した1枚の画像のみから様々な情報を認識するシステムである。しかし、これらの処理の多くは、例えばカメラの位置を原点とした座標系を世界座標系として周囲の物体の位置を計算する構成となっている。すなわちカメラを備えたロボットが周囲の物体に衝突しないように走行可能とするなどを主目的とするものであり、カメラによって撮影された画像に含まれる様々な認識対象、例えば「壁」、「テーブル」、「ソファ」などの認識結果物についての詳細な解析、具体的には各認識対象物の三次元形状の解析などを行なうものではなく、自立走行のための環境マップに過ぎないものであった。

【特許文献7】特開2003-269937号公報

【発明の開示】

【発明が解決しようとする課題】

【0005】

本発明は、上述の問題点に鑑みてなされたものであり、カメラによって撮影された画像に基づく環境マップの作成を行なうものであり、撮影画像に含まれる様々なオブジェクトの解析を実行して、より詳細な情報を含む環境マップの作成を行なう情報処理装置、および情報処理方法、並びにコンピュータ・プログラム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】

【0006】

本発明の第1の側面は、
環境マップの生成処理を実行する情報処理装置であり、
画像を撮影するカメラと、
前記カメラの取得した画像に基づいてカメラの位置および姿勢を検出する自己位置検出部と、
前記カメラの取得した画像からオブジェクトを検出する画像認識処理部と、
前記自己位置検出部の検出したカメラの位置姿勢情報と、前記画像認識処理部の検出したオブジェクト情報を入力して、環境マップの生成または更新処理を実行するデータ構築部と、
オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録し

10

20

30

40

50

た辞書データを格納した辞書データ記憶部を有し、

前記画像認識処理部は、

前記辞書データを参照して前記カメラの取得した画像からオブジェクトを検出する処理を実行し、

前記データ構築部は、

前記辞書データに登録された三次元形状データを適用して環境マップへのオブジェクト配置を実行する構成であることを特徴とする情報処理装置にある。

【0007】

さらに、本発明の情報処理装置の一実施態様において、前記画像認識処理部は、前記カメラの取得した画像に含まれるオブジェクトの特徴点の位置を識別し、前記データ構築部に出力する処理を実行する構成であり、前記データ構築部は、前記画像認識処理部から入力する特徴点情報と、前記自己位置検出部から入力するカメラの位置姿勢情報とに基づいて、オブジェクトの世界座標系における位置および姿勢を算出して環境マップに登録する処理を実行する構成であることを特徴とする。

10

【0008】

さらに、本発明の情報処理装置の一実施態様において、前記自己位置検出部は、前記カメラから入力する画像内の特徴点を利用して世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラ位置(Cw)、カメラの回転行列(Rw)を算出して前記データ構築部に出力する処理を実行する構成であることを特徴とする。

【0009】

さらに、本発明の情報処理装置の一実施態様において、前記自己位置検出部は、前記カメラから入力する画像内の特徴点の位置と、カメラの位置姿勢を併せて検出するSLAM(simultaneous localization and mapping)を実行する構成であることを特徴とする。

20

【0010】

さらに、本発明の情報処理装置の一実施態様において、前記情報処理装置は、さらに、前記画像認識処理部におけるオブジェクト検出対象フレームより以前の画像フレーム中のオブジェクト中の特徴点位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換する座標変換部を有し、前記画像認識処理部は、前記座標変換部の提供する特徴点情報を前記データ構築部に出力する処理を実行する構成であることを特徴とする。

30

【0011】

さらに、本発明の情報処理装置の一実施態様において、前記情報処理装置は、さらに、前記画像認識処理部において検出した特徴点と、前記自己位置検出部において検出した特徴点との位置が予め定めた閾値以内の距離にある場合に、この特徴点を特徴点データベースに登録する特徴点照合部を有し、前記座標変換部は、前記特徴点データベースに登録された特徴点の位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換して画像認識処理部において利用可能とする処理を行なう構成であることを特徴とする。

【0012】

さらに、本発明の情報処理装置の一実施態様において、前記データ構築部は、生成済みの環境マップを格納した環境マップデータベースと、前記環境マップデータベースから環境マップを取得する環境情報取得部と、前記環境情報取得部の取得した環境マップと、前記画像認識処理部から入力するオブジェクト検出情報とを比較し、比較結果を環境マップ更新部に出力する認識結果比較部と、前記認識結果比較部から入力する比較結果に基づいて、前記環境マップデータベースに格納された環境マップの更新処理を実行する環境マップ更新部を有する構成であることを特徴とする。

40

【0013】

さらに、本発明の情報処理装置の一実施態様において、前記環境情報取得部は、前記環境マップデータベースから環境マップを取得し、取得した環境マップに含まれる特徴点の

50

位置を含む特徴点情報を生成する特徴点情報生成部を有し、前記認識結果比較部は、前記画像認識処理部から入力するオブジェクト対応の特徴点情報と、前記特徴点情報生成部の生成した特徴点情報を比較して、比較情報を前記環境マップ更新部内の画像内変更領域抽出部へ出力する特徴点情報比較部を有し、前記環境マップ更新部は、前記特徴点情報比較部から前記比較情報を入力し、マッチした特徴点の距離が予め設定した閾値より近い領域以外を環境マップの更新領域として抽出する画像内変更領域抽出部と、前記環境マップデータベースに格納された環境マップに対して、前記画像内変更領域抽出部の抽出した更新領域のみを更新対象として、前記画像認識処理部から入力する特徴点情報を用いた更新処理を実行する環境マップ更新部と、を有する構成であることを特徴とする。

【0014】

さらに、本発明の情報処理装置の一実施態様において、前記環境情報取得部は、前記環境マップデータベースから環境マップを取得し、取得した環境マップに含まれるオブジェクトの重心・姿勢の存在確率分布を抽出する存在確率分布抽出部を有し、前記認識結果比較部は、前記画像認識処理部から入力するオブジェクト認識結果を入力して、認識オブジェクトの存在確率分布を生成する存在確率分布生成部と、前記存在確率分布抽出部が環境マップに基づいて生成した認識オブジェクトの存在確率分布と、前記存在確率分布生成部が前記画像認識処理部の認識結果に基づいて生成したオブジェクトの存在確率分布との比較を行い、比較情報を前記環境マップ更新部へ出力する確率分布比較部を有し、前記環境マップ更新部は、前記存在確率分布比較部から入力する比較情報に基づいて環境マップの更新領域を決定し、前記環境マップデータベースに格納された環境マップに対して、前記更新領域のみを更新対象として、前記画像認識処理部から入力する特徴点情報を用いた更新処理を実行する構成であることを特徴とする。

【0015】

さらに、本発明の情報処理装置の一実施態様において、前記確率分布比較部は、前記存在確率分布抽出部が環境マップに基づいて生成した認識オブジェクトの存在確率分布と、前記存在確率分布生成部が前記画像認識処理部の認識結果に基づいて生成したオブジェクトの存在確率分布との差分を示すマハラノビス距離 (s) を算出して前記環境マップ更新部へ出力する構成であり、前記環境マップ更新部は、前記マハラノビス距離 (s) が、予め設定した閾値より大きい場合に、前記環境マップデータベースに登録された環境マップを更新する処理を実行する構成であることを特徴とする。

【0016】

さらに、本発明の第2の側面は、
環境マップに基づいてオブジェクト検索範囲を特定する処理を実行する情報処理装置であり、
オブジェクトの隣接領域に特定オブジェクトが存在する可能性を示すオントロジーデータ (意味情報) を格納した記憶部と、
前記オントロジーデータ (意味情報) に基づいて、特定オブジェクトの探索領域を決定する画像認識処理部と、
を有することを特徴とする情報処理装置にある。

【0017】

さらに、本発明の第3の側面は、
情報処理装置において、環境マップの生成処理を実行する情報処理方法であり、
カメラが画像を撮影する画像撮影ステップと、
自己位置検出部が、前記カメラの取得した画像に基づいてカメラの位置および姿勢を検出する自己位置検出ステップと、
画像認識処理部が、前記カメラの取得した画像からオブジェクトを検出する画像認識処理ステップと、
データ構築部が、前記自己位置検出部の検出したカメラの位置姿勢情報と、前記画像認識処理部の検出したオブジェクト情報を入力して、環境マップの生成または更新処理を実行するデータ構築ステップとを有し、

10

20

30

40

50

前記画像認識処理ステップは、

オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを参照して前記カメラの取得した画像からオブジェクトを検出する処理を実行し、

前記データ構築ステップは、

前記辞書データに登録された三次元形状データを適用して環境マップへのオブジェクト配置を実行することを特徴とする情報処理方法にある。

【0018】

さらに、本発明の情報処理方法の一実施態様において、前記画像認識処理ステップは、前記カメラの取得した画像に含まれるオブジェクトの特徴点の位置を識別し、前記データ構築部 10 10 に出力する処理を実行するステップであり、前記データ構築ステップは、前記画像認識処理部から入力する特徴点情報と、前記自己位置検出部から入力するカメラの位置姿勢情報とに基づいて、オブジェクトの世界座標系における位置および姿勢を算出して環境マップに登録する処理を実行するステップであることを特徴とする。

【0019】

さらに、本発明の情報処理方法の一実施態様において、前記自己位置検出ステップは、前記カメラから入力する画像内の特徴点を利用して世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラ位置 (Cw)、カメラの回転行列 (Rw) を算出して前記データ構築部に出力する処理を実行するステップであることを特徴とする。

【0020】

さらに、本発明の情報処理方法の一実施態様において、前記自己位置検出ステップは、前記カメラから入力する画像内の特徴点の位置と、カメラの位置姿勢を併せて検出する SLAM (simultaneous localization and mapping) を実行するステップであることを特徴とする。

【0021】

さらに、本発明の情報処理方法の一実施態様において、前記情報処理方法は、さらに、座標変換部が、前記画像認識処理部におけるオブジェクト検出対象フレームより以前の画像フレーム中のオブジェクト中の特徴点位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換する座標変換ステップを有し、前記画像認識処理ステップは、前記座標変換部の提供する特徴点情報を前記データ構築部に出力する処理を実行するステップであることを特徴とする。

【0022】

さらに、本発明の情報処理方法の一実施態様において、前記情報処理方法は、さらに、特徴点照合部が、前記画像認識処理部において検出した特徴点と、前記自己位置検出部において検出した特徴点との位置が予め定めた閾値以内の距離にある場合に、この特徴点を特徴点データベースに登録する特徴点照合ステップを有し、

前記座標変換ステップは、前記特徴点データベースに登録された特徴点の位置を、前記画像認識処理部におけるオブジェクト検出対象フレームの画像対応の座標上に変換して画像認識処理部において利用可能とする処理を行なうステップであることを特徴とする。

【0023】

さらに、本発明の情報処理方法の一実施態様において、前記データ構築ステップは、環境情報取得部が、生成済みの環境マップを環境マップデータベースから取得する環境情報取得ステップと、認識結果比較部が、環境マップデータベースから取得した環境マップと、前記画像認識処理部から入力するオブジェクト検出情報とを比較し、比較結果を環境マップ更新部に出力する認識結果比較ステップと、環境マップ更新部が、前記認識結果比較部から入力する比較結果に基づいて、前記環境マップデータベースに格納された環境マップの更新処理を実行する環境マップ更新ステップを有することを特徴とする。

【0024】

さらに、本発明の情報処理方法の一実施態様において、前記環境情報取得ステップは、特徴点情報生成部が、前記環境マップデータベースから環境マップを取得し、取得した環 50

境マップに含まれる特徴点の位置を含む特徴点情報を生成する特徴点情報生成ステップを有し、前記認識結果比較ステップは、特徴点情報比較部が、前記画像認識処理部から入力するオブジェクト対応の特徴点情報と、前記特徴点情報生成部の生成した特徴点情報を比較して、比較情報を前記環境マップ更新部内の画像内変更領域抽出部へ出力する特徴点情報比較ステップを有し、前記環境マップ更新ステップは、画像内変更領域抽出部が、前記特徴点情報比較部から前記比較情報を入力し、マッチした特徴点の距離が予め設定した閾値より近い領域以外を環境マップの更新領域として抽出する画像内変更領域抽出ステップと、環境マップ更新部が、前記環境マップデータベースに格納された環境マップに対して、前記画像内変更領域抽出部の抽出した更新領域のみを更新対象として、前記画像認識処理部から入力する特徴点情報を用いた更新処理を実行する環境マップ更新ステップとを有することを特徴とする。

10

【0027】

さらに、本発明の第4の側面は、
情報処理装置において、環境マップの生成処理を実行させるコンピュータ・プログラムであり、

カメラに画像を撮影させる画像撮影ステップと、
自己位置検出部に、前記カメラの取得した画像に基づいてカメラの位置および姿勢を検出させる自己位置検出ステップと、

画像認識処理部に、前記カメラの取得した画像からオブジェクトを検出させる画像認識処理ステップと、

20

データ構築部に、前記自己位置検出部の検出したカメラの位置姿勢情報と、前記画像認識処理部の検出したオブジェクト情報を入力して、環境マップの生成または更新処理を実行させるデータ構築ステップとを有し、

前記画像認識処理ステップは、

オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを参照して前記カメラの取得した画像からオブジェクトを検出するステップであり、

前記データ構築ステップは、

前記辞書データに登録された三次元形状データを適用して環境マップへのオブジェクト配置を実行するステップであることを特徴とするコンピュータ・プログラムにある。

30

【0028】

なお、本発明のコンピュータ・プログラムは、例えば、様々なプログラム・コードを実行可能な汎用コンピュータ・システムに対して、コンピュータ可読な形式で提供する記憶媒体、通信媒体によって提供可能なコンピュータ・プログラムである。このようなプログラムをコンピュータ可読な形式で提供することにより、コンピュータ・システム上でプログラムに応じた処理が実現される。

【0029】

本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。なお、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

40

【発明の効果】**【0030】**

本発明の一実施例の構成によれば、カメラの取得した画像に基づいてカメラの位置および姿勢を検出する自己位置検出処理と、カメラの取得画像からオブジェクトを検出する画像認識処理と、カメラの位置姿勢情報とオブジェクト情報と、オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを適用して環境マップの生成または更新処理を実行する構成としたので、様々なオブジェクトの三次元データを反映した環境マップを、1つのカメラによって取得した画像に基づいて効率的に生成することが可能となる。

50

【発明を実施するための最良の形態】

【0031】

以下、図面を参照しながら本発明の実施形態に係る情報処理装置、および情報処理方法、並びにコンピュータ・プログラムの詳細について説明する。

【0032】

[実施例1]

本発明の一実施形態にかかる情報処理装置の構成について図1を参照して説明する。本発明の情報処理装置は、カメラによって撮影された画像データに基づいて環境マップを構築する情報処理装置であり、図1に示すように、周囲環境を撮影するカメラ101と、カメラ101の撮影画像を入力して画像認識を行う画像認識処理部102と、カメラ101の撮影画像を入力してカメラの位置姿勢を推定する自己位置検出部103と、画像認識処理部102において生成された画像認識結果データと、自己位置検出部103において検出されたカメラ101の位置および姿勢情報を入力して、ある世界座標系で表現された環境マップの作成処理を実行するデータ構築部110、さらに、画像認識処理部102における画像認識処理や、データ構築部110における環境マップ作成処理に利用する辞書データを格納した辞書データ記憶部104を有する。

【0033】

画像認識処理部102と自己位置検出部103には少なくとも入力としてカメラ映像を用いる。画像認識処理部102はカメラ101によって撮影される様々なオブジェクト(検出対象)の画像内の位置姿勢情報を含む認識結果をデータ構築部110に出力する。ただし、認識結果はカメラ座標系で表現されている。自己位置検出部103はカメラ101の位置・姿勢情報をデータ構築部110に出力する。データ構築部110では、画像認識処理部102から入力する認識結果と、自己位置検出部103から入力するカメラ位置姿勢情報に基づいて、世界座標系(環境マップの座標系)における様々なオブジェクトの位置姿勢を求め、環境マップを更新し、最新の環境マップを出力する。

【0034】

データ構築部110は、画像認識処理部102からの認識結果と、自己位置検出部103からのカメラ位置姿勢情報を入力して世界座標系(環境マップの座標系)における様々なオブジェクトの位置姿勢を求め環境マップを更新する環境マップ更新111と、更新された環境マップを格納する環境マップデータベース112を有する。データ構築部110の生成する環境マップは、カメラ101によって撮影された画像に含まれる様々なオブジェクト、例えば「テーブル」、「椅子」などの様々なオブジェクトに関する詳細な情報、具体的には三次元形状や位置、姿勢情報などの詳細情報を含むデータとして構成される。

【0035】

上記構成において、画像認識処理部102は、カメラ101の撮影画像を入力して画像認識を実行して、カメラ101によって撮影される様々なオブジェクト(検出対象)の画像内の位置姿勢情報を含む認識結果を生成して、認識結果をデータ構築部110に出力する。この画像認識処理部102において実行する画像認識処理は、例えば、特開2006-190191号公報や、特開2006-190192号公報に記載された画像認識処理をベースとした画像認識処理を実行する。具体的な処理例については後述する。

【0036】

また、自己位置検出部103は、文献[Andrew J. Davison, "Real-time simultaneous localisation and mapping with a single camera", Proceedings of the 9th International Conference on Computer Vision, Ninth, (2003)]に記載された構成をベースに開発したものであり、カメラ101によって撮影される画像内の局所領域(以下、特徴点)の位置のフレーム間の変化を元に、三次元空間の特徴点の位置とカメラ位置を同時に毎フレーム推定する処理を行なう。なお、本実施例の自己位置検出部103の入力はカメラ101の映像だけの設定としているが、他のセンサを利用して自己位置推定のロバスト性

10

20

30

40

50

を上げる構成としてもよい。

【 0 0 3 7 】

画像認識処理部 1 0 2 における画像認識処理や、データ構築部 1 1 0 における環境マップ作成処理に利用する辞書データを格納した辞書データ記憶部 1 0 4 の格納データについて図 2 を参照して説明する。

【 0 0 3 8 】

辞書データには、カメラ 1 0 1 によって撮影されると推定されるオブジェクト（認識対象）に関する情報が、オブジェクトごとにまとめて保存されている。1つのオブジェクト（認識対象）に対して、

- (1) オブジェクト名称 10
- (2) 特徴点情報（主要な視点からの二次元情報。詳しくは図 2 を参照）
- (3) 物体形状情報（特徴点の位置情報を含む）
- (4) オントロジー情報（オブジェクトの属するカテゴリ、接する可能性が高い対象、動体・静体などの情報を含む）

【 0 0 3 9 】

図 2 は、カップに関する辞書データの登録情報例を示している。

- (1) オブジェクト名称：「バラ柄マグカップ」
- (2) 特徴点情報：上下前後左右の 6 方向からの視点からオブジェクトを観察した場合の特徴点情報が記録される。なお、各視点からの画像情報としてもよい。
- (3) 物体形状情報：オブジェクトの三次元形状情報が記録される。特徴点の位置についても記録される。 20
- (4) オントロジー情報
 - (4 - 1) オブジェクトの属性：[食器]、
 - (4 - 2) オブジェクトに接する可能性の高いオブジェクト情報：[机] [食器洗い機]
 - (4 - 3) オブジェクトに接する可能性の低いオブジェクト情報：[本棚]
 これらの各情報が記録される。

辞書データ記憶部 1 0 4 には、様々なオブジェクトに対応する上記情報が記録されている。

【 0 0 4 0 】 30

以下、図 1 を参照して説明した本発明の情報処理装置において実行する環境マップの生成処理の具体例について説明する。

【 0 0 4 1 】

図 1 に示す情報処理装置のカメラ 1 0 1 の撮影した画像を用いて環境マップを構築する処理例について説明する。カメラ 1 0 1 はピンホールカメラモデルにあうものを利用した。ピンホールカメラモデルは、三次元空間の点の位置とカメラ像平面の画素位置との射影変換を表しており、以下の式で表される。

【 数 1 】

$$\lambda \tilde{\mathbf{m}} = \mathbf{AR}_w (\mathbf{M} - \mathbf{C}_w) \quad \dots \text{(数式 1)} \quad 40$$

【 0 0 4 2 】

上記式の意味について、図 3、図 4 を参照して説明する。上記式は、カメラの撮影画像 2 1 0 に含まれるオブジェクト 2 1 1 の点 (m) のカメラ像平面の画素位置 2 1 2、すなわち、カメラ座標系によって表現されている位置と、世界座標系におけるオブジェクト 2 0 0 の三次元位置 (M) 2 0 1 との対応関係を示す式である。

【 0 0 4 3 】

カメラ像平面の画素位置 2 1 2 はカメラ座標系によって表現されている。カメラ座費用 50

系は、カメラの焦点を原点 C として、像平面が X_c, Y_c の二次元平面、奥行きを Z_c とした座標系であり、カメラの動きによって原点 C は移動する。

【0044】

一方、オブジェクト 200 の三次元位置 (M) 201 は、カメラの動きによって移動しない原点 O を有する XYZ 三軸からなる世界座標系によって示される。この異なる座標系でのオブジェクトの位置の対応関係を示す式が上述のピンホールカメラモデルとして定義される。

【0045】

この式に含まれる A, C_w, R_w は、図 4 に示すように、

A : 正規化パラメータ

A : カメラ内部パラメータ、

C_w : カメラ位置、

R_w : カメラ回転行列、

を意味している。

【0046】

カメラ内部パラメータ A には、以下の値が含まれる。

f : 焦点距離

k_u : 画像軸の直交性 (理想値は 90°)

k_u : 縦軸のスケール (三次元位置のスケールから二次元画像のスケールへの変換)

k_v : 横軸のスケール (三次元位置のスケールから二次元画像のスケールへの変換)

(u_0, v_0) : 画像中心位置

【0047】

このピンホールカメラモデル、すなわち、三次元空間の点の位置とカメラ像平面の画素位置との射影変換を表す式を用いて、図 1 に示すデータ構築部 110 では、カメラ 101 の撮影画像から得られるカメラ座標系の情報を世界座標系の情報に変換した環境マップを生成する。例えば、カメラ 101 をユーザ (物理エージェント) が持ち、自由に移動させて動画像を撮影して画像認識処理部 102 と、自己位置検出部 103 に撮影画像を入力する場合の処理例について説明する。

【0048】

自己位置検出部 103 は、カメラ 101 から入力する映像内の特徴点を利用して、特徴点の位置とカメラ位置を毎フレーム修正し、自己位置検出部 103 で決めた世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラの位置 (C_w) およびカメラの回転行列 (R_w) をデータ構築部 110 に出力する。なお、この処理は、前述の論文 [Andrew J. Davison, "Real-time simultaneous localisation and mapping with a single camera", Proceedings of the 9th International Conference on Computer Vision, Ninth, (2003)] に記載された方法を用いることができる。

【0049】

また、カメラ 101 の撮影画像は画像認識処理部 102 にも送られる。画像認識処理部 102 では、特徴点の特徴を記述しているパラメータ (以下、特徴量) を利用して画像認識を行う。画像認識処理部 102 で得られる認識結果は、オブジェクトの属性 (コップ A や椅子 B など) と画像内の認識対象の位置・姿勢である。画像認識で得られた結果は、世界座標系で表現されていないため、過去の認識結果との直接的な統合は出来ない。

【0050】

そのため、データ構築部 110 は、前述したピンホールカメラモデル、すなわち、三次元空間の点の位置とカメラ像平面の画素位置との射影変換を表す式と事前情報として取得済みのオブジェクトの形状データを含む辞書データを辞書データ記憶部 104 から取得して三次元空間へ射影変換する。

【0051】

10

20

30

40

50

まず、前述の数式 1 から、以下の (数式 2)、すなわち、カメラ 1 0 1 の撮影画像に含まれるオブジェクトのある点、すなわち特徴点 (m) の世界座標系での三次元位置 (M) を示す式が導出される。

【数 2】

$$\mathbf{M} = \mathbf{C}_w + \lambda \cdot \mathbf{R}_w^T \mathbf{A}^{-1} \tilde{\mathbf{m}} = \mathbf{C}_w + d \cdot \mathbf{R}_w^T \frac{\mathbf{A}^{-1} \tilde{\mathbf{m}}}{\|\mathbf{A}^{-1} \tilde{\mathbf{m}}\|} \dots (\text{数式 2})$$

10

【0052】

上述の式において、

d は、カメラとオブジェクトの特徴点との実世界における距離を表す。距離 (d) が求まれば、カメラ座標系内の物体の位置を求めることが出来る。なお、前述したように、(数式 1) の () は正規化パラメータであるが、この () は、カメラとオブジェクトの距離 (奥行) に全く関係の無いある正規化用変数であり、距離 (d) の変化に関係なく入力データにより変化する。

【0053】

カメラ 1 0 1 からオブジェクトの特徴点との距離の算出処理について、図 5 を参照して説明する。図 5 は、オブジェクト 3 0 0 をカメラ 1 0 1 によって撮影した画像 3 1 0 を示している。撮影画像 3 1 0 にはオブジェクト 3 0 0 の画像であるオブジェクト画像 3 1 1 が含まれる。

20

【0054】

画像認識処理部 1 0 2 は、撮影画像 3 1 0 内のオブジェクト画像 3 1 1 から図に示すように、 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) の 4 つの特徴点を得る。画像認識処理部 1 0 2 は、辞書データ記憶部 1 0 4 から辞書データを抽出して、オブジェクト画像 3 1 1 と、辞書データとの照合を実行して、撮影されたオブジェクト画像 3 1 1 を特定し、その特定されたオブジェクトに対応して辞書データとして登録された特徴点情報に対応するオブジェクト画像 3 1 1 中の特徴点位置を決定する。この特徴点位置が図に示すように直方体の 4 つの頂点 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) 、 (x_4, y_4) であるとする。

30

【0055】

データ構築部 1 1 0 は、画像認識処理部 1 0 2 からこれらの特徴点 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) 、 (x_4, y_4) 情報を取得し、前述したピンホールカメラモデル、すなわち、三次元空間の点の位置とカメラ像平面の画素位置との射影変換を表す式と事前情報として取得済みのオブジェクトの形状データを含む辞書データを辞書データ記憶部 1 0 4 から取得して三次元空間へ射影変換する。具体的には、カメラからの各特徴点 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) 、 (x_4, y_4) までの距離 (奥行) を算出する。

【0056】

辞書データには、前述したように各オブジェクトの三次元形状情報が登録されている。画像認識処理部 1 0 2 は、この三次元形状情報に基づいて以下に示す (数式 3) を適用して、カメラからの各特徴点 (x_1, y_1) 、 (x_2, y_2) 、 (x_3, y_3) 、 (x_4, y_4) までの距離 (奥行) を算出する。

40

【数3】

$$\left\{ \begin{array}{l} \|d_\alpha \mathbf{X}_\alpha - d_\beta \mathbf{X}_\beta\| = d_{obj1} + \varepsilon_1 \\ \|d_\beta \mathbf{X}_\beta - d_\gamma \mathbf{X}_\gamma\| = d_{obj2} + \varepsilon_2 \\ \|d_\gamma \mathbf{X}_\gamma - d_\delta \mathbf{X}_\delta\| = d_{obj3} + \varepsilon_3 \\ \|d_\delta \mathbf{X}_\delta - d_\alpha \mathbf{X}_\alpha\| = d_{obj4} + \varepsilon_4 \end{array} \right. \dots (数式3)$$

10

【0057】

上記式(数式3)において、

d : カメラからオブジェクト300の特徴点 までの距離

d : カメラからオブジェクト300の特徴点 までの距離

d : カメラからオブジェクト300の特徴点 までの距離

d : カメラからオブジェクト300の特徴点 までの距離

dobj1 : オブジェクト300の特徴点 間の距離(辞書データの登録情報)

dobj2 : オブジェクト300の特徴点 間の距離(辞書データの登録情報)

dobj3 : オブジェクト300の特徴点 間の距離(辞書データの登録情報)

dobj4 : オブジェクト300の特徴点 間の距離(辞書データの登録情報)

である。

20

【0058】

また、

X : カメラから特徴点 へのベクトル

であり以下の式(数式4)によって算出される。

【数4】

$$\mathbf{X}_\alpha = \mathbf{R}_w^T \frac{\mathbf{A}^{-1} \tilde{\mathbf{m}}_\alpha}{\|\mathbf{A}^{-1} \tilde{\mathbf{m}}_\alpha\|} \dots (数式4)$$

30

【0059】

同様に、

X : カメラから特徴点 へのベクトル

X : カメラから特徴点 へのベクトル

X : カメラから特徴点 へのベクトル

これらは、いずれも画像内の各特徴点の画素位置情報(m)とカメラパラメータ(A)およびカメラ回転行列(R)とに基づいて算出可能な値である。

40

【0060】

また、式(数式3)に含まれる 1 ~ 4 は計測誤差である。これらの計測誤差を表す
が最小になるように、式3に含まれる4つの未知数、すなわち、カメラからオブジェク
ト300の特徴点までの距離: d , d , d , d を算出する。例えば最小二乗法を
用いて計測誤差 が最小になるようにカメラからオブジェクト300の特徴点までの各距

50

離： d_1, d_2, d_3, d_4 を算出する。

【0061】

この処理によってカメラからオブジェクト300の各特徴点までの距離（奥行）を計算することができ、世界座標系にオブジェクト認識結果を正確にマッピングすることが可能となる。但し、上記式（数式3）を用いた方法では、マッピングを行うには最低4点は必要となる。もし5点以上特徴点を得られた場合は、式を増やして誤差の総和が最小になるように計算する。

【0062】

図5を参照して説明した方法では、オブジェクト300の全体が撮影画像310内に記録されている。すなわちオブジェクト画像311はオブジェクト300の全体を含む設定である。オブジェクトの全体が撮影画像に含まれず一部だけが撮影されている場合、本発明の情報処理装置の構成では、カメラ位置と辞書データを利用して認識対象全体をマッピングすることが可能である。

10

【0063】

すなわち、図6に示すようにオブジェクト300の全体は撮影画像321に含まれず、オブジェクト画像321が、オブジェクト300の一部として撮像されている場合でも、本発明の情報処理装置の構成では、カメラ位置と辞書データを利用してオブジェクト（認識対象）全体をマッピングすることが可能である。図6の例では、特徴点（P）がオブジェクト画像321に含まれず、オブジェクト画像321の上面の4頂点に対応する特徴点 P_1, P_2, P_3, P_4 のみがオブジェクト画像321に含まれている。

20

【0064】

この場合も、図5を参照して説明したように前述の式（数式3）から、カメラからオブジェクト300の特徴点までの各距離： d_1, d_2, d_3, d_4 を算出することが可能となり、さらに、辞書データ中のオブジェクト300の三次元形状データに基づいて撮影されていない特徴点Pまでのカメラからの距離を算出することができ、世界座標系におけるオブジェクト300のマッピングを行なうことが可能となる。

【0065】

さらに、上述の手法は、1フレームのカメラ撮影画像のみを利用した処理として説明したが、複数のフレーム画像を適用した処理としてもよい。すなわち、過去に抽出した特徴点も利用する処理である。但し、特徴点の三次元空間の位置が必要なので、同じ特徴点を用いている自己位置検出部の特徴点の位置を利用する。

30

【0066】

[実施例2]

図7を参照して、複数のフレーム画像を適用した環境マップの生成処理構成を持つ情報処理装置について説明する。本実施例の構成においても、実施例1と同様、カメラによって撮影された画像データに基づいて環境マップを構築する情報処理装置であり、図7に示すように、周囲環境を撮影するカメラ101と、カメラ101の撮影画像を入力して画像認識を行う画像認識処理部102と、カメラ101の撮影画像を入力してカメラの位置姿勢を推定する自己位置検出部103と、画像認識処理部102において生成された画像認識結果データと、自己位置検出部103において検出されたカメラ101の位置および姿勢情報を入力して、ある世界座標系で表現された環境マップの作成処理を実行するデータ構築部110、さらに、画像認識処理部102における画像認識処理や、データ構築部110における環境マップ作成処理に利用する辞書データを格納した辞書データ記憶部104を有する。データ構築部110は環境マップ更新部111、環境マップデータベース112を有する。これらの構成は、図1を参照して説明した実施例1の構成と同様の構成である。

40

【0067】

本実施例2においては、さらに、特徴点照合部401、特徴点データベース402、座標変換部403を有する。自己位置検出部103ではカメラ101によって撮影された画像から得られる特徴点の三次元位置とカメラ位置を同時に求める処理を行なう。

50

【 0 0 6 8 】

エージェント (agent) の位置や方向などのポーズ (位置姿勢) の確認として実行される自己ポーズ (位置姿勢) 同定 (ローカリゼーション) に併せて、エージェント (agent) 周囲の地図 (環境地図) の作成 (mapping)、すなわち環境マップの作成を行なう処理は SLAM (simultaneous localization and mapping) と呼ばれる。本実施例 2 では、この SLAM 技法をベースにしている。

【 0 0 6 9 】

自己位置検出部 103 は、カメラ 101 から入力する映像内の特徴点を利用して、特徴点の位置とカメラ位置を解析フレーム毎に算出し、自己位置検出部 103 で決めた世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラ位置 (Cw)、カメラの回転行列 (Rw) をデータ構築部 110 に出力する。さらに、自己位置検出部 103 は、解析対象とするフレームに対応する特徴点の三次元位置情報を特徴点照合部 401 に出力する。なお、解析対象フレームは、カメラ 101 の撮影する画像フレームの連続する全てのフレームとしてよいし、予め設定した間隔のフレームとしてもよい。

10

【 0 0 7 0 】

特徴点照合部 401 は、画像認識処理部 102 で求めた特徴点と、自己位置検出部 103 から入力する特徴点とを撮影された画像空間の距離、例えば画素数で比較する。但し自己位置検出部 103 が、世界座標系で特徴点位置情報を算出している場合は、前述の数式 (数式 1) を用いて画像内の特徴点位置を求める。

20

【 0 0 7 1 】

特徴点照合部 401 では、自己位置検出部 103 があるフレーム画像から検出したある 1 つの特徴点位置と、画像認識処理部 102 が同じフレーム画像から検出したある 1 つの特徴点位置との距離が予め定めた閾値以下であるか否かを検証し、閾値以下ならば、その特徴点同士は同一のものであるとし、特徴点の特徴量と対応する自己位置検出部 103 が検出した特徴点のカメラ座標系における位置情報と識別子を特徴点データベース 402 に登録する。

【 0 0 7 2 】

座標変換部 403 は、特徴点データベース 402 に登録された特徴点の各々について現在のカメラ画像座標系に位置情報を変換して、画像認識処理部 102 に出力する。位置情報の変換処理には、前述の式 (数式 1) と、自己位置検出部 103 から入力する世界座標系で表されたカメラの推定位置および姿勢情報としてのカメラ位置 Cw、カメラの回転行列 (Rw) を用いる。

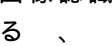
30

【 0 0 7 3 】

画像認識処理部 102 では、特徴点解析対象としている現フレームの特徴点位置情報に併せて、座標変換部 403 から入力する先行する過去フレームから求められた特徴点位置情報を追加して、データ構築部 110 に出力することができる。このようにして複数フレームから抽出されたより多くの特徴点情報からなる認識結果を生成してデータ構築部 110 に出力することが可能となり、データ構築部 110 ではより多くの特徴点情報に基づく正確な環境マップを生成することが可能となる。

40

【 0 0 7 4 】

図 8 は、過去フレームから取得した特徴点情報の利用例について示す図である。撮影画像 420 は、画像認識処理部 102 において解析中のフレームである撮影画像を示している。撮影画像 420 には、オブジェクト 400 の一部のみが記録されている。特徴点 P は、過去のフレームにおいて検出された特徴点位置である。座標変換部 403 は、過去フレームから抽出された特徴点 P の位置を現在のフレームである撮影画像 420 のどの位置に対応するかを算出し、画像認識処理部 102 に提供する。画像認識処理部 102 では、撮影画像 420 に含まれる、 の 4 つの特徴点の他、過去フレームにおいて抽出された特徴点 P を追加して、これらの特徴点情報をデータ構築部 110 に出力することができる。このようにして複数フレームから抽出されたより多くの特徴点情報からなる認識

50

結果を生成してデータ構築部 110 に出力することが可能となり、データ構築部 110 ではより多くの特徴点情報に基づく正確な環境マップを生成することが可能となる。

【0075】

[実施例3]

次に、本発明の実施例3としてカメラ位置とカメラ画像の認識結果を用いて、環境マップを更新する処理例について説明する。図1に示すデータ構築部110では、画像認識処理部102から入力する特徴点情報と、事前情報として取得済みのオブジェクトの形状データを含む辞書データを辞書データ記憶部104から取得して、各オブジェクトを世界座標系で表される環境マップ上に配置する処理を行って環境マップを生成する。

【0076】

環境マップを構築した後、環境の一部が変更された場合を想定する。変更された場合、過去に構築した環境マップを破棄して、新たに入力する特徴点情報に基づいて新たな環境マップを構築するという手法もあるが、この処理は非常に手間がかかる。そこで、本実施例では、過去に構築された環境マップを利用して、変更された部分を識別して識別した変更部分を更新することで新たな環境マップを生成する構成としたものである。

【0077】

本実施例の情報処理装置の構成および処理について、図9を参照して説明する。本実施例の構成においても、実施例1と同様、カメラによって撮影された画像データに基づいて環境マップを構築する情報処理装置であり、図9に示すように、周囲環境を撮影するカメラ101と、カメラ101の撮影画像を入力して画像認識を行う画像認識処理部102と、カメラ101の撮影画像を入力してカメラの位置姿勢を推定する自己位置検出部103と、画像認識処理部102において生成された画像認識結果データと、自己位置検出部103において検出されたカメラ101の位置および姿勢情報を入力して、ある世界座標系で表現された環境マップの作成処理を実行するデータ構築部510、さらに、画像認識処理部102における画像認識処理や、データ構築部510における環境マップ作成処理に利用する辞書データを格納した辞書データ記憶部104を有する。データ構築部510は環境マップ更新部511、環境マップデータベース512を有する。これらの構成は、図1を参照して説明した実施例1の構成と同様の構成である。

【0078】

さらに、本実施例の構成ではデータ構築部510が、環境情報取得部513、認識結果比較部514を有する。環境情報取得部513は、自己位置検出部103で得られたカメラの自己位置情報を用いて、環境マップデータベース512に格納済みの環境マップ(例えば過去の画像に基づいて生成された環境マップ)を利用してカメラ101に写ると推定される様々なオブジェクトからなる環境情報を検出する。

【0079】

認識結果比較部514は、環境情報取得部513が、生成済みの環境マップに基づいて推定した現在のカメラ位置に応じた撮影画像から取得されると判断した環境情報(特徴点情報など)と画像認識処理部102において、実際の撮影画像から得られた認識結果(特徴点情報)とを比較する。

【0080】

環境マップ更新部511は、この認識結果比較部514の比較情報を入力し、比較情報を利用して、環境マップデータベース512に格納済みの過去に作成された環境マップを修正してマップ更新処理を実行する。

【0081】

本実施例で利用する各モジュールについて詳しく説明する。

環境情報取得部513では自己位置検出部103において取得されるカメラ位置姿勢情報を入力して、環境マップデータベース512に記録済みの環境マップのどの部分が撮影画像として取得され、その推定した撮影画像に含まれる特徴点の位置などを判断する。

【0082】

なお、環境マップデータベース512に記録済みの環境マップは、各オブジェクトが世

10

20

30

40

50

界座標系で登録されているので、前述の式（数式1）を利用して各特徴点位置をカメラ座標系に変換して撮影画像のどの位置に特徴点が見えるかを決定する。前述の（数式1）において、カメラの推定位置および姿勢情報としてのカメラ位置（ C_w ）、カメラの回転行列（ R_w ）は自己位置検出部103からの入力情報を適用する。（数式1）の左辺の m が、新たに撮影され画像認識処理部において解析対象としている画像内に入っているか否かを判断する。

【0083】

認識結果比較部514では、環境情報取得部513が生成済みの環境マップに基づいて推定した現在のカメラ位置に応じた撮影画像から取得されると判断した環境情報（特徴点情報など）と画像認識処理部102において、実際の撮影画像から得られた認識結果（特徴点情報）とを比較する。なお、比較情報としては、様々なパラメータがあるが本例では、

10

（1）特徴点同士の比較

（2）ある確率分布モデルで表現された認識対象の重心と姿勢同士の比較

これらの2通りの比較例について説明する。

【0084】

（1）特徴点同士の比較

まず、認識結果比較部514において特徴点比較を行なう構成例について図10を参照して説明する。本実施例では図10に示すように、環境情報取得部513には、特徴点情報生成部521が設定され、認識結果比較部514には、特徴点情報比較部522が設けられ、環境マップ更新部511には画像内変更領域抽出部523、環境マップ更新部524が設けられている。

20

【0085】

環境情報取得部513の特徴点情報生成部521は、自己位置検出部103において取得されるカメラ位置姿勢情報を入力して、環境マップデータベース512に記録済みの環境マップのどの部分が撮影画像として取得され、その推定した撮影画像に含まれる特徴点の位置などの特徴点情報を取得する。因みに特徴点情報とは、特徴点の位置と特徴量で構成される情報をさす。特徴点情報生成部521は、環境マップとカメラ位置姿勢と辞書データをもとに、カメラに写ると推定される特徴点全ての特徴点情報を生成する。

【0086】

認識結果比較部514の特徴点比較部522は、特徴点情報生成部521が過去に蓄積された環境マップから求めた特徴点情報を入力し、さらに画像認識処理部102において、実際の撮影画像から得られた認識結果（特徴点情報）を入力してこれらを比較する。特徴点比較部522は、マッチした特徴点の位置と特徴点同士の距離を求める。これらの情報は、環境マップ更新部511の画像内変更領域抽出部523に出力される。

30

【0087】

画像内変更領域抽出部523は、マッチした特徴点の距離が予め設定した閾値より近い領域以外を環境マップの更新領域として抽出する。この抽出領域情報は、環境マップ更新部524に出力される。

【0088】

環境マップ更新部524は、過去に蓄積された環境マップを取得して、画像内変更領域抽出部523において更新領域として選択された抽出領域についてのみ、新たに画像認識処理部102から入力する特徴点情報を用いたマップ更新を実行し、その他の領域は、作成済みの環境マップの情報を更新せずに利用する。この処理により、効率的な環境マップ更新が可能となる。

40

【0089】

（2）ある確率分布モデルで表現された認識対象の重心と姿勢同士の比較

次に認識結果比較部514において、ある確率分布モデルで表現された認識対象の重心と姿勢同士の比較を行なう構成例について図11を参照して説明する。本実施例では図11に示すように、環境情報取得部513には、存在確率分布抽出部541が設定され、認

50

識結果比較部 5 1 4 には、存在確率分布生成部 5 4 2、確率分布比較部 5 4 3 が設けられている。

【 0 0 9 0 】

認識結果比較部 5 1 4 の存在確率分布生成部 5 4 2 は、画像認識処理部 1 0 2 からの認識結果を入力して、画像認識結果にある認識オブジェクトの存在確率分布を生成する。ちなみに、存在確率分布は、平均（三次元位置とクオータニオン）共分散行列で表現された多次元正規分布に従うとする。画像認識では、存在確率分布は求まらないためユーザーが定数で与える。

【 0 0 9 1 】

環境情報取得部 5 1 3 の存在確率分布抽出部 5 4 1 は、環境マップデータベース 5 1 2 から取得した環境マップを利用して、カメラ 1 0 1 の撮影した画像に含まれる認識オブジェクトの重心・姿勢の存在確率分布を抽出する（平均共分散行列）。

【 0 0 9 2 】

認識結果比較部 5 1 4 の存在確率分布比較部 5 4 3 は、環境情報取得部 5 1 3 の存在確率分布抽出部 5 4 1 が環境マップに基づいて生成した認識オブジェクトの存在確率分布と、認識結果比較部 5 1 4 の存在確率分布生成部 5 4 2 が画像認識処理部 1 0 2 からの認識結果に基づいて生成したオブジェクトの存在確率分布との比較を行う比較には、例えば以下の式（数式 5）を用いる。

【数 5】

$$s = (\mathbf{z} - \hat{\mathbf{z}})^T (\Sigma_z + \Sigma_{\hat{z}})^{-1} (\mathbf{z} - \hat{\mathbf{z}}) \quad \dots \text{(数式 5)}$$

【 0 0 9 3 】

上記式（数式 5）の s はマハラノビス距離を表しており、環境マップに基づいて生成した認識オブジェクトの存在確率分布と、画像認識処理部 1 0 2 からの認識結果に基づいて生成したオブジェクトの存在確率分布との差分に応じた値となる。認識結果比較部 5 1 4 の存在確率分布比較部 5 4 3 は、上記式（数式 5）の s の値を環境マップ更新部 5 1 1 に出力する。

【 0 0 9 4 】

環境マップ更新部 5 1 1 は、認識結果比較部 5 1 4 の存在確率分布比較部 5 4 3 から入力する上記式（数式 5）の s の値が予め設定した閾値より小さい場合、環境マップデータベース 1 1 2 に記録済みのオブジェクトと、新たな撮影画像に基づく画像認識結果として得られた認識オブジェクトは同一のものであると判断し、環境マップデータベースの修正は行わない。しかし、 s が閾値より大きい場合は異なるオブジェクトが出現したものと判断して、環境マップデータベース 5 1 2 に登録された環境マップを更新する。

【 0 0 9 5 】

[実施例 4]

上述の実施例においては、環境マップに登録される各オブジェクトは移動しない静オブジェクトであることを前提とした処理であった。移動するオブジェクトが撮影画像に含まれ、これらの移動オブジェクトを環境マップに登録する場合には、移動オブジェクト（動物体）と、静止オブジェクト（静物体）を区別した処理が必要となる。実施例 4 は、このような移動オブジェクト（動物体）と、静止オブジェクト（静物体）を区別した処理を行って環境マップの生成、更新を実行する処理例である。

【 0 0 9 6 】

本実施例について、図 1 2 を参照して説明する。図 1 2 に示す例は、先に図 1 1 を参照

10

20

30

40

50

して説明したある確率分布モデルで表現された認識対象の重心と姿勢同士の比較を行なう構成例をベースとして、動物体判定部561を追加した構成である。

【0097】

図10を参照して説明した構成との差異は、動物体判定部561が、画像認識処理部102からの認識結果を入力して、画像認識結果としてのオブジェクトが動物体かどうか判断する処理が含まれる点である。動物体判定部561は、辞書データ記憶部104の格納データに基づいて、画像認識結果としてのオブジェクトが動物体か静物体かを判断する。この処理例では、辞書データ記憶部104の格納データである辞書データには、各オブジェクトが動物体であるか静物体であるかを示すオブジェクト属性データが登録されているものとする。

10

【0098】

画像認識処理部102からの認識結果に含まれるオブジェクトに動物体として認識されるオブジェクトが含まれる場合、このオブジェクト情報は、直接、環境マップ更新部511に出力され、環境マップ更新部511は、オブジェクト情報に基づいて動物体を環境マップ上に登録する。

【0099】

画像認識処理部102からの認識結果に含まれるオブジェクトに静物体として認識されるオブジェクトが含まれる場合、このオブジェクト情報は、認識結果比較部514に出力される。この後の処理は、図11を参照して説明したとど羽陽の処理が実行される。なお、環境マップ更新部511は、動物体の情報は毎フレーム自動的に消去する。

20

【0100】

[実施例5]

次に、実施例5として、環境マップを用いて画像認識処理部の認識対象・探索範囲を限定する構成例について説明する。

【0101】

環境マップは、例えば上述した実施例1~4のいずれかの処理によって生成、更新することが可能である。このようにして生成した環境マップを利用して、画像認識処理部において認識するオブジェクトの種類や探索範囲を限定することを可能とした構成を実施例5として説明する。

【0102】

事前に物体間のオントロジーデータ(意味情報)をまとめたオントロジーマップを用意し、記憶部に格納する。オントロジーデータ(意味情報)は、環境に存在する対象を体系的に分類し、その関係を記述するものである。例えば、

「椅子やテーブルは床に接している可能性が高い」、

「電源が入っているテレビのモニタの前には顔がある可能性が高い」

「トイレにテレビがある可能性は低い」

というような情報がオントロジーデータである。対象物体毎にオントロジーデータをまとめたものがオントロジーマップである。

【0103】

本実施例5では、このようなオントロジーマップを用いて、画像認識処理部において認識するオブジェクトの種類や探索範囲を限定する。本実施例5の情報処理装置は、環境マップに基づいてオブジェクト検索範囲を特定する処理を実行する情報処理装置であり、オブジェクトの隣接領域に特定オブジェクトが存在する可能性を示すオントロジーデータ(意味情報)を格納した記憶部と、オントロジーデータ(意味情報)に基づいて、特定オブジェクトの探索領域を決定する画像認識処理部とを有する情報処理装置である。

30

40

【0104】

例えば、図13に示す環境で顔の位置を検出する場合について説明する。但し、既に環境マップは構築されているものとする。最初に、床面を検出し天井方向を推定する。地面に接している可能性が非常に高いものは(自動販売機やテーブル、車など)。図13中では、

50

- ・テーブル
- ・ソファ

があげられる。環境マップより、テーブルとソファの位置・姿勢がわかっているため、それから床の平面方程式を算出する。これで、床面が識別され、天井方向が識別される。

【0105】

次に、顔の位置がありそうな場所を、オントロジーマップをもとに探索する。図13の環境では、

- ・ [顔がある可能性が非常に高い場所] ソファの上、テレビの前、
- ・ [顔がある可能性が非常に低い場所] テレビの真上、本棚の中、カレンダーより向こうの壁

10

このように区分することが出来る。この区分を三次元空間に対応させて範囲を指定し、その範囲を前述の式(数式1)で画像に反映させる。例えば図14に示すような区分情報711, 712が設定できる。この反映結果を画像認識モジュールに渡す。この処理により、画像認識率を向上させ、画像認識の高速化を図ることが可能となる。

【0106】

以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、特許請求の範囲の欄を参酌すべきである。

20

【0107】

また、明細書中において説明した一連の処理はハードウェア、またはソフトウェア、あるいは両者の複合構成によって実行することが可能である。ソフトウェアによる処理を実行する場合は、処理シーケンスを記録したプログラムを、専用のハードウェアに組み込まれたコンピュータ内のメモリにインストールして実行させるか、あるいは、各種処理が実行可能な汎用コンピュータにプログラムをインストールして実行させることが可能である。例えば、プログラムは記録媒体に予め記録しておくことができる。記録媒体からコンピュータにインストールする他、LAN(Local Area Network)、インターネットといったネットワークを介してプログラムを受信し、内蔵するハードディスク等の記録媒体にインストールすることができる。

30

【0108】

なお、明細書に記載された各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。また、本明細書においてシステムとは、複数の装置の論理的集合構成であり、各構成の装置が同一筐体内にあるものには限らない。

【産業上の利用可能性】

【0109】

以上、説明したように、本発明の一実施例の構成によれば、カメラの取得した画像に基づいてカメラの位置および姿勢を検出する自己位置検出処理と、カメラの取得画像からオブジェクトを検出する画像認識処理と、カメラの位置姿勢情報とオブジェクト情報と、オブジェクトに対応する少なくとも三次元形状データを含むオブジェクト情報を登録した辞書データを適用して環境マップの生成または更新処理を実行する構成としたので、様々なオブジェクトの三次元データを反映した環境マップを、1つのカメラによって取得した画像に基づいて効率的に生成することが可能となる。

40

【図面の簡単な説明】

【0110】

【図1】本発明の一実施例に係る情報処理装置の構成および処理について説明する図である。

【図2】辞書データ記憶部の格納データである辞書データについて説明する図である。

【図3】ピンホールカメラモデル、すなわちカメラ座標系によって表現されている位置と

50

、世界座標系におけるオブジェクトの三次元位置との対応関係を示す式の意味について説明する図である。

【図4】ピンホールカメラモデル、すなわちカメラ座標系によって表現されている位置と、世界座標系におけるオブジェクトの三次元位置との対応関係を示す式の意味について説明する図である。

【図5】カメラからオブジェクトの特徴点との距離の算出処理について説明する図である。

【図6】オブジェクトが一部撮像されている場合におけるカメラ位置と辞書データを利用したオブジェクト（認識対象）全体のマッピング処理について説明する図である。

【図7】本発明の一実施例に係る情報処理装置の構成および処理について説明する図である。

10

【図8】過去フレームから取得した特徴点情報の利用例について示す図である。

【図9】本発明の一実施例に係る情報処理装置の構成および処理について説明する図である。

【図10】本発明の一実施例に係る情報処理装置の構成および処理について説明する図である。

【図11】本発明の一実施例に係る情報処理装置の構成および処理について説明する図である。

【図12】本発明の一実施例に係る情報処理装置の構成および処理について説明する図である。

20

【図13】顔の位置を検出する画像認識処理例について説明する図である。

【図14】顔の位置を検出する画像認識処理例について説明する図である。

【符号の説明】

【0111】

101 カメラ

102 画像認識処理部

103 自己位置検出部

104 辞書データ記憶部

110 データ構築部

111 環境マップ更新部

112 環境マップデータベース

401 特徴点照合部

402 特徴点データベース

403 座標変換部

510 データ更新部

511 環境マップ更新部

512 環境マップデータベース

513 環境情報取得部

514 認識結果比較部

521 特徴点情報生成部

522 特徴点情報比較部

523 画像内変更領域抽出部

524 環境マップ更新部

541 存在確率分布抽出部

542 存在確率分布生成部

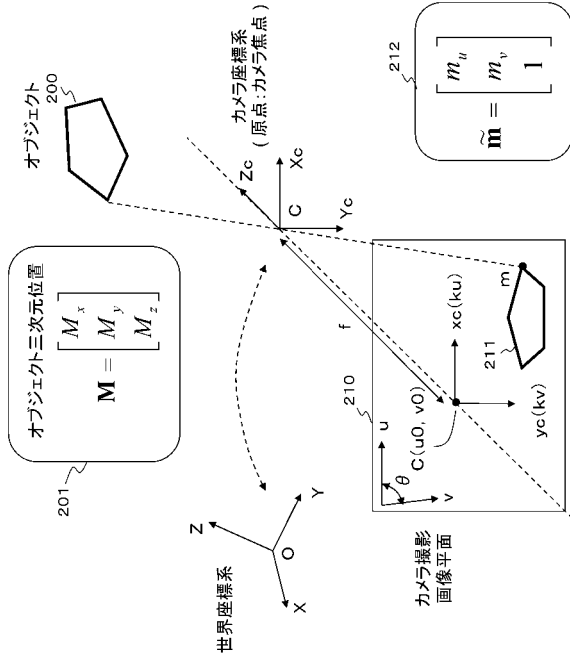
543 確率分布比較部

561 動物体判定部

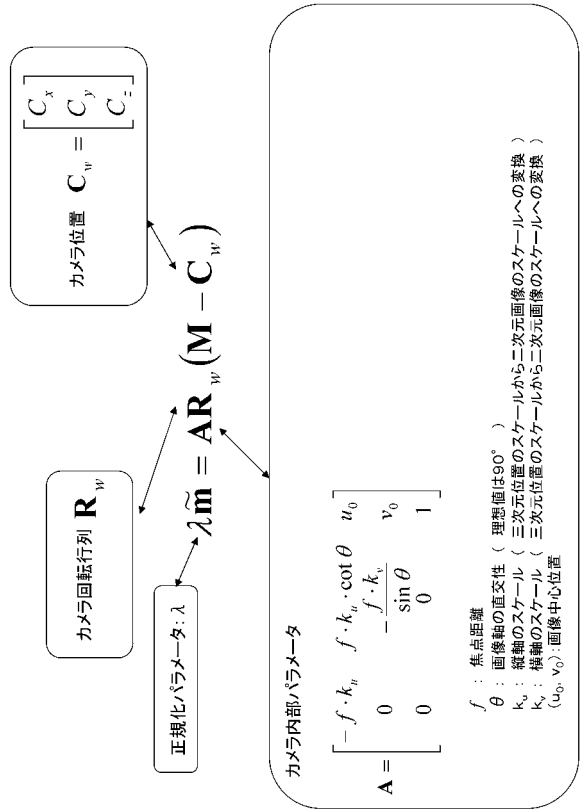
30

40

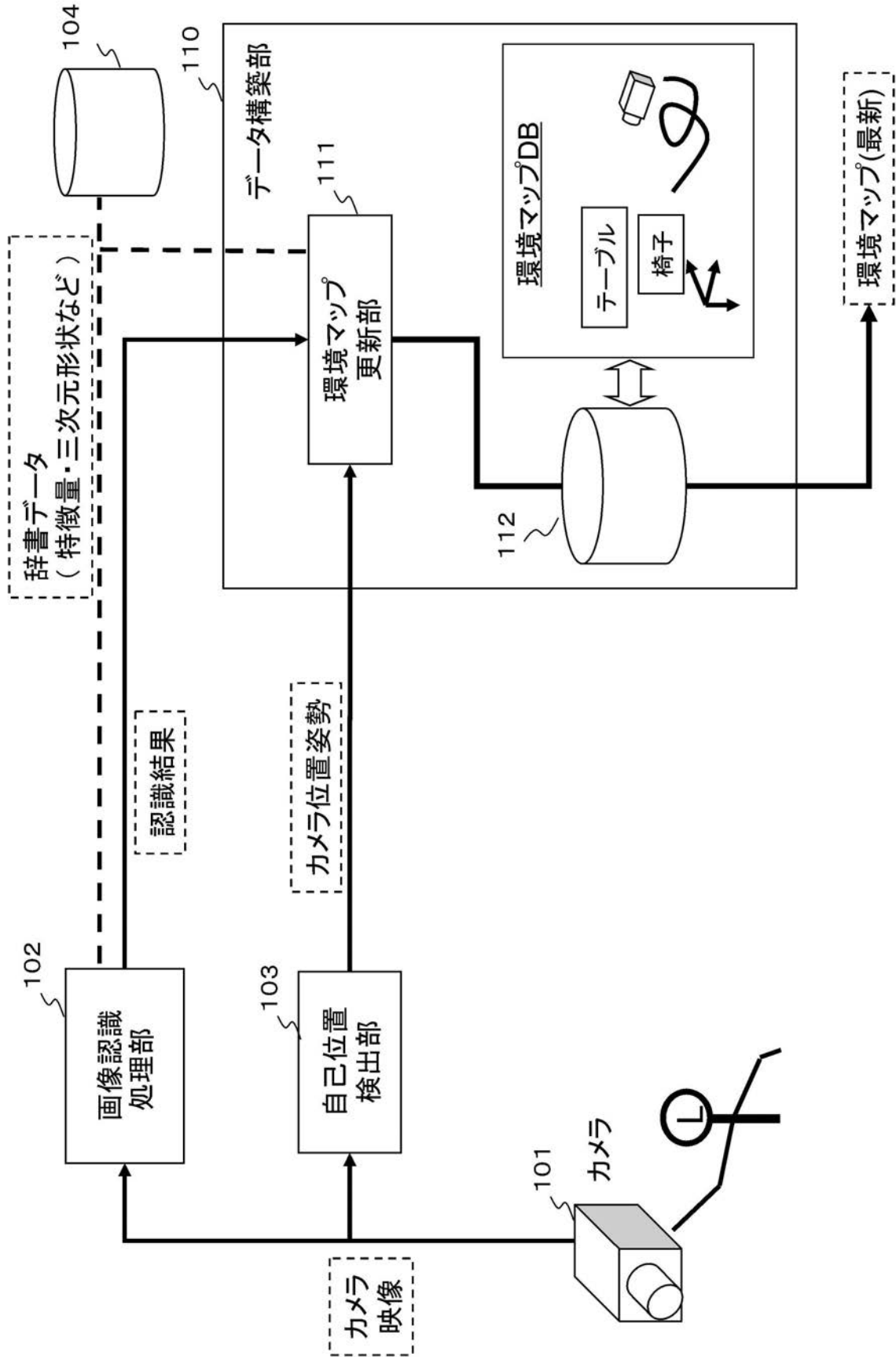
【 図 3 】



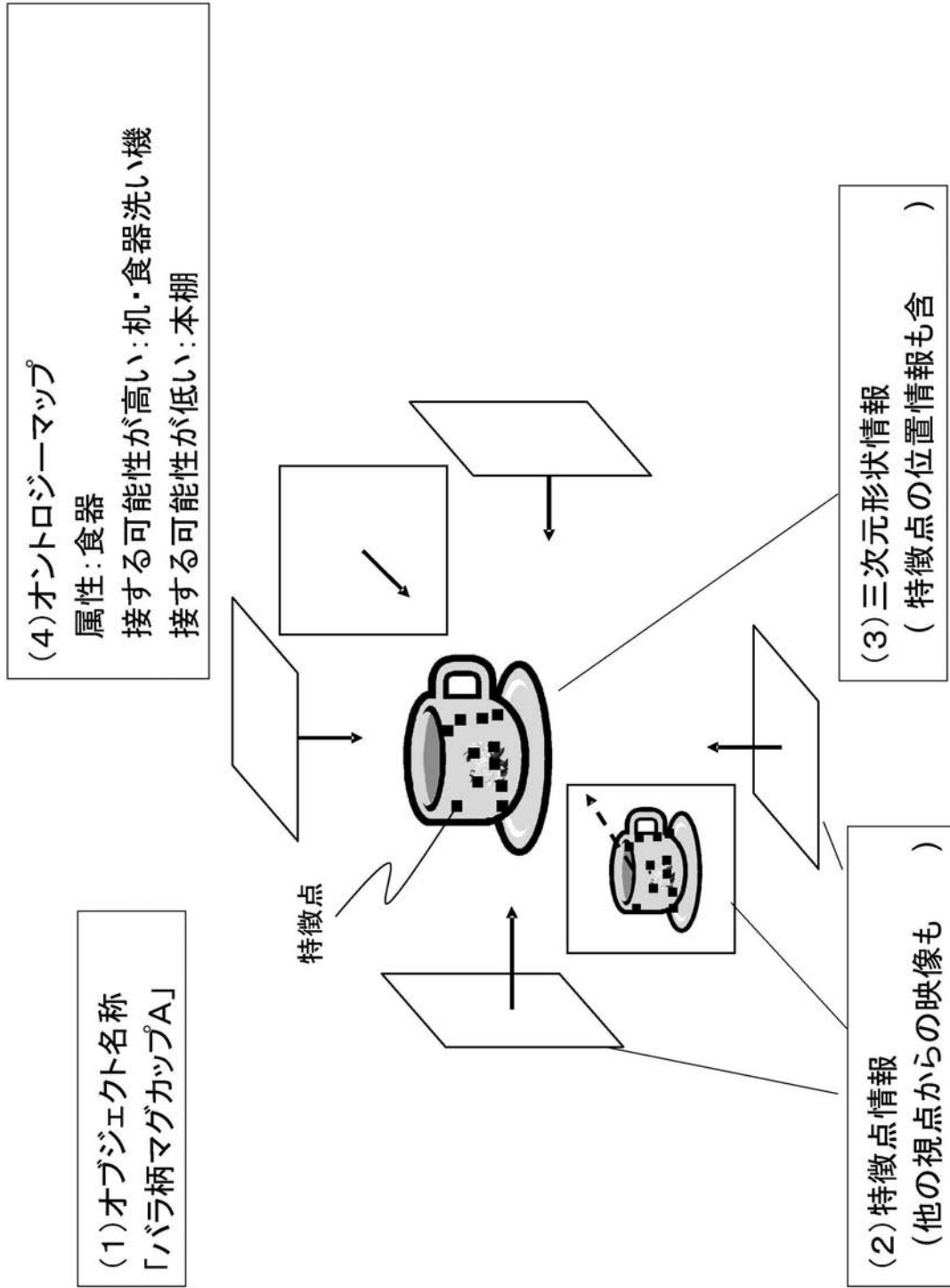
【 図 4 】



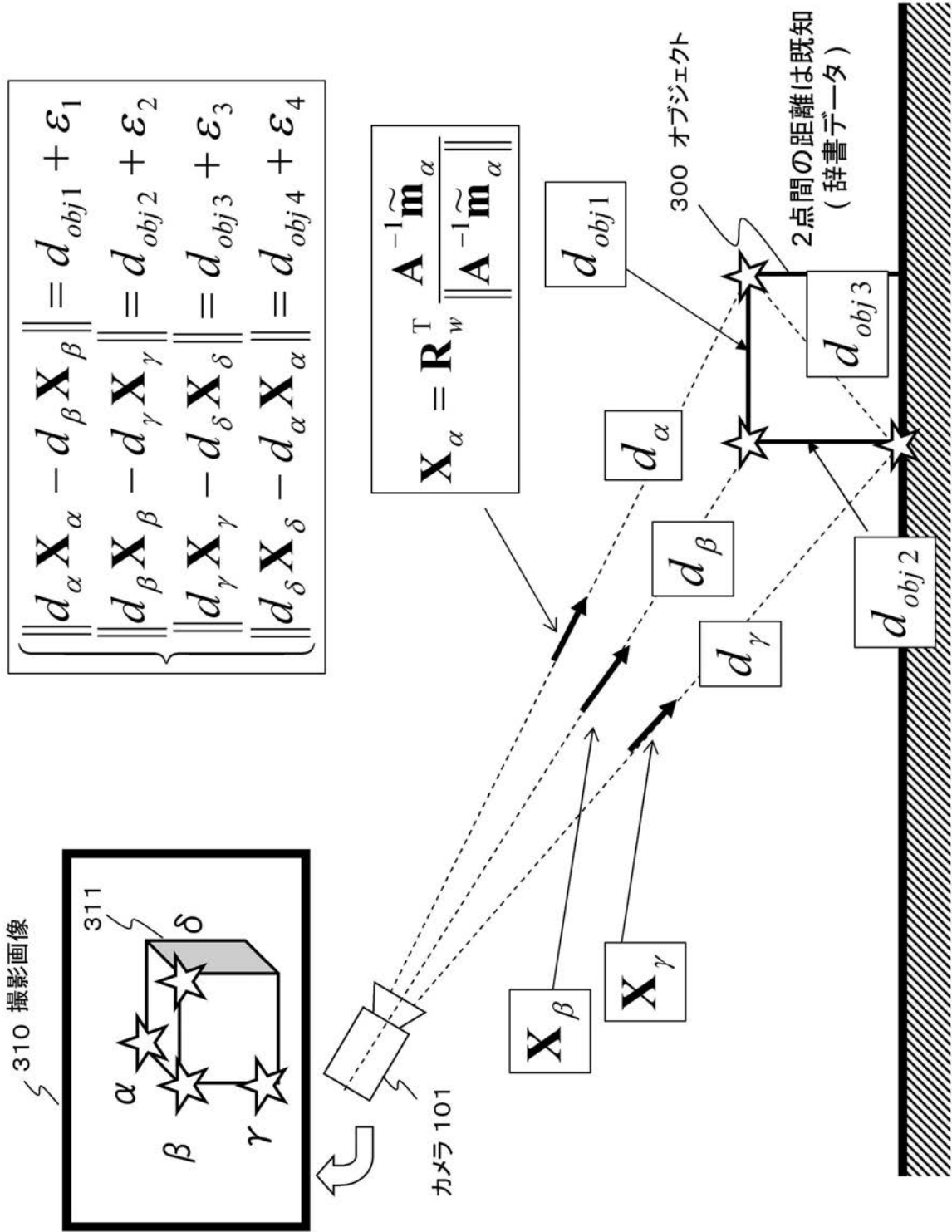
【図1】



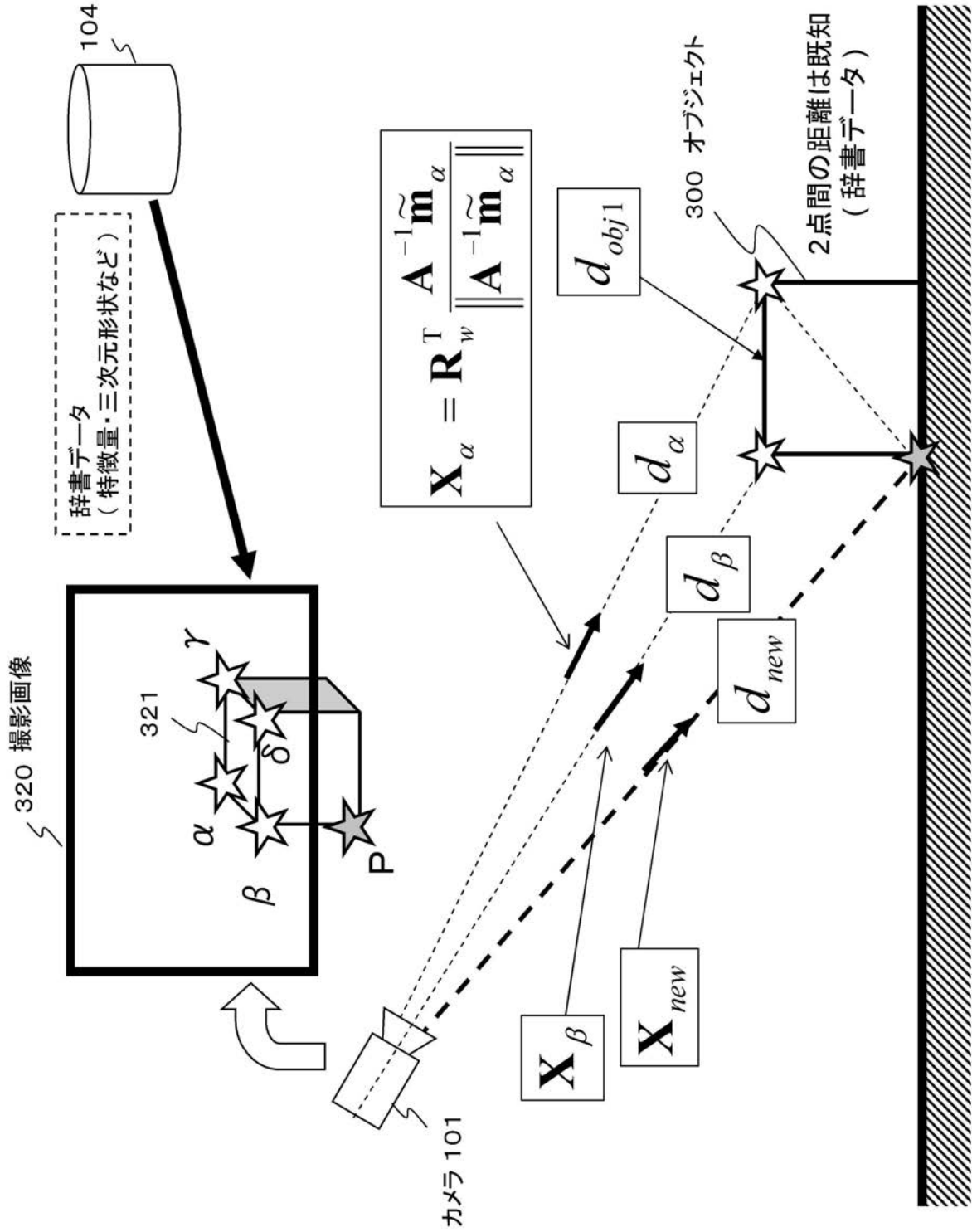
【 図 2 】



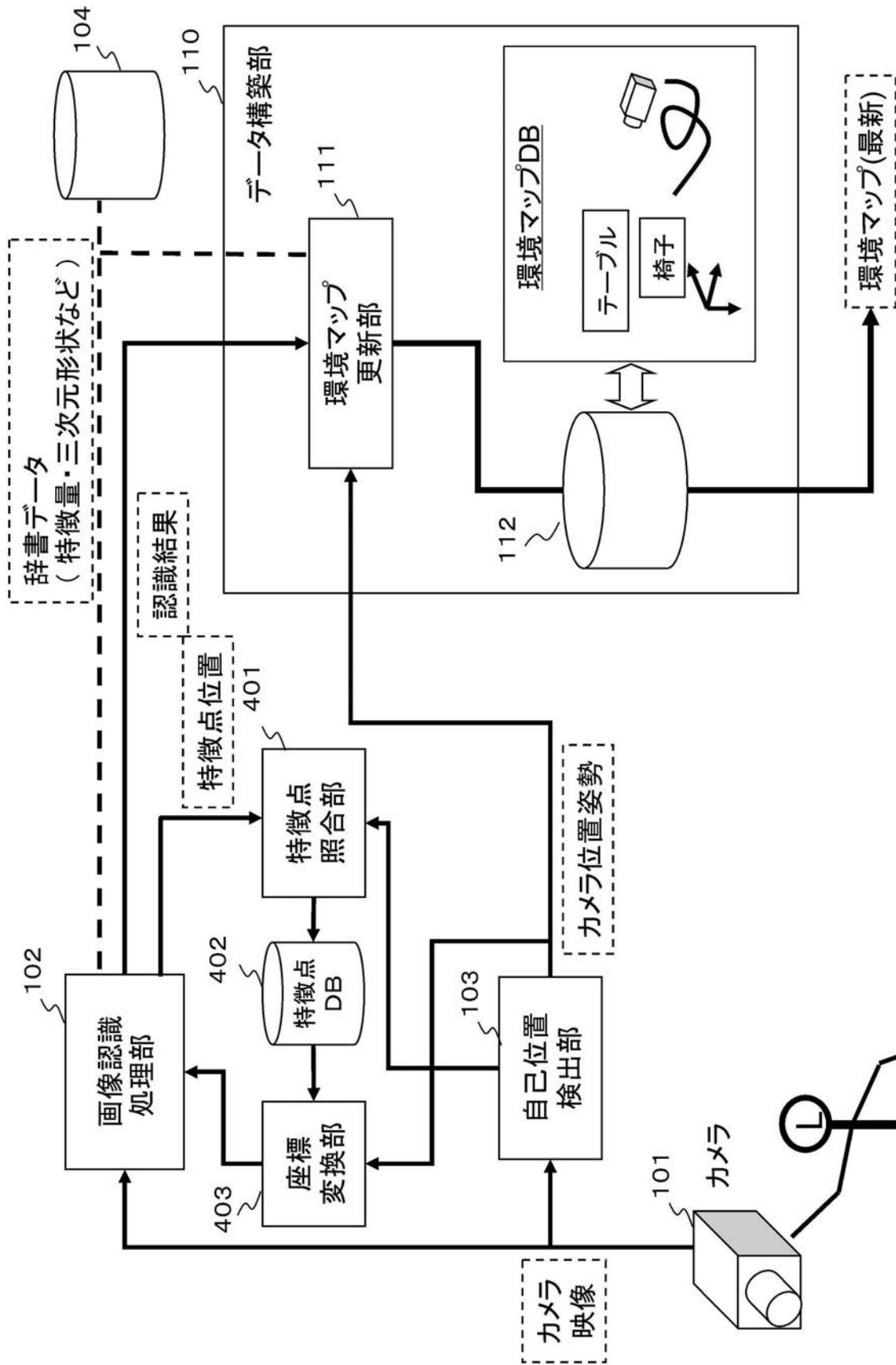
【図5】



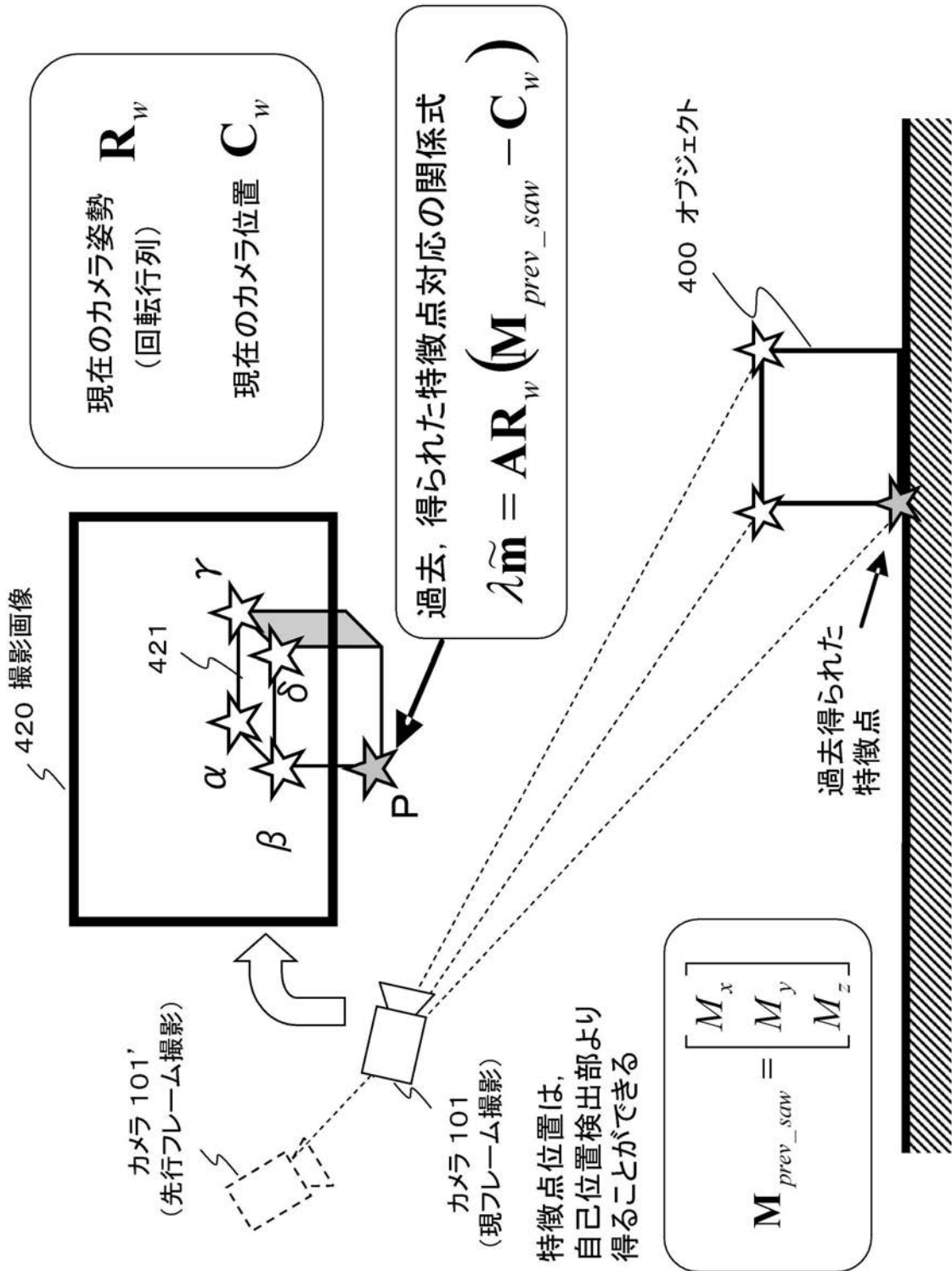
【図6】



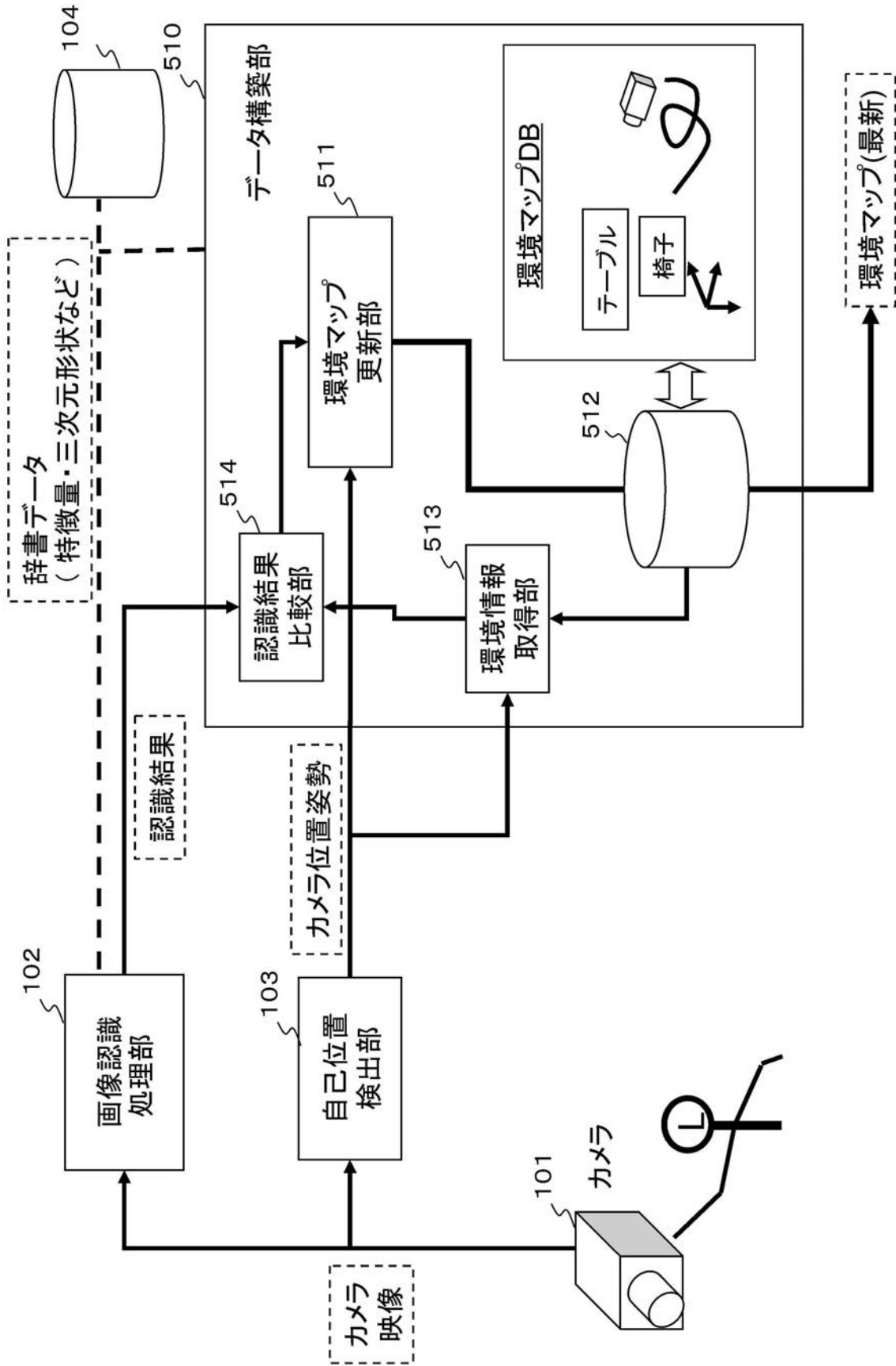
【図7】



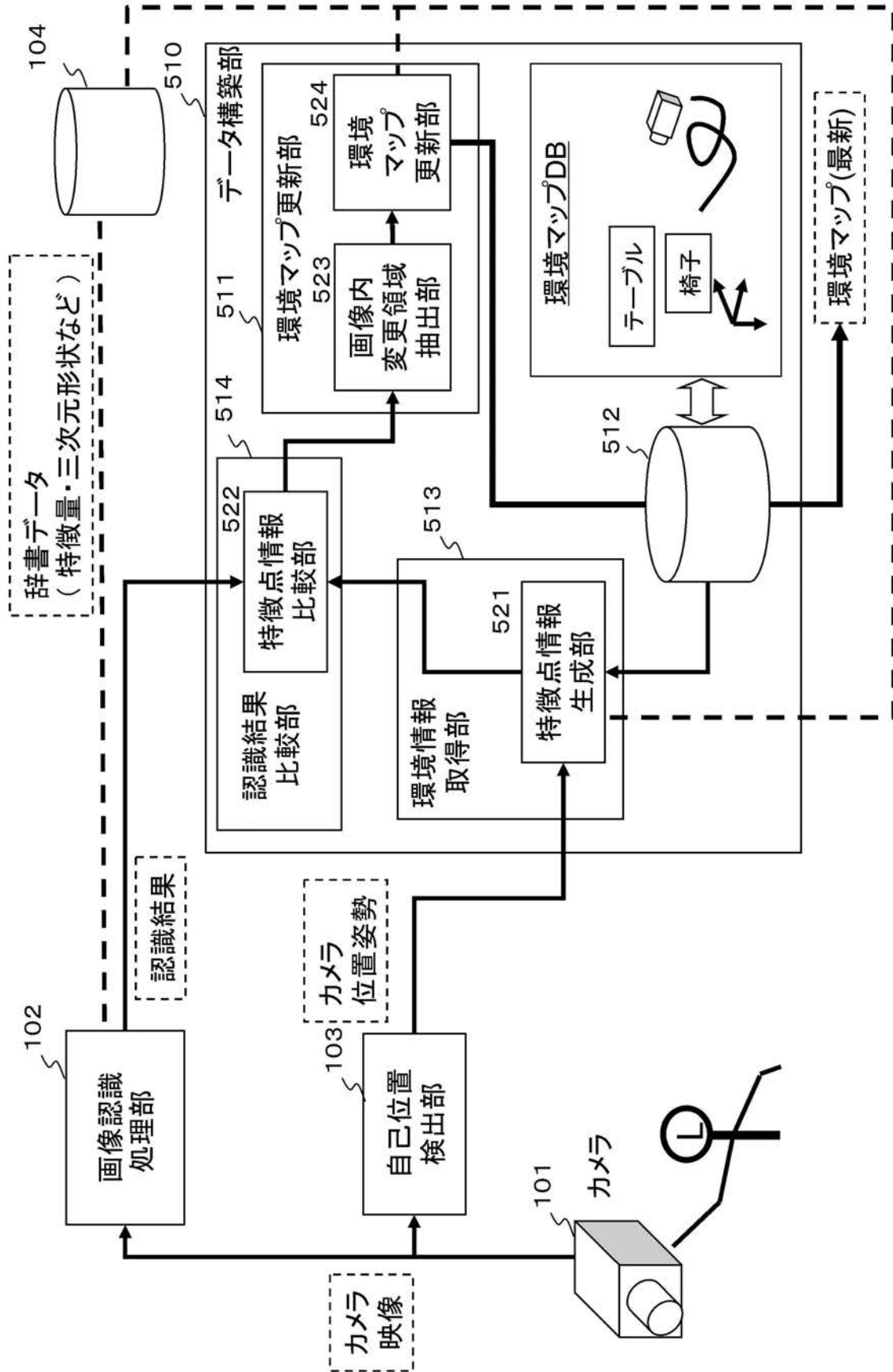
【 図 8 】



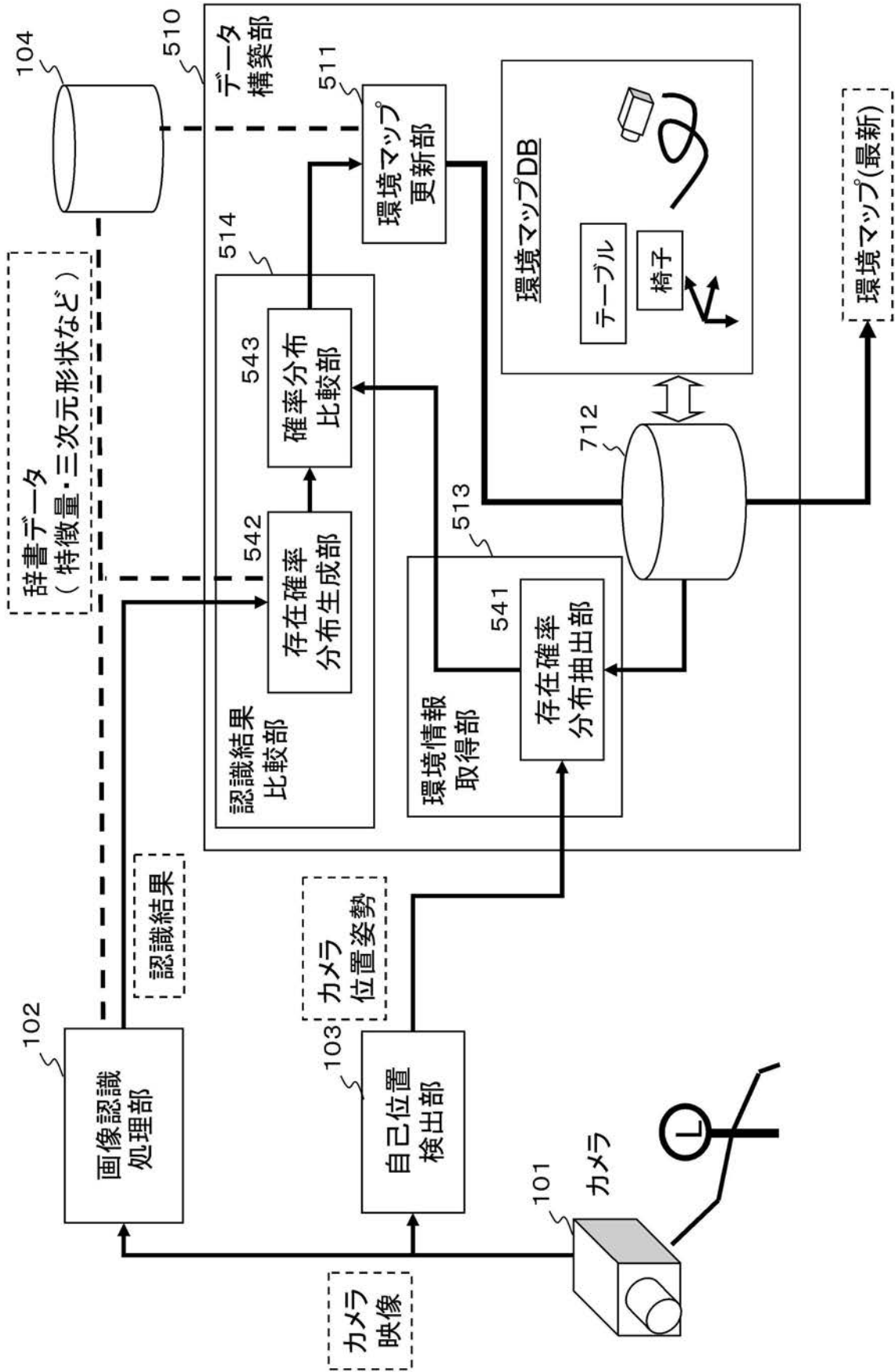
【図9】



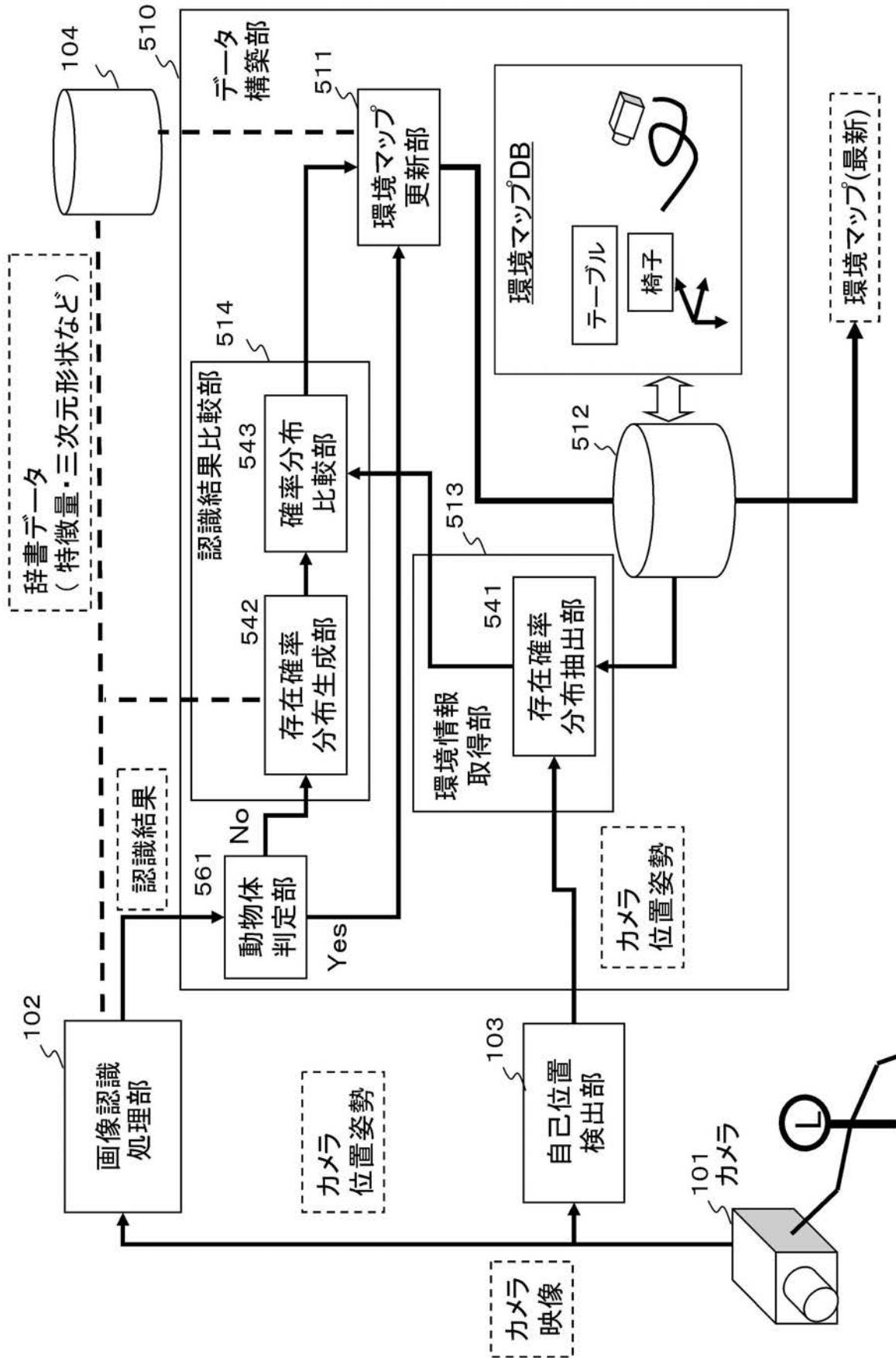
【図10】



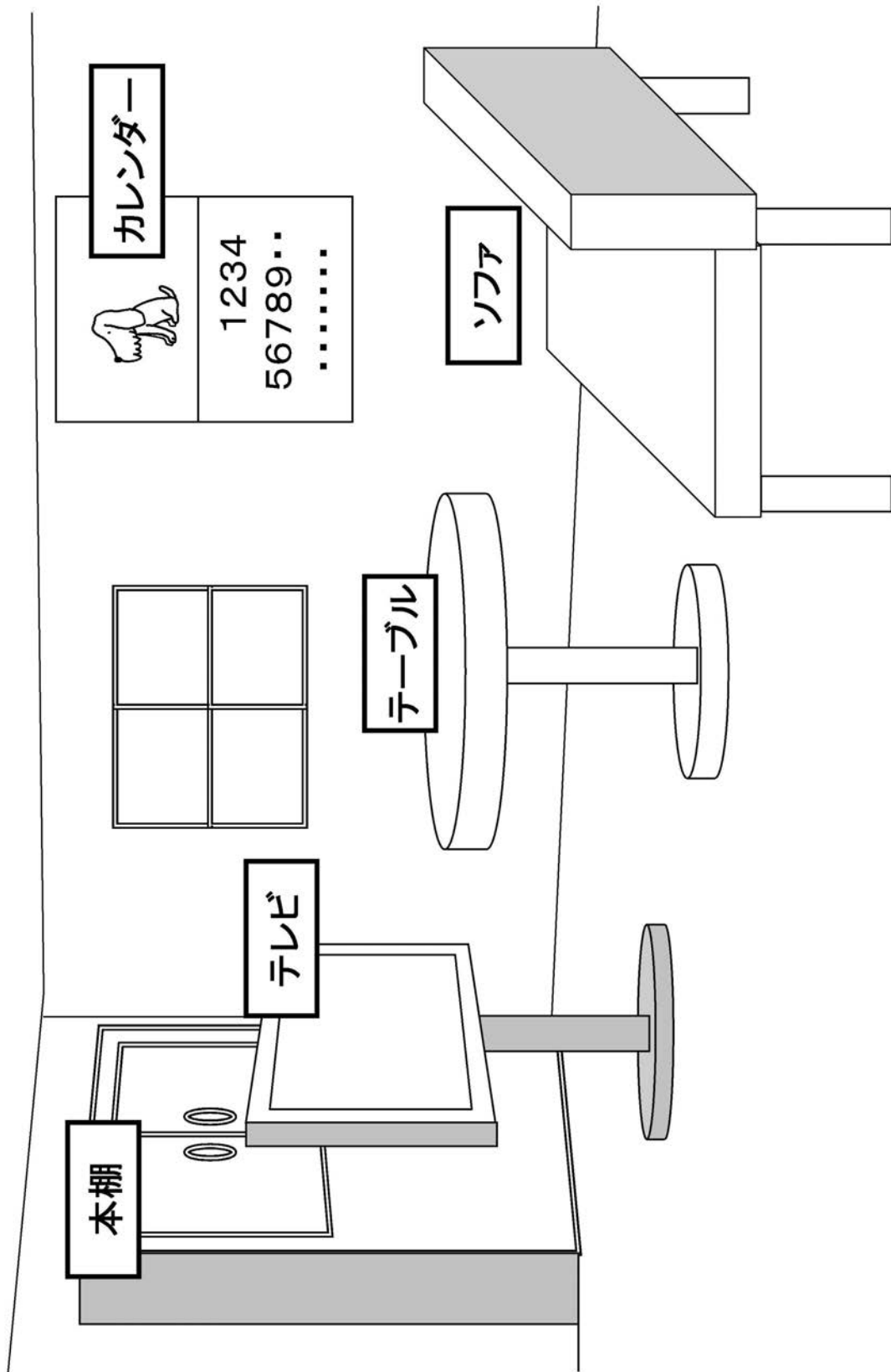
【図11】



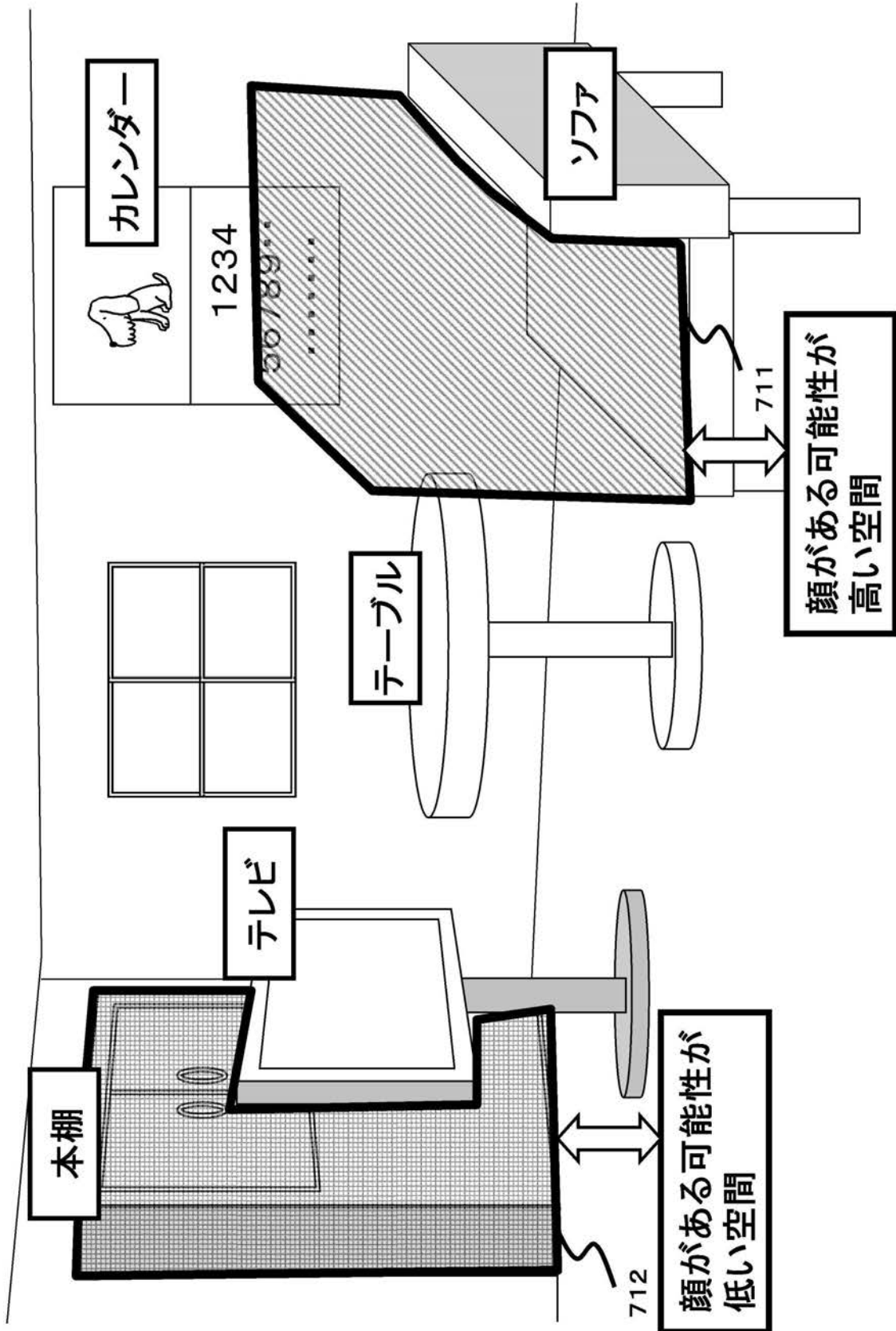
【図12】



【 図 1 3 】



【 図 1 4 】



フロントページの続き

- (72)発明者 グットマン ステフェン
東京都港区港南1丁目7番1号 ソニー株式会社内
- (72)発明者 芦ヶ原 隆之
東京都港区港南1丁目7番1号 ソニー株式会社内

審査官 うし 田 真悟

- (56)参考文献 特開2004-110802(JP,A)
特表2007-535765(JP,A)
特開2004-030445(JP,A)
友納正裕, 形状情報と認識情報を統合した物体モデルによる3次元物体マップの構築, 第11回ロボティクスシンポジウム予稿集, ロボティクスシンポジウム運営委員会, 2006年 3月16日, p.246-p.251

(58)調査した分野(Int.Cl., DB名)

G01B 11/00 - 11/30
G05D 1/00 - 1/12
G06T 1/00
G06T 7/00 - 7/60
JSTPlus(JDreamIII)