

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6144828号
(P6144828)

(45) 発行日 平成29年6月7日(2017.6.7)

(24) 登録日 平成29年5月19日(2017.5.19)

(51) Int.Cl.

F I

G 0 6 T 7 / 2 0 (2017.01)

G 0 6 T 7 / 2 0

B

請求項の数 30 (全 17 頁)

(21) 出願番号	特願2016-520066 (P2016-520066)	(73) 特許権者	507364838
(86) (22) 出願日	平成26年9月25日 (2014. 9. 25)		クアルコム, インコーポレイテッド
(65) 公表番号	特表2016-536668 (P2016-536668A)		アメリカ合衆国 カリフォルニア 921
(43) 公表日	平成28年11月24日 (2016. 11. 24)		21 サン ディエゴ モアハウス ドラ
(86) 国際出願番号	PCT/US2014/057522		イブ 5775
(87) 国際公開番号	W02015/050773	(74) 代理人	100108453
(87) 国際公開日	平成27年4月9日 (2015. 4. 9)		弁理士 村山 靖彦
審査請求日	平成29年2月13日 (2017. 2. 13)	(74) 代理人	100163522
(31) 優先権主張番号	61/887, 211		弁理士 黒田 晋平
(32) 優先日	平成25年10月4日 (2013. 10. 4)	(72) 発明者	アンドリュー・フィリップ・ジー
(33) 優先権主張国	米国 (US)		アメリカ合衆国・カリフォルニア・921
(31) 優先権主張番号	14/494, 947		21-1714・サン・ディエゴ・モアハ
(32) 優先日	平成26年9月24日 (2014. 9. 24)		ウス・ドライブ・5775
(33) 優先権主張国	米国 (US)		
早期審査対象出願		最終頁に続く	

(54) 【発明の名称】 動的に構築された環境地図データに基づくオブジェクト追跡

(57) 【特許請求の範囲】

【請求項 1】

オブジェクト認識システムにおけるターゲットオブジェクトの追跡のコンピュータ実装方法であって、

カメラで複数の画像を取得するステップと、

同時に前記ターゲットオブジェクトを追跡して、前記複数の画像から環境地図データを動的に構築するステップであって、前記ターゲットオブジェクトを追跡するステップが、前記複数の画像のうちの少なくとも1つに基づいて、およびターゲット地図データに基づいて、前記カメラに対する前記ターゲットオブジェクトのターゲットポーズを推定しようとするステップを含む、ステップと、

前記カメラに対する前記ターゲットオブジェクトの前記追跡が成功したかどうかを決定するステップと、失敗した場合、

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測するステップと

を備える、コンピュータ実装方法。

【請求項 2】

前記カメラに対する前記ターゲットオブジェクトの追跡が成功したと決定されると、前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測するステップと、

前記推定されたターゲットポーズと前記推測されたターゲットポーズとを融合するステ

ップと

をさらに備える、請求項1に記載のコンピュータ実装方法。

【請求項3】

前記推定されたターゲットポーズと前記推測されたターゲットポーズとを融合するステップが、カルマンフィルタを適用するステップを含む、請求項2に記載のコンピュータ実装方法。

【請求項4】

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測するステップが、

前記カメラに対する前記環境の環境ポーズを推定するステップと、

前記環境に対する前記ターゲットオブジェクトのターゲットポーズを推定するステップと、

前記環境に対する前記推定されたターゲットポーズに基づいて、前記カメラに対する前記ターゲットポーズを推測するステップと

を含む、請求項1に記載のコンピュータ実装方法。

【請求項5】

前記環境地図データが、前記カメラで取得された1つまたは複数の画像に対応する1つまたは複数のキーフレームに基づく、請求項1に記載のコンピュータ実装方法。

【請求項6】

前記環境地図データを構築するステップが、キーフレームSLAM(Simultaneous Localization And Mapping)を含む、請求項5に記載のコンピュータ実装方法。

【請求項7】

前記環境地図データを構築するステップが、候補キーフレームと、前記カメラによって取得された画像に対応する他のオンラインキーフレームとの間のポーズ差を計算することによって、前記1つまたは複数のキーフレームに前記候補キーフレームを追加するステップを含む、請求項5に記載のコンピュータ実装方法。

【請求項8】

前記候補キーフレームと他のオンラインキーフレームとの間の前記ポーズ差を計算するステップが、前記ポーズ差を計算する際にオフラインキーフレームをスキップするステップを含む、請求項7に記載のコンピュータ実装方法。

【請求項9】

前記ポーズ差が、前記候補キーフレームと他のオンラインキーフレームとの間の最小ポーズ差であり、前記ポーズ差がポーズしきい値を上回る場合、前記候補キーフレームを前記他のオンラインキーフレームに追加するステップをさらに備える、請求項7に記載のコンピュータ実装方法。

【請求項10】

オブジェクト認識システムにおけるターゲットオブジェクトを追跡するためのプログラムコードを記憶した非一時的コンピュータ可読記憶媒体であって、前記プログラムコードが、

カメラで複数の画像を取得する命令と、

同時に前記ターゲットオブジェクトを追跡して、前記複数の画像から環境地図データを動的に構築する命令であって、前記ターゲットオブジェクトを追跡する命令が、前記複数の画像のうちの少なくとも1つに基づいて、およびターゲット地図データに基づいて、前記カメラに対する前記ターゲットオブジェクトのターゲットポーズを推定しようとする命令を含む、命令と、

前記カメラに対する前記ターゲットオブジェクトの追跡が成功したかを決定する命令と、失敗した場合、

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測する命令と

を備える、非一時的コンピュータ可読記憶媒体。

10

20

30

40

50

【請求項 1 1】

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測する前記命令が、

前記カメラに対する前記環境の環境ポーズを推定する命令と、

前記環境に対する前記ターゲットオブジェクトのターゲットポーズを推定する命令と、

前記環境に対する前記推定されたターゲットポーズに基づいて、前記カメラに対する前記ターゲットポーズを推測する命令と

を含む、請求項10に記載の非一時的コンピュータ可読記憶媒体。

【請求項 1 2】

前記環境地図データが、前記カメラで取得された1つまたは複数の画像に対応する1つまたは複数のキーフレームに基づく、請求項10に記載の非一時的コンピュータ可読記憶媒体。

10

【請求項 1 3】

前記環境地図データを構築する前記命令が、キーフレームSLAM(Simultaneous Localization And Mapping)を含む、請求項12に記載の非一時的コンピュータ可読記憶媒体。

【請求項 1 4】

前記環境地図データを構築する前記命令が、前記1つまたは複数のキーフレームに候補キーフレームを追加して、前記候補キーフレームと、前記カメラによって取得された画像に対応する他のオンラインキーフレームとの間のポーズ差を計算する命令を含む、請求項12に記載の非一時的コンピュータ可読記憶媒体。

20

【請求項 1 5】

前記候補キーフレームと他のオンラインキーフレームとの間の前記ポーズ差を計算する前記命令が、前記ポーズ差を計算する際にオフラインキーフレームをスキップする命令を含む、請求項14に記載の非一時的コンピュータ可読記憶媒体。

【請求項 1 6】

前記ポーズ差が、前記候補キーフレームと他のオンラインキーフレームとの間の最小ポーズ差であり、前記ポーズ差がポーズしきい値を上回る場合、前記候補キーフレームを前記他のオンラインキーフレームに追加する命令をさらに備える、請求項14に記載の非一時的コンピュータ可読記憶媒体。

【請求項 1 7】

30

装置であって、

オブジェクト認識システムにおけるターゲットオブジェクトを追跡するためのプログラムコードを記憶するように適合されたメモリと、

前記プログラムコード内に含まれる命令にアクセスして実行するように適合された処理ユニットと

を備え、前記命令が前記処理ユニットによって実行されると、前記処理ユニットが前記装置に、

カメラで複数の画像を取得することと、

同時に前記ターゲットオブジェクトを追跡して、前記複数の画像から環境地図データを動的に構築することであって、前記ターゲットオブジェクトを追跡することが、前記複数の画像のうちの少なくとも1つに基づいて、およびターゲット地図データに基づいて、前記カメラに対する前記ターゲットオブジェクトのターゲットポーズを推定しようとすることを含むことと、

40

前記カメラに対する前記ターゲットオブジェクトの追跡が成功したかを決定することと、失敗した場合、

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測することと

を行わせるように指示する、装置。

【請求項 1 8】

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測する

50

前記命令が、

前記カメラに対する前記環境の環境ポーズを推定する命令と、

前記環境に対する前記ターゲットオブジェクトのターゲットポーズを推定する命令と、

前記環境に対する前記推定されたターゲットポーズに基づいて、前記カメラに対する前記ターゲットポーズを推測する命令と

を含む、請求項17に記載の装置。

【請求項 19】

前記環境地図データが、前記カメラで取得された1つまたは複数の画像に対応する1つまたは複数のキーフレームに基づく、請求項17に記載の装置。

【請求項 20】

前記環境地図データを構築する前記命令が、キーフレームSLAM(Simultaneous Localization And Mapping)を含む、請求項19に記載の装置。

【請求項 21】

前記環境地図データを構築する前記命令が、前記1つまたは複数のキーフレームに候補キーフレームを追加して、前記候補キーフレームと、前記カメラによって取得された画像に対応する他のオンラインキーフレームとの間のポーズ差を計算する命令を含む、請求項19に記載の装置。

【請求項 22】

前記候補キーフレームと他のオンラインキーフレームとの間の前記ポーズ差を計算する前記命令が、前記ポーズ差を計算する際にオフラインキーフレームをスキップする命令を含む、請求項21に記載の装置。

【請求項 23】

前記ポーズ差が、前記候補キーフレームと他のオンラインキーフレームとの間の最小ポーズ差であり、前記ポーズ差がポーズしきい値を上回る場合、前記候補キーフレームを前記他のオンラインキーフレームに追加する命令をさらに備える、請求項21に記載の装置。

【請求項 24】

前記複数の画像を取得するための前記カメラをさらに備える、請求項17に記載の装置。

【請求項 25】

カメラで複数の画像を取得するための手段と、

同時に、ターゲットオブジェクトを追跡し、かつ、前記複数の画像から環境地図データを動的に構築するための手段であって、前記ターゲットオブジェクトを追跡するための前記手段が、前記複数の画像のうちの少なくとも1つに基づいて、およびターゲット地図データに基づいて、前記カメラに対する前記ターゲットオブジェクトのターゲットポーズを推定しようとするための手段を含む、手段と、

前記カメラに対する前記ターゲットオブジェクトの前記追跡が成功したかどうかを決定するための手段と、

前記カメラに対する前記ターゲットオブジェクトの前記追跡が失敗した場合、前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測するための手段と

を備える、装置。

【請求項 26】

前記カメラに対する前記ターゲットオブジェクトの追跡が成功したと決定されると、前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測するための手段と、

前記推定されたターゲットポーズと前記推測されたターゲットポーズとを融合するための手段と

をさらに備える、請求項25に記載の装置。

【請求項 27】

前記環境地図データに基づいて、前記カメラに対する前記ターゲットポーズを推測するための前記手段が、

前記カメラに対する前記環境の環境ポーズを推定するための手段と、
前記環境に対する前記ターゲットオブジェクトのターゲットポーズを推定するための手段と、

前記環境に対する前記推定されたターゲットポーズに基づいて、前記カメラに対する前記ターゲットポーズを推測するための手段と
を含む、請求項25に記載の装置。

【請求項28】

前記環境地図データが、前記カメラで取得された1つまたは複数の画像に対応する1つまたは複数のキーフレームに基づき、前記環境地図データを構築するための前記手段が、キーフレームSLAM(Simultaneous Localization And Mapping)を含む、請求項25に記載の装置。

10

【請求項29】

前記環境地図データが、前記カメラで取得された1つまたは複数の画像に対応する1つまたは複数のキーフレームに基づき、前記環境地図データを構築するための前記手段が、候補キーフレームと、前記カメラによって取得された画像に対応する他のオンラインキーフレームとの間のポーズ差を計算することによって、前記1つまたは複数のキーフレームに前記候補キーフレームを追加するための手段を含む、請求項25に記載の装置。

【請求項30】

前記ポーズ差が、前記候補キーフレームと他のオンラインキーフレームとの間の最小ポーズ差であり、前記装置が、前記ポーズ差がポーズしきい値を上回る場合、前記候補キーフレームを前記他のオンラインキーフレームに追加するための手段をさらに備える、請求項29に記載の装置。

20

【発明の詳細な説明】

【技術分野】

【0001】

関連出願の相互参照

本出願は、2013年10月4日に出願された米国仮出願第61/887,211号の利益を主張する。

【0002】

本開示は、一般に、コンピュータビジョンベースのオブジェクト認識アプリケーションに関し、詳細には、限定はしないが、拡張現実システムにおけるオブジェクト追跡に関する。

30

【背景技術】

【0003】

モバイル電話または他のモバイルプラットフォーム上で拡張現実(AR)を可能にするための課題は、オブジェクトをリアルタイムに検出して追跡する問題である。ARアプリケーションのためのオブジェクト検出には、非常に厳しい要件がある。それは、完全な6自由度を提供し、所与の座標系に対して絶対的な測定値を与え、非常に堅牢であり、リアルタイムに実行しなければならない。興味深いのは、コンピュータビジョン(CV)ベースの手法を使用して、カメラポーズを計算する方法であり、その方法は、カメラの視野内のオブジェクトを第1に検出すること、続いてそれを追跡することに依存する。一態様では、検出動作は、デジタル画像内に含まれるフィーチャ(feature)のセットを検出することを含む。フィーチャは、その領域を取り巻くエリアと比較して明るさや色などの特性の点で異なる、デジタル画像内の領域を指すことができる。一態様では、フィーチャは、いくつかの特性が一定である、または所定の値の範囲内で変動する、デジタル画像の領域である。

40

【0004】

次いで、検出されたフィーチャは、画像内に現実世界のオブジェクトが存在するかどうかを決定するために、フィーチャデータベースに含まれる、既知のフィーチャと比較される。したがって、視覚ベースのARシステムの動作における重要な要素は、フィーチャデータベースの構成である。いくつかのシステムでは、フィーチャデータベースは、様々な既知の視点から、既知のターゲットオブジェクトの複数のサンプル画像をキャプチャするこ

50

とによって、実行時以前に構築される。次いで、これらのサンプル画像からフィーチャが抽出されて、フィーチャデータベースに追加される。

【 0 0 0 5 】

最近では、拡張現実システムは、カメラによってキャプチャされた色またはグレースケール画像データに基づく、モデルベースの追跡アルゴリズム、またはSLAM(Simultaneous Localization and Mapping)アルゴリズムに変化している。SLAMアルゴリズムは、カメラによってキャプチャされた、入ってくる画像シーケンスから、3次元(3D)ポイントを再構築し、リアルタイムに場面の3D地図(すなわち、SLAM地図)を構築するために使用される。再構築された地図から、現在の画像フレーム内のカメラの6DoF(自由度)ポーズを位置特定することが可能である。

10

【 0 0 0 6 】

いくつかのシステムでは、ターゲットオブジェクトのSLAM地図は、実行時以前に、およびオブジェクトから近距離で生成される。実行時に、入ってくるビデオフレームからオブジェクトに対するカメラの6DoFポーズを推定するために、生成されたオブジェクトのSLAM地図が使用される。

【 0 0 0 7 】

既存の方法では、追跡性能は、カメラの視野内のオブジェクトの外観と、その大きさに依存する。ターゲットオブジェクトが小さい場合、部分的に遮蔽されている場合、または独特の視覚的フィーチャに欠けている場合、推定されたカメラポーズは精度を失い、また重要な追跡ジッタ(jitter)を示す場合がある。さらに極端な状況では、非常に遠くのオブジェクトと、現在の視野の外にあるオブジェクトとをまったく追跡することができないので、ターゲットに登録されたあらゆる仮想拡張も失われてしまう。

20

【 先行技術文献 】

【 非特許文献 】

【 0 0 0 8 】

【 非特許文献 1 】 IS-95標準

【 非特許文献 2 】 IS-2000標準

【 非特許文献 3 】 IS-856標準

【 非特許文献 4 】 IEEE802.11x

【 非特許文献 5 】 IEEE802.15x

30

【 発明の概要 】

【 課題を解決するための手段 】

【 0 0 0 9 】

本明細書に記載のいくつかの実施形態は、任意の追跡可能なオブジェクトの改善された追跡堅牢性を提供し、ターゲットの可視範囲を超えて追跡可能なポーズの範囲を拡張することができる。本明細書に記載の実施形態は、参照ターゲットの追跡に基づく任意の拡張現実(AR)システムに適用され得る。一態様では、追跡システムは、ターゲット地図データに基づいてカメラフレーム内のターゲットポーズを推定し、並行して、ポーズ計算機は、ターゲットを取り巻く環境の動的に生成された3D表現(すなわち、環境地図データ)に基づいてターゲットポーズの推定を決定する。追跡がターゲット地図データに基づいて失敗した場合、ポーズ計算機によって生成された、推定されたターゲットポーズが使用され得る。

40

【 0 0 1 0 】

たとえば、SLAM(Simultaneous Localization And Mapping)システムは、カメラが環境を移動するときに、ターゲットを取り巻く環境の3D表現(任意でターゲット自体も含み得る)を構築し得る。次に、SLAM地図データに基づいて環境ポーズが推定される。カメラのフレーム内のターゲットポーズと環境ポーズとがともに環境内のターゲットポーズの推定を提供し、それらはより多くの観察が記録されるにつれて改良され得る。ターゲットの追跡が失敗した場合、たとえば、ターゲットが視野の外に出るか、小さくなった場合、環境内のターゲットポーズの推定は、追跡された環境ポーズからターゲットポーズを推測する

50

ために使用される。一実施形態では、ターゲットの追跡が不安定になる場合、たとえば、小さいサイズまたは画像内のフィーチャのカバレッジ不良による低精度のためにジッタになると、堅牢性を向上させるために、推測されたターゲットポーズと実際のターゲットポーズとが融合される。

【0011】

本開示の上記および他の態様、目的、ならびに特徴は、添付の図面に関連して与えられる様々な実施形態の以下の説明から明らかになるであろう。

【0012】

本発明の非限定的かつ非網羅的な実施形態を、以下の図面を参照して説明する。図面において、特に断らない限り、様々な図を通じて同様の参照番号は同様の部分を指す。

10

【図面の簡単な説明】

【0013】

【図1A】異なる距離のターゲットオブジェクトを含む場面の画像を示す図である。

【図1B】異なる距離のターゲットオブジェクトを含む場面の画像を示す図である。

【図2】環境地図データを構築するプロセスを示す流れ図である。

【図3】場面内のターゲットオブジェクトの追跡のプロセスを示す流れ図である。

【図4】オブジェクト検出および追跡のための処理ユニットの機能ブロック図である。

【図5】本明細書に記載のプロセスを実行することが可能なモバイルプラットフォームの機能ブロック図である。

【図6】オブジェクト認識システムの機能ブロック図である。

20

【発明を実施するための形態】

【0014】

本明細書全体を通して「一実施形態」、「ある実施形態」、「一例」、または「ある例」への参照は、実施形態または実施例に関連して説明される特定の特徴、構造、または特性が、本発明の少なくとも1つの実施形態に含まれることを意味する。したがって、本明細書全体を通して様々な箇所における「一実施形態では」、または「ある実施形態では」という語句の出現は、必ずしもすべて同じ実施形態を指すわけではない。さらに、特定の特徴、構造、または特性は、1つまたは複数の実施形態において任意の適切な方法で組み合わせられ得る。本明細書に記載の任意の実施例または実施形態は、他の実施例または実施形態よりも好ましい、または有利であると解釈されるべきではない。

30

【0015】

一態様では、ターゲットオブジェクトを追跡する際に環境地図データを使用することによって、あらゆる追跡可能なオブジェクトの追跡堅牢性を改善し、ターゲットの可視範囲を越えて追跡可能なポーズの範囲を拡張し、環境に対して移動しているオブジェクトの追跡性能を高めることさえできるようになる。

【0016】

一実施形態では、環境地図データは地図フィーチャ(map feature)とキーフレーム画像とを含み、各フィーチャは、キーポイントの位置と、その位置の少なくとも1つの対応する記述子とを含む。地図フィーチャは、場面から抽出されたフィーチャ(たとえば、2D/3Dポイント、エッジ、プロブ(blob)など)を含み得る。キーフレーム画像は、場面の画像を含み得、そこから地図フィーチャが抽出される。たとえば、3Dポイントは、ポイントが可視であるキーフレーム間の三角測量によって再構築され得る。地図フィーチャの記述子は、地図ポイントが観察されるキーフレームから抽出される(たとえば、画像パッチ記述子、およびライン記述子)。

40

【0017】

一例では、ターゲットオブジェクトのターゲット地図データはオフラインで生成され、したがって「オフライン地図データ」と呼ばれることがあり、オフライン地図データのキーフレームは、通常、オブジェクトの近距離から生成される。オフライン地図データは、ターゲットの検出および追跡を実行するために十分に「良い」と考えられる。ユーザがターゲットオブジェクトとの対話を開始すると、そのオフライン地図データが拡張現実アプ

50

リケーションにロードされて、カメラ画像からターゲットオブジェクトが検出されて、追跡される。

【 0 0 1 8 】

ターゲットオブジェクトがカメラビデオフレームから検出されると、それが追跡されて、カメラの6自由度ポーズがリアルタイムに取得される。図1Aは、フィーチャ106を含む、近距離にあるターゲットオブジェクト104を含む場面102の取得された画像100Aを示している。したがって、ターゲットオブジェクト104はカメラ画像100Aの大部分を仮定し、追跡は通常正常に動作する。しかしながら、図1Bに示されるように、カメラがターゲットオブジェクト104から離れるにつれて、画像100B内のターゲット104のスケールはすぐに小さくなってしまふ。オブジェクトのスケールがカメラ画像内で小さくなると、あるいは画像の外に出てしまふと、追跡が失敗する可能性がある。

10

【 0 0 1 9 】

したがって、本明細書に開示された実施形態は、ユーザがオブジェクトの周りでカメラを移動させる間、動的に取得された場面の環境地図データを使用することによって、追跡におけるこの制限を克服する。環境のキーフレームが環境地図データに追加されると、追跡は堅牢になり、安定する。一実施形態では、環境地図データは、「オンライン地図データ」と呼ばれ得る。

【 0 0 2 0 】

ユーザがオブジェクトとその環境の周りにカメラを移動させると、新しいキーフレームが環境地図データに追加される。従来のSLAMフレームワークは、候補のキーフレームのポーズ(すなわち、現在のカメラ画像)と、既存のキーフレームのポーズとを比較することによって、キーフレームを追加する。候補のキーフレームのポーズが既存のキーフレームのポーズの1つに類似している場合、それは無視される。このように、従来のSLAMフレームワークは、近い視点を有するあまりにも多くの重複キーフレームを作成することを回避する。しかしながら、この従来の方式はオンラインキーフレームを追加すること妨げ、オフライン地図を迅速に拡張することを困難にする。

20

【 0 0 2 1 】

したがって、本開示のいくつかの実施形態は、オンラインキーフレームを追加するときにポーズ比較ステップでオフラインキーフレームをスキップすることによってこの問題に対処する。環境地図データを生成する場合、このポーズの比較においてオンラインキーフレームだけが考慮される。最初のオンラインキーフレームが追加されようとしている場合、比較するオンラインキーフレームがない。この場合、すべてのオフラインキーフレームとのポーズ差が計算されて、最大値がポーズ差測定値として使用される。一実施形態では、作成後にオフラインキーフレームの多くが破棄されて、3Dポイントの地図および記述子だけが保持される。したがって、この例では、オンライン地図データは、任意のオフラインキーフレームなしに構築され、したがって、最初のオンラインキーフレームが自動的に受け入れられて、環境地図に追加される。図2は、オンライン地図データを構築する例示的なプロセス200を示す流れ図である。

30

【 0 0 2 2 】

処理ブロック205で、キーフレームの作成が始まる。既存のオンラインキーフレームが環境地図データに含まれていない場合、プロセス200は処理ブロック220に進み、オフラインキーフレームを使用してポーズ差が計算される。しかしながら、決定ブロック210で、実際に既存のオンラインキーフレームがあると決定されると、次いで、処理ブロック215は、オンラインキーフレームだけからのポーズ差(たとえば、ポーズ差)を計算することを含む。決定ブロック225で、ポーズ差が十分に大きいかどうかを決定するために、計算されたポーズ差がポーズしきい値 $POSE_{TH}$ と比較される。ポーズ差が十分に大きい場合、環境地図データ内に新しいキーフレームが作成される(すなわち、処理ブロック230)。ポーズ差が十分に大きくはない場合、プロセス200は処理ブロック235に進み、キーフレーム作成が終了する。プロセス200は、カメラが場面を移動するにつれて、キャプチャされた画像のうちの1つまたは複数のために繰り返され得る。

40

50

【 0 0 2 3 】

新しいキーフレームが既存の環境地図データに追加されると、地図データをグローバルに最適化するために、すべての地図フィーチャの座標とすべてのキーフレームの6自由度ポーズがSLAMフレームワーク内で更新される。したがって、キーフレームが追加されると、オブジェクト表面上の地図フィーチャは最適化の結果に応じてわずかに変化する。この戦略は、地図の品質をグローバルな方法で向上させるために有用である。しかしながら、オフライン地図データを修正することは、近距離でのターゲットの検出および追跡の品質を低下させる可能性があるため、現在の実施形態では望ましいものではない。これを防止するために、本明細書に開示された実施形態は、オフライン地図データを「定数」として設定し、プロセス200においてオフライン地図データを更新しない。

10

【 0 0 2 4 】

図3は、場面内のターゲットオブジェクトの追跡のプロセス300を示す流れ図である。プロセス300は、ターゲットオブジェクトがすでに検出されており、追跡が実装されている、処理ブロック305で開始する。したがって、処理ブロック305で次の画像が受信される。次に、処理ブロック310で、オブジェクトトラッカ(たとえば、追跡アルゴリズム)が、検出されたオブジェクトを追跡して、カメラに対するターゲットの推定されたポーズを提供しようとすることができる。追跡が成功した場合、オブジェクトトラッカは、この推定されたターゲットポーズをARエンジンに渡して、ARエンジンは、拡張現実などの他のジョブを実行し得る(すなわち、処理ブロック345)。

20

【 0 0 2 5 】

処理ブロック310と並列に実行されて、プロセス300は、環境に対するターゲットの推定されたポーズを計算するステップを含む、処理ブロック325~335を含む。たとえば、処理ブロック325で、環境の環境地図データが、上述のSLAM技法を適用することによって、受信された画像に基づいて構築される。図2のプロセス200は、処理ブロック325の1つの可能な実装形態である。次いで、処理ブロック330で、カメラに対する環境のポーズを推定するために環境地図データが使用される。次いで、処理ブロック335で、環境内のターゲットポーズの推定を提供するために、処理ブロック310の推定されたターゲットポーズと、処理ブロック330の推定された環境ポーズとがともに使用され、それらはより多くの画像が受信されるにつれて改良され得る。一例では、改良は、ターゲットが環境に対して移動された場合を取り扱うことを含み得る(たとえば、移動するターゲットなどのように)。

30

【 0 0 2 6 】

決定ブロック315で、追跡が失敗した場合、プロセス300は処理ブロック340に進み、カメラに対するターゲットのポーズが、処理ブロック335の推定されたターゲットポーズに基づいて推測される。次いで、処理ブロック345で、推測されたターゲットポーズがARエンジンに渡される。

【 0 0 2 7 】

いくつかの実施形態では、ターゲットの追跡が不安定になる場合、たとえば、小さいサイズまたは画像内のフィーチャのカバレッジ不良による低精度のためにジッタになると、堅牢性を向上させるために、推測されたターゲットポーズと実際のターゲットポーズとが融合され得る。推測されたターゲットポーズと実際のターゲットポーズとの融合は、カルマンフィルタなどのフィルタを適用することを含み得る。

40

【 0 0 2 8 】

図4は、オブジェクト検出および追跡のための処理ユニット400の機能ブロック図である。一実施形態では、処理ユニット400は、プログラムコードの指示の下で、上述のプロセス200および/または300を実行し得る。たとえば、画像402の時間的なシーケンスが処理ユニット400によって受信される。オブジェクト検出器406は、画像のうちの少なくとも1つに含まれるターゲットオブジェクトを検出して、オブジェクトトラッカ408が画像402のシーケンスに基づいてオブジェクトを追跡する。画像が取得されると、地図データ生成器404は、選択キーフレームを環境地図データ414に追加し得る(たとえば、プロセス200参照)。

50

【 0 0 2 9 】

オブジェクト検出器406は、受信された画像に基づいて、およびターゲット地図データ416に基づいて、ターゲットオブジェクトに対するカメラの現在のポーズを決定し得る。上述のように、オブジェクトトラッカ408がターゲットオブジェクトを追跡すると、ポーズ計算機412が、並行して、環境に対するターゲットポーズの推定を生成し得る。ターゲットオブジェクトの追跡が失われると、オブジェクトトラッカ408は、ポーズ計算機412によって生成されたターゲットポーズに基づいて、カメラに対するターゲットポーズを推測し得る。

【 0 0 3 0 】

拡張現実(AR)エンジン410は、オブジェクト検出器406および/またはオブジェクトトラッカ408によって決定されたカメラポーズに基づいて、拡張現実に関連する任意の動作を実行し得る。

【 0 0 3 1 】

図5は、本明細書に記載のプロセスを実行することが可能なモバイルプラットフォーム500の機能ブロック図である。本明細書で使用するモバイルプラットフォームは、セルラーまたは他のワイヤレス通信デバイス、パーソナル通信システム(PCS)デバイス、パーソナル航法デバイス(PND)、個人情報マネージャ(PIM)、携帯情報端末(PDA)、ラップトップ、あるいはワイヤレス通信および/または航法位置決め信号などの航法信号を受信することが可能な他の適したモバイルデバイスなどのデバイスを指す。「モバイルプラットフォーム」という用語は、衛星信号受信、支援データ受信、および/または位置関連処理がデバイスで発生するかパーソナル航法デバイス(PND)で発生するかに関わらず、短距離ワイヤレス、赤外線、有線接続、または他の接続などによって、PNDと通信するデバイスを含むことも意図する。また、「モバイルプラットフォーム」は、インターネット、Wi-Fi、または他のネットワークを介するなどしてサーバと通信することが可能である、および、衛星信号受信、支援データ受信、および/または位置関連処理が、デバイスで、サーバで、またはネットワークに関連付けられる別のデバイスで発生するかに関わらず、ワイヤレス通信デバイス、コンピュータ、ラップトップ等を含む、すべてのデバイスを含むことを意図する。さらに、「モバイルプラットフォーム」は、拡張現実(AR)、仮想現実(VR)、および/または複合現実(MR)アプリケーションが可能なすべての電子デバイスも含み得る。上記の任意の動作可能な組合せも「モバイルプラットフォーム」と考えられる。

【 0 0 3 2 】

モバイルプラットフォーム500は、カメラ502、ならびにカメラ502によってキャプチャされた画像を表示することが可能なディスプレイ522を含む任意のユーザインターフェース506を任意で含み得る。ユーザインターフェース506はまた、キーパッド524、またはユーザが情報をモバイルプラットフォーム500に入力できる他の入力デバイスを含み得る。必要に応じて、キーパッド524は、タッチセンサ付きディスプレイ522に仮想キーパッドを統合することによって除去され得る。ユーザインターフェース506はまた、マイクロフォン526とスピーカ528とを含み得る。

【 0 0 3 3 】

モバイルプラットフォーム500はまた、カメラ502およびユーザインターフェース506がある場合はそれらに接続され、またそれらと通信する、制御ユニット504を含む。制御ユニット504は、カメラ502から、および/またはネットワークアダプタ516から受信された画像を受け入れて、処理する。制御ユニット504は、処理ユニット508および関連メモリ514、ハードウェア510、ソフトウェア515、ならびにファームウェア512によって提供され得る。

【 0 0 3 4 】

図4の処理ユニット400は、上述のように、オブジェクト検出および追跡のための処理ユニット508の1つの可能な実装形態である。制御ユニット504は、たとえば、必要に応じてディスプレイ522に所望のデータをレンダリングするためのゲームエンジンであり得る、グラフィックスエンジン520をさらに含み得る。処理ユニット508およびグラフィックスエ

10

20

30

40

50

ンジン520は、明確にするために別々に示されているが、単一のユニットでもよく、および/または、処理ユニット508内で実行されているソフトウェア515内の命令に基づいて処理ユニット508に実装されてもよい。処理ユニット508、ならびにグラフィックスエンジン520は、必ずしも必要ないが、1つまたは複数のマイクロプロセッサ、組込みプロセッサ、コントローラ、特定用途向け集積回路(ASIC)、デジタル信号プロセッサ(DSP)などを含み得る。プロセッサおよび処理ユニットという用語は、特定のハードウェアではなくシステムによって実装される機能を説明する。さらに、本明細書で使用されるように、「メモリ」という用語は、長期メモリ、短期メモリ、またはモバイルプラットフォーム500に関連付けられる他のメモリを含む、任意のタイプのコンピュータストレージ媒体を指し、任意の特定のタイプまたは数のメモリや、メモリが記憶されるタイプの媒体に限定されない。

10

【0035】

本明細書に記載のプロセスは、アプリケーションに応じて様々な手段によって実装され得る。たとえば、これらのプロセスは、ハードウェア510、ファームウェア512、ソフトウェア515、またはそれらの任意の組合せに実装され得る。ハードウェア実装形態の場合、処理ユニットは、1つまたは複数の特定用途向け集積回路(ASIC)、デジタル信号プロセッサ(DSP)、デジタル信号処理デバイス(DSP)、プログラマブルロジックデバイス(PLD)、フィールドプログラマブルゲートアレイ(FPGA)、プロセッサ、コントローラ、マイクロコントローラ、マイクロプロセッサ、電子デバイス、本明細書に記載の機能を実行するように設計された他の電子ユニット、またはそれらの組合せ内に実装され得る。

【0036】

20

ファームウェアおよび/またはソフトウェア実装形態の場合、プロセスは、本明細書に記載の機能を実行するモジュール(たとえば、プロシージャ、関数等)で実装され得る。命令を具体的に実施する任意のコンピュータ可読媒体は、本明細書に記載のプロセスを実装する際に使用され得る。たとえば、プログラムコードはメモリ515に記憶されて、処理ユニット508によって実行され得る。メモリは、処理ユニット808内に実装されてもよく、処理ユニット508内の外部に実装されてもよい。

【0037】

ファームウェアおよび/またはソフトウェア内に実装されると、機能は1つまたは複数の命令あるいはコードとしてコンピュータ可読媒体上に記憶され得る。この例には、データ構造によって符号化された非一時的コンピュータ可読媒体、およびコンピュータプログラムによって符号化されたコンピュータ可読媒体が含まれる。コンピュータ可読媒体には、物理的コンピュータストレージ媒体が含まれる。ストレージ媒体は、コンピュータによってアクセスされ得る任意の入手可能な媒体であり得る。例として、限定はしないが、そのようなコンピュータ可読媒体は、RAM、ROM、フラッシュメモリ、EEPROM、CD-ROM、または他の光ディスクストレージ、磁気ディスクストレージまたは他の磁気ストレージデバイス、あるいは所望のプログラムコードを命令またはデータ構造の形で記憶するために使用され得、またコンピュータからアクセスされ得る他の任意の媒体を備え得、本明細書で使用する場合、ディスク(disk)およびディスク(disc)には、コンパクトディスク(CD)、レーザーディスク(登録商標)、光ディスク、デジタル多用途ディスク(DVD)、フロッピー(登録商標)ディスク、およびブルーレイディスクが含まれ、ディスク(disk)は通常、データを磁気的に再生し、一方、ディスク(disc)はデータをレーザーによって光学的に再生する。前記の組合せも、コンピュータ可読媒体の範囲内に含まれるべきである。

30

40

【0038】

図6は、オブジェクト認識システム600の機能ブロック図である。図示されるように、オブジェクト認識システム600は、オブジェクト614を含む場面の画像をキャプチャすることが可能なカメラを含む(現在のビューには図示されていない)、例示的なモバイルプラットフォーム602を含む。フィーチャデータベース612は、環境(オンライン)地図データとターゲット(オフライン)地図データとを含む、上述の地図データのうちのいずれかを含み得る。

【0039】

50

モバイルプラットフォーム602は、カメラによってキャプチャされた画像を表示するためのディスプレイを含み得る。モバイルプラットフォーム602はまた、たとえば、衛星ビークル606、あるいはセルラータワー604またはワイヤレス通信アクセスポイント605を含む、位置を決定するための他の任意の適切なソースを含む、衛星測位システム(SPS)からの信号を使用してその緯度および経度を決定することに基づいて、航法のために使用され得る。モバイルプラットフォーム602はまた、モバイルプラットフォーム602の方位を決定するために使用され得る、デジタルコンパス、加速度計、またはジャイロ스코ープなどの、方位センサを含み得る。

【0040】

衛星測位システム(SPS)は、通常、エンティティが、送信機から受信された信号に少なくとも部分的に基づいて、地球上のまたは地球上空のそれらの位置を決定することを可能にするように配置された送信機のシステムを含む。そのような送信機は、通常、設定された数のチップの繰返し擬似ランダム雑音(PN)コードでマークされた信号を送信し、地上ベースの制御局、ユーザ機器、および/または宇宙船上に配置され得る。特定の例では、そのような送信機は、地球周回衛星ビークル(SV)606に配置され得る。たとえば、全地球測位システム(GPS)、ガリレオ、グロナス、またはコンパスなどの全地球的航法衛星システム(GNSS)のコンステレーション内のSVは、コンステレーション内の他のSVによって送信されたPNコードから区別可能なPNコードでマークされた信号を送信することができる(たとえば、GPSにおけるように衛星ごとに異なるPNコードを使用するか、グロナスにおけるように異なる周波数上で同じコードを使用して)。

【0041】

特定の態様によれば、本明細書で提示される技法は、SPSのための地球規模のシステム(たとえばGNSS)には限定されない。たとえば、本明細書で提供される技法は、たとえば、日本の準天頂衛星システム(QZSS)、インドのインド地域航法衛星システム(IRNSS)、中国の北斗などのような様々な地域システム等、および/あるいは、1つまたは複数の全地球および/または地域航法衛星システムと関連し得る、あるいは場合によってはこれらとともに使用できるようにされ得る、様々な補強システム(たとえば、静止衛星型衛星航法補強システム(SBAS))に対して適用されてもよく、またはそれらのシステムにおいて使用できるようにされてもよい。限定ではなく例として、SBASは、広域補強システム(WAAS)、欧州静止衛星航法オーバーレイサービス(EGNOS)、多機能衛星補強システム(MSAS)、GPS支援静止補強ナビゲーションまたはGPSおよび静止補強ナビゲーションシステム(GAGAN)、ならびに/または同様のものなどの、完全性情報、微分補正などを提供する補強システムを含み得る。したがって、本明細書で使用する場合、SPSは1つもしくは複数の全地球および/または地域航法衛星システムならびに/あるいは補強システムの任意の組合せを含み得、またSPS信号はSPS信号、SPS様信号、および/またはそのような1つもしくは複数のSPSに関連する他の信号を含み得る。

【0042】

位置決定技法は、ワイヤレスワイドエリアネットワーク(WWAN)、ワイヤレスローカルエリアネットワーク(WLAN)、ワイヤレスパーソナルエリアネットワーク(WPAN)など、セルラータワー604を含んでおりワイヤレス通信アクセスポイント605からの様々なワイヤレス通信ネットワークに関連して実装され得るので、モバイルプラットフォーム602は、位置決定のためのSPSでの使用に限定されるものではない。さらに、モバイルプラットフォーム602は、セルラータワー604を介して、またワイヤレス通信アクセスポイント605から様々なワイヤレス通信ネットワークを使用して、あるいは、必要に応じて衛星ビークル606を使用して、データベース612からオンライン地図データおよび/またはオフライン地図データなどのデータを取得するために、1つまたは複数のサーバ608にアクセスし得る。「ネットワーク」および「システム」という用語は、しばしば互換可能に使用される。WWANは、符号分割多元接続(CDMA)ネットワーク、時分割多元接続(TDMA)ネットワーク、周波数分割多元接続(FDMA)ネットワーク、直交周波数分割多元接続(OFDMA)ネットワーク、シングルキャリア周波数分割多元接続(SC-FDMA)ネットワーク、ロングタームエボリューション(LTE)

などであり得る。CDMAネットワークは、cdma2000、Wideband-CDMA(W-CDMA(登録商標))などの1つまたは複数の無線アクセス技術(RAT)を実装し得る。cdma2000はIS-95標準、IS-2000標準、IS-856標準を含む。TDMAネットワークは、グローバルシステムフォーモバイルコミュニケーションズ(GSM(登録商標))、デジタルアドバンストモバイルフォンシステム(D-AMPS)、または何らかの他のRATを実装し得る。GSM(登録商標)およびW-CDMA(登録商標)は、「第3世代パートナーシッププロジェクト」(3GPP)という名称の組織からの文書に記述されている。cdma2000は、「第3世代パートナーシッププロジェクト2」(3GPP2)という名称の組織からの文書に記載されている。3GPP文書および3GPP2文書は公的に入手可能である。WLANはIEEE802.11xネットワークであり得、WPANはブルートゥース(登録商標)ネットワーク、IEEE802.15x、または他の何らかのタイプのネットワークであり得る。また、これらの技法は、WWAN、WLAN、および/またはWPANの任意の組合せとともに、実装され得る。

10

【0043】

図6に示されるように、システム600は、フィーチャデータベース612に含まれる地図データに基づいて検出されて追跡されるべきオブジェクト614の画像をキャプチャするモバイルプラットフォーム602を含む。図示されるように、モバイルプラットフォーム602は、たとえばセルラータワー604またはワイヤレス通信アクセスポイント605を介して、サーバ608に結合されたワイヤレスワイドエリアネットワーク(WWAN)などのネットワーク610にアクセスすることができ、サーバ608は、ターゲットオブジェクトおよびそれらの画像に関連する情報を記憶するデータベース612に接続されている。図6は1つのサーバ608を示しているが、複数のサーバ、ならびに複数のデータベース612が使用され得ることが理解されるべきである。モバイルプラットフォーム602は、サーバ608からデータベース612のうちの少なくとも一部を取得して、ダウンロードされた地図データをモバイルプラットフォーム602内のローカルデータベースに記憶することによって、図6に示されるように、オブジェクト検出および追跡自体を実行し得る。サーバ608から取得されたデータベースの一部は、モバイルプラットフォームの測位システムによって決定されたモバイルプラットフォームの地理的位置に基づき得る。さらに、サーバ608から取得されたデータベースの一部は、モバイルプラットフォーム602上のデータベースを必要とする具体的な適用例に依存し得る。モバイルプラットフォーム602は、キャプチャされたクエリ画像からフィーチャを抽出して、クエリフィーチャを、ローカルデータベースに記憶されたフィーチャに一致させ得る。クエリ画像は、カメラからのプレビューフレーム内の画像でもよく、カメラによってキャプチャされた画像でもよく、ビデオシーケンスから抽出されたフレームでもよい。オブジェクト検出は、クエリフィーチャごとに決定された信頼レベルに少なくとも部分的に基づき得、次いで異常値の除去に使用され得る。モバイルプラットフォームの地理的位置に基づいてデータベース612のわずかな部分をダウンロードすることと、モバイルプラットフォーム602上でオブジェクト検出を実行することとによって、ネットワーク遅延の問題を回避することができ、オーバーザエア(OTA)帯域幅の使用量が、クライアント(すなわ

20

30

ち、モバイルプラットフォーム)側のメモリ要件とともに減少される。しかしながら、必要に応じて、オブジェクト検出および追跡は、サーバ608(または、他のサーバ)によって実行され得、クエリ画像自体、またはクエリ画像から抽出されたフィーチャは、モバイルプラットフォーム602によってサーバ608に提供される。一実施形態では、オンライン地図データは、モバイルプラットフォーム602によってローカルに記憶され、オフライン地図データは、データベース612内のクラウドに記憶される。

40

【0044】

処理ブロックのうちのいくつかまたはすべてが上述の各プロセスに表示される順序は、限定されるべきではない。むしろ、本開示の利益を有する当業者は、処理ブロックのうちのいくつかは、示されていない様々な順序で実行され得ることを理解するであろう。

【0045】

当業者は、本明細書に開示された実施形態に関連して説明される様々な例示的な論理ブ

50

ロック、モジュール、エンジン、回路、およびアルゴリズムステップが、電子ハードウェア、コンピュータソフトウェア、または両方の組合せとして実装され得ることをさらに理解するであろう。ハードウェアとソフトウェアとの、この互換性を明確に示すために、様々な例示的な構成要素、ブロック、モジュール、エンジン、回路、およびステップについて、上記では概してそれらの機能に関して説明した。そのような機能がハードウェアとして実装されるか、またはソフトウェアとして実装されるかは、具体的な適用例および全体的なシステムに課される設計制約に依存する。当業者は、説明された機能を具体的な適用例ごとに様々な方法で実装することができるが、そのような実装の決定は、本発明の範囲からの逸脱を生じるものと解釈されるべきではない。

【 0 0 4 6 】

10

本明細書に開示された実施形態への様々な修正が当業者には容易に明らかになり、本明細書で定義された一般原理は、本発明の趣旨または範囲から逸脱することなしに他の実施形態に適用され得る。したがって、本発明は、本明細書に示される実施形態に限定されることが意図されるものではなく、本明細書に開示された原理および新規のフィーチャに矛盾しない最も広い範囲を与えられるべきである。

【 符号の説明 】

【 0 0 4 7 】

100A 画像
 100B 画像
 102 場面
 104 ターゲットオブジェクト
 106 フィーチャ
 200 プロセス
 300 プロセス
 400 処理ユニット
 402 画像
 404 地図データ生成器
 406 オブジェクト検出器
 408 オブジェクトトラッカ
 410 拡張現実(AR)エンジン
 412 ボーズ計算機
 414 環境地図データ
 416 ターゲット地図データ
 500 モバイルプラットフォーム
 502 カメラ
 504 制御ユニット
 506 ユーザインターフェース
 508 処理ユニット
 510 ハードウェア
 512 ファームウェア
 514 関連メモリ
 515 ソフトウェア
 516 ネットワークアダプタ
 520 グラフィックスエンジン
 522 ディスプレイ
 524 キーパッド
 526 マイクロフォン
 528 スピーカ
 600 オブジェクト認識システム
 602 モバイルプラットフォーム

20

30

40

50

- 604 セルラータワー
- 605 ワイヤレス通信アクセスポイント
- 606 地球周回衛星ビークル
- 608 サーバ
- 610 ネットワーク
- 612 フィーチャデータベース
- 614 オブジェクト

【図 1 A】

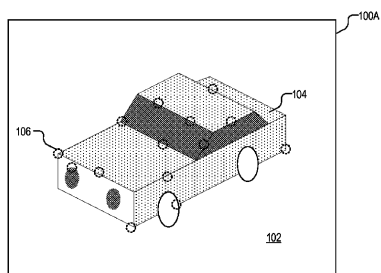


FIG. 1A

【図 1 B】

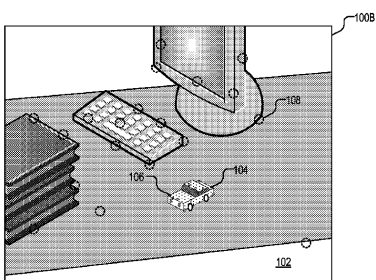
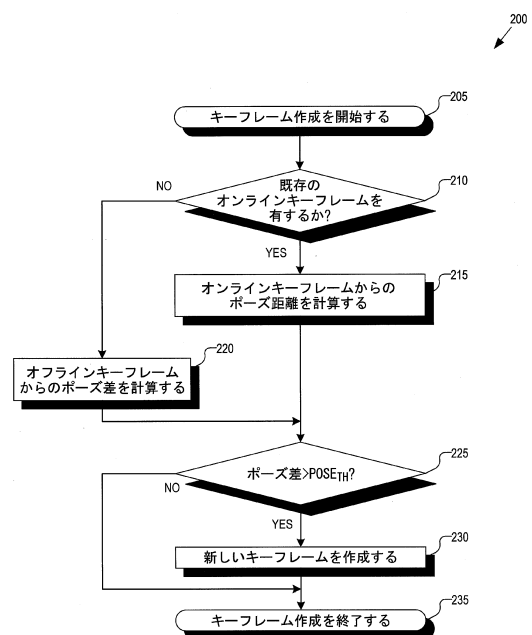
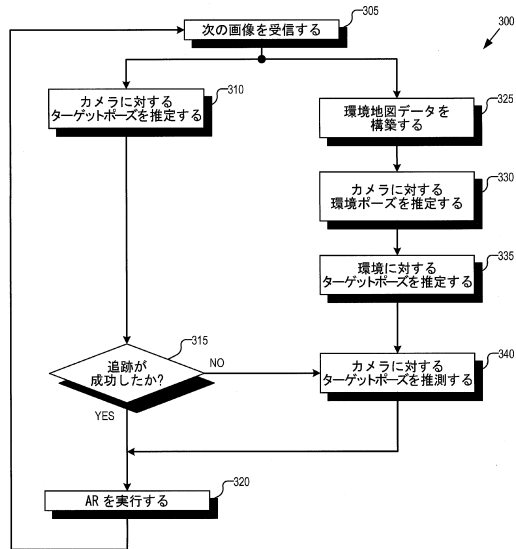


FIG. 1B

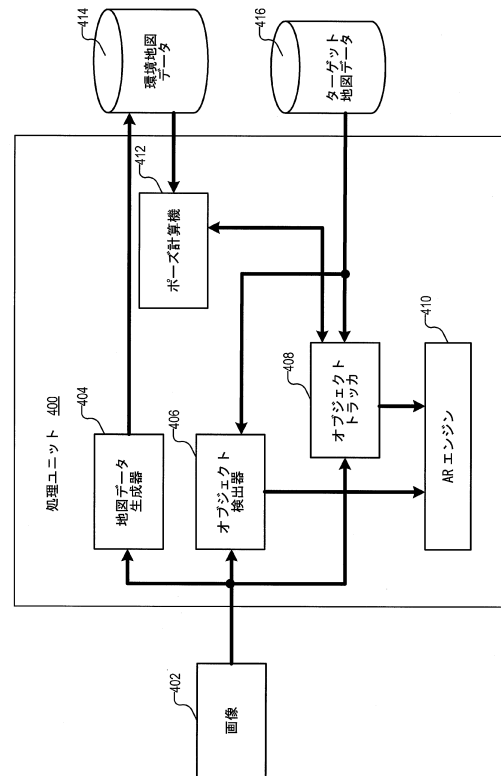
【図 2】



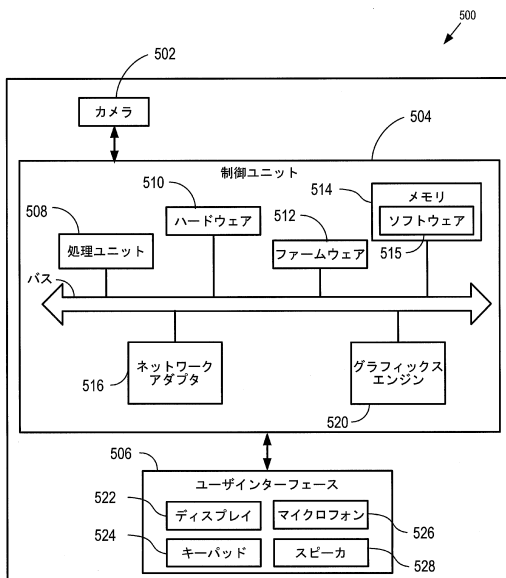
【図 3】



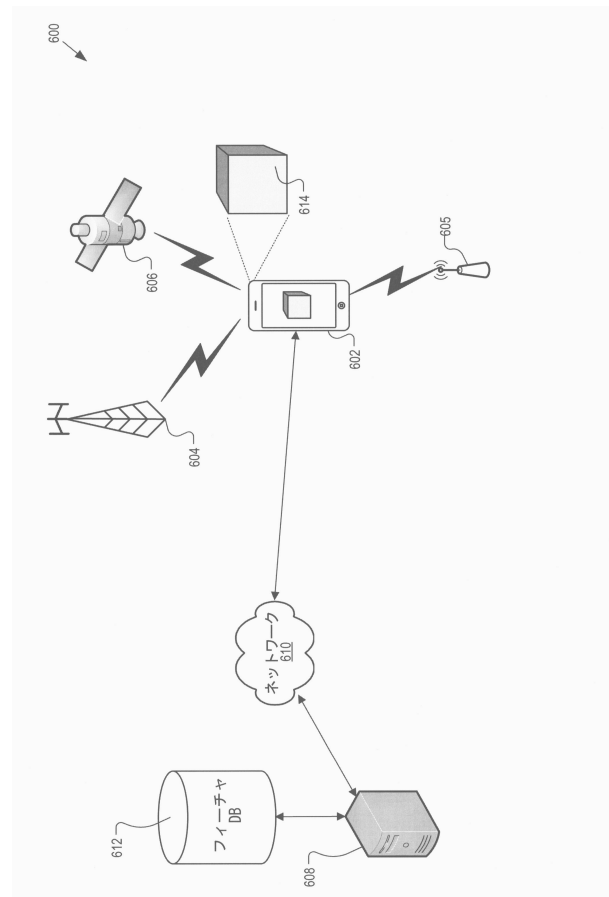
【図 4】



【図 5】



【図 6】



フロントページの続き

(72)発明者 チ・パン

アメリカ合衆国・カリフォルニア・9 2 1 2 1 - 1 7 1 4・サン・ディエゴ・モアハウス・ドライ
ヴ・5 7 7 5

(72)発明者 マイケル・ゲルヴォーツ

アメリカ合衆国・カリフォルニア・9 2 1 2 1 - 1 7 1 4・サン・ディエゴ・モアハウス・ドライ
ヴ・5 7 7 5

審査官 佐藤 実

(56)参考文献 一刈良介 外5名, マーカレス・カメラ・トラッキングを強化するリハーサル・パス法とその実
世界運用, 日本バーチャルリアリティ学会論文誌, 日本バーチャルリアリティ学会, 2 0 1 2 年
9 月 3 0 日, 第17巻 第3号, 第241 - 252頁

(58)調査した分野(Int.Cl., D B 名)

G 0 6 T 7 / 2 0