

(19) 日本国特許庁(JP)

(12) 公表特許公報(A)

(11) 特許出願公表番号

特表2010-532465  
(P2010-532465A)

(43) 公表日 平成22年10月7日(2010.10.7)

(51) Int.Cl.		F I	テーマコード (参考)
<b>GO1B</b>	<b>11/26</b>	<b>(2006.01)</b>	GO1B 11/26 H 2F065
<b>GO6T</b>	<b>1/00</b>	<b>(2006.01)</b>	GO6T 1/00 315 5B050
<b>GO6T</b>	<b>17/40</b>	<b>(2006.01)</b>	GO6T 17/40 G 5B057
<b>GO1B</b>	<b>11/00</b>	<b>(2006.01)</b>	GO1B 11/00 H

審査請求 有 予備審査請求 未請求 (全 28 頁)

(21) 出願番号 特願2010-509290 (P2010-509290)  
 (86) (22) 出願日 平成20年5月22日 (2008.5.22)  
 (85) 翻訳文提出日 平成22年1月7日 (2010.1.7)  
 (86) 国際出願番号 PCT/N02008/000178  
 (87) 国際公開番号 W02008/143523  
 (87) 国際公開日 平成20年11月27日 (2008.11.27)  
 (31) 優先権主張番号 20072616  
 (32) 優先日 平成19年5月22日 (2007.5.22)  
 (33) 優先権主張国 ノルウェー (NO)

(71) 出願人 509321549  
 メタイオ ゲゼルシャフト ミット ベシ  
 ユレンクテル ハフツング  
 ドイツ, ミュンヘン, インファンテリーシ  
 ユトラーセ 19, ハウス 3, 2 オー  
 ゲー  
 (74) 代理人 100096459  
 弁理士 橋本 剛  
 (74) 代理人 100092613  
 弁理士 富岡 潔  
 (72) 発明者 エンゲダル, トールビョルン  
 ノルウェー, クリスティアンサン エス,  
 スヴァルトヒュンヴァイアン 16

最終頁に続く

(54) 【発明の名称】 拡張リアリティイメージのためのカメラポーズ推定装置および方法

(57) 【要約】

シーンに対する3Dカメラポーズを、該カメラから得た2D画像フレームの2D画像データから推定する装置である。候補2Dキーポイント・デテクターは、2D画像フレームから候補2Dキーポイントを決定する。デテクテド・3D観測デテクターは、予記録シーンデータと候補2Dキーポイントとから、デテクテド・3D観測を決定する。デテクテド・3Dカメラポーズ・エスティメータは、カメラデータとデテクテド・3D観測と候補2Dキーポイントとから、デテクテド・3Dカメラポーズ推定を出力する。第1ストレージは、デテクテド・候補2Dキーポイントと2D画像データとを記憶し、かつ3Dカメラポーズ推定出力にตอบสนองして、前の3Dカメラポーズ推定出力に関連した前の2D画像データおよび候補2Dキーポイントを出力する。第2ストレージは、前の3Dカメラポーズ推定を記憶し、かつ出力する。トラックド・3D観測デテクターは、2D画像データと、候補2Dキーポイントと、カメラデータと、前の2D画像データおよび候補2Dキーポイントと、前の3Dカメラポーズ推定と、3Dシーンモデルデータと、から、トラッ

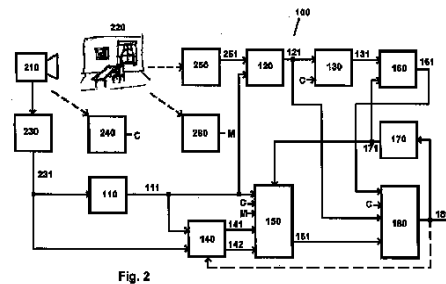


Fig. 2

## 【特許請求の範囲】

## 【請求項 1】

シーンに対する 3 D カメラポーズを、該カメラから得た 2 D 画像フレームの 2 D 画像データから推定する装置であって、

( a ) 入力としてカメラデータソースからのカメラデータと 2 D 画像データとを受けるとともに、候補 2 D キーポイントを出力する候補 2 D キーポイント・デテクターと、

( b ) 入力として予記録シーンデータと候補 2 D キーポイントとを受けるとともに、デテクテド・3 D 観測を出力するデテクテド・3 D 観測デテクターと、

( c ) 入力としてカメラデータとデテクテド・3 D 観測と候補 2 D キーポイントとを受けるとともに、デテクテド・3 D カメラポーズ推定を出力するデテクテド・3 D カメラポーズ・エスティメータと、

10

( d ) 入力として検出した候補 2 D キーポイントと 2 D 画像データとを受けるとともに、制御入力として 3 D カメラポーズ推定出力に関連した制御信号を受け、3 D カメラポーズ推定出力に関連する 2 D 画像データおよび候補 2 D キーポイントを記憶するとともに、前の 3 D カメラポーズ推定出力に関連した前の 2 D 画像データおよび候補 2 D キーポイントを出力する、2 D 画像データおよび候補 2 D キーポイントのストレージと、

( e ) 入力として 3 D カメラポーズ推定出力を受け、かつこの 3 D カメラポーズ推定出力を記憶するとともに、前の 3 D カメラポーズ推定出力を出力するように構成された前の 3 D カメラポーズ推定のストレージと、

( f ) 入力として、2 D 画像データと、候補 2 D キーポイントと、カメラデータと、前の 2 D 画像データおよび候補 2 D キーポイントと、前の 3 D カメラポーズ推定と、3 D シーンモデルデータソースからの 3 D シーンモデルデータと、を受けるとともに、出力としてトラックド・3 D 観測を提供するトラックド・3 D 観測デテクターと、

20

( g ) 入力としてデテクテド・カメラポーズ推定と前の 3 D カメラポーズ推定を受け、これらのデテクテド・3 D カメラポーズ推定と前の 3 D カメラポーズ推定の一方を、選択後 3 D カメラポーズ推定として出力するように構成されたポーズ推定セクタと、

( h ) 入力として、カメラデータと、デテクテド・3 D 観測と、トラックド・3 D 観測と、選択後 3 D カメラポーズ推定と、を受けるとともに、出力として、3 D カメラポーズ推定出力を提供する 3 D カメラポーズ・エスティメータと、

を備えてなる装置。

30

## 【請求項 2】

上記の予記録シーンデータの受け取りは、予トレーニングした 2 D キーポイント分類ツリーあるいは 2 D キーポイントデータベースからなることを特徴とする請求項 1 に記載の装置。

## 【請求項 3】

上記の予トレーニングした 2 D キーポイント分類ツリーあるいは 2 D キーポイントデータベースは、2 D キーポイントを該 2 D キーポイントの 3 D 座標にリンクさせるように構成され、上記 3 D 座標は 3 D シーンモデルの座標系に参照されるものであることを特徴とする請求項 2 に記載の装置。

## 【請求項 4】

上記カメラデータは、カメラレンズの焦点距離およびカメラの主点を含むことを特徴とする請求項 1 ~ 3 のいずれかに記載の装置。

40

## 【請求項 5】

上記デテクテド・3 D カメラポーズ・エスティメータは上記デテクテド・3 D カメラポーズ推定とともに第 1 の推定品質指標を出力するように構成され、上記 3 D カメラポーズ・エスティメータは上記 3 D カメラポーズ推定とともに第 2 の推定品質指標を出力するように構成され、上記 3 D カメラポーズセクタは、第 1 および第 2 の推定品質指標に応じて選択後 3 D カメラポーズ推定を出力するように構成されていることを特徴とする請求項 1 ~ 4 のいずれかに記載の装置。

## 【請求項 6】

50

上記の前の2D画像データ、前の候補2Dキーポイント、前の3Dカメラポーズ出力は、いずれも、上記2D画像データに先行する直前の2D画像フレームに関連するものであることを特徴とする請求項1～5のいずれかに記載の装置。

【請求項7】

上記デテクト・3D観測デテクターおよびトラックド・3D観測デテクターは、デテクト・3D観測についてのスコアおよびトラックド・3D観測についてのスコアを決定し、かつこれらのデテクト・3D観測スコアおよびトラックド・3D観測スコアをそれぞれデテクト・3D観測およびトラックド・3D観測とともに出力するように構成されていることを特徴とする請求項1～6のいずれかに記載の装置。

【請求項8】

上記3Dカメラポーズ・エスティメータは、

(i) デテクト・3D観測スコアおよびトラックド・3D観測スコアを正規化する手段と、

(ii) これらの正規化したスコアに応じて選択した上記デテクト・3D観測およびトラックド・3D観測の一方を提供する手段と、

を含むことを特徴とする請求項7に記載の装置。

【請求項9】

上記3Dカメラポーズ・エスティメータは、該3Dカメラポーズ・エスティメータに入力された選択後3Dカメラポーズ推定の有効性を判定し、入力された選択後3Dカメラポーズ推定が有効でないとは判定した場合には、3Dカメラポーズ推定として前の3Dカメラポーズ推定を維持するように構成されていることを特徴とする請求項1～8のいずれかに記載の装置。

【請求項10】

メモリを有するコンピュータ装置を、請求項1～9のいずれかに記載の装置として構成するように機能するコンピュータプログラム製品。

【請求項11】

第1オブジェクトを画像化するカメラによって見たときに第1オブジェクトに対し静止しているとみなされる第2オブジェクトの画像を生成する方法であって、

(i) カメラによる第1オブジェクトの $n$ 番目の画像の取得に基づき、 $n$ 番目のグラブド・画像フレーム $IF_n$ を提供する第1ステップと、

(ii) カメラによる第1オブジェクトの $n+1$ 番目の画像の取得に基づき、 $n+1$ 番目のグラブド・画像フレーム $IF_{n+1}$ を提供するとともに、請求項1～10のいずれかに記載の装置を用いて、 $n$ 番目のグラブド・画像フレームから $n$ 番目の第1の新しい3Dカメラポーズ推定 $CP_n$ を生成する第2ステップと、

(iii) カメラによる第1オブジェクトの $n+2$ 番目の画像の取得に基づき、 $n+2$ 番目のグラブド・画像フレーム $IF_{n+2}$ を提供するとともに、請求項1～10のいずれかに記載の装置を用いて、 $n+1$ 番目のグラブド・画像フレームから $n+1$ 番目の第2の新しい3Dカメラポーズ推定 $CP_{n+1}$ を生成し、かつ第2オブジェクトの $n$ 番目の画像フレーム $GF_n$ を、 $n$ 番目の第1の新しい3Dカメラポーズ推定 $CP_n$ により示された3Dカメラ位置におけるカメラから見たものとして生成する第3ステップと、

を備えてなる方法。

【請求項12】

さらに、生成された $n$ 番目の画像フレーム $GF_n$ を、 $n$ 番目、 $n+1$ 番目、もしくは $n+2$ 番目の取得した画像の少なくとも一つと結合することで、 $n$ 番目の結合画像フレーム $CF_n$ を生成することを特徴とする請求項11に記載の方法。

【請求項13】

第1オブジェクトを画像化しつつカメラを動かすとともに上記ステップ(i)(ii)(iii)を $n=1, 2, 3, \dots, k$ について繰り返すことにより、第1オブジェクトを画像化した2Dカメラから見たものとして、第2オブジェクトの動画用の第2オブジェクトの一連の画像を生成することをさらに備えてなる請求項11または請求項12に記載の方法。

10

20

30

40

50

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

この発明は画像の分野に関し、特に、あるシーンに関するカメラのポーズを、カメラが取り込んだシーンの画像から得た情報と、シーンの既知の特徴の既定の位置と、から判定することに関する。

## 【背景技術】

## 【0002】

## 1 問題

解決すべき問題は、リアルタイムに、あるいは記録された一連の画像に基づいて、物理的カメラのポーズを推定することである。このように物理的カメラのポーズを知ろうとする動機は、画像ストリームの中にバーチャルオブジェクトをリアルタイムに埋め込み、拡張リアリティ効果を得ることが可能となるためである。典型的な一式の装置としては、ユーザが身に付けるヘッド装着型のディスプレイ（HMD）およびカメラであり、カメラからの画像が該カメラのポーズについてはHMDのポーズを再構築するために利用される。このポーズは、バーチャルオブジェクトをHMDのスクリーン上に（カメラ画像の上（ビデオ・シースルー）もしくは光学的に透けて見えるスクリーン上のいずれかに）正確に整合させるために利用され、これにより、カメラで撮影された現実のシーンには実際には存在しないバーチャルオブジェクト/エレメントが、ユーザを囲む現実の中の実在のオブジェクトのように現れる。

10

20

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0003】

## 2 従来技術

上記の問題についての従来技術は、カメラ画像からカメラポーズをリアルタイムに判定するために用いられる方法やアルゴリズムが遅く、かつ効率的でないことから、リアルタイムの利用に適しておらず、あるいは、得られた結果の不正確性、低いリピータビリティ、ドリフト、ポーズの回復の利用不可能、ポーズ推定の非ロバスト性、といった欠点があることが知られている。

## 【0004】

30

## 2.1 環境の既知の3Dモデルを用いたトラッキング

トラッキングもしくはフレーム・トゥ・フレーム・トラッキングについては、4.23章に述べている。一般に、フレーム・トゥ・フレーム・トラッキング法は、ドリフトおよび（実際上は）ポーズ法の回復の利用不可能という欠点がある。カメラのごく僅かな動きでは、フレーム・トゥ・フレーム・トラッキングは良好に機能するが、現実のシナリオでは、ユーザがカメラを制御するので、これは、実際に適用することができない制約となる。

実際の世界では、限界のない処理能力や限界のないフレーム速度は利用不可能であるから、フレーム・トゥ・フレーム・トラッキングは、また、ルージング・トラックという欠点があることが知られている。カメラが過度に動くと、次のフレームは、ポーズ推定を実行すべく前のフレームにおける特徴点とマッチし得る十分な特徴点を含まないものとなり得る。カメラの早い動きは、新しいフレームにおけるぼけ（動きによるぼけ）を招き、新しいフレームと信頼性をもってマッチする前のフレームの特徴点の数が減少する結果となる。もしトラッキングがそのポーズを手離してしまうと、トラッキングの再初期化のために、ポーズの回復法を実行する必要がある。種々の方法が提案されかつ発表されているが、その大部分は、ポーズの回復およびトラッキングの再初期化のために、ユーザがカメラを特定の位置や方向へと移動ないし回転させる必要があるものである。仮に適当な回復法を実行したとしても、フレーム・トゥ・フレーム・トラッキングはドリフトの欠点があることが知られている。

40

US 2002/0191862 A1 (US 6,765,569 B1) は、ユーザが特定の方向を見るつまり特定の方向にカメラを向けることによって、システムの利用を開始し、かつこのシステムが使用

50

中の可能な作業領域を拡張するようにした方法を開示している。この開示された方法は、フレーム・トゥ・フレーム・トラッキングに類似しているが、先の既知の3Dモデルを利用しない方法であり、つまり、特徴点の3D位置を探索するために後続のフレームにおける特徴点の位置の三角法に依存している。この方法は、使用中に特徴点の3D情報を記憶し、かつ特徴点が発見されるたびに、三角法に基づいて、3D位置を修正(refine)していく。この3D位置の修正が十分なレベルに達したら、特徴点は校正されたものとみなされる。この方法は、ドリフトの問題を多少は軽減するが、実際には、開始点から大きく動いたときに、まだ大きなドリフトが残る。新しい特徴点の正確な判定は、前の3D特徴点位置の3D位置の正確な推定に依存し、これはさらに、前のカメラポーズの正確な推定に依存している。

10

上記のような提案された方法においては、常に、演算/推定のたびに多少の数値的および数学的な誤差が加わる、ということがよく知られている。このことは、開始点から遠く離れた特徴点はドリフトのために大きな誤差を有する、ということの意味する。さらに、もし単一の特徴点の3D位置が「校正された」ものとみなされ、かつこの位置が実際には不良であった場合には、特徴点のその後のトラッキングおよび校正は、より大きな誤差の影響を受ける。

さらに、特徴点の3D位置の正確な推定および修正を確実にを行うために、ユーザは、特定の方法で動くことを要求され、実用上は、面倒でかつ扱いが困難なシステムとなる。US 6,765,569 B1は、画像中に十分な特徴点を認識できなかった場合にポーズを回復する方法の開示を含んでいるが、この方法は、満足な推定がなされた特徴点が存在する方向をユーザに見せ、これらの点を「校正した」特徴点とみなしうようにする、ということにまだ依存している。この方法は、カメラで観測し得るような環境の3Dモデルは使用せず、このことは、後続のフレームで三角法を実行しつつ特徴点の3D位置を演算することを意味する。

20

#### 【0005】

### 2.2 特徴点の局所的な検出およびマッチング(デテクション)

ドリフトおよびポーズの喪失という問題を除去する良いやり方は、システムの実行前に、環境/オブジェクト/シーンの特徴点に基づくシステムのトレーニングを行う方法を用いることである。3D位置とともにトレーニング/分類の情報を記憶することで、現在のフレームで検出された特徴点を、データベースに保存されたものあるいは分類装置を通して分類されたものとリアルタイムにマッチングすることができるようになる。十分な数の特徴点がマッチすれば、幾何学的計算を介して物理的カメラポーズの推定が可能である。数値最小化を通して、さらなる修正が得られる。これは、前のフレームの推定の結果に依存することなく、単一のフレームの各々ごとに実行できる。実際には、これは、「ポーズの回復」が単一のフレームの各々で実行されることを意味する。

30

十分な数の特徴点がキーポイントデータベースの内容に対して分類されあるいはマッチしている限りは、デテクションではドリフトおよびポーズの喪失は問題とならない。しかしながら、この種の方法は、偽デテクションの問題があることが知られており、つまり、リアルタイムに検出された特徴点が、データベースに記憶された特徴点に対し、偽の分類あるいは偽のマッチとなる。カメラポーズの推定のいくつかの方法では、ある程度の数の偽マッチ(アウト라이어とも呼ばれる)を処理することができるが、いくつかのアウト라이어があっただけでも、推定ポーズの最終的な結果は、一般にリピータビリティが低いものとなる。このようなデテクション法の通常のケースでは、少なくとも15個のアウト라이어が存在する。これでは、カメラが完全に静止している場合であっても、一般に、低いリピータビリティとなる。この問題は、拡張リアリティシステムにおいては、パーティクルオブジェクトが所定位置に維持されず、周囲でバウンドしたり震えたりするように現れる、という影響を与える。

40

US 2006/0233423 A1は、特徴点をリアルタイムにデータベースとマッチングする方法、および、校正/トレーニングした特徴点についての関連情報をトレーニング/記憶する方法を示唆している。上述した記載との関係では、この示唆されている方法は、デテクショ

50

ン法であると考えられる。

V. Lepetit、P. LaguerreおよびP. Fuaによる「Randomized Trees for Real-Time Keypoint Recognition」(<http://cvlab.epfl.ch/research/augm/detect.php>)は、単一のフレーム内でオブジェクトを検出するために、上記US 2006/0233423 A1に開示されたものと類似しかつこれよりも早期の方法を開示している。データベースを用いる代わりに、「分類ツリー」(後述の4.13章を参照)を用いることが提案されている。分類ツリーは、トレーニングフェーズにおいて構築され、実行時に各キーポイントが「ドロップダウン」する複数のツリーであって、最終的にキーポイントがあるスコア(信頼度)とマッチする。この方法であっても、デテクション法の長所と短所とを示す。

【0006】

10

### 2.3 記録した画像ストリーム

記録した画像ストリームに基づいて物理的カメラのポーズを推定する他の方法およびアルゴリズムが多数公開され、かつ特許されている。画像ストリームに基づく物理的カメラポーズの推定とは、「現在」のフレームのカメラポーズを算出し得るようにするために、ストリームの全てないしいくつかの画像の「早送り」および「巻き戻し」を適時になし得ることが必要であることを意味している。映画の製造においては、この種のいくつかの方法が特許されかつ公開されているが、上述したように、かかる分野においては、その方法をリアルタイムに実行するのではなく、また満足 of いく結果を得るために、完全な画像ストリームに依存している。これらの使用は、リアルタイムの拡張リアリティを意図したものではなく、「ハリウッド映画」の後工程的な画像効果を意図している。

20

【課題を解決するための手段】

【0007】

### 3 本発明

本発明の装置および方法は、3Dモデルを用いたフレーム・トゥ・フレーム・トラッキングの長所と局所的な特徴点の検出およびマッチングの長所とを組み合わせ、満足 of いく精度で物理的カメラのポーズの推定を得るようにしたものである。ドリフトや偽マッチによるポーズの喪失ならびに震えといった問題は、デテクションの長所とトラッキングの長所とを組み合わせた革新的かつユニークな方法および装置によって解消される。デテクションとトラッキングとを組み合わせることで、リアルタイムの利用には遅すぎる解法を実行するというリスクが回避されるものであり、本発明の装置および方法では、特にマルチCPUシステムにおいて、早くかつ効率的な実行を確実なものとするために、サブタスクを複数のスレッドおよびスレッド・タスクに分割するというユニークな機会を提供する。今日のコンピュータでは、マルチCPUやデュアルコアCPUのテクノロジーが広く利用可能でかつ手頃なものとなっており、これは、本発明によって提供される解法の実行に非常に適している。本発明によって提供される解法のいくつかの重要な利点は、ステップ毎に数字を付すと、下記の通りである。

30

1. 局所的な特徴点の検出およびマッチングを通じたロバストな推定。これは、前のフレームでの推定に依存しない各フレーム毎の「ポーズの回復」を意味する。本発明により提供される解法は、2005年6月のカルフォルニア州サンディエゴでの「Conference on Computer Vision and Pattern Recognition」におけるV. Lepetit、P. LaguerreおよびP. Fuaによる「Randomized Trees for Real-Time Keypoint Recognition」に開示された方法を用いることにより、有利に実行される。

40

2. フレーム・トゥ・フレーム・トラッキングを通じた精度およびリピータビリティが、前に推定したポーズと環境の3Dモデルとの組み合わせを用いることで得られる。

本発明により提供される解法では、さらに、一般的なフレーム・トゥ・フレーム・トラッキングに新規なアイデアを導入しており、最後に正常に(successfully)算出されたポーズが記憶され、かつトラッキングに用いられる。これは、もし前のフレームにおいてカメラポーズが正常に算出されなかったとしても、本発明による解法では、最後の正常なポーズに基づいてトラッキングの実行を試みる、ということの意味する。これにより、本発明による解法では、上記ステップ1に示すデテクション法が仮に失敗した場合や直

50

前のいくつかのフレームにおいて最終的なポーズが役に立たないような場合であっても、ポーズの推定を再初期化することができる。ただし、この新規なアイデアは、最後のポーズの推定が正常になされたところに近似した位置/向きに、ユーザが戻ることを必要とする。

3. カメラポーズの推定のための数値最小化への入力パラメータに運動モデル観測を含むことによって、精度およびリピータビリティがさらに向上する。

4. 「ロバスト・エスティメータ (estimator)」を通したロバストな演算により、アウトライアーの影響が軽減する。我々はTukeyエスティメータを用いるが、他のものも使用できる。

5. 比類のない速度および性能が、我々のユニークな並行処理アルゴリズムによって得られる。

【図面の簡単な説明】

【0008】

【図1】コンピュータのメインループの説明図。

【図2】本発明が組み込まれた装置のブロック図。

【図3】結合モードの説明図。

【図4】分割モードの説明図。

【図5】スレッドの説明図。

【図6】本発明の実施に適合したコンピュータの説明図。

【図7】P3Pの説明図。

【発明を実施するための形態】

【0009】

### 3.1 システムの概説

図6は、本発明の実施に適合したものとして、デュアルコアの中央処理ユニット(CPU)、ランダムアクセスメモリ(RAM)グラフィックカードおよびハードドライブ(HDD)が、マザーボード上に取り付けられてなる標準的なコンピュータを示している。さらに、カメラをコンピュータに接続するためのデータインターフェースがある。カメラは、興味のあるシーンのような現実環境の画像を取り込むために、レンズを有する。この画像は、センサ、通常はCCDセンサないしCMOSセンサ、の上でデジタル化される。通常、画像はカメラの内部メモリに一旦記憶され、次いで、接続Aを通してコンピュータに転送されるが、カメラインターフェースがこれを解釈し、かつマザーボードと通信して、システムRAMの目的に応じて配置された領域へと送る。

CPUは、典型的には、IntelのCore 2 Duoとすることができる。我々のシステムは、シングルコアCPUを用いることもできるが、その場合の性能は低下するであろう。RAMは、典型的には、標準の512MBのDDR2 SDRAMである。ハードドライブは、典型的には、標準の40GB SATA 7200 rpmのハードドライブである。データインターフェースは、典型的には、USB 2.0、FireWire IEEE 1394、ワイヤレスLAN(WLAN)等とすることができる。接続Aは、データインターフェースに応じて、有線もしくは無線となり得る。

ソフトウェアは、マシン可読コードをもってハードドライブ(HDD)に記憶されており、実行開始時にRAMに読み込まれる。画像はカメラを用いて取り込まれ、接続Aを介して、RAMのメモリ領域に転送される。次いで、このRAMに記憶した画像に対し、1つあるいは複数のCPUないしCPUコアを用いて、アルゴリズムが実行される。このアルゴリズムの結果となる物理的カメラの推定ポーズは、取り込んだ画像の上にバーチャルオブジェクトを現実環境に正しく整合した形で描くために用いられる。そして、この画像および/またはバーチャルオブジェクトは、通常はグラフィックカード内にレンダリングされ、その結果がスクリーン上に表示される。スクリーンとしては、透明スクリーン(画像がミラー、ガラス等に投影される)あるいは不透明スクリーンとすることができる。不透明スクリーンとする場合には、拡張リアリティ効果を得るために、バーチャルオブジェクトをレンダリングする前に、カメラで取り込んだ画像を背景として描く必要がある。スクリーンは、カメラを一体化したヘッド装着型ディスプレイとすることができ、カメラは

10

20

30

40

50

該ヘッド装着型ディスプレイの前方に取り付けられる。カメラを前方に備えた他のディスプレイとしては、PDAや携帯電話ディスプレイ（将来的には携帯電話の内部コンピュータを処理に利用し得る可能性もある）、タブレットPCディスプレイ、等がある。

最近のパワフルなグラフィックカードは、GPU（グラフィカル処理ユニット）上で数学的機能を実行することもでき、このことは、我々のシステムにおけるタスクおよびアルゴリズムを、CPU/コアに加えてGPU上でも実行し得ることを意味している。Ageia PhysXのような物理カードであっても、我々のソフトウェアのタスクを実行するために利用し得る。

【0010】

### 3.2 ソフトウェアの実現 (implementation)

本発明の装置および方法（本章では一部、アルゴリズムと記す）は、基本的な数学的方法が利用可能なあらゆる種類のコンピュータソフトウェア環境において有利に実現することができる。もし基本的な数学的方法が利用可能でなければ、最初から開発 (scratch) することができる。MATLAB、FORTRAN、C/C++、C#などが適したプログラム言語である。ソフトウェアは、Linuxオペレーティングシステム、Windows、Unix、Xbox、Playstation、Symbian等で動作するコンピュータ上で実現することができる。

必要な数学的方法としては、特に（但し、これらには限定されないが）、数値最小化と、フィルタ、エッジ検出、コーナ検出等の基本的なコンピュータビジョンアルゴリズムと、行列操作の方法（行列の反転、転置、乗算、等）と、がある。

【0011】

### 3.3 メインループの概説

本発明を実現する例示的なシステム、例えば図1に示したようなコンピュータにおいて、そのメインループとしては、実際の画像をカメラから抽出するとともにシステムRAMの循環メモリバッファに記憶し、カメラポーズを推定し、かつ最終的に結果をスクリーン上にレンダリングする。

最初のフレームの取り込みがなされかつポーズ推定タスクが開始すると、直ちに、新しいフレームを取り込むスレッドAが開始する。最初のポーズ推定が完了すると、新フレームが既に並行して取り込まれている場合には、直ちに、新フレームにおけるカメラポーズの推定の開始が可能である。最初のレンダリングが完了すると、新フレームおよびカメラポーズがスレッドA、Bから利用可能である場合には、直ちに、新しいレンダリングが開始可能である。これは、スレッドAがスレッドB、Cと並行して働き、新規なフレームを取り出し、スレッドBが、前のタスクが完了するや否やこの新規なフレームを利用することができる、ということの意味する。スレッドBも同様であり、このスレッドBはスレッドCと並行して働き、新規なポーズ推定を行い、スレッドCは、前のレンダリングが完了するや否やこの新規なポーズ推定を利用することができる。このような方式は、複数のCPU、GPUあるいはマルチコアCPUを備えたコンピュータを利用した非常に効率的な方法を提供する。結果として、フレーム速度および実行時間が実質的に速くなる。スレッドBは、複数の並行したタスクに分割され、これにより、物理的カメラポーズの推定が非常に効率よく実現される。スレッドBについての詳細な情報は3.4章を参照されたい。

開始時に最初の画像/フレームが取り込まれかつRAMの保存先となるメモリ領域に転送された後に、このフレームにおける物理的カメラポーズの推定が可能である。ポーズ推定が達成された後に、カメラ画像およびポーズは、拡張リアリティ効果を得るべく最初の画像中にバーチャルオブジェクトをレンダリングするために利用できる。

スレッドA (TA)、スレッドB (TB) およびスレッドC (TC) は、全て、1つあるいは複数のシングルコアCPU、1つあるいは複数のマルチコアCPU、および/または、1つあるいは複数のGPU、あるいは他の特化したコンピューティングユニット、の上で動作する。これらのタスクは、特定のCPU/コア上で動作するように特に設定することもでき、また、オペレーティングシステム (OS) によってスレッドタスクを実行するCPUが自動的に定まるものであってもよい。

本発明の一実施例が組み込まれた装置の例が、図2にブロック図として模式的に描かれ

10

20

30

40

50



ている。

この装置は、シーン220の2次元(2D)画像を示す画像フレームの画像データ231をフレームグラバ230から受ける入力(通常は上記画像フレームはカメラ210によって取り込まれ、かつフレームグラバによって取得される)と、カメラデータソース240から、カメラレンズデータ、主点、等のカメラデータCを受ける入力と、シーンモデルデータソース260から3次元(3D)シーンモデルデータMを受ける入力と、シーンデータソース250から予記録シーンデータ251(例えば予め記録したシーン画像の処理により得た3D特徴点データ等)を受ける入力と、を備える。この装置は、上記の中の1つあるいは複数の入力に基づいて演算を行い、カメラによって撮像されているシーンに関するカメラの幾何学的3Dポーズの現在の推定181を出力するように構成されている。有利には、本装置の幾何学的3Dカメラポーズ・エスティメータ(estimator)は、幾何学的3Dカメラポーズ推定の「信頼性のレベル」を判定するように構成され、幾何学的3Dカメラポーズ推定が算出されたところの信頼性のレベルを示す3Dポーズスコア値を出力するように構成し得る。

10

この装置は、さらに、フレームグラバ230から画像データ231を受け取るとともにこれを処理するように構成された候補2Dキーポイント・デテクター110を備えており、この候補2Dキーポイント・デテクターは、3D位置が既知であると考えられるキーポイントを示す特定の画像特徴部を検出し、この検出した特定の画像特徴部を示すデータを候補2Dキーポイント111として出力するように構成されている。この特定の特徴部の例としては、オブジェクトが特異な色で出現するようなシーンの特徴部、あるいは明瞭に識別可能な単一もしくは1組の幾何学的特徴、例えばコーナやエッジ、を示すシーンの一部、である。

20

この装置は、さらに、候補2Dキーポイント111および予記録シーンデータ251を受け取るデテクテド・3D観測デテクター120を備えており、このデテクテド・3D観測デテクターは、予記録シーンデータ中に対応する3D特徴点を有すると考えられる候補2Dキーポイントの各々についてのデテクテド・3D観測121を出力するように構成されている。上記デテクテド・3D観測デテクターは、さらに、デテクテド・3D観測の「信頼性のレベル」を判定するように構成され、デテクテド・3D観測が判定されたところの信頼性のレベルの指標としてスコア値を出力するように構成し得る。

この装置は、さらに、デテクテド・3D観測121およびカメラデータCを受け取るデテクテド・3Dカメラポーズ・エスティメータ(estimator)130を備えており、このデテクテド・3Dカメラポーズ・エスティメータは、デテクテド・3Dカメラポーズ推定131を演算によって決定し、かつ出力するように構成されている。有利には、このデテクテド・3Dカメラポーズ・エスティメータは、デテクテド・3Dカメラポーズ推定の「信頼性のレベル」を判定するように構成され、デテクテド・3Dカメラポーズ推定が演算されたところの信頼性のレベルの指標となるポーズスコア値を出力するように構成し得る。

30

この装置は、さらに、グラブド・フレーム画像データ231および候補2Dキーポイント111を記憶するための第1ストレージ手段140を備えており、この第1ストレージ手段140は、シーンの2次元(2D)画像を示す画像フレームのフレームグラバ画像データを受け取るように構成されたものであって、カメラの幾何学的ポーズについての現在の推定出力181を検出し、これに応答して、前の幾何学的カメラポーズ推定が基礎としたストアド・グラブド・フレーム画像データ141およびストアド・候補2Dキーポイント142を出力するように構成されている。

40

この装置は、さらに、カメラの幾何学的ポーズについての前の推定161を記憶するとともに出力する第2ストレージ手段160を備えており、この前の推定は、シーンに関するカメラの幾何学的ポーズについての現在の推定の出力181よりも前に装置の出力に表されたものである。

この装置は、さらに、トラックド・3D観測デテクター150を備えており、このトラックド・3D観測デテクター150は、カメラの幾何学的ポーズについてのストアド・前

50

推定 1 4 1 と、候補 2 D キーポイント 1 1 1 と、前の幾何学的カメラポーズ推定が基礎としたストアド・グラブド・フレーム 1 4 1 および候補 2 D キーポイント 1 4 2 と、カメラデータ C と、3次元(3D)シーンモデルデータ M と、を受け取るように構成されている。このトラックド・3D観測デテクターは、前の幾何学的カメラポーズ推定に存在する対応する3D特徴部を有するものと考えられる候補2Dキーポイントの各々についてトラックド・3D観測151を出力するように構成されている。上記トラックド・3D観測デテクターは、さらに、トラックド・3D観測の「信頼性のレベル」を判定するように構成され、トラックド・3D観測が判定されたところの信頼性のレベルの指標となるスコア値を出力するように構成し得る。

この装置は、さらに、3Dカメラポーズ推定セクタ170を備え、この3Dカメラポーズ推定セクタ170は、記憶した前の幾何学的3Dカメラポーズ推定161とデテクテド・3Dカメラポーズ推定131とをそれぞれ受け取る複数の入力を有するとともに、選択した3Dカメラポーズ推定171を提供する1つの出力を有する。有利には、この3Dカメラポーズ推定セクタは、入力の各々が受け取った3Dカメラポーズ推定とともに表されたスコア値に回答して、その出力に、選択後3Dカメラポーズ推定として、入力が受け取った3Dポーズ推定の中の一方、あるいは、両者を結合したものを提供する。

この装置は、さらに、幾何学的3Dカメラポーズ・エスティメータ(estimator)180を備え、この幾何学的3Dカメラポーズ・エスティメータ180は、選択後3Dカメラポーズ推定171と、デテクテド・3D観測121と、トラックド・3D観測151と、カメラデータCと、を受け取るように構成されているとともに、現在の幾何学的3Dカメラポーズ推定181を演算によって決定し、かつ出力するように構成されている。有利には、上記幾何学的3Dカメラポーズ・エスティメータは、現在の幾何学的3Dカメラポーズ推定の「信頼性のレベル」を判定するように構成され、現在の幾何学的3Dカメラポーズ推定が演算されたところの信頼性のレベルの指標となるポーズスコア値を出力するように構成し得る。

【0012】

### 3.4 物理的カメラポーズ推定(スレッドB)

本発明のコンピュータによる有利な実施は、下記の通りである。

本発明のコンピュータで実行する物理的カメラポーズ推定方法(ここでは物理的カメラポーズ推定アルゴリズムとも呼ぶ)は、有利には、2つのモードの中の1つとして実行されるものであり、スタートアップ時に(A)「結合」モードと(B)「分割」モードから選択される。スレッドBは、そのタスクを、別のスレッドやスレッドタスクに分配し、複数のシングルコアCPUやGPU、1つあるいは複数のマルチコアCPUやGPU、を備えてなるコンピュータにおける実行時間性能を最適化する。

上記「結合」モードでは、デテクションおよびトラッキングの一部を並行して実行し、その後、結果を結合かつ処理して、最終的な数値最適化を行う。上記「分割」モードでは、デテクションおよびトラッキングをより切り離して実行し、最終的な結果を最終カメラポーズとして分析かつ処理する。本発明における「結合」モードおよび「分割」モードの実行は、図3および図4にそれぞれ図示されている。

【0013】

結合モード(A)と分割モード(B)の共通事項

【0014】

#### 3.4.1.1 回転(Rotation)および並進(Translation)のパラメータ化(外部パラメータ)

我々のPROSACの利用における再射影誤差(およびヤコビ行列)の演算および数値最小化(Levenberg-Marquardt)方法の際には、回転が、「exponential map」を用いてパラメータ化される。推定の前および後における回転行列と「exponential map」との間の相互の変換に、Rodriguesの公式が用いられる。

数値最小化の間、状態ベクトルは、6要素のベクトルであり、「exponential map」および並進ベクトルとしての回転および並進を含む。

10

20

30

40

50

「exponential map」における特異点を回避するために、我々は、「exponential map」（ベクトル） $\|w\|$ の長さが、 $2^{-n}$ （但し、 $n$ は3.14、 $n$ は0よりも大きい整数）に近いかどうかチェックする。もし長さが $2^{-n}$ に近い場合は、これを「 $(1 - 2^{-n} / \|w\|)w$ 」（但し、 $w$ はexponential mapである）に置き換えることができる。（F. Sebastian Grassia「Practical parameterization of rotations using the exponential map」Journal of Graphics Tools, 3(3):29-48, 1998年より）

【0015】

#### 3.4.1.2 キーポイントの検出

カメラからの新鮮な生の画像に対し、可能性のある特徴点/キーポイントの検索がなされる。これらのキーポイントは、さらなるデテクションおよびトラッキングに用いるための候補点となる。添付の図に示すように、キーポイントの検出は、トラッキングおよびデテクションの双方に対し演算される。

10

【0016】

#### 3.4.1.3 レンズ歪曲の補償

カメラから取り込んだ画像は多少のレンズ歪曲を含む。我々のシステムにおいて、我々は、カメラおよびレンズの予校正を通して得た内部カメラパラメータにおけるラジアル方向歪曲および接線方向歪曲を補償する。

本発明の有利な実施例において内部カメラパラメータを補償しかつパラメータ化するためには、Z. Zhangによる「A flexible new technique for camera calibration」IEEE Transactions on Pattern Analysis and Machine Intelligence, 22(11):1330-1334, 2000年 に記載されている方法が用いられる。

20

レンズ歪曲の補償においては、入力画像の全体を歪曲除去してもよく、あるいは、検出されるキーポイントだけ歪曲除去してもよい。画像全体を歪曲除去するのは、時間が掛かるタスクであり、検出されるキーポイントを歪曲除去の方がより効率的である。

もし歪曲したカメラ画像をバーチャルグラフィックとオーバーレイするために使用したとすると、バーチャルグラフィックは背景画像と正しく整合しない。これを補償するためには、バーチャルグラフィックを内部カメラパラメータに従って歪曲させる必要がある。もし、ビデオ・シースルー式の装置としてカメラ画像を背景に用いない場合には、このことは問題とはならない。

30

【0017】

#### 3.4.1.4 初期ポーズの推定

このタスクは、双方のモードにおいて実行され、キーポイントとデータベースとのマッチングに基づくラフな初期ポーズの推定（デテクション）からなる。

初期ポーズの推定の間に用いられる位置および方向のパラメータ化については、3.4.1.1章を参照されたい。

初期ポーズの推定のステップは次の通りである。

1. 画像中に検出したキーポイントを、2D-3D点相関関係を得るために、分類する。有利には、本発明では、分類およびトレーニングのためのアルゴリズムとして、V. Lepetit、P. LaguerreおよびP. Fuaによる「Randomized Trees for Real-Time Keypoint Recognition」という刊行物に示されているアルゴリズムを用いる。これ以外の特徴点の分類/デテクション法にも適当なものがあり得る。

40

2. 所定の閾値（例えば0.5）を超えるスコアでもってマッチするものが4つ以上存在するか否か判定し、

- a. YESであれば、PROSACアルゴリズムを実行し（4.14章を参照）、
  - i. PROSACからの結果が満足できるものか判定して、
    - 1. YESであれば、成功として終了し、
    - 2. NOであれば、失敗として終了し、
  - b. NOであれば、失敗として終了する。

この初期のラフなポーズは、最小化アルゴリズムが包括的な最小値を確実に見つけるために、数値最小化の開始点として用いられる。間違った/悪い初期ポーズを用いると、最

50

小化は、最終的に、局所的な最小値、換言すれば不正解な解、となる。初期ポーズを包括的な最小値の近くに得るためには、要求される再射影誤差およびラフ初期ポーズの推定におけるインライアの最小数に、適切な制約条件を設定することが重要である。最大の再射影誤差は、5ピクセルから25ピクセルの間とすべきである。インライアの最小数は、15~40とすべきである。

我々の実施においては、我々はまた、インライアの最小数を、観測の総回数（インライアとアウトライアの双方）のパーセンテージとして毎回演算するように、指定することができる余地を与えている。これにより、多数の観測が存在する場合に、アルゴリズムが自動的に要求に適合できるものとなる。このことの危険性は、多数のアウトライアがある場合である。もしインライアの最小パーセンテージが過度に高いと、PROSACアルゴリズムは、十分なインライアでもってポーズを演算することに失敗してしまう。

【0018】

#### 3.4.1.5 数値最小化

本発明のさらなる有利な実施例においては、Levenberg-Marquardtアルゴリズムが数値最小化のために用いられるが、他の方法も利用可能である。

数値最小化の間に用いられる位置および方向のパラメータ化については、3.4.1.1を参照されたい。再射影誤差の合計が、数値最小化の間に最小化される。4.14章に記載したPROSAC法のように、数値最小化におけるアウトライアを無視するために、ロバスト・エスティメータもまた再射影誤差の計算の間に適用することができる。有利には、本発明は、Tukeyエスティメータを用いるが、他のロバスト・エスティメータを用いることもできる。さらに、現在および前のフレームにおけるポーズの間での並進および回転の最小化のために、運動モデル観測が用いられる。

上記の記述は、モードAおよびモードB双方の数値最小化サブタスクについて該当する。

【0019】

#### 3.4.2 結合モード(A)

この章では、図3を参照して、結合モードについて説明する。

一般に、「新フレーム」は歪曲のないものと仮定され、つまり、カメラ/レンズの歪曲が補償されていることを前提とする。

2Dキーポイントのみが歪曲除去される場合には、図は多少変更する必要がある。しかし、コンピュータビジョンの技術者には、どこで2Dキーポイントの歪曲除去を適用するか認識することが可能である、と思われる。

「2D-3D相関関係の結合および予備処理」および「運動モデル観測の評価および準備」の2つのプロシージャは、互いに独立しており、さらなる性能最適化のために2つの別のスレッドにおいて、並行して実行されるように設定可能なものである。しかし、この2つのプロシージャは特に時間の掛かるものではなく、これらのタスクを別々のスレッドに分割するためのオーバヘッドは、タスクを並行に処理することにより得られる向上代と同等かそれ以上となり得る。

【0020】

##### 3.4.2.1 キーポイントの検出、デテクションおよびトラッキングのスレッドの開始

キーポイントが検出された後に、結合したスレッドA(CTA)が結合したスレッドC(CTC)と並行して開始する。

CTAは下記を実行する。

1. 2D-3D点相関関係を得るために、Lepetit法を用いて、現在の画像中に検出したキーポイントを、予トレーニングした分類ツリー/データベースで分類する。

2. 3.4.1.4章に記載した方法を用いて、2D-3D点相関関係に基づき初期ポーズを推定する。

CTCは、前に算出に成功したポーズ(最後の良ポーズ)が存在するのであれば、下記を実行する。

1. 現在の画像中に検出したキーポイントを、「最後の良ポーズ」中に検出したキー

10

20

30

40

50

ポイントとマッチングする。

平均輝度閾値を適用することで、平均輝度の差が所定の閾値未満のキーポイント間でのみマッチングを許容するようにすることができる。低い値の設定（例えば可能性のある255の中の10）により、類似した平均輝度を有するキーポイントのみが許容される。これにより、キーポイントのマッチングの信頼性が向上する。

2. 対応マッチ・スコアが所定の閾値よりも高い全てのキーポイントについて、レイ・トレーシングを実行する。このレイ・トレーシングでは、前のフレームのカメラポーズと、3Dモデルと、内部カメラパラメータと、を用いて、マッチしたキーポイントの3D位置を演算する（4章のレイ・トレーシングの説明を参照されたい）。

我々のシステムにおいて、CTAからの「デテクション観測」のみに依存する場合と、CTCからの「トラッキング観測」にも依存する場合と、をシステムに自動的に選択させるための重要な特徴は、下記の通りである。

トラッキング・スレッド（CTC）におけるキーポイントマッチングの実行においては、フレーム間の2Dキーポイント位置の最大差異を設定することが可能である。これにより、2つのフレームで互いに近接しているか否かに依存してマッチングされる、可能性のあるキーポイントを限定することができる。u方向およびv方向の最大差異を低い数値、例えば10ピクセル（但し、システムが1秒あたりに20～30フレームを実行する場合）に設定することで、我々のシステムは、カメラが素早く動くときに、自動的に、デテクションによる観測のみに依存するようになる。カメラがゆっくりと動くようになると、直ちに、より多数のマッチングが許容されることから、フレーム・トゥ・フレーム・トラッキングからの観測が数値最小化により強く影響するようになる。このことは、我々のシステムの他の特徴と組み合わせられて、システムを、速いカメラの動きに対しロバストなものとし、かつ同時に、カメラの動きが小さいときには、高いリピータビリティを有する正確なものとしている。

【0021】

#### 3.4.2.2 2D - 3D 相関関係の結合および予備処理

CTAおよびCTCからの2つのセットの相関関係ないし観測は、2つの異なるタイプのスコアを有する。デテクション・セット（CTA）は、予トレーニングしたデータベースに対するマッチング/分類により与えられるスコアを有するが、トラッキング・セット（CTC）は、現在のフレームにおけるキーポイントが前のフレームに対しどの程度良好にマッチするかを示す対応マッチ・スコアを有する。これらの2つのタイプの観測スコアを制御された形で確実に取り扱うために、デテクション観測（CTA）がトラッキング観測（CTC）と等しく「重要」になるように、全てのスコアが正規化される。

さらに、正規化したスコアに基づき、正規化したスコアにスカラを乗算することで、2つのタイプの観測の一方を「増進」もしくは「低下」させることが可能であり、これにより、選択されたタイプの観測がより「重要」にあるいはより「重要」ではないものとなる。

最後に、一方のセットもしくは他方のセットの観測の数の最小数および最大数を設定するフィルタリングが実行される。一般に、デテクション・セット（CTA）から最大のn個の観測を許容したり、トラッキング・セット（CTC）から最大のk個の観測を許容したりすることは、決してしないことが望ましい。これは、2つのタイプの観測の「重要性」をさらに制御するためになされる。

観測のタイプを互いに評価および制限/制御した後に、全ての観測が数値最小化に加えられる。もし、前のフレームでのポーズの演算が成功していなかった場合には、トラッキング・タスク（CTC）は実行されておらず、CTAからのデテクション観測のみが加えられる。

【0022】

#### 3.4.2.3 運動モデル観測の評価および比較

運動モデル観測が用意される。この運動モデル観測は、前のフレームにおいてポーズが推定されていた場合に、現在のフレームと前のフレームとにおけるカメラポーズのx, y

、 $z$  軸に沿った並進およびこれらを中心とした回転の差からなる。これらの観測は、数値最小化が、 $x$ 、 $y$ 、 $z$  軸に沿った並進およびこれらを中心とした回転を最小化するように作用し、つまり、前のポーズに基づくポーズの結果を抑制しようとする。

運動モデルは、また、新フレームでの運動を予測するアドバンス型予測運動モデルともなる。

運動モデル観測の評価は、所定の値を設定し、この値は、これらの観測がどの程度数値最小化に影響するか、ということ左右する。

再射影誤差に関しては、運動モデル観測に対しやはりロバスト・エスティメータ (estimator) が用いられる。これは、例えば、2つの連続したフレームの間での $x$ 軸に沿った非常に大きな運動がアウトライアー観測である、とみなされることを意味する。インライアー/アウトライアーを決定するための閾値は、システム全体のフレーム速度(カメラがどれくらい速く動くことが予想されるか、を含めて)に応じて設定される。

10

【0023】

#### 3.4.2.4 数値最小化の実行

初期の推定ポーズ(3.4.1.4章を参照)が、数値最小化の開始点として用いられ、これにより、最終結果が、確実に、実際の/正確な包括的な解となるようにしている。ピクセルにおける2D再射影誤差は、2D-3D点相関関係のために最小化され、運動モデル観測は、できるだけ最良のポーズを得るために最小化される。2D-3D点相関関係および運動モデル観測の双方に対し、ロバスト・エスティメータが実行され、アウトライアーに対し確実にロバストなものとしている。

20

最小化は、制限された数の繰り返しだけ実行され、あるいは、満足できる結果が得られるまで繰り返し実行される。

【0024】

#### 3.4.2.5 現在のカメラポーズ、フレーム、および2Dキーポイントの記憶保持

次のフレームへの準備を行うための後処理は下記の通りである。

インライアーの数を最小とする要求に基づいて、ポーズの推定が成功したか？

a. YESの場合

i. 現在のフレーム、2Dキーポイントおよびポーズを、次のフレームのトラッキング・スレッドにおいて用いるために、「最後の良い」フレーム、2Dキーポイントおよびポーズとして、記憶し、

30

ii. 現在のポーズを、次のフレームの運動モデル観測に用いるために、「前のカメラポーズ」として記憶する。

b. NOの場合。運動モデル観測に用いられる「前のカメラポーズ」が無効であると、マーク付けする。

【0025】

#### 3.4.3 分割モード(B)

図4を参照して、分割モードを説明する。分割モードにおいては、結合モードと同様、最初のステップは現在の画像中のキーポイントを検出することである。

次に、タスクは2つのスレッド、つまりデテクション・タスクを実行する分割スレッドA(STA)と、トラッキング・タスクを実行する分割スレッドC(CTC)と、に分割される。

40

STAは下記を実行する。

1. 2D-3D点相関関係を得るために、現在の画像中に検出したキーポイントを予トレーニングしたデータベースとマッチングする。

2. 2D-3D点相関関係に基づき、初期ポーズを推定する。

3. 観測を、制限し、評価し、かつ制御する。

4. デテクションに関する数値最小化に観測を加える。

5. もし、前のポーズにおいてポーズの推定が成功していれば、運動モデル観測をデテクション・最小化に加える(結合モード中として)。

6. 初期ポーズを開始点として用いて数値最小化を実行し、2D再射影誤差および運

50

動モデル観測（これが追加されていれば）を最小化する。

前のポーズにおいてポーズの推定が成功していれば、これと並行して、S T C は下記を実行する。

1．現在の画像中に検出したキーポイントを、前の画像中に検出したキーポイントとマッチングする。

2．対応マッチ・スコアが所定の閾値よりも高い全てのキーポイントについて、レイ・トレーシングを実行する。マッチしたキーポイントの3D位置を、前のフレームでのカメラポーズと3Dモデルとを用いて演算する（4章のレイ・トレーシングについての説明を参照）。

3．トラッキングに関する数値最小化に観測を加える。

4．（前のポーズの推定が成功していれば）S T A と同様に運動モデルをトラッキング・最小化に加える。

5．前のポーズを開始点として用いて数値最小化を実行し、2D再射影誤差および運動モデル観測（これが追加されていれば）を最小化する。

a．前のポーズを用いる代わりに、S T A が初期ポーズを演算するのを待ち、この初期ポーズを前のポーズの代替として使用することも可能である。これにより、性能上、多少のオーバーヘッドが生じるが、カメラの動きが速い間に、より良好な結果を導く。

各々の最小化は、最終的に別々に推定したカメラポーズとなり、各々、いくつかのインライナー観測を有するとともに、最小化からのトータルの残差を有する。2つのポーズの「良好性」を分析するためのステップが実行される。もし一方の最小化からの結果が他方に比べて顕著に良好であれば、その最良の一方が最終結果として選択される。もし、2つの最小化からの結果がほぼ同程度に良好なものであれば、これらの結果の平均が用いられる。所望であれば、最終結果を得るために、加重平均を用いることができる。

【0026】

### 3.5 他の改良点

本発明により提供される解は、画像中に現れた特徴点、つまり高いコントラストを有するコーナやエッジ等に依存している。そのため、現実のシーンが、カメラポーズの推定のためにカメラで取り込まれるフレーム中に十分に特徴点を検出できるようなテクスチャ（質感）を含んでいることが必要である。室内において、例えば白い壁のみで塗装や他のオブジェクト/イラスト（これらは取り込まれるフレームにテクスチャおよびコントラストを与える）がない場合には、特徴点を提供する必要がある。

エッジ・デテクションを環境の3Dモデルと組み合わせることで実施することにより、本システムは、このような種類の環境においても機能する。このようなアプローチを本システムに統合するために、同じ性能（毎秒のフレーム数）を確保するように、別のスレッドないしスレッド・タスクを並列に実行する必要がある。

この新しいスレッド・タスクは、別のセットの観測を生成する。これらの観測は、一般には、取り込んだフレーム中のエッジ/ラインである。そして、数値最小化によって、取り込んだ画像中に検出/測定したラインと、現在推定されているカメラポーズ（現在とは、最小化プロセスの各々の繰り返し毎のポーズ）を用いてスクリーンに射影された環境の3Dモデルにおけるラインと、の間の距離が最小化される。

結合モード（3.4.2章）において、このことは、4つのタイプの観測が存在することを意味する。

1．デテクテド・2D - 3D点相関関係（2Dピクセル再射影誤差の最小化）

2．トラックド・2D - 3D点相関関係（2Dピクセル再射影誤差の最小化）

3．運動モデル観測（前のフレームでポーズが推定されていれば、前のフレームに対する並進および回転の最小化）

4．新規：エッジ観測（取り込んだ画像中のエッジと、現在のカメラポーズを用いてスクリーンに射影した3Dモデルにおけるライン/エッジと、の間の距離の最小化）

これらの4つのタイプの観測を加えることによって、本システムは、テクスチャが少ないもしくは無いけれどもエッジ/ラインは存在するような環境においても、機能すること

10

20

30

40

50

ができる。ライン/エッジは、典型的には、壁の間のエッジ、壁と床の間のエッジ、壁と天井との間のエッジ、ドアと壁との間のエッジ、等である。

【 0 0 2 7 】

4 本発明の開示に用いられている用語の定義

【 0 0 2 8 】

#### 4.1 ポーズ

3 D位置 (  $x, y, z$  ) + 3 D方向 (  $x, y, z$  軸回りの回転 )。方向は、多くの場合、回転行列、オイラー角、四元数マップもしくは「exponential map」を用いてパラメータ化される。位置は、通常、ベクトル ( 3次元もしくは相次 (homogenous) 4次元 ) としてパラメータ化される。

【 0 0 2 9 】

#### 4.2 パーチャル・カメラ

内部カメラパラメータおよび外部カメラパラメータを含む物理的カメラを数学的に表したものの。

【 0 0 3 0 】

#### 4.3 内部カメラパラメータ

焦点距離、画像中心、解像度、レンズの歪曲パラメータ ( ラジアル方向歪曲、接線方向歪曲 )。

【 0 0 3 1 】

#### 4.4 外部カメラパラメータ

特定の座標系における物理的カメラのポーズ。

【 0 0 3 2 】

#### 4.5 キーポイントないし特徴点

画像の中の関心のある点。キーポイントの検出に利用できるアルゴリズムは、いくつか存在する。特徴点とは、コーナ、コントラストの急変、エッジ、などである。画像から特徴点を抽出するために用いられるいくつかの方法が存在する。我々の方法は、コーナを検出するものであり、2005年の「Foundations and Trend in Computer Graphics and Vision」におけるVincent LepetitおよびPascal Fuaによる「Monocular Model-based 3d Tracking of Rigid Objects」の4.4.1章の記述に基づいている。

【 0 0 3 3 】

#### 4.6 数値最小化

数値最小化とは、変数の制約条件を受ける数学的関数の最小化を意味する。数値最小化の一例は、Levenberg-Marquardt法である。

【 0 0 3 4 】

#### 4.7 Levenberg-Marquardt

Wikipedia.orgによれば、「Levenberg-Marquardtアルゴリズムは、一般に非線形の関数について該関数のパラメータの間隔に亘る関数の最小化の数学問題に数値解を与える。この最小化問題は、特に、最小自乗曲線フィッティングにおいて起こる。Levenberg-Marquardtアルゴリズム (LMA) は、ガウス - ニュートン法 (GNA) と勾配法とを融合したものである。」とある。最小化の状態に応じて、LMAはGNAと勾配法から最良のものをとる。

【 0 0 3 5 】

#### 4.8 2 D - 3 D の点相関関係

画像中の検出された点は、所定のイメージピクセル座標における2 D位置を有する。同じ点が、実世界においては、3 D位置を有する。最小の3つの点およびこれらの2 D - 3 D相関関係を有していれば、画像の撮影に用いた物理的カメラのポーズの推定が可能である (例えば、P3Pアルゴリズムを通して)。

【 0 0 3 6 】

#### 4.9 観測 (Observation)

数値最小化を論ずる際に、観測の一例は、2 D - 3 D点相関関係である。物理的カメラのポーズの推定のために、運動モデルを2 D - 3 D点相関関係と組み合わせて用いる場合

10

20

30

40

50



には、前回の繰り返し (iteration) / フレームから現在の繰り返し / フレームの間での 3 軸の 1 つあるいはいくつかにける位置の変化やいずれかの軸の回りの回転変化もまた観測となり得る。

【 0 0 3 7 】

#### 4.10 アウトライアー (Outlier)

無効な観測である。観測として 2 D - 3 D 点相関関係を論ずる場合には、アウトライアーは、3 D 位置が 2 D 位置に誤ってマッチングしている誤った 2 D - 3 D 点相関関係である。

【 0 0 3 8 】

#### 4.11 インライアー

インライアーとは観測である。

【 0 0 3 9 】

#### 4.12 キーポイントデータベース

トレーニングフェーズの間に記憶されるキーポイントおよびそのプロパティのデータベースである。

【 0 0 4 0 】

#### 4.13 分類ツリー

分類を試みるいくつかのテストを含むツリー状階層化構造であり、我々の事例では、キーポイント / 特徴点を分類する。このツリーは、「さかさま」に使われ、つまり、各キーポイントが「ドロップダウン」型のツリーであり、ルートから始まって葉において終了する。各ノードにおいてテストがなされ、このテストによりキーポイントが他の 1 つの枝へと下方へ進む。最終的にキーポイントは 1 つの葉 ( 枝の末端 ) に達する。この枝は、キーポイントを識別するために必要な情報を含む。キーポイントの認識のために分類ツリーをどのように用いるか、については、2005 年 6 月のカルフォルニア州サンディエゴでの「Conference on Computer Vision and Pattern Recognition」における V. Lepetit、P. Lagger および P. Fua による「Randomized Trees for Real-Time Keypoint Recognition」というタイトルの開示を参照されたい。

【 0 0 4 1 】

#### 4.14 ロバスト・エスティメータ (estimator) を備えた PROSAC

PROSAC は、多数のアウトライアーデータが存在するモデルのロバストな当てはめ (fitting) のためのアルゴリズムである。我々の事例では、大まかな初期カメラポーズの推定のために PROSAC を用いるが、ここでは、多数の観測 ( 2 D - 3 D 点相関関係 ) がアウトライアーである。

観測として 2 D - 3 D 点相関関係のセットが与えられたときに、我々の PROSAC の各ステップは次のように実行される。

1 . 観測をそのスコアに従ってソートし、最高スコアを最上部に一致させる。

2 . 観測を選ぶために、観測の配列の最上部からの観測の数である  $n$  を、2 にセットする。

3 .  $n = n + 1$  にセットする。

4 . もし  $n$  が観測の数よりも大であれば、エラーとして中止する。最小の再射影誤差および最小のインライアーの個数とする所定の要求条件では、満足できる初期ポーズは検出されていない。

5 .  $n$  個の上部の列に限定して、配列の上部から 3 個の観測をランダムに選択する。

6 . この 3 個のランダムに選択した観測を入力として、P3P を用いて、カメラポーズを演算する。特異点を回避するために、P3P から得た回転行列を「exponential map」に変換する。「exponential map」については、3.4.1.1 章を参照されたい。P3P 以外のアルゴリズムも、カメラポーズの演算に適しているかもしれない。

7 . 配列の中の全ての観測のために、

a . 現在演算されたカメラポーズを用いて、再射影誤差 ( 4.26 章を参照 ) を演算し、

10

20

30

40

50

b . アウトライアーを無視するために、ロバスト・エスティメータ ( estimator ) を適用する。例：Tukey...もし再射影誤差が所定の閾値 ( 例えば 1 5 ピクセル ) よりも大であれば、観測がアウトライアーであると記録する。

8 . 全てのインライアーの再射影誤差の和をインライアーの個数で除して、残差を算出する ( 残差 = インライアーの再射影誤差の平均 ) 。

9 . 残差が、前回の繰り返しにおいて算出された最小の残差よりも低いか否か判定し、

a . Y E S であれば、

i . 現在のカメラポーズを最良のポーズとして記憶し、かつ対応する残差を記憶し、

ii . インライアーの数がインライアーの所定の個数 ( 例えば 2 0 個 ) 以上であるかチェックし、

1 . N O であれば、ステップ 3 ~ 7 を繰り返し、

2 . Y E S であれば、ステップ 4 で演算されたポーズを初期ポーズとして許容して、処理を中止し、

b . N O であれば、ステップ 3 ~ 7 を繰り返す。

PROSAC についてのさらなる情報については、2 0 0 5 年の C V P R での Chum、Ondrej、Matas、Jiri による「Matching with PROSAC - Progressive Sample Consensus」を参照されたい。

【 0 0 4 2 】

4.15 P 3 P

「Perspective-Three-Point」：幾何学的推定の 1 つの方法であり、例えば、内部カメラパラメータが既知のときに、最小の 3 つの 2 D - 3 D 点相関関係に基づき、物理的カメラのポーズを推定する。包括的な解を確実に発見するためには、4 つの 2 D - 3 D 点相関関係が必要である。

以下の P 3 P の説明について、ここで、添付した図 7 を参照する。

$M_i$  および  $M_j$  の各対の相関関係は、( 未知の ) カメラ - 点間の距離  $x_i = ||M_i - C||$  および  $x_j = ||M_j - C||$  に制約条件を与える。

$$d_{ij}^2 = x_i^2 + x_j^2 - 2 x_i x_j \cos \theta_{ij}$$

但し、

$d_{ij} = ||M_i - M_j||$  は、 $M_i$  と  $M_j$  の間の ( 既知の ) 距離、

$\theta_{ij}$  は、カメラ中心において  $M_i$  と  $M_j$  で定まる角度である。

アルゴリズム

1 . 距離  $x_i$  について多項式を解く。

$$d_{ij}^2 = x_i^2 + x_j^2 - 2 x_i x_j \cos \theta_{ij}$$
$$f_{ij}(x_i, x_j) = x_i^2 + x_j^2 - 2 x_i x_j \cos \theta_{ij} - d_{ij}^2 = 0$$

2 . 点  $M_i$  のカメラ座標系における位置  $M_i^C$  は演算可能である。

3 . 「R」および「T」は、 $M_i$  から  $M_i^C$  へのユークリッド変位として演算される。

最良の最小二乗回転は、四元数を用いたクローズドフォームにおいて得られる。

( 1 9 8 7 年 4 月、Horn, B.K.P. 「Closed Form Solutions of Absolute Orientation Using Unit Quaternions」JOSA-A(4), No. 4, 6 2 9 ~ 6 4 2 頁 )

定義、図および説明は、Vincent Lepetit による「Selected Topics in Computer Vision」の講座 1 からのものである。

我々が実施する P 3 P は、また、2 0 0 3 年 8 月の P A M I の I E E E 会報 vol.25, No . 8 の X. S. Gao, X.-R. Hou, J. Tang, H.-F. Chang による「Complete Solution Classification for the Perspective-Three-Point Problem」に基づいている。

【 0 0 4 3 】

4.16 レイ・トレーシング

物理的カメラのポーズおよび内部カメラパラメータが既知で、かつ現実環境の 3 D モデ

10

20

30

40

50

ルが与えられれば、画像中に検出したいかなる2Dキーポイントの位置も演算可能である。これは、カメラ中心を始点として画像中の2Dキーポイント座標を通りかつ現実環境の3Dモデルに衝突する3次元ベクトルを、数学的に記述することでなされる。このようなベクトルは、「光線」においても見ることができ、この光線が3Dモデルと衝突するまでトレースされるので、それ故、この処理は、「レイ・トレーシング」と呼ばれる。そして、カメラの内部パラメータおよび外部パラメータが既に演算されていることを条件として、その衝突点が、画像中の検出した2Dキーポイントの現実世界の3D座標を与えるのである。

【0044】

#### 4.17 運動モデル

運動モデルとは、オブジェクトや物理的カメラ等の物理的運動を制約および/又は予測する数学的モデルである。これは、今回推定したポーズの「正確さ (correctness)」を改善するために、今回および前回の繰り返し (iteration) / フレーム / 画像からの情報を用いる。

【0045】

#### 4.18 ドリフト (Drifting)

ドリフトは、フレーム・トゥ・フレーム・トラッキングに共通の問題である。良好かつ正確なポーズを有するフレームから開始しても、いくつか後のフレームでは、正しい解から離れるようにポーズが「ドリフト」し、つまり、時間が経過する (将来のフレームが処理される) に従って、ポーズの推定結果はさらにドリフトし、正しい解からさらに離れる。

ドリフトは、この種の方法の基本が所定のフレームから開始し、この「最初のフレーム」に基づいて後続のフレームでのポーズを演算するために、生じる。新しいフレームを処理する各々のたびに、多少の数値的誤差や未知数 (uncertainties) が推定したポーズに加わる。誤差は、各フレーム毎に「それ自身」の上に加わり、つまり、誤差の影響は、僅か数フレームのトラッキングの後に急激に増加する。

【0046】

#### 4.19 ロバスト性

カメラが素早く動くラフな環境の下では、カメラポーズを演算する能力は、部分的ないし完全に妨げられ、不正確な特徴点 / キーポイントが推定に用いられる。

【0047】

#### 4.20 精度 (Accuracy)

ポーズ推定の観点では、精度とは、精密かつ正確に物理的カメラポーズを演算する能力である。

【0048】

#### 4.21 リピータビリティ

ポーズ推定の観点では、リピータビリティとは、カメラがごく僅か動き、実際 / 現実のポーズが微小値だけ変化した後続のフレーム / 画像において、殆ど同一の物理的カメラポーズを演算し得る能力である。

【0049】

#### 4.22 デテクションもしくはオブジェクト・デテクション

デテクションもしくはオブジェクト・デテクションとは、推定すべき前のフレームの情報を要求することなく、個々のフレームにおけるオブジェクト (あるいはオブジェクトに沿った1セットのキーポイント) を検出する処理のことである。

【0050】

#### 4.23 トラッキングもしくはフレーム・トゥ・フレーム・トラッキング

カメラのフレーム毎のトラッキングを実行し、前に推定したカメラポーズを、キーポイントのマッチングを介して、新しい / 将来のフレームにおけるカメラポーズの演算に利用する処理を、トラッキングもしくはフレーム・トゥ・フレーム・トラッキングと呼ぶ。

この方法は、あるフレームにおける既知のポーズを、後続のフレームにおけるポーズの

10

20

30

40

50

推定に用いる。最初の既知となるポーズを見つけるためには、種々の方法がある。後続のフレームにおけるポーズの推定は、下記のような手順でなされる。

1. できるだけ良い特徴点を検出し、次のフレームで使用するために、保存する。
2. 次のフレームにおいて、新たなできるだけ良い特徴点を検出する。
3. 前のフレームにおいてポーズが正しく推定 / 提供されていれば、特徴点を前のフレームとマッチングする（特徴点のマッチングについては、4.25章を参照されたい）。
4. マッチした各々の特徴点について

a. 前のカメラポーズおよび内部パラメータを用いて、特徴点の3D位置を演算する。これは、前のポーズにおけるカメラ中心から、前のフレームで検出した特徴点の2D位置を通り、かつ環境の3Dモデルへと、レイ・トレーシングを実行することにより行う。

b. 前のフレームを用いて見いだした3D位置を、現在のフレームにおける同じ特徴点の2D位置とともに、ポーズ推定の基礎として用いる。

5. 全てのマッチした特徴点の2D - 3D点相関関係を、数値最適化法例えばLevenberg-Marquardt法による物理的カメラポーズの推定のために用いる。

6. アルゴリズムが不良となるまで、ステップ2～5を繰り返す。アルゴリズムがトラッキングから外れると不良となるが、これは、前のフレームと今回のフレームとの間で十分なマッチした特徴点を見いだせなかった場合あるいは前のフレームでのポーズが不正確な推定であったり推定されていなかった場合、に生じる。

ステップ1は、初期化ステップであり、ここでは、カメラポーズが既知であると仮定している。

ステップ2～5は、トラッキングの要求条件が成立している限り、新しいフレーム毎に繰り返される。この要求条件は、前のフレームについてポーズが正しく推定つまり提供されていること、そして、有効なポーズを推定し得るように、4つより多い特徴点が現在のフレームと前のフレームとの間でマッチしていること、である。

【0051】

#### 4.24 残差 (Residual)

「残差とは、観測不能な誤差の観測可能な推定量である。残差は観測可能であるが、誤差は観測できない。」

([http://en.wikipedia.org/wiki/Errors\\_and\\_residuals\\_in\\_statistics](http://en.wikipedia.org/wiki/Errors_and_residuals_in_statistics))

【0052】

#### 4.25 キーポイントマッチング

画像間でキーポイントをマッチングするには、いくつかのやり方がある。処理の例は、1995年のZ. Zhang, R. Deriche, O. Faugeras, Q. Luongによる「A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry」Artificial Intelligence, vol.78の87～119頁に説明されている。

本発明は、近接状態および輝度の相互対応関係に基づくマッチングを用いる。運動の推定のために、似た（通常は連続した）視点からの2つの画像から抽出した関心を引く点「 $m_i$ 」および「 $m'_j$ 」のセットのマッチングを行うことができる。第1の画像の各点「 $m_i$ 」に対して、第2の画像において「 $m_i$ 」の周囲で点「 $m'_j$ 」を探索する。この探索は、各点を中心とした局所的な画像ウィンドウの類似性に基づく。この類似性は、ゼロ正規化された相互対応関係を用いて評価され、これは、局所的な画像輝度のアフィン変化に対し不変であって、照度変化に対し処理をロバストなものとしている。より信頼できるマッチングのセットを得るために、我々は、2つの画像の立場を逆にして上記の処理を繰り返すようにしている。互いに選択された点の相関関係  $m_i$   $m'_j$  のみが保持される。

(2005年の「Foundations and Trend in Computer Graphics and Vision」におけるVincent LepetitおよびPascal Fuaによる「Monocular Model-based 3d Tracking of Rigid Objects」の4.4.2章の方法および記述に基づく)。

【0053】

#### 4.26 再射影誤差

10

20

30

40

50

再射影誤差とは、特徴点の測定 / 検出した 2 D 座標と、その対応する 3 D 座標の射影 2 D 座標と、の間の距離である。再射影誤差は下記のように記述できる。

$$||A[R|T]M_i - m_i||^2$$

ここで、A は、3 × 3 の内部カメラパラメータ行列、[R|T] は、変換行列 (R : 回転行列、T : 並進ベクトル)、M<sub>i</sub> は特徴点の 3 D 位置、m<sub>i</sub> は特徴点の測定 / 検出した 2 D 座標、である。

A :

$$\begin{array}{c|ccc|}
 & fx & 0 & cx & \\
 & 0 & fy & cy & \\
 & 0 & 0 & 1 & \\
 \hline
 \end{array}$$

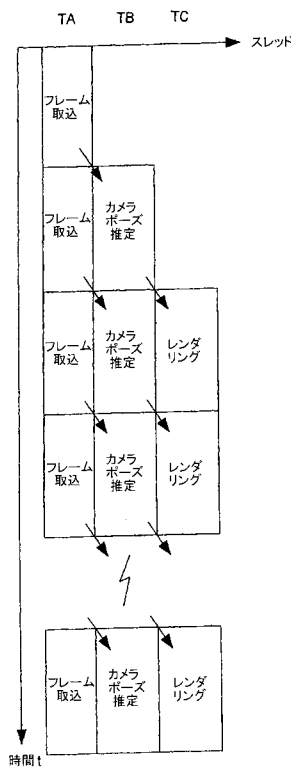
但し、fx、fy は焦点距離、cx、cy は主点 (画像中心) である。

【 0 0 5 4 】

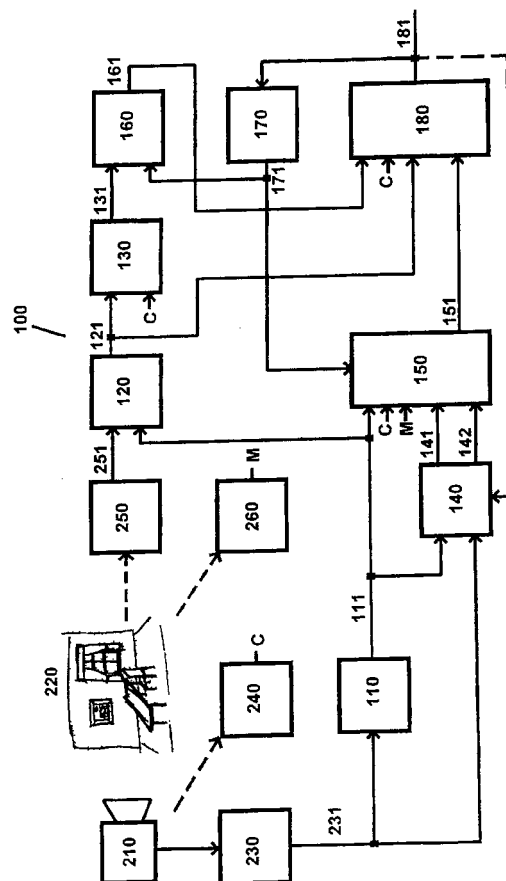
#### 4.27 再射影誤差の最小化

数値最小化の各繰り返しの際に外部パラメータを変更し、全ての測定 2 D キーポイントと射影 2 D キーポイントとの間の距離の平均が最小となる外部パラメータを記憶することで、最適解が得られる。各繰り返しの際の外部パラメータの導関数は、やはり演算される。一般に、導関数によって、より小さい平均の再射影誤差を得るために次の繰り返しにおける外部パラメータを、どのように、またどの程度に、変更すべきか、が決定される。

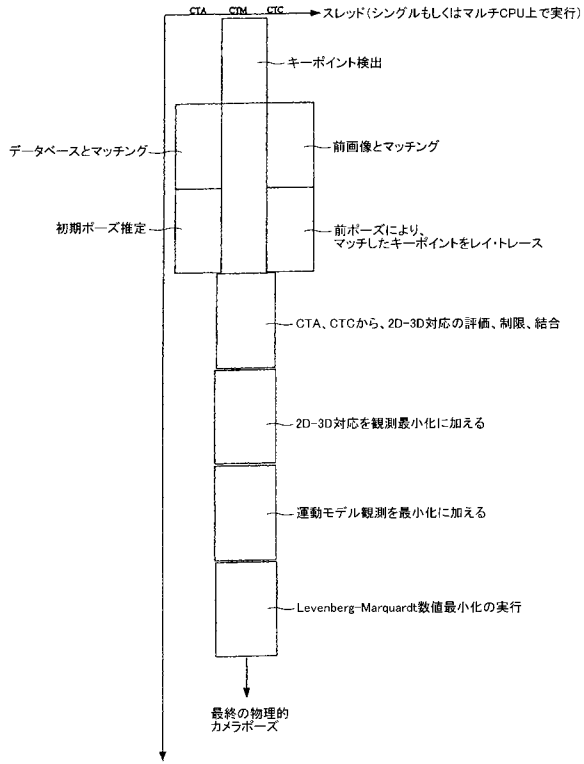
【 図 1 】



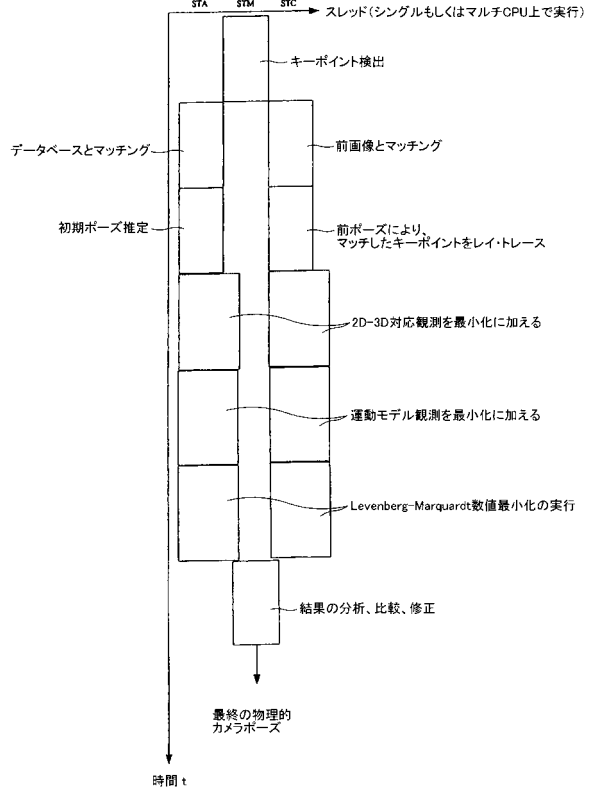
【 図 2 】



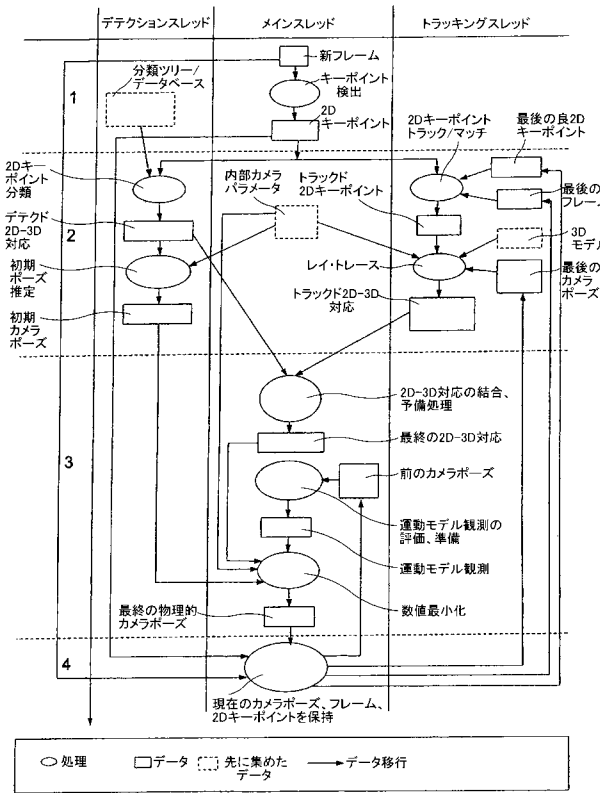
【 図 3 】



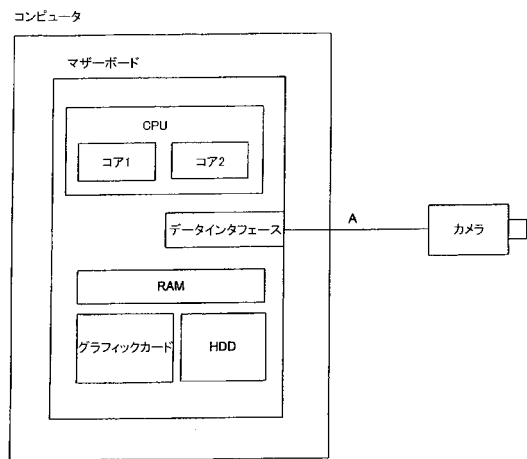
【 図 4 】



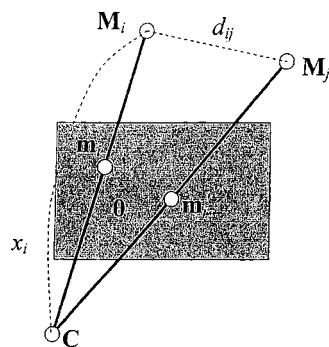
【 図 5 】



【 図 6 】



【 図 7 】



## 【 国際調査報告 】

1

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/NO2008/000178

<b>A. CLASSIFICATION OF SUBJECT MATTER</b>		
IPC: see extra sheet According to International Patent Classification (IPC) or to both national classification and IPC		
<b>B. FIELDS SEARCHED</b>		
Minimum documentation searched (classification system followed by classification symbols)		
IPC: H04N, G06T, G06K		
Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched		
SE,DK,FI,NO classes as above		
Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)		
EPO-INTERNAL, WPI DATA, PAJ, INSPEC		
<b>C. DOCUMENTS CONSIDERED TO BE RELEVANT</b>		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,A	US 20070122001 A1 (WANG ET AL), 31 May 2007 (31.05.2007)	1-13
	--	
A	US 20050078178 A1 (BROWN ET AL), 14 April 2005 (14.04.2005)	1-13
	--	
A	US 20060188131 A1 (ZHANG ET AL), 24 August 2006 (24.08.2006)	1-13
	--	
A	US 20040175039 A1 (MILLER), 9 Sept 2004 (09.09.2004)	1-13
	--	
<input checked="" type="checkbox"/> Further documents are listed in the continuation of Box C. <input checked="" type="checkbox"/> See patent family annex.		
* Special categories of cited documents: "A" document defining the general state of the art which is not considered to be of particular relevance "E" earlier application or patent but published on or after the international filing date "L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified) "O" document referring to an oral disclosure, use, exhibition or other means "P" document published prior to the international filing date but later than the priority date claimed "T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention "X" document of particular relevance: the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone "Y" document of particular relevance: the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art "&" document member of the same patent family		
Date of the actual completion of the international search		Date of mailing of the international search report
8 October 2008		15-10-2008
Name and mailing address of the ISA/ Swedish Patent Office Box 5055, S-102 42 STOCKHOLM Facsimile No. +46 8 666 02 86		Authorized officer Henrik Andersson/PR Telephone No. +46 8 782 25 00

## INTERNATIONAL SEARCH REPORT

International application No.

PCT/NO2008/000178

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT		
Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 20040190775 A1 (MILLER), 30 Sept 2004 (30.09.2004)  --	1-13
A	US 20030055335 A1 (SAUER ET AL), 20 March 2003 (20.03.2003)  --	1-13
A	US 6571024 B1 (SAMHNEY ET AL), 27 May 2003 (27.05.2003)  --	1-13
A	US 20030012410 A1 (NAVAB ET AL), 16 January 2003 (16.01.2003)  -- -----	1-13



## INTERNATIONAL SEARCH REPORT

International application No. PCT/NO2008/000178
--

**International patent classification (IPC)****G06T 7/00** (2006.01)**G06K 9/00** (2006.01)**G06T 7/20** (2006.01)**G06K 9/62** (2006.01)**H04N 7/18** (2006.01)**Download your patent documents at [www.prv.se](http://www.prv.se)**

The cited patent documents can be downloaded at [www.prv.se](http://www.prv.se) by following the links:

- In English/Searches and advisory services/Cited documents (service in English) or
- e-tjänster/anförda dokument (service in Swedish).

Use the application number as username.

The password is **IYCUGBINXX**.

Paper copies can be ordered at a cost of 50 SEK per copy from PRV InterPat (telephone number 08-782 28 85).

Cited literature, if any, will be enclosed in paper form.

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

30/08/2008

International application No.

PCT/NO2008/000178

US	20070122001	A1	31/05/2007	NONE		
US	20050078178	A1	14/04/2005	US	7336296 B	26/02/2008
US	20060188131	A1	24/08/2006	NONE		
US	20040175039	A1	09/09/2004	CA	2524185 A	09/12/2004
				CN	1717717 A	04/01/2006
				EP	1599828 A	30/11/2005
				EP	1599829 A	30/11/2005
				EP	1599830 A	30/11/2005
				EP	1634274 A	15/03/2006
				JP	2006520054 T	31/08/2006
				JP	2006520055 T	31/08/2006
				JP	2006522411 T	28/09/2006
				JP	2007528037 T	04/10/2007
				KR	20060004689 A	12/01/2006
				US	20040175041 A	09/09/2004
				US	20040190775 A	30/09/2004
				US	20040243415 A	02/12/2004
				WO	2004081853 A	23/09/2004
				WO	2004081854 A	23/09/2004
				WO	2004081855 A	23/09/2004
				WO	2004107315 A	09/12/2004
				US	6801424 B	05/10/2004
US	20040190775	A1	30/09/2004	CA	2524185 A	09/12/2004
				CN	1717717 A	04/01/2006
				EP	1599828 A	30/11/2005
				EP	1599829 A	30/11/2005
				EP	1599830 A	30/11/2005
				EP	1634274 A	15/03/2006
				JP	2006520054 T	31/08/2006
				JP	2006520055 T	31/08/2006
				JP	2006522411 T	28/09/2006
				JP	2007528037 T	04/10/2007
				KR	20060004689 A	12/01/2006
				US	20040175039 A	09/09/2004
				US	20040175041 A	09/09/2004
				US	20040243415 A	02/12/2004
				WO	2004081853 A	23/09/2004
				WO	2004081854 A	23/09/2004
				WO	2004081855 A	23/09/2004
				WO	2004107315 A	09/12/2004
				US	6801424 B	05/10/2004
US	20030055335	A1	20/03/2003	US	6612991 B	02/09/2003
				US	6695779 B	24/02/2004
				US	7079132 B	18/07/2006
				US	7190378 B	13/03/2007
				US	20030060706 A	27/03/2003
				US	20030063132 A	03/04/2003
				US	20030117393 A	26/06/2003
				US	20030120155 A	26/06/2003
US	6571024	B1	27/05/2003	NONE		

**INTERNATIONAL SEARCH REPORT**  
Information on patent family members

30/08/2008

International application No.  
**PCT/NO2008/000178**

US 20030012410 A1 16/01/2003 NONE

---

## フロントページの続き

(81)指定国 AP(BW, GH, GM, KE, LS, MW, MZ, NA, SD, SL, SZ, TZ, UG, ZM, ZW), EA(AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), EP(AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MT, NL, NO, PL, PT, RO, SE, SI, SK, TR), OA(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG), AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IS, JP, KE, KG, KM, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LT, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RS, RU, SC, SD, SE, SG, SK, SL, SM, SV, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW

(特許庁注：以下のものは登録商標)

- 1 . L i n u x
- 2 . W I N D O W S
- 3 . U N I X

Fターム(参考) 2F065 AA04 AA12 AA17 AA22 AA31 CC16 FF01 JJ03 JJ26 QQ08  
 QQ18 QQ24 QQ28 QQ29 RR08 SS13  
 5B050 BA09 BA11 EA18 EA19 EA26  
 5B057 BA02 CA08 CA12 CA16 CB13 CB18 CC03 CD14 CH08 CH11  
 CH18 DA06 DA17 DB02 DB09 DC02 DC08 DC32 DC36

## 【要約の続き】

クド・3D観測を決定する。ポーズ推定セクタは、デテクトド・3Dカメラポーズ推定と前の3Dカメラポーズ推定の中の選択した一方を出力する。3Dカメラポーズ・エスティメータは、カメラデータと、デテクトド・3D観測と、トラックド・3D観測と、選択後3Dカメラポーズ推定と、から、3Dカメラポーズ推定を演算し、かつ出力する。