

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2016-192132

(P2016-192132A)

(43) 公開日 平成28年11月10日(2016.11.10)

(51) Int.Cl.			F I			テーマコード (参考)		
<b>G06T</b>	<b>7/20</b>	<b>(2006.01)</b>	G06T	7/20		B	5B050	
<b>G06T</b>	<b>19/00</b>	<b>(2011.01)</b>	G06T	19/00		600	5L096	
<b>G06T</b>	<b>7/00</b>	<b>(2006.01)</b>	G06T	7/00		300D		

審査請求 未請求 請求項の数 11 O L (全 16 頁)

(21) 出願番号 特願2015-72514 (P2015-72514)  
 (22) 出願日 平成27年3月31日 (2015. 3. 31)

(71) 出願人 000208891  
 K D D I 株式会社  
 東京都新宿区西新宿二丁目3番2号  
 (74) 代理人 100092772  
 弁理士 阪本 清孝  
 (74) 代理人 100084870  
 弁理士 田中 香樹  
 (74) 代理人 100119688  
 弁理士 田邊 壽二  
 (72) 発明者 小林 達也  
 埼玉県ふじみ野市大原二丁目1番15号  
 株式会社K D D I 研究所内  
 (72) 発明者 加藤 晴久  
 埼玉県ふじみ野市大原二丁目1番15号  
 株式会社K D D I 研究所内

最終頁に続く

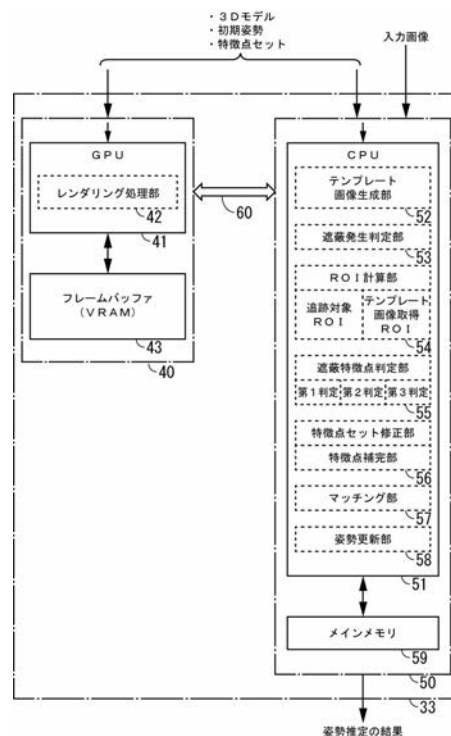
(54) 【発明の名称】 画像認識AR装置並びにその姿勢推定装置及び姿勢追跡装置

(57) 【要約】

【課題】 追跡対象ごとに設定した多数の特徴点のマッチングによる追跡中にオクルージョンが発生し、特徴点の一部が遮蔽されても、高速かつ高精度な追跡を継続する。

【解決手段】 レンダリング処理部42は、複数の追跡対象の3Dモデル、初期姿勢および特徴点に基づく3Dレンダリングにより各追跡対象のレンダリング結果をフレームバッファ43上に展開する。テンプレート画像生成部52はレンダリング結果をバス経由で読み取ってテンプレート画像を取得する。遮蔽特徴点判定部55は、オクルージョンにより遮蔽される特徴点を判定する。特徴点セット修正部56は、遮蔽された特徴点をマッチング用の特徴点から除外する。マッチング部57はフレーム画像と各テンプレート画像との間で特徴点マッチングを実行する。姿勢更新部58はマッチング結果に基づいて追跡対象の姿勢を更新する。レンダリング処理部42は更新後の姿勢を次フレームの初期姿勢に反映する。

【選択図】 図2



**【特許請求の範囲】****【請求項 1】**

3Dレンダリングにより得られた追跡対象のテンプレート画像をバス経由で読み取り、フレーム画像とテンプレート画像とのマッチング結果に基づいて追跡対象の姿勢を更新する姿勢追跡装置において、

複数の追跡対象の3Dモデルを、初期姿勢および特徴点に基づいてフレームバッファ上にレンダリングするレンダリング手段と、

前記レンダリング結果をバス経由で読み取ってメインメモリ上にテンプレート画像を生成するテンプレート画像生成手段と、

前記テンプレート画像を対象に、オクルージョンにより遮蔽される特徴点を判定する遮蔽特徴点判定手段と

前記遮蔽された特徴点をマッチング用の特徴点から除外する特徴点修正手段と、

フレーム画像と各テンプレート画像との間で前記修正後の特徴点を用いたマッチングを実行するマッチング手段と、

前記マッチング結果に基づいて追跡対象の姿勢を更新する姿勢更新手段とを具備し、

前記レンダリング手段は、更新後の姿勢を次フレームの初期姿勢に反映することを特徴とする姿勢追跡装置。

**【請求項 2】**

複数の追跡対象の論理和領域を一のテンプレート画像取得ROIに設定するROI計算手段をさらに具備し、

前記テンプレート画像生成手段は、前記一のテンプレート画像取得ROIをバス経由で読み取ってテンプレート画像とすることを特徴とする請求項 1 に記載の姿勢追跡装置。

**【請求項 3】**

前記特徴点を遮蔽された追跡対象の特徴点群から、当該遮蔽された特徴点の代替えとなる特徴点を選択して前記マッチング用の特徴点を補完する手段をさらに具備し、

前記マッチング手段は、補完後の特徴点を用いてマッチングを実行することを特徴とする請求項 1 または 2 に記載の姿勢追跡装置。

**【請求項 4】**

前記遮蔽特徴点判定手段は、デプスバッファを使用して各特徴点の遮蔽判定を行う第 1 判定方法、特徴点の幾何変換を利用して特徴点全体の前後関係を一括判定する第 2 判定方法、および3Dモデルの頂点の幾何変換を利用して特徴点全体の前後関係を一括判定する第 3 判定方法のいずれかにより、前記オクルージョンにより遮蔽される特徴点を判定することを特徴とする 1 ないし 3 のいずれかに記載の姿勢追跡装置。

**【請求項 5】**

前記遮蔽特徴点判定手段は、3Dモデルの頂点数が所定数よりも少なければ第 3 判定方法を採用することを特徴とする請求項 4 に記載の姿勢追跡装置。

**【請求項 6】**

3Dレンダリングにより得られた追跡対象のテンプレート画像をバス経由で読み取り、フレーム画像とテンプレート画像とのマッチング結果に基づいて姿勢推定を行う姿勢推定装置において、

追跡対象の姿勢ごとに特徴点を含む画像特徴を蓄積する画像特徴DBと、

フレーム画像から特徴点を検出する手段と、

前記蓄積された特徴点と前記検出された特徴点との間で特徴点マッチングを行うマッチング手段と、

前記マッチング結果に基づいて追跡対象の初期姿勢を推定する手段と、

追跡対象ごとに、その初期姿勢、3Dモデルおよび特徴点に基づいて姿勢追跡を実行する姿勢追跡装置とを具備し、

前記姿勢追跡装置は、

複数の追跡対象の3Dモデルを、初期姿勢および特徴点に基づいてフレームバッファ上にレンダリングするレンダリング手段と、

10

20

30

40

50

前記レンダリング結果をバス経由で読み取ってメインメモリ上にテンプレート画像を生成するテンプレート画像生成手段と、

前記テンプレート画像を対象に、オクルージョンにより遮蔽される特徴点を判定する遮蔽特徴点判定手段と

前記遮蔽された特徴点をマッチング用の特徴点から除外する特徴点修正手段と、

フレーム画像と各テンプレート画像との間で前記修正後の特徴点を用いたマッチングを実行するマッチング手段と、

前記マッチング結果に基づいて追跡対象の姿勢を更新する姿勢更新手段とを具備し、

前記レンダリング手段は、更新後の姿勢を次フレームの初期姿勢に反映することを特徴とする姿勢推定装置。

10

【請求項 7】

複数の追跡対象の論理和領域を一のテンプレート画像取得ROIに設定するROI計算手段をさらに具備し、

前記テンプレート画像生成手段は、前記一のテンプレート画像取得ROIをバス経由で読み取ってテンプレート画像とすることを特徴とする請求項 6 に記載の姿勢推定装置。

【請求項 8】

前記特徴点を遮蔽された追跡対象の特徴点群から、当該遮蔽された特徴点の代替えとなる特徴点を選択して前記マッチング用の特徴点を補完する手段をさらに具備し、

前記マッチング手段は、補完後の特徴点を用いてマッチングを実行することを特徴とする請求項 6 または 7 に記載の姿勢推定装置。

20

【請求項 9】

フレーム画像を取得する手段と、

フレーム画像内で追跡対象の姿勢を推定する手段と、

AR情報が蓄積されたAR情報DBと、

前記AR情報を前記取得画像に重畳表示する重畳表示手段とを具備し、

前記姿勢を推定する手段は、

追跡対象の姿勢ごとに特徴点を含む画像特徴を蓄積する画像特徴DBと、

フレーム画像から特徴点を検出する手段と、

前記蓄積された特徴点と前記検出された特徴点との間で特徴点マッチングを行うマッチング手段と、

30

前記マッチング結果に基づいて追跡対象の初期姿勢を推定する手段と、

追跡対象ごとに、その初期姿勢、3Dモデルおよび特徴点に基づいて姿勢追跡を実行する姿勢追跡装置とを具備し、

前記姿勢追跡装置(33)は、

複数の追跡対象の3Dモデルを、初期姿勢および特徴点に基づいてフレームバッファ上にレンダリングするレンダリング手段と、

前記レンダリング結果をバス経由で読み取ってメインメモリ上にテンプレート画像を生成するテンプレート画像生成手段と、

前記テンプレート画像を対象に、オクルージョンにより遮蔽される特徴点を判定する遮蔽特徴点判定手段と

40

前記遮蔽された特徴点をマッチング用の特徴点から除外する特徴点修正手段と、

フレーム画像と各テンプレート画像との間で前記修正後の特徴点を用いたマッチングを実行するマッチング手段と、

前記マッチング結果に基づいて追跡対象の姿勢を更新する姿勢更新手段とを具備し、

前記レンダリング手段は、更新後の姿勢を次フレームの初期姿勢に反映することを特徴とする画像認識AR装置。

【請求項 10】

複数の追跡対象の論理和領域を一のテンプレート画像取得ROIに設定するROI計算手段をさらに具備し、

前記テンプレート画像生成手段は、前記一のテンプレート画像取得ROIをバス経由で読

50

み取ってテンプレート画像とすることを特徴とする請求項 9 に記載の画像認識AR装置。

【請求項 11】

前記特徴点を遮蔽された追跡対象の特徴点群から、当該遮蔽された特徴点の代替えとなる特徴点を選択して前記マッチング用の特徴点を補完する手段をさらに具備し、

前記マッチング手段は、補完後の特徴点を用いてマッチングを実行することを特徴とする請求項 9 または 10 に記載の画像認識AR装置。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像認識AR装置並びにその姿勢推定装置および姿勢追跡装置に係り、特に、追跡対象ごとに設定した複数の追跡用の特徴点のマッチングによる追跡中にオクルージョンが発生して一部の特征点が追跡不能に陥っても、高速かつ高精度な追跡を継続できる画像認識AR装置並びにその姿勢推定装置及び姿勢追跡装置に関する。

10

【背景技術】

【0002】

近年、現実空間の画像（映像）をコンピュータで処理して仮想情報を重畳する拡張現実感（AR：Augmented Reality）技術が注目を集めている。AR技術を用いることで、ユーザの行動を支援することや、ユーザに直観的な情報提示を行うことが可能となる。

【0003】

例えば、身の回りの看板や広告にAR技術を適用することで、限られたスペースでは伝えることのできない詳細な情報や動画、3Dコンテンツ等を付加することや、場所や時間、閲覧者の属性等によって提示する情報を適切に変更することが可能となる。また、複雑な操作手順を必要とする装置のガイドにAR技術を適用することで、作業の効率化や、安全性の向上を実現することが可能となる。

20

【0004】

AR技術の主なプラットフォームは、以前はWebカメラを接続したラップトップPCであったが、近年は撮像装置（カメラ）や各種センサ、表示ディスプレイを搭載し、十分な処理性能を備えたタブレットデバイスやスマートフォン、HMD（Head Mounted Display）に代表されるウェアラブルデバイス等、様々な高性能モバイル端末が普及しつつあり、AR技術の主要なプラットフォームを担いつつある。

30

【0005】

これらAR技術を構成する要素技術の中では、仮想情報を正しい位置に重畳するための「位置合わせ技術」が特に重要な役割を担っている。仮想情報があたかも現実空間の重畳位置に存在するかのように見せるためには、カメラ及びディスプレイの移動に合わせて、高速かつ高精度に位置合わせ処理を実行する必要がある。

【0006】

特許文献1では、GPSや電子コンパス（方位センサ）、加速度センサといった各種センサデバイスを利用し、現実空間とカメラの相対的な位置関係（位置及び方向の6自由度のパラメータで決定される。以降「姿勢」と表現と表現する。）を求め、位置合わせ（以降「姿勢推定」と表現する。）を実現する技術が開示されている。

40

【0007】

この技術を利用したARサービスは既に広く提供されており、代表的なサービス例として、ナビゲーションサービスや、観光案内サービス等が挙げられる。センサに基づく姿勢推定は処理負荷が少ない利点がある一方、姿勢推定精度が十分でなく、異なる対象物（建物など）に仮想情報を重畳してしまうことが起こりうる。また、現実空間内に固定配置されたオブジェクトに対してのみ有効（仮想情報を重畳可能）という制約が存在する。

【0008】

特許文献2、3および非特許文献1には、カメラのプレビュー画像に対する画像解析処理によって姿勢を推定する技術が開示されている。これら画像解析による姿勢推定は、正確な姿勢推定と移動物体に対するARを実現できるという点が大きな利点である。推定精度

50

については、既存技術の時点で、仮想情報の表示位置の誤差を数ピクセル程度（ユーザには知覚できないレベル）に抑えることが可能であるため、作業ガイドの重畳表示等、正確な重畳が求められる場合に相性が良い。

【0009】

特に特許文献2では、エッジや特徴点等の画像特徴を利用し、事前登録された三次元オブジェクトの姿勢をリアルタイム推定するための技術が開示されている。この手法は任意形状のオブジェクトを対象物として扱うことができるものの、処理負荷が高く、現在のARの主要なプラットフォームであるモバイル端末上でのリアルタイム処理を実現することが出来ない。

【0010】

一方、特許文献3や非特許文献1では、事前登録された平面/三次元オブジェクトの姿勢をモバイル端末でリアルタイム推定するための技術が開示されている。追跡中のオブジェクトの登録画像を、過去の姿勢から予測（例えば線形予測）した姿勢に基づき変換することで、追跡用のテンプレート画像を生成し、これを入力画像とマッチングする（追跡する）手法が採られている。

【0011】

対象物の移動が極端に大きい場合で無ければ、テンプレート画像と入力画像はかなり類似した画像となるため、視点変化に頑健かつ高速な追跡が可能になる。さらに、テンプレート画像と入力画像の両方をダウンサンプリングすることで画像ピラミッドを構成し、ピラミッド上位の低解像度画像から順に追跡処理を行うことで、処理負荷を削減する手法が採られている。

【先行技術文献】

【特許文献】

【0012】

【特許文献1】特開2010-238096号公報

【特許文献2】特許5291605号公報

【特許文献3】特表2013-508844号公報

【非特許文献】

【0013】

【非特許文献1】D. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, and D. Schmalstieg, "Real-time detection and tracking for augmented reality on mobile phones," IEEE Trans. on Visualization and Computer Graphics, 2010.

【発明の概要】

【発明が解決しようとする課題】

【0014】

上記の従来技術は、追跡の対象物が1体であれば処理負荷および精度の両面で優れるものの、追跡対象が複数になった際に、処理負荷が線形に増加するという課題がある。また、対象物が互いに重なり合うオクルージョンへの対策が無く、頑健性が低下する技術課題がある。

【0015】

すなわち、追跡対象間のオクルージョンにより一方（下側）の追跡対象の特徴点が他方（上側）の追跡対象により遮蔽されてしまい、追跡用に使用できなくなっても、後段のマッチング機能はそれを認識できない。したがって、一方の追跡対象については、追跡できない特徴点のマッチングに要する無駄時間が発生し、また特徴点が増加することで追跡精度が低下するという技術課題があった。

【0016】

本発明の目的は、上記の技術課題を解決し、追跡対象ごとに設定した多数の特徴点のマッチングによる追跡中にオクルージョンが発生し、一方の追跡対象の特徴点が他方の追跡対象により遮蔽されても、高速かつ高精度な追跡を継続できる画像認識AR装置並びにその姿勢推定装置及び姿勢追跡装置を提供することにある。

10

20

30

40

50

**【課題を解決するための手段】****【0017】**

上記の目的を達成するために、本発明は、以下のような構成を具備した点に特徴がある。

**【0018】**

(1) 本発明の姿勢追跡装置は、複数の追跡対象の3Dモデルを、初期姿勢および特徴点に基づいてフレームバッファ上にレンダリングする手段と、レンダリング結果をバス経由で読み取ってメインメモリ上にテンプレート画像を生成する手段と、テンプレート画像を対象に、オクルージョンにより遮蔽される特徴点を判定する手段と、遮蔽された特徴点をマッチング用の特徴点から除外する手段と、フレーム画像と各テンプレート画像との間で前記修正後の特徴点を用いたマッチングを実行する手段と、マッチング結果に基づいて追跡対象の姿勢を更新する手段とを具備し、レンダリング手段は、更新後の姿勢を次フレームの初期姿勢に反映するようにした。

10

**【0019】**

(2) 本発明の姿勢推定装置は、追跡対象の姿勢ごとに特徴点を含む画像特徴を蓄積する画像特徴DBと、フレーム画像から特徴点を検出する手段と、蓄積された特徴点と検出された特徴点との間で特徴点マッチングを行う手段と、マッチング結果に基づいて追跡対象の初期姿勢を推定する手段と、追跡対象ごとに、その初期姿勢、3Dモデルおよび特徴点に基づいて姿勢追跡を実行する姿勢追跡装置として、上記(1)の姿勢追跡装置を用いた。

**【0020】**

(3) 本発明の画像認識AR装置は、フレーム画像を取得する手段と、フレーム画像内で追跡対象の姿勢を推定する手段と、AR情報が蓄積されたAR情報DBと、AR情報を取得画像に重畳表示する手段とを具備し、姿勢を推定する手段として、上記(2)の姿勢推定装置を用いた。

20

**【0021】**

(4) 複数の追跡対象の論理和領域を一のテンプレート画像取得ROIに設定し、テンプレート画像生成手段は、一のテンプレート画像取得ROIをバス経由で読み取ってテンプレート画像とするようにした。

**【発明の効果】****【0022】**

本発明によれば、3Dレンダリングにより得られた追跡対象のテンプレート画像をバス経由で読み取り、フレーム画像とテンプレート画像とのマッチング結果に基づいて追跡対象の姿勢を更新、推定する装置において、以下のような効果が達成される。

30

**【0023】**

(1) 対象物間のオクルージョンおよび遮蔽の状況を判定し、マッチング用の特徴点を修正するようにしたので、特徴点マッチングの頑健性を向上させることができる。

**【0024】**

(2) 複数の対象物を同時に追跡する際の3Dレンダリング画像の転送時間を短縮することで処理時間のボトルネックを解消し、処理負荷の増大を抑制したので、低スペックのモバイル端末による大規模なAR表示を実現することができる。

40

**【図面の簡単な説明】****【0025】**

【図1】本発明の一実施形態に係る画像認識AR装置1の主要部の構成を示した機能ブロック図である。

【図2】本発明の一実施形態に係る姿勢追跡装置の主要部の構成を示した機能ブロック図である。

【図3】テンプレート画像取得ROIの計算方法を示した図である。

【図4】オクルージョンの判定方法を模式的に表現した図である。

【図5】姿勢追跡の方法を示したフローチャートである。

**【発明を実施するための形態】**

50

## 【0026】

以下、図面を参照して本発明の実施の形態について説明する。図1は、本発明の一実施形態に係る画像認識AR装置1の主要部の構成を示した機能ブロック図であり、デスクトップPC等の据え置き型のコンピュータをはじめ、ラップトップPC、タブレットデバイス、携帯電話機、スマートフォン、携帯ゲーム機、HMD、腕時計型デバイス等、あらゆる携帯型の情報端末に搭載することができる。

## 【0027】

プレビュー取得部21は、携帯端末等に標準装備されているカメラモジュールあるいは外付けされたWEBカメラ等の撮像部から、姿勢追跡の対象物(以下、追跡対象と表現する場合もある)が写ったプレビュー画像をフレーム単位で取得する。

10

## 【0028】

姿勢推定装置22は、姿勢検出部31、画像特徴データベース(DB)32および姿勢追跡装置33を含み、今回フレームのプレビュー画像を対象に、後に詳述する特徴点の追跡を実施し、追跡対象とAR装置1との相対的な位置および姿勢の関係を推定する。

## 【0029】

AR情報データベース(DB)23は、ハードディスクドライブや半導体メモリモジュール等により構成された記憶装置であり、AR装置1が追跡対象の位置を認識した際に液晶ディスプレイ等の表示部に重畳表示するCGや二次元画像等のAR情報を保持している。

## 【0030】

重畳表示部24は、前記姿勢推定装置22が推定したカメラの姿勢パラメータに対応するAR情報を前記AR情報DB23から取得し、フレーム単位で連続的に提供されるプレビュー画像に重畳して表示部へ出力する。

20

## 【0031】

前記姿勢推定装置22において、画像特徴DB32には、オフラインで追跡対象の3Dモデルを様々な代表的な姿勢(視点)で投影した多数の画像から視点に関わらず顕著性、識別性の高い二次元特徴点を抽出し、更に各二次元特徴点を三次元逆投影によって三次元座標に変換する処理を通じて作成/記録された多数の画像特徴A(特徴点の画像特徴量及びその三次元座標)が蓄積される。

## 【0032】

なお、特徴点には追跡対象の姿勢によって視認可能(visible)な特徴点と、裏側に回るなどして視認不可能(invisible)な特徴点とが存在し、視認不可能な特徴点は追跡に利用できない。したがって、追跡に利用する特徴点は追跡対象の姿勢に応じて使い分けことが望ましい。

30

## 【0033】

そこで、本実施形態では各代表的な姿勢で視認可能な画像特徴の集合を当該各姿勢に固有の特徴点セットとして予め多数用意し、追跡対象の姿勢に応じて特徴点セットを使い分けようとしている。

## 【0034】

姿勢検出部31において、画像特徴検出部31aは、プレビュー画像から追跡対象の画像特徴Bを検出する。画像特徴の検出には、Harris、Hessian、SIFT、SURF、FAST、BRIEF、ORB、BRISK、FREAK等の周知のアルゴリズムを用いることができる。

40

## 【0035】

マッチング部31bは、検出された画像特徴Bと前記画像特徴DB32に蓄積されている画像特徴Aとのマッチングを行う。姿勢推定部31cは、マッチング結果に基づいて追跡対象の初期姿勢を推定する。

## 【0036】

マッチング結果である2D-3D対応点(画像特徴Bの二次元座標と画像特徴Aの三次元座標との間の対応点)から姿勢を推定するアルゴリズムとしては、DLT(Direct Linear Transform)法やPNP(Perspective-N-Points)法等が一般に知られており、本実施形態ではPnP法によって推定された姿勢を追跡対象の初期姿勢とする例を想定する。

50

## 【 0 0 3 7 】

姿勢追跡装置 3 3 は、図 2 に示したように、GPU 4 1 およびフレームバッファ 4 3 を含むグラフィックボード 4 0 と、CPU 5 1 およびメインメモリ 5 9 を含む主装置 5 0 とを主要な構成とし、グラフィックボード 4 0 と主装置 5 0 とは、バス 6 0 により命令及びその応答を送受する。

## 【 0 0 3 8 】

CPU 5 1 のテンプレート画像生成部 5 2 は、GPU 4 1 に対して、OpenGLのインタフェース経由で、カメラに対する追跡対象の姿勢を表す姿勢行列、カメラの画角や焦点距離を表す投影行列、および描画の解像度を定めるビューポートの設定と、描画命令、描画結果の転送命令とを行う。

## 【 0 0 3 9 】

GPU 4 1 のレンダリング処理部 4 2 は、前記描画命令に応答して、追跡対象の3Dモデルに3Dレンダリングを実行し、フレームバッファ 4 3 にレンダリング結果（一般には、RGBAの4チャンネルデータ）を展開する。

## 【 0 0 4 0 】

追跡対象が複数の場合、それぞれの3Dモデルおよびその姿勢行列を用いて3Dレンダリングが行われる。この際、GPU 4 1 のデプステスト機能を有効にして各画像の上下関係を識別することで、複数の追跡対象の重なりによる遮蔽（オクルージョン）を考慮したレンダリング結果をフレームバッファ 4 3 上へ展開できる。

## 【 0 0 4 1 】

CPU 5 1 のテンプレート画像生成部 5 2 は、フレームバッファ 4 3 のレンダリング結果を、バス 6 0 経由でメインメモリ 5 9 へ読み出すことでテンプレート画像を取得する。読み出した画像に対しては、入力画像に合わせて適宜リサイズやグレースケール化が行われる。

## 【 0 0 4 2 】

ここで、3Dレンダリングで利用する姿勢行列には初期姿勢を設定する。初期姿勢としては、追跡中の対象物であれば、過去の複数フレーム分の追跡結果を線形予測やカルマンフィルタに適用して反映させたものを採用できる。また、過去の追跡結果が存在しない、すなわち追跡を開始して1フレーム目であれば、姿勢検出部 3 1 の推定した姿勢を初期値としても良いし、あるいは初期値となる姿勢行列を外部から取得しても良い。

## 【 0 0 4 3 】

投影行列には、実際に入力画像を撮像するカメラモジュールのパラメータ（内部パラメータ）を設定することが望ましい。一般に内部パラメータを取得する方法として、チェスボードを利用したキャリブレーション手法が知られているので、これを利用すれば良い。

## 【 0 0 4 4 】

内部パラメータが未知の場合は、別のカメラの内部パラメータで代用しても良い。その場合、焦点距離の誤差（真値とのズレ）は、推定する姿勢のスケール（距離）の誤差に直結する。そのため、対象物とカメラの間の距離計測等に利用する場合はキャリブレーションが必須である。

## 【 0 0 4 5 】

ただし、ARの用途に限れば、AR情報を重畳する際の投影行列と揃えることで相殺され、AR情報と現実空間の位置合わせの誤差には影響しない（レンダリング結果には影響しない）ため、キャリブレーションの必要は無い。また、zNear、zFarの値は追跡対象が撮影され得る範囲の距離（例えば1cm～500cm等）に設定する。

## 【 0 0 4 6 】

ビューポートは、入力画像と同じサイズを設定する。入力画像は、カメラから撮像された画像に対して、扱いやすいサイズ（例えばVGAやQVGA）へのリサイズと、グレースケール変換を施した画像を想定する。

## 【 0 0 4 7 】

図 2 へ戻り、CPU 5 1 のROI計算部 5 4 は、レンダリング結果の高速転送を実現するため

10

20

30

40

50

の処理として、追跡対象が実際にレンダリングされた追跡対象ROI (Region Of Interest) を計算し、更には複数の追跡対象ROIをフレームバッファ43から一のレンダリング結果として取得するためのテンプレート画像取得ROIを計算する。

【0048】

前記追跡対象ROIは、テンプレート画像生成部52がフレームバッファ43からレンダリング結果を取得する際に、取得する画像サイズを限定することで高速転送を実現するためのものである。

【0049】

前記テンプレート画像取得ROIは、テンプレート画像生成部52がフレームバッファ43からレンダリング結果を取得する際に、複数の追跡対象ROIのレンダリング結果を一つにまとめて一括転送することで高速転送を実現するためのものである。

10

【0050】

テンプレート画像生成部52によるフレームバッファ43からの画像転送に関して、一般に、グラフィックボード40上のレンダリング画像を主装置部50がバス60経由で読み出す処理時間は長く、全体のボトルネックになる場合が多い。また、画像データの総量が同じ場合、複数の追跡対象のレンダリング画像を別々に生成して順次に読み出すよりも、全ての追跡対象のレンダリング画像を含むように1枚にまとめて読み出す方が、主装置部50への転送処理を短時間で行うことができる。

【0051】

そこで、前記ROI計算部54は、追跡対象が複数であれば、図3に示したように、各追跡対象のレンダリング画像を包含する各追跡対象ROI (Z1, Z2) の (論理和) 領域Zorを計算してテンプレート画像取得ROIとし、テンプレート画像生成部52は、当該テンプレート画像取得ROI内の画像をフレームバッファ43から一度の転送により取得する。追跡対象が一つの場合は、テンプレート画像取得ROIは追跡対象ROIと等しくなる。

20

【0052】

また、グラフィックボード40上のレンダリング画像を主装置部50が読み出す際、ビューポート全体を読み出すより、追跡対象が実際にレンダリングされた領域のみを読み出せば高速転送が可能になる。これはカメラと追跡対象との距離が離れ、追跡対象が占める領域のサイズが小さくなる程効果が大きい。

【0053】

そこで、前記ROI計算部54は、ビューポートからテンプレート画像取得ROIのみを読み出すこととし、更にテンプレート画像取得ROIを構成する個々の追跡対象ROIを計算する方法として、3Dモデルの頂点を利用する第1の計算方法、および特徴点セットを利用する第2の計算方法、の2種類を使い分ける。

30

【0054】

第1の計算方法では、3Dモデルの各頂点 $[X, Y, Z]^T$ を次式(1)で二次元に投影して、投影後の各頂点のピクセル座標 $P=[u, v]^T$ を計算する。全ピクセル座標を包含する矩形領域が追跡対象ROIとなる。この方法は3Dモデルの頂点数が少ない場合に有効である。

【0055】

【数1】

40

$$\begin{bmatrix} u \\ v \\ 1 \end{bmatrix} \doteq A \begin{bmatrix} R & \mathbf{t} \\ \mathbf{0} & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} = \begin{bmatrix} f_x & 0 & c_x & 0 \\ 0 & f_y & c_y & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \dots(1)$$

【0056】

ここで、Aは内部パラメータである。R( $r_{11} \sim r_{33}$ )は、三次元空間内の回転を表すパラメータであり、各パラメータはオイラー角等の表現によって三パラメータで表現することが可能である。t( $t_1 \sim t_3$ )は、三次元空間内の平行移動を表すパラメータである。X、Y、

50

Zは、3Dモデルのモデル座標系におけるX座標、Y座標及びZ座標である。u、vは、画面中のピクセル座標系におけるu座標及びv座標である。

【0057】

第2の計算方法では、前記3Dモデルの頂点の代わりに特徴点セットを投影し、同様にこれらを含む矩形領域を追跡対象ROIとする。複雑な3Dモデルの場合、特徴点の数の方が頂点より大幅に数が少ないため、第1の計算方法よりも高速な処理が期待できる。ただし特徴点に配置的な偏りがある場合、追跡対象ROIが対象物領域を正確に包含していない場合が起こり得る。そのため、投影点の包含領域より一定幅広い領域を追跡対象ROIに設定することが望ましい。

【0058】

本実施形態では、3Dモデルの頂点数に応じて、第1および第2の計算方法を使い分ける。より具体的には、3Dモデルの頂点数が閾値より大きい場合は第2の計算方式を利用し、それ以外の場合は第1の計算方式を利用する。このような使い分けにより、精度と速度の両立を実現可能である。

【0059】

図2へ戻り、主装置部50のCPU51は、テンプレート画像生成部52、遮蔽発生判定部53、ROI計算部54、遮蔽特徴点判定部55、特徴点セット修正部56、マッチング部57および姿勢更新部58の各機能により姿勢追跡を実行する。メインメモリ59は、CPU51へワークエリアを提供し、前記グラフィックボード40のフレームバッファ43からバス60経由で読み取ったレンダリング結果をテンプレート画像として展開する。

【0060】

前記遮蔽発生判定部53および遮蔽特徴点判定部55は、前記レンダリング処理部42と同様に、各追跡対象の初期姿勢、3Dモデルおよび初期姿勢に対応した特徴点セットの入力を受けて、テンプレート画像内で複数の追跡対象が重なり合う遮蔽（オクルージョン）領域を特定し、更にはオクルージョン領域における各対象物の前後関係を判定する。

【0061】

オクルージョン領域では、他の追跡対象により遮蔽された特徴点セットは、実際には見えなくなるために追跡することができず、マッチング精度劣化の要因となる。そのため、遮蔽特徴点判定部55の役割は、各追跡対象の特徴点セットについて、遮蔽された特徴点を除外することで頑健性を向上することである。

【0062】

前記遮蔽発生判定部53は、図4に示したように、各追跡対象ROIの（論理積）領域Zandの有無を確認することで、オクルージョンが発生している領域の有無を判定する。領域Zandが存在しない場合は、オクルージョンが発生していないと判定し、以降の処理を終了する。また、第1の計算方法で追跡対象ROIを計算する場合は、各頂点の投影点Pの凸包となる輪郭Cを計算し、この輪郭の重なる領域Candの有無でオクルージョンを判定しても良い。

【0063】

遮蔽特徴点判定部55は、前記領域Zandが存在すると判定されると、領域Zandを有する各追跡対象のそれぞれの特徴点セットについて、当該領域Zandの内部の特徴点のみを対象に上式(1)で二次元投影し、領域Zandの内外判定を行う。

【0064】

本実施形態では、遮蔽判定精度および処理負荷の異なる3種類の判定方法として、デブスバッファを使用して各特徴点の遮蔽判定を行う第1判定方法と、特徴点セットの幾何変換を利用して特徴点セット全体の前後関係を一括判定する第2判定方法と、3Dモデルの頂点の幾何変換を利用して特徴点セット全体の前後関係を一括判定する第3判定方法とを使い分けることができる。

【0065】

オクルージョンの判定精度と処理負荷とはトレードオフの関係があり、第1判定方法では厳密な判定が可能だが処理負荷が大きく、第2判定方法では高速な判定可能だが厳密さ

10

20

30

40

50

に欠ける。また、第3判定方法は3Dモデルの頂点数が少ない場合に有効である。

【0066】

第1判定方法では、テンプレート画像生成部52の処理と同様に、フレームバッファ上の追跡対象ROIの領域Zandから、デプスバッファに記憶されている値（一般には、32bit浮動小数点データ）をメインメモリ59に読み出す。デプスバッファ値は、次式(2)で三次元座標のZ座標に変換することができる。

【0067】

【数2】

$$Z = -\frac{zFar \times zNear}{d(zFar - zNear) - zFar} \quad \dots(2)$$

10

【0068】

ここで、zNear, zFarは、テンプレート画像生成部52の投影行列で指定した奥行の手前と奥、dはデプスバッファ値、Zはdから変換される、カメラ基準のZ座標値である。本実施形態では、基本的に座標系の取り方には依存しないが、Z軸は、奥行Zが大きくなるほど、正の大きい値を取る方向に向いていると仮定する。

【0069】

領域Zandの内部の特徴点について、(2)式でZ座標を計算する。さらに、初期姿勢と(3)式により、当該特徴点をカメラ基準の座標系に幾何変換し、幾何変換によるZ'座標とデプスバッファよるZ座標との比較を行う。

20

【0070】

【数3】

$$\begin{bmatrix} X' \\ Y' \\ Z' \\ 1 \end{bmatrix} = \begin{bmatrix} r_{11} & r_{12} & r_{13} & t_1 \\ r_{21} & r_{22} & r_{23} & t_2 \\ r_{31} & r_{32} & r_{33} & t_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \quad \dots(3)$$

【0071】

遮蔽されていない場合、これらは一致するか、近い数字になるが、遮蔽されている場合、デプスバッファは遮蔽物体の値になるので一致しない。そこで、Z'とZの比較により、当該特徴点が遮蔽されているか否かを判定し、遮蔽されている特徴点は特徴点セットから除外する。より具体的には、差分Z'-Zが所定の閾値以上になる場合に、当該特徴点が遮蔽されていると判定する。

30

【0072】

第2判定方法では、領域内部のそれぞれの追跡対象の特徴点セットについて、(3)式によりそれぞれの初期姿勢で幾何変換を行ってZ'座標を取得する。それぞれの追跡対象のZ'座標を平均し、平均の小さい方の対象物（3体以上の場合は最も小さい対象物）の特徴点セットを残し、それ以外の特徴点セットは遮蔽されていると判定して除外する。複雑に前後が入れ替わるような形状の対象物で無い限りは、この方法でも概ね正しい判定が見込める。

40

【0073】

第3判定方法は、ROI計算部54が第1判定方法で追跡対象ROIを計算した場合にのみ利用できる。3Dモデルの各頂点の投影位置の凸包となる輪郭の領域の輪郭Candを計算し、Cand内の特徴点セットについて、第2判定方法と同様の方法で前後関係を判定する。領域を矩形で計算するより、輪郭で計算する第3判定方法の方が、より高精度な判定が可能である。

【0074】

本実施形態では、端末の処理能力及び3Dモデルの頂点数に応じて第1、第2および第3判定方法を使い分けることとし、処理能力が高い場合は第1判定方法を利用し、低い場合

50

は第2判定方法を利用する。または、3Dモデルの頂点数が少ない場合に第3判定方法を利用し、多い場合に第1、第2判定方法を利用する。これにより、端末の処理能力に応じて、処理負荷と精度のバランスを取ることが可能となる。

【0075】

なお、追跡対象が三体以上であると、個々の追跡対象ROIの領域が複数領域存在する場合は起こりうる。このような場合、第1判定方法では、デブスバッファの読み出しが複数回となるので効率が悪い。したがって、複数の領域を包含する矩形領域を計算し、複数領域のデブスバッファを一括で読み出すことが望ましい。

【0076】

特徴点セット修正部56は、前記遮蔽判定の結果に基づいて、遮蔽された特徴点を特徴点セットから除外する。また、除外された特徴点セットと同数の未選択の特徴点セットを新たに選択して追跡用の特徴点セットに追加する。

10

【0077】

マッチング部57は、前記遮蔽発生判定によって修正された特徴点セットを利用して、追跡対象ごとに、そのテンプレート画像と入力画像との間で特徴点マッチングを行い、2D-3D対応点を取得する。

【0078】

2D-3D対応点とは、入力画像中の二次元ピクセル座標と、3Dモデルの三次元座標との間の点对応のことであり、三次元座標を初期姿勢でテンプレート画像内の二次元ピクセル座標に投影し、入力画像内で、投影された特徴点の対応箇所を探索することで取得する。

20

【0079】

本実施形態では、特許文献3や非特許文献1に開示されているマッチング処理を適用できる。この場合、各特徴点の周囲(本実施形態では、 $8 \times 8$ の矩形領域)の画素情報(輝度値)をテンプレート画像から取得し、これを切り出し画像とする。次いで、プレビュー画像中で切り出し画像と一致する領域を探索する。本実施形態ではNCC(Normalized Cross Correlation)やZNCC(Zero-mean Normalized Cross Correlation)を用いたテンプレートマッチングによる探索を想定する。

【0080】

マッチングによって高い相関を持つ領域が求めれば、当該領域の中心のピクセル座標が追跡位置となる。マッチングは各切り出し画像の追跡の初期位置を中心とした固定範囲(例えば、 $5 \times 5$ )に限定して行われる。

30

【0081】

なお、画像ピラミッドを利用する場合、ピラミッドの高段で比較的広い探索範囲(例えば $5 \times 5$ )で荒く推定した姿勢をさらに初期値として、低段でより狭い探索範囲(例えば $2 \times 3$ )で高精度なマッチング処理を行うことで、追跡精度の低下を抑えることが可能になる。

【0082】

なお、高い相関を持つ領域が探索範囲内で見つからなかった場合、マッチング部57は当該特徴点のマッチングに失敗したと判定する。例えば、追跡対象、又はカメラが急速に移動するなどして、特徴点の移動量が探索幅を超えた場合はマッチングに失敗する。そして、マッチングに成功した(相関が閾値以上の箇所が求められた)特徴点数が所定の閾値以下となれば、マッチングを中断して追跡失敗のメッセージを出力する。

40

【0083】

姿勢更新部58は、前記マッチング部57が取得した2D-3D対応点に基づき、以下に詳述するように、入力画像中の各追跡対象の姿勢を更新する。

【0084】

2D-3D対応点間では上式(1)が成り立つ。姿勢の推定値は、初期姿勢からスタートし、非線形の最適化手法を用いて次式(4)で求める。

【0085】

【数 4】

$$\hat{W} = \arg \min_W \sum_i \rho(\|m_i - m'_i\|^2) \quad \dots(4)$$

【0086】

ここで、 $m'_i$  は  $m_i$  の画像上の追跡位置を表す。 $m_i$  は特徴点セットの姿勢  $W$  による投影位置を表す。 $(\cdot)$  はロバスト推定の重み係数を表す。上式(4)は、特徴点セットの投影位置が追跡位置に最も近づく(再投影誤差が最も小さくなる)姿勢  $W$  を求めることに相当する。ここで、姿勢の推定値は、最終的な推定値として出力しても良いし、既存手法と同様、画像ピラミッドの下段の画像のマッチングにおける、初期値として扱っても良い。その場合、ピラミッドの下段方向に向かって、姿勢追跡の処理が反復的に実行される。

10

【0087】

このようにして更新された初期姿勢は、次フレームで検出された追跡対象のテンプレート画像を生成する際の初期姿勢に反映される。

【0088】

図5は、本発明の一実施形態の動作を示したフローチャートであり、主に姿勢追跡装置33による姿勢追跡の動作を示している。

【0089】

ステップS1では、GPU41のレンダリング処理部42が、追跡対象の3Dモデルおよび初期姿勢を取得し、さらに当該初期姿勢に対応した特徴点セットを前記画像特徴DB22から取得し、その一部を今回の追跡用に決定する。

20

【0090】

ステップS2では、ROI計算部54により、レンダリング結果のバス経由での高速転送を実現するために、前記追跡対象ROIおよびテンプレート画像取得ROI(x, y, width, height)が計算される。

【0091】

ステップS3では、レンダリング処理部42が、追跡対象の3Dモデルに対して前記初期姿勢および追跡用の特徴点に基づいて3Dレンダリングを実行し、そのレンダリング結果をフレームバッファ43上に展開する。

【0092】

ステップS4では、テンプレート画像生成部52が前記テンプレート画像取得ROIの範囲内のレンダリング結果をフレームバッファ43からバス経由で取得し、これをメインメモリ59上にテンプレート画像として展開する。ステップS5では、追跡対象の数に応じて処理が分岐する。

30

【0093】

追跡対象が複数であればステップS6へ進み、前記遮蔽発生判定部53によりオクルージョン領域が計算される。ステップS7では、前記計算結果であるオクルージョン領域の有無により処理が分岐し、オクルージョンが発生していればステップS8へ進む。

【0094】

ステップS8では、前記遮蔽特徴点判定部55により、前記オクルージョン領域における各追跡対象ROIの前後(上下)関係が判定される。ステップS9では、前(上)側の追跡対象ROIにより遮蔽されて視認不可能となった後(下)側の追跡対象ROIの特徴点セットが、前記特徴点セット修正部56により識別されて追跡用の特徴点セットから除外される。

40

【0095】

このとき、必要に応じて、除外された特徴点セットと同数の特徴点セットを、今回の初期姿勢の対応付けられた未選択の特徴点セット集合から選択し、新たに追跡用の特徴点セットとして追加するようにしても良い。

【0096】

ステップS10では、前記マッチング部57が、前記除外されなかった特徴点セットお

50

よび追加された特徴点セットを用いて、各テンプレート画像と入力画像との間でマッチングを行う。ステップS 1 1では、マッチングが取れた特徴点セットが最も多く帰属する姿勢が現在の姿勢として更新登録される。このようにして更新された追跡姿勢は、次フレームで検出された対象物のテンプレート画像を生成する際の初期姿勢として用いられる。

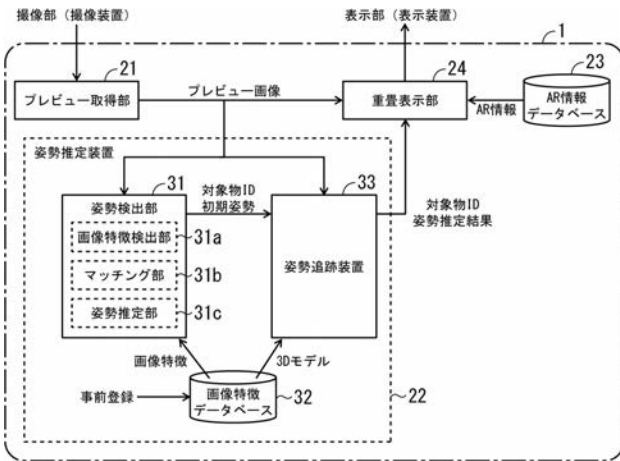
【符号の説明】

【0097】

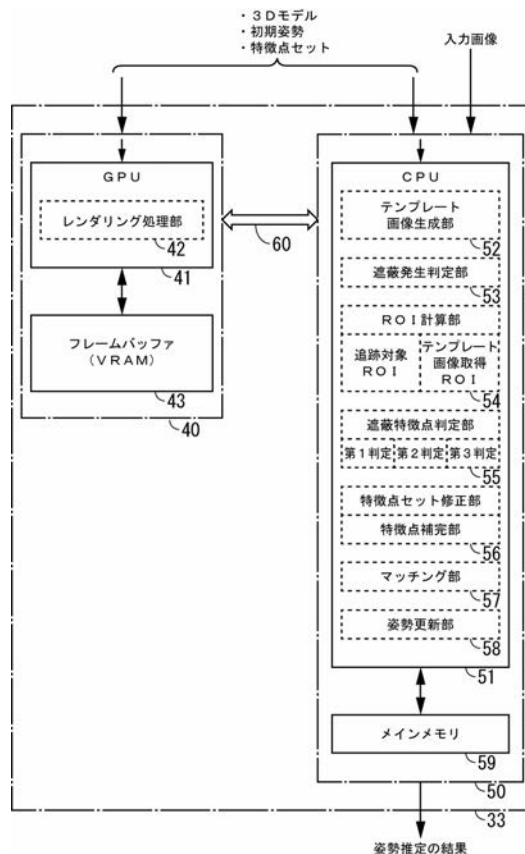
2 1 ... プレビュー取得部, 2 2 ... 姿勢推定装置, 2 3 ... AR情報DB, 2 4 ... 重畳表示部, 3 1 ... 姿勢検出部, 3 1 a ... 画像特徴検出部, 3 1 b ... マッチング部, 3 1 c ... 姿勢推定部, 3 2 ... 画像特徴DB, 3 3 ... 姿勢追跡装置, 4 0 ... グラフィックボード, 4 1 ... GPU, 4 2 ... レンダリング処理部, 4 3 ... 一括転送準備部, 4 3 ... フレームバッファ, 5 0 ... 主装置部, 5 1 ... CPU, 5 2 ... テンプレート画像生成部, 5 3 ... 遮蔽発生判定部, 5 4 ... ROI計算部, 5 5 ... 遮蔽特徴点判定部, 5 6 ... 特徴点セット修正部, 5 7 ... マッチング部, 5 8 ... 姿勢更新部, 5 9 ... メインメモリ, 6 0 ... バス

10

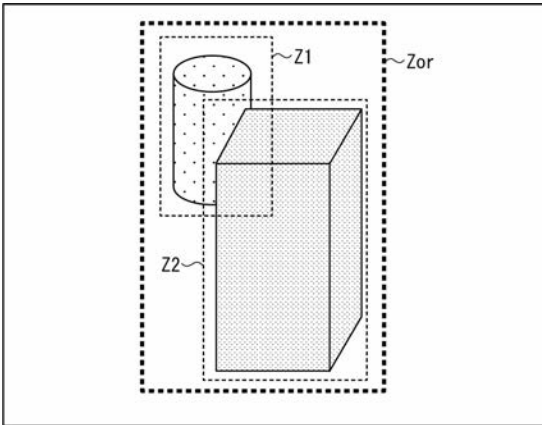
【図 1】



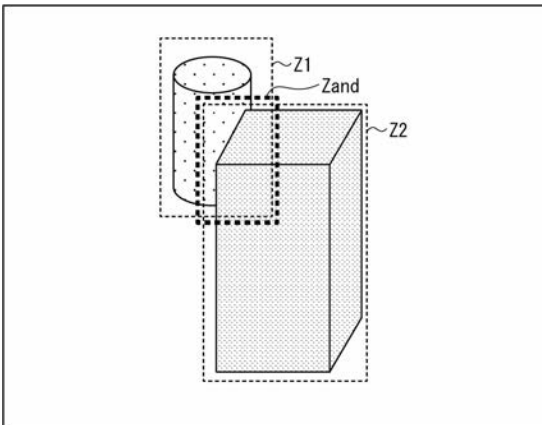
【図 2】



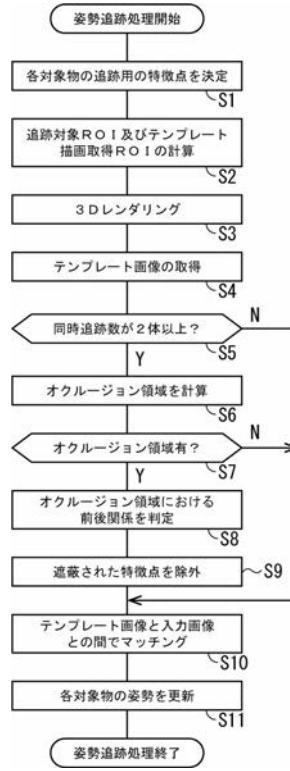
【 図 3 】



【 図 4 】



【 図 5 】



フロントページの続き

Fターム(参考) 5B050 AA10 BA08 BA09 BA11 BA18 EA06 EA07 EA19 EA24 EA29  
FA02  
5L096 AA06 AA09 FA09 HA05 JA09