

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5981990号

(P5981990)

(45) 発行日 平成28年8月31日(2016. 8. 31)

(24) 登録日 平成28年8月5日(2016. 8. 5)

(51) Int.Cl.

F I

G O 6 T 13/40 (2011. 01)

G O 6 T 13/40

G O 6 T 15/00 (2011. 01)

G O 6 T 15/00 5 0 1

請求項の数 5 (全 47 頁)

(21) 出願番号	特願2014-517093 (P2014-517093)	(73) 特許権者	314015767
(86) (22) 出願日	平成24年6月19日 (2012. 6. 19)		マイクロソフト テクノロジー ライセン
(65) 公表番号	特表2014-522018 (P2014-522018A)		シング, エルエルシー
(43) 公表日	平成26年8月28日 (2014. 8. 28)		アメリカ合衆国 ワシントン州 9805
(86) 国際出願番号	PCT/US2012/043154		2 レッドモンド ワン マイクロソフト
(87) 国際公開番号	W02012/177643		ウェイ
(87) 国際公開日	平成24年12月27日 (2012. 12. 27)	(74) 代理人	100140109
審査請求日	平成27年6月3日 (2015. 6. 3)		弁理士 小野 新次郎
(31) 優先権主張番号	13/172, 255	(74) 代理人	100075270
(32) 優先日	平成23年6月29日 (2011. 6. 29)		弁理士 小林 泰
(33) 優先権主張国	米国 (US)	(74) 代理人	100101373
			弁理士 竹内 茂雄
		(74) 代理人	100118902
			弁理士 山本 修

最終頁に続く

(54) 【発明の名称】 全自動動的有関節モデル校正

(57) 【特許請求の範囲】

【請求項 1】

物体の有関節(articulated)身体部分の姿勢をモデル化するための方法を実行するように少なくとも1つのプロセッサをプログラムするためのコンピューター可読ソフトウェアがその上に具体化される、コンピューター可読記憶デバイスであって、前記方法が、

1つまたは複数のフレーム内において前記有関節身体部分の深度(depth)ピクセルを獲得するステップと、

前記物体の前記有関節身体部分を識別するために、前記1つまたは複数のフレームの前記深度ピクセルを処理するステップと、

モデルにアクセスするステップであって、前記モデルが、前記物体の前記有関節身体部分に対応する有関節身体部分を含む、ステップと、

初期マッチ(initial match)を提供するために、前記モデルの前記有関節身体部分を前記物体の前記識別された有関節身体部分にマッチさせるステップと、

前記初期マッチにおいて距離制約(distance constraints)が前記モデルの前記有関節身体部分によって違反された(violated)程度(extent)を評価する(evaluating)ステップであって、前記距離制約が、前記モデルの前記有関節身体部分の間の距離に課される制約を含む、ステップと、

前記距離制約が違反された程度に基づいて、前記モデルの前記有関節身体部分の少なくとも1つについての長さを調整するステップを含む、前記距離制約を満たそうとする試みにおいて前記モデルを調整するステップと

10

20

を含む、コンピューター可読記憶デバイスであって、
実行される前記方法が、

前記モデルの深度ピクセルを提供するために、前記モデルをラスタ化するステップと

、
少なくとも1つの比較平面内(comparison plane)において前記モデルの前記深度ピクセルとオーバーラップしていない、前記1つまたは複数のフレームの非オーバーラップ深度ピクセルを、前記1つまたは複数のフレームの前記深度ピクセルの中から識別するために、前記モデルの前記深度ピクセルを前記1つまたは複数のフレームの前記深度ピクセルと比較するステップと、

前記モデルの前記有関節身体部分の前記少なくとも1つについての寸法を調整することによって、前記モデルを前記1つまたは複数のフレームの前記非オーバーラップ深度ピクセルにより密接にマッチさせるために、前記比較ステップに基づいて前記モデルを調整するステップとをさらに含み、

前記比較ステップに基づいて前記モデルを調整する前記ステップが、前記モデルの前記有関節身体部分の前記少なくとも1つの横側に(lateral)隣接する、前記1つまたは複数のフレームの非オーバーラップ深度ピクセルの数に従って、前記モデルの前記有関節身体部分の前記少なくとも1つについての幅を増やすステップを含む、

コンピューター可読記憶デバイス。

【請求項2】

物体の有関節(articulated)身体部分の姿勢をモデル化するための方法を実行するように少なくとも1つのプロセッサをプログラムするためのコンピューター可読ソフトウェアがその上に具体化される、コンピューター可読記憶デバイスであって、前記方法が、

1つまたは複数のフレーム内において前記有関節身体部分の深度(depth)ピクセルを獲得するステップと、

前記物体の前記有関節身体部分を識別するために、前記1つまたは複数のフレームの前記深度ピクセルを処理するステップと、

モデルにアクセスするステップであって、前記モデルが、前記物体の前記有関節身体部分に対応する有関節身体部分を含む、ステップと、

初期マッチ(initial match)を提供するために、前記モデルの前記有関節身体部分を前記物体の前記識別された有関節身体部分にマッチさせるステップと、

前記初期マッチにおいて距離制約(distance constraints)が前記モデルの前記有関節身体部分によって違反された(violated)程度(extent)を評価する(evaluating)ステップであって、前記距離制約が、前記モデルの前記有関節身体部分の間の距離に課される制約を含む、ステップと、

前記距離制約が違反された程度に基づいて、前記モデルの前記有関節身体部分の少なくとも1つについての長さを調整するステップを含む、前記距離制約を満たそうとする試みにおいて前記モデルを調整するステップと

を含む、コンピューター可読記憶デバイスであって、

前記有関節身体部分が、手の手のひらおよび指を含み、実行される前記方法が、

前記モデルにおいて前記モデルの少なくとも2つの隣接する指によって衝突制約(collision constraints)が違反される程度を評価するステップと、

前記モデルの前記少なくとも2つの隣接する指の少なくとも1つによって前記衝突制約が違反される前記程度に基づいて、前記モデルの前記少なくとも2つの隣接する指の前記少なくとも1つについての寸法を調整するステップを含む、前記衝突制約を満たそうとする試みにおいて前記モデルを調整するステップとをさらに含み、

前記衝突制約を満たそうとする前記試みにおいて前記モデルを調整する前記ステップが、前記モデルの前記少なくとも2つの隣接する指の前記少なくとも1つによって前記衝突制約が違反される前記程度に基づいて、前記モデルの前記少なくとも2つの隣接する指の前記少なくとも1つについての幅を減らすステップを含む、

コンピューター可読記憶デバイス。

10

20

30

40

50

【請求項 3】

物体の有関節(articulated)身体部分の姿勢をモデル化するための方法を実行するように少なくとも1つのプロセッサをプログラムするためのコンピューター可読ソフトウェアがその上に具体化される、コンピューター可読記憶デバイスであって、前記方法が、

1つまたは複数のフレーム内において前記有関節身体部分の深度(depth)ピクセルを獲得するステップと、

前記物体の前記有関節身体部分を識別するために、前記1つまたは複数のフレームの前記深度ピクセルを処理するステップと、

モデルにアクセスするステップであって、前記モデルが、前記物体の前記有関節身体部分に対応する有関節身体部分を含む、ステップと、

初期マッチ(initial match)を提供するために、前記モデルの前記有関節身体部分を前記物体の前記識別された有関節身体部分にマッチさせるステップと、

前記初期マッチにおいて距離制約(distance constraints)が前記モデルの前記有関節身体部分によって違反された(violated)程度(extent)を評価する(evaluating)ステップであって、前記距離制約が、前記モデルの前記有関節身体部分の間の距離に課される制約を含む、ステップと、

前記距離制約が違反された程度に基づいて、前記モデルの前記有関節身体部分の少なくとも1つについての長さを調整するステップを含む、前記距離制約を満たそうとする試みにおいて前記モデルを調整するステップと

を含む、コンピューター可読記憶デバイスであって、

実行される前記方法が、

前記モデルの深度ピクセルを提供するために、前記モデルをラスタ化(rasterizing)するステップと、

少なくとも1つの比較平面内において前記1つまたは複数のフレームの前記深度ピクセルとオーバーラップしていない、前記モデルの非オーバーラップ深度ピクセルを、前記1つまたは複数のフレームの前記深度ピクセルの中から識別するために、前記モデルの前記深度ピクセルを前記1つまたは複数のフレームの前記深度ピクセルと比較するステップと、

前記モデルの前記有関節身体部分の前記少なくとも1つについての寸法を調整することによって、前記モデルを前記1つまたは複数のフレームの前記非オーバーラップ深度ピクセルにより密接にマッチさせるために、前記比較ステップに基づいて前記モデルを調整するステップとをさらに含み、

前記比較ステップに基づいて前記モデルを調整する前記ステップが、前記モデルの前記有関節身体部分の少なくとも1つの前記少なくとも1つの横側に隣接する、前記モデルの非オーバーラップ深度ピクセルの数に従って、前記モデルの前記有関節身体部分の前記少なくとも1つについての幅を減らすステップを含む、

コンピューター可読記憶デバイス。

【請求項 4】

ユーザーの手の姿勢をモデル化するためのプロセッサ実施方法であって、

1つまたは複数のフレーム内において前記手の深度ピクセルを獲得するプロセッサ実施ステップと、

前記手の有関節部分を識別するために、前記1つまたは複数のフレームの前記深度ピクセルを処理するプロセッサ実施ステップと、

前記手の前記有関節部分のモデルにアクセスするプロセッサ実施ステップであって、前記モデルの前記手の前記有関節部分が、指セグメントを含む、手のひらおよび指を含む、プロセッサ実施ステップと、

初期マッチを提供するために、前記モデルの前記手の前記有関節部分を前記1つまたは複数のフレームの前記深度ピクセルからなる前記手の前記識別された有関節部分にマッチさせるプロセッサ実施ステップと、

前記初期マッチにおいて距離制約が前記指の少なくとも1つによって違反された程度を

10

20

30

40

50

評価するプロセッサ実施ステップであって、前記距離制約が、前記指の前記少なくとも1つについての指セグメントの間の距離に課される制約を含む、プロセッサ実施ステップと、

前記モデルの深度ピクセルを提供するために、前記モデルをラスタ化し、前記プロセッサ実施ステップと、

少なくとも1つの比較平面内において前記モデルの前記指の前記少なくとも1つについての前記深度ピクセルとオーバーラップしていない、前記1つまたは複数のフレームの非オーバーラップ深度ピクセルを、前記1つまたは複数のフレームの前記深度ピクセルの中から識別するために、前記指の前記少なくとも1つについての前記深度ピクセルを前記1つまたは複数のフレームの前記深度ピクセルと比較するプロセッサ実施ステップと、

10

(a) 前記距離制約が前記指の前記少なくとも1つによって違反された程度に基づいて、前記モデルの前記指の少なくとも1つについての少なくとも1つの指セグメントの長さを調整することを含む、前記距離制約を満たそうとする試みにおいて、および(b) 前記モデルの前記指の前記少なくとも1つについての前記指セグメントの前記少なくとも1つについての幅を増やすことによって、前記モデルを前記1つまたは複数のフレームの前記非オーバーラップ深度ピクセルにより密接にマッチさせるために、前記比較ステップに基づいて、前記モデルを調整するプロセッサ実施ステップと

を含むプロセッサ実施方法。

【請求項5】

前記少なくとも1つの比較平面内において前記1つまたは複数のフレームの前記深度ピクセルとオーバーラップしていない、前記モデルの前記手の前記有関節部分の前記少なくとも1つについての非オーバーラップ深度ピクセルを、前記1つまたは複数のフレームの前記深度ピクセルの中から識別するために、前記指の前記少なくとも1つ以外の、前記手の前記有関節部分の少なくとも1つについての前記深度ピクセルを前記1つまたは複数のフレームの前記深度ピクセルと比較するステップと、

20

前記幅を減らすことによって、前記モデルを前記1つまたは複数のフレームの前記深度ピクセルにより密接にマッチさせるために、前記比較ステップに基づいて前記モデルを調整するステップと

をさらに含む、請求項4に記載のプロセッサ実施方法。

【発明の詳細な説明】

30

【技術分野】

【0001】

本願発明の一実施例は、例えば、全自動動的有関節モデル校正に関する。

【背景技術】

【0002】

[0001] 撮像システムは、物理空間内の人間または他の対象などの有関節物体 (articulated object) の位置、姿勢、および／または動きに関するデータを獲得し、そのデータをコンピューティングシステム内のアプリケーションへの入力として使用することができる。軍事、娯楽、スポーツ、および医療目的など、多くの用途が可能である。例えば、物体の姿勢および／または運動を使用して、アニメーションのキャラクターまたはアバターを制御すること、ゲームに参加すること、および画面上のコマンドのメニューから選択を行うことなどができる。姿勢は、深度ピクセル (depth pixel) からなるフレーム内などにおける、ある時点の物体の配置または形状を表すことができ、一方、物体の運動は、深度ピクセルからなる複数のフレーム内などにおける、一連の姿勢によって表すことができる。撮像システムは、カメラを使用して視野内における物体の存在を検出する光学系を含むことができる。しかし、より高い忠実度で物体を撮像するためには、さらなる改良が必要とされる。特に、手などの有関節身体部分 (articulated body portion) を高度な忠実度で撮像することが望ましい。

40

【発明の概要】

【発明が解決しようとする課題】

50

【0003】

本願発明の一実施例は、例えば、全自動動的有関節モデル校正に関する。

【課題を解決するための手段】

【0004】

[0002]ユーザーの手など有関節身体部分を改善された忠実度で撮像するための、プロセッサ実施方法、撮像システム、および有形なコンピューター可読記憶デバイスが提供される。例えば、ユーザーは、ハンドジェスチャーを行って、メニューを操作すること、ブラウジングもしくはショッピングエクスペリエンスにおいて対話すること、プレーするゲームを選択すること、または友人にメッセージを送信するなど通信機能にアクセスすることができる。ユーザーは、手を使用して、カーソルを制御し画面上のメニューから項目を選択すること、または3D仮想世界においてアバターの動きを制御することができる。一般に、手の位置および姿勢を検出し、撮像システム内のアプリケーションへの制御入力として使用することができる。

10

【0005】

[0003]撮像システムの能力を高めて、有関節身体部分を正確に識別するために、有関節身体部分のモデルが提供される。モデルを深度センサーから獲得された対応する深度ピクセルにマッチさせて、初期マッチを提供する。その後、マッチングプロセスは、距離制約 (distance constraint) と、衝突制約 (collision constraint) と、角度制約 (angle constraint) と、ラスタ化モデル (rasterized model) を使用するピクセル比較とを使用して改善される。

20

【0006】

[0004]一実施形態では、(ユーザー/人、動物、ロボットなど) 物体の有関節身体部分の姿勢をモデル化するための方法を実行するように少なくとも1つのプロセッサをプログラムするためのコンピューター可読ソフトウェアがその上に具体化される、コンピューター可読記憶デバイスが提供される。方法は、有関節身体部分の深度ピクセルを獲得するステップを含む。深度ピクセルは、例えば1つまたは複数のフレーム内において、深度センサーを使用して獲得することができる。深度ピクセルは、センサーデータである。方法は、有関節身体部分を識別するために、1つまたは複数のフレームの深度ピクセルを処理するステップをさらに含む。方法は、モデルにアクセスするステップであって、モデルが、物体の有関節身体部分に対応する有関節身体部分を含む、ステップをさらに含む。物体の識別された有関節身体部分にモデルの有関節身体部分をマッチさせて、初期マッチを提供する。1つの可能な手法では、このマッチングプロセスは、モデルの各有関節身体部分の代表的なアトラクトポイント (attract point) を、1つまたは複数のフレームの深度ピクセルのそれぞれの重心 (centroid) にマッチさせる。

30

【0007】

[0005]方法は、初期マッチにおいて距離制約がモデルの有関節身体部分によって違反された程度を評価するステップであって、距離制約が、モデルの有関節身体部分の間の距離に課される制約を含む、ステップをさらに含む。モデルは、距離制約を満たそうとする試みにおいて調整され、それには、距離制約が違反された程度に基づいて、モデルの有関節身体部分の少なくとも1つについての長さを調整することが含まれる。一例として、有関節身体部分の長さを増やすこと、または減らすことができる。

40

【0008】

[0006]衝突制約は、モデルの少なくとも2つの隣接する指セグメントが、例えば互いに平行といった指定された相対位置にあると判定された場合など、モデルが指定された条件を満たす場合に実施することができる。モデルは、モデルの少なくとも2つの隣接する指セグメントの幅を減らすことなどによって、衝突制約を満たそうとする試みにおいて調整することができる。

【0009】

[0007]ピクセル比較は、1つまたは複数のフレームの非オーバーラップ深度ピクセル (

50

non-overlapping depth pixel) および手の有関節部分の少なくとも1つについての非オーバーラップ深度ピクセルを識別するために、手の有関節部分の少なくとも1つについての深度ピクセルを1つまたは複数のフレームの深度ピクセルと比較することを含むことができる。モデルは、モデルを1つまたは複数のフレームの深度ピクセルにより密接にマッチさせるために、モデルの有関節身体部分の少なくとも1つについての寸法を減らすこと、および／または増やすことなどによって、比較ステップに基づいて調整することができる。

【0010】

[0008] 上述の要約は、本明細書において以下でさらに説明される概念のうちのいくつかを選んで簡潔な形で紹介するために提供された。上述の要約は、特許請求される本発明の主要な特徴または必須の特徴を識別することを意図しておらず、特許請求される本発明の範囲を限定するために使用することも意図していない。

【0011】

[0009] 図面では、同様の番号が付された要素は互いに一致する。

【図面の簡単な説明】

【0012】

【図1】 [0010] 撮像システムの例示的な一実施形態を示す図である。

【図2】 [0011] 図1の撮像システムの例示的なブロック図である。

【図3】 [0012] 図1の撮像システムにおいて使用できるコンピューティング環境の例示的なブロック図である。

【図4】 [0013] 図1の撮像システムにおいて使用できるコンピューティング環境の別の例示的なブロック図である。

【図5】 [0014] 関節とアトラクトポイントとを含む、ユーザーの手のモデルを示す図である。

【図6】 [0015] 撮像システムにおいて改善された忠実度で有関節身体部分の姿勢を検出するための方法を示す図である。

【図7A】 [0016] 図6のステップ700、720、および750のさらなる詳細を示す図である。

【図7B】 [0017] 図6のステップ760のさらなる詳細を示す図である。

【図8A】 [0018] ユーザーの手の深度ピクセルを含む、深度ピクセルからなるフレームを示す図である。

【図8B】 [0019] 背景ピクセルを除去した、図8Aのユーザーの手の深度ピクセルを示す図である。

【図9A】 [0020] 図8Bのユーザーの手の深度ピクセルからなる末端部指セグメントおよび末端部親指セグメントについての、図7Aの見本ステップ712からの確率値を示す図である。

【図9B】 [0021] 図8Bのユーザーの手の深度ピクセルからなる中間部指セグメントおよび基部親指セグメントについての、図7Aのステップ721において生成される重心を示す図である。

【図9C】 [0022] 図8Bのユーザーの手の深度ピクセルからなる基部指セグメントについての、図7Aのステップ721において生成される重心を示す図である。

【図9D】 [0023] 図8Bのユーザーの手の深度ピクセルからなる手のひらについての、図7Aのステップ721において生成される重心を示す図である。

【図9E】 [0024] 図8Bのユーザーの手の深度ピクセルからなる手首についての、図7Aのステップ721において生成される重心を示す図である。

【図10】 [0025] 図7Aのステップ721において、ユーザーの手のテストデータから生成された重心を示す図である。

【図11】 [0026] 重心対アトラクトポイントマッチング (centroid-to-attract point matching) についての、図7Aのステップ734の詳細を示す図である。

10

20

30

40

50

【図 1 2 A】[0027]重心へのアトラクトポイントの剛体変換についての、図 7 A のステップ 7 3 6 の詳細を示す図である。

【図 1 2 B】[0028]アトラクトポイントを重心のより近くに移動させるための、図 7 A のステップ 7 5 2 の詳細を示す図である。

【図 1 2 C】[0029]モデルをラスタ化して、モデルの深度ピクセルを提供するための、図 7 B のステップ 7 6 4 の詳細を示す図である。

【図 1 2 D】[0030]親指エリアについて、モデルの深度ピクセルをセンサーの深度ピクセルと比較するための、図 7 B のステップ 7 6 6 ~ 7 7 2 の詳細を示す図である。

【図 1 2 E】[0031]図 1 2 D と一致する、ピクセルベースの調整ベクトル (adjustment vector) に基づいたモデルの関節の調整を示す、図 7 B のステップ 7 7 8 の詳細を示す図である。

10

【図 1 2 F】[0032]図 1 2 E において使用されるピクセルベースの調整ベクトルを示す図である。

【図 1 2 G】[0033]図 1 2 E と一致する、モデルの更新を示す、図 7 B のステップ 7 9 2 の詳細を示す図である。

【図 1 3】[0034]図 1 3 A は比較平面 (comparison plane) が深度センサーの深度軸を横断する、モデルの深度ピクセルをセンサーの深度ピクセルと比較するための、図 7 B のステップ 7 6 8、7 7 0、および 7 7 2 のさらなる詳細を示す図である。

[0035]図 1 3 B は図 1 3 A の比較平面 1 3 1 0 のさらなる詳細を示す図である。

【図 1 4】[0036]距離制約、衝突制約、およびピクセル比較を使用して、撮像システムにおいて改善された忠実度で有関節身体部分の姿勢を検出するための方法を示す図である。

20

【図 1 5】[0037]図 1 4 の方法とともに使用するための、ユーザーの手の深度ピクセル例を示す図である。

【図 1 6 A 1】[0038]距離制約が実施されず、モデルが望ましいものよりも小さい場合の、図 1 5 の深度ピクセルに対する図 5 のモデルのマッチングを示す図である。

【図 1 6 A 2】[0039]親関節および子関節が示された、図 1 6 A 1 の代替を示す図である。

【図 1 6 B 1】人差し指についての図 1 6 A 1 の画像の詳細を示す図である。

【図 1 6 B 2】[0040]親関節および子関節が示された、図 1 6 B 1 の代替を示す図である。

30

【図 1 6 C】[0041]距離制約を実施した後の、図 1 6 B 1 のモデルに対する調整を示す図である。

【図 1 6 D】[0042]ラスタ化および深度センサーの深度ピクセルとの比較の後の、図 1 6 C のモデルを示す図である。

【図 1 6 E】[0043]ピクセル比較に基づいた、図 1 6 D のモデルに対する調整を示す図である。

【図 1 7 A】[0044]距離制約が実施されず、モデルが望ましいものよりも大きい場合の、図 1 5 の深度ピクセルに対する図 5 のモデルのマッチングの別の例を示す図である。

【図 1 7 B】[0045]親関節および子関節が示された、図 1 7 A の代替を示す図である。

【図 1 8】[0046]図 1 8 A は衝突を示す、図 5 のモデルの 2 つの隣接する指を示す図である。[0047]図 1 8 B は衝突制約を実施した後の、図 1 8 A の 2 つの隣接する指を示す図である。

40

【発明を実施するための形態】

【0013】

[0048]本明細書で提供される技法は、人間対コンピューターの様々な対話タスクのための入力のソースとして消費できる提案を生成するために、確率情報および制約を用いて、自由度の高い有関節物体の状態（姿勢および／または形状）の推定をリアルタイムに提供する。有関節物体は、たいていは 6 自由度を有する多くの関節を有する 3 D 形状物体として知ることができる。状態は、知られた有関節物体のすべての部分の、並進、方向、スケール、および他の形状記述情報である。

50

【0014】

[0049]全身トラッキングといったより限定的なケースのためのこの問題のソリューションは、手足および頭などの身体の特徴を識別するには十分であるが、より精緻なレベルの詳細が望まれる場合には一般に不十分である。例えば、ユーザーの手を検出する場合、ピクセル空間内でユーザーが手を素早く動かすことができるせいで、程度の低いフレーム対フレームのコヒーレンスしか得られないことがある。ユーザーが深度センサーに近い場合など、ユーザーの手が視野の大部分を満たす場合は、特にそうである。さらに、全身トラッキングと比較して、手は一般に、自己遮蔽に起因するより大きな曖昧性を有し、例えば指など、見掛けが似た部分を多く有する。しかし、本明細書で提供されるソリューションは、全身トラッキングにも姿勢推定にも同様に適している。

10

【0015】

[0050]いくつかのケースでは、トポロジーが一定の有関節モデルが使用される。しかし、変化する寸法をモデルの各関節部分が有することが有利であり得る。このケースでは、モデルのすべての部分の寸法を動的に良好に推定し、リアルタイムに誤差を補正するために寸法を調整することが、モデルフィッティング (model fitting) などのタスクには有益である。有関節姿勢推定方法は、直接測定を通して、またはユーザーがいくつかの事前に定められた較正姿勢を取る前処理ステップにおいて計算されて、モデルサイズパラメータがあらかじめ分かっていることを前提にすることができる。しかし、そのような較正プロセスはユーザーには不便であり得るので、それを避けるのが望ましい。本明細書で提供されるソリューションは、リアルタイムに有関節モデルのすべての部分の寸法を動的に較正し、調整する。これは、有関節モデルの各部分の寸法に関する情報の複数のソースを組み合わせ、重み付けを行うことによって行われる。提案されるソリューションは、ユーザー入力、事前較正ステップ、またはユーザーが簡単な姿勢を取ることを必要としない。

20

【0016】

[0051]ある程度の連続的な期間にわたって異なるソースから獲得した情報に重み付けを行い、組み合わせることによって、既知の有関節モデルを正しい寸法に動的に調整することができる。これは、静的初期化ステップとして、またはより柔軟な有関節物体に対処するための、もしくはともに同じ既知のタイプに属する異なる有関節物体のモデル化を切り換えるための継続調整ステップとして、行うことができる。

30

【0017】

[0052]ソリューションの特徴は、(1) 距離制約違反量を使用して、有関節部分の長さを推定すること、(2) 衝突制約違反量を使用して、有関節部分幅の過大評価を見つけること、(3) モデルフィッティングからのピクセル分類を使用して、有関節部分の幅を推定すること、および(4) 剛体3D点群変換 (rigid 3D point cloud transformation) を使用して、1Dから3Dスケールを使用する有関節モデルの全体サイズを推定することを含む。

【0018】

[0053]図1は、ユーザー8がアプリケーションと対話する、撮像システム10の例示的な一実施形態を示している。これは、ユーザーの自宅内などにおける、撮像システムの現実世界配置を示している。撮像システム10は、ディスプレイ196と、深度カメラシステム20と、コンピューティング環境または装置12とを含む。深度カメラシステム20は、画像カメラコンポーネント22または深度センサーを含むことができ、画像カメラコンポーネント22は、赤外 (IR) 光放射器24、赤外光カメラ26、および赤緑青 (RGB) カメラ28を有する。ユーザー8は、人またはプレーヤーとも呼ばれ、深度センサーの視野6内に立つ。ライン2および4は、視野6の境界線を示す。この例では、深度カメラシステム20およびコンピューティング環境12が、ディスプレイ196上のアバター197がユーザー8の動きをトラッキングする、アプリケーションを提供する。例えば、アバターは、ユーザーが腕を上げたときに、腕を上げることができる。ユーザーは、手でジェスチャーを行って、制御入力をアプリケーションに提供することもできる。アバタ

40

50

ー197は、3D仮想世界内の道路198上に立っている。深度カメラシステム20の焦点距離に沿って例えば水平に伸びるz軸と、垂直に伸びるy軸と、(z軸を横断して)横方向に水平に伸びるx軸とを含む、デカルト世界座標系を定義することができる。図面のパースペクティブは、単純化として、ディスプレイ196が、y軸方向に垂直に広がり、z軸が、y軸およびx軸に直交し、ユーザー8が立つ地面に平行して、深度カメラシステムから手前に伸びるように変更されていることに留意されたい。

【0019】

[0054]一般に、撮像システム10は、人ターゲットを認識し、解析し、および／またはトラッキングするために使用される。コンピューティング環境12は、コンピューター、またはゲームシステムもしくはコンソールなどの他、アプリケーションを実行するためのハードウェアコンポーネント、および／またはソフトウェアコンポーネントも含むことができる。

10

【0020】

[0055]深度カメラシステム20は、片手または両手の姿勢、ジェスチャー、および／または動きを含む、ユーザーによって実行される姿勢、ジェスチャー、および／または動きをキャプチャーし、解析し、トラッキングして、アバターもしくは画面上のキャラクターを動かし、またはユーザーインターフェース(UI)内においてメニュー項目を選択するなどの、1つまたは複数の制御またはアクションをアプリケーション内において実行できるように、ユーザー8など1人または複数人の人を視覚的にモニタリングするために使用されるカメラを含むことができる。手の姿勢は、深度ピクセルからなるフレーム内などにおける、ある時点での、ユーザーの手の配置に相当し、ジェスチャーは、深度ピクセルからなる複数のフレーム内などにおける、一連の姿勢に相当する。複数のユーザーを同様にモニタリングすることができる。

20

【0021】

[0056]撮像システム10は、例えば、テレビ、モニター、もしくは高解像度テレビ(HDTV)などの、または視覚および聴覚出力をユーザーに提供する壁面もしくは他の面上への投影のことさえある、ディスプレイ196などの視聴覚デバイスに接続することができる。聴覚出力は、別個のデバイスを介して提供することもできる。ディスプレイを駆動するため、コンピューティング環境12は、アプリケーションに関連する視聴覚信号を提供する、グラフィックスカードなどのビデオアダプター、および／またはサウンドカードなどのオーディオアダプターを含むことができる。ディスプレイ196は、例えば、Sビデオケーブル、同軸ケーブル、HDMI(登録商標)ケーブル、DVIケーブル、またはVGAケーブルなどを介して、コンピューティング環境12に接続することができる。

30

【0022】

[0057]ユーザーの姿勢、ジェスチャー、および／または動きが、キャプチャーされ、アバターもしくは画面上のキャラクターを動かすために使用されるように、および／またはコンピューティング環境12によって実行されるアプリケーションへの入力制御として解釈されるように、深度カメラシステム20を使用して、ユーザー8をトラッキングすることができる。

【0023】

[0058]ユーザー8のいくつかの動きは、アバターの制御以外のアクションに対応することができる制御として解釈することができる。例えば、一実施形態では、プレーヤーは、姿勢および／または動きを使用して、ゲームを終了、停止、またはセーブすること、レベルを選択すること、高得点を見ること、および友人と通信することなどができる。プレーヤーは、姿勢および／または動きを使用して、メインユーザーインターフェースからゲームもしくは他のアプリケーションを選択すること、またはそれとは別に選択肢のメニューを操作することができる。したがって、1つまたは複数のアプリケーションと対話するために、ユーザー8のあらゆる姿勢および／または動きを、任意の適切な方法で利用可能とすること、使用すること、および解析することができる。

40

【0024】

50

[0059] 撮像システム 10 はさらに、ゲームならびに娯楽およびレジャーを目的とする他のアプリケーションの分野の外にある、オペレーティングシステムおよび／またはアプリケーションの制御入力として、ターゲットの姿勢および／または動きを解釈するために使用することができる。例えば、実際に、オペレーティングシステムおよび／またはアプリケーションの制御可能な側面はいずれも、ユーザー 8 の姿勢および／または動きによって制御することができる。

【0025】

[0060] 図 2 は、図 1 a の撮像システム 10 の例示的なブロック図である。深度カメラシステム 20 は、深度ピクセル値を含むことができる深度画像を含む深度情報を有するビデオを、例えば、飛行時間 (time-of-flight)、構造化照明 (structured light)、またはステレオ画像 (stereo image) などを含む、任意の適切な技法を介して、キャプチャーするように構成することができる。深度カメラシステム 20 は、深度情報を「Z レイヤ」に、すなわち、深度カメラからその視線に沿って伸びる Z 軸に直交できるレイヤに組織することができる。

10

【0026】

[0061] 深度カメラシステム 20 は、物理空間内のシーンの深度画像をキャプチャーする深度カメラなどの、画像カメラコンポーネント 22 を含むことができる。深度画像は、キャプチャーされたシーンの 2 次元 (2D) ピクセル領域を含むことができ、2D ピクセル領域内の各ピクセルは、画像カメラコンポーネント 22 からの直線距離を表す、関連する深度値を有する。

20

【0027】

[0062] 画像カメラコンポーネント 22 は、シーンの深度画像をキャプチャーするために使用できる、赤外 (IR) 光放射器 24 と、赤外光カメラ 26 と、赤緑青 (RGB) カメラ 28 とを含むことができる。赤外光放射器 24 および赤外光カメラ 26 の組み合わせによって、3D カメラが形成される。例えば、飛行時間解析では、IR 光放射器 24 が、物理空間に赤外光を放射し、赤外光カメラ 26 が、物理空間内の 1 つまたは複数のターゲットおよび物体の表面からの後方散乱光を検出する。いくつかの実施形態では、出射光パルスと対応する入射光パルスの間の時間が測定され、深度カメラシステム 20 から物理空間内のターゲットまたは物体上の特定の位置までの物理距離を決定するために使用されるように、パルス赤外光を使用することができる。出射光波の位相を入射光波の位相と比較して、位相シフトを決定することができる。その後、位相シフトを使用して、深度カメラシステムからターゲットまたは物体上の特定の位置までの物理距離を決定することができる。

30

【0028】

[0063] 例えば散乱光パルス撮像 (shuttered light pulse imaging) を含む、様々な技法を介して、光の反射ビームの強度を経時的に解析することによって、深度カメラシステム 20 からターゲットまたは物体上の特定の位置までの物理距離を間接的に決定するために、飛行時間解析を使用することもできる。

【0029】

[0064] 別の例示的な実施形態では、深度カメラシステム 20 は、構造化照明を使用して、深度情報をキャプチャーすることができる。そのような解析では、パターン化された光 (すなわち、グリッドパターンまたはストライプパターンなど既知のパターンで表示される光) を、例えば IR 光放射器 24 を介して、シーン上に投影することができる。シーン内の 1 つまたは複数のターゲットまたは物体の表面に当たると、それに応じて、パターンが変形されることがある。パターンのそのような変形は、例えば赤外光カメラ 26 および／または RGB カメラ 28 によって、キャプチャーすることができ、その後、深度カメラシステムからターゲットまたは物体上の特定の位置までの物理距離を決定するために解析することができる。

40

【0030】

[0065] 深度カメラシステム 20 は、深度情報を生成するために分解できる視覚的ステレ

50

オデータを獲得するために、シーンを異なる角度から見る事ができる2つ以上の物理的に離れたカメラを含むことができる。

【0031】

[0066]深度カメラシステム20は、音波を受け取り、電気信号に変換する、例えばトランスデューサーまたはセンサーを含む、マイクロフォン30をさらに含むことができる。加えて、コンピューティング環境12によって実行されるアプリケーションを制御するために、マイクロフォン30を使用して、人によって提供される音などのオーディオ信号を受け取ることができる。オーディオ信号は、話された言葉、口笛、叫び、および他の発声などの、人の発声音の他、手を叩く、または足を踏み鳴らすなどの、非発声音も含むことができる。

10

【0032】

[0067]深度カメラシステム20は、画像カメラコンポーネント22と通信するプロセッサ32を含むことができる。プロセッサ32は、例えば、深度画像を受け取るため、深度画像に基づいてボクセルからなるグリッドを生成するため、ボクセルからなるグリッド内に含まれる背景を除去して、人ターゲットに関連する1つもしくは複数のボクセルを分離するため、分離された人ターゲットの1つもしくは複数の四肢の位置もしくは場所を決定するため、1つもしくは複数の四肢の位置もしくは場所に基づいてモデルを調整するための命令、または以下でより詳細に説明される他の任意の適切な命令を含む命令を実行できる、標準化されたプロセッサ、専用プロセッサ、またはマイクロプロセッサなどを含むことができる。

20

【0033】

[0068]深度カメラシステム20は、プロセッサ32によって実行される命令を記憶できる他、3DカメラもしくはRGBカメラによってキャプチャされた画像もしくは画像のフレーム、または他の任意の適切な情報もしくは画像なども記憶できる、メモリーコンポーネント34をさらに含むことができる。例示的な一実施形態によれば、メモリーコンポーネント34は、ランダムアクセスメモリー(RAM)、リードオンリーメモリー(ROM)、キャッシュ、フラッシュメモリー、ハードディスク、または他の任意の適切な有形なコンピューター可読記憶コンポーネントを含むことができる。メモリーコンポーネント34は、バス21を介して画像キャプチャコンポーネント22およびプロセッサ32と通信する、別個のコンポーネントとすることができる。別の実施形態によれば、メモリーコンポーネント34は、プロセッサ32および／または画像キャプチャコンポーネント22内に組み込むことができる。

30

【0034】

[0069]深度カメラシステム20は、通信リンク36を介してコンピューティング環境12と通信することができる。通信リンク36は、有線および／または無線接続とすることができる。一実施形態によれば、コンピューティング環境12は、深度カメラシステム20の視野内にある物理空間からいつ画像データをキャプチャすべきかを知らせるクロック信号を、通信リンク36を介して深度カメラシステム20に提供することができる。

【0035】

[0070]加えて、深度カメラシステム20は、例えば3Dカメラ26および／もしくはRGBカメラ28によってキャプチャされた深度情報および画像、ならびに／または深度カメラシステム20によって生成できるスケルタルモデル(skeletal model)を、通信リンク36を介してコンピューティング環境12に提供することができる。その後、コンピューティング環境12は、モデル、深度情報、およびキャプチャされた画像を使用して、アプリケーションを制御することができる。例えば、図2に示されるように、コンピューティング環境12は、(ユーザーが姿勢を取った、または動いたときに)モデルによって実行され得るジェスチャー／姿勢に関する情報を各々が有する、ジェスチャー／姿勢フィルターの集まりなどの、ジェスチャー／姿勢ライブラリ190を含むことができる。例えば、様々な手の姿勢および／またはジェスチャーのためのジェスチャー／姿勢フィルターを提供することができる。手の例示的な姿勢は、「ストップ」姿勢にお

40

50

ける手のひらを前に向けて手を上げること、および数量を示すために1本または複数本の指を立てることなどを含む。手の例示的なジェスチャーは、手をスワイプする（swipe）こと、または振り立てる（fling）ことを含む。検出された姿勢または運動を各フィルターと比較することによって、人によって実行された特定の姿勢またはジェスチャーを識別することができる。

【0036】

[0071]深度カメラシステム20によってキャプチャーされてモデルの形式を取るデータは、アプリケーションへの様々な制御入力に関連付けられた1つまたは複数の特定の姿勢または動きを、（スケルタルモデルによって表される）ユーザーがいつ実行したかを識別するために、ジェスチャー／姿勢ライブラリ190内のジェスチャー／姿勢フィルターと比較することができる。

10

【0037】

[0072]コンピューティング環境は、メモリー194内に記憶された命令を実行して、オーディオビデオ出力信号を表示デバイス196に提供するための、および本明細書で説明される他の機能を達成するための、プロセッサ192も含むことができる。

【0038】

[0073]メモリー34は、本明細書で説明されるようなユーザーの手または他の有関節身体部分の姿勢をモデル化する方法を実行するために、少なくとも1つのプロセッサをプログラミングするための、コンピューター可読ソフトウェアがその上に具体化される、有形なコンピューター可読記憶デバイスとすることができる。また、プロセッサ32は、ユーザーの姿勢をモデル化するためのプロセッサ実施方法を可能にする、1つまたは複数のプロセッサとすることができ、方法は、本明細書で説明されるプロセッサ実施ステップを含む。

20

【0039】

[0074]図3は、図1の撮像システムにおいて使用できるコンピューティング環境の例示的なブロック図を示している。コンピューティング環境は、1つまたは複数のジェスチャーまたは他の動きを解釈し、それに応じて、ディスプレイ上の視覚空間を更新するために使用することができる。上で示されたコンピューティング環境12などのコンピューティング環境は、ゲームコンソールなどの、マルチメディアコンソール100を含むことができる。マルチメディアコンソール100は、レベル1キャッシュ102、レベル2キャッシュ104、およびフラッシュROM（リードオンリーメモリー）106を有する、中央処理装置（CPU）101を有する。レベル1キャッシュ102およびレベル2キャッシュ104は、データを一時的に記憶し、その結果、メモリアクセスサイクルの回数を減らし、それによって、処理スピードおよびスループットを改善する。2つ以上のコアを、したがって、追加のレベル1キャッシュ102およびレベル2キャッシュ104有する、CPU101を提供することができる。フラッシュROMなどのメモリー106は、マルチメディアコンソール100に電源が入れられたときに、ブートプロセスの初期フェーズ中にロードされる、実行可能コードを記憶することができる。

30

【0040】

[0075]グラフィックス処理ユニット（GPU）108およびビデオエンコーダー／ビデオコーデック（エンコーダー／デコーダー）114は、高速および高解像度グラフィックス処理のためのビデオ処理パイプラインを形成する。データは、バスを介してグラフィックス処理ユニット108からビデオエンコーダー／ビデオコーデック114に搬送される。ビデオ処理パイプラインは、データを、テレビまたは他のディスプレイに送るために、A/V（オーディオ／ビデオ）ポート140に出力する。RAM（ランダムアクセスメモリー）などの様々なタイプのメモリー112へのプロセッサアクセスを容易にするために、メモリーコントローラー110が、GPU108に接続される。

40

【0041】

[0076]マルチメディアコンソール100は、好ましくはモジュール118上に実装される、I/Oコントローラー120、システム管理コントローラー122、オーディオ処理

50

ユニット123、ネットワークインターフェース124、第1のUSBホストコントローラー126、第2のUSBコントローラー128、および前面パネルI/Oサブアセンブリー130を含む。USBコントローラー126、128は、周辺機器コントローラー142(1)~142(2)、無線アダプター148、および外部メモリーデバイス146(例えば、フラッシュメモリー、外部CD/DVD ROMドライブ、着脱可能媒体など)に対するホストとして機能する。ネットワークインターフェース(NWIF)124および/または無線アダプター148は、ネットワーク(例えば、インターネット、ホームネットワークなど)へのアクセスを提供し、イーサネット(登録商標)カード、モデム、Bluetooth(登録商標)モジュール、およびケーブルモデムなどを含む、多種多様な様々な有線または無線アダプターコンポーネントのいずれかとすることができる。

10

【0042】

[0077]ブートプロセス中にロードされるアプリケーションデータを記憶するために、システムメモリー143が提供される。媒体ドライブ144が提供され、媒体ドライブ144は、DVD/CDドライブ、ハードドライブ、または他の着脱可能媒体ドライブを含むことができる。媒体ドライブ144は、マルチメディアコンソール100に内蔵すること、または外付けすることができる。アプリケーションデータは、実行、再生などのために、マルチメディアコンソール100によって、媒体ドライブ144を介してアクセスすることができる。媒体ドライブ144は、シリアルATAバスまたは他の高速接続などのバスを介して、I/Oコントローラー120に接続される。

【0043】

20

[0078]システム管理コントローラー122は、マルチメディアコンソール100の利用可能性の保証に関する様々なサービス機能を提供する。オーディオ処理ユニット123およびオーディオコーデック132は、高忠実度およびステレオ処理を行う、対応するオーディオ処理パイプラインを形成する。オーディオデータは、通信リンクを介して、オーディオ処理ユニット123とオーディオコーデック132の間で搬送される。オーディオ処理パイプラインは、データを、オーディオ機能を有する外部オーディオプレーヤーまたはデバイスによる再生のために、A/Vポート140に出力する。

【0044】

[0079]前面パネルI/Oサブアセンブリー130は、電源ボタン150および排出ボタン152の機能の他、マルチメディアコンソール100の外面上に露出する任意のLED(発光ダイオード)または他のインジケータをサポートする。システム電源モジュール136は、マルチメディアコンソール100のコンポーネントに電力を提供する。ファン138は、マルチメディアコンソール100内の回路を冷却する。

30

【0045】

[0080]CPU101、GPU108、メモリーコントローラー110、およびマルチメディアコンソール100内の他の様々なコンポーネントは、様々なバスアーキテクチャのいずれかを使用する、シリアルおよびパラレルバス、メモリーバス、周辺バス、ならびにプロセッサまたはローカルバスを含む、1つまたは複数のバスを介して相互接続される。

【0046】

40

[0081]マルチメディアコンソール100に電源が入れたとき、アプリケーションデータを、システムメモリー143からメモリー112および/またはキャッシュ102、104にロードし、CPU101上で実行することができる。アプリケーションは、マルチメディアコンソール100上で利用可能な異なる媒体タイプに移動したときにも一貫したユーザーエクスペリエンスを提供する、グラフィカルユーザーインターフェースを提示することができる。動作中、媒体ドライブ144内に含まれるアプリケーションおよび/または他のメディアを、媒体ドライブ144から起動または再生して、マルチメディアコンソール100に追加の機能を提供することができる。

【0047】

[0082]マルチメディアコンソール100は、システムをテレビまたは他のディスプレー

50

に単純に接続することによって、スタンドアロンシステムとして動作させることができる。このスタンドアロンモードでは、マルチメディアコンソール100は、1人または複数人のユーザーが、システムと対話すること、映画を観ること、または音楽を聴くことを可能にする。しかし、ネットワークインターフェース124または無線アダプター148を介して利用可能になるブロードバンド接続を組み込むことによって、マルチメディアコンソール100はさらに、より大規模なネットワークコミュニティへの参加者として動作させることができる。

【0048】

[0083]マルチメディアコンソール100に電源が入れられたとき、指定された量のハードウェアリソースが、マルチメディアコンソールオペレーティングシステムによるシステム使用のために予約される。これらのリソースは、メモリー（例えば16MB）、CPUおよびGPUサイクル（例えば5%）、ネットワーク帯域幅（例えば8kbs）などの予約を含むことができる。これらのリソースはシステムブート時間に予約されるので、予約されたリソースは、アプリケーションの見地からは存在していない。

【0049】

[0084]特に、メモリー予約は、起動カーネル、並列システムアプリケーションおよびドライバを含むのに十分な大きさであることが好ましい。CPU予約は、予約されたCPU使用がシステムアプリケーションによって使用されない場合は、アイドルスレッドが任意の未使用サイクルを消費するように、好ましくは一定である。

【0050】

[0085]マルチメディアコンソール100がブートされ、システムリソースが予約された後、システム機能を提供するために、並列システムアプリケーションが実行される。システム機能は、上で説明した予約されたシステムリソース内で実行される、1組のシステムアプリケーション内にカプセル化される。オペレーティングシステムカーネルは、ゲームアプリケーションスレッドに対してシステムアプリケーションスレッドであるスレッドを識別する。システムアプリケーションは、アプリケーションからシステムリソースが一貫した見え方をするように、好ましくは所定の時間および間隔でCPU101上において動作するようにスケジューリングされる。スケジューリングは、コンソール上で動作しているゲームアプリケーションが被るキャッシュ中断（cache disruption）を最低限に抑えるためである。

【0051】

[0086]並列システムアプリケーションがオーディオを必要とする場合、時間に対して鋭敏であるため、ゲームアプリケーションとは非同期にオーディオ処理がスケジューリングされる。（以下で説明される）マルチメディアコンソールアプリケーションマネージャーは、システムアプリケーションがアクティブである場合、ゲームアプリケーションのオーディオレベルを抑制する（例えば、消音、減衰）。

【0052】

[0087]入力デバイス（例えばコントローラー142（1）、142（2））は、ゲームアプリケーションおよびシステムアプリケーションによって共用される。入力デバイスは、予約リソースではないが、システムアプリケーションとゲームアプリケーションの間に切り換えられて、各々がデバイスを占有する（have a focus of the device）ようにすべきである。アプリケーションマネージャーは、好ましくは、ゲームアプリケーションの知識を知ることなく、入力ストリームの切り換えを制御し、ドライバが占有の切り換えに関する状態情報を維持する。コンソール100は、カメラ26および28を含む、図2の深度カメラシステム20から、追加の入力を受け取ることができる。

【0053】

[0088]図4は、図1の撮像システムにおいて使用できるコンピューティング環境の別の例示的なブロック図を示している。

[0089]撮像システムでは、コンピューティング環境は、1つまたは複数のジェスチャー

10

20

30

40

50

または姿勢を解釈し、それに応じて、ディスプレイ上の視覚空間を更新するために使用することができる。コンピューティング環境220は、コンピューター241を含み、コンピューター241は、一般に様々な有形なコンピューター可読記憶媒体またはデバイスを含む。これは、コンピューター241によってアクセスできる任意の利用可能な媒体とすることができる、揮発性および不揮発性媒体、着脱可能および着脱不能媒体をともに含む。システムメモリ222は、リードオンリーメモリ（ROM）223およびランダムアクセスメモリ（RAM）260など、揮発性および／または不揮発性メモリの形態を取るコンピューター記憶媒体を含む。基本入出力システム224（BIOS）は、スタートアップ中などにコンピューター241内の要素間における情報の転送を助ける基本ルーチンを含み、一般にROM223内に記憶される。RAM260は、処理ユニット259から直ちにアクセス可能な、および／または処理ユニット259によって現在操作されている、データおよび／またはプログラムモジュールを一般に含む。グラフィックスインターフェース231は、GPU229と通信する。限定することなく、例を挙げると、図4は、オペレーティングシステム225、アプリケーションプログラム226、他のプログラムモジュール227、およびプログラムデータ228を示している。

【0054】

[0090]コンピューター241は、例えば、着脱不能な不揮発性の磁気媒体から読み取り、それに書き込む、ハードディスクドライブ238、着脱可能な不揮発性の磁気ディスク254から読み取り、それに書き込む、磁気ディスクドライブ239、およびCD ROMまたは他の光媒体などの着脱可能な不揮発性の光ディスク253から読み取り、それに書き込む、光ディスクドライブ240など、他の着脱可能／着脱不能、揮発性／不揮発性のコンピューター記憶媒体も含むことができる。例示的な動作環境において使用できる、他の着脱可能／着脱不能、揮発性／不揮発性の有形なコンピューター可読記憶媒体またはデバイスは、磁気テープカセット、フラッシュメモリーカード、デジタル多用途ディスク、デジタルビデオテープ、ソリッドステートRAM、およびソリッドステートROMなどを含むが、それらに限定されない。ハードディスクドライブ238は、一般に、インターフェース234などの着脱不能メモリーインターフェースを介して、システムバス221に接続され、磁気ディスクドライブ239および光ディスクドライブ240は、一般に、インターフェース235などの着脱可能メモリーインターフェースを介して、システムバス221に接続される。

【0055】

[0091]上で説明され、図4に示される、ドライブおよび関連するコンピューター記憶媒体は、コンピューター241のために、コンピューター可読命令、データ構造、プログラムモジュール、および他のデータの記憶を可能にする。例えば、ハードディスクドライブ238は、オペレーティングシステム258、アプリケーションプログラム257、他のプログラムモジュール256、およびプログラムデータ255を記憶するものとして示されている。これらのコンポーネントは、オペレーティングシステム225、アプリケーションプログラム226、他のプログラムモジュール227、およびプログラムデータ228と同じであることも、または異なることもできることに留意されたい。オペレーティングシステム258、アプリケーションプログラム257、他のプログラムモジュール256、およびプログラムデータ255には、少なくとも、それらが異なるコピーであることを示すために、異なる番号が与えられている。ユーザーは、キーボード251、およびマウス、トラックボール、またはタッチパッドと一般には呼ばれるポインティングデバイス252などの、入力デバイスを介して、コンピューター241にコマンドおよび情報を入力することができる。他の入力デバイス（図示されず）は、マイクロフォン、ジョイスティック、ゲームパッド、サテライトディッシュ、またはスキャナーなどを含むことができる。これらおよび他の入力デバイスは、しばしばシステムバスに結合されるユーザー入力インターフェース236を介して、処理ユニット259に接続されるが、パラレルポート、ゲームポート、またはユニバーサルシリアルバス（USB）などの、他のインターフェースおよびバス構造によって接続されることもある。カメラ26および28を含む、図2

の深度カメラシステム 20 は、コンソール 100 のための追加の入力デバイスを定義することができる。モニター 242 または他のタイプのディスプレイも、ビデオインターフェース 232 などのインターフェースを介して、システムバス 221 に接続される。モニターに加えて、コンピューターは、出力周辺インターフェース 233 を介して接続できる、スピーカー 244 およびプリンター 243 などの、他の周辺出力デバイスも含むことができる。

【0056】

[0092] コンピューター 241 は、リモートコンピューター 246 などの 1 つまたは複数のリモートコンピューターへの論理接続を使用して、ネットワーク環境において動作することができる。リモートコンピューター 246 は、パーソナルコンピューター、サーバー、ルーター、ネットワーク PC、ピアデバイス、または他の共通ネットワークノードとすることができ、一般に、コンピューター 241 に関連して上で説明された要素の多くまたはすべてを含むが、図 4 ではメモリー記憶デバイス 247 だけが示されている。論理接続は、ローカルエリアネットワーク (LAN) 245 およびワイドエリアネットワーク (WAN) 249 を含むが、他のネットワークも含むことができる。そのようなネットワーク環境は、オフィス、企業規模のコンピューターネットワーク、イントラネット、およびインターネットにおいて一般的である。

【0057】

[0093] LAN ネットワーク環境において使用される場合、コンピューター 241 は、ネットワークインターフェースまたはアダプター 237 を介して、LAN 245 に接続される。WAN ネットワーク環境において使用される場合、コンピューター 241 は、インターネットなどの WAN 249 を介する通信を確立するためのモデム 250 または他の手段を含む。モデム 250 は、内蔵または外付けとすることができ、ユーザー入力インターフェース 236 または他の適切なメカニズムを介して、システムバス 221 に接続することができる。ネットワーク環境では、コンピューター 241 に関連して示されたプログラムモジュールまたはその一部は、リモートメモリー記憶デバイス内に記憶することができる。限定することなく、例を挙げると、図 4 は、メモリーデバイス 247 上に存在するものとして、リモートアプリケーションプログラム 248 を示している。示されるネットワーク接続は例示的なものであり、コンピューター間において通信リンクを確立する他の手段も使用できることが理解されよう。

【0058】

[0094] コンピューティング環境は、本明細書で説明されるようなユーザーの手または他の有関節身体部分の姿勢をモデル化する方法を実行するために、少なくとも 1 つのプロセッサをプログラミングするための、コンピューター可読ソフトウェアがその上に具体化される、有形なコンピューター可読記憶デバイスを含むことができる。有形なコンピューター可読記憶デバイスは、例えば、コンポーネント 222、234、235、230、253、および 254 の 1 つまたは複数を含むことができる。さらに、コンピューティング環境の 1 つまたは複数のプロセッサは、ユーザーの手または他の有関節身体部分の姿勢をモデル化するためのプロセッサ実施方法を可能にすることができ、方法は、本明細書で説明されるプロセッサ実施ステップを含む。プロセッサは、例えば、コンポーネント 229 および 259 の 1 つまたは複数を含むことができる。

【0059】

[0095] 図 5 は、(菱形によって示される) 関節と、(四角形によって示される) アトラクトポイントとを含む、ユーザーの手のモデルを示している。姿勢が検出されるべき所定の有関節身体部分に対するモデルを提供することができる。1 つの手法では、手領域のモデルが提供される。モデル 500 は、腕 504 と、手のひら 502 と、指 510、520、530、540 と、親指 550 とを含むことができる。指は、末端部セグメント (distal segment) 512、522、532、542 と、中間部セグメント (middle segment) 514、524、534、544 と、基部セグメント (proximal segment) 516、526、536、546 とを含む。親指は、

10

20

30

40

50

末端部セグメント 5 5 2 と、基部セグメント 5 5 4 とを含む。モデルの指セグメントおよび親指セグメントは、人体の指骨に対応する。腕／手首と手のひらの間の関節 5 0 3、末端部親指部分と基部親指部分の間の関節 5 6 0、および基部親指部分と手のひらの間の関節 5 6 4 など、隣接する身体部分の間には関節が設けられる。これらの関節は、定められた範囲の動きを行う、人体内の関節に一般に対応する。この動きは、主として角度的なものであるが、同じく並進的なものでもあり得る。指先（末端部指セグメントの先端および末端部親指セグメントの先端）などにおける追加のポイントは、末端部指セグメントの位置を定義するために使用することができる。これらの先端ポイントは、人体骨格の関節に対応しないが、便宜的に、それらも「関節」と呼ばれる。末端部親指セグメントにおける関節 5 5 6 が、一例である。したがって、関節 5 5 6 など、いくつかの関節は、人体骨格における関節に対応しないことがあり、関節 5 6 0、5 6 4 など、いくつかの関節は、人体骨格における関節に対応する。身体部分の方向は、身体部分の少なくとも 2 つの関節の位置によって定めることができる。例えば、指セグメントなどの細長い身体部分の場合、その身体部分の対向端に関節が存在することがある。

10

【0060】

[0096]さらに、各身体部分に対して、1 つまたは複数の代表的なアトラクトポイントを指定することができる。例えば、指 5 3 0 の末端部指セグメント、指 5 4 0 の末端部指セグメント、末端部親指部分 5 5 2、基部親指部分 5 5 4、手のひら 5 0 2、および腕 5 0 4 に対して、それぞれ、例示的な代表的なアトラクトポイント 5 3 1、5 4 1、5 5 8、5 6 2、5 0 1、および 5 0 5 が提供される。例えば、手のひらなど、相対的に大きい、および／または相対的により複雑な範囲の動きを行う、身体部分に対しては、複数のアトラクトポイントを同様に使用することができる。指 5 4 0 の中間部および基部指セグメントに対して、それぞれ、アトラクトポイント 5 4 3 および 5 4 5 が提供される。

20

【0061】

[0097]モデルの身体部分の 1 つまたは複数が、アトラクトポイントを有さないことも可能である。

[0098]一般に、この指および他の指では、小さな四角形が、代表的なアトラクトポイントを表す。代表的なアトラクトポイントは、モデルの身体部分を代表するポイントとすることができる。「アトラクトポイント」という用語は、マッチングプロセスにおいて、アトラクトポイントが深度センサーデータの方へ移動される、または引き付けられる (a t t r a c t e d) ことを示している。1 つの手法では、代表的なアトラクトポイントは、身体部分の中心点または重心に存在する。重心は、身体部分の内部または身体部分の表面に存在することができる。表面は、深度軸に沿って、または深度カメラに向かう視線に沿って、深度カメラの方を向くことができる。重心は、身体部分の中心点であると見なすことができる。

30

【0062】

[0099]図 7 A のステップ 7 3 4 に関連して以下で説明される、重要度の軸 (a x i s o f i m p o r t a n c e) 5 0 7 は、重心対アトラクトポイントマッチングのために使用することができる。

【0063】

40

[00100]モデルは、様々な身体部分の許容可能な範囲の動きを示すことができる。各身体部分は、隣接する身体部分に対して、許容可能な角度範囲の動きを行うことができる。例えば、各指の基部セグメントは、手のひらに対して、許容可能な角度範囲の動きを行うことができ、各指の中間部セグメントは、指の基部セグメントに対して、許容可能な角度範囲の動きを行うことができるなどである。これらは、角度制約によって定義することができる。身体部分のいくつかの相対的な並進運動も許容されることがある。異なる身体部分が衝突できないこと、例えば、全体的または部分的に、同じ時間に同じ位置に存在できないことを示す、衝突制約を提供することができる。

【0064】

[00101]モデルは、身体部分が一般に独立しており、自由に動けるように定義すること

50

ができるが、互いに隣接していると思われる身体部分が実際に隣接している配置を得るのに有利な距離制約を設定することができる。隣接性は、隣接する身体部分の関節が実質的に同じ位置にあることを含むことができる。この手法では、各身体部分は、その先端に、隣接する身体部分の先端における関節と実質的に一致する、関節を有することができる。一例として、末端部親指部分と基部親指部分の間の関節 5 6 0 は、基部親指部分 5 5 4 の末端にある 1 つの関節と、末端部親指部分 5 5 2 の基端にある 1 つの関節とを表すことができる。末端部親指部分と基部親指部分が互いに離れている場合に、距離制約違反の一例が生じる。

【0065】

[00102]別の手法では、例えば、末端部親指部分と基部親指部分の間に 1 つの関節が存在し、距離制約は、その関節と親関節の間の距離である。モデルのこの手法では、メッシュが、子関節に達するほど十分に長くはないとしても、子関節の方を指して親関節に配置される。親関節または子関節は、それぞれ、親身体部分または子身体部分に関連付けることができる。1 つの可能な手法では、子身体部分は、手のひらから伸びる指など、より大きな身体部分から伸びるより小さな身体部分とすることができる。または、子身体部分は、腕から指先に向かう方向など、指定された方向に別の身体部分から伸びる身体部分とすることができる。例えば、基部親指部分 5 5 4 は、末端部親指部分 5 5 2 の親とすることができる。

【0066】

[00103]身体部分のモデルは、所望のレベルの詳細さを有することができる。例えば、手のひらを単一エンティティとしてモデル化する代わりに、中手骨および関節をモデル化することなどによって、さらに詳細に手のひらをモデル化することができる。モデルは、長方形、立方体、楕円体、および球体など、個別の体積／形状を含むことができる。これらの体積は、所望のレベルの粒度を提供するために、サイズおよび形状を変えることができる。代替的または追加的に、適切にリギング (r i g) およびスキニング (s k i n) が行われた連続的な面を形成するメッシュによって、モデルを表すことができる。リギングおよびスキニングとは、スケルトンを 3 D メッシュ内に配置し、スケルトンを動かした場合にスムーズに変形する身体部分を生成するために、スケルタルセグメントをメッシュの対応する部分にマッピングすることである。モデルの複雑さは、所望のリアリズムの程度および利用可能な計算リソースなどの要因を考慮すべきである。

【0067】

[00104]さらに、有関節身体部分の一例として人間の手について説明したが、その概念は、既知の範囲の動きを行う物体の他の有関節身体部分に拡張することができる。物体は、人もしくは動物、またはロボットもしくは他の無生物物体などの、生物／物体を含むことができる。1 つまたは複数のそのような物体を同時にトラッキングすることができる。さらに、アプリケーションの部品として深度センサーが使用される場合、深度センサーによって撮像される人は、アプリケーションの能動的なユーザー／参加者、または受動的な見物人とすることができる。

【0068】

[00105]図 6 は、撮像システムにおいて改善された忠実度で有関節身体部分の姿勢を検出するための方法を示している。ステップ 6 0 0 は、先に言及した深度カメラシステムからなど、深度センサーからユーザーの深度ピクセルを獲得するステップを含む。例えば、これは、深度ピクセルからなるフレームとすることができる。ステップ 7 0 0 は、図 7 A のステップ 7 0 2 ～ステップ 7 1 4 に関連してさらに説明されるような、プリマッチング処理を実行するステップを含む。ステップ 7 2 0 は、図 7 A のステップ 7 2 1 ～ステップ 7 3 6 に関連してさらに説明されるような、重心を使用してマッチングを実行するステップを含む。ステップ 7 5 0 は、図 7 A のステップ 7 5 2 ～ステップ 7 5 6 に関連してさらに説明されるような、重心を使用してマッチングを改善するステップを含む。ステップ 7 6 0 は、図 7 B に関連してさらに説明されるような、ラスタ化モデルを使用してマッチングを改善するステップを含む。1 つの手法では、ステップ 7 0 0、7 2 0、7 5 0、お

10

20

30

40

50

との間の2部マッチング(bipartite matching)と、確率に基づいた、重心の信頼度尺度(confidence measure)とを含むことができる。

【0073】

[00110]ステップ734は、重心対アトラクトポイントマッチングを実行する。図11が、さらなる詳細を提供する。重心セット728、730およびモデルの代表的なアトラクトポイント744(図5を参照)を入力として提供することができる。代表的なアトラクトポイントは、各身体部分について、深度センサーデータのマッチング重心と一致するように理想的に作成された、有関節モデル内のポイントである。ステップ734は、様々な要因を使用して計算されたマッチスコアを使用して、重心セット内のすべての重心とモデルのすべてのアトラクトポイントとの間の1つまたは複数の最良マッチを見つけることを含むことができる。

10

【0074】

[00111]1つの要因は、各マッチした重心とアトラクトポイントの間の距離に関し、その距離は、例えば、重心とアトラクトポイントとを接続するベクトルの大きさによって決定することができる。アトラクトポイントにより近い重心には、アトラクトポイントからより遠い重心よりも、より高いスコアを割り当てることができ、より高いスコアは、適切なマッチである可能性がより高いことに関連する。

【0075】

[00112]別の要因は、重要度の軸に沿った重心の順序付けに関する。例えば、手のモデルの場合、重要度の軸は、手のひらと指の基部セグメントとの間の関節に沿って伸び、例えば指510、520、530、540などの指の(小指の付根から人差し指の付根に向かう)予想される順序を示すことができる。図5の例示的な重要度の軸507を参照されたい。例えば、重心をこの軸に直交する方向に動かして、重心が軸上のどこに落ちるかを決定することによって、重心をこの軸上に投影することができる。点の直線上への投影は、閉形式で解析的に計算することができる。その後、軸に沿った重心の順序を決定することができる。おそらく指を交差させている場合を除いて、一般に、指は予想される順序を取る。重要度の軸に沿って予想される順序を取る重心には、より高いスコアを割り当てることができ、より高いスコアは、適切なマッチである可能性がより高いことに関連する。重要度の軸に沿って予想される順序を取らない重心には、より低いスコアを割り当てることができ、より低いスコアは、適切なマッチである可能性がより低いことに関連する。重要度の軸を別の方向に取ることができ、2つ以上の重要度の軸を同様に使用することもできる。

20

30

【0076】

[00113]別の要因は、隣接する重心または身体部分との接続性に関する。例えば、指の末端部セグメントの重心は、別の指のセグメントの重心ではなく、同じ指の中間部セグメントの重心に連結するピクセル経路(pixel path)内にあることが予想される。隣接する身体部分の重心に連結する経路内にある重心には、より高いスコアを割り当てることができ、より高いスコアは、適切なマッチである可能性がより高いことに関連する。隣接する身体部分の重心に連結する経路内にない重心には、より低いスコアを割り当てることができ、より低いスコアは、適切なマッチである可能性がより低いことに関連する。この要因は、比喩的に深度マップを一直線に歩いて、ピクセルデータ内に何らかの不連続が存在するかどうかを決定することによって、評価することができる。この要因は、ノイズおよび他のエッジケースに対してより頑健になるための、経路がどれだけ直線からそれ得るかについての制約を有する、重心の間の深度マップに沿った経路探索などの、より複雑な方法を使用して、評価することもできる。例えば、指の末端部セグメントから同じ指の中間部セグメントまで歩こうと試みたときに、ピクセルデータ内にギャップが検出された場合、これは、マッチが誤りである可能性が高いことを示す。すなわち、指セグメントの一方または両方が、正確に識別されていない。例えば、各指について、指の末端部と中間部の重心の間に接続を見つけることによって、接続性を評価することができる。

40

【0077】

50

[00114]別の要因は、モデルの身体部分の間の衝突チェックに関する。衝突をもたらさない重心対アトラクトポイントマッチには、より高いスコアを割り当てることができ、より高いスコアは、適切なマッチである可能性がより高いことに関連し、一方、異なる身体部分が同時に同じ位置に存在する、衝突をもたらす重心対アトラクトポイントマッチには、より低いスコアを割り当てることができ、より低いスコアは、適切なマッチである可能性がより低いことに関連する。

【0078】

[00115]重心とアトラクトポイントの間の最良マッチは、これらまたは他の要因の1つまたは複数の何らかの組み合わせに基づくことができる。異なる要因には、例えばテスト結果から決定された相対的な重要度に基づいて、異なる重みを与えることができる。

10

【0079】

[00116]ステップ736は、重心へのアトラクトポイントの剛体変換を実行する。一例として図12Aを参照されたい。剛体変換は、モデルの並進、回転、およびスケーリングの1つまたは複数を含むことができる。このステップは、アトラクトポイントの最適3D点群変換の並進、回転、およびスケーリングの計算と、すべてのアトラクトポイントに対する変換の適用とを含むことができる。並進は、フレームまたは座標系に対するモデルの移動を含むことができる。座標系は、直交するx、y、z軸（図1）などの、任意の軸を含むことができる。変換を表すために、オイラー角（Euler angle）を使用することができる。回転は、軸の1つまたは複数の周りのモデルの回転を含むことができる。スケーリングは、身体部分の相対的な方向または比率を変化させることなく、モデルの全体サイズを変化させる。1つの手法では、全体としての剛体変換は、身体部分の相対的な方向を変化させない。身体部分の相対的な方向は、子身体部分とその親身体部分の間の局所変換として定義することができる。1つの可能な手法では、子身体部分は、手のひらから伸びる指など、より大きな身体部分から伸びるより小さな身体部分とすることができる。立方体の剛体変換の一例は、3D空間内で1つの位置から別の位置に動かされた、より大きなまたはより小さな立方体をもたらすことができる。剛体変換中に身体部分の相対的な方向を維持する一例は、剛体変換の前および後のどちらにおいても、第3の指の長さの4分の3の第1の指を有し、また第3の指に対して特定の角度で伸びる第1の指を有することである。

20

【0080】

[00117]剛体変換の例示的な実施は、アトラクトポイントと重心の間の剛体変換推定を使用して、モデルのグローバルスケーリングを計算する。2つの対応するポイントセットの間の剛体変換（回転、並進、およびすべての次元における一様スケーリング）の例示的な閉形式計算は、参照により本明細書に組み込まれる、Horn, B. K. P.、「Closed-form solution of absolute orientation using unit quaternions（単位四元数を使用する絶対的な方向の閉形式解法）」、Journal of the Optical Society of America、4（4）、629～642（1987）において説明されている。Hornは、次のように述べている。「2つの座標系における多数の点の座標の測定値のペアを使用して両系間の関係を見出すことは、写真測量法における古典的な課題である。それは、実体写真測量およびロボット工学において用途を見出す。私はここで、3つ以上の点についての最小2乗問題に対する閉形式解法を提示する。現在、様々な経験的な方法、グラフィカルな方法、および数値反復法が使用されている。解の導出は、回転を表す単位四元数の使用によって簡略化される。私は、この問題に対する解法が有すべき対称性を強調する。最良の並進オフセットは、一方の系における座標の重心と他方の系における座標の回転およびスケーリングが施された重心との間の差である。最良のスケールは、2つの系における座標の、それぞれの重心からの2乗平均平方根偏差（root-mean-square deviation）の比に等しい。これらの正確な結果は、選択された数個の点の測定に基づいた近似法よりも好ましい。最良の回転を表す単位四元数は、対称4×4行列の最大の正固有値に関連する固有ベクトルである。この行列の要素は

30

40

50

、点の対応する座標の積の和の結合 (combinations of sums of products) である」。

【0081】

[00118]剛体変換の別の例示的な実施は、A. Balan、「Restricted Affine Registration of 3D Point Clouds (3D点群の制限付きアフィンレジストレーション)」、<http://priorartdatabase.com/IPCOM/000207623>、2011年6月7日において説明されている。この実施は、重心とアトラクションポイント (attraction point) の間の剛体変換を計算し、1つの重要な特徴は、モデルのx軸、y軸、およびz軸に沿った異なるスケーリングパラメータを推定できることである。それは、回転、並進、および異なる次元 (幅、高さ、長さ) ごとに独立した引き伸ばしを行う引き伸ばしによって、センサー観測に合わせて、物体の幾何学的モデルのレジストレーションを行う問題に対処する。この問題は、物体トラッキングおよびサイズ推定のためのモデルフィッティングにおいて用途を有する。形状は、モデルとセンサー観測との間でペアごとにマッチさせられる、3D表面点の集まりとして表される。それは、対応するセンサー測定点までの距離を最小化する、モデル点の制限付きアフィン変換 (affine transformation) を計算する、反復的な解法を説明している。

10

【0082】

[00119]ステップ752～ステップ756は、図7Aのステップ750に対応する。ステップ752、754、および756は、重心対アトラクトポイント磁力プロセス (magnetism process) の一部と見なすことができる。ステップ752は、アトラクトポイントと重心のマッチングペアについて、アトラクトポイントを重心のより近くに移動させることを含む。例えば、アトラクトポイントと重心のマッチングペアが、モデルの基部親指部分554のアトラクトポイント562と深度センサーデータの重心916、またはモデルの末端部親指部分552のアトラクトポイント558と深度センサーデータの重心904を含むことができる、図12Aおよび図12Bを参照されたい。

20

【0083】

[00120]ステップ752は、ステップ736から獲得された変換された代表的なアトラクトポイントのどれが、重心の対応する1つからそれぞれの閾値距離よりも遠いかを決定すること、および重心の対応する1つからそれぞれの閾値距離よりも遠い、変換された代表的なアトラクトポイントを、重心の対応する1つからそれぞれの閾値距離よりも近くなるように調整することを含むことができる。1つの手法では、それぞれの閾値距離は、各代表的なアトラクトポイントに固有の調整された距離とすることができる。または、それぞれの閾値距離は、異なる代表的なアトラクトポイントに共通とすることができる。

30

【0084】

[00121]ステップ754は、アトラクトポイントの移動に基づいて、モデルの関節を動かすことを含む。アトラクトポイントの移動は、モデルの1つまたは複数の関節の動きと一致するので、ステップ752および754は、一緒に実行することができる。1つまたは複数のアトラクトポイントの移動によって、関節の動きに影響を及ぼすことができる。関節の動きは、動きの3D空間内における方向および大きさを定義するベクトルまたは力によって定義することができる。ステップ754への入力、関節データ746であり、それは、3D空間内のモデルの各関節の位置を定義することができる。判断ステップ756において、例えば次の身体部分など、アトラクトポイントの次のペアが存在する場合、ステップ752および754が繰り返される。判断ステップ756が偽である場合、プロセスは、継続ポイント「E」において図7Bに引き継がれる。モデルの連続面を表すメッシュデータ748も、関節データ746に基づいて提供することができる。

40

【0085】

[00122]次に説明される図7Bでは、継続ポイント「A」によって示されるように、確率マップが使用され、継続ポイント「B」によって示されるように、BGR深度フレームが使用され、継続ポイント「C」によって示されるように、関節データ746が使用され

50

、継続ポイント「D」によって示されるように、メッシュデータ748が使用される。

【0086】

[00123]図7Bは、図6のステップ760のさらなる詳細を示している。先に図7Aにおいて作成された重心対アトラクトポイントマッチが、反復によって改善される。1つの手法では、反復は、深度センサーデータの重心を使用せず、モデルのラスタ化から獲得された深度ピクセルを使用する。モデルの関節など、モデルに適用される調整ベクトルを決定するために、モデルの深度ピクセルが、深度センサーの深度ピクセルと比較される。新しい状態のモデルは、次の反復においてラスタ化され、そのプロセスが、停止基準が満たされるまで繰り返される。

【0087】

[00124]ステップ762が、反復を開始する。ステップ764は、入力として、確率マップ、BGR深度フレーム、および関節データを受け取り、モデルをラスタ化して、モデルの深度ピクセルを提供する。例えば、モデルのメッシュ面を、3次元(3-D)方法でラスタ化することができる。1つの手法では、モデルの深度ピクセルは、x-y平面内のデータを提供する2Dバッファと、z軸沿いの深度データを提供するz軸バッファとにおいて提供される。モデルの深度ピクセルは、3D空間内の点である。ラスタ化は、深度センサーの深度ピクセルと同じまたは異なる任意の方向に行われ、必ずしもx軸およびy軸に沿う必要はないことに留意されたい。さらに、ラスタ化は、深度センサーの深度ピクセルと同じ解像度、またはそれよりも細かいもしくは粗い解像度を有することができる。手のモデルをラスタ化することによって獲得されるピクセルの一例については、図12Cのピクセル1220を参照されたい。

【0088】

[00125]ステップ766は、比較平面内において、モデルの深度ピクセルを深度センサーの深度ピクセルと比較する。図13Aおよび図13Bに例示される1つの手法では、平面は、深度軸を横断/直交するx-y平面である。別の手法では、平面は、深度センサーの局所的な視線を横断し、深度軸の横断からは僅かにずれている。ピクセルは、比較平面を横断する方向に比較平面に投影され、またはすでに比較平面に存在する。1つの手法では、比較平面は、 $z > 0$ と深度カメラに最も近いモデルのピクセルの深度との間のどこかに存在することができる。ピクセルの3つのカテゴリーを定義することができる。カテゴリーの2つは、深度センサーの深度ピクセルについてのものであり(ステップ768および770)、カテゴリーの1つは、モデルの深度ピクセルについてのものである(ステップ772)。

【0089】

[00126]具体的には、ステップ768は、モデルの深度ピクセルとオーバーラップする深度センサーのオーバーラップ深度ピクセルを識別する。「オーバーラップ」という用語は、比較平面内において、ピクセルが、少なくとも部分的に、同じ領域を占有することを指す。図12Dのピクセル1230が一例である。ピクセルをオーバーラップと分類するのに十分な、ピクセル間の最小度のオーバーラップを定義することができる。これらは、モデルが存在する場所に近似的に存在する、深度センサーの深度ピクセルであり(言い換えると、モデルは、この深度センサーデータが存在する場所に存在し)、モデルがこの深度センサーデータと良好に一致していることを示している。

【0090】

[00127]ステップ770は、モデルの深度ピクセルとオーバーラップしない深度センサーの非オーバーラップ深度ピクセルを識別する。図12Dのピクセル1240が一例である。これらは、モデルが存在しない場所に近似的に存在する、深度センサーの深度ピクセルであり(言い換えると、モデルは、この深度センサーデータが存在する場所に存在せず)、この深度センサーデータとより良好に一致するようにモデルを調整すべきことを示している。

【0091】

[00128]ステップ772は、深度センサーの深度ピクセルとオーバーラップしないモデ

10

20

30

40

50

ルの非オーバーラップ深度ピクセルを識別する。図12Dのピクセル1220が一例である。これらは、深度センサーデータが存在しない場所に近似的に存在する、モデルの深度ピクセルであり（言い換えると、モデルは、深度センサーデータが存在しない場所に存在し）、深度センサーデータとより良好に一致するようにモデルを調整すべきことを示している。

【0092】

[00129]比較に基づいて、ステップ774は、データ776として記憶されるピクセルベースの調整ベクトルを決定する。モデルをピクセルの方に引く／調整する傾向にある3D調整ベクトルを計算することができる。3つのピクセルタイプの各々に基づいて、調整ベクトルを生成し、その後、合算して、関節のための全体的な調整ベクトルを獲得することができる。ピクセルベースの調整ベクトルには、それらがステップ768、ステップ770、またはステップ772のいずれからのものかに基づいて、異なる重みを割り当てることができる。例えば、深度センサーの特性およびアプリケーションの要求に応じて、調整の1つにより重い重みを与えることが望ましいことがある。例えば、横方向解像度が最大の関心事である場合、ステップ770および／またはステップ772の調整により重い重みを与えることができる。深度解像度が最大の関心事である場合、ステップ768の調整により重い重みを与えることができる。

10

【0093】

[00130]ピクセルの3つのタイプの各々について、調整ベクトルは、調整するのに最適な関節であるとして識別されたモデルの関節を調整することができる。この最適な関節は、必ずしもピクセルに最も近い関節である必要はない。最適な関節は、ピクセルからの関節までの距離に基づいて、モデルの異なる関節のスコアを計算することによって、識別することができる。ピクセルにより近い関節は、より高いスコアを有し、選択される確率がより高い。ピクセルからより遠い関節は、より低いスコアを有し、選択される確率がより低い。

20

【0094】

[00131]1つの手法では、深度センサーの非オーバーラップまたはオーバーラップ深度ピクセルの場合、スコアは、確率マップからのピクセルに関連する確率（例えば、先に説明したPr1、Pr2、Pr3、...）に基づくこともできる。ピクセルが相対的に高い確率で関連する身体部分と同じ身体部分に関連する関節は、より高いスコアを有し、選択される確率がより高い。スコアは、関連の確率に比例して高くなることができる。ピクセルが相対的に高い確率で関連する身体部分と異なる身体部分に関連する関節は、より低いスコアを有し、選択される確率がより低い。

30

【0095】

[00132]図13Aの調整ベクトル1320は、深度センサーの非オーバーラップ深度ピクセルの例示的な調整ベクトルである。ベクトルは、関節j1から深度センサーの非オーバーラップ深度ピクセルp3に向かう方向に伸びる。深度センサーのオーバーラップ深度ピクセルの場合、調整ベクトルは、深度センサーの非オーバーラップ深度ピクセルの場合と同じ方法で決定することができる。

【0096】

40

[00133]調整ベクトルは、身体部分の動きの方向に依存する成分を有する。例えば、身体部分がz方向において深度センサーから主に遠ざかるように動かされる場合、調整ベクトルは、z方向において最も大きい成分を有する。身体部分がx-y平面内において主に動かされる場合、調整ベクトルは、x-y平面内において1つまたは複数の最も大きい成分を有する。

【0097】

[00134]1つの手法では、調整ベクトルは、最良マッチ関節からピクセルに向かって伸びる。図13Aの調整ベクトル1328が一例である。ベクトルは、関節j1から深度センサーのオーバーラップ深度ピクセルp2に向かう方向に伸びる。この例では、同じ関節j1が最良マッチ関節であると見なされるが、異なる最良マッチ関節が、深度センサーの

50

非オーバーラップ深度ピクセル、および深度センサーのオーバーラップ深度ピクセルと関連することができる。別の手法では、調整ベクトルは、モデルのオーバーラップされる深度ピクセルから深度センサーのオーバーラップ深度ピクセルに向かって伸びる。例えば、図13Aの調整ベクトル1326は、モデルのオーバーラップ深度ピクセルp1から深度センサーのオーバーラップ深度ピクセルp2に向かう方向に、基本的には深度軸に沿って伸びる。このケースでは、ベクトル1326は、関節j1に作用するベクトル1330まで並進させることができる。

【0098】

[00135] 1つの手法では、最良マッチングアトラクトポイントの並進には中間ステップが存在し、それは、そのアトラクトポイントによって影響を受けるすべての関節を並進させる。ピクセルのベクトル/デルタを、ある調整/調節された力で、関節に適用することができる。関節に影響を与える多くのピクセルは、あまりにも大きなデルタをもたらすことがあるので、関節にすべてのデルタを完全に適用すると、その関節を飛び去らせる可能性があり、通常は、デルタの平均が役に立つ。

10

【0099】

[00136] 1つの手法では、ベクトル1326とベクトル1330は、同じ方向にあり、同じ大きさを有する。しかし、通常はピクセルよりもはるかに少数の関節しか存在しないので、このようなことはめったにない。そのため、2つのオーバーラップピクセルと関節が単一の直線内に揃うことは、一般的ではない。調整ベクトルは、例えば、ベクトル1328とベクトル1330の間の補間に基づくこともできる。

20

【0100】

[00137] 1つの手法では、調整ベクトルは、モデルを引かない。代わりに、調整ベクトルは、アトラクトポイントを引き、今度は、アトラクトポイントが、影響を受ける関節を引く。その後、関節上のモデルが最も意味をなす場所に、モデルが配置される。例えば、基部指セグメントは、手のひらと指を接続する関節のところに基部があり、基部指セグメントが中間部指セグメントの関節の方を向くように、配置することができる。

【0101】

[00138] オーバーラップするカメラ深度ピクセルとモデル深度ピクセルの間の調整ベクトルに対して、そのベクトルを使用して動かすのに適切な関節を見つけ、その関節への並進を累積し、その後、それを平均することができる。これは、z方向調整だけを含むことができる。別のケースでは、深度ピクセルと関節の間に調整ベクトルが存在し、それも累積して、平均することができる。これは、x、y、およびz方向調整を含むことができる。これらの調整は、必要に応じて組み合わせることができる。

30

【0102】

[00139] したがって、深度センサーの非オーバーラップ深度ピクセルおよびモデルの非オーバーラップ深度ピクセルに対して、選択された関節をそれがあるべき場所に動かすために、ピクセルが寄与する必要がある力または調整を計算することができる。

【0103】

[00140] 深度センサーのオーバーラップ深度ピクセルに対して、深度センサーの非オーバーラップ深度ピクセルおよびモデルの非オーバーラップ深度ピクセルと同様に、ピクセルから関節までの距離から計算されるスコアに基づいて、動かすのに最良の関節を見つけることによって、調整ベクトルを決定することができる。モデルの非オーバーラップ深度ピクセルがオーバーラップされる位置に関節を動かす傾向にある調整を計算して、次の反復において、それらが、深度センサーのオーバーラップ深度ピクセルになるようにすることもできる。しかし、この計算は、モデルのすべての非オーバーラップ深度ピクセルにわたって最適化され、それは、関節における対立調整 (conflicting adjustment) を課すことができる。例えば、モデルの腕は、センサーデータによって示される腕の幅よりも広いことがある。この場合、最適な解決策は、身体部分の少なくとも1つとしての腕の外周の周りのモデルの非オーバーラップ深度ピクセルの数を等しくする傾向をもつことができる。

40

50

【0104】

[00141] 1つの手法では、調整ベクトルは、モデルの非オーバーラップ深度ピクセルからセンサーデータの最も近い深度ピクセルに向かって伸びる。図13Aの調整ベクトル1322は、ピクセルp2がモデルの非オーバーラップ深度ピクセルp4に最も近いことを仮定した、一例である。ベクトルは、p4からp2に向かう方向に伸びる。別の手法では、調整ベクトルは、センサーデータの最も近い深度ピクセルの平面1312に向かってなど、p4から深度軸に沿って伸び、p4と平面1312の間の距離に基づいた大きさを有する。図13Aの調整ベクトル1324が、一例である。j1が動かすのに最良の関節であると仮定される場合、ベクトル1322またはベクトル1324は、関節j1に作用するベクトル（図示されず）まで並進させることができる。

10

【0105】

[00142] 図7Bにおいて、ステップ778は、説明したように、調整ベクトルを関節に適用する。このステップへの入力は、関節データ746である。図12Eおよび図12Fが、一例を提供する。その後、ステップ780における衝突制約、ステップ782における距離制約、およびステップ784における角度制約など、1つまたは複数の制約を、先に説明したように適用することができる。ステップ786は、制約ベースの調整ベクトルを決定し、対応する制約ベースの調整ベクトルデータ788を提供する。制約ベースの調整ベクトルには、タイプに基づいて、異なる重みを割り当てることができる。例えば、衝突制約は、より高い重みを受け取ることができ、距離制約は、次に高い重みを受け取ることができ、角度制約は、次に高い重みを受け取ることができる。

20

【0106】

[00143] ステップ790は、制約ベースの調整ベクトルをモデルの関節に適用する。このステップへの入力、関節データ746である。判断ステップ791は、例えば、一定の数のパスが実行されたとき、または関節の動きが一定の閾値を下回ったときといった、停止基準が満たされているかどうかを判定する。判断ステップ791が偽である場合、ステップ780～ステップ790が繰り返される。判断ステップ791が真である場合、ステップ792に進む。したがって、すべての制約を同時に最も良く満たすように、ステップ780～ステップ790において適用される制約を複数回反復することができる。

【0107】

[00144] ステップ792は、ステップ780～ステップ790の調整に基づいて、モデルメッシュを更新する。判断ステップ794は、停止基準が満たされているかどうかを判定する。例えば、停止基準は、(a) 最大許容回数の反復が実行されたときに反復が停止するなど、実行された反復の回数、(b) 調整ベクトルの大きさが十分に小さく、閾値を下回ったときに反復が停止するなど、調整ベクトルの大きさに基づいた収束基準、(c) 最大許容回数のCPUサイクルが消費されたときに反復が停止するなど、実行された計算の量に基づいた基準のうちの1つまたは複数に基づくことができる。判断ステップ794において停止基準が満たされない場合、次の反復がステップ796において実行され、ステップ762に進む。モデルは先行する反復からの調整された状態にあるので、モデルに対してステップ764のラスタライズが実行される。図12Gが、モデルの調整された状態の一例を提供する。判断ステップ794が真である場合、反復がステップ798において終了する。この時点で、モデルは、最適に調整されており、先に説明した姿勢またはジェスチャー推定ライブラリなど、より高いレベルのプロセス／アプリケーションによって使用することができる。1つの手法では、モデルは、深度データの各フレームに対して最適に調整され、制御入力としてアプリケーションに提供される。

30

40

【0108】

[00145] 一度に1つまたは複数の深度フレームを処理できることに留意されたい。バッチ処理モードが、一度に複数の深度フレームを処理し、複数のフレームからの情報を一緒に使用することを可能にする。最初に各フレームに対して調整ベクトルを決定することができ、その後、平均された調整ベクトルがフレームのバッチに適用されるように、調整ベクトルをフレームのバッチにわたって平均すること、または他の方法で組み合わせること

50

ができる。

【0109】

[00146] 1つの手法では、図7Bは、モデルの身体部分の変換（回転および並進）を変更するが、相対的な比率は変更しない。別の手法では、相対的な比率が変更される。

[00147] 図8Aは、ユーザーの手の深度ピクセルを含む、深度ピクセルからなるフレーム800を示している。フレームは、図7Aの深度フレームデータ702として使用することができる。フレームは、ユーザーの片手もしくは両手、または複数のユーザーの1つもしくは複数の手を含むことができる。1人または複数人のユーザーの他の身体部分も、同様に含むことができる。基本的に、フレームは、深度センサーの視野内の物体を含む。

【0110】

[00148] 図8Bは、背景ピクセルを除去した、図8Aのユーザーの手の深度ピクセル810である。説明したように、図7Aの背景除去ステップ704は、（図8Aのフレーム800など）深度フレームデータ702を受け取り、背景ピクセルを除去する。

【0111】

[00149] 図9Aは、図8Bのユーザーの手の深度ピクセルからなる末端部指セグメントおよび末端部親指セグメントについての、図7Aの見本ステップ712からの確率値を示している。これは、確率値の2Dにおける簡略化グラフィカル表示である。言及したように、各ピクセルは、モデルの各身体部分に対して、ピクセルがその身体部分の一部である尤度を示す異なる確率値（Pr1、Pr2、Pr3、...）を有する。この図は、5つの身体部分についての結果を一緒に示している。太い実線によって輪郭が描かれた影なしピクセルの各エリアは、その局所的な身体部分の一部である確率が最も高いピクセルを表す。点パターンを有するピクセルは、その局所的な身体部分の一部である確率がより低く、斜線パターンを有するピクセルは、その局所的な身体部分の一部である確率が最も低い。さらに、各三角形は、身体部分の深度ピクセルの重心を表す。身体部分は、第1の指または小指930と、第2の指932と、第3の指934と、第4の指または人差し指936と、親指938とを含む。

【0112】

[00150] 例えば、第3の指934の場合、エリア910内のピクセルは、末端部指セグメントに割り当てられ、関連する重心912が決定される。第4の指936の場合、エリア906内のピクセルは、末端部指セグメントに割り当てられ、関連する重心908が決定される。親指938の場合、エリア902内のピクセルは、末端部親指セグメントに割り当てられ、関連する重心904が決定される。

【0113】

[00151] 図9A～図9Eでは、影なしピクセル、点パターンピクセル、および斜線パターンピクセルが、一貫して使用される。

[00152] 図9Bは、図8Bのユーザーの手の深度ピクセルからなる中間部指セグメントおよび基部親指セグメントについての、図7Aのステップ721において生成される重心を示している。例えば、親指の場合、エリア914内のピクセルは、基部親指セグメントに割り当てられ、基部親指セグメントの重心916が決定される。

【0114】

[00153] 図9Cは、図8Bのユーザーの手の深度ピクセルからなる基部指セグメントについての、図7Aのステップ721において生成される重心を示している。

[00154] 図9Dは、図8Bのユーザーの手の深度ピクセルからなる手のひら940についての、図7Aのステップ721において生成される重心を示している。

【0115】

[00155] 図9Eは、図8Bのユーザーの手の深度ピクセルからなる腕942についての、図7Aのステップ721において生成される重心を示している。

[00156] 図10は、図7Aのステップ721において、ユーザーの手のテストデータから生成された重心を示している。テストデータは、一般にノイズを示す。各画像において、より暗い領域は、より高い確率に対応する。さらに、四角形は、重心を表す。例えば、

10

20

30

40

50

画像１０００は、末端部指および親指セグメントについてのものである。暗い領域が、重心の四角形アイコンによって部分的に覆われている。画像１００２は、中間部指セグメントについてのものであり、画像１００４は、基部指セグメントについてのものであり、画像１００６は、基部親指セグメントについてのものであり、画像１００８は、手のひらについてのものであり、画像１０１０は、前腕の下部についてのものであり、画像１０１２は、前腕の上部についてのものである。

【０１１６】

[00157]図１１は、重心対アトラクトポイントマッチングについての、図７Ａのステップ７３４の詳細を示している。言及したように、各三角形は、太い実線で輪郭が描かれたそれぞれの身体部分の重心を表す。モデル５００に加えて、深度センサーからのピクセルデータ９００が示されている。一例として、末端部指セグメント９１０を表すピクセルの場合、重心９１２は、モデル５００の対応する末端部指セグメント５３２のアトラクトポイント５３１にマッチさせられる。同様に、末端部指セグメント９０６を表すピクセルの場合、重心９０８は、モデル５００の対応する末端部指セグメント５４２のアトラクトポイント５４１にマッチさせられる。末端部親指セグメント９０２を表すピクセルの場合、重心９０４は、モデル５００の対応する末端部親指セグメント５５２のアトラクトポイント５５８にマッチさせられる。基部親指セグメント９１４を表すピクセルの場合、重心９１６は、モデル５００の対応する基部親指セグメント５５４のアトラクトポイント５６２にマッチさせられる。腕９１８を表すピクセルの場合、重心９２０は、モデル５００の対応する腕５０４のアトラクトポイント５０５にマッチさせられる。

【０１１７】

[00158]各重心は、それぞれのアトラクトポイントにマッチさせることができる。場合によっては、必ずしもすべての重心を識別すること、および／またはそれぞれのアトラクトポイントにマッチさせることができず、その場合は、モデルは、利用可能な情報に基づいて調整される。利用可能な情報が不十分である場合、モデルを調整しないという決定を、フレームごとに下すことができる。

【０１１８】

[00159]図１２Ａは、重心へのアトラクトポイントの剛体変換についての、図７Ａのステップ７３６の詳細を示している。モデル１２００は、すべてのアトラクトポイントについて、深度ピクセルデータの重心に対するモデルのアトラクトポイントの最良マッチを提供するために変換される。この例では、ピクセルデータに対するモデルのマッチは、親指を除いて、かなり良好である。特に、親指の基部セグメントのアトラクトポイント５６２は、関連する重心９１６からいくぶん離れて配置されており、親指の末端部セグメントのアトラクトポイント５５８は、関連する重心９０４からより大きく離れて配置されている。

【０１１９】

[00160]図１２Ｂは、アトラクトポイントを重心のより近くに移動させるための、図７Ａのステップ７５２の詳細を示している。ここでは、アトラクトポイント５６２を関連する重心９１６のより近くに移動させ、アトラクトポイント５５８を関連する重心９０４のより近くに移動させるように、モデル１２１０が改善される。

【０１２０】

[00161]図１２Ｃは、モデルをラスタ化して、モデルの深度ピクセル１２２０を提供するための、図７Ｂのステップ７６４の詳細を示している。これらのピクセルは、比較平面内などにおいては、末端部セグメントのピクセル９００の上に見られる。モデルのピクセルは、影なしであり、一方、末端部セグメントのピクセル９００は、斜線パターンを有する。基部親指セグメント５５４についてのモデルのピクセルの上よりも下に、深度センサーのピクセルがより多く存在するので、基部親指セグメントの位置は最適ではない。また、末端部親指セグメント５５２についてのモデルのピクセルの下よりも上に、深度センサーのピクセルがより多く存在するので、末端部親指セグメントの位置も最適ではない。同様に、腕および手のひらの外周の周り、左側よりも右側に、深度センサーのピクセルが

より多く存在するので、腕 5 0 4 および手のひら 5 0 2 の位置も最適ではない。ラスタ化を使用する改善は、モデルと深度センサーとのピクセルデータを比較することを可能にすることによって、これらの非最適な状態を識別することができる。

【0 1 2 1】

[00162]図 1 2 D は、親指エリアについて、モデルの深度ピクセルをセンサーの深度ピクセルと比較するための、図 7 B のステップ 7 6 6 ~ 7 7 2 の詳細を示している。細部を示すために、図 1 2 C の親指エリアを独立させている。しかし、モデルのすべての身体部分について、比較を実行することができる。影なしのピクセルからなるエリア 1 2 3 0 は、ステップ 7 6 8 の深度センサーのオーバーラップ深度ピクセルを表す。斜線パターンを有するピクセルからなるエリア 1 2 4 0 は、ステップ 7 7 0 の深度センサーの非オーバーラップ深度ピクセルを表す。点パターンを有するピクセルからなるエリア 1 2 2 0 は、ステップ 7 7 2 のモデルの非オーバーラップ深度ピクセルを表す。

10

【0 1 2 2】

[00163]図 1 2 E は、図 1 2 D と一致する、ピクセルベースの調整ベクトルに基づいたモデルの関節の調整を示す、図 7 B のステップ 7 7 8 の詳細を示している。関節位置 5 6 4、5 6 0、および 5 5 6 が、それぞれ、位置 1 2 6 2、1 2 5 8、および 1 2 5 4 に調整される。結果として、アトラクトポイントも、同様に、位置 5 6 2 および 5 5 8 から、それぞれ、位置 1 2 6 0 および 1 2 5 6 に調整される。

【0 1 2 3】

[00164]図 1 2 F は、図 1 2 E において使用されるピクセルベースの調整ベクトルを示している。関節位置 5 6 4、5 6 0、および 5 5 6 を、それぞれ、位置 1 2 6 2、1 2 5 8、および 1 2 5 4 に調整するために、それぞれ、調整ベクトル 1 2 7 8、1 2 7 4、および 1 2 7 0 が使用される。調整ベクトルは、3 D とすることができるが、この例では簡潔にするために、2 D として示されている。

20

【0 1 2 4】

[00165]図 1 2 G は、図 1 2 E と一致する、モデルの更新を示す、図 7 B のステップ 7 9 2 の詳細を示している。図 1 2 E では、モデルの末端部親指セグメント 5 5 2 が、セグメント 1 2 5 0 になるように調整され、モデルの基部親指セグメント 5 5 2 が、セグメント 1 2 5 2 になるように調整される。この場合、親指は、手のひらに対しておおよそ下方に移動させられるとともに、回転させられる。これは、深度センサーの深度ピクセルとより緊密にマッチする、モデルの新しい状態である。

30

【0 1 2 5】

[00166]図 1 3 A は、比較平面 1 3 1 0 が深度センサー 1 3 0 0 の深度軸 1 3 0 2 を横断または直交する、モデルの深度ピクセルをセンサーの深度ピクセルと比較するための、図 7 B のステップ 7 6 8、7 7 0、および 7 7 2 のさらなる詳細を示している。軸 1 3 0 4、1 3 0 6、および 1 3 0 8 も、(表記「|」によって示されるように) 深度軸 1 3 0 2 と平行である。比較平面 1 3 1 0 は、ラスタ化モデルの 2 つの例示的なピクセル p 1 および p 4 を含み、一方、平面 1 3 1 2 は、深度センサーからの 2 つの例示的なピクセル p 2 および p 3 を含む。1 つの手法では、比較平面 1 3 1 0 および 1 2 1 2 は、x 軸および y 軸に平行に広がり、したがって、深度軸 1 3 0 2 を横断/直交する。別の手法では、比較平面 1 3 1 0 および 1 2 1 2 は、x 軸および y 軸から僅かにずれており、深度軸 1 3 0 2 に対して角度 α をなして伸びる、深度カメラからの視線 1 3 0 5 を横断する。

40

【0 1 2 6】

[00167]p 2 は、例えば軸 1 3 0 6 に沿って、p 1 とオーバーラップするので、深度センサーのオーバーラップ深度ピクセルである(図 7 B、ステップ 7 6 8)。p 1 も、モデルの対応するオーバーラップ深度ピクセルと考えることができる。j 1 は、モデルの例示的な関節である。1 つの手法では、調整ベクトル 1 3 2 8 が、最良マッチ関節 j 1 から p 2 に向かう方向に伸びる。別の手法では、調整ベクトル 1 3 2 6 が、p 1 から p 2 に向かう方向に伸びる。さらに、ベクトル 1 3 2 6 は、関節 j 1 に作用するベクトル 1 3 3 0 まですら並進させることができる。

50

【0127】

[00168] p 3 は、例えば軸 1 3 0 8 に沿って、モデルのピクセルとオーバーラップしないので、深度センサーの非オーバーラップ深度ピクセルである（図 7 B、ステップ 7 7 0）。図 7 B のステップ 7 7 8 に関連して先に説明したように、p 3 についての例示的な調整ベクトル 1 3 2 0 は、関節 j 1 から p 3 に向かう方向に伸びる。

【0128】

[00169] p 4 は、例えば軸 1 3 0 4 に沿って、深度センサーデータのピクセルとオーバーラップしないので、モデルの非オーバーラップ深度ピクセルである（図 7 B、ステップ 7 7 2）。対応する調整ベクトル 1 3 2 2 は、p 4 から、この例では p 2 であると仮定される、センサーデータの最も近い深度ピクセルに向かって伸びる。別の手法では、p 4 についての調整ベクトル 1 3 2 4 は、平面 1 3 1 2 のセンサーデータの最も近い深度ピクセルなどに向かって、p 4 から深度軸 1 3 0 2 / 1 3 0 4 に沿って伸びる。ベクトル 1 3 2 2 またはベクトル 1 3 2 4 は、j 1 が p 4 のために動かすのに最良な関節であると仮定した場合には、関節 j 1 に作用するベクトル（図示されず）まで並進させることができる。

【0129】

[00170] 図 1 3 B は、図 1 3 A の比較平面 1 3 1 0 のさらなる詳細を示している。p 1 および p 2 は、オーバーラップピクセルであり、深度センサーの p 3 は、モデルのピクセルとオーバーラップしておらず、モデルの p 4 は、深度センサーのピクセルとオーバーラップしていない。|| x および || y は、それぞれ、x 軸および y 軸と平行な軸である。

【0130】

[00171] 図 1 4 は、距離制約、衝突制約、およびピクセル比較を使用して、撮像システムにおいて改善された忠実度で有関節身体部分の姿勢を検出するための方法を示している。図 1 4 の方法は、図 7 A および図 7 B の方法と併せて実行することができる。方法は、一般に、初期姿勢推定を獲得するステップと、それに続いて、距離制約、衝突制約、角度制約、およびラスタライズモデルを使用するピクセル比較のうちの 1 つまたは複数に基づいて、姿勢推定を改善するステップとを含む。1 つの手法では、初期姿勢推定は、図 7 A の方法によって提供することができる。1 つの手法では、ラスタライズモデルを使用するピクセル比較は、図 7 B の方法のステップ 7 6 2 ~ ステップ 7 9 0 によって提供することができる。姿勢推定の改善は、全体的または部分的に、順次的または並列的に実行することができる。

【0131】

[00172] その中に多数の区別可能な特徴点を有する既知のトポロジーを有する有関節モデルのパラメトリック幾何モデル（例えばメッシュ）を有することを仮定することができる。特徴検出器は、モデル内の点に対応する深度センサーデータから（関節および他の参照など）できるだけ多くの区別可能な特徴点を見つける。既知の有関節モデルのメッシュは、較正され、動的に調整される。この方法は、これを達成するために、有関節モデルの寸法についてのローカルおよびグローバル情報の複数のソースを組み合わせて、それらに重み付けを行う。

【0132】

[00173] ステップ 1 4 0 0 は、例えば、図 6 のステップ 6 0 0 と同様に、深度センサーからユーザーの深度ピクセルを獲得するステップを含む。ステップ 1 4 0 2 は、例えば、図 6 のステップ 7 0 0 と同様に、深度センサーの深度ピクセルに対するモデルの初期マッチを決定するステップを含む。図 6 のステップ 7 5 0 も、同様に含むことができる。これは、先に説明したような、いくつかのタイプの代表的なアトラクトポイント対重心マッチング、または他の任意の技法を含むことができる。さらに、図 7 A のステップ 7 3 6 に関連して説明したような変換を使用することができる。これは、例えば、3 D スケールを用いる 3 D 点群の剛体変換など、3 D 点群の制約付きアフィンレジストレーションを含むことができる。モデルと深度センサー観測との間の特徴点対応を与えると、単一のスケーリングまたは各次元（長さ、幅、および高さ）で別々のスケーリングを用いて、最良の剛体変換を計算することによって、有関節物体の各身体部分のための個別のスケーリングパラ

メータ、またはすべての身体部分のためのグローバルなスケーリングパラメータを直ちに（おそらくは先行フレームにおいて推定された身体比率を考慮して）計算することができる。

【0133】

[00174]ステップ1404は、距離制約を使用して、モデルの初期マッチを改善する。ステップ1406は、各有关節身体部分について距離制約が違反された程度を決定するステップを含む。ステップ1404およびステップ1406に関する例は、図16A1～図16Cおよび図17Aに関連して提供される。

【0134】

[00175]ステップ1408は、衝突制約を使用して、モデルの初期マッチを改善する。ステップ1410は、衝突制約を実施することを決定するステップを含む。例えば、衝突制約の実施をもたらす特定の姿勢にモデルがあると決定された場合に、この決定を下すことができる。一例は、モデルの少なくとも2つの隣接する指セグメントが、例えば互いに実質的に平行など、特定の相対的な位置にあると決定された場合である。この条件は、例えば、数度の閾値角度よりも小さい角度をなして伸びる2つの隣接する指セグメントの縦軸によって、検出することができる。ステップ1412は、各有关節身体部分について衝突制約が違反された程度を決定するステップを含む。ステップ1408～ステップ1412に関する例は、図18Aおよび図18Bに関連して提供される。

【0135】

[00176]ステップ1414は、ラスタ化モデルを使用してマッチングを改善する。ステップ1416は、例えば、図7Bのステップ764と同様に、モデルをラスタ化して、モデルの深度ピクセルを提供する。ステップ1418は、例えば、図7Bのステップ766と同様に、比較平面内において、モデルの深度ピクセルを深度センサーの深度ピクセルと比較する。ステップ1420は、例えば、図7Bのステップ770と同様に、モデルの深度ピクセルとオーバーラップしない、深度センサーの非オーバーラップ深度ピクセルを識別する。ステップ1422は、例えば、図7Bのステップ772と同様に、深度センサーの深度ピクセルとオーバーラップしない、モデルの非オーバーラップ深度ピクセルを識別する。ステップ1414～ステップ1422に関する例は、図16Dおよび図16Eに関連して提供される。

【0136】

[00177]ステップ1424は、様々な制約を満たそうと試みて、モデルを調整する。判断ステップ1426において、停止基準が満たされている場合、プロセスは、ステップ1428において停止する。判断ステップ1426が偽の場合、プロセスの別の反復が実行される。停止基準が満たされるまで、1つまたは複数の反復において、方法のステップのいずれかを反復的に実行することができる。一般に、すべての制約を適度に満たすために、複数の反復が使用される。さらに、モデル調整は、各制約についてのすべての利用可能な情報に基づくことができる。ステップ1426における調整は、各制約の寄与を組み合わせ、重み付けを行うことができる。代替として、例えば、ステップ1406、1412、1420、および1422の各々の後、別々に各制約に基づいて、モデルを調整することもできる。

【0137】

[00178]距離制約を満たそうと試みるモデルの調整は、例えば、指もしくは親指セグメントまたは手のひらもしくは手首の長さまたは幅を増やしたまたは減らすことによって、有関節身体部分の少なくとも1つについての寸法を調整することを含むことができる。衝突制約を満たそうと試みるモデルの調整は、例えば、指もしくは親指セグメントまたは手のひらもしくは手首の長さまたは幅を減らすことによって、衝突に係わる有関節身体部分の少なくとも1つについての寸法を調整することを含むことができる。場合によっては、処理時間もしくは他の限界、または他の制約のために、必ずしもすべての制約を完全に満たすことはできず、その場合は、衝突制約を満たすために、最善の努力が行われる。

【0138】

[00179] 深度センサーの非オーバーラップ深度ピクセルに基づいたモデルの調整は、深度センサーの非オーバーラップ深度ピクセルの方向に、モデル部分の寸法（例えば、高さ、幅）を増やすことを含むことができる。与えられた方向に寸法を増やす程度は、例えば、与えられた方向におけるオーバーラップピクセルの数に基づくなど、オーバーラップの程度に比例することができる。モデルの非オーバーラップ深度ピクセルに基づいたモデルの調整は、深度センサーの非オーバーラップ深度ピクセルから離れる方向にモデル部分の寸法を減らすことを含むことができる。最適な解決策は、モデルの非オーバーラップ深度ピクセルに隣接する、身体部分の外周の周りのモデルの非オーバーラップ深度ピクセルの数を等しくする傾向をもつことができる。与えられた方向に寸法を減らす程度は、例えば、与えられた方向におけるオーバーラップピクセルの数に基づくなど、オーバーラップの程度に比例することができる。

10

【0139】

[00180] 図15は、図14の方法とともに使用するための、ユーザーの手の深度ピクセル1500の一例を示している。この例は、図16A1～図16Eおよび図17Aにおいて使用される。

【0140】

[00181] 図16A1は、距離制約が実施されず、モデルが望ましいものよりも小さい場合の、図15の深度ピクセルに対する図5のモデルのマッチングを示している。このマッチングは、深度センサーデータのそれぞれの部分に最適にマッチさせるために、モデルの各部分が、隣接部分に取り付けられることなく、独立して浮かんでいることを可能にする

20

【0141】

[00182] 距離制約は、モデルの部分の間の距離、およびモデルの部分の相対的な配置に関する。例えば、距離制約は、指セグメントが端部と端部とを接して配置され、指セグメントの一方の端部が隣接する指セグメントの近接端から伸びることを要求することができる。さらに、深度センサーは、基部指セグメントが手のひらの指定された関節から伸びることを要求することができる。別の距離制約は、手のひらが手首から伸びることを要求することができる。身体部分の関節の位置に基づいて、距離制約が違反されたかどうか、および違反の程度を決定することができる。この図および他の図では、例示的な関節が、菱形によって表されている。1つの手法では、距離制約は、モデルの1つの部分の関節がモデルの隣接する部分の関節と実質的に（例えば閾値距離以内で）一致することを要求する。例えば、モデル部分516の下部の関節は、手のひらの関連する最も左の関節と一致すべきであり、モデル部分516の上部の関節は、身体部分514の下部の関節と一致すべきであり、モデル部分5146の上部の関節は、身体部分512の下部の関節と一致すべきである。関節の間のギャップは、距離制約の違反を表す。より大きなギャップまたは距離は、違反の程度がより大きいことを表す。

30

40

【0142】

[00183] この例でも、モデルは、深度ピクセル1500からなる手の表現と比較して、全体的により小さい。

[00184] 距離制約は、関節などモデル内の特徴点の間の直線によって表すことができる。各直線は、圧縮または伸張状態にあると見なすことができる。センサーが検出したものよりもモデルが小さい場合、関節を互いに引き離そうとする傾向にある張力が存在する。センサーが検出したものよりもモデルが大きい場合、関節を互いの方に押しやろうとする傾向にある圧縮力が存在する。モデルを調整するために、有関節モデルの各部分に関連する個々の張力および／または圧縮力情報が使用される。

50

【0143】

[00185]図16A2は、親関節および子関節が示された、図16A1の代替である。別の手法では、例えば、末端部親指部分と基部親指部分の間に1つの関節560が存在し、距離制約は、関節560と関節564などの親関節との間の距離である。関節564の親関節は、関節503とすることができる。例えば、関節560は、末端部親指セグメント552の基部端に存在することができ、関節564は、基部親指セグメント554の基部端に存在することができる。モデルのこの手法では、メッシュが、子関節に達するほど十分に長くはないとしても、子関節の方を指して親関節に配置される。親または子関節は、それぞれ、親または子身体部分に関連付けることができる。1つの可能な手法では、子身体部分は、手のひらから伸びる指など、より大きな身体部分から伸びるより小さな身体部分とすることができる。または、子身体部分は、腕から指先の方向など、指定された方向に別の身体部分から伸びる身体部分とすることができる。例えば、基部親指部分554は、末端部親指部分552の親とすることができる。同様に、基部指セグメント546の基部端の関節553は、基部指セグメントと手のひら502の間にあり、手のひら502の関節503から距離dにある、スケルタル関節を表す。中間部指セグメント544の基部端の関節551は、中間部指セグメント544と基部指セグメント546の間にあるスケルタル関節を表す。末端部指セグメント542の基部端の関節549は、末端部指セグメント542と中間部指セグメント544の間にあるスケルタル関節を表す。

10

【0144】

[00186]図16B1は、人差し指についての図16A1の画像の詳細を示している。人差し指540では、深度センサーのピクセルは、太い実線によって示されるように、末端部指セグメントピクセル領域1620と、中間部指セグメントピクセル領域1630と、基部指セグメントピクセル領域1640とにグループ化される。指の末端部、中間部、および基部モデル部分は、それぞれ、542、544、および546である。モデル部分542は、上部関節（菱形547）と、下部関節549と、（ピクセル領域1620の重心である三角形1622に揃えられる）代表的なアトラクトポイント541と、幅wdm（幅、末端部、モデル）と、高さhdm（高さ、末端部、モデル）と、下にあるピクセル領域の左端からの間隔dl（末端部、左）と、下にあるピクセル領域の右端からの間隔dr（末端部、右）とを含む。末端部指セグメントに割り当てられた下にあるピクセル領域は、幅wdおよび高さHDP（高さ、末端部、ピクセル）を有する。

20

30

【0145】

[00187]モデル部分544は、上部関節1632と、下部関節551と、（ピクセル領域1620の重心である三角形1634に揃えられる）代表的なアトラクトポイント543と、幅wmm（幅、中間部、モデル）と、高さhmm（高さ、中間部、モデル）と、下にあるピクセル領域の左端からの間隔ml（中間部、左）と、下にあるピクセル領域の右端からの間隔mr（中間部、右）とを含む。中間部指セグメントに割り当てられた下にあるピクセル領域は、幅wdおよび高さHMP（高さ、中間部、ピクセル）を有する。

【0146】

[00188]モデル部分546は、上部関節1642と、下部関節553と、（ピクセル領域1640の重心である三角形1644に揃えられる）代表的なアトラクトポイント545と、幅wpm（幅、基部、モデル）と、高さhpm（高さ、基部、モデル）と、下にあるピクセル領域の左端からの間隔pl（基部、左）と、下にあるピクセル領域の右端からの間隔pr（基部、右）とを含む。基部指セグメントに割り当てられた下にあるピクセル領域は、幅wdおよび高さHPPを有する。

40

【0147】

[00189]手のひら502は、モデル部分546の基部端がそこから伸びる、関節1650を有する。

[00190]寸法は、モデルの大域座標系において定義することができ、または各身体部分は、独自の座標系を有することができる。高さおよび幅などの用語は、任意の方向の寸法を表すことができる。

50

【0148】

[00191] 関節 549 と関節 544 の間の距離 $m d$ (中間部から末端部) は、制約が、距離がゼロであること、またはゼロではないが一定の許容値よりも小さいことである場合、違反された距離制約を表す。制約が身体部分 542 および 544 の両方によって違反されたと言うことができる。関節 551 と関節 1642 の間の距離 $p m$ (基部から中間部) は、今回は身体部分 544 および 546 によって違反された、別の距離制約を表す。関節 553 と関節 1650 の間の距離 $p p$ (基部から基部) は、今回は身体部分 546 および 502 によって違反された、別の距離制約を表す。関節の間の距離制約は、3D 距離とすることができることに留意されたい。本例は、簡潔にするため、2D である。

【0149】

10

[00192] 距離 $d t$ (末端部から上端) は、ピクセル領域 1620 の上端からの関節 547 の隔たりである。図 16D および図 16E を参照されたい。

[00193] この例では、次に説明するように、指の方向に伸ばすことによって、モデル部分を調整すべきである。

【0150】

[00194] 図 16B2 は、親関節および子関節が示された、図 16B1 の代替を示す図である。この場合、距離制約は、(a) $h d m$ 、(b) $h m m + m d$ 、(c) $h p m + p m$ 、および (d) d によって表すことができる。この場合、関節 1632、1642、および 1650 は使用されない。距離 d は、他の距離とともにスケールすべきでなく、したがって、断線によって示されている。角度制約などの他の制約も、関節 503 と関節 553 の間の直線などに適用することができる。

20

【0151】

[00195] 図 16C は、距離制約を実施した後の、図 16B1 のモデルに対する調整を示している。指セグメントが $H D P + H M P + H P P$ の全長を有するように、各指セグメントの長さを伸ばして、距離制約を実施することができる。この長さは、直線である軸に沿うことができ、または 3D においてはカーブすることができる。距離制約を実施した結果は、関節 549 と関節 1632、関節 551 と関節 1642、および関節 553 と関節 1650 が、実質的に (許容範囲内で) 一致することである。身体部分 542 の長さは、 $h d m$ から $h d m'$ に増やされ、身体部分 544 の長さは、 $h m m$ から $h m m'$ に増やされ、身体部分 546 の長さは、 $h p m$ から $h p m'$ に増やされる。モデル調整の結果として、重心および代表的なアトラクトポイントの相対的な位置は移動できる。1つの手法では、各身体部分は、長さが増やされる隣接する他の身体部分と比べた自らの長さに基づいて、また身体部分の間の距離に基づいて、長さが増やされる。他の例では、指のすべてのセグメントのうちの一部が、寸法を調整される。他の例では、手の他の部分が、寸法を調整される。他の例では、モデルは、手以外の別の身体部分についての、またはその挙動をモデル化できる別の有関節物体についてのものである。

30

【0152】

[00196] 親関節および子関節が使用される場合も (例えば図 16B2)、同様の結果が達成される。

[00197] 図 16D は、ラスタ化および深度センサーの深度ピクセルとの比較の後の、図 16C のモデルを示している。モデル部分 542、544、および 546 は、ピクセル領域として示されている。モデルのピクセルとオーバーラップしない深度センサーの深度ピクセルを識別することができる。具体的には、領域 1650、1651、および 1652 内の深度ピクセルは、モデル部分 542 に対して非オーバーラップであり、領域 1660 および 1662 内の深度ピクセルは、モデル部分 544 に対して非オーバーラップであり、領域 1670 および 1672 内の深度ピクセルは、モデル部分 546 に対して非オーバーラップである。非オーバーラップピクセルは、点ボックスに囲われ、点パターンで示されている。この比較は、モデルの指セグメントが狭すぎることを示している。モデル部分 542 は、左に量 $d l$ だけ、右に量 $d r$ だけ、広げることができる。モデル部分 544 は、左に量 $m l$ だけ、右に量 $m r$ だけ、広げることができる。モデル部分 546 は、左に

40

50

量 p_l だけ、右に量 p_r だけ、広げることができる。モデル部分 542 の上端のオーバーラップ d_t に関して、これは、モデル部分 542 の長さのみを d_t だけ増やすことによって、または指のモデル部分の各々についての長さを増やすことによって、対処することができる。一般に、モデルを調整する距離は、長さメトリックまたはピクセルの数によって表すことができる。

【0153】

[00198] この例では、各モデル部分の右側および左側に一定のオーバーラップが存在する。モデル部分の外周に沿ってオーバーラップが変化する場合、オーバーラップに対して最も良く一致するように、モデル部分を調整することができる。1つの手法では、平均オーバーラップが決定され、モデル部分を調整するために使用される。

10

【0154】

[00199] 図 16E は、ピクセル比較に基づいた、図 16D のモデルに対する調整を示している。この例では、モデル部分 542 の上端のオーバーラップ d_t は、指のモデル部分の各々についての長さを増やすことによって対処されている。身体部分 542 の長さは、 h_{dm}' から h_{dm}'' に増やされ、身体部分 544 の長さは、 h_{mm}' から h_{mm}'' に増やされ、身体部分 546 の長さは、 h_{pm}' から h_{pm}'' に増やされる。同様に、モデル部分に隣接する横方向の非オーバーラップピクセルのため、身体部分 542 の幅は、 w_{dm} から w_{dm}' に増やされ、身体部分 544 の幅は、 w_{mm} から w_{mm}' に増やされ、身体部分 546 の幅は、 w_{pm} から w_{pm}' に増やされる。

【0155】

20

[00200] 親関節および子関節が使用される場合も（例えば図 16B2）、同様の結果が達成される。

[00201] 図 17A は、距離制約が実施されず、モデルが望ましいものよりも大きい場合の、図 15 の深度ピクセルに対する図 5 のモデルのマッチングの別の例を示している。モデルのピクセルは、明瞭にするために示されていない。

【0156】

[00202] 深度センサーのピクセルとオーバーラップしないモデルの深度ピクセルを識別することができる。具体的には、領域 1750 および 1752 内の深度ピクセルは、モデル部分 542 に対して非オーバーラップであり、領域 1760 および 1762 内の深度ピクセルは、モデル部分 544 に対して非オーバーラップであり、領域 1770 および 1772 内の深度ピクセルは、モデル部分 546 に対して非オーバーラップである。非オーバーラップピクセルは、点ボックスに囲われ、点パターンで示されている。この比較は、指セグメントが広すぎることを示している。モデル部分 542 は、左に量 d_l だけ、右に量 d_r だけ、狭めることができる。モデル部分 544 は、左に量 m_l だけ、右に量 m_r だけ、狭めることができる。モデル部分 546 は、左に量 p_l だけ、右に量 p_r だけ、狭めることができる。

30

【0157】

[00203] この例では、各モデル部分の右側および左側に一定または一様のオーバーラップが存在する。モデル部分の外周に沿ってオーバーラップが変化する場合、オーバーラップに対して最も良く一致するように、モデル部分を調整することができる。1つの手法では、平均オーバーラップが決定され、モデル部分を調整するために使用される。

40

【0158】

[00204] 与えられたモデル部分は、1つまたは複数のフレームにおいて、反対の調整を施され得ることに留意されたい。同様に、隣接するモデル部分は、1つまたは複数のフレームにおいて、反対の調整を施され得る。例えば、指セグメントは、1つのフレームでは、幅を広げられ、別のフレームでは、幅を狭められ得る。これは、例えば、調整の反復性、ノイズ、深度センサーの解像度、ユーザーの動き、計算限界、および他の要因によって引き起こされることがある。モデル部分は、例えば、反対または対立する距離制約調整、衝突制約調整、および／またはピクセル比較調整のせいで、1つのフレームにおいても、反対または対立する調整を施され得る。

50

【0159】

[00205]距離制約は、(a) md 、(b) pm 、(c) pp' によって表すことができる。

[00206]図17Bは、親関節および子関節が示された、図17Aの代替を示している。この場合、距離制約は、(a) hdm 、(b) $hmm-md$ 、(c) $hpm-pm-pp'$ 、および(d) d' によって表すことができる。この場合、関節1632、1642、および1650は使用されない。距離 d' は、他の距離とともにスケーリングすべきでなく、したがって、断線によって示されている。

【0160】

[00207]図18Aは、衝突を示す、図5のモデルの2つの隣接する指を示している。衝突制約は、モデルの異なる部分が同時に同じ位置に存在することを回避することである。例えば、この例は、モデルの隣接する指530および540を示しており、指の影つき部分1820が、同一場所に配置されている。この簡略化された例では、指530のモデル部分532、534、および536は、幅 w_{dm1} を有し、指540のモデル部分542、544、および546は、幅 w_{dm2} を有する。距離 pd （浸食深さ（*penetration depth*））は、例えば衝突制約の違反の程度など、衝突の程度を表す。中心縦軸が、各モデル部分内に、例えば、軸1800、1802、および1804が、それぞれ、モデル部分532、534、および536内に、また軸1810、1812、および1814が、それぞれ、モデル部分542、544、および546内に伸びている。

【0161】

[00208]軸1830は、指540の外縁に沿って伸びる、指530の浸食軸であると見なすことができ、軸1832は、指530の外縁に沿って伸びる、指540の浸食軸であると見なすことができる。この場合、 pd は、浸食軸1830と浸食軸1832の間の距離である。

【0162】

[00209]図18Bは、衝突制約を実施した後の、図18Aの2つの隣接する指を示している。衝突制約を満たすため、指530および540の一方または両方が調整される。ことによると、1つの解決策は、指どうしをより遠くに引き離すことである。しかし、距離制約が、手のひら内の指定された関節から指が伸びることを要求することがあり、そのため、この解決策は、好ましくない。別の解決策は、指の一方のみの幅を減らすことであるが、これも、最適である可能性は高くない。最適な解決策は、一般に、衝突する指または他の身体部分の両方の幅を減らすことである。1つの手法では、指530および540は、衝突を防止するのに十分な、例えば同じパーセンテージなど、同じ比率だけ幅を減らすことができる。この場合、指530の幅は、 w_{dm1} から w_{dm1}' に減らされ、指540の幅は、 w_{dm2} から w_{dm2}' に減らされる。

【0163】

[00210]モデルが衝突制約の実施に特に適した姿勢にあると判断された場合に、衝突制約を実施する決定を下すことができる。したがって、衝突制約は、姿勢が適しているがどうかに応じて、フレームに対して実施することができる。1つの可能な手法では、姿勢が適していない場合、そのフレームに対して衝突制約は実施されない。

【0164】

[00211]1つの手法では、初期モデルの改善は、距離制約を満たそうとする試みにおいてモデルを調整することと、その後、衝突制約を実施することを決定することを含む。衝突制約を実施することを決定したことに応じて、初期マッチにおいて、モデルの指セグメントなどの少なくとも2つの隣接するモデル部分によって衝突制約が違反された程度に関する評価を行うことができる。モデルは、モデルの少なくとも2つの隣接する指セグメントの幅を減らすことによって、モデルの少なくとも2つの隣接する指セグメントによって衝突制約が違反された程度に基づいて、衝突制約を満たそうとする試みにおいてモデルを調整することができる。衝突制約を実施することの決定は、モデルの少なくとも2つの隣接する指セグメントが指定された相対的な位置にあることを決定することを含むことが

10

20

30

40

50

できる。1つの手法では、モデルの少なくとも2つの隣接する指セグメントは、モデルの少なくとも2つの隣接する指セグメントのそれぞれの縦軸が、図18Aにおけるように、指定された角度閾値内にあり、実質的に平行である場合に、指定された相対的な位置にある。この角度閾値は、例えば、軸1830と軸1832の間の角度とすることができる。モデルの少なくとも2つの隣接する指セグメントによって衝突制約が違反された程度は、それぞれの縦軸の間の距離に基づくことができる。

【0165】

[00212]この例では、2つの隣接する指または指セグメントが、実質的に平行で、著しく衝突している場合、これは、指の幅が太すぎであり、減らすべきであることを示している。1つの手法では、各指セグメントは、円筒としてモデル化され、その場合、円筒の長さ軸が平行であるかどうかを判定することができる。

【0166】

[00213]本明細書の技術の上述の詳細な説明は、例示および説明の目的で提示された。それは、網羅的であること、または開示された形通りのものに技術を限定することを意図していない。上述の教示に照らして、多くの変更および変形が可能である。説明された実施形態は、技術の原理およびその実際的な適用を最も良く説明し、それによって、様々な実施形態において、企図される特定の用途に適するように様々な変更を施して、当業者が技術を最も良く利用することを可能にするように選択された。技術の範囲は、本明細書に添付された特許請求の範囲によって確定されることが意図されている。

10

【図1】

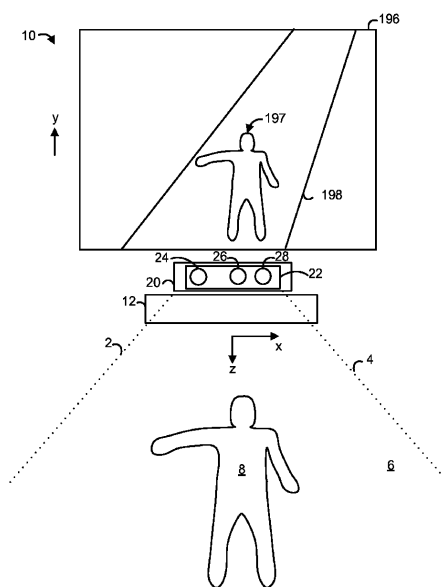
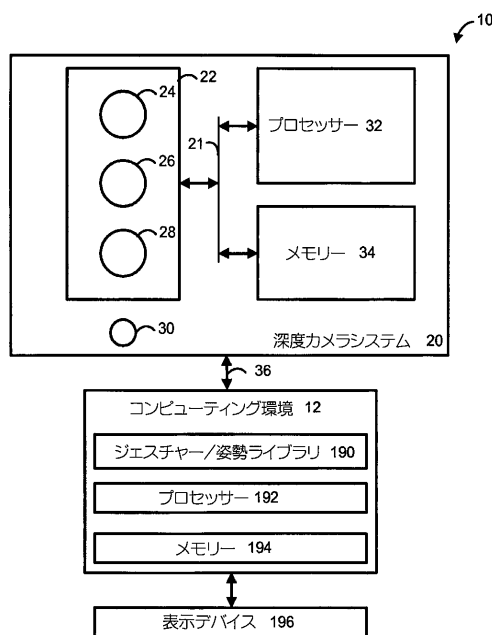
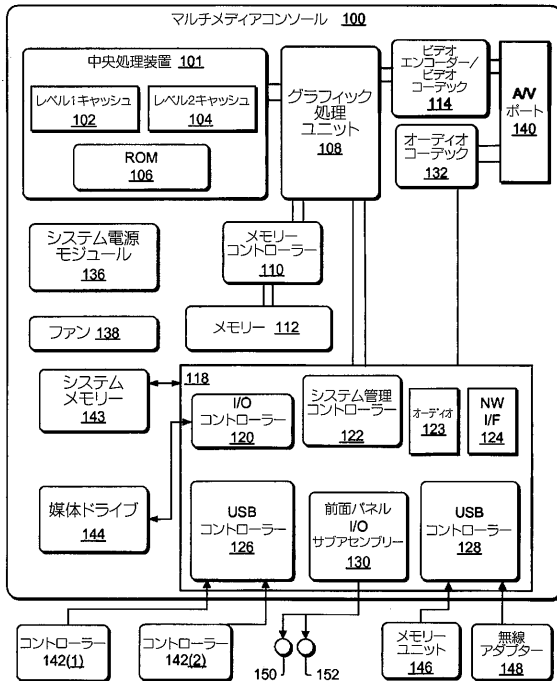


Fig. 1

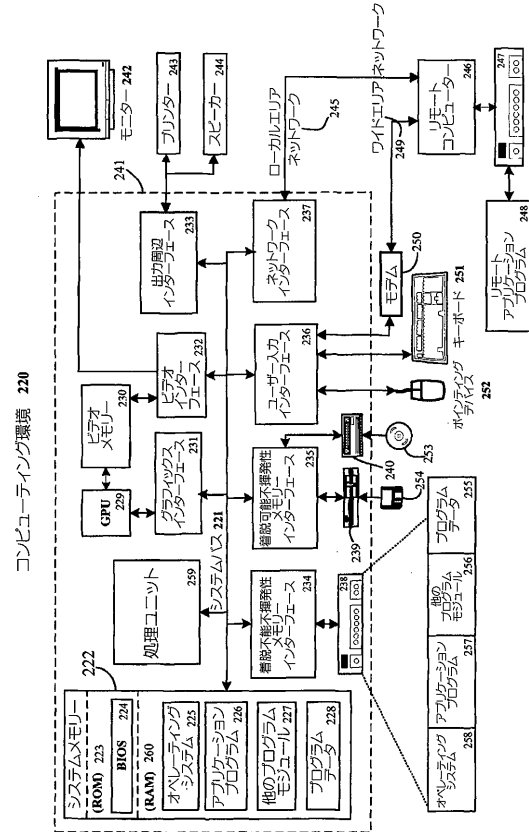
【図2】



【図 3】



【図 4】



【図 5】

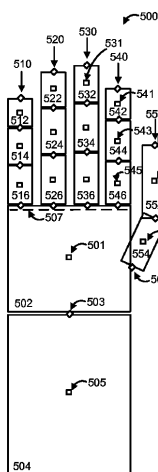
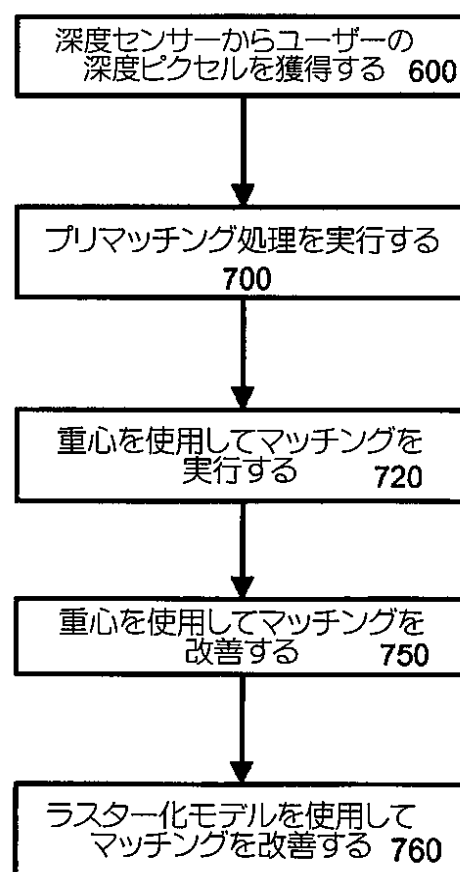
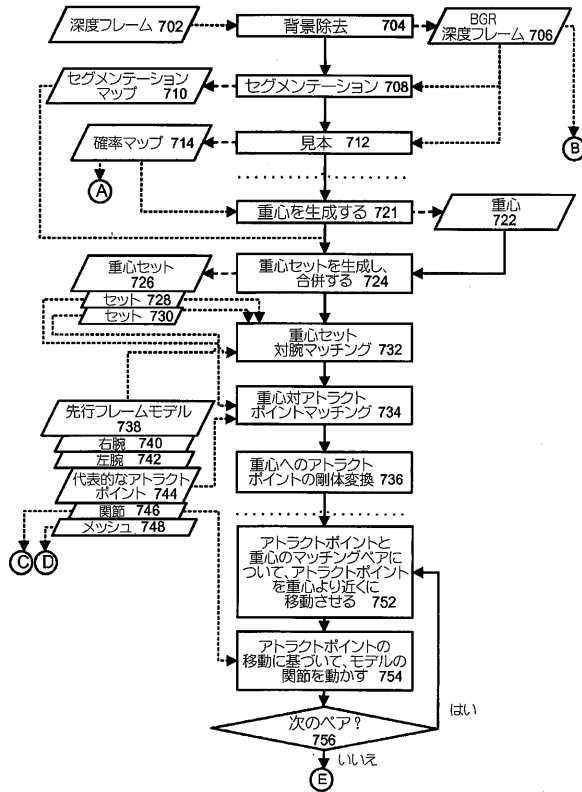


Fig. 5

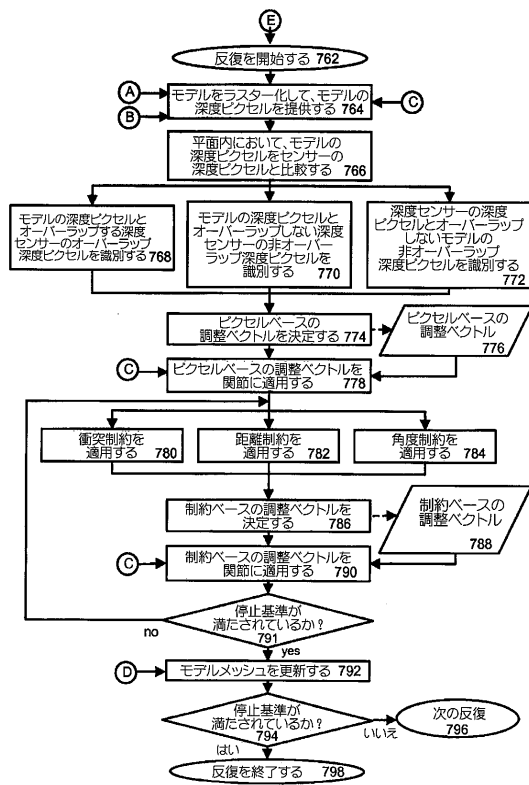
【図 6】



【図 7 A】



【図 7 B】



【図 8 A】

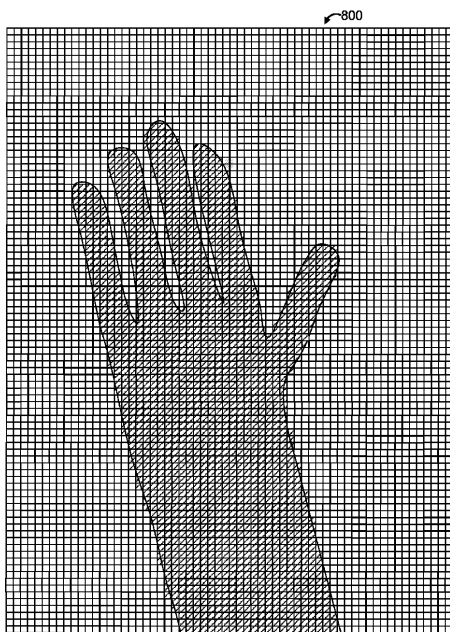


Fig. 8A

【図 8 B】

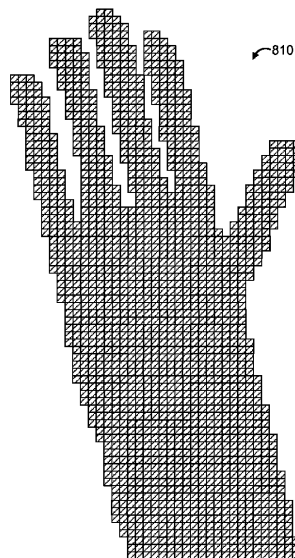
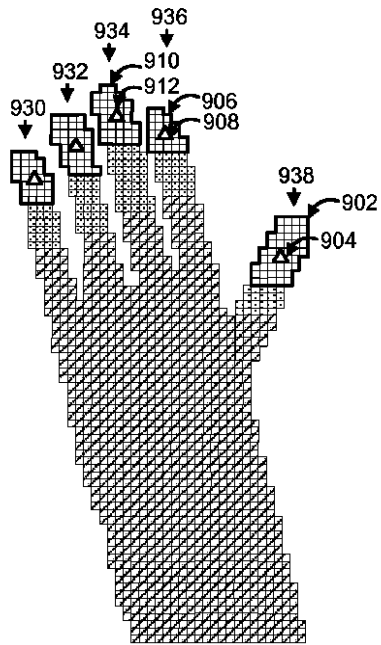
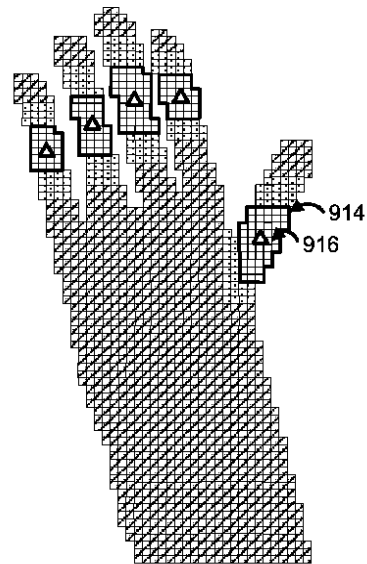


Fig. 8B

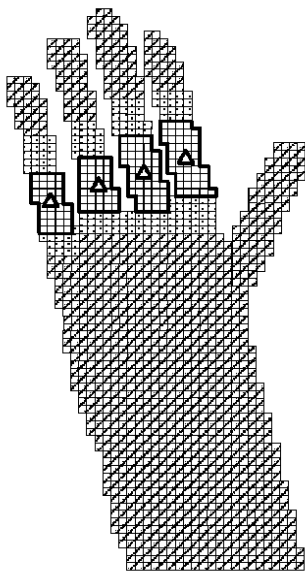
【図 9 A】

**Fig. 9A**

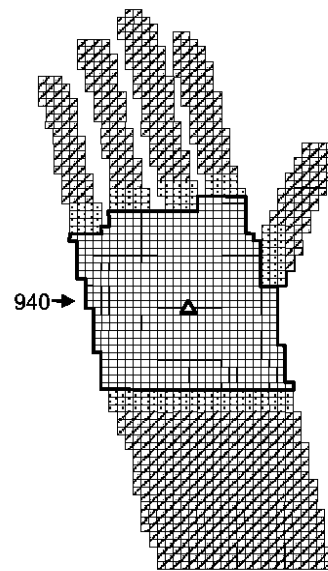
【図 9 B】

**Fig. 9B**

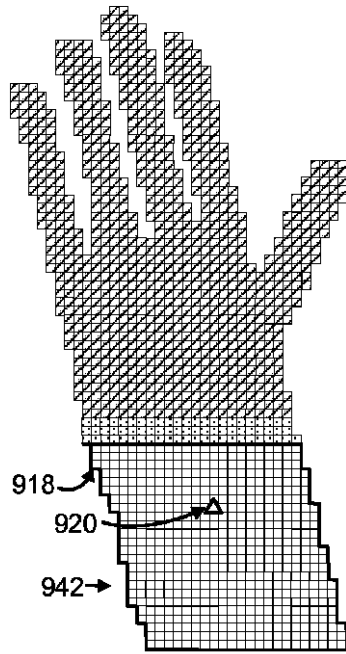
【図 9 C】

**Fig. 9C**

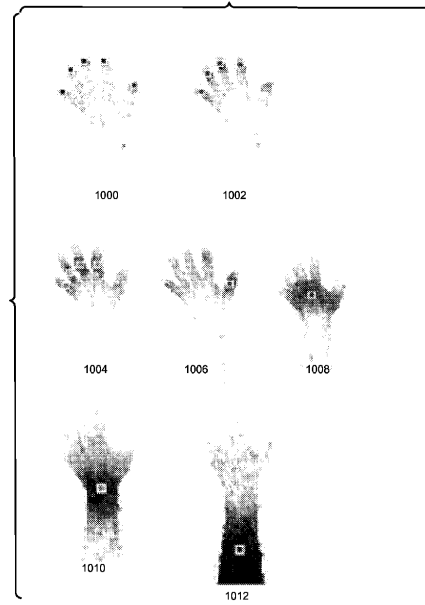
【図 9 D】

**Fig. 9D**

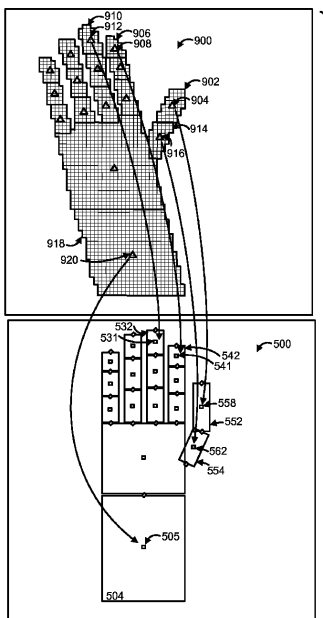
【図 9 E】

**Fig. 9E**

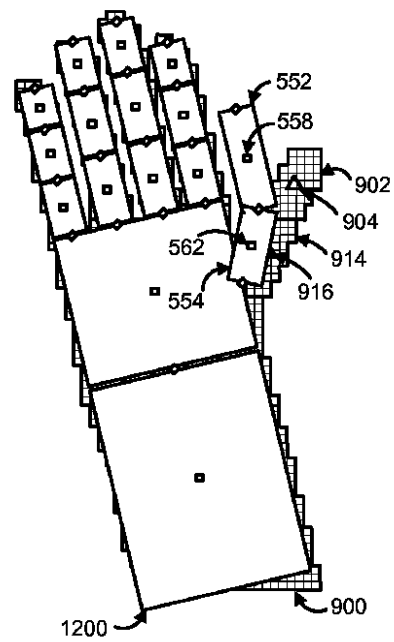
【図 1 0】

**Fig. 10**

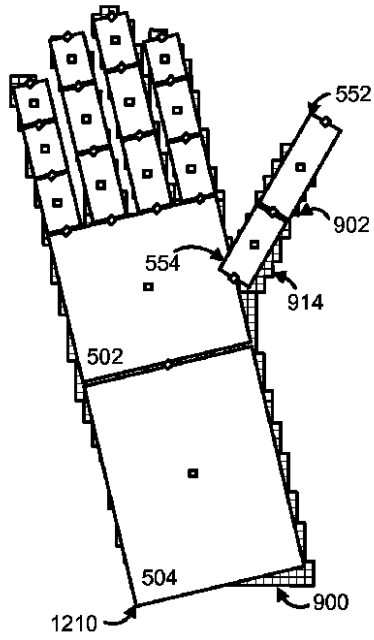
【図 1 1】

**Fig. 11**

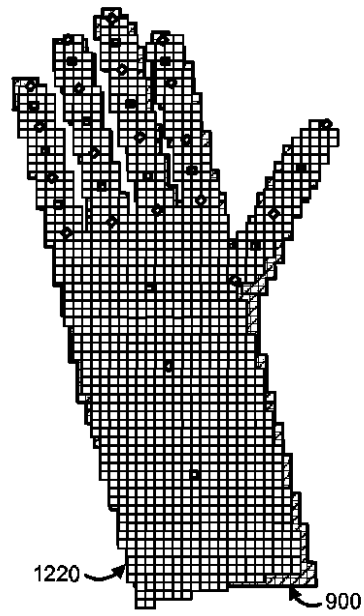
【図 1 2 A】

**Fig. 12A**

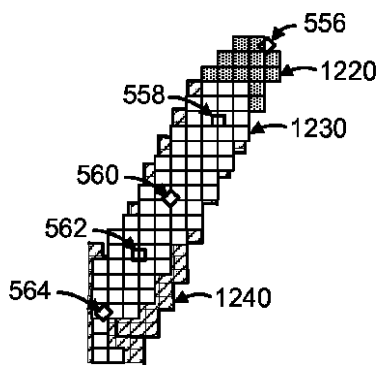
【図 1 2 B】

**Fig. 12B**

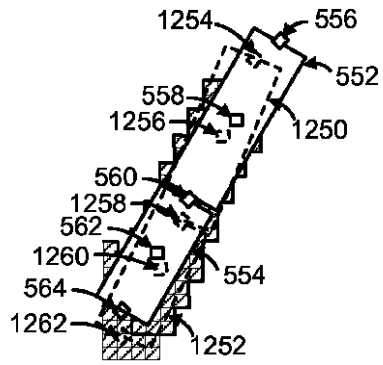
【図 1 2 C】

**Fig. 12C**

【図 1 2 D】

**Fig. 12D**

【図 1 2 E】

**Fig. 12E**

【図 1 2 F】

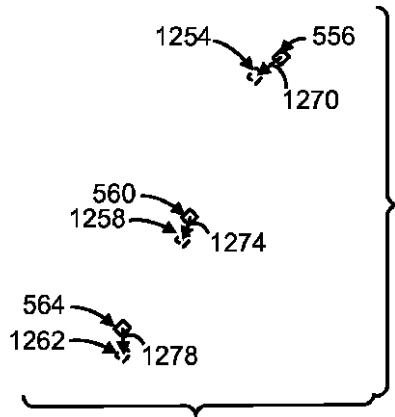


Fig. 12F

【図 1 2 G】

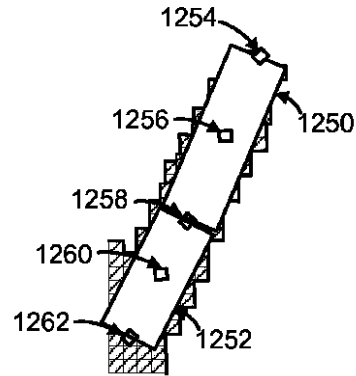
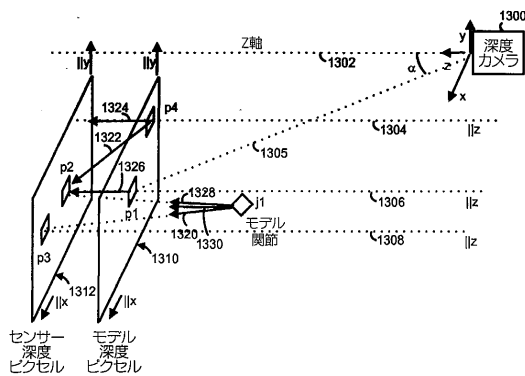


Fig. 12G

【図 1 3 A】



【図 1 3 B】

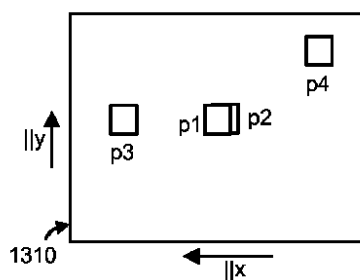
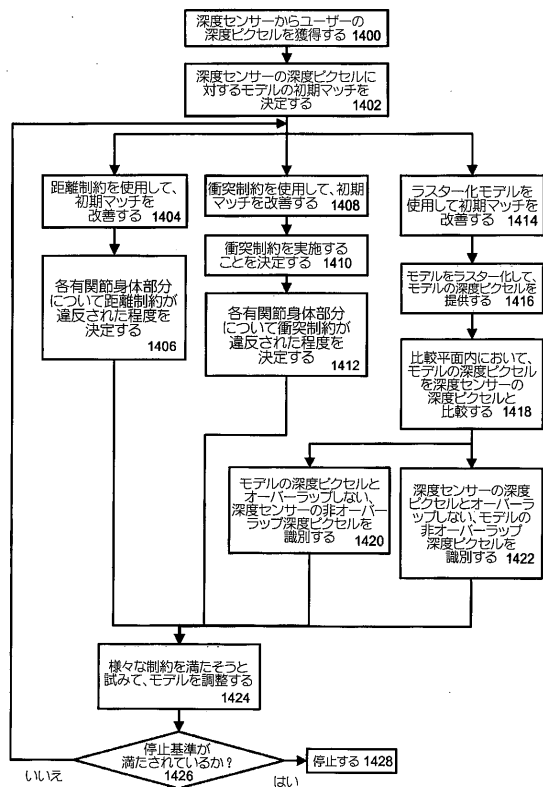


Fig. 13B

【図 1 4】



【図 16 B 2】

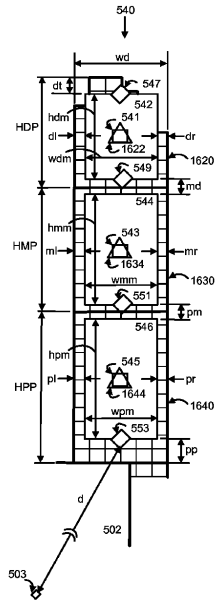


Fig. 16B2

【図 16 C】

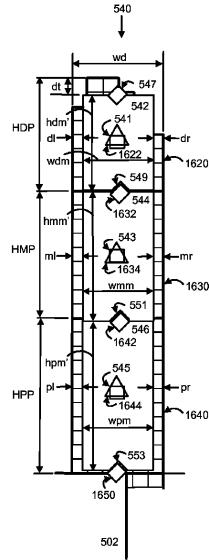


Fig. 16C

【図 16 D】

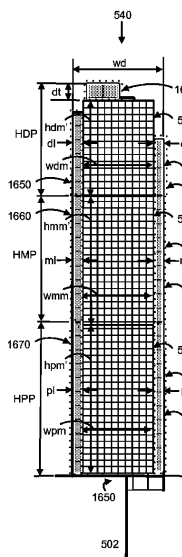


Fig. 16D

【図 16 E】

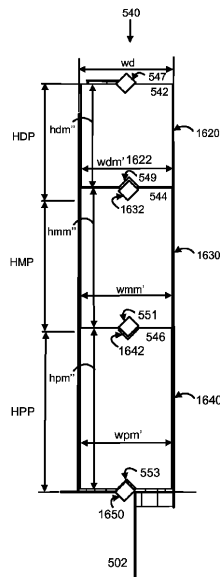


Fig. 16E

【図 17 B】

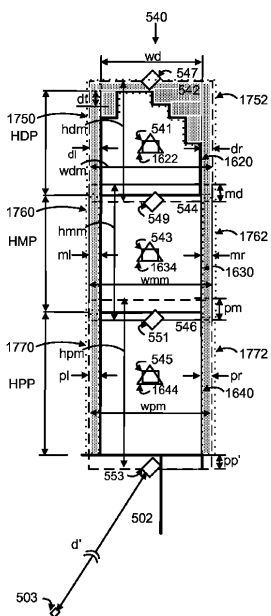


Fig. 17B

【図 1 8 B】

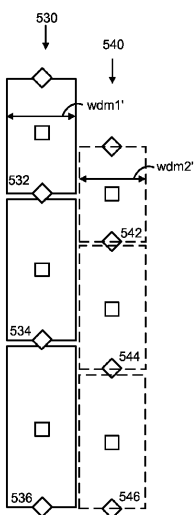


Fig. 18B

フロントページの続き

- (74)代理人 100153028
弁理士 上田 忠
- (74)代理人 100120112
弁理士 中西 基晴
- (74)代理人 100196508
弁理士 松尾 淳一
- (74)代理人 100147991
弁理士 鳥居 健一
- (74)代理人 100119781
弁理士 中村 彰吾
- (74)代理人 100162846
弁理士 大牧 綾子
- (74)代理人 100173565
弁理士 末松 亮太
- (74)代理人 100138759
弁理士 大房 直樹
- (72)発明者 リー, キュンスク・デーヴィッド
アメリカ合衆国ワシントン州98052-6399, レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイインターナショナル・パテント
- (72)発明者 バラン, アレクサンドル
アメリカ合衆国ワシントン州98052-6399, レッドモンド, ワン・マイクロソフト・ウェイ, マイクロソフト コーポレーション, エルシーエイインターナショナル・パテント

審査官 千葉 久博

- (56)参考文献 特開2007-333690 (JP, A)
特開2006-209445 (JP, A)
特表2011-501316 (JP, A)
米国特許出願公開第2011/0069866 (US, A1)
国際公開第2011/005784 (WO, A1)
国際公開第2010/088032 (WO, A1)
平尾公男, 外3名, "多関節CGモデルと距離画像による上半身の姿勢推定", 電子情報通信学会技術研究報告, 日本, 社団法人電子情報通信学会, 2005年 1月14日, 第104巻, 第573号, p.79-84
- (58)調査した分野(Int.Cl., DB名)
G06T 13/00-13/80, 19/00-19/20
G06T 15/00-15/87