

(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2014-29656  
(P2014-29656A)

(43) 公開日 平成26年2月13日(2014.2.13)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G06F 3/01 (2006.01)</b>	G06F 3/01 310C	2K103
<b>G09G 5/00 (2006.01)</b>	G09G 5/00 510B	5B087
<b>G09G 5/36 (2006.01)</b>	G09G 5/36 510V	5C082
<b>G03B 21/00 (2006.01)</b>	G09G 5/00 550A	5E555
<b>G06F 3/0346 (2013.01)</b>	G09G 5/00 510H	

審査請求 未請求 請求項の数 6 O L (全 22 頁) 最終頁に続く

(21) 出願番号 特願2012-171408 (P2012-171408)  
 (22) 出願日 平成24年8月1日(2012.8.1)  
 (31) 優先権主張番号 特願2012-144691 (P2012-144691)  
 (32) 優先日 平成24年6月27日(2012.6.27)  
 (33) 優先権主張国 日本国(JP)

特許法第30条第2項適用申請有り 平成24年2月2日発表の修士論文にて公開した。

(71) 出願人 598123138  
 学校法人 創価大学  
 東京都八王子市丹木町1丁目236番  
 (74) 代理人 100094053  
 弁理士 佐藤 隆久  
 (72) 発明者 今村 弘樹  
 東京都八王子市丹木町1-236 創価大学内  
 (72) 発明者 和田 英行  
 東京都八王子市丹木町1-236 創価大学内

Fターム(参考) 2K103 AA16 AA22 AA27 AB10 CA54  
 5B087 AA07 AE03 BC32  
 5C082 AA03 AA27 BA12 BA35 BA46  
 CB01 MM08

最終頁に続く

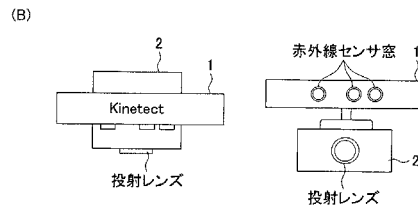
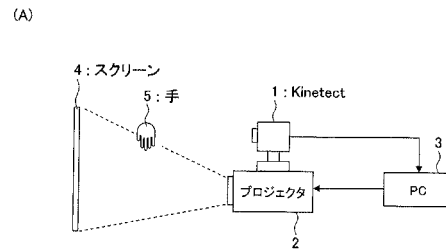
(54) 【発明の名称】 画像処理装置および画像処理方法

(57) 【要約】 (修正有)

【課題】 3次元画像の3次元的な画像処理を指示する操作手段として、たとえば、人間の手を、Kinectなどの撮像手段を用いて撮像し、その撮像結果を分析して手の位置、姿勢などを検出して、手の位置の変化、姿勢などに応じて3次元画像に画像処理を施して、表示する。

【解決手段】 画像処理装置は、スクリーン4に投影されている3次元画像に対する3次元の操作を指示する、たとえば、手5を撮像して、その色信号および深度信号を出力する、撮像手段、たとえば、Kinect1と、Kinect1から出力された色信号および深度信号から手5の動きを検出し、検出した結果に応じて3次元画像に対する画像処理を行なうPC3と、PC3から出力された画像信号をスクリーン4に投影するプロジェクタ2とを有する。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

表示された、または、スクリーンに投影された 3 次元画像に対して 3 次元的な画像処理を指示する 3 次元操作対象を撮像し、当該 3 次元操作対象の色信号および深度信号を出力する撮像手段と、

当該撮像手段から出力された色信号および深度信号を入力し、当該入力した色信号および深度信号を解析して前記 3 次元操作対象の操作内容を検出し、当該検出した結果に応じて前記 3 次元画像に対して画像処理を行なって表示手段に出力する、信号処理手段とを有する、  
画像処理装置。

10

**【請求項 2】**

上記表示手段は、上記信号処理手段から出力された画像処理された映像信号に応じた 3 次元画像をスクリーンに向けて投影（投射）する、投影手段を含む、  
請求項 1 に記載の画像処理装置。

**【請求項 3】**

上記撮像手段は、Kinectを含む、  
請求項 1 または 2 に記載の画像処理装置。

**【請求項 4】**

上記投影手段は、プロジェクタを含む、  
請求項 2 または 3 のいずれかに記載の画像処理装置。

20

**【請求項 5】**

上記手による操作指示は、フリックによるスクロール、ピンチイン、ピンチアウトを含む 3 次元画像の回転、移動などに対応する操作指示である、  
請求項 1 ~ 4 のいずれかに記載の画像処理装置。

**【請求項 6】**

表示手段に表示された、または、スクリーンに投影された 3 次元画像に対して 3 次元的な画像処理を指示する 3 次元操作対象を撮像し、当該 3 次元操作対象の色信号および深度信号を出力する撮像手段から出力された色信号および深度信号を解析して前記 3 次元操作対象の操作内容を検出し、

当該検出した結果に応じて前記 3 次元画像に対して画像処理を行なって表示手段に出力する、画像処理方法。

30

**【発明の詳細な説明】****【技術分野】****【0001】**

本発明は、画像処理装置およびその方法に関する。

より特定的には、本発明は、たとえば、3 次元画像（3 次元オブジェクト、3 D オブジェクト）に対する、回転などの信号処理を指示する操作手段、たとえば、人間の手の、3 次元位置、形状、動きを検出して、それに応じた 3 D オブジェクトに対する画像処理を行う、画像処理装置に関する。また本発明はそのような信号処理を行なう画像処理方法に関する。

40

**【背景技術】****【0002】**

ジェスチャーなど人間の操作に応じた画像処理を行う発明が種々提案されている（たとえば、特許文献 1 ~ 4 を参照）。

**【0003】**

また、近年、いわゆる、「スマートフォン」、タブレット PC が普及しており、これまでのキーボード、マウスによる入力操作ではなく、タッチパネル上で人間の指を動かすことにより、直観的に利用者の希望する情報をスマートフォン、タブレット PC などに入力することが可能となっている。すなわち、タッチパネルを有するデバイスに、文字、図形等を表示し、それらの情報をユーザがタッチしたり、スライドさせることにより、映像に

50

対する信号処理をすることが行われている。

このようなデバイスを用いると、従来のマウスを使用した場合に比較して、フリックによるスクロール、ピンチイン、ピンチアウトによる図形の3次元的な拡大または縮小（拡大・縮小）などが、ユーザの指の操作で直観的に行なうことができる。

【0004】

しかしながら、上述したデバイスにおける操作は、2次元画像に対する2次元的な操作であり、ユーザの手の位置、姿勢などの3次元的な情報を直接入力することができず、そのような3次元操作に対する画像の3次元的な処理を行なうことには限界がある。

また、上述したデバイスでは、個人を想定しているため、情報を表示する範囲が狭く、情報表示そのものに限界がある。

10

【0005】

他方、コンピュータゲームなどへの適用を想定した、Kinect（非特許文献1）を用いたマイクロソフト社によるHoloDeskが知られている。HoloDeskを用いると、Kinectで取得した手の形状情報に応じて反射板にコンピュータグラフィック（CG）を投影して、希望する操作を行うことができる。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開平10-177449号公報

【特許文献2】特開2001-246161号公報

【特許文献3】特開2010-039607号公報

【特許文献4】特開2010-079662号公報

【非特許文献】

【0007】

【非特許文献1】ウィキペディア（Kinect）

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、上記HoloDeskにおいては、下記に例示する課題または限界に遭遇している。

30

（1）手の形状情報に対して表示されたCGに対して3次元的な回転処理を行うことができない。

（2）情報を表示する範囲は、個人を想定しているため、狭く、情報表示に限界がある。

（3）特殊な反射板を用いており、枠があるところでしか使用出来ない。

【0009】

そこで、画像に対する3次元的な操作を直接指示する3次元操作手段として、たとえば、人間の手などを用い、その3次元操作手段、たとえば、手の操作指示内容を、たとえば、Kinectなどの3次元撮像手段（3次元形状取得手段）を用いて撮像し、撮像結果に基づく3次元操作手段の色信号および深度信号から3次元操作手段（たとえば、手）の位置、姿勢などを分析して、その分析結果に基づいて3次元画像に対して信号処理して表示することが望まれている。

40

【0010】

さらに、好ましくは、3次元画像に対して3次元的な画像処理された画像として、たとえば、プロジェクタなどの投射装置からスクリーンに投影（投射）可能とし、任意の大きさの画像で多数の聴衆に提示可能とすることが望まれている。

【0011】

上記3次元操作手段としての、たとえば、手は、フリックによるスクロール、ピンチイン、ピンチアウトなどを含む3次元画像の回転、移動などに対応する操作を指示する。

【0012】

50

また、たとえば、ミーティングなどにおいて、一人または複数のユーザ（聴衆）が、直観的な操作で、たとえば、インタラクティブな3次元画像の画像処理を行なうことが望まれている。

【課題を解決するための手段】

【0013】

上記要望を達成するため、本発明によれば、表示された、または、スクリーンに投影されて3次元画像に対して3次元的な画像処理を指示する3次元操作対象を撮像し、当該3次元操作対象の色信号および深度信号を出力する撮像手段と、当該撮像手段から出力された色信号および深度信号を入力し、当該入力した色信号および深度信号を解析して前記3次元操作対象の操作内容を検出し、当該検出した結果に応じて前記3次元画像に対して画像処理を行なって表示手段に出力する信号処理手段とを有する画像処理装置が提供される。

10

【0014】

好ましくは、上記表示手段は、プロジェクタなどの画像投影手段であり、当該画像投影手段は、上記信号処理手段から出力された画像処理された映像信号に応じた3次元画像をスクリーンに向けて投影（投射）する。

【0015】

また本発明によれば、表示された、または、スクリーンに投影されて3次元画像に対して3次元的な画像処理を指示する3次元操作対象を撮像した結果得られた当該3次元操作対象の色信号および深度信号を解析して前記3次元操作対象の操作内容を検出し、当該検出した結果に応じて前記3次元画像に対して画像処理を行なって表示手段に出力する信号処理方法が提供される。

20

【発明の効果】

【0016】

本発明によれば、画像に対する3次元的な操作、たとえば、フリック、ピンチイン、ピンチアウトを含み、回転、移動などを指示する3次元操作手段として、たとえば、人間の手などを用いる。これにより、ユーザが直接的に操作を行なうことが可能となる。

また本発明によれば、手などの操作指示内容を、たとえば、Kinectなどの3次元撮像手段（3次元形状取得手段）を用いて撮像し、撮像結果に基づく3次元操作手段の色信号および深度信号から3次元操作手段（たとえば、手）の位置、姿勢などを分析して、その分析結果に基づいて3次元画像に対して信号処理して表示するので、直接的かつインタラクティブな操作が可能となる。

30

【0017】

好ましくは、本発明によれば、3次元画像に対して3次元的な画像処理された画像として、たとえば、プロジェクタなどの投射装置からスクリーンに投影（投射）可能としており、任意の大きさの3次元画像を多数の聴衆に同時に提示することができる。

【0018】

さらに、本発明によれば、たとえば、ミーティングなどにおいて、一人または複数のユーザ（聴衆）が、直観的な操作で、たとえば、インタラクティブな3次元画像の画像処理を行なうことができる。

40

【図面の簡単な説明】

【0019】

【図1】図1は本発明の画像処理装置の1実施の形態としての構成を示す図であり、図1（A）は全体構成図、図1（B）は図1（A）におけるプロジェクタとKinectの平面図、図1（C）は図1（A）におけるプロジェクタとKinectの正面図である。

【図2】図2は図1に図解した画像処理装置の動作の形態を示す図である。

【図3】図3は図1（A）に図解したパーソナル・コンピュータ（PC）の処理の流れを示すフローチャートである。

【図4】図4は図1（A）に図解したスクリーン4の4隅に4個のランドマーク（基準マ

50

ーク)を配置する例を示す図である。

【図5】図5は、スクリーンの4隅に投影されたランドマークをKinectで撮像して、4隅の座標位置を検出する方法を示す図である。

【図6】図6はKinectから出力される深度情報の初期値の例を示す図である。

【図7】図7はスクリーンの投射面に手を置いた時に、それを撮像したKinectの手の深度情報の例を示す図である。

【図8】図8(A)~(C)は手移動抽出範囲を認識することを示す図である。

【図9】図9はプロジェクタからスクリーンに投影(投射)する画像領域を抽出することを示す図である。

【図10】図10はスクリーンにおける手の中心の座標を取得する方法を示す図である。

【図11】図11はスクリーンにおける腕の中心の座標を取得する方法を示す図である。

【図12】図12はスクリーンに腕が入る状況での手の中心位置の座標を示す図である。

【図13】図13は手の垂直成分を示す図である。

【図14】図14は、図3に図解したステップ4の分岐処理を詳細に図解したフローチャートである。

【図15】図15は両手の認識を行う方法を示す図である。

【図16】図16は手がグー(グー、または、パー)のいずれかであることを認識することを示す図である。

【図17】図17は手の3次元座標系から3次元画像(対象物)の3次元座標系への座標変換を示す図である。

【図18】図18は手の指の角度を検出することを示す図である。

【図19】図19は手の平面を推定する方法を示す図である。

【図20】図20は両手の中心座標を示す図である。

【発明を実施するための形態】

【0020】

#### 装置構成

図1を参照して本発明の実施の形態の画像処理装置10について述べる。

図1は本発明の画像処理装置の1実施の形態としての構成を示す図であり、図1(A)は全体構成図であり、図1(B)は図1(A)に図解したプロジェクタとKinectの平面図であり、図1(C)は図1(A)に図解したプロジェクタとKinectの正面図である。

【0021】

画像処理装置10は、Kinect(1)と、プロジェクタ2と、パーソナル・コンピュータ(PC)3と、スクリーン4とを有する。

Kinect(1)は、3次元画像に対する3次元操作を指示する操作対象としての手5を、実時間で撮像し、手5の色信号、たとえば、各ピクセル毎のRGB信号、および、たとえば、図6に例示した各ブロックごとの深度信号を出力する、本発明の撮像手段(または、3次元形状取得手段)の1例である。

【0022】

パーソナル・コンピュータ(PC)3は、下記に詳述する信号処理を行なう演算処理部(CPU)、各種データを記憶するメモリ部(記憶部)、Kinect(1)からの色情報と深度情報を入力する入力インターフェース、信号処理した画像(映像)信号をプロジェクタ2に出力する出力インターフェースなどを有して構成されている。

PC3は、Kinect(1)から出力された色信号および深度信号を実時間で入力してメモリ部に記憶し、入力した色信号および深度信号を解析して手5で指示された画像処理のための操作内容を検出し、当該検出した操作結果に応じた画像処理を、スクリーン4に表示される3次元画像に対して行なって映像信号を実時間で表示手段に出力する、本発明の信号処理手段の1例である。

【0023】

なお、3次元画像の原画像と、上画像処理した修正(更新)画像とをPC3のメモリ部

10

20

30

40

50

に記憶している。

また、Kinect(1)から出力されたRGB信号から輝度信号を得ることができる。

特に、下記に述べる手5の内容を分析するためには、RGB信号そのものを用いず、たとえば、0~255の階調を持つ白色または黒色の輝度信号として扱う。

#### 【0024】

プロジェクタ2は、PC3などの信号処理手段から出力された映像信号をスクリーン4に投影(投射)する、本発明の投影(投射)手段の1例である。

#### 【0025】

本実施の形態において、3次元オブジェクトとしてのスクリーン4に投射される3次元画像に対する3次元画像操作を指示する3次元操作手段は、人間の手5である。

しかしながら、指を備えた手5に相当する義手、または、それに相当する手段など、たとえば、フリックして画像をスクロールさせる手段、ピンチイン、ピンチアウトさせる手段、あるいは、後述する移動、回転などをさせるための他の手段を用いることができる。

#### 【0026】

Kinect(1)はスクリーン4を指向して配設されている。その理由は、本実施形態においては、手5がスクリーン4の前の手移動可能領域に規定されており、手5がスクリーン4の前とKinect(1)との間で操作されるためである。

Kinect(1)は、Kinect(1)とスクリーン4との間に位置して3次元画像に対する画像処理操作を指示する手5を、たとえば、赤外線センサ窓を介して入射した赤外線を用いて実時間で撮像して、手5の色信号および深度信号を実時間でPC3に出力する。

#### 【0027】

色信号としては、各ピクセル毎の、たとえば、R(赤)G(緑)B(青)の3原色信号である。また、深度信号は、たとえば、図6に例示したように、各ブロックごとの深度信号である。

なお、この例示では、Kinect(1)はプロジェクタ2の上に載置されている。

#### 【0028】

PC3は、図2に画像処理装置10の処理の概要を示したように、Kinect(1)から出力されたRGB画像信号と深度信号とを実時間で入力して、手5の3次元座標および形状(姿勢)を示す情報を取得して、手5の3次元画像に対する基本的な3次元操作、たとえば、移動、回転、拡大、縮小などの処理を解析して、その操作に応じた画像処理を行う。

PC3は、その結果の映像信号を実時間でプロジェクタ2に出力する。

PC3の処理内容については図3を参照して後述する。

#### 【0029】

プロジェクタ2は、PC3から出力された画像データに対応した3次元画像を投射(投影)レンズからスクリーン4に投射(投影)する。

#### 【0030】

Kinect(1)による手5の撮像、Kinect(1)に対するPC3による信号処理、PC3の処理結果をスクリーン4に投射するプロジェクタ2の処理が、実時間で反復されることにより、ユーザは、自己の手5の操作に応じた3次元画像をスクリーン4に、実時間で、連続的に、投射(投影)させることができる。

#### 【0031】

図3を参照してPC3の処理内容を述べる。

#### 【0032】

##### ステップ1、初期設定

PC3は図4~図6を参照して述べる下記の初期設定を行う。

この初期設定は、特に、プロジェクタ2と、スクリーン4との距離を任意に設定可能と

10

20

30

40

50

しているためであり、プロジェクタ2とスクリーン4との距離が固定している場合は、不要である。なお、この初期設定は、必ずしも、実時間で行なう必要はない。

#### 【0033】

##### (1) ランド(基準)マークの投影

図4は、図1(A)に図解したスクリーン4の4隅に隅(角)を指向した4個のランドマーク(基準マーク)を配置する例を示す図である。

ユーザは、スクリーン4の投影面(投射面)の4隅に隅(角)を指向したランドマークを配置する。

なお、4個のランドマークは、スクリーン4に対する手5の操作範囲(手移動抽出範囲)を規定し、さらに、PC3からプロジェクタ2を介して出力される画像のスクリーン4

10

における有効範囲を規定する。  
すなわち、4個のランドマークで規定される領域にプロジェクタ2からの画像が投影され、この領域(手移動抽出範囲)において手5が操作される。

#### 【0034】

##### (2) ランドマーク画像の取得

図5はスクリーン4の4隅に配置されたランドマークをKinect(1)で撮像して、4隅の座標位置を検出する方法を示す図である。

Kinect(1)は、ランドマークが配置されたスクリーン4を撮像して、RGB信号および、深度信号を生成して、PC3に出力する。

PC3はKinect(1)から出力されたRGB信号を、PC3内のメモリ部に記憶する。なお、PC3は手5の操作を処理するとき、RGB信号としてではなく、たとえば、0~255の階調を持つ輝度信号として用いるので、PC3のメモリ部には輝度信号も記憶される。メモリ部は、各ピクセル毎に輝度信号をも記憶している。

20

#### 【0035】

PC3は、Kinect(1)から出力された4隅に配置したランドマークのテンプレートマッチングを行い、RGB画像上の4隅のランドマークの輝度からランドマークの座標を取得して、メモリ部に記憶する。

この処理以降、メモリ部に記憶したランドマークの座標の範囲内で、PC3が画像処理を行う。

#### 【0036】

##### (3) 深度情報の初期値の記憶

図6はKinectから出力されるランドマークの深度情報の初期値の例を示す図である。

スクリーン4の投射面にユーザの手5が存在しない状態で、Kinect(1)でスクリーン4を撮像して、Kinect(1)からのスクリーン4上のランドマークの深度状態を取得する。この深度情報は、Kinect(1)の先端からスクリーン4までの距離を意味する。

PC3はその深度情報、すなわち、スクリーン4までの距離情報をPC3内のメモリ部に記憶する。

30

#### 【0037】

この処理は、後述するスクリーン4の投射面に手5を置いた場合の深度情報と対比して、スクリーン4の投影面からKinect(1)に向かう距離が、たとえば、20mm以上の領域を手5の移動する領域(手移動可能領域)として認識するための処理である。上記例示した20mmを、手移動可能領域の下限値という。

40

なお、この例示では、Kinect(1)から、ピクセル毎ではなく、横8×縦8=64ブロック(区分)の各ブロックごとに深度情報が出力される。

この例示では、スクリーン4が下から上に向かって傾斜しており、上列の深度情報が大きく、下列の深度情報が小さい。

#### 【0038】

以上の初期値設定により、スクリーン4とKinect(1)との間の距離、およ

50

び/または、手5とスクリーン4またはKinect(1)との間の距離が任意であっても、画像処理装置10を使用可能となる。

換言すれば、スクリーン4に表示する3次元画像の大きさ、手移動可能領域の大きさ等に応じて、スクリーン4とプロジェクタ2の距離、または、スクリーン4とKinect(1)の距離を設定した場合に、上述した初期設定を行えば、希望するスクリーン4とKinect(1)またはプロジェクタ2の距離に応じた3次元画像(3次元オブジェクト)を得ること、および、手5の操作が可能となる。

#### 【0039】

##### ステップ2、投影画像上の手領域の抽出

図7~図9を参照して、ユーザの手5の3次元座標を取得するために、PC3は、Kinect(1)と協働して、スクリーン4の投影画像上の手5の領域(手移動可能領域)の抽出を行う。

#### 【0040】

##### (1)手移動可能領域の認識

図7はスクリーンの投射面にユーザが手5を置いた時に、それを撮像したKinect(1)の手の深度情報の例を示す図である。

図7に例示したように、ユーザがスクリーン4の投影面に手5を置いた状態で、Kinect(1)で手5を撮像し、Kinect(1)は撮像による深度情報をPC3に出力する。PC3はその深度情報をメモリ部に記憶する。

#### 【0041】

PC3は、図5に例示したスクリーン4の投射面に手5がないときの深度情報の初期値と、図7に例示したスクリーン4の投射面に手5を置いたときのKinect(1)で取得した現在の深度情報との差を、各ブロック(区分)ごとに求める。

なお、図7に例示した深度情報と図5に例示した深度情報とを対比すると、図7の深度情報は手5があるブロック(区分)の深度情報の値が小さい。

スクリーン4の投射面における手5の深度情報の差が、スクリーン4とKinect(1)との間の手移動可能領域の下限值、たとえば、20mmより大きいとき、PC3は、有効な手移動可能領域とする。この例示では、手5はスクリーン4の投射面から20mm以上離して操作する。

#### 【0042】

図8(A)~(C)は手移動可能領域を認識することを示す図である。図8(A)は図6を参照して述べたスクリーン4の投射面の前に手5がないときの深度情報の初期値を示し、図8(B)は図7を参照して述べたスクリーン4の投射面の前にユーザの手5を置いたときの深度情報を示す。図8(C)は、上述したPC3の処理により、ユーザの手5の部分の、図8(B)の深度情報と図8(A)の深度情報との差が手移動可能領域の下限值である20mm以上ある領域を明示した図である。

#### 【0043】

次に、PC3は図9に図解した投影画像領域を抽出する。

図9はプロジェクタ2からスクリーン4に3次元画像を投影(投射)する画像領域を抽出することを示す図である。図9において、左側の上部にプロジェクタ2側から肉眼で見た画面を示し、左側の下部にKinect(1)で撮像した画像を示す。また、図9において、右側の上部に左側の上部の肉眼で見た画面を4隅のランドマーク内に展開した図を示し、右側の下部にスクリーン4の投影画像上の手移動可能領域を示す。

PC3は、図9の左側の上部に図解した、初期設定で取得したスクリーン4へのランドマークの投影画像の4隅の座標を用いて、座標変換のための射影変換行列を算出し、手移動可能領域から、図9の右側に図解した、手移動可能領域を抽出する。

#### 【0044】

参考までに、座標変換のための射影変換について述べる。

$n + 1$ 次元アフィン空間  $A^{n+1}$  の一次変換  $F$  を考える。

$$F : A^{n+1} \rightarrow A^{n+1} (u \rightarrow MU)$$

10

20

30

40

50



ただし、 $u$  は、縦ベクトル表示してあるものとし、  
 $M$  は、 $n + 1$  次正方形行列である。

原点を通る直線は  $F$  によって再び原点を通る直線に移る。したがって、 $F$  は射影空間の変換を引きおこしている。

$$F : P^n \rightarrow P^n \quad ([u] \rightarrow M[u] = [Mu])$$

これを射影変換という。

なお、 $n + 1$  次元の縦ベクトル  $u ( 0 )$  に対して、 $[u]$  は対応する  $n$  次元射影空間の点とする。

【0045】

以上の処理により、初期設定で取得したスクリーン4へのランドマークの投影画像内における手移動可能領域が抽出できる。

10

【0046】

#### ステップ3、手の3次元座標取得

図10～図13を参照して、手5の3次元座標 ( $x, y, z$ ) の取得を行う。

(1) 手の中心座標の取得

図10は、スクリーン4における手5の中心の座標を取得する方法を示す図である。

PC3は、図10の左側に図解した手5の手移動可能領域を抽出した画像を、図10において、白画素から最も近い黒画素までの距離で正規化する。そして、PC3は、輝度値が最大の輝度、本例では、255の画素を検出し、図10の右側に図解したように、その座標を手5の中心座標とする。

20

【0047】

なお、本例では、手5のある部分(ピクセル)は白画素または白画素に近い画素として表し、手5が存在しない部分、たとえば、3次元画像、手5が存在しないスクリーン4の白い色の部分を、黒画素として図解している場合について、述べている。

【0048】

図11はスクリーン4の手移動可能領域に、手5だけでなく、腕も入った時の、腕の中心の座標を取得する方法を示す図である。

図11の左側に図解したように、手5の他に腕も手移動可能領域に入った場合には、その状態でスクリーン4の投影画像上の手移動可能領域を、白画素から最も近い黒画素まで距離で正規化して、最大輝度値、たとえば、255の画素を検出すると、図11の右側に図解したように、腕の中心座標を検出してしまうという問題が起こる。

30

【0049】

この問題を解決するため、図12に図解したように、腕の画像を排除する処理を行う。

図12はスクリーン4の前の手移動抽出範囲に腕が入る状況での手の中心位置の座標を示す図である。

PC3は、腕がスクリーン4の前の手移動可能領域に位置する状態で手の中心座標を検出する。

まず、PC3は、腕の画像を削除するため、PC3のメモリ部に記憶してある画像データについて、取得した腕の中心座標から輪郭までの最短距離を計測する。

PC3は、その距離を半径として、図12の左側に図解したように、腕の中心座標を中心とする黒い円を描く。そして、黒い円の輝度は低いので、それら低い輝度を排除して、腕の画像を廃棄する。これにより、手移動可能領域に入り込んだ腕の情報を削除することができる。

40

したがって、ユーザは、手5の操作に対して、特別注意をする必要がない。

【0050】

次に、PC3は、白画素から最も近い黒画素までの距離で正規化し、最大輝度値、たとえば、255の画素を検出する。

この最大輝度値の画素の部分が、図12の右側に図解した、手5の中心座標 ( $x, y$ ) である。

【0051】

50

PC3は、さらに、図13に図解したように、取得した手5の中心座標(x, y)の深度情報の初期値と現在の深度情報の差を算出して、垂直成分zとする。

この垂直成分は、手5までの距離を意味している。

以上の処理により、ステップ3の手の3次元座標を取得する処理が終了する。

#### 【0052】

##### ステップ4、分岐処理

手5の操作に応じて、スクリーン4に表示されている3次元画像(すなわち、3Dオブジェクト)の移動、回転、拡大、縮小操作の切り替えを一意的に行うため、操作の分岐処理を行う。

図14は、図4に図解したステップ4の分岐処理を詳細に図解したフローチャートである。

図4においては、3Dオブジェクトの移動操作(ステップ4a)、3次元画像(3Dオブジェクト)の回転操作(ステップ4b)、3Dオブジェクトの拡大、縮小操作(ステップ4c)と、簡略して図解している。

他方、図14においては、片手または両手の操作(ステップS41)、片手のゲー操作またはパー操作(ステップS42)、両手のゲー操作またはパー操作(ステップS43)、Kinect(1)から見て、スクリーン4に表示されている3次元画像(3Dオブジェクト)の上に(または、前に)手5がある処理(ステップS44)、移動操作(ステップS45)、z軸回転操作(ステップS46)、xy軸回転操作(ステップS47)、拡大・縮小操作(ステップS48)、3Dオブジェクトの表示(ステップS49)の操作に分類している。

#### 【0053】

##### 3Dオブジェクトの移動操作

3次元画像(3Dオブジェクト)の移動操作は、手5が片手で(ステップS41)、手5の形状が「ゲー(握りこぶし)」のとき(ステップS42、ステップS45)、その手5の移動方向、移動量、移動速度などに対応して行う(ステップS45)。

PC3が行う3次元画像の移動操作としては、スクリーン4において3次元画像の上下左右への移動、スクリーン4から3次元画像がKinect(1)に向かって近づく、または、Kinect(1)からスクリーン4に向かって遠ざかるなどの操作である。

なお、PC3は、3次元画像がスクリーン4に向かって移動するとき、3次元画像をその距離に応じて縮小させることもできる。他方、PC3は、3次元画像がスクリーン4から離れてKinect(1)に向かって接近するとき、3次元画像をその距離に応じて拡大させることもできる。

#### 【0054】

そのため、PC3は、手5の移動を、メモリ部に記憶している、Kinect(1)で撮像した連続する複数の画像を示す、色信号(輝度信号)および深度情報を参照して、判断する。

以下、各画面におけるKinect(1)の撮像データに基づく、PC3の処理について述べる。

#### 【0055】

##### ステップS41、片手または両手の認識

図15は両手の認識を行う方法を示す図である。

PC3は、図15に図解したように、Kinect(1)から出力されて、メモリ部に記憶しているRGB信号に基づく輝度値のうち、スクリーン4の投影画像上の画像について、たとえば、図7、図8に例示した区分(領域)のピクセル(画素)ごとに、画像処理において、画像処理において通常行われている「ラベリング処理」を行い、輝度が第1の輝度値、たとえば、250、より大きな最大領域を抽出して、最大領域を認識する。

#### 【0056】

次いで、PC3は、上述した手移動可能領域に腕が入ったときの排除処理と同様、上記

10

20

30

40

50

抽出した最大領域の画像データを黒くして、次の処理の対象から外す。

P C 3 は、再度、「ラベリング処理」を行い、輝度が上記第 1 の輝度値より小さな第 2 の輝度値、たとえば、240、より大きな大きい領域が所定量のピクセル、たとえば、100ピクセル以上存在すれば、両手であると認識する。

両手でなければ、P C 3 は、片手であると判断する。

【0057】

ステップ S 4 2、S 4 3、手の形状認識（ゲー、または、パーの検出）

図 1 6 は手がゲーパー（ゲー、または、パー）のいずれかであることを認識することを示す図である。

まず、P C 3 が、上記処理で、手移動可能領域にあるものが、片手であると認識した場合を例示して述べる。

【0058】

（1）輪郭検出

P C 3 は、スクリーン 4 の投影画像上の輪郭を検出する。この輪郭の検出は、隣接する二次元方向の領域（ピクセル）の輝度の差が所定の値以上の部分として（あるいは、二次元方向の微分値が所定値以上の部分として）検出することができる。その結果を、図 1 6 において、細い白線で示している。

【0059】

（2）輪郭線の曲率の算出

P C 3 は次いで、輪郭線の曲率を算出する。

輪郭線の曲率は、変曲点を意味しており、手 5 が開いている場合は、握りこぶしするときより、変曲点は多くなる。

このような曲率の演算は、通常の数学モデルを用いて行なうことができる。

【0060】

（3）ゲーまたはパーの検出

P C 3 は、曲率の高い部分が所定量以下、たとえば、5 個以下の場合には、手 5 が「ゲー」であると判断する。他方、P C 3 は、曲率の高い部分が所定量以上、たとえば、6 個以下の場合には、手 5 が「パー」であると（検出）認識する。

【0061】

以上、P C 3 が片手であると認識した場合について述べたが、P C 3 が両手であると認識した場合についても、片手ごとに上述した処理を行う。

それにより、両手の場合についても、片手ごとに、「ゲー」または「パー」を検出することができる。

【0062】

動き検出

P C 3 は、Kinect (1) で撮像した結果得られた、ある時刻における RGB 信号および深度信号を用いて、上述した処理を行う。

P C 3 は、Kinect (1) から時々刻々と連続して出力される、RGB 信号および深度信号についての上述した処理結果を追跡している。

それにより、P C 3 は総合的に、（1）手 5 の動き、たとえば、手 5 全体の動き、手 5 の指の広がり程度、（2）さらに、それらの動きの速度、（3）動きの方向などを検出することができる。

すなわち、時々刻々と連続して出力される、RGB 信号および深度信号についての上述した処理結果の変化を追跡していくことにより、（1）手 5 の動き、たとえば、手 5 全体の動き、手 5 の指の広がり程度、（2）さらに、それらの動きの速度、（3）動きの方向などを検出することができる。

【0063】

3次元画像の移動操作としては、たとえば、上記手 5 の動き（操作）に応じた、スクリーン 4 において上下左右への移動、スクリーン 4 から Kinect (1) に向かって遠ざかる、または、Kinect (1) からスクリーン 4 に向かって近づくなどの操

10

20

30

40

50

作である。

【0064】

ステップS45、3Dオブジェクトの3次元移動操作

図17は手の3次元座標系(x, y, z)から3Dオブジェクト(3次元画像)の3次元座標系(X, Y, Z)への変換を示す図である。

【0065】

(1) x y 軸への変換

PC3は、図17に例示した手移動可能領域について、下記式(1)、(2)に基づいてx y 軸への変換を行う。

【0066】

(2) z 軸への変換

PC3は、次いで、図17に例示した範囲について、下記式(3)に基づいて、z 軸への変換を行う。

【0067】

【数1】

$$X = \frac{8}{320}x - 4 \quad \dots(1)$$

10

20

$$Y = -\frac{6}{240}y + 3 \quad \dots(2)$$

30

$$Z = 20\frac{z}{S} - 10 \quad \dots(3)$$

【0068】

以上の処理により、3次元画像(3Dオブジェクト)について、手5の3次元座標系(x, y, z)から、3次元座標系(X, Y, Z)への変換を行うことができる。

【0069】

PC3は、このようにして、座標変換を行い、手5が片手で(ステップS41)、手5の形状が「グー(握りこぶし)のとき(ステップS42、ステップS45)、その手5の移動に応じて量、および、その手5の移動方向に対応して、スクリーン4に投影されている3次元画像(3Dオブジェクト)の移動操作を行う(ステップS45)。

40

【0070】

この処理は、実際は、PC3は、現在、スクリーン4に投影されている3次元画像に対応する、PC3のメモリ部に記憶されている画像データに対して、移動操作に応じた処理を行うことを意味する。

移動操作以外の他の操作についても、PC3は、スクリーン4に投影されている3次元画像に対応するメモリ部に記憶されている画像データについて、対応する信号処理を行う。

【0071】

50

ステップ S 4 9、表示

PC 3 は、上記結果に基づく画像処理した画像データをプロジェクタ 2 に出力する。  
プロジェクタ 2 は、その画像データをスクリーン 4 に、3 次元画像として投影する。

【 0 0 7 2 】

ステップ S 4 4、4 6、3 D オブジェクトの垂直軸 ( Z 軸 ) の回転操作

図 1 8 は手の指の角度を検出することを示す図である。

スクリーン 4 に投影されている 3 次元画像 ( 3 D オブジェクト ) 上に手 5 がある場合、  
PC 3 は、手 5 の角度を用いて、垂直軸 ( Z 軸 ) の回転操作を行う。

PC 3 は、図 1 8 に例示したように、投影画像上の手移動可能領域の画像信号 ( R G B  
画像 ) において中心座標を腕の中心座標として、その半径を手の中心座標 ( x , y ) まで  
の距離で黒い円を描く。

10

PC 3 は次に、手の中心座標から最も遠い白画素を検出する。検出した座標を、中指の  
座標 ( F x , F y ) とする。

PC 3 は、中指の座標 ( F x , F y ) と手の中心座標 ( x , y ) の角度 F を、下記式  
( 4 ) により、算出する。

【 0 0 7 3 】

【 数 2 】

$$F\theta = \tan^{-1} \left( \frac{Fy - y}{Fx - x} \right) \quad \dots(4)$$

20

【 0 0 7 4 】

PC 3 は、算出した手の角度 F に応じて、3 D オブジェクトの垂直軸 ( z 軸 ) の回転  
操作を行う。

PC 3 におけるこの回転操作は、たとえば、CG などで行う変換処理と同様に行うこと  
ができる。

【 0 0 7 5 】

30

ステップ S 4 9、表示

PC 3 は、回転操作の結果に基づく画像データをプロジェクタ 2 に出力する。

プロジェクタ 2 は、その画像データをスクリーン 4 に 3 次元画像として投影する。

【 0 0 7 6 】

ステップ S 4 4、4 7、x y 軸回転操作

図 1 9 は手の平面を推定する方法を示す図である。

PC 3 は、手 5 がスクリーン 4 に投影されている 3 D オブジェクト ( 3 次元画像 ) の上  
にない場合、手の法線ベクトルを用いて、3 D オブジェクトの平面 ( x y 軸面 ) の回転操  
作を行う。

【 0 0 7 7 】

40

( 1 ) 手の 3 点の検出

PC 3 は、まず、手の法線ベクトル ( H ) を算出するため、図 1 9 に例示した手の 3 点  
を検出する。

PC 3 は、まず、手の中心座標 O ( x , y , z ) の点を第 1 の点とする。PC 3 はこの  
点を中心に手の幅を検出する。

PC 3 は次に、中指の方向に、手の中心座標 O ( x , y , z ) から手の幅の分離れた点  
を検出する。

PC 3 は次に、手の中心座標 O ( x , y , z ) を中心に、中指の方向の点を + 4 5 度  
の点を第 2 の点とし、その座標を R 1 ( R<sub>1</sub> x , R<sub>1</sub> y , R<sub>1</sub> z ) とする。

また、PC 3 は、手の中心座標 O ( x , y , z ) を中心に、中指の方向の点を - 4 5 度

50

の点を第3の点とし、その座標を  $R_2(x, R_2 y, R_2 z)$  とする。

以上の3点を用いて、下記式(5)、(6)、(7)を用いて、手5の法線ベクトル( $H$ )を算出する。

【0078】

【数3】

$$\vec{R}_1 = (R_1x - x, R_1y - y, R_1z - z) \quad \dots(5)$$

$$\vec{R}_2 = (R_2x - x, R_2y - y, R_2z - z) \quad \dots(6)$$

10

$$\vec{H} = \vec{R}_1 \times \vec{R}_2 \quad \dots(7)$$

【0079】

次に、PC3は、手の法線ベクトル( $H$ )と、z軸ベクトル( $Z$ )の法線ベクトル( $V$ )を、下記式(8)に従って算出する。

【0080】

【数4】

$$\vec{V} = \vec{H} \times \vec{Z} \quad \dots(8)$$

20

【0081】

PC3は、算出した法線ベクトル( $V$ )を、平面軸( $xy$ 軸)の回転操作の回転軸とする。

次に、PC3は、手の法線ベクトル( $H$ )とz軸ベクトルとの間の角度  $R$  を下記式(9)を用いて算出する。

この角度  $R$  を平面軸( $xy$ 軸)回転操作の回転角度とする。

30

【0082】

【数5】

$$R\theta = \cos^{-1} \frac{\vec{H} \cdot \vec{Z}}{|\vec{H}| |\vec{Z}|} \quad \dots(9)$$

【0083】

PC3は、手の法線ベクトルによって算出した回転軸( $V$ )、回転角度  $R$  とを下記式(10)に代入して回転行列  $M$  を算出する。

40

【0084】

【数 6】

$$M = \begin{pmatrix} V_x^2(1 - \cos R\theta) + \cos R\theta & V_x V_y(1 - \cos R\theta) + V_z \sin R\theta & V_x V_z(1 - \cos R\theta) - V_y \sin R\theta \\ V_x V_y(1 - \cos R\theta) - V_z \sin R\theta & V_y^2(1 - \cos R\theta) + \cos R\theta & V_y V_z(1 - \cos R\theta) + V_x \sin R\theta \\ V_x V_z(1 - \cos R\theta) + V_y \sin R\theta & V_y V_z(1 - \cos R\theta) - V_x \sin R\theta & V_z^2(1 - \cos R\theta) + \cos R\theta \end{pmatrix}$$

---(10)

【0085】

10

PC3は、上記算出した回転行列Mを用いてスクリーン4の投射されている3Dオブジェクト(3次元画像)の平面軸(x,y軸)の回転操作を行う。

【0086】

ステップS49、表示

PC3は、上記回転操作の結果に基づく画像データをプロジェクタ2に出力する。

プロジェクタ2は、その画像データをスクリーン4に3次元画像(3Dオブジェクト)として投影する。

【0087】

ステップS48、3Dオブジェクトの任意軸方向の拡大・縮小

図20は両手の中心座標を示す図である。

20

両手がゲーの場合(ステップS43)、PC3は、両手間のベクトルと距離に応じて、3Dオブジェクトの任意軸方向の拡大・縮小操作を行う。

【0088】

(1) 両手の座標の検出

PC3は、まず、図20に例示した、両手の中心座標を検出する。右手の中心座標S1を(S1x, S1y, S1z)とし、左手の中心座標S2を(S2x, S2y, S2z)とする。

【0089】

(2) 両手間のベクトルと距離の算出

PC3は、式(11)、(12)を用いて、両手間のベクトル(u)と、距離λを算出する。

30

【0090】

【数7】

$$\vec{u} = (S_1x - S_2x, S_1y - S_2y, S_1z - S_2z) \quad \dots(11)$$

$$\lambda = \sqrt{(S_1x - S_2x)^2 + (S_1y - S_2y)^2 + (S_1z - S_2z)^2} \quad \dots(12)$$

40

【0091】

(3) 拡大・縮小行列の算出

PC3は次に、式(13)を用いて、任意軸方向の拡大・縮小行列Tを算出する。

【0092】

【数 8】

$$T = (\lambda - 1) \begin{bmatrix} u_1 u_1 & u_1 u_2 & u_1 u_3 \\ u_2 u_1 & u_2 u_2 & u_2 u_3 \\ u_3 u_1 & u_3 u_2 & u_3 u_3 \end{bmatrix} + \begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix} \dots (13)$$

【0093】

PC3は、上記算出した任意軸方向の拡大・縮小行列Tを用いて、スクリーン4に投影されている3Dオブジェクト(3次元画像)に対して、任意軸方向の拡大・縮小処理を行う。

10

【0094】

ステップS49、表示

PC3は、上記任意軸方向の拡大・縮小操作の結果に基づく、3Dオブジェクトの画像データをプロジェクタ2に出力する。

プロジェクタ2は、その画像データを3次元画像として、スクリーン4に投影する。

【0095】

ステップS5、終了判定

PC3は、ユーザが終了指示をするまで、上述したステップS2～ステップS4の処理を反復する。

20

【0096】

ステップS6、終了

PC3は、ユーザが終了指示をしたとき、上記処理を終了する。

【0097】

上記実施の形態においては、スクリーン4に投影された3次元画像に対する3次元的な操作、たとえば、フリック、ピンチイン、ピンチアウトを含み、回転、移動などを指示する手5を用いる。これにより、ユーザが手5を用いて3次元画像に直接的に操作を行なうことが可能となる。

【0098】

また上記実施の形態本発明によれば、手5操作指示内容を、たとえば、Kinectなどの3次元撮像手段(3次元形状取得手段)を用いて撮像し、撮像結果に基づく3次元操作手段の色信号および深度信号から3次元操作手段(たとえば、手)の位置、姿勢などを分析して、その分析結果に基づいて3次元画像に対して信号処理して、実時間で表示するので、直接的かつインタラクティブな操作が可能となる。

30

【0099】

好ましくは、本実施の形態によれば、スクリーン4に投影された3次元画像に対して3次元的な画像処理された画像として、プロジェクタなどの投射装置からスクリーンに投影(投射)可能としており、任意の大きさの3次元画像を多数の聴衆に同時に提示することができる。

【0100】

さらに、本実施の形態によれば、たとえば、ミーティングなどにおいて、一人または複数のユーザ(聴衆)が、直観的な操作で、たとえば、インタラクティブな3次元画像の画像処理を行なうことができる。

40

【0101】

3次元形状取得装置、または、撮像手段としてのKinect(1)は、既に市販されており、実時間で、撮像結果をRGB信号および深度情報として提供するので、3次元形状取得装置として新規に設計することが不要である。

また、プロジェクタ2も既存のものを使用することができる。

したがって、以上の実施の形態においては、上述した処理を行うパーソナル・コンピュータ(PC)3のみを新規に開発すればよく、実施で容易であった。

50



## 【0102】

スクリーン4と、Kinect(1)またはプロジェクタ2との間隔を適宜設定することにより、手移動可能領域に挿入する手5の数を任意にすることができる。その結果、大勢の聴衆の前で、多数のユーザがプロジェクタ2からの投影画面を、手を用いて、フリック、ピンチイン、ピンチアウトなどの操作をすることにより、ユーザが直観的に操作することができる。

## 【0103】

さらに、プロジェクタ2から投影される画面の寸法が大きいため、複数のユーザが同時に投影画面上の表示を操作することが可能となり、たとえば、複数の人間によるミーティングなどにおいても、手などの直観的な操作で、PC3を介して、所望の画像処理を行うことが可能となる。

10

## 【0104】

以上の実施の形態においては、3次元形状取得装置、または、撮像手段として、Kinect(1)を用いた場合について述べたが、Kinect(1)に代えて、Kinect(1)と同等の機能、性能を有する撮像手段を用いることができる。

## 【0105】

第2実施の形態

また、上記実施の形態において、PC3からの映像を3次元画像として表示する表示手段として、プロジェクタ2を用いた場合について述べたが、その表示手段をPC3の表示画面、または、他のコンピュータの表示画面とすることもできる。PC3の表示画面、または、他のコンピュータの表示画面は、プロジェクタ2からの投射画像とは異なるが、PC3のユーザ、または、他の聴衆に視認することができる。

20

なお、ユーザは、PC3の表示画面の前に規定される手移動可能領域、または、他のコンピュータの表示画面の前に規定される手移動可能領域において、手5による操作を行う。

したがって、本発明の表示手段としては、プロジェクタ2には限定されない。

## 【0106】

本発明の実施の形態に際しては、上述した例示に限定されず、種々の変形態様を取ることができる。

## 【符号の説明】

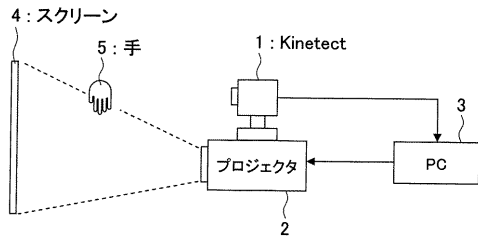
30

## 【0107】

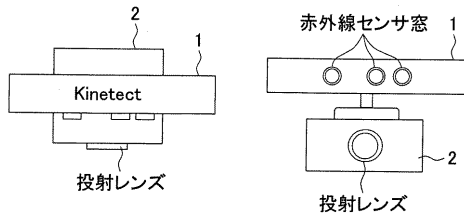
10...画像処理装置、1...Kinect、2...プロジェクタ、パーソナル・コンピュータ、4...スクリーン、5...手。

【 図 1 】

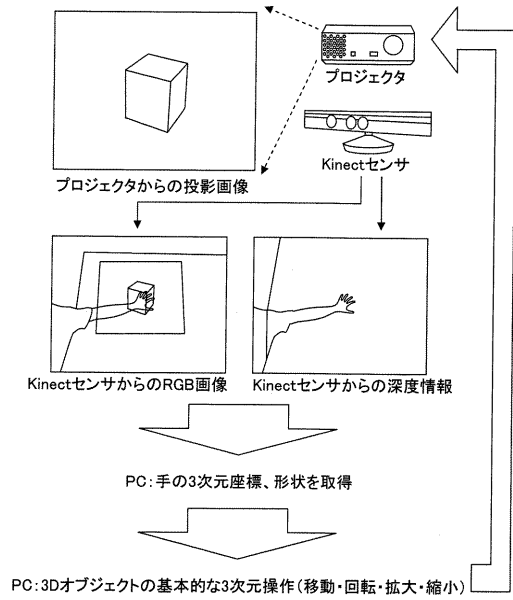
(A)



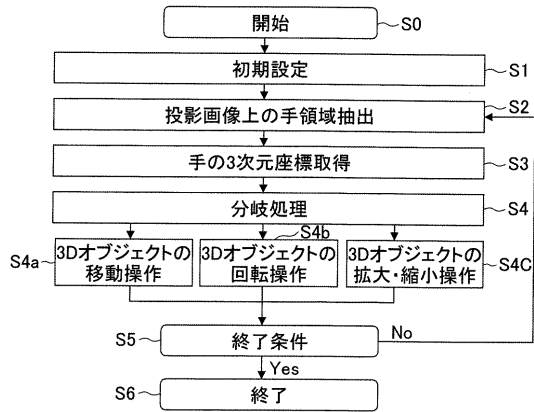
(B)



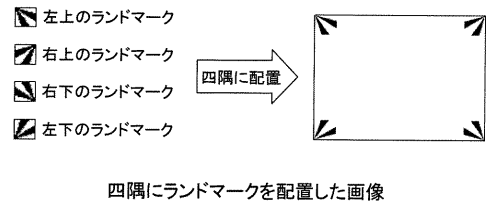
【 図 2 】



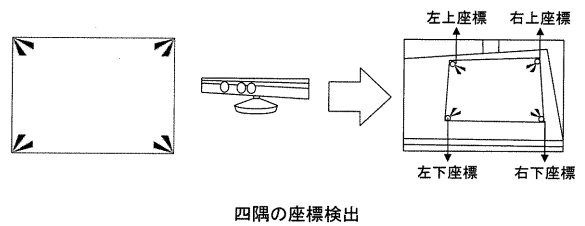
【 図 3 】



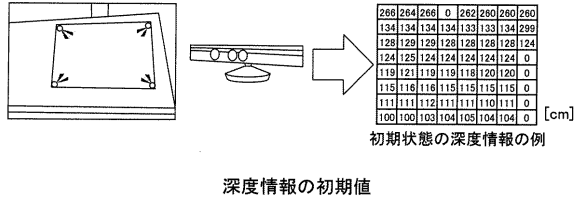
【 図 4 】



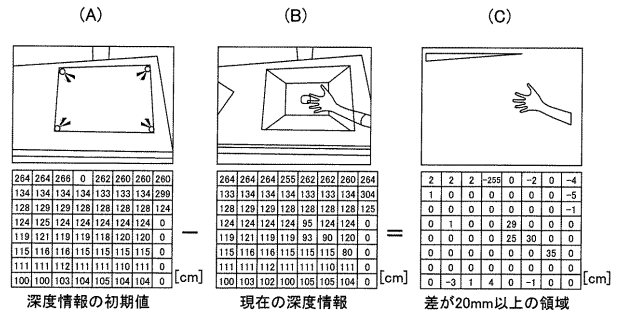
【 図 5 】



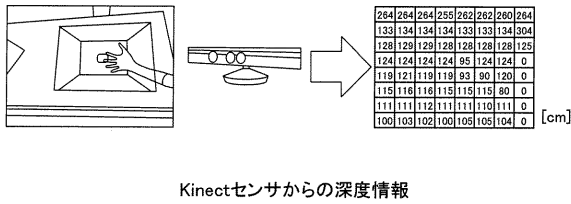
【 図 6 】



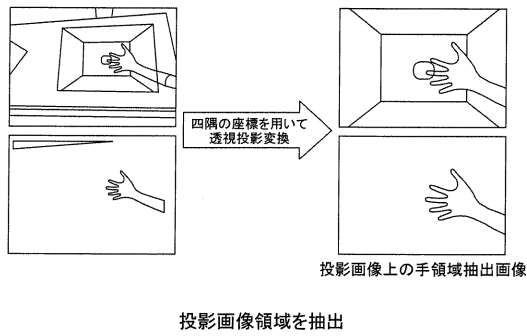
【 図 8 】



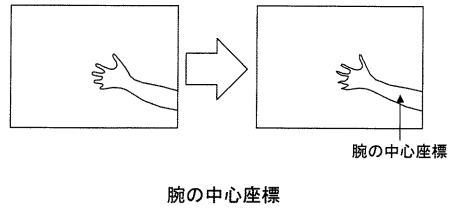
【 図 7 】



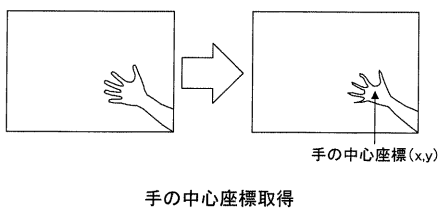
【 図 9 】



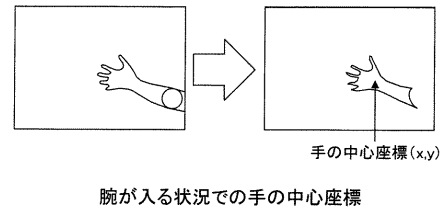
【 図 1 1 】



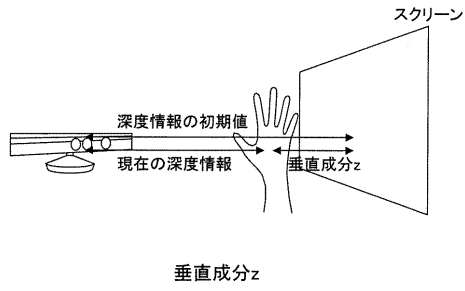
【 図 1 0 】



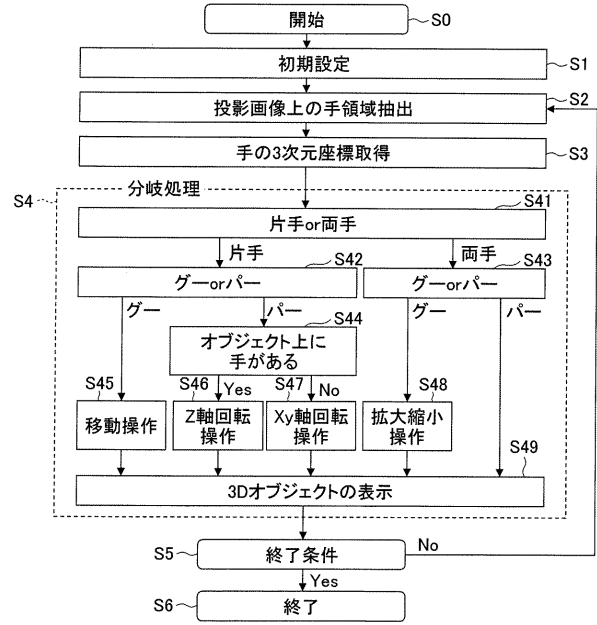
【 図 1 2 】



【図13】

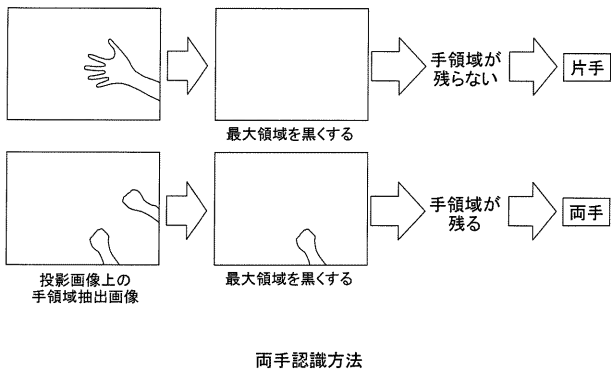


【図14】



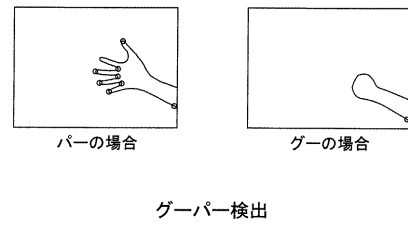
移動・回転・拡大・縮小操作の分岐処理

【図15】



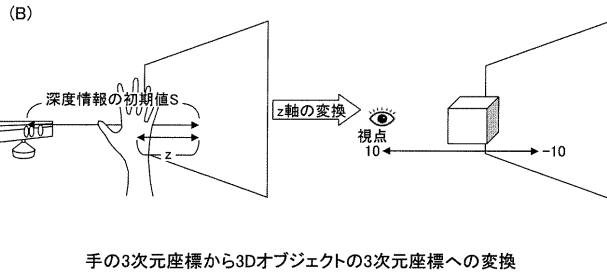
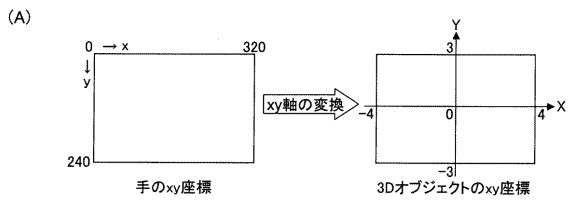
両手認識方法

【図16】

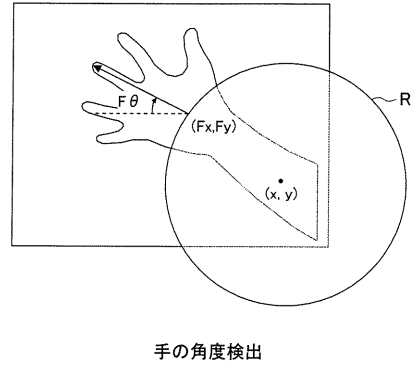


ゲーパー検出

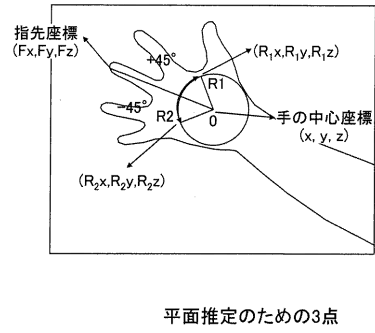
【 図 1 7 】



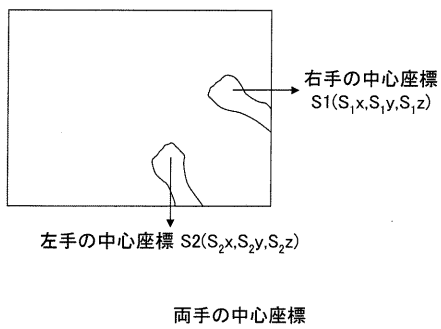
【 図 1 8 】



【 図 1 9 】



【 図 2 0 】



---

フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 3 B 21/00 D  
G 0 6 F 3/033 4 2 2

Fターム(参考) 5E555 AA01 BA29 BA86 BB02 BC18 BE16 CA29 CA42 CB46 CB66  
CC22 CC24 CC25 CC26 CC27 DA03 DA11 DB56 DC19 DC24  
DC26 DC27 DC30 DC43 EA22 FA30