

(19) 日本国特許庁(JP)

## (12) 特許公報(B2)

(11) 特許番号

特許第6253981号  
(P6253981)

(45) 発行日 平成29年12月27日(2017.12.27)

(24) 登録日 平成29年12月8日(2017.12.8)

(51) Int.Cl.	F 1
HO4N 13/02 (2006.01)	HO4N 13/02
HO4N 5/225 (2006.01)	HO4N 5/225
HO4N 5/232 (2006.01)	HO4N 5/232

請求項の数 38 (全 25 頁)

(21) 出願番号	特願2013-524182 (P2013-524182)
(86) (22) 出願日	平成23年8月9日(2011.8.9)
(65) 公表番号	特表2013-539273 (P2013-539273A)
(43) 公表日	平成25年10月17日(2013.10.17)
(86) 国際出願番号	PCT/US2011/047126
(87) 国際公開番号	W02012/021541
(87) 国際公開日	平成24年2月16日(2012.2.16)
審査請求日	平成25年4月8日(2013.4.8)
審判番号	不服2016-1258 (P2016-1258/J1)
審判請求日	平成28年1月28日(2016.1.28)
(31) 優先権主張番号	13/205,481
(32) 優先日	平成23年8月8日(2011.8.8)
(33) 優先権主張国	米国(US)
(31) 優先権主張番号	61/371,979
(32) 優先日	平成22年8月9日(2010.8.9)
(33) 優先権主張国	米国(US)

(73) 特許権者	595020643 クアルコム・インコーポレイテッド QUALCOMM INCORPORATED アメリカ合衆国、カリフォルニア州 92121-1714、サン・ディエゴ、モアハウス・ドライブ 5775
(74) 代理人	100108855 弁理士 蔵田 昌俊
(74) 代理人	100109830 弁理士 福原 淑弘
(74) 代理人	100158805 弁理士 井関 守三
(74) 代理人	100194814 弁理士 奥村 元宏

最終頁に続く

(54) 【発明の名称】立体カメラのためのオートフォーカス

## (57) 【特許請求の範囲】

## 【請求項 1】

第1の視点に関連付けられた第1の画像を受信することと、  
第2の視点に関連付けられた第2の画像を受信することと、

前記第1の画像に基づいて第1の複数のキーポイントを決定することと、ここで、前記第1の複数のキーポイントは、前記第1の画像の固有な一部を識別し、前記第1の複数のキーポイントの数は、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない、

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることと、

疎な深度マップを形成するために、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定することと、

前記複数の視差の平均、前記第1の視点の位置、および前記第2の視点の位置に基づいて焦点深度を決定すること

を含む、画像センサのための焦点深度を決定するための立体キャプチャ・デバイスを有する電子デバイスにおける方法。

## 【請求項 2】

前記第2の画像に基づいて第2の複数のキーポイントを決定することをさらに含み、前記第2の複数のキーポイントは、前記第2の画像の一部を識別し、前記第2の複数のキ-

ポイントの数は、前記識別された前記第2の画像の一部に含まれるピクセルの数よりも少ない、請求項1に記載の方法。

【請求項3】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の複数のキーポイントからのキーポイントを前記第2の複数のキーポイントからのキーポイントと相関させることを含む、請求項2に記載の方法。

【請求項4】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第2の画像における検索範囲内のピクセルを介して繰り返すことを含む、請求項1に記載の方法。

10

【請求項5】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の画像および前記第2の画像におけるピクセル間の平均2乗誤差を決定することを含む、請求項4に記載の方法。

【請求項6】

前記第1の画像に基づいて第1の複数のキーポイントを決定することは、前記第1の画像に基づいてスケール不变特徴変換(SIFT)キーポイントを決定することを含む、請求項1に記載の方法。

【請求項7】

前記第1の画像に基づいて第1の複数のキーポイントを決定することは、前記第1の画像をサブサンプリングすることと、前記サブサンプリングされた画像に高域フィルタを適用することと、前記サブサンプリングされた画像力の値を計算することと、前記サブサンプリングされた画像を閾値処理することを含む、請求項1に記載の方法。

20

【請求項8】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、リアルタイムで生じる、請求項1に記載の方法。

【請求項9】

前記電子デバイスは、携帯電話を含む、請求項1に記載の方法。

【請求項10】

立体キャプチャ・デバイスに動作可能に接続されたコンピュータに、

30

第1の視点に関連付けられた第1の画像を受信することと、

第2の視点に関連付けられた第2の画像を受信することと、

前記第1の画像に基づいて第1の複数のキーポイントを決定することと、ここで、前記第1の複数のキーポイントは、前記第1の画像の固有な一部を識別し、前記第1の複数のキーポイントの数は、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない、

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることと、

疎な深度マップを形成するために、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定することと、

40

前記複数の視差の平均、前記第1の視点の位置、および前記第2の視点の位置に基づいて焦点深度を決定すること

のステップを行わせるように構成される命令を含む、コンピュータ可読記憶媒体。

【請求項11】

前記第2の画像に基づいて第2の複数のキーポイントを決定することをさらに含み、前記第2の複数のキーポイントは、前記第2の画像の一部を識別し、前記第2の複数のキーポイントの数は、前記識別された前記第2の画像の一部に含まれるピクセルの数よりも少ない、請求項10に記載のコンピュータ可読記憶媒体。

【請求項12】

50

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の複数のキーポイントからのキーポイントを前記第2の複数のキーポイントからのキーポイントと相関させることを含む、請求項11に記載のコンピュータ可読記憶媒体。

【請求項13】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第2の画像における検索範囲内のピクセルを介して繰り返すことを含む、請求項10に記載のコンピュータ可読記憶媒体。

【請求項14】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の画像および前記第2の画像におけるピクセル間の平均2乗誤差を決定することを含む、請求項13に記載のコンピュータ可読記憶媒体。 10

【請求項15】

前記第1の画像に基づいて第1の複数のキーポイントを決定することは、前記第1の画像に基づいてスケール不变特徴変換(SIFT)キーポイントを決定することを含む、請求項10に記載のコンピュータ可読記憶媒体。

【請求項16】

前記第1の画像に基づいて第1の複数のキーポイントを決定することは、前記第1の画像をサブサンプリングすることと、前記サブサンプリングされた画像に高域フィルタを適用することと、前記サブサンプリングされた画像力の値を計算することと、前記サブサンプリングされた画像を閾値処理することを含む、請求項10に記載のコンピュータ可読記憶媒体。 20

【請求項17】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、リアルタイムで生じる、請求項10に記載のコンピュータ可読記憶媒体。

【請求項18】

前記コンピュータは、携帯電話に位置する、請求項10に記載のコンピュータ可読記憶媒体。

【請求項19】

立体キャプチャ・デバイスの焦点を合わせるためのシステムであって、前記システムは、  
第1の視点に関連付けられた第1の画像を生成するように構成される第1の画像センサと、

第2の視点に関連付けられた第2の画像を生成するように構成される第2の画像センサと、

前記第1の画像に基づいて第1の複数のキーポイントを決定するように構成される特徴生成モジュールと、ここで、前記第1の複数のキーポイントは、前記第1の画像の固有な一部を識別し、前記第1の複数のキーポイントの数は、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない。 40

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させるように構成されるキーポイント相関性モジュールと、

疎な深度マップを形成するために、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定するように構成される視差決定モジュールと、

前記複数の視差の平均、前記第1の視点の位置、および前記第2の視点の位置に基づいて焦点深度を決定するように構成される深度決定モジュール

を含む、システム。

【請求項20】

前記特徴生成モジュールは、前記第2の画像に基づいて第2の複数のキーポイントを決

10

20

30

40

50

定するように構成され、前記第2の複数のキーポイントは、前記第2の画像の一部を識別し、前記第2の複数のキーポイントの数は、前記識別された前記第2の画像の一部に含まれるピクセルの数よりも少ない、請求項19に記載のシステム。

【請求項21】

キーポイントを相関させるように構成される前記モジュールは、前記第1の複数のキーポイントからのキーポイントを前記第2の複数のキーポイントからのキーポイントと相関させるように構成される、請求項20に記載のシステム。

【請求項22】

キーポイントを相関させるように構成される前記モジュールは、前記第2の画像における検索範囲内のピクセルを介して繰り返すように構成される、請求項19に記載のシステム。

10

【請求項23】

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の画像および前記第2の画像におけるピクセル間の平均2乗誤差を決定することを含む、請求項22に記載のシステム。

【請求項24】

前記特徴生成モジュールは、前記第1の画像に基づいてスケール不变特徴変換(SIFT)キーポイントを決定するように構成される、請求項19に記載のシステム。

【請求項25】

前記特徴生成モジュールは、前記第1の画像をサブサンプリングし、前記サブサンプリングされた画像に高域フィルタを適用し、前記サブサンプリングされた画像の値を計算し、前記サブサンプリングされた画像を閾値処理するように構成される、請求項19に記載のシステム。

20

【請求項26】

キーポイントを相関させるように構成される前記モジュールは、前記第1の複数のキーポイントからのキーポイントをリアルタイムで前記第2の画像における位置と相関させる、請求項19に記載のシステム。

【請求項27】

前記立体キャプチャ・デバイスは、携帯電話に位置する、請求項19に記載のシステム。

30

【請求項28】

焦点深度を決定するように構成される前記モジュールは、視差ヒストグラムを含む、請求項19に記載のシステム。

【請求項29】

立体キャプチャ・デバイスの焦点を合わせるためのシステムであって、前記システムは

、  
第1の視点に関連付けられた第1の画像を受信するための手段と、

第2の視点に関連付けられた第2の画像を受信するための手段と、

前記第1の画像に基づいて第1の複数のキーポイントを決定するための手段と、ここで、前記第1の複数のキーポイントは、前記第1の画像の固有な一部を識別し、前記第1の複数のキーポイントの数は、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない、

40

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させるための手段と、

疎な深度マップを形成するために、前記識別された前記第1の画像の固有な一部に含まれるピクセルの数よりも少ない前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定するための手段と、

前記複数の視差の平均、前記第1の視点の位置、および前記第2の視点の位置に基づいて焦点深度を決定するための手段

を含む、システム。

50

**【請求項 3 0】**

第1の画像を受信するための前記手段は、第1のセンサを含み、第2の画像を受信するための前記手段は、第2のセンサを含み、第1の複数のキーポイントを決定するための前記手段は、特徴生成モジュールを含み、相関させるための前記手段は、キーポイント相関性モジュールを含み、複数の視差を決定するための前記手段は、視差決定モジュールを含み、焦点深度を決定するための前記手段は、深度決定モジュールを含む、請求項29に記載のシステム。

**【請求項 3 1】**

第1の複数のキーポイントを決定するための前記手段は、前記第2の画像に基づいて第2の複数のキーポイントを決定するように構成され、前記第2の複数のキーポイントは、前記第2の画像の一部を識別し、前記第2の複数のキーポイントの数は、前記識別された前記第2の画像の一部に含まれるピクセルの数よりも少ない、請求項29に記載のシステム。  
10

**【請求項 3 2】**

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させるための前記手段は、前記第1の複数のキーポイントからのキーポイントを前記第2の複数のキーポイントからのキーポイントと相関させるように構成される、請求項31に記載のシステム。

**【請求項 3 3】**

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させるための前記手段は、前記第2の画像における検索範囲内のピクセルを介して繰り返すように構成される、請求項29に記載のシステム。  
20

**【請求項 3 4】**

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させるための前記手段は、前記第1の画像および前記第2の画像におけるピクセル間の平均2乗誤差を決定するように構成される、請求項33に記載のシステム。

**【請求項 3 5】**

第1の複数のキーポイントを決定するための前記手段は、前記第1の画像に基づいてスケール不变特徴変換（SIFT）キーポイントを決定するように構成される、請求項29に記載のシステム。  
30

**【請求項 3 6】**

第1の複数のキーポイントを決定するための前記手段は、前記第1の画像をサブサンプリングし、前記サブサンプリングされた画像に高域フィルタを適用し、前記サブサンプリングされた画像力の値を計算し、前記サブサンプリングされた画像を閾値処理するように構成される、請求項29に記載のシステム。

**【請求項 3 7】**

キーポイントを相関させるための前記手段は、前記第1の複数のキーポイントからの前記キーポイントをリアルタイムで前記第2の画像における位置と相関させる、請求項29に記載のシステム。

**【請求項 3 8】**

前記立体キャプチャ・デバイスは、携帯電話に位置する、請求項29に記載のシステム。  
40

**【発明の詳細な説明】****【関連出願に対する相互参照】****【0001】**

本願は、2011年3月23日に出願された、「立体画像のためのオートフォーカス（AUTOFOCUS FOR STEREO IMAGES）」と題された、米国特許仮出願番号第61/489,231号、および、2010年8月9日に出願された、「立体画像からの瞬時オートフォーカス（INSTANTANEOUS AUTOFOCUS FROM STEREO IMAGES）」と題された、米国特許仮出願番号第61/371,979号、の利益を主張する、2011年8月8日に出願された、米  
50

国特許出願第13/205,481号に対する優先権を主張し、それらの出願はここでの参照により組み込まれる。

【技術分野】

【0002】

本実施形態は、立体画像キャプチャ (stereoscopic image capture) に関し、特に、立体画像キャプチャ・デバイスに適した焦点深度 (focal depth) を決定するための方法、装置、およびシステムに関する。

【背景技術】

【0003】

立体視 (Stereopsis) は、人間の脳が左目および右目から見られるようなオブジェクトの相対的変位 (relative displacement) に基づいて、オブジェクトの深度を解釈する処理である。立体感 (stereoscopic effect) は、第1および第2の横方向オフセットの視聴位置 (laterally offset viewing positions) から第1および第2のシーンの画像をとることによって、および、左および右目の各々に画像を別々に表示することによって、人為的に引き起こされ得る。時間内に一連の立体画像ペア (stereoscopic image pairs) をキャプチャすることによって、画像ペアは、ユーザに3次元を有しているように見える立体ムービーを形成するように、目に連続的に表示され得る。

【0004】

2つの従来のカメラは、立体画像ペアの画像の各々を取得するために使用されることがある。従来のカメラは、異なる焦点深度で複数の画像をキャプチャするオートフォーカス手順を使用して適切に焦点を合わせることができる。最も高い頻度容量 (the highest frequency content) に対応する焦点深度は、次に、次の画像キャプチャのために使用される。従来のムービー・カメラは、ビデオ・キャプチャの間、オートフォーカスするために、この方法を使用することができる。しかしながら、オートフォーカス機能が行われる間、フレームのキャプチャは、定期的に遅らせられる必要があるだろう。このオートフォーカス技術は、従来の單一カメラで2D画像をキャプチャすることに適しているが、立体画像キャプチャに対しては不適当であり得る。特に、前記技術は、ビデオ・ストリームを中断し、カメラの動き、たとえば、ユーザの手の動き、によって影響を受け得る。

【発明の概要】

【0005】

ある特定の実施形態は、画像センサのための焦点深度を決定するための電子デバイスにおける方法について考察する。それら方法は、第1の視点に関連付けられた第1の画像を受信することと、第2の視点に関連付けられた第2の画像を受信することと、第1の画像に基づいて第1の複数のキーポイントを決定することと、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることと、第1の複数のキーポイントの各々に関連付けられた複数の視差 (disparities) を決定することと、複数の視差、第1の視点の位置、および第2の視点の位置に基づいて焦点深度を決定することを含むことができる。

【0006】

ある特定の実施形態において、前記方法は、第2の画像に基づいて第2の複数のキーポイントを決定することをさらに含むことができる。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、第1の複数のキーポイントからのキーポイントを第2の画像におけるキーポイントを第2の複数のキーポイントからのキーポイントと相關させることを含む。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、第2の画像における検索範囲内のピクセルを介して繰り返すことを含む。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、第1の画像および第2の画像におけるピクセル間の平均2乗誤差を決定することを含む。いくつかの実施形態において、第1の画像に基づいて第1の複数のキーポイントを決定することは、第1の画像に基づいてスケール不变特徴変換 (SIFT : Scale Invariant

10

20

30

40

50

t Feature Transform) キーポイントを決定することを含む。いくつかの実施形態において、第1の画像に基づいて第1の複数のキーポイントを決定することは、第1の画像をサブサンプリングすることと、第1の画像に高域フィルタ (high-pass filter) を適用することと、第1の画像力 (power of the first image) を計算することと、第1の画像を閾値処理することを含む。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、リアルタイムで生じる。いくつかの実施形態において、電子デバイスは、携帯電話を含む。

#### 【0007】

ある特定の実施形態は、コンピュータに、第1の視点に関連付けられた第1の画像を受信し、第2の視点に関連付けられた第2の画像を受信し、第1の画像に基づいて第1の複数のキーポイントを決定し、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させ、第1の複数のキーポイントの各々に関連付けられた複数の視差を決定し、複数の視差、第1の視点の位置、および第2の視点の位置に基づいて焦点深度を決定するステップを行わせるように構成される命令を含むコンピュータ可読媒体について考察する。

10

#### 【0008】

いくつかの実施形態において、前記命令は、また、プロセッサに、第2の画像に基づいて第2の複数のキーポイントを決定させるように構成される。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、第1の複数のキーポイントからのキーポイントを第2の複数のキーポイントからのキーポイントと相關させることを含む。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、第2の画像における検索範囲内のピクセルを介して繰り返すことを含む。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、第1の画像および第2の画像におけるピクセル間の平均2乗誤差を決定することを含む。いくつかの実施形態において、第1の画像に基づいて第1の複数のキーポイントを決定することは、第1の画像に基づいてスケール不变特徴変換 (SIFT) キーポイントを決定することを含む。いくつかの実施形態において、第1の画像に基づいて第1の複数のキーポイントを決定することは、第1の画像をサブサンプリングすることと、第1の画像に高域フィルタ (high-pass filter) を適用することと、第1の画像力を計算することと、第1の画像を閾値処理することを含む。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させることは、リアルタイムで生じる。いくつかの実施形態において、コンピュータは、携帯電話に位置する。

20

#### 【0009】

ある特定の実施形態は、立体キャプチャ・デバイスの焦点を合わせるためのシステムについて考察する。前記システムは、第1の視点に関連付けられた第1の画像を生成するように構成される第1の画像センサと、第2の視点に関連付けられた第2の画像を生成するように構成される第2の画像センサと、第1の画像に基づいて第1の複数のキーポイントを決定するように構成される特徴生成モジュールと、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相關させるように構成されるキーポイント相關性モジュールと、第1の複数のキーポイントの各々に関連付けられた複数の視差を決定するように構成される視差決定モジュールと、複数の視差、第1の視点の位置、および第2の視点の位置に基づいて焦点深度を決定するように構成される深度決定モジュールを含むことができる。

30

#### 【0010】

いくつかの実施形態において、特徴生成モジュールは、第2の画像に基づいて第2の複数のキーポイントを決定するように構成され得る。いくつかの実施形態において、キーポイントを相關させるように構成されるソフトウェア・モジュールは、第1の複数のキーポイントからのキーポイントを第2の複数のキーポイントからのキーポイントと相關させる

40

50

ように構成される。いくつかの実施形態において、キーポイントを相関させるように構成されるソフトウェア・モジュールは、第2の画像における検索範囲内のピクセルを介して繰り返すように構成される。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相関させることは、第1の画像および第2の画像におけるピクセル間の平均2乗誤差を決定することを含む。

#### 【0011】

いくつかの実施形態において、特徴生成モジュールは、第1の画像に基づいてスケール不变特徴変換（SIFT）キーポイントを決定するように構成される。いくつかの実施形態において、特徴生成モジュールは、第1の画像をサブサンプリングし、第1の画像に高域フィルタを適用し、第1の画像力を計算し、第1の画像を閾値処理するように構成される。10

#### 【0012】

いくつかの実施形態において、キーポイントを相関させるように構成されるソフトウェア・モジュールは、第1の複数のキーポイントからのキーポイントをリアルタイムで前記第2の画像における位置と相関させる。いくつかの実施形態において、立体キャプチャ・デバイスは、携帯電話に位置する。いくつかの実施形態において、焦点深度を決定するように構成されるソフトウェア・モジュールは、視差ヒストグラム（disparity histogram）を含む。

#### 【0013】

ある特定の実施形態は、立体キャプチャ・デバイスの焦点を合わせるためのシステムを考察し、前記システムは、第1の視点に関連付けられた第1の画像を受信するための手段と、第2の視点に関連付けられた第2の画像を受信するための手段と、第1の画像に基づいて第1の複数のキーポイントを決定するための手段と、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相関させるための手段と、第1の複数のキーポイントの各々に関連付けられた複数の視差を決定するための手段と、複数の視差、第1の視点の位置、および第2の視点の位置に基づいて焦点深度を決定するための手段を含む。20

#### 【0014】

いくつかの実施形態において、第1の画像を受信するための手段は、第1のセンサを含み、第2の画像を受信するための手段は、第2のセンサを含み、第1の複数のキーポイントを決定するための手段は、特徴生成モジュールを含み、相関させるための手段は、キーポイント相関性モジュールを含み、複数の視差を決定するための手段は、視差決定モジュールを含み、焦点深度を決定するための手段は、深度決定モジュールを含む。いくつかの実施形態において、第1の複数のキーポイントを決定するための手段は、第2の画像に基づいて第2の複数のキーポイントを決定するように構成される。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相関させるための手段は、第1の複数のキーポイントからのキーポイントを第2の複数のキーポイントからのキーポイントと相関させるように構成される。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相関させるための手段は、第2の画像における検索範囲内のピクセルを介して繰り返すように構成される。いくつかの実施形態において、第1の複数のキーポイントからのキーポイントを第2の画像における位置と相関させるための手段は、第1の画像および第2の画像におけるピクセル間の平均2乗誤差を決定するように構成される。いくつかの実施形態において、第1の複数のキーポイントを決定するための手段は、第1の画像に基づいてスケール不变特徴変換（SIFT）キーポイントを決定するように構成される。いくつかの実施形態において、第1の複数のキーポイントを決定するための手段は、第1の画像をサブサンプリングし、第1の画像に高域フィルタを適用し、第1の画像力を計算し、第1の画像を閾値処理するように構成される。いくつかの実施形態において、キーポイントを相関させるための手段は、第1の複数のキーポイントからのキーポイントをリアルタイムで第2の画像における位置と相関させる。いくつかの実施形態において、立体キャプチャ・デバイスは4050

、携帯電話に位置する。

【図面の簡単な説明】

【0015】

開示される態様は、以下において、同一符号が同一要素を示す添付図面と連携して説明され、開示された観点を限定することなく例示するために提供される。

【図1】図1は、立体画像のキャプチャを容易にするセンサ配置 (sensor arrangement) を含む1つの可能なモバイル・デバイスを図示する一般化されたブロック図である。

【図2】図2は、図1のモバイル・デバイスのような、モバイル・デバイス中のある特定のコンポーネントのブロック図である。

【図3】図3は、キャプチャ・デバイスの立体写真 (stereo pair) を使用して第1および第2の位置でオブジェクトをキャプチャすることを図示する。

10

【図4】図4は、特定のカメラ配置のためのオブジェクトの距離とピクセル視差 (pixel disparity) の関係を図示するグラフである。

【図5A】図5Aは、立体感を得るために位置付けられた2つの画像キャプチャ・センサおよび任意のシーンのトップダウン・ビュー (top-down view) を図示するブロック図である。

【図5B】図5Bは、オーバーレイされた (overlaid) シーンにおけるオブジェクト視差の方向および規模 (magnitude) を備えた図5Aのシーンが撮影された立体画像ペアのうちの1つを図示する。

【図6】図6は、図4のグラフを図示するが、オブジェクト視差ヒストグラムおよびオーバーレイされた対応するオブジェクト深度ヒストグラムを備える。

20

【図7】図7は、ある特定の実施形態が新しい焦点深度を決定する処理のためのフローチャートを図示する。

【図8】図8は、ある特定の実施形態がキーポイントを決定する処理のためのフローチャートを図示する。

【図9】図9は、キーポイントが画像の各々の間と相關している領域および立体画像ペアを図示する。

【詳細な説明】

【0016】

実施形態は、立体カメラにおける構成データを決定または設定するシステムおよび方法に関する。一実施形態において、構成データは、立体カメラの2つのレンズの適切な焦点距離に関する。一実施形態において、第1のカメラはシーンから第1の画像を受信し、第2のカメラは同一のシーンの第2の画像を受信する。キーポイントのセットは、第1の画像の分析から決定されている。キーポイントは、たとえば、画像の一部から一貫して再現されることができる任意のデータ構造を含むことができ、これにより画像部分の固有の識別を許可する。いくつかの実施形態において、キーポイントは、画像の一部に対応する複数のピクセルを含むことができる。キーポイントは、画像における位置に関連付けられる。第1の画像におけるキーポイントを決定した後、システムは、第2の画像において類似した位置を探す。一度、第2の画像において類似した位置が識別されると、システムは、第1の画像におけるキーポイントと、第2の画像における対応する位置との間の違いを計算する。これは、システムに、立体レンズの位置に加えて、両方のフレームにおける同一のキーポイントの位置の間の視差を周知することでシーンの焦点深度を決定することを可能にする。

30

【0017】

本実施形態は、少なくとも立体画像ペアに基づいてセンサに適した焦点深度を決定するためのシステム、装置、および方法について考察する。特に、ある特定の実施形態は、各画像のためのキーポイントを決定することと、キーポイント間の相關性を識別することと、相關性からオブジェクトの距離を引き出すことについて考察する。当業者は、これらの実施形態がハードウェア、ソフトウェア、ファームウェア、またはそれらの任意の組み合せで実現され得ることを理解するだろう。立体視システム (stereoscopic system) は

40

50

、モバイル・ワイヤレス通信デバイス、携帯情報端末（PDA's）、ラップトップ・コンピュータ、デスクトップ・コンピュータ、デジタル・カメラ、デジタル記録デバイス、および同類のものを含む、広範囲の電子デバイスで実現され得る。

【0018】

図1は、立体画像のキャプチャ、および画像を受信するための他の手段を容易にするセンサ配置を含むモバイル・デバイス100を図示する。そのようなデバイスは、携帯電話、携帯情報端末、ゲーミング・デバイス、または同様のものであり得る。デバイス100は、距離dによって分離された第1のセンサ101aおよび第2のセンサ101bを含むことができる。デバイスは、また、ユーザ入力制御102およびディスプレイ103を含むことができる。いくつかの実施形態において、ユーザが立体画像またはムービーをキャプチャするためにデバイス100を手に持つ場合、センサ101aおよび101bは、それらは垂直ではなく、水平オフセットのように位置することができる。

【0019】

この特定のデバイスは2つのセンサ101aおよび101bを図示するが、当業者は、おおよそ2つの画像センサを含む立体画像キャプチャ・デバイスを容易に想像することができる。たとえば、単一センサのみを備えたデバイスは、すばやく連続してセンサ101aおよび101bの位置で2つの画像を得るために、一連のレンズまたは反射面との組み合わせで動作することができる。この配置は、同様に、下記に説明される方法で使用するための立体画像ペアを得ることができ、単一センサはそれに従って焦点を合わせられることができる。従って、本願において議論される方法およびシステムは、これらの視点が画像シーンの立体描写を容易にする限り、第1および第2の視点から2つの画像を得る任意のシステムに適応できるだろう。従って、画像センサのペアへの言及は、2つの視点から画像を受信する単一画像センサの可能性を除外すると考えられるべきではない。

【0020】

図2は、図1に図示されたモバイル・デバイス100のような、モバイル・デバイス中のある特定のコンポーネントのブロック図である。センサ101aは、立体画像ペアの第1の画像を受信し、センサ101bは、立体画像ペアの第2の画像を受信する。いくつかの実施形態において、センサは画像を同時に受信することができる。デバイスは、ビデオ・フロント・エンド102およびメモリ103を含むことができる。ビデオ・フロント・エンド102は、センサ101aおよび101bから入ってくる生データ（incoming raw image data）を処理し、データをメモリ103に記憶することができる。メモリ103は、また、モバイル・デバイス100のためのさまざまなアプリケーションおよびソフトウェア・ドライバを含むことができる。たとえば、ディスプレイ・ドライバ・モジュール104は、ディスプレイ103と通信することができる。ユーザ入力モジュール106は、同様にユーザ・インターフェース102と通信することができる。ワイヤレス通信ドライバ・モジュール107は、ワイヤレス通信ハードウェア112と通信することができる。

【0021】

メモリは、また、汎用プロセッサ113と通信することができる。汎用プロセッサ113は、アドバンストRISCマシン（ARM：Advanced RISC Machine）、デジタル・シグナル・プロセッサ（DSP）、またはグラフィカル処理ユニット（GPU）のような、サブ処理ユニットまたはサブプロセッサを含むことができる。これらのプロセッサは、さまざまな動作を処理するとき、ローカル・メモリと通信することができる。

【0022】

ある特定の本実施形態は、システム・アーキテクチャへの「焦点深度分析モジュール（Focal Depth Analysis Module）」115a、115bの追加について考察する。いくつかの実施形態において、モジュールは、専用プロセッサ115a、または汎用プロセッサに位置するプロセッサの一部の形をとることができる。いくつかの実施形態において、モジュールは、メモリ103のようなコンピュータ可読媒体に記憶されるソフトウェア・コード115bを含むことができる。いくつかの実施形態は、ファームウェア、またはソフトウェア・ハードウェアの組み合わせとしてのメモリ115bおよび専用プロセッサ11

10

20

30

40

50

5 aにおいて、モジュールの一部を設置することができる。いくつかの実施形態において、モジュールは、SIFT特徴生成システムのような特徴生成システム、およびセンサ101aおよび101bへのアクセスを許可する図2において任意の位置に存在することができる。従って、モジュールは、特徴生成および/または検出のために構成される既存のハードウェアまたはソフトウェアを活用することができる。当業者は、下記に説明される実施形態が汎用プロセッサ113におけるサブプロセッサを使用して実現されることができ、メモリ103において分離したアプリケーションとして記憶されることができることを理解するだろう。いくつかの実施形態において、SIFT特徴生成システムがソフトウェアにおいて発見されることがあるのに対して、他の実施形態において、SIFT特徴生成システムはハードウェアにおいて発見されることがある。

10

#### 【0023】

ある特定の本実施形態は、立体画像キャプチャの幾何学的特性 (geometric properties) を活用するオートフォーカス機能を提供する。図3は、トップダウン・ビューを経由して、立体カメラ配置を使用した第1の位置300a、および第2の位置300bにおけるオブジェクト304の立体画像キャプチャを図示する。第1の画像キャプチャ・デバイス301aは、第2の位置に位置する第2のキャプチャ・デバイス301bから横方向に分離される第1の位置に位置することができる。第1のキャプチャ・デバイス301aは、第1の位置から第1のシーンの画像をキャプチャすることができ、第2のキャプチャ・デバイス301bは、第2の位置から第2のシーンの画像をキャプチャすることができる。第1および第2の画像は、それに応じて、キャプチャ・デバイス301aおよび301bの位置、および方位に基づいて、第1および第2のシーンの視点に関連付けられるだろう。オブジェクト304は、両方の画像に現れ得る。いくつかの実施形態において、キャプチャ・デバイス301aおよびキャプチャ・デバイス301bは、図1のセンサ101aおよび101bとそれぞれ同一であり得る。キャプチャ・デバイス301a、301bは、垂直視差(vertical disparity)を有さず、かなり近い焦点距離を所持すると較正され(calibrated)得る。

20

#### 【0024】

デバイス301aの視点の中心は、線302aに沿っていく。同様に、デバイス301bの視点の中心は、線302bに沿っていく。これら2つの中心線は、位置303で交差する。述べたように、オブジェクト304は第1および第2の画像の各々で現れる。しかしながら、位置300aに関して、オブジェクト304は、量(amount)305aによって中心線302aの右に現れ、量305bによって中心線302bの左に現れる。反対に、位置300bにおいて、オブジェクト304は、量306aによって中心線302aの左に現れ、量306bによって中心線302bの右に現れる。この方法において、z方向におけるオブジェクトの相対的位置は、左および右の画像の各々における相対的変位によって反映される。

30

#### 【0025】

オブジェクト視差は、第2の画像におけるオブジェクトの位置と比較して第1の画像におけるオブジェクトの位置との間の違いとして定義され得る。キャプチャ・デバイス間の垂直視差がないところでは、視差は、1つの画像における位置から別の位置への横方向オフセットのみを含むことができる。当業者(one)は、左と右、または右と左の画像の間の違いとして視差を任意にとることができる。本明細書の目的について、視差は、(図3に示されるようなx方向の正(x-direction positive)を備えた)センサ301aからの画像におけるオブジェクトの位置を引いた、センサ301bからの画像におけるオブジェクトの位置として定義される。従って、負の視差(negative disparity)は、位置300aにおけるオブジェクト304の描写によりもたらされ、正の視差は、位置300bにおけるオブジェクト304の描写によりもたらされる。

40

#### 【0026】

センサの位置および相対的方位の知識を用いて、当業者は、カメラ配置から観測される視差、およびオブジェクトの距離、または深度の関係のグラフを構成することができる。

50

たとえば、図 4 は、この 1 つの特定のセンサ配置のための関係のグラフである。視差 401 が増加するにつれて、オブジェクトの距離 402 も増加する。初期の負の視差は、非常にカメラ配置に近いオブジェクトのために存在することができ、すなわち、これらは図 3 の z 方向における深度 402 をほとんど有していない。オブジェクトがカメラ配置からさらに移動する（すなわち、深度が増加する）につれて、視差は次第に正になり、相当な距離でオブジェクトのために横ばい状態となり始める。当業者は、図 4 のチャートがセンサ 301a、301b が垂直である角度に依存し得るということを理解することができる。同様に、センサは図 1 および 3 のように互いに平行であり得るが、センサ間の z および y 方向における変位は、また、グラフへの修正をもたらす。そのようなグラフは、デバイスにおけるメモリ、またはクイック・リファレンス (quick reference) に類似した記憶構造において記憶されることができる。

10

#### 【0027】

図 5 a は、いくつかのオブジェクトを含むシーンのトップダウン・ビューである。また、画像キャプチャ・デバイス 301a および 301b は、画像 501a および 501b をそれぞれ得るために、使用され得る。オブジェクト 502 - 504 は、シーン内のさまざまな深度に位置する。その結果、画像 501a および 501b におけるオブジェクトの位置の間の視差が観測されるだろう。図 5 B は、それらが生じるピクセル位置で示されるある特定の視差の規模および方向を用いて画像 501a を図示する。たとえば、複数の正の視差 510 は、遠くのオブジェクト 504 に対して出現し、複数の負の視差 511 は、より近いオブジェクト 502 に対して現れる。図 4 のグラフを参照して、自動システムは、各視差に関連付けられた深度を決定することができる。たとえば、図 6 に示されるように、同一の規模の視差は、視差ヒストグラム 601 を形成するために、蓄積され、プロット (plotted) されてきた。深度ヒストグラム 602 を生成するために、対応する深度 602 は、センサの関係から導出されることができる。深度ヒストグラム 602 は、示された深度で各最大の領域における 1 つ以上のオブジェクトの存在を示唆するだろう。

20

#### 【0028】

オートフォーカス動作は、センサ 101a、101b のうちの 1 つまたは両方に適した焦点深度の決定を含む。いくつかの実施形態において、適切な焦点深度は、オブジェクト深度ヒストグラム (object depth histogram) (または図 4 のようなグラフに関連したオブジェクト視差ヒストグラム) の平均値 (mean)、中央値 (median)、または類似した統計値をとることによって決定されることができる。特殊な順序統計フィルタ (special order statistic filter) が特定のアプリケーションを収容するために使用されることができる一方で、中央統計値は、かけ離れた値 (outlying value) に対していくつかのロバスト性 (robustness) を提供することができる。選択された統計値は、非常に遠くの、および非常に近くのオブジェクトが与えられる相対的な重みに依存し得る。たとえば、焦点の質は、1 つの深度の範囲を通しておおよそ同じであり得るが、第 2 の範囲において劇的に変化する。これらの変化は、下記でより詳しく議論されている。

30

#### 【0029】

当業者は、第 1 および第 2 の画像の各々において発見されたオブジェクトごとのすべてのピクセルの視差を決定することによって視差ヒストグラム 601 を生成することができるが、これは、モバイル・デバイスでは計算的に費用がかかり、実用的ではない。すべてのピクセルの相関性が相当な数のピクセルを通して繰り返すことを必要とするだけでなく、各画像は、個別のピクセルの識別および各画像におけるオブジェクトへのその相関性を難解 (difficult) にする、同一の値の複数のピクセルを含み得る。

40

#### 【0030】

すべてのピクセルを分析する代わりに、ある特定の本実施形態は、画像内容の「疎な」対応する深度マップまたは「疎な」視差マップを作り出することを考察する。ある特定の実施形態において、キーポイントは、画像の各々において決定されることができ、画像におけるピクセルのすべてまたは大部分の間ではなく、キーポイント間、またはキーポイントとピクセル間の視差が、オブジェクトの深度を推測するために使用されることができる

50

。ピクセルよりも少ないキーポイントがあるので、結果として生じる視差または深度のマップは「疎」である。キーポイントは、画像の一部から一貫して再現することができ、これにより画像部分の固有の識別を許可する任意のデータ構造を含むことができる。キーポイントは、画像における位置に関連付けられ得る。キーポイントの固有の決定は、第2の画像に類似した、または同一部分から識別されるべきキーポイントを許可する。いくつかの実施形態において、キーポイントは、スケール不变特徴変換 (SIFT) キーポイント、または、類似した特徴生成モジュールのキーポイントを含むことができる。いくつかの実施形態において、システムは、キーポイントを決定するためのサブプロセッサまたは汎用プロセッサ 113 において前から存在しているマシン・ビジョン・コンポーネント (machine vision components) を再利用することができる。たとえば、高域フィルタ・ブロックがキーポイント検出のために再利用されることができる。あるいは、メモリ 103 に記憶されるマシン・ビジョン動作を行うためのソフトウェア・ライブラリは、キーポイントを生成するために使用されることができる。この方法において、ある特定の実現は、オートフォーカスを行うためのキーポイントを生成するために他のアプリケーションに関連付けられた機能を無駄なく活用することができる。SIFT 以外のアルゴリズムを適用する特徴生成モジュールのような複数のキーポイントを決定するための代替の手段が、下記でより詳しく説明されている。10

#### 【0031】

図 7 は、立体画像キャプチャのためのオートフォーカス処理 700 を図示するフローチャートであり、それは、ある特定の本実施形態によって実現されることができる。処理は、少なくとも立体画像ペアを得ること、または受信すること 702 によって、開始する 701。ある特定の実施形態は、計算時間を減らすために、画像の各々から関心領域 (region of interest) をトリミングする (cropping) ことを考察する。一度、画像が受信されると、システムは、第1の画像からキーポイントを決定する 703。述べたように、いくつかの実施形態において、これらのキーポイントは、SIFT または他の特徴検出ハードウェア、ファームウェア、またはソフトウェアを使用して決定されることができる。いくつかの実施形態において、システムは、また、第2の画像においてキーポイントを決定することができる。システムは、次に、第1の画像からのキーポイントを第2の画像における (特定のピクセル位置のような) ピクセル領域と相關させることができる 704。「キーポイント相關性」ソフトウェア、ファームウェア、またはハードウェア・モジュールは、この動作を行うように構成されることができる。動作のある一部は、キーポイントを相關させるための他の手段を作り出し、他のモジュール (ファームウェア、ハードウェア、またはソフトウェア) にわたって分散されることができる。この動作は、第2の画像における第1の画像における同一の画像領域を識別するためにサーブ (serve) することができる。20

#### 【0032】

視差 D は、次に、第1の画像の各キーポイントの位置と第2の画像の相關ピクセル位置との間で計算されることができる 705。キーポイントが両方の画像のために計算されたところで、キーポイント間の視差は、相關キーポイントの各々の相対的位置を減算することによって決定されることができる。視差は、次に、図 6 の 601 に類似した視差ヒストグラムとして編成され、ヒストグラム 602 に類似した対応する深度ヒストグラムが決定されることができる。深度ヒストグラムは、次に、選択された統計値に基づいて、センサのための最大焦点深度を決定するために使用されることができる。「視差決定」ソフトウェア、ファームウェア、またはハードウェア・モジュールは、この動作を行うように構成されることができる。動作のある一部は、視差を決定するための他の手段を作り出す、他のモジュール (ファームウェア、ハードウェア、またはソフトウェア) にわたって分散されることができる。30

#### 【0033】

処理 700 を実現する実施形態において、計算効率を改善するために、処理 700 は、各視差のための深度に変換し、次に、深度の平均を決定するのではなく、視差の統計値 (span style="float: right;">40

このケースでは、平均)を決定する 706。次に、単一の統計値の深度のみが、図 4 のそれと類似したグラフに関して決定される必要がある 707。この深度は、次に、その次の画像キャプチャの間、新しいカメラの焦点深度として使用されること 708 ができる。

#### 【0034】

述べられたように、他の実施形態は、その代わりに、視差の各々を深度に変換することができ、次に、深度を平均する。他の実施形態は、所望の焦点深度を決定するために、平均値、中央値、またはいくつかの他の統計値を代わりにとることができる。「深度決定」ソフトウェア、ファームウェア、またはハードウェア・モジュールは、この動作を行うように構成されることができる。モジュールは、視差ヒストグラムと連動して動作することができる。動作のある一部は、焦点深度を決定するための他の手段を作り出し、他のモジュール(ファームウェア、ハードウェア、またはソフトウェア)にわたって分散されることができる。一度、焦点深度が決定されたとすると、画像センサ 101a は調整され得る。センサ 101b は、また、プロセッサ 113 によって調整されることができる、または、センサ 101b はセンサ 101a の焦点深度を独自に追尾することができる。上記に述べられたように、ある特定の実施形態において、単一のセンサのみが、決定された焦点深度に基づいて調整されることができる。

#### 【0035】

これらの実施形態の変化において、システムは、単に視差の平均をとるのではなく、深度を決定するために、シーンからの情報を代わりに使用することができる。たとえば、平均をとる代わりに、キーポイント視差は、単一のオブジェクトにおけるそれらの存在、および照明条件に基づいて、重み付けされ得る。たとえば、ヒストグラムは、ある特定の焦点距離に関連付けられた焦点の質(focal quality)によるヒストグラムから各ポイントを重み付けることによって向上されることができる。ある特定のカメラ配置において、焦点が 3 メートルにセットされた場合、2 メートルと無限大(infinity)との間のオブジェクトは、良質な焦点を有することができ、1 メートルから 2 メートル間のオブジェクトは、相当の焦点を有することができ、0.5 から 1 メートル間のオブジェクトは劣った焦点を有し得る。図 6 のヒストグラムは、他の範囲よりも頻繁に優先焦点範囲が選択されるように、それに応じて重み付けされる。これは、ある特定の実施形態において、「領域の重み付け特性(region weighted saliency)」と称されることができる。他の変化において、画像からの頻度情報(frequency information)が、キーポイント選択に組み込まれることができる。テクスチャを含むオブジェクトは、テクスチャなしで、またはわずかなテクスチャで、オブジェクトよりも多くのキーポイントを生成することができ、それにより平均に影響を及ぼす。従って、テクスチャ・オブジェクトに関連付けられたキーポイントは、非テクスチャ・オブジェクトから異なる重みを受信することができる。一変化において、テクスチャ内の領域が検出されることができ、これらの領域は、次に、その領域におけるキーポイントの重みを少なくするために使用される。

#### 【0036】

立体ムービーをキャプチャする場合、処理 700 が、単一フレーム、すなわち、単一の立体画像ペアに適用され得る。次に、カメラ配置またはシーンが適切な焦点の再評価を必要とするように修正されるまで、決定された焦点深度は、次の画像キャプチャの間、画像センサによって使用され得る。従って、動作 700 は、任意の状態依存である必要はないというメリットを有する。すなわち、従来のオートフォーカス・システムは、ムービー・キャプチャ処理を定期的にタイムアウトする必要があり、焦点を再評価するために複数の焦点深度をキャプチャするだろう。対照的に、処理 700 は、それがフレーム遅延を生じさせないという点で「瞬時の」であり得る。これは、シームレス・フォーカス・トラッキング(seamless focus tracking)を容易にする。処理は、現在の焦点位置と焦点位置推定(focus position estimation)との間にフィードバック(または、依存性)はないので、システムの安定性をさらに保証することができる。さらに、焦点動作は単一フレームで達成され得るので、ユーザの手の動きが、任意のぶれ(blur)を生成する可能性は高くないだろう。

10

20

30

40

50

## 【0037】

述べられたように、処理700のステップ703および704において生成されたキーポイントは、第2の画像に適用された場合、画像の一部にアイデンティティを割り当てることができ、一貫して理解される任意のデータ構造を含むことができる。述べられたように、いくつかの実施形態において、キーポイントは、SIFT特徴生成モジュールから生成されるスケール不变特徴変換(SIFT)キーポイントを含むことができる。図8は、特徴生成モジュールにおけるキーポイントを生成するための別の可能な処理を図示する。

## 【0038】

処理800は、生の立体画像ペアのうちの1つを受信すること802によって開始する801。次に、画像は、できる限り雑音に対するアルゴリズムのロバスト性を改善するため、および、計算要求を減らすために、サブサンプルされる803。次に、画像は、水平な高域フィルタを通過し得る804。いくつかの実施形態において、フィルタは、

## 【数1】

$$h = \begin{bmatrix} -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{bmatrix}$$

## 【0039】

によって与えられた応答を備えた $3 \times 4$ カーネル(kernel)を含むことができる。

## 【0040】

次に、処理は、すなわち、各値を2乗することによって、画像力を計算すること805ができる。最終的に、処理は、雑音および低電力値(low-power values)を除去するために、値を閾値処理すること806ができる。システムは、閾値を上回る残りの値の中から、「極大値(maximum values)」を識別するだろう。いくつかの実施形態において、極大値が、閾値を上回るそれらの画像部分であり得るのに対して、他の実施形態において、極大値は、それらの局地的な隣接値(local neighbors)に関連して定義され得る。たとえば、閾値を上回る隣接ピクセル間のデルタ(delta)は、極大値を識別するために使用され得ることができる。識別された極大値は、上記に説明された視差決定ステップのために使用され得るキーポイントの位置を示す。システムは、終了する808前に、これらのキーポイントの位置を記憶すること807ができる。この方法で画像をサブサンプリングおよび閾値処理することによって、キーポイントとしてサーブされ得るピクセル位置を決定することが必要とされる計算時間が、削減され得る。これらの実施形態におけるキーポイントがピクセル位置を含むので、ピクセル位置は、時折「キーポイント」と称され得ることができる。しかしながら、当業者は、変化を容易に理解するだろう。ここにおいて、キーポイントは、位置と、数多くの隣接ピクセル値および位置との両方を含む。キーポイントは、また、ピクセル値またはピクセル位置ではなく、画像部分の頻度容量またはピクセル値のグラディエント(gradients)を直接言及することができる。たとえば、SIFTキーポイントは、ピクセル・グラディエントを示すベクトルを含むことができる。

## 【0041】

一度、第1の画像におけるキーポイントが決定されたとすると、それは、画像部分の間の視差が決定され得るように、キーポイントを第2の画像における位置と相關させるために依然として残り得る。図8の処理800によって生成されたキーポイントのようなキーポイントかどうかを決定するための可能な方法は、画像における位置と相互に関連付けられ、図9に関して説明されるだろう。図9は、立体画像ペア900aおよび900bを図示する。画像900aおよび900bの画像ペアをキャプチャするために使用される2つのセンサの構成に基づいて、各キーポイント901周囲の検索領域902を決定することができる。検索領域902は、キャプチャ・デバイス構成の結果として、キーポイントが左右の画像において変位され得る最大距離(すなわち、最大予測視差(maximum expected

10

20

30

40

50

disparity) )を、指定することができる。いくつかの実施形態において、画像センサは垂直な視差が欠如しているので、キーポイント 901 は、一般に、シーンにおける垂直エッジ (vertical edges) に位置することができる。

#### 【 0042 】

検索領域 902 は、画像 900a および 900b の各々の同一の絶対位置 (absolute position) に位置する。図 9において、検索領域 902 は、画像 900a および 900b の間に垂直な視差が存在しないと仮定される (すなわち、同列におけるピクセルのみが考慮される) ので、単一ピクセルの高さを有する長方形 (rectangle) を含む。領域 902 の高さは、画像 900a および 900b の間に存在し得る垂直な視差の量に関して増加され得る。システムは、第 2 の画像 900b の検索領域 902 における各ピクセルを通して繰り返すことができる、キーポイント 901 を囲む画像 900a の一部とのピクセルの対応を決定することができる。これは、下記にさらに詳しく説明される相関メトリック (correlation metric) を使用していくつかの実施形態において達成され得る。

#### 【 0043 】

ある特定の実施形態において、キーポイントは、単に画像 900a のためではなく、画像 900a と画像 900b の両方のために決定され得る。画像間のキーポイントを相關させることを試みる場合、システムは、第 2 の画像 900b における検索領域 902 内でキーポイントを識別することができる。画像 900b のうちの 1 つのキーポイントのみが検索領域 902 において発見された場合、このキーポイントは、第 1 の画像からのキーポイントと相互に関連付けられることができる。画像 900b の 2 つ以上のキーポイントが領域 902 に存在するところで、システムは、どのキーポイントが第 1 の画像からのキーポイント 901 に最も良く対応するかを決定するために、検索領域 902 における画像 900b の各キーポイントに対して相関メトリックを適用することができる。キーポイントが 1 つの画像のみのためにとられる場合、メトリックが適用されるのと同様に、メトリックは、キーポイント 901 および 901b が、画像 900a および 900b の各々における同一のシーンの一部を参照する可能性が高いことを確認するために、キーポイント 901 および 901b のピクセル位置に隣接するピクセル値を考慮し得る。キーポイントが両方の画像のために作り出されているところで、領域内の各ピクセル間ではなく、領域 902 におけるキーポイント間で繰り返すことのみが必要であり得る。

#### 【 0044 】

上記説明された実施形態において、システムは、決定されたキーポイント 901 に対応する検索領域 902 のある特定のピクセルを通して繰り返す。システムは、領域 902 における各ピクセルに相関メトリックを適用することができる。キーポイント 901 の位置を囲む領域との最大相關性を有する領域 902 におけるピクセルは、次に、キーポイント 901 と相互に関連付けられ得る。領域 902 における画像 900b の各ピクセルを通して繰り返すための計算コスト (computational cost) は、画像 900b のすべてのためのキーポイントを計算し、各キーポイント間の相關性を決定するためのコストよりも少ない。しかしながら、いくつかの実施形態において、ほんのわずかなキーポイントのみが生成されたところで、システムは、1 つの画像のキーポイントに関連付けられた領域 902 の間で繰り返すのではなく、すべてのキーポイント間の相關性を直接決定することができる。

#### 【 0045 】

ある特定の実施形態において、900a におけるキーポイントに対応する画像 900b におけるキーポイントまたはピクセル位置を識別するために使用される相關性メトリックは、考慮中の画像 900b における位置を囲むピクセル、および画像 900a におけるキーポイントの位置を囲むピクセルための平均 2 乗誤差の計算を含むことができる。すなわち、画像 900b における検索領域 902 における位置のための隣接ピクセルおよび第 1 画像のピクセル隣接キーポイント 901 の平均 2 乗誤差は、相關性メトリックとして使用されることができる。平均 2 乗誤差は、

10

20

30

40

## 【数2】

$$R(\Delta) = \sum_{i=-3}^3 \sum_{j=-3}^3 \left( S_{left}(i+M, j+N) - S_{right}(i+M+\Delta, j+N) \right)^2$$

## 【0046】

として計算されることができ、ここで、Rは、平均2乗誤差であり、 $S_{left}$ は、画像900aにおけるピクセル値を含み、 $S_{right}$ は、画像900bにおけるピクセル値を含み、MおよびNは、検査中の領域902のピクセル位置または現在のキーポイントのための領域902に対する画像の中に、水平および垂直オフセットを含み、 $\Delta$ は、検索領域902における現在の位置のために適用された水平偏移(horizontal shift)を含む( $S_{left}$ および $S_{right}$ に対する第1のパラメータは、列の位置(column position) / x軸、および第2の行の位置(row position) / y軸である)。平均2乗誤差は上記例の $7 \times 7$ ウィンドウ(window)内であるが、当業者は、画像解像度および適用されたサブサンプリングによって、ウィンドウ寸法(window dimensions)の範囲を容易に思い描くことができる。さらに、センサ101aおよび101bは、上記の例において1つも垂直な視差を有することができないと推定されるので、検索領域902は、水平にのみ拡大し、 $\Delta$ は、x軸 / 列の方向(column direction)にのみ現れる。より多くのロバスト・システムは、検索領域902の高さを増加させること、および垂直方向の $\Delta$ を含むことによって、センサの位置付け(sensor positioning)におけるエラーを補償することができる。画像がダウンサンプルされる(downsampled)につれて、サブピクセル解像度は、キーポイント901と相互に関連付けられた画像900bの領域902におけるピクセルのより正確な決定を容易にするために、いくつかの実施形態において、多項式補間(polynomial interpolation)のような、補間を使用して決定することができる。すなわち、センサ101aに関連するセンサ101bの変位は、正確なピクセルの整数であることができない。したがって、特にサブサンプリング後、正確なキーポイント901の相関性は、検索領域902におけるピクセル間の位置を含むことを必要とし得る。キーポイント901と最大限に相互に関連付けられた画像900bの領域902における位置は、補間されたポイントにおいて、ピクセル位置の中間に位置する(fall between)。

## 【0047】

ここに開示された実現に関連して記載されたさまざまな例示的な論理ブロック、モジュール、および回路は、汎用プロセッサ、デジタル信号プロセッサ(DSP)、特定用途向け集積回路(ASIC)、フィールド・プログラマブル・ゲート・アレイ(FPGA)または他のプログラマブル論理デバイス、離散ゲートまたはトランジスタ論理、離散ハードウェア・コンポーネント、または、ここに説明される機能を実行するように設計されたこれら任意の組み合わせとともに実現または実行することができる。汎用プロセッサは、マイクロ・プロセッサであることができるが、あるいは、プロセッサは、任意の従来のプロセッサ、コントローラ、マイクロコントローラ、または状態機械(state machine)であることができる。プロセッサは、また、たとえばDSPとマイクロ・プロセッサの組み合わせ、複数のマイクロ・プロセッサ、DSPコアと連動する1つ以上のマイクロ・プロセッサ、または任意の他のこののような構成であるコンピューティング・デバイスの組み合わせとして実現することができる。

## 【0048】

ここで開示された実現に関連して説明された方法または処理のステップは、ハードウェアで直接的に、プロセッサによって実行されるソフトウェア・モジュールで、または両者の組み合わせで具現化することができる。ソフトウェア・モジュールは、RAMメモリ、フラッシュ・メモリ、ROMメモリ、EPROMメモリ、EEPROMメモリ、レジスタ、ハード・ディスク、リムーバブル・ディスク、CD-ROM、または当該技術で周知の任意の他の形態の非一時的な記憶媒体に存在することができる。例示的なコンピュータ可読記憶媒体は、プロセッサがコンピュータ可読記憶媒体から情報を読み出し、コンピュ

10

20

30

40

50

一タ可読記憶媒体に情報を書き込むことができるように、プロセッサに結合される。あるいは、記憶媒体は、プロセッサと一体化することができる。プロセッサおよび記憶媒体は、A S I C に存在することができる。A S I C は、ユーザ端末、カメラ、または他のデバイスに存在することができる。あるいは、プロセッサおよび記憶媒体は、ユーザ端末、カメラ、または他のデバイスに離散コンポーネントとして存在することができる。

**【 0 0 4 9 】**

見出し (headings) は、参考のために、およびさまざまなセクションの位置を特定することを支援するために、ここに含まれる。これらの見出しは、それについて説明された概念の範囲を限定するように意図されてはいない。そのような概念は、明細書全体を通して適用性を有することができる。

10

**【 0 0 5 0 】**

開示の実現の先の説明は、いずれの当業者でも本発明を作り出し、使用することを可能にさせるために提供されている。これらの実現に対するさまざまな変更は、当業者に容易に理解され、ここで定義される一般的な原理は、本発明の精神または範囲から逸脱することなく、他の実現に適応することができる。このように、本発明は、ここに説明された実現に限定されることは意図しておらず、ここに開示される原理および新規な特徴と一致する可能性がある最も広い範囲が付与されるべきである。

以下に本願発明の当初の特許請求の範囲に記載された発明を付記する。

**[ C 1 ]**

第 1 の視点に関連付けられた第 1 の画像を受信することと、  
第 2 の視点に関連付けられた第 2 の画像を受信することと、  
前記第 1 の画像に基づいて第 1 の複数のキーポイントを決定することと、  
前記第 1 の複数のキーポイントからのキーポイントを前記第 2 の画像における位置と相関させることと、  
前記第 1 の複数のキーポイントの各々に関連付けられた複数の視差を決定することと、  
前記複数の視差、前記第 1 の視点の前記位置、および前記第 2 の視点の前記位置に基づいて焦点深度を決定すること  
を含む、画像センサのための焦点深度を決定するための電子デバイスにおける方法。

20

**[ C 2 ]**

前記第 2 の画像に基づいて第 2 の複数のキーポイントを決定することをさらに含む、C 1 に記載の方法。

30

**[ C 3 ]**

前記第 1 の複数のキーポイントからのキーポイントを前記第 2 の画像における位置と相関させることは、前記第 1 の複数のキーポイントからのキーポイントを前記第 2 の複数のキーポイントからのキーポイントと相関させることを含む、C 2 に記載の方法。

**[ C 4 ]**

前記第 1 の複数のキーポイントからのキーポイントを前記第 2 の画像における位置と相関させることは、前記第 2 の画像における検索範囲内のピクセルを介して繰り返すことを含む、C 1 に記載の方法。

40

**[ C 5 ]**

前記第 1 の複数のキーポイントからのキーポイントを前記第 2 の画像における位置と相関させることは、前記第 1 の画像および前記第 2 の画像におけるピクセル間の平均 2 乗誤差を決定することを含む、C 4 に記載の方法。

**[ C 6 ]**

前記第 1 の画像に基づいて第 1 の複数のキーポイントを決定することは、前記第 1 の画像に基づいてスケール不变特徴変換 ( S I F T ) キーポイントを決定することを含む、C 1 に記載の方法。

**[ C 7 ]**

前記第 1 の画像に基づいて第 1 の複数のキーポイントを決定することは、前記第 1 の画

50

像をサブサンプリングすることと、前記第1の画像に高域フィルタを適用することと、前記第1の画像力を計算することと、前記第1の画像を閾値処理することを含む、C 1に記載の方法。

[ C 8 ]

前記第1の複数のキーポイントからのキーポイントをリアルタイムで生じる前記第2の画像における位置と相関させる、C 1に記載の方法。

[ C 9 ]

前記電子デバイスは、携帯電話を含む、C 1に記載の方法。

[ C 1 0 ]

コンピュータに、

10

第1の視点に関連付けられた第1の画像を受信する、

第2の視点に関連付けられた第2の画像を受信する、

前記第1の画像に基づいて第1の複数のキーポイントを決定する、

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させる、

前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定する、

前記複数の視差、前記第1の視点の前記位置、および前記第2の視点の前記位置に基づいて焦点深度を決定するステップを行わせるように構成される命令を含む、コンピュータ可読媒体。

[ C 1 1 ]

20

前記第2の画像に基づいて第2の複数のキーポイントを決定することをさらに含む、C 1 0に記載のコンピュータ可読媒体。

[ C 1 2 ]

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の複数のキーポイントからのキーポイントを前記第2の複数のキーポイントからのキーポイントと相関させることを含む、C 1 1に記載のコンピュータ可読媒体。

[ C 1 3 ]

30

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第2の画像における検索範囲内のピクセルを介して繰り返すことを含む、C 1 0に記載のコンピュータ可読媒体。

[ C 1 4 ]

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の画像および前記第2の画像におけるピクセル間の平均2乗誤差を決定することを含む、C 1 3に記載のコンピュータ可読媒体。

[ C 1 5 ]

前記第1の画像に基づいて第1の複数のキーポイントを決定することは、前記第1の画像に基づいてスケール不变特徴変換(SIFT)キーポイントを決定することを含む、C 1 0に記載のコンピュータ可読媒体。

[ C 1 6 ]

40

前記第1の画像に基づいて第1の複数のキーポイントを決定することは、前記第1の画像をサブサンプリングすることと、前記第1の画像に高域フィルタを適用することと、前記第1の画像力を計算することと、前記第1の画像を閾値処理することを含む、C 1 0に記載のコンピュータ可読媒体。

[ C 1 7 ]

前記第1の複数のキーポイントからのキーポイントをリアルタイムで生じる前記第2の画像における位置と相関させる、C 1 0に記載のコンピュータ可読媒体。

[ C 1 8 ]

前記コンピュータは、携帯電話に位置する、C 1 0に記載のコンピュータ可読媒体。

[ C 1 9 ]

50

立体キャプチャ・デバイスの焦点を合わせるためのシステムであって、前記システムは

、第1の視点に関連付けられた第1の画像を生成するように構成される第1の画像センサと、

第2の視点に関連付けられた第2の画像を生成するように構成される第2の画像センサと、

前記第1の画像に基づいて第1の複数のキーポイントを決定するように構成される特徴生成モジュールと、

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させるように構成されるキーポイント相関性モジュールと、

前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定するように構成される視差決定モジュールと、

前記複数の視差、前記第1の視点の前記位置、および前記第2の視点の前記位置に基づいて焦点深度を決定するように構成される深度決定モジュール

を含む、システム。

[ C 2 0 ]

前記特徴生成モジュールは、前記第2の画像に基づいて第2の複数のキーポイントを決定するように構成される、C 1 9に記載のシステム。

[ C 2 1 ]

キーポイントを相関させるように構成される前記モジュールは、前記第1の複数のキーポイントからのキーポイントを前記第2の複数のキーポイントからのキーポイントと相関させるように構成される、C 2 0に記載のシステム。

[ C 2 2 ]

キーポイントを相関させるように構成される前記モジュールは、前記第2の画像における検索範囲内のピクセルを介して繰り返すように構成される、C 1 9に記載のシステム。

[ C 2 3 ]

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相関させることは、前記第1の画像および前記第2の画像におけるピクセル間の平均2乗誤差を決定することを含む、C 2 2に記載のシステム。

[ C 2 4 ]

前記特徴生成モジュールは、前記第1の画像に基づいてスケール不变特徴変換(SIFT)キーポイントを決定するように構成される、C 1 9に記載のシステム。

[ C 2 5 ]

前記特徴生成モジュールは、前記第1の画像をサブサンプリングし、前記第1の画像に高域フィルタを適用し、前記第1の画像力を計算し、前記第1の画像を閾値処理するように構成される、C 1 9に記載のシステム。

[ C 2 6 ]

キーポイントを相関させるように構成される前記モジュールは、前記第1の複数のキーポイントからのキーポイントをリアルタイムで前記第2の画像における位置と相関させる、C 1 9に記載のシステム。

[ C 2 7 ]

前記立体キャプチャ・デバイスは、携帯電話に位置する、C 1 9に記載のシステム。

[ C 2 8 ]

前記モジュールは、視差ヒストグラムを含む焦点深度を決定するように構成される、C 1 9に記載のシステム。

[ C 2 9 ]

立体キャプチャ・デバイスの焦点を合わせるためのシステムであって、前記システムは

、第1の視点に関連付けられた第1の画像を受信するための手段と、

第2の視点に関連付けられた第2の画像を受信するための手段と、

10

20

30

40

50

前記第1の画像に基づいて第1の複数のキーポイントを決定するための手段と、  
前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と  
相関させるための手段と、  
前記第1の複数のキーポイントの各々に関連付けられた複数の視差を決定するための  
手段と、  
前記複数の視差、前記第1の視点の前記位置、および前記第2の視点の前記位置に基  
づいて焦点深度を決定するための手段  
を含む、システム。

[ C 3 0 ]

第1の画像を受信するための前記手段は、第1のセンサを含み、第2の画像を受信する  
ための前記手段は、第2のセンサを含み、第1の複数のキーポイントを決定するための前  
記手段は、特徴生成モジュールを含み、相関させるための前記手段は、キーポイント相関  
性モジュールを含み、複数の視差を決定するための前記手段は、視差決定モジュールを含  
み、焦点深度を決定するための前記手段は、深度決定モジュールを含む、C 2 9に記載の  
システム。

10

[ C 3 1 ]

第1の複数のキーポイントを決定するための前記手段は、前記第2の画像に基づいて第  
2の複数のキーポイントを決定するように構成される、C 2 9に記載のシステム。

[ C 3 2 ]

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相  
関させるための前記手段は、前記第1の複数のキーポイントからのキーポイントを前記第  
2の複数のキーポイントからのキーポイントと相関させるように構成される、C 3 1に記  
載のシステム。

20

[ C 3 3 ]

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相  
関させるための前記手段は、前記第2の画像における検索範囲内のピクセルを介して繰り  
返すように構成される、C 2 9に記載のシステム。

[ C 3 4 ]

前記第1の複数のキーポイントからのキーポイントを前記第2の画像における位置と相  
関させるための前記手段は、前記第1の画像および前記第2の画像におけるピクセル間の  
平均2乗誤差を決定するように構成される、C 3 3に記載のシステム。

30

[ C 3 5 ]

第1の複数のキーポイントを決定するための前記手段は、前記第1の画像に基づいてス  
ケール不变特徴変換(SIFT)キーポイントを決定するように構成される、C 2 9に記  
載のシステム。

[ C 3 6 ]

第1の複数のキーポイントを決定するための前記手段は、前記第1の画像をサブサンプ  
リングし、前記第1の画像に高域フィルタを適用し、前記第1の画像力を計算し、前記第  
1の画像を閾値処理するように構成される、C 2 9に記載のシステム。

40

[ C 3 7 ]

キーポイントを相関させるための前記手段は、前記第1の複数のキーポイントからの前  
記キーポイントをリアルタイムで前記第2の画像における位置と相関させる、C 2 9に記  
載のシステム。

[ C 3 8 ]

前記立体キャプチャ・デバイスは、携帯電話に位置する、C 2 9に記載のシステム。

【図1】

図1

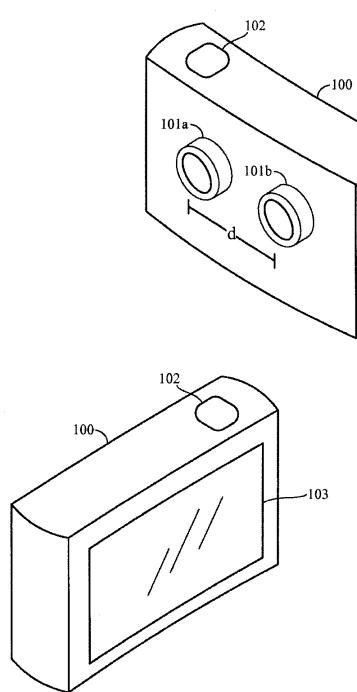


FIG. 1

【図2】

図2

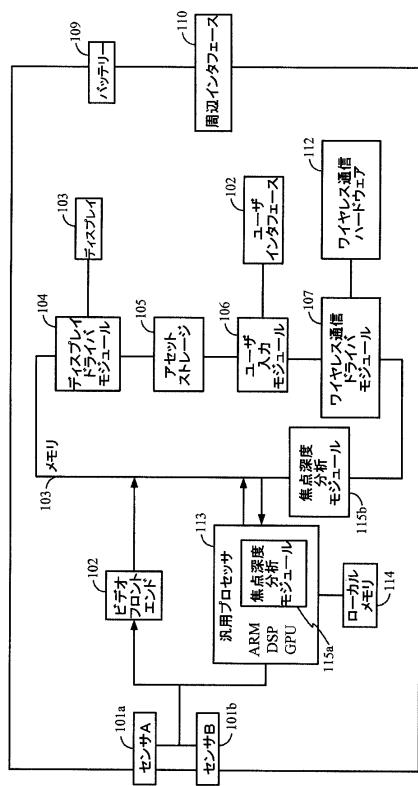


FIG. 2

【図3】

図3

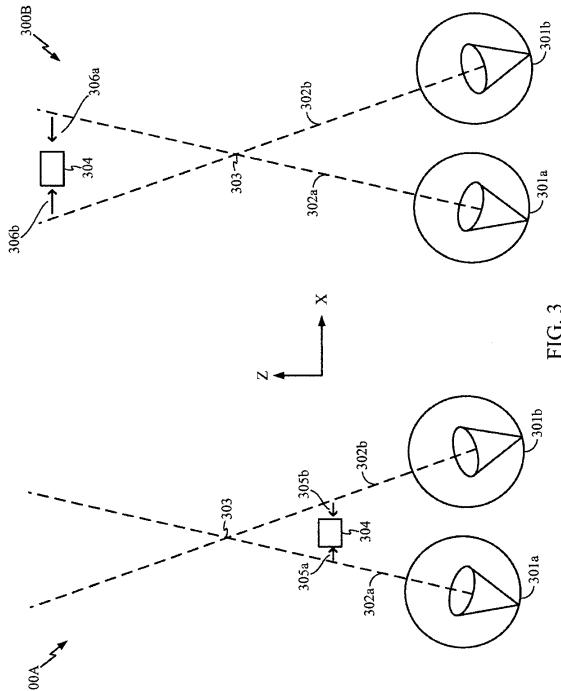


FIG. 3

【図4】

図4

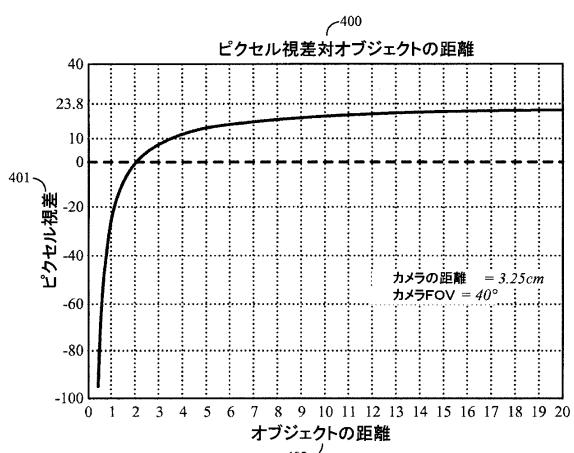


FIG. 4

【図5A】

図5A

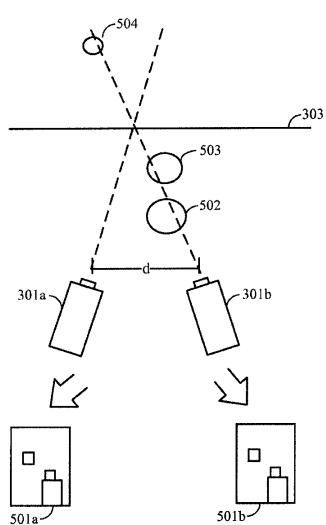


FIG. 5A

【図5B】

図5B

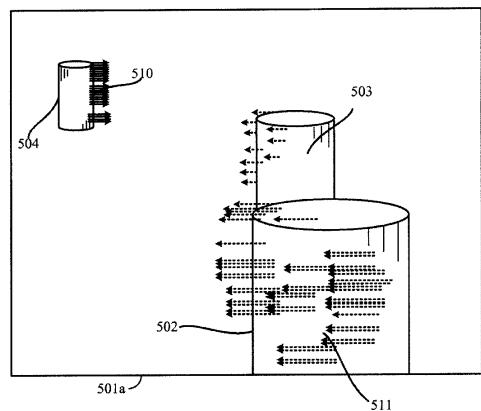


FIG. 5B

【図6】

図6

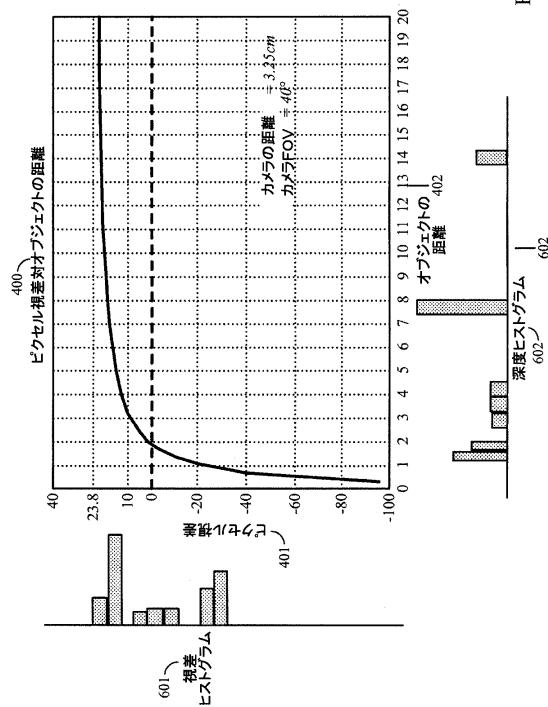


FIG. 6

【図7】

図7

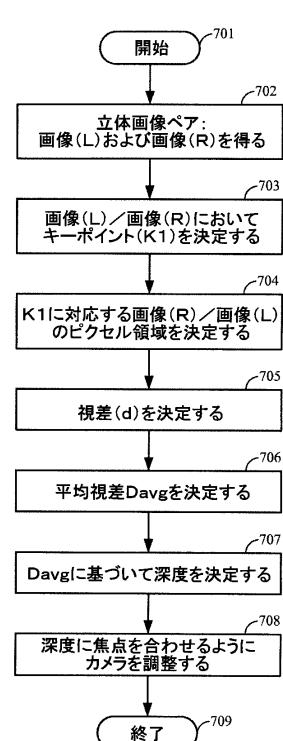


FIG. 7

【図8】

図8

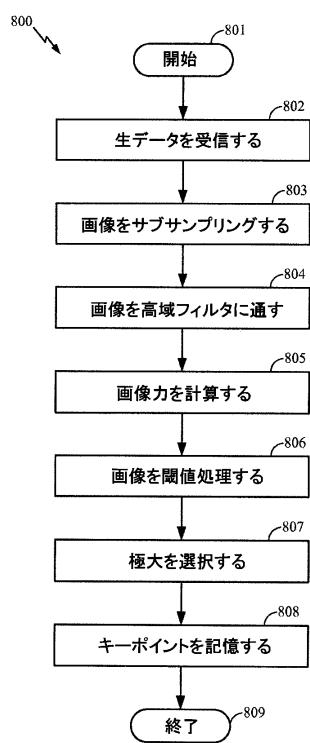


FIG. 8

【図9】

図9

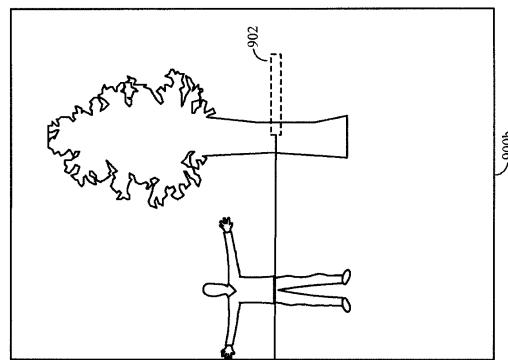
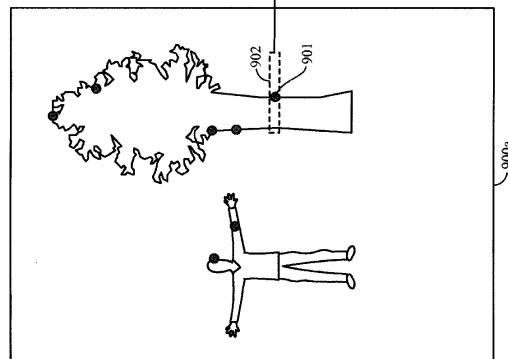


FIG. 9



900a

---

フロントページの続き

(31)優先権主張番号 61/489,231

(32)優先日 平成23年5月23日(2011.5.23)

(33)優先権主張国 米国(US)

(72)発明者 アタナソブ、カリン・エム .

アメリカ合衆国、カリフォルニア州 92121、サン・ディエゴ、モアハウス・ドライブ 57  
75

(72)発明者 ゴマ、セルジウ・アール .

アメリカ合衆国、カリフォルニア州 92121、サン・ディエゴ、モアハウス・ドライブ 57  
75

(72)発明者 ラマチャンドラ、ピカス

アメリカ合衆国、カリフォルニア州 92121、サン・ディエゴ、モアハウス・ドライブ 57  
75

合議体

審判長 鳥居 稔

審判官 篠原 功一

審判官 渡辺 努

(56)参考文献 特開2010-41381(JP, A)

特開平8-47001(JP, A)

(58)調査した分野(Int.Cl., DB名)

H04N 13/00-17/06

H04N 5/222-5/257