

(19)日本国特許庁(JP)

**(12)特許公報(B2)**

(11)特許番号  
**特許第7078392号**  
**(P7078392)**

(45)発行日 令和4年5月31日(2022.5.31)

(24)登録日 令和4年5月23日(2022.5.23)

(51)国際特許分類

G 0 1 B	11/22 (2006.01)	F I	G 0 1 B	11/22	H
G 0 6 T	1/00 (2006.01)		G 0 6 T	1/00	3 1 5

請求項の数 15 外国語出願 (全20頁)

(21)出願番号	特願2017-248111(P2017-248111)	(73)特許権者	500102435 ダッソー システムズ DASSAULT SYSTEMES フランス国 78140 ベリジー ピラ クブレー リュ マルセル ダッソー 10 110001243
(22)出願日	平成29年12月25日(2017.12.25)	(74)代理人	特許業務法人 谷・阿部特許事務所
(65)公開番号	特開2018-109976(P2018-109976 A)	(72)発明者	モハメド アミーネ アヤリ フランス 78140 ベリジー - ピラ ブレー リュ ラスパイユ 5 ヴィンセント ギテニー フランス 75010 パリ ブールバード ド マゼンタ 33
(43)公開日	平成30年7月12日(2018.7.12)	(72)発明者	佐田 宏史
審査請求日	令和2年12月7日(2020.12.7)	審査官	

最終頁に続く

(54)【発明の名称】 深度センサノイズ

**(57)【特許請求の範囲】****【請求項1】**

深度センサタイプのノイズを入力深度マップに付加するように構成された関数を決定するためのコンピュータによって実施される方法であって、

- 学習データセットを形成する(S20-S60)ステップであって、前記学習データセットは、各々がそれぞれのノイズ有り深度マップに関連付けられた複数のノイズ無し深度マップを含み、各ノイズ無し深度マップおよびそれとのノイズ有り深度マップが1または複数のシーンの複数の深度センサポジショニングのうちの同じポジショニングに対応し、前記形成は、

前記深度センサタイプの物理的インスタンスを用いて各ノイズ有り深度マップを取得する(S40)ステップと、

各ノイズ無し深度マップを仮想的に算出する(S604)ステップとを含む、形成するステップと、

- 前記学習データセットに基づいて前記関数を学習する(S70)ステップとを備えることを特徴とするコンピュータによって実施される方法。

**【請求項2】**

各ノイズ無し深度マップを仮想的に算出する(S604)ステップは、前記対応する深度センサポジショニングと前記対応するシーンの所定のモデルとに基づくことを特徴とする請求項1に記載の方法。

**【請求項3】**

前記形成は、前記それぞれのノイズ有り深度マップから前記対応する深度センサポジショニングを決定する（S602）ステップをさらに含むことを特徴とする請求項1または2に記載の方法。

**【請求項4】**

前記それぞれのノイズ有り深度マップから前記対応する深度センサポジショニングを決定する（S602）ステップは、前記それぞれのノイズ有り深度マップにおいてキャリブレーションオブジェクトおよび／またはキャリブレーションパターンを特定することを含むことを特徴とする請求項3に記載の方法。

**【請求項5】**

前記複数の深度センサポジショニングは、各々が同じ領域からのそれぞれの深度に対応する1または複数のサブ複数の深度センサポジショニングを含むことを特徴とする請求項1ないし4のいずれか一項に記載の方法。 10

**【請求項6】**

各々が同じ領域からのそれぞれの深度に対応する各サブ複数の深度センサポジショニングは、10深度センサポジショニングより多く含み、好適には50深度センサポジショニングより多く含むことを特徴とする請求項5に記載の方法。

**【請求項7】**

前記複数のノイズ有り深度マップは、1または複数のビデオを形成することを特徴とする請求項5または6に記載の方法。

**【請求項8】**

前記関数の学習は、回帰学習によって実行されることを特徴とする請求項1ないし7のいずれか一項に記載の方法。 20

**【請求項9】**

前記関数の学習は、畳み込み層および／または逆畳み込み層を備えるネットワークの空間内で実行されることを特徴とする請求項1ないし7のいずれか一項に記載の方法。

**【請求項10】**

前記畳み込み層および／または逆畳み込み層は、サイズのそれが $10 \times 10$ より下位のフィルタ、好適には $5 \times 5$ より下位のフィルタ、好適には $3 \times 3$ に等しいフィルタを実装することを特徴とする請求項9に記載の方法。

**【請求項11】**

前記学習した関数をデータストレージ媒体に格納するステップをさらに備えることを特徴とする請求項1ないし10のいずれか一項に記載の方法。 30

**【請求項12】**

シーンの深度センサポジショニングと前記シーンの所定のモデルとに基づく深度マップを生成するためのコンピュータによって実施される方法であって、

- 前記シーンの前記深度センサポジショニングと前記所定のモデルとに基づいてノイズ無し深度マップを仮想的に算出する（S100）ステップと、
- 前記算出されたノイズ無し深度マップに、請求項1ないし10のいずれか一項に従う学習可能な関数を適用する（S200）ステップを備えることを特徴とするコンピュータによって実施される方法。

**【請求項13】**

請求項1ないし10および12のいずれか一項に記載の方法を実行する命令を備えることを特徴とするコンピュータプログラム。

**【請求項14】**

請求項13に記載のコンピュータプログラムが格納されているメモリを備えることを特徴とするコンピュータシステム。

**【請求項15】**

前記メモリに結合されたプロセッサをさらに備えることを特徴とする請求項14に記載のコンピュータシステム。

**【発明の詳細な説明】**

10

20

30

40

50

**【技術分野】****【0001】**

本発明は、コンピュータプログラムおよびシステムの分野に関し、より詳細には、深度センサタイプのノイズに関する方法、プログラムおよび製品に関する。

**【背景技術】****【0002】**

オブジェクトの設計、エンジニアリングおよび製造のための多数のシステムおよびプログラムが市場に出されている。C A Dは、コンピュータ支援設計(Computer-Aided Design)の頭字語であり、例えば、それは、オブジェクトを設計するためのソフトウェアソリューションに関連する。C A Eは、コンピュータ支援エンジニアリング(Computer-Aided Engineering)の頭字語であり、例えば、それは、将来の製品の物理的振る舞いをシミュレートするためのソフトウェアソリューションに関連する。C A Mは、コンピュータ支援製造(Computer-Aided Manufacturing)の頭字語であり、例えば、それは、製造のプロセスおよびオペレーションを定義するためのソフトウェアソリューションに関連する。このようなコンピュータ支援設計システムにおいて、グラフィカルユーザインターフェースは、技術の効率性に関して重要な役割を果たす。これらの技術は、製品ライフサイクル管理(P L M)システム内に組み込まれ得る。P L Mは、企業が、拡張企業の概念に沿って製品データを共有し、共通のプロセスを適用し、および構想から製品寿命の終わりに至る製品開発のための企業知識を活用するのに役立つビジネス戦略を指す。ダッソーシステムズによって提供されるP L Mソリューション(C A T I A、E N O V I AおよびD E L M I Aの商標に基づく)は、製品エンジニアリング知識を構造化するエンジニアリングハブ、製造エンジニアリング知識を管理する製造ハブ、およびエンジニアリングハブと製造ハブとの両方への企業統合および企業接続を可能にする企業ハブ(Enterprise Hub)を提供する。全体的に、システムは、製品、プロセス、資源をリンクして動的で知識ベースの製品創出を可能にするオープンオブジェクトモデル、および最適化された製品定義、製造準備、生産およびサービスを推進する意思決定サポートを遂行する。

10

20

30

**【0003】**

このフレームワークにおいて、深度センサは現在、3 D再構成、拡張現実、ヒューマンコンピュータインターフェースおよびビデオゲームなど、多くのアプリケーションに関与している。深度センサは、深度情報を実時間および高フレームレートで提供する。主な既存の深度センサ技術は、タイムオブフライト(T o F)深度センサとストラクチャードライト(S L)深度センサを含む。

**【0004】**

タイムオブフライト深度センサは、光信号がカメラと被写体の間に要する飛行時間を測定する。これは、その時点の被写体の深度を提供する。これらのセンサは、変調された赤外光の放出に基づき、光はその後、シーンのオブジェクトに反射される。その信号の位相シフトが決定され、従って深度は

**【0005】****【数1】**

$$Z = \frac{c\varphi}{4\pi\omega},$$

40

**【0006】**

で計算され、ここに c は、光の速度であり、 $\varphi$  は、変調周波数である。

**【0007】**

構造化された光深度センサは、ステレオペアを形成する、1つのカメラと1つのレーザベースのI Rプロジェクタを有する。I Rプロジェクタは、固定されたグリッド光パターンを被写体に送出し、被写体は、赤外線カメラでキャプチャされた、このグリッドの歪んだバージョンを与える。深度は、正確なグリッドに対して歪んだグリッドを三角測量するこ

50

とによって算出される。新しい画像の場合、各ピクセルにおいて深度を算出することを求める。IR画像の各ピクセルの場合、小さい相關ウィンドウ( $9 \times 9$ または $9 \times 7$ )を使用して、そのピクセルのローカルパターンをそのピクセルの記憶されたパターンおよび水平ウィンドウの64の近隣ピクセルと比較する。ベストマッチが周知の深度からのオフセットを得る。ピクセルの観点から、これを格差と呼ぶ。従って、深度は

【0008】

【数2】

$$Z = \frac{bf}{d}$$

10

【0009】

で計算され、ここに $Z$ は、深度(メートル)であり、 $b$ は、カメラとプロジェクタの間(メートル)の水平ベースラインであり、 $f$ は、カメラの焦点距離(ピクセル)であり、および $d$ は、格差(ピクセル)である。

【0010】

これらのセンサの低成本で使い易さが高い評価を得ているが、このようなセンサは、高レベルのノイズに悩まされる。一部の研究は、この問題を例えれば、ノイズの多い深度測定に適用されるフィルタリング/ノイズ除去技術を用いて改善することに専念しているが、今のところノイズの無い深度センサは知られていない。

20

【0011】

この文脈において、深度センサによって実行される測定のノイズに関して改善されたソリューションがなおも必要である。

【先行技術文献】

【非特許文献】

【0012】

【文献】M. Gschwandtner, R. K. (2011). : Blensor: Blender Sensor Simulation Toolbox. Advances in Visual Computing: 7th International Symposium. Las Vegas, Nevada, USA  
 Chuong V. Nguyen, S. I. (2012). Modeling Kinect Sensor Noise for Improved 3D Reconstruction and Tracking. 3DIM / 3DPVT  
 Avishhek Chatterjee, V. M. (2015). Noise in Structured-Light Stereo Depth Cameras. arXiv: 1505.01936.  
 Amira Belhedi, A. B. - B. (2012). Noise Modeling and Uncertainty Propagation. ECCV.

30

【発明の概要】

【課題を解決するための手段】

40

【0013】

そのため、深度センサタイプのノイズを入力深度マップに付加するように構成された関数を決定するためのコンピュータ実装方法が提供される。

【0014】

本方法は、学習データセットを形成することを備える。学習データセットは、各々がそれぞれのノイズ有り深度マップに関連付けられた複数のノイズ無し深度マップを含む。各ノイズ無し深度マップとそれぞれのノイズ有り深度マップは、1または複数のシーンの複数の深度センサポジショニング(positionings)のうち同じポジショニングに対応する。その形成は、深度センサタイプの物理的インスタンスを用いて各ノイズ有り深度マップを取得することを含む。その形成はまた、各ノイズ無し深度マップを仮想的に算出することも

50

含む。本方法はまた、学習データセットに基づいて関数を学習することも含む。

#### 【0015】

例において、本方法は、以下のうちの1または複数を備えることができる。

- 各ノイズ無し深度マップを仮想的に算出することは、対応する深度センサポジショニングと対応するシーンの所定のモデルとに基づく。
- 形成は、それぞれのノイズ有り深度マップから対応する深度センサポジショニングを決定することをさらに含む。
- 形成は、それぞれのノイズ有り深度マップから対応する深度センサポジショニングを決定することをさらに含む。
- それぞれのノイズ有り深度マップから対応する深度センサポジショニングを決定することは、それぞれのノイズ有り深度マップにおいてキャリブレーションオブジェクトおよび/またはキャリブレーションパターンを特定することを含む。
- 複数の深度センサポジショニングは、各々が同じシーン領域からのそれぞれの深度に対応する1または複数のサブ複数(sub-pluralities)の深度センサポジショニングを含む。
- 各々が同じシーン領域からのそれぞれの深度に対応する各サブ複数の深度センサポジショニングは、10個の深度センサポジショニングより多く含み、好適には50個の深度センサポジショニングより多く含む。
- 複数のノイズ有り深度マップは、1または複数のビデオを形成する。
- 関数の学習は、回帰学習によって実行される。

関数の学習は、畳み込み層および/または逆畳み込み層を備えるネットワーク空間内で実行される。および/または

- 畳み込み層および/または逆畳み込み層は、サイズのそれが $10 \times 10$ より下位のフィルタ、好適には $5 \times 5$ より下位のフィルタ、好適には $3 \times 3$ に等しいフィルタを実装する。

#### 【0016】

上記に提供された方法のいずれかに従って学習可能な関数がさらに提供される。関数は、上記の方法に従って学習され得る。関数は、データ構造として提供される。

#### 【0017】

データ構造は、データストレージ媒体に記録され得る。

#### 【0018】

シーンの深度センサポジショニングとそのシーンの所定のモデルとに基づいて深度マップを生成するためのコンピュータ実装方法がさらに提供される。方法は、深度センサポジショニングとそのシーンの所定のモデルとに基づいてノイズ無し深度マップを仮想的に算出することを備える。方法はまた、算出されたノイズ無し深度マップに上記に提供された関数を適用することも含む。

#### 【0019】

上記に提供された方法のいずれか1つまたは任意の組み合わせを実行する命令を備えるコンピュータプログラムがさらに提供される。

#### 【0020】

上記の関数および/または上記のコンピュータプログラムが格納されているメモリを備える製品がさらに提供される。

#### 【0021】

製品は、前記メモリで構成されてもよく、それによってデータストレージ媒体を形成することができる。

#### 【0022】

製品は代替として、メモリに結合されたプロセッサなど、他の要素を備えることができ、それによってコンピュータシステムを形成することができる。

#### 【0023】

例において、システムは、深度センサ、例えば、(例えば、ワイヤレス)ネットワーク経由で、深度センサに結合されているプロセッサをさらに備えることができる。

10

20

30

40

50

**【図面の簡単な説明】****【0024】**

発明の実施形態は、限定されない例として、添付図面を参照してこれより説明される。

【図1】提供されるソリューションを示す図である。

【図2】提供されるソリューションを示す図である。

【図3】提供されるソリューションのパイプラインの例を示す図である。

【図4】学習データセットを形成する例を示す図である。

【図5】学習方法の例のフローチャートの図である。

【図6】ノイズ付加方法の例のフローチャートの図である。

【図7】回帰学習の例を示す図である。

【図8】システムの例を示す図である。

**【発明を実施するための形態】****【0025】**

関数を決定するためのコンピュータ実装方法が提供され、その関数は、ノイズを入力深度マップに付加するように構成される。その方法は、「学習方法」または「オンライン」モードまたは方法と呼ぶこともできる。

**【0026】**

学習方法に従って学習可能な関数、即ち、学習方法によって出力される関数に対応する（例えば、ストレージ媒体に記録可能な）データ構造がさらに提供される。例において、関数は、学習方法に従って効果的に学習する。その関数は、「ノイズ付加関数」と呼ぶことができる。

**【0027】**

付加されるノイズは、深度センサタイプのノイズである。深度センサは、深度データを取得する（即ち、深度測定を実行する／キャプチャする）時に、それらが出力される深度マップが、正確な予測結果から（例えば、少なくともわずかに）外れるような、不完全性を示す。正確な予測結果と実取得との差は、「ノイズ」を呼ばれる。実深度マップは、「ノイズ無し」と定義される正確な深度マップとは対照的に「ノイズ有り」深度マップである。例えば、完全に明瞭な直線エッジを有するオブジェクトをキャプチャするノイズ無し深度マップは、明瞭な直線エッジに対応する完全な直線セグメントを特徴付けるのに対して、同じオブジェクトをキャプチャするノイズ有り深度マップは、同じ明瞭な直線エッジに対応する、概ね直線であるが完全には直線でない形状を特徴付けるであろう。

**【0028】**

「深度センサタイプ」という表現は、同様のまたは少なくとも実質的に同じ構成の深度センサのカテゴリを指し、それらが同様のまたは少なくとも実質的に同じノイズを示せるようにする。例えば、所与のカテゴリの深度センサは、同じ確立分布  $f$  に従うノイズを示し、ここに  $f$  は、関数のパラメータを指す。

**【0029】**

深度センサタイプの物理的インスタンスは、そのようなカテゴリの範囲内に収まる任意の深度センサになり得る。カテゴリの深度センサは、同じ物理的パラメータおよびキャリブレーションパラメータを示して、同じブランドの同じモデルが、同じコンストラクタによって生産されている、および／または同じ製造オペレーションのインスタンスによって同じ生産ラインで生産されているようにできる。カテゴリのすべての物理的インスタンスに関して、カテゴリが制限されるほどますます、学習したノイズ付加関数が正確になる。

**【0030】**

例において、学習方法は、所与の深度センサのノイズを入力深度マップに付加するように構成された関数を決定するための方法であり、その形成に含まれるノイズ有り深度マップの各取得は、前記所与の深度センサを用いて実行される。言い換えれば、深度センサタイプは、1つの所与の深度センサまで削減されて、使用できる深度センサタイプの唯一の物理的インスタンスが、その所与の深度センサ自身になるようにする。そのような場合、ノイズ付加関数は特に、前記所与の深度センサに関して正確であり、それはまた、前記所

10

20

30

40

50

との深度センサと同様のまたは少なくとも実質的に同じ構成の深度センサに関して、それらの付加関数が同様のまたは少なくとも実質的に同じノイズを示すように、正確である。

#### 【 0 0 3 1 】

例において、学習方法は、所定の複数の深度センサのいずれか1つのノイズを入力深度マップに付加するように構成された関数を決定するための方法であり、前記所定の複数は、任意の方法、例えば、ユーザによる事前選択で決定され、その形成に含まれるノイズ有り深度マップの各取得は、前記所定の複数の深度センサのうちの1つを用いて実行され、前記所定の複数の深度センサの各深度センサは、少なくとも1つのそのような取得に関与する。そのような場合、ノイズ付加関数は、前記所定の複数の深度センサに関する平均ノイズをシミュレートし、ノイズ付加関数は、前記所定の複数の深度センサに関してそれらが示すノイズが、互いに同様になる程度に正確である。

10

#### 【 0 0 3 2 】

学習方法は、学習データセットを形成することを備える。それ自体が知られているように、学習データセットは、あるタイプのデータの値を別のタイプのデータの値に（例えば、1対1に）関連付けるデータセットである。学習データセットは、一方のタイプ（またはそれぞれの他方のタイプ）の任意の値を、他方のタイプ（またはそれぞれの一方のタイプ）の対応する値に変換する関数の（例えば、マシン）学習を実行するように構成され、学習した関数によって演算される一致は、初期の関連付けを重視する。初期の関連付けは、関連付けられたデータ間の実際の一致（即ち、物理的な関連付けに従った一致）に従ってもよく、学習データセットは、データの量および／または可変性（即ち、ダイバーシティ）の観点から学習した関数が正確に、例えば、「正確に」という語は、所定の値未満の誤り率を指す、そのような実際の一致を実質的に伝達するように構成されてもよい。

20

#### 【 0 0 3 3 】

本事例において、学習方法によって形成される学習データセットは、特に複数のノイズ無し深度マップを含む。各ノイズ無し深度マップは、それぞれのノイズ有り深度マップに関連付けられる。例えば、学習データセットは、複数のノイズ無し深度マップを備えることができ、各ノイズ無し深度マップは、それぞれのノイズ有り深度マップをポイントする。

#### 【 0 0 3 4 】

学習方法は、学習データセットに基づいて関数を学習することを備える。いずれの学習技術も実装され得る。例において、学習データセットは、訓練データセットとテストデータセットに分けられ、そしてテストフェーズが後に続く訓練フェーズによって形成されるシーケンスの発生は、テストフェーズが満たされる（例えば、テストフェーズが検証基準に合う）まで実行される。他の例において、学習データセットは、例えば、テストフェーズがないあるいはテストフェーズが他のデータで実行される、訓練データセットを構成する。

30

#### 【 0 0 3 5 】

学習したノイズ付加関数は従って、学習データセットに従って、任意のノイズ無し深度マップを対応するノイズ有り深度マップに変換するように構成された関数である。言い換えれば、学習したノイズ付加関数を特に学習データセットの所与のノイズ無し深度マップに適用することにより、データセットの所与のノイズ無し深度マップに関連付けられたそれぞれのノイズ有り深度マップに少なくとも近い結果、例えば、所定の距離の閾値未満となる。

40

#### 【 0 0 3 6 】

各ノイズ無し深度マップおよびそのそれぞれのノイズ有り深度マップは、複数の深度センサポジショニングの同じ深度センサポジショニングに対応する。複数の深度センサポジショニングは、1または複数のシーンのうちの1つに各々が配置された深度センサポジショニングから成る。言い換えれば、深度センサポジショニングは、シーンの深度センサの位置および方向から成る。「シーン」という語は、例えば、建物、街路および／または人々を特徴付ける建築または都市環境、または例えば、壁、ガラスなどのオブジェクト、および／または可動オブジェクトを特徴付ける建物内部環境（部屋など）のような、区別可能な要素を特徴付ける任意の環境を指す。

50

**【 0 0 3 7 】**

学習方法は特に、建物内部環境において効率がよく、そのためシーンが情報の良好なダイバーシティを提供する。1または複数のシーンは従って、1または複数の建物内部環境のシーンを含むことができる。建物内部環境のシーンのうちの1または複数は、（例えば、少なくとも2または5の）サイズの異なるオーダーのオブジェクト（例えば、2つのオブジェクト間のサイズの異なるオーダーは、あるオブジェクトが1番目のオブジェクトの体積の少なくとも5倍、10倍または20倍であるという意味である）および／または（例えば、少なくとも2または5の）異なる材料のオブジェクトおよび／または（例えば、少なくとも2または5の）異なる色のオブジェクトを含むことができる。

**【 0 0 3 8 】**

学習データセットの各ノイズ無し深度マップおよびそのそれぞれのノイズ有り深度マップが、同じ深度センサポジショニングに対応するので、両方の深度マップは、ノイズ有り深度マップがノイズを特徴付ける一方、ノイズ無し深度マップが正確な表現を提供することを除いては同じシーンの領域（それはまさに両方の深度センサに共通するポジショニングに使用できるキャプションに対応する領域である）の同じ深度表現を提供することになる。このように、学習データセットは、ノイズ付加関数の正確な学習を可能にさせる。

**【 0 0 3 9 】**

表現されるシーンは、（例えば、全学習データセットの）ノイズ無し深度マップとそのそれぞれのノイズ有り深度マップとの両方に対して全く同じにすることができる。「全く同じ」とは、全く同じ要素が両方の表現に提示されることを意味する。あるいは、わずかな差が提示される場合もあるが、学習に支障をきたすものではない。例えば、ノイズ有り深度マップで表現されるシーンに提示される一部の要素がノイズ無し深度マップで表現されるシーンに出ないようにすることができ、および／またはその逆も同じである。そのような場合、方法は、何らかの方法、例えば、シーンを両方の表現に共通する要素の最大セットにまで削減することによってそのような差を特定して処理することを備えることができる。これは、実装のわずかな問題に過ぎない。

**【 0 0 4 0 】**

形成は、深度センサタイプの物理的インスタンスを用いて各ノイズ有り深度マップを（物理的に）取得することを含む。深度センサタイプの同じ固有の物理的インスタンスを方法全体にわたって使用することができる。このような取得は特に、人間工学的にすることができる。選択肢として、いくつかのインスタンスを使用することができる。

**【 0 0 4 1 】**

例において、すべてのノイズ有り深度マップの取得は、実質的に連続にことができる。例えば、ユーザは、- 例えば、（深度カメラなどの）深度センサを含む任意のデバイスを通じて - 深度センサを操作し、そして動き回る間に深度マップを実質的に連続的にキャプチャすることができ、場合によりプロセス中に数回中断が起こる。各キャプションをビデオキャプションにすることでき、即ち、ユーザがビデオキャプションコマンドを起動し、その後単純に動き回り、各深度マップが自動的に取得される。複数のノイズ有り深度マップは従って、1または複数の（一連の）ビデオ（複数）を形成することができる。このような取得は特に、人間工学的にすることができる。

**【 0 0 4 2 】**

そのために必ずしも必要ではないが、ビデオの取得は、いくつかの画像が各々、それぞれの深度で同じシーン領域から人間工学的に取得されるようにさせ、従って学習データセットに、各々が同じシーン領域からのそれぞれの深度に対応するサブ複数の深度センサポジショニングを含ませる。

**【 0 0 4 3 】**

シーン領域は、シーンの一部である。同じ領域の異なる深度（即ち、距離）に深度センサをポジショニングし、その深度センサが毎度その領域の方向に向けられることによって、異なる深度でそのシーンの同じ要素（それらはまさにその領域内部の要素である）のいくつかの画像を取得することができる。

10

20

30

40

50

**【 0 0 4 4 】**

これにより情報の冗長を可能にし、特に学習方法の文脈において役立つ。実際に、深度センサタイプのノイズは、そのような取得の繰り返しに関してランダムに振る舞う物理値であり、繰り返しによって学習データセットを形成できるようにし、次にノイズの学習が特に上手くできるようにさせる。

**【 0 0 4 5 】**

これは、ユーザが深度センサデバイスを手に持ち、デバイスをシーン領域の方向に向かせるおよび／または例えば、連続的で滑らかな方法で、デバイスをシーン領域に近づけるおよび／またはデバイスをシーン領域から遠ざけることによって容易に実装される。

**【 0 0 4 6 】**

同じシーン領域からの異なる深度における 10 より多い、好適には 50 より多い取得は、情報の良好な量を提供する。取得の数を最大値（例えば、10,000 または 1000）未満に維持して取得時間を削減することもでき、例えば、100 の大きさのオーダーにすることができる。

**【 0 0 4 7 】**

例において、学習データセットの全セットの深度画像に対応する深度のセットは、深度センサタイプによって許可される深度の実質的に全領域に広がる値を含む。例えば、深度センサタイプが深度値  $d_{min}$  と  $d_{max}$  との間の（例えば、上記の所定の閾値の品質の）取得を許可すると、学習データセットは - 自ら課した制約によって - 、 $(d_{max} - d_{min}) / k$  より大きいサイズ  $d_2 - d_1$  の領域  $[d_1, d_2]$  を特徴付けず、画像内部に対応する深度を有しない。ここに  $k$  は 4、8 または 16 に等しい。これによって学習データセットの情報の高いダイバーシティが可能にする。

**【 0 0 4 8 】**

さらに、2、3 または 5 より大きいシーン領域の数  $N$  は（例えば、 $N$  領域はすべて、1 に等しいまたは 2、3 または 5 より大きいシーンの数で広がる）情報の最良の可変性を提供するように企図され得る。シーン領域の数を最大値（例えば、100 または 50）未満に維持することもでき、例えば、10 の大きさのオーダーにすることができる。

**【 0 0 4 9 】**

以下において、「シーン」と「シーン領域」の概念は、そのような概念の定義が、常にシーンをシーン領域と定義し得るので混同する恐れがある。

**【 0 0 5 0 】**

形成はまた、各ノイズ無し深度マップを仮想的に算出することをも含む。言い換えれば、各ノイズ無し深度マップは、数値データに基づいて、即ち、物理的取得を用いずに完全に決定される。各仮想算出は、対応する深度センサポジショニングに基づくことができる（即ち、ノイズ有り深度マップの取得の深度センサポジショニングが、学習データセットのノイズ無し深度マップに関連付けられるようにする）。例えば、各ノイズ有り深度マップの取得は、ノイズ無し深度マップの同じポジショニングに対する仮想算出を（例えば、実時間またはプログラム経由で後に実行するように）起動することができる。

**【 0 0 5 1 】**

例において、ノイズ無し深度マップの仮想算出は、対応するシーン（即ち、学習データセットのノイズ無し深度マップに関連付けられる、ノイズ有り深度マップの取得の深度センサポジショニングに対応するシーン）の所定のモデルに基づく。シーンのモデルは、そのシーンを表す任意のデータセットであり、そのデータセットから深度マップは例えば、シーンの 3D 形状の表現に計算され得る。このような計算は、任意の古典的手法で実行され得るが、ここでは論じない。シーンのモデルは、例えば、ユーザ（例えば、物理的インスタンスの深度センサを操作する同じユーザ）によるまたはライブラリから自動的に読み出される任意の方法で提供され得る。

**【 0 0 5 2 】**

ノイズ付加関数は、任意の入力深度マップに適用され得る。しかしながら、ノイズ付加関数の適用は、入力深度マップがノイズ無しである場合に最も関連する。

10

20

30

40

50

**【 0 0 5 3 】**

従って、ノイズ付加関数を入力ノイズ無し深度マップに適用することを備える深度マップを生成するためのコンピュータ実装方法がさらに提供される。このような方法を「ノイズ付加方法」または「オンライン」モードもしくは方法と呼ぶこともできる。

**【 0 0 5 4 】**

ノイズ付加方法は、ノイズ無し深度マップを仮想的に算出することを備える。その算出は、深度センサポジショニングとシーンの所定のモデルとに基づく。ノイズ付加方法は、ノイズ付加関数を算出されたノイズ無し深度マップに適用することをさらに備える。

**【 0 0 5 5 】**

ノイズ付加関数とノイズ付加方法はそれによって、ノイズ有り深度マップの生成が可能となり、そこでのノイズは、学習方法で使用される深度センサタイプに対応する。ノイズ付加方法は、ノイズ有り深度マップの仮想合成を実行する。ノイズ付加方法は、- 例えば、完全にデジタルで、即ち、実測定を全く用いず - 深度センサタイプの物理的インスタンスに実行される実測定によって出力されるであろう深度マップに比較的近い深度マップを出力する。所与の深度センサタイプに対応するノイズ付加関数と、所与のシーンの所定のモデルとに基づき、従ってノイズ付加方法は、前記所与のシーンの前記所与のタイプの深度センサの振る舞いをシミュレートする。

10

**【 0 0 5 6 】**

ノイズ付加方法によって入手可能な深度マップ、即ち、ノイズ付加方法によって出力される深度マップに対応するデータ構造がさらに提供される。深度マップはそれによってノイズが多くなり、「ノイズが付加された深度マップ」と呼ぶこともできる。

20

**【 0 0 5 7 】**

深度マップの 1 または複数の実世界の取得を自動車両運転プロセスなど、1 または複数のシーンの所与のタイプの深度センサに関与させる、実世界プロセスをシミュレートするシミュレーションプロセスがさらに提供される。シミュレーションプロセスは、所与の深度センサタイプと、1 または複数のシーンのそれぞれの所定のモデルとに基づくノイズ付加方法によって入手可能な 1 または複数のノイズが付加された深度マップを提供することを備え得る。

**【 0 0 5 8 】**

提供される方法、プログラム、データ構造、媒体およびシステムは、深度マップの仮想生成を、とりわけ、そのような文脈でマシン学習パラダイムを実装し、それによってその優位性を獲得することによって改善する。

30

**【 0 0 5 9 】**

椅子の例を示す図 1 について、提供される方法、プログラム、データ構造、媒体およびシステムは、仮想的に算出されたノイズ無し深度マップ 1 2 からノイズ有り深度マップ 1 4 の仮想生成を可能にさせる。

**【 0 0 6 0 】**

提供される方法、プログラム、データ構造、媒体およびシステムは、3 D 深度カメラシミュレーション(3D Depth Camera Simulation)の分野に関連し、カメラの取得ノイズをモデル化することによって実(深度)カメラの深度ストリームをシミュレートするのに適用され得る。このようなノイズは、平面のスキャン 2 2 および同じ平面のノイズレベルにおけるズーム 2 4 を示す、図 2 で示されている。カメラのノイズモデルが分かると、その振る舞いを合成深度データにエミュレートすることができる。

40

**【 0 0 6 1 】**

提供される方法、プログラム、データ構造、媒体およびシステムは、多くのアプリケーション、例えば、あらゆるマシン学習ベースのアプリケーションから、合成データから現実的な訓練データを例えば、自動運転シミュレータ(Autonomous Driving Simulators)で人工的に生成させる 3 D センサ訓練データオーグメンテーションまで有することができる。

**【 0 0 6 2 】**

ノイズセンサモデリングに関する既存の先行技術は、深度センサに関連する文献で見つけ

50

ることができる。実際に、深度センサノイズのモデリングに関して、以下の例のようないくつかの研究がある。

#### 【0063】

- 論文“M. Gschwandtner, R. K. (2011).: Blensor: Blender Sensor Simulation Toolbox. Advances in Visual Computing: 7th International Symposium. Las Vegas, Nevada, USA”において深度取得技術ToFおよびSLをシミュレートするセンサシミュレーションフレームワークである、ブレンサ(Blensor)を提案している。センサの技術に応じて、ブレンサは、全取得パイプラインを各自シミュレートする。例えば、ToFセンサの場合、ブレンサは、信号のセットをレイキヤストして、ランダムなガウスノイズを各自進んだ距離に均一に付加する。このような均一なノイズ分布が非常に単純化されたままであることに留意されたい。10

#### 【0064】

- 論文“Chuong V. Nguyen, S. I. (2012). Modeling Kinect Sensor Noise for Improved 3D Reconstruction and Tracking. 3DIM / 3DPVT”において深度測定の実験分析を使用して経験的に導かれたノイズモデルを提案している。論文では、ノイズを距離とセンサのポーズの両方の関数にモデル化する。実験構成(experimental setup)で、両方：各深度レベル(0.4から3.0メートルの範囲)と各々変化する回転角とで測定された1000の異なる深度マップの回転面の深度測定から側方と軸方向のノイズ確率分布を測定する。著者は、側方のノイズが深度に対して大きく変化しないことを見つけた。対照的に、軸方向のノイズは、深度とともに明らかに増大し、特に深度測定の二次関数に従う。20

#### 【0065】

- 論文“Avishhek Chatterjee, V. M. (2015). Noise in Structured-Light Stereo Depth Cameras:, arXiv:1505.01936.”においてSLノイズ特性を研究して、ノイズを深度の二次関数として推定する理論モデルを手作業で定義する。著者は、この理論モデルを経験的深度測定によって認証する。30

#### 【0066】

- 論文“Amira Belhedi, A. B. - B. (2012). Noise Modelling and Uncertainty Propagation. ECCV.”においてToFノイズ分布をガウス分布として特徴付け、その標準偏差がピクセルの位置および深度に従って変化することを示す。著者は、3D薄板スプラインをノイズの標準偏差モデルとして使用することを提案している。このノイズモデルは、深度ピクセルの変動測定の研究に基づき、各々深度レベル100において深度マップが記録されている。それらの値に従って、著者は、ノイズの正規性について結論付けている。著者は、ノイズ分散をピクセルの位置および深度の関数にモデル化する。著者は、ノイズ分散が画像中央から離れて境界に増加して、画像中央で最高の精度が測定されると結論付けている。ノイズ分散は、3D薄板スプライン(3D Thin-Plate-Spline)平滑化関数によってモデル化される。40

#### 【0067】

このような既存の深度センサモデリング方法は、以下が欠如している：

精度。モデル化が難しい物理的取得プロセスをモデル化しなければならない。ほとんどの場合、精密さに欠ける実験的測定に基づく非常に単純化した物理的取得モデルを使用する理由による。

一般性。深度センサの各新しいタイプとは別に、各深度センサの取得パイプラインをモデル化しなければならない理由による。

#### 【0068】

提供される方法、プログラム、データ構造、媒体およびシステムは、これら2つの制限を克服する。50

**【 0 0 6 9 】**

第1に、提供される学習方法は、物理的ノイズモデルを数学的に特定する必要がないため、より正確である。マシンにマシン自身でモデルをデータから直接学習するようにさせることができる。既存のモデリング方法とは対照的に、提供される学習方法は、十分に構築された学習データセットに依存する。

**【 0 0 7 0 】**

第2に、提供される学習方法は、あらゆる深度センサに働く汎用パイプラインを提供する。提供される学習方法はそれによって、あらゆるセンサのノイズを数学的にモデル化することを不要にする。パイプラインは、同じであり、ノイズが学習することになる深度センサタイプに従って学習データセットの変更を自身で制限することができる。

10

**【 0 0 7 1 】**

さらに、既存のアルゴリズムと異なり、ノイズ付加方法は、実空間と合成空間との間のマッピングがすでに学習したオフラインであるので、より高速であり、実時間性能を獲得する。

**【 0 0 7 2 】**

提供される方法、プログラム、データ構造、媒体およびシステムは、任意の深度センサをシミュレートすることを学習する。信号伝搬および反射物理特性を使用して物理的深度取得パイプラインをモデル化する代わりに、例示的な手法は、このステップを避けて、合成深度マップの空間と実「ノイズ有り」深度マップの空間との間のシフトを直接学習する。

20

**【 0 0 7 3 】**

実深度マップと合成深度マップの学習データセットから開始して、例示的な方法は、深層学習技術を通じて2つのモダリティ間のマッピング関数を学習する。後に論じるように、従来のニューラルネットワーク(CNN)技術が特にこの文脈において効果的である。

**【 0 0 7 4 】**

この処理は、オフラインで行われる。ひとたび合成空間と実空間とのマッピング関数が学習されると、それを直接オンライン方法で使用して合成深度データを実深度データにマップすることができる。このパイプラインは、図3に示されている。

**【 0 0 7 5 】**

学習方法は、例において2つの主なステップから成る。

## 1 . 学習データセット生成

30

## 2 . 回帰学習

**【 0 0 7 6 】**

学習データセット生成(Learning Dataset Generation)の例をこれより論じる。

**【 0 0 7 7 】**

手法の第1のステップは、実深度マップと合成深度マップのセットを生成することである。深度センサ(例えば、シミュレートしたい実深度センサ)を使用して、N個の厳選されたシーン、即ち、3Dで分かる形状をn個の異なるビューポイントでスキャンすることができる。静止識別3Dキャリブレーションオブジェクト(例えば、3Dチェス盤)および/または2Dキャリブレーションパターンを入力して、シーンをスキャンしている間は常にカメラに可視にすることができる。これを用いてカメラの3Dポーズを各主要なビューポイントにおいて計算することができ、取得された深度マップにおいて前記キャリブレーション要素を特定することによって、対応するノイズ無し深度マップを合成するようになる。カメラポーズの推定に、3Dキャリブレーションオブジェクトを使用する代わりに、周知の3Dシーンとそのシーンが十分テクスチャされる/識別されるかどうかスキャンされたシーンとの間の直接マッチングを使用する。これは、例えば、平面オブジェクトをマッチさせようとしても上手くいかない恐れがあることに留意されたい。3Dオブジェクトのシーンの代替手段は、異なる深度レベルにおいて、そのシーンの上にペーストされる2Dパターンの白い壁(例えば、チェス盤)をスキャンすることであろう。

40

**【 0 0 7 8 】**

図4は、データベース生成チャート(Database Generation Chart)の学習を例に沿って

50

示している。

**【 0 0 7 9 】**

各ビューポイントのカメラポーズが分かると、学習方法は、この同じポーズを、同じ実力メラの内部パラメータを使用して仮想シーンのカメラに適用することができる。学習方法はその後、各ビューポイントにおいて合成深度マップを抽出することができる。実深度マップとそれらに対応する合成深度マップのセットは、学習データセットを構成する。

**【 0 0 8 0 】**

図5は、学習方法の例のフローチャートを示している。

**【 0 0 8 1 】**

ユーザはS10において、コンピュータシステムのプロセッサに通信可能およびワイヤレスに結合された深度センサを提供される。深度センサは、物理的にユーザに取り付けられ、例えば、彼の／彼女の手に持つまたは彼の頭に取り付けられる。あるいは、ロボットデバイスが方法を実行することができる。他の例において、深度センサの同じモデルのいくつかのインスタンスをS10において例えば、数人のユーザに提供することができる。図5の方法のインスタンスは次に、例において並行して実行され得る。

10

**【 0 0 8 2 】**

ユーザはS20において、学習方法を開始する関数を起動し、これにより学習データセットを初期化する。これは、1または複数のシーンに実装されるスキームS30-S60をトリガする。図5のフローチャートは、フィードバックS25を示すが、学習方法はまた、単一のシーンにも実装され得る。

20

**【 0 0 8 3 】**

各シーンに対し、ユーザはS30において、シーンを入力して、シーンモデルをコンピュータシステムに提供する。シーンモデルは代替として、S30においてコンピュータシステムによって自動的に取得され得るか、またはS40など、それ以降に、例えば、データベースからの読み出しを経て深度測定に基づいて取得され得る。

**【 0 0 8 4 】**

学習方法の例はその後、1または複数のシーン領域に実行されるスキームS40-S50を備える。図5のフローチャートは、フィードバックS35を示すが、学習方法はまた、単一のシーン領域にも実装され得る。

30

**【 0 0 8 5 】**

各シーン領域に対し、ユーザは、シーン領域に近づいて、領域の深度ビデオを連続的に取得する（例えば、領域の方向に向かせて、領域に近づくおよび／または領域から後退する）。これは、同じシーン領域の異なる深度における取得S40をオペレートする。ビデオ取得は、シーン領域間で停止または維持され得る。取得された深度マップフレームは、S50において（例えば、自動的に）コンピュータシステムに送信される。

**【 0 0 8 6 】**

方法はその後、すべて受信されたノイズ有り深度マップに実行されるスキームS60を備える。スキームS60のインスタンスは、反復的に（即ち、連続して）および／または平行して実行され得る。

40

**【 0 0 8 7 】**

スキームS60は、コンピュータシステムによって自動的に実行され得る。スキームS60は、深度センサポジショニングの決定S602を備える。これは、それぞれの受信されたノイズ有り深度マップにおいて前述したようにキャリブレーションオブジェクトおよび／またはキャリブレーションパターンを特定することを通じて実行され得る。スキームS60はその後、そのようなポジショニングと提供されたシーンモデルとに基づいて（このシーンモデルは再度、ユーザによってすでに提供された可能性があるまたはシステムによって自動的に読み出されるか、またはこのシーンモデルは、このポイントまたはS50において送信および受信された深度マップに基づく任意のポイントにおいて読み出されることがある）、それぞれのノイズ有り深度マップに対応するノイズ無し深度マップを完全に仮想的に算出することS604ができる。スキームS60は、S606において算出され

50

たノイズ無し深度マップのそれぞれをそれに対応する取得されたノイズ有り深度マップに関連付けることによって学習データセットのフォーメーションを終了する。

#### 【 0 0 8 8 】

最終的に、図 5 の方法は、任意の方法で、形成された学習データセットに基づいて S 7 0 においてノイズ付加関数を学習する。

#### 【 0 0 8 9 】

図 6 は、ノイズ付加方法の例のフローチャートを示している。

#### 【 0 0 9 0 】

図 6 のノイズ付加方法は、深度センサポジショニングとシーンの所定のモデルに基づいてノイズ無し深度マップを仮想的に算出すること S 1 0 0 を備える。これは、ユーザによってまたは深度データ取得をシミュレートするシミュレーションプロセス内でトリガされ得る。所定のモデルおよび前記モデル内の深度センサポジショニングは、任意の方法で提供され得る。方法はその後、算出されたノイズ無し深度マップに、すでに学習したおよび企図された深度センサタイプに対応するノイズ付加関数を適用すること S 2 0 0 を備える。

10

#### 【 0 0 9 1 】

学習方法の関数学習ステップ（例えば、S 7 0 ）の例をこれより論じる。

#### 【 0 0 9 2 】

ノイズ付加関数の学習は、回帰学習によって実行され得る。このようなよく知られた学習は、簡単に実行することができ、従って高速でロバストである。

#### 【 0 0 9 3 】

ノイズ付加関数の学習はとりわけ、畳み込み層および / または逆畳み込み層を備える（例えば、ニューラル）ネットワーク空間内で実行され得る。言い換えれば、学習（例えば、S 7 0 ）は、信号 - 即ち、学習データセットのノイズ無し深度マップ - を符号化する畳み込み層のある数（例えば、3 以上、例えば、3 または 4 ）（例えば、およびまたは後続する）符号化された信号を復号化する逆畳み込み層に対応する（例えば、同じ）数で作られたネットワークを探索するように構成される。符号化 - 復号化パイプラインは、深度ノイズを付加するように構成される。

20

#### 【 0 0 9 4 】

それ自体が知られているように、学習は、探索を実行するパラメータを自由に変える。例において、層の数、層のシーケンス（例えば、畳み込み層の固定および所定の数、例えば、3 または 4 、その後同じ数の逆畳み込み層）、および / または各層によって実装されるフィルタの数およびサイズ（例えば、層単位で 9 × 9 フィルタおよび / またはそれがサイズ 3 × 3 のフィルタ）はすべて、学習の固定パラメータになり得る。学習の（例えば、唯一の）自由なパラメータはその後、フィルタ内部の係数になり得る。

30

#### 【 0 0 9 5 】

畳み込み層は、深度マップのノイズの学習に上手く適応する、局所的特徴の良好な抽出を可能にさせる。実際に、深度ノイズは、少なくともほとんどの場合、深度マップの局所撮動である。畳み込み層および / または逆畳み込み層は、例えば、サイズのそれぞれが 10 × 10 より下位のフィルタ、好適には 5 × 5 より下位のフィルタ、好適には 3 × 3 に等しいフィルタを実装することができる。これにより特に、特徴の抽出の良好な局所性が可能となり、従って正確で効率的な学習を可能にする。

40

#### 【 0 0 9 6 】

このような回帰学習の例を、図 7 を参照してこれよりさらに論じる。

#### 【 0 0 9 7 】

学習モデルは、エンドツーエンドの深層学習ベースのニューラルネットワークモデルにすることができる。このようなネットワークは、深度マップ画像を符号化する / 復号化するこの特定のタスクに対して上手く設計されている。ネットワークは、合成深度マップからノイズ有り深度マップを作り出すように学習され得る。

#### 【 0 0 9 8 】

ネットワークは、合成畳み込み層と逆畳み込み層のチェーンから成り得る。畳み込み層は

50

、特徴抽出および深度マップの符号化部分として機能する一方、逆畳み込み層は、デコーダおよび深度マップの再構築部分として機能する。

#### 【 0 0 9 9 】

層のパイプラインは、ブーリング層を有しない完全な畳み込みになり得る。その学習は、小さい畳み込み受容フィールドを用いて、小さい画像変化および薄いノイズのある輪郭エリアをカバーすることができる。

#### 【 0 1 0 0 】

実深度マップと合成深度マップとの間のユークリッド損失  $L$  の誤差逆伝搬を使用する、完全な教師付き(supervised)方法でネットワークを訓練することができる。

#### 【 0 1 0 1 】

#### 【 数 3 】

$$L = \frac{1}{2wh} \sum_{i=1}^w \sum_j^h \| \widehat{D}_y - D_{ij} \|$$

#### 【 0 1 0 2 】

ここに

#### 【 0 1 0 3 】

#### 【 数 4 】

$\widehat{D}, D$

10

#### 【 0 1 0 4 】

は、ノイズ有り深度マップ画像と合成深度マップ画像である。

#### 【 0 1 0 5 】

シーンの各モデルは、3Dモデル化オブジェクト、例えば、CADオブジェクトであってもよい。

#### 【 0 1 0 6 】

モデル化オブジェクトは、例えば、データベースに格納されたデータで定義される任意のオブジェクトである。拡大解釈すると、「モデル化オブジェクト」という表現は、データ自体を指す。使用時のシステムのタイプに従って、モデル化オブジェクトは、データの異なる種類によって定義され得る。システムは、実際にはCADシステム、CAEシステム、CAMシステム、PDMシステムおよび/またはPLMシステムの任意の組み合わせにすることができる。それらの異なるシステムにおいて、モデル化オブジェクトは、対応するデータによって定義される。それに応じてCADオブジェクト、PLMオブジェクト、PDMオブジェクト、CAEオブジェクト、CAMオブジェクト、CADデータ、PLMデータ、PDMデータ、CAMデータ、CAEデータを示してもよい。しかしながら、これらのシステムは、モデル化オブジェクトがこれらのシステムの任意の組み合わせに対応するデータによって定義され得るので、一方が他方を排除するシステムではない。システムは従って、以下に提供されるそのようなシステムの定義から明らかであるように、CADとPLMの両方のシステムであるのも当然である。

30

#### 【 0 1 0 7 】

CADシステムとは、付加的には、少なくともCATIAなど、モデル化オブジェクトのグラフィカル表現を基礎にしてモデル化オブジェクトを設計するために適応される任意のシステムを意味する。この場合、モデル化オブジェクトを定義するデータは、モデル化オブジェクトの表現を可能にさせるデータを備える。CADシステムは、例えば、エッジまたはライン、ある場合には面または表面を使用してCADモデル化オブジェクトの表現を提供することができる。ライン、エッジ、または表面は、さまざまな方法、例えば、非一様有理Bスプライン(NURBS)で表現され得る。具体的には、CADファイルは、形

40

50

状が生成され得る仕様を包含し、その仕様は次に、生成される表現を可能にさせる。モデル化オブジェクトの仕様は、単一の C A D ファイルまたは複数の C A D ファイルに格納され得る。C A D システムのモデル化オブジェクトを表すファイルの典型的なサイズは、1 パーツ当たり 1 メガバイトの範囲内である。そしてモデル化オブジェクトは典型的には、数千パーティから成るアセンブリにことができる。

#### 【 0 1 0 8 】

C A D の文脈において、モデル化オブジェクトは典型的には、3 D モデル化され得る。「3 D モデル化オブジェクト」とは、その 3 D 表現を可能にさせるデータによってモデル化される任意のオブジェクトを意味する。3 D 表現は、あらゆる角度からのパーティングを可能にさせる。例えば、3 D モデル化オブジェクトは、3 D 表現される時、操作されてその軸のいずれかを回転するか、または表現が表示されるスクリーンの任意の軸を回転する。これはとりわけ、3 D モデル化されない、2 D アイコンを除外する。3 D 表現の表示は、設計を容易にする（即ち、設計者が自分達のタスクを統計的に達成する速度が増大する）。

10

#### 【 0 1 0 9 】

方法は、コンピュータ実装される。これは、方法のステップ（または実質的にすべてのステップ）は、少なくとも 1 つのコンピュータ、または同様の任意のシステムによって実施されることを意味する。従って、方法のステップは、コンピュータによって場合により全自动的に、または半自動的に実行される。例において、方法のステップの少なくとも一部のトリガリングは、ユーザ - コンピュータインタラクションを介して実行され得る。必要とされるユーザ - コンピュータインタラクションのレベルは、予見される自動性のレベルに応じて異なり、ユーザの望みを実装する必要性とバランスを取ることができる。例において、このレベルは、ユーザ定義および / または事前定義にされてもよい。

20

#### 【 0 1 1 0 】

方法のコンピュータ実装の典型的な例は、この目的に適応されるシステムで方法を実行することである。システムは、メモリに結合されたプロセッサとグラフィカルユーザインターフェース（G U I ）を備えることができ、メモリは、方法を実行する命令を備えるコンピュータプログラムを記録している。メモリはまた、データベースを格納することもできる。メモリは、そのような格納に適応される任意のハードウェアであり、場合によりいくつかの物理的に別個の部分（例えば、1 つはプログラム用、場合により 1 つはデータベース用）を備える。システムは、プロセッサに結合されて 3 D 再構築のインスタンスの深度マップを取得するように構成される深度センサをさらに備えることができる。システムはまた、取得中にポーズするインスタンスのポージングスポットを備えることもできる。

30

#### 【 0 1 1 1 】

図 8 は、コンピュータシステムが深度センサに接続される時の、システムの例を示している。

#### 【 0 1 1 2 】

例示的なシステムは、内部通信 B U S 1 0 0 0 に接続された中央処理装置（C P U ）1 0 1 0 、またも B U S に接続されたランダムアクセスメモリ（R A M ）1 0 7 0 を備える。システムは、B U S に接続されたビデオランダムアクセスメモリ 1 1 0 0 と関連付けられるグラフィック処理装置（G P U ）1 1 1 0 がさらに備わる。ビデオ R A M 1 1 0 0 はまた、当業者にはフレームバッファとして知られている。マスクトレージデバイスコントローラ 1 0 2 0 は、ハードドライブ 1 0 3 0 など、マスマモリデバイスへのアクセスを管理する。コンピュータプログラム命令およびデータを有形に具現化するのに適したマスマモリデバイスは、あらゆる形態の不揮発性メモリを含み、例として E P R O M 、E E P R O M 、およびフラッシュメモリデバイスなど、半導体メモリデバイスと、内部ハードディスクおよびリムーバブルディスクなど、磁気ディスクと、磁気光ディスクと、C D - R O M ディスク 1 0 4 0 とを含む。上述のいずれも、特別に設計されたA S I C （特定用途向け集積回路）を補足するか、または組み込まれ得る。ネットワークアダプタ 1 0 5 0 は、ネットワーク 1 0 6 0 へのアクセスを管理する。システムはまた、深度センサ 1 0 9 0 を含

40

50

むこともできる。

【0113】

コンピュータプログラムは、コンピュータによって実施可能な命令を備えることができ、その命令は、上記システムに方法を実行させる手段を備える。プログラムは、システムのメモリを含む、任意のデータストレージ媒体に記録可能にすることができる。プログラムは例えば、デジタル電子回路に実装されるか、またはコンピュータハードウェア、ファームウェア、ソフトウェアに実装されるか、またはそれらの組み合わせにおいて実装されてもよい。プログラムは、装置、例えば、プログラマブルプロセッサによって実施されるためのマシン可読ストレージデバイスに有形に具現化される製品として実装され得る。方法のステップは、入力データを演算して出力を生成することによって方法の関数を実行する命令のプログラムを実施するプログラマブルプロセッサによって実行され得る。プロセッサは従って、データストレージシステム、少なくとも1つの入力デバイス、および少なくとも1つの出力デバイスからデータおよび命令を受信してそれらにデータおよび命令を送信するようにプログラマブルに結合され得る。アプリケーションプログラムは、ハイレベルの手続き型またはオブジェクト指向のプログラミング言語か、または必要に応じてアセンブリ言語またはマシン言語に実装され得る。いずれの場合も、言語は、コンパイル言語またはインタプリタ言語になり得る。プログラムは、完全インストールプログラムまたは更新プログラムになり得る。システムにプログラムを適用した結果、いずれの場合も方法を実行する命令が生じる。

10

20

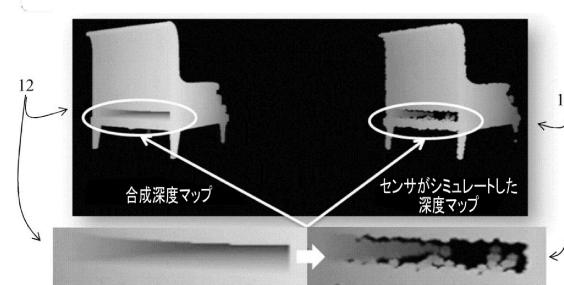
30

40

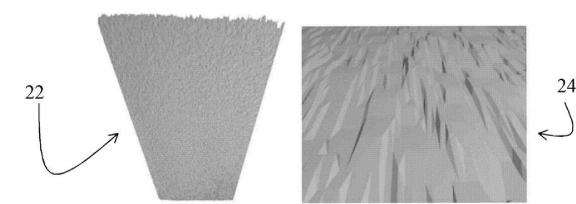
50

## 【図面】

## 【図 1】

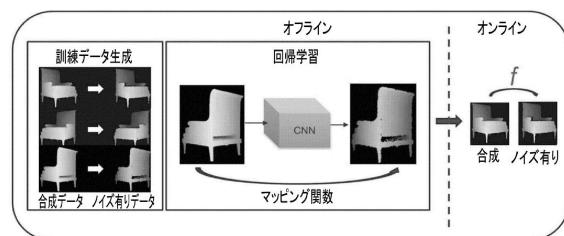


## 【図 2】

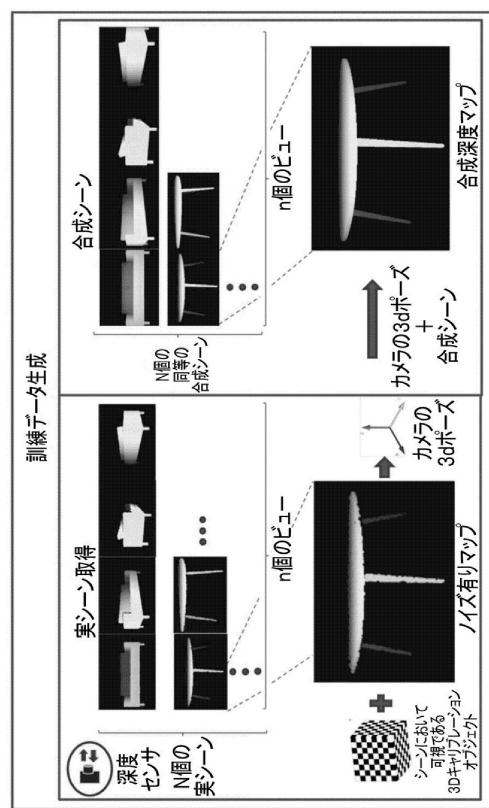


10

## 【図 3】



## 【図 4】



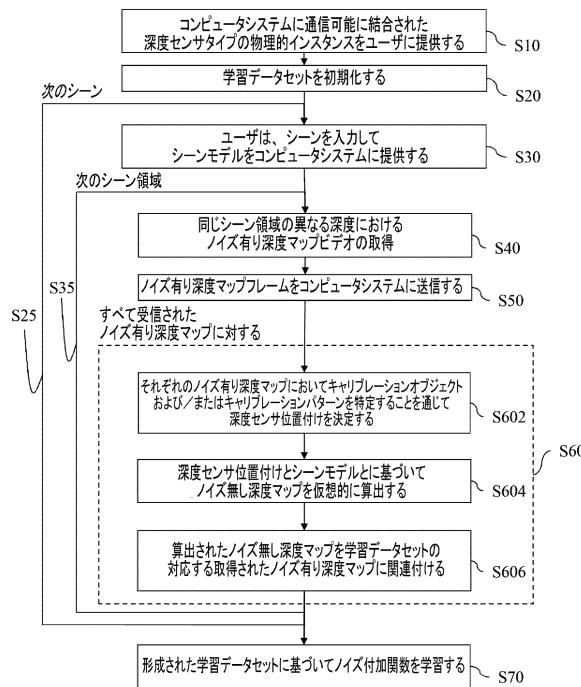
20

30

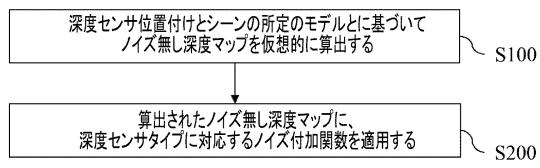
40

50

【図 5】



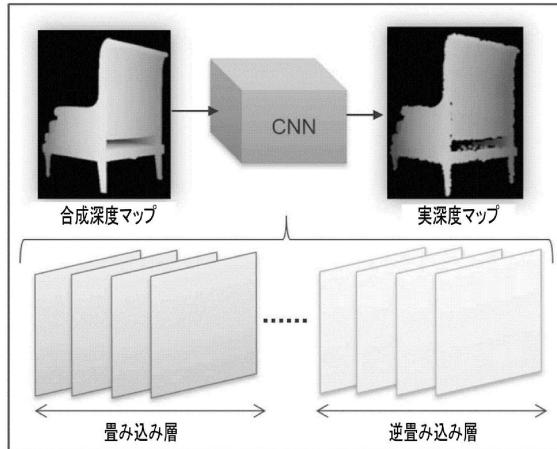
【図 6】



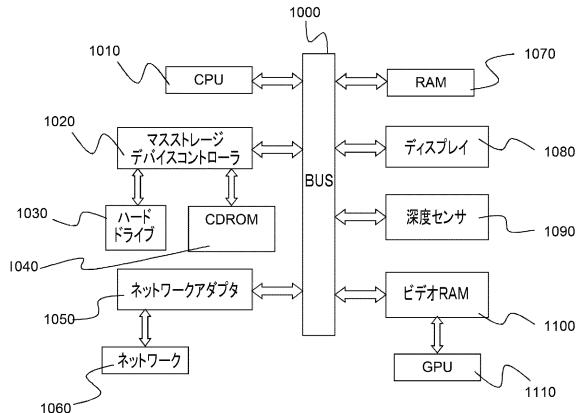
10

20

【図 7】



【図 8】



30

40

50

---

フロントページの続き

(56)参考文献

特開2014-199584(JP,A)  
特開2008-109640(JP,A)  
米国特許出願公開第2016/0239725(US,A1)  
金子 将也、外4名，“決定木を用いた距離画像からの高速なエッジ検出”，SSII2014 第20回画像センシングシンポジウム 講演論文集，日本，画像センシング技術研究会，2014年06月11日，pp.1-8  
Edward Johns et al. , "Deep learning a grasp function for grasping under gripper pose uncertainty" , 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) , 米国 , IEEE , 2016年10月09日 , pp.4461-4468  
Ankur Handa et al. , "A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM" , 2014 IEEE International Conference on Robotics and Automation (ICRA) , 米国 , IEEE , 2014年05月31日 , pp.1524-1531  
Michael J. Landau et al. , "Simulating Kinect Infrared and Depth Images" , IEEE Transactions on Cybernetics , 米国 , IEEE , 2015年11月13日 , Vol.46, No.12 , p.3018-3031

(58)調査した分野

(Int.Cl. , D B 名)

G 01 B 11 / 00 - 11 / 30  
G 06 T 1 / 00 , 7 / 00 - 7 / 90  
G 06 N 3 / 02 - 3 / 08 , 20 / 00