

(19)日本国特許庁(JP)

(12)公開特許公報(A)

(11)公開番号
特開2025-71766
(P2025-71766A)

(43)公開日 令和7年5月8日(2025.5.8)

(51)国際特許分類

F I

テーマコード (参考)

G 0 6 T 9/00 (2006.01)

G 0 6 T 9/00

5 C 1 5 9

H 0 4 N 19/597(2014.01)

H 0 4 N 19/597

審査請求		未請求	請求項の数	20	O L	外国語出願	(全25頁)
(21)出願番号	特願2024-105877(P2024-105877)		(71)出願人	000006013			
(22)出願日	令和6年7月1日(2024.7.1)			三菱電機株式会社			
(31)優先権主張番号	18/491,934			東京都千代田区丸の内二丁目 7 番 3 号			
(32)優先日	令和5年10月23日(2023.10.23)		(74)代理人	110001195			
(33)優先権主張国・地域又は機関	米国(US)			弁理士法人深見特許事務所			
			(72)発明者	ミラルド, ペドロ			
				アメリカ合衆国、マサチューセッツ州、			
				ケンブリッジ、ブロードウェイ 2 0 1			
				、ケアオブ・ミツビシ・エレクトリック			
				・リサーチ・ラボラトリーズ・インコー			
				ポレイテッド			
			(72)発明者	バイス, ゴンサロ			
				アメリカ合衆国、マサチューセッツ州、			
				ケンブリッジ、ブロードウェイ 2 0 1			
				、ケアオブ・ミツビシ・エレクトリック			
				最終頁に続く			

(54)【発明の名称】 3 D陰関数表現のための有向格子エンコーダ

(57)【要約】

【課題】ニューラルネットワークをトレーニングすることによって3次元(3D)オブジェクトを含む3Dシーンの陰関数表現を生成するための人工知能システムを提供する。

【解決手段】ニューラルネットワークは、オブジェクトの形状の3D点群を示すデータを、オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダと、デコーダとを含み、システムは、プロセッサと、命令が格納されたメモリとを備え、命令は、プロセッサに(i)3Dオブジェクトを含む3Dシーンの有向点群を示す入力データを受信させ、入力データは、3D点群の点の3D位置と、点の3D位置に近接する位置における3Dオブジェクトの表面に対する法線を定義する点の向きとを示し、命令は、さらに、プロセッサに、(i i)エンコーダおよびデコーダをトレーニングさせて、3Dオブジェクトの陰関数表現を生成させる。

【選択図】図 1

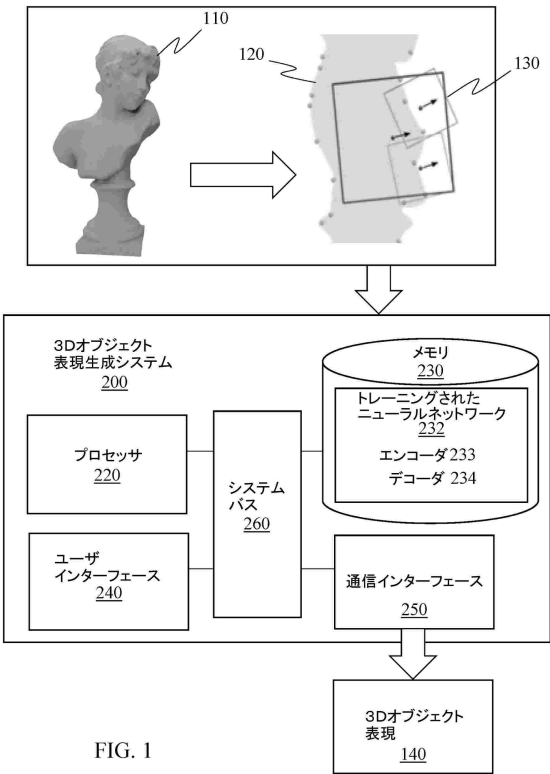


FIG. 1

【特許請求の範囲】

【請求項 1】

ニューラルネットワークをトレーニングすることによって 3 次元 (3 D) オブジェクトを含む 3 D シーンの陰関数表現を生成するための人工知能 (A I) システムであって、前記ニューラルネットワークは、前記オブジェクトの形状の 3 D 点群を示すデータを、前記オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダと、前記格子ベースの特徴を、前記 3 D シーンにおける任意の点から前記オブジェクトまでの距離にデコードするよう構成されるデコーダとを含み、前記 A I システムは、

少なくとも 1 つのプロセッサと、命令を格納したメモリとを備え、前記命令は、前記 A I システムの前記少なくとも 1 つのプロセッサに、

3 D オブジェクトを含む 3 D シーンの有向点群を示す入力データを受信させ、前記入力データは、前記 3 D 点群の点の 3 D 位置と、前記点の前記 3 D 位置に近接する位置における前記 3 D オブジェクトの表面に対する法線を定義する前記点の向きとを示し、前記命令は、さらに、前記 A I システムの前記少なくとも 1 つのプロセッサに、

前記 3 D オブジェクトの陰関数表現を生成するために、前記点の前記位置および前記点の前記向きの両方を使用して前記エンコーダおよび前記デコーダをトレーニングさせ、

前記エンコーダおよび前記デコーダを含む前記 3 D オブジェクトの前記陰関数表現を、有線または無線通信チャネルを介して送信させる、 A I システム。

【請求項 2】

前記エンコーダは、前記 3 D 点群の前記点の位置および前記点の前記向きのうちの 1 つまたは組合せを、前記オブジェクトの複数の分解能を捕捉する前記格子ベースの特徴に変換するようトレーニングされ、前記デコーダは、前記デコーダによって生成される前記 3 D シーン内の点から前記オブジェクトまでの距離とグラウンドトゥールス距離との間の誤差の損失関数を低減するよう、前記格子ベースの特徴の補間上でトレーニングされる、請求項 1 に記載の A I システム。

【請求項 3】

前記エンコーダは、前記 3 D 点群の前記点の位置を前記格子ベースの特徴に変換するようトレーニングされ、前記デコーダは、ネスト化された形状のセット内での前記補間上においてトレーニングされ、前記ネスト化された形状のセットは、前記格子ベースの特徴を囲み、対応するネスト化された形状の付近の前記点の前記向きに基づいて方向付けされる、請求項 2 に記載の A I システム。

【請求項 4】

前記エンコーダは、前記点の前記位置を、前記オブジェクトの複数の分解能を捕捉する前記格子ベースの特徴としてエンコードするようトレーニングされ、前記デコーダは、ネスト化された形状のセット内で補間によって表される方向付けされた特徴に基づいてトレーニングされ、前記ネスト化された形状のセットは、前記格子ベースの特徴を囲み、対応するネスト化された形状の付近の前記点の前記向きに基づいて方向付けされる、請求項 1 に記載の A I システム。

【請求項 5】

前記エンコーダおよび前記デコーダをトレーニングするために、前記プロセッサは、前記エンコーダを使用して、前記入力データを、前記 3 D オブジェクトの形状の複数の分解能を捕捉する特徴の八分木表現にエンコードするよう構成され、

前記八分木表現の各特徴を、軸周りに回転対称性を有する方向付けされた形状で囲むよう構成され、特徴を囲む方向付けされた形状の次元は、前記八分木表現上の囲まれた前記特徴のレベルによって支配され、前記方向付けされた形状の軸の向きは、囲まれた前記特徴の座標の近傍の点のサブセットの表面に対する法線によって支配され、前記プロセッサはさらに、

ボリュメトリック補間を用いて各方向付けされた形状内の特徴を補間して、前記八分木表現の前記特徴を更新するよう構成され、

10

20

30

40

50

前記デコーダを使用して、前記特徴の更新された前記八分木表現をデコードして、前記距離関数を生成するよう構成され、

前記ニューラルネットワークのパラメータを更新して、前記デコーダによって生成される前記 3 D シーン内の点から前記オブジェクトまでの距離とグラウンドトゥールス距離との間の誤差の損失関数を最小にするよう構成される、請求項 1 に記載の A I システム。

【請求項 6】

回転対称性を有する前記方向付けされた形状は、円筒および球体のうちの 1 つ以上を含む、請求項 5 に記載の A I システム。

【請求項 7】

回転対称性を有する前記方向付けされた形状の各々は円筒であり、前記円筒は、1 つまたは複数の格子ベースの特徴を囲み、前記円筒の各々の軸が、前記円筒の次元および囲まれた前記特徴の位置によって支配される表面の領域に対する法線に整列するように方向付けされた円筒である、請求項 5 に記載の A I システム。 10

【請求項 8】

回転対称性を有する前記方向付けされた形状の各々は、1 つまたは複数の格子ベースの特徴を囲む円筒であり、前記プロセッサは、前記円筒の次元および囲まれた前記特徴の位置によって支配される表面の領域に対する法線に前記円筒の軸を整列させるように前記円筒を方向付けるよう構成される、請求項 7 に記載の A I システム。

【請求項 9】

前記補間はボリュメトリック補間であり、前記プロセッサは、前記円筒形表現の末端に対する点の近接性を測定する円筒補間係数を求めるよう構成される、請求項 7 に記載の A I システム。 20

【請求項 10】

前記円筒補間係数は、(i) 前記円筒の頂面に対する前記点の距離と、前記円筒と前記円筒の対称軸に対する前記点の距離との体積における差とから計算される第 1 の係数と、(i i) 前記円筒の底面に対する前記点の距離と、前記円筒と前記円筒の前記対称軸に対する前記点の距離との体積における差とから計算される第 2 の係数と、(i i i) 前記円筒の残りから計算される第 3 の係数とを含む、請求項 9 に記載の A I システム。

【請求項 11】

前記ニューラルネットワークの前記エンコーダのトレーニング中に、前記プロセッサは、入力点群の複数のサンプリングされた点について、前記円筒補間係数を求めるよう構成される、請求項 10 に記載のシステム。 30

【請求項 12】

前記プロセッサは、前記 3 D オブジェクトの前記陰関数表現を使用して、表示装置上に前記 3 D オブジェクトの画像をレンダリングするよう構成される、請求項 1 に記載の A I システム。

【請求項 13】

請求項 1 に記載の前記 A I システムに、前記有線または無線通信チャネルを介して、動作するよう接続される画像処理システムであって、前記画像処理システムは、前記 3 D オブジェクトの前記陰関数表現を使用して、表示装置上に前記 3 D オブジェクトの画像をレンダリングするよう構成される、画像処理システム。 40

【請求項 14】

前記 3 D オブジェクトの前記画像は、変動する視野角に対してレンダリングされる、請求項 13 に記載の画像処理システム。

【請求項 15】

前記 3 D オブジェクトの前記画像は、仮想現実またはゲームアプリケーション内で、変動する視野角に対してレンダリングされる、請求項 14 に記載の画像処理システム。

【請求項 16】

請求項 1 に記載の前記 A I システムに、前記有線または無線通信チャネルを介して、動作するよう接続されるロボットシステムであって、前記ロボットシステムは、前記 3 D オ 50

プロジェクトの前記陰関数表現を使用してタスクを実行するよう構成される、ロボットシステム。

【請求項 17】

請求項 1 に記載の前記 AI システムに、前記有線または無線通信チャネルを介して、動作するよう接続される表示装置であって、前記プロセッサは、前記 3D オブジェクトの前記陰関数表現を使用して、前記 3D オブジェクトの画像をレンダリングするよう構成され、前記表示装置は、前記 3D オブジェクトのレンダリングされた前記画像を表示するよう構成される、表示装置。

【請求項 18】

3 次元 (3D) オブジェクトの画像を、前記 3D オブジェクトの陰関数表現を使用してディスプレイ上にレンダリングするよう構成される画像処理システムであって、

トレーニングされたニューラルネットワークを備え、前記トレーニングされたニューラルネットワークは、前記 3D オブジェクトの形状の 3D 点群を示すデータを、前記オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダと、前記格子ベースの特徴を、前記 3D オブジェクトを含む 3D シーン内の任意の点から前記オブジェクトまでの距離にデコードするよう構成されるデコーダとを含み、前記画像処理システムはさらに、

少なくとも 1 つのプロセッサと、命令を格納したメモリとを備え、前記命令は、前記少なくとも 1 つのプロセッサに、

3D オブジェクトを含む 3D シーンの有向点群を示す入力データを受信させ、前記入力データは、前記 3D 点群の点の 3D 位置と、前記点の前記 3D 位置に近接する位置における前記 3D オブジェクトの表面に対する法線を定義する前記点の向きとを示し、前記命令は、さらに、前記少なくとも 1 つのプロセッサに、

前記エンコーダで、前記点の前記位置および前記点の前記向きの両方を使用して、前記 3D オブジェクトの陰関数表現を生成させ、

前記エンコーダで、前記 3D オブジェクトの前記陰関数表現を使用して、前記 3D オブジェクトの画像をレンダリングさせ、

レンダリングされた前記画像をディスプレイに表示させる、画像処理システム。

【請求項 19】

前記エンコーダは、前記 3D 点群の前記点の位置を前記格子ベースの特徴に変換するようトレーニングされ、前記デコーダは、ネスト化された形状のセット内での前記補間上においてトレーニングされ、前記ネスト化された形状のセットは、前記格子ベースの特徴を囲み、対応するネスト化された形状の付近の前記点の前記向きに基づいて方向付けされる、請求項 18 に記載の画像処理システム。

【請求項 20】

前記 3D オブジェクトの前記画像は、変動する視野角に対してレンダリングされる、請求項 18 に記載の画像処理システム。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、概して、3次元シーンの陰関数表現を生成するための方法およびシステムを対象とする。

【背景技術】

【0002】

3次元 (3D) 表面を表現するには、多くの異なる方法がある。陰関数曲面表現では、座標 x 、 y 、および z を有する点は、 $F(x, y, z) = 0$ である場合に、あるオブジェクトに属し、関数 $F(\cdot)$ はそのオブジェクトを定義する。このタイプの 3D 表現は、簡潔で連続性を保証するので有利である。大抵の学習ベースの 3D 陰関数表現は、3D 点をエンコードすることから始まり、次いで、それらの特徴を選択された表現にデコードし、 $F(\cdot)$ を定義する。

10

20

30

40

50

【 0 0 0 3 】

2 種類のエンコーダ、すなわち、(1)ここでは位置エンコーダとして示される、各点の 3 D 座標のみを、より高次元のベクトル空間にマッピングすること、および(2)格子ベースと呼ばれる、それらの近傍についての情報を収集する 3 D 点、が通常並列に使用される。多層パーセプトロン(M L P)は、通常、デコーダのための好適な選択と見なされる。以前の、幾何学的エンコーダを使用する技法は、オブジェクトの基底の幾何学的特性の一部を考慮せず、その空間的局所化のみを利用し、したがって、不満足な 3 D 表現を形成する。

【 発明の概要 】

【 0 0 0 4 】

10

したがって、陰関数曲面表現を生成するときにオブジェクトの基底の幾何学的特徴のすべてを考慮する方法およびシステムに対する、継続的な、まだ対処されていないニーズが存在する。様々な実施形態および実現形態は、ニューラルネットワークをトレーニングすることによって 3 次元(3 D)オブジェクトを含む 3 D シーンの陰関数表現を生成するための方法およびシステムを対象とし、当該ネットワークは、当該 3 D オブジェクトの形状の 3 D 点群を示すデータを、当該オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダと、当該格子ベースの特徴を、当該 3 D シーンにおける任意の点から当該オブジェクトまでの距離にデコードするよう構成されるデコーダとを含む。

【 0 0 0 5 】

20

ある態様によれば、人工知能(A I)システムが提供される。本 A I システムは、ニューラルネットワークをトレーニングすることによって 3 次元(3 D)オブジェクトを含む 3 D シーンの陰関数表現を生成し、当該ニューラルネットワークは、当該オブジェクトの形状の 3 D 点群を示すデータを、当該オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダと、当該格子ベースの特徴を、当該 3 D シーンにおける任意の点から当該オブジェクトまでの距離にデコードするよう構成されるデコーダとを含み、当該 A I システムは、少なくとも 1 つのプロセッサと、命令を格納したメモリとを備え、当該命令は、当該 A I システムの当該少なくとも 1 つのプロセッサに、(i) 3 D オブジェクトを含む 3 D シーンの有向点群を示す入力データを受信させ、当該入力データは、当該 3 D 点群の点の 3 D 位置と、当該点の当該 3 D 位置に近接する位置における当該 3 D オブジェクトの表面に対する法線を定義する当該点の向きとを示し、当該命令は、さらに、当該 A I システムの当該少なくとも 1 つのプロセッサに、(i i) 当該 3 D オブジェクトの陰関数表現を生成するために、当該点の当該位置および当該点の当該向きの両方を使用して当該エンコーダおよび当該デコーダをトレーニングさせ、(i i i) 当該エンコーダおよび当該デコーダを含む当該 3 D オブジェクトの当該陰関数表現を、有線または無線通信チャネルを介して送信させる。

30

【 0 0 0 6 】

この入力データは、3 D オブジェクトを含む 3 D シーンの有向点群を示し、いくつかの方法で取得することができる。ある実施形態によれば、それは、点群を出力する R G B - D センサから取得される。3 D 点の向きは、いくつかの方法で取得することができ、例えば、近傍の点を平面によって局所的に近似することによって、またはニューラルネットワークを使用して、取得することができる。同じ種類の 3 D 点群は、他の可能性の中でもとりわけ、ピクセル当たりの深度を予測するニューラルネットワークを有するステレオカメラまたは単眼カメラなどの他のタイプのセンサから得ることができる。

40

【 0 0 0 7 】

入力データを取得する別の方法は、拡張現実 / 仮想現実に関連する。この実施形態によれば、オブジェクトを定義する三角メッシュがあり、3 D 点群は三角メッシュの角部から直接得ることができる。それぞれの 3 D 点の向きは、それぞれの 3 D 点近傍三角形の法線の平均によって与えられる。

【 0 0 0 8 】

50

一実施形態によれば、当該エンコーダは、当該3D点群の当該点の位置および当該点の当該向きのうちの1つまたは組合せを、当該オブジェクトの複数の分解能を捕捉する当該格子ベースの特徴に変換するようトレーニングされ、当該デコーダは、当該デコーダによって生成される当該3Dシーン内の点から当該オブジェクトまでの距離とグラウンドトゥールス距離との間の誤差の損失関数を低減するよう、当該格子ベースの特徴の補間上でトレーニングされる。

【0009】

ある実施形態によれば、当該エンコーダは、当該3D点群の当該点の位置を当該格子ベースの特徴に変換するようトレーニングされ、当該デコーダは、ネスト化された形状のセット内での当該補間上においてトレーニングされ、当該ネスト化された形状のセットは、当該格子ベースの特徴を囲み、対応するネスト化された形状の付近の当該点の当該向きに基づいて方向付けされる。

10

【0010】

ある実施形態によれば、当該エンコーダは、当該点の当該位置を、当該オブジェクトの複数の分解能を捕捉する当該格子ベースの特徴としてエンコードするようトレーニングされ、当該デコーダは、ネスト化された形状のセット内で補間によって表される方向付けされた特徴に基づいてトレーニングされ、当該ネスト化された形状のセットは、当該格子ベースの特徴を囲み、対応するネスト化された形状の付近の当該点の当該向きに基づいて方向付けされる。

20

【0011】

ある実施形態によれば、当該エンコーダおよび当該デコーダをトレーニングするために、当該プロセッサは、当該エンコーダを使用して、当該入力データを、当該3Dオブジェクトの形状の複数の分解能を捕捉する特徴の八分木表現にエンコードするよう構成され、当該八分木表現の各特徴を、軸周りに回転対称性を有する方向付けされた形状で囲むよう構成され、特徴を囲む方向付けされた形状の次元は、当該八分木表現上の囲まれた当該特徴のレベルによって支配され、当該方向付けされた形状の軸の向きは、囲まれた当該特徴の座標の近傍の点のサブセットの表面に対する法線によって支配され、当該プロセッサはさらに、ボリユメトリック補間を用いて各方向付けされた形状内の特徴を補間して、当該八分木表現の当該特徴を更新するよう構成され、当該デコーダを使用して、当該特徴の更新された当該八分木表現をデコードして、当該距離関数を生成するよう構成され、当該ニューラルネットワークのパラメータを更新して、当該デコーダによって生成される当該3Dシーン内の点から当該オブジェクトまでの距離とグラウンドトゥールス距離との間の誤差の損失関数を最小にするよう構成される。

30

【0012】

一実施形態によれば、回転対称性を有する当該方向付けされた形状は、円筒および球体のうちの1つ以上を含む。

【0013】

一実施形態によれば、回転対称性を有する当該方向付けされた形状の各々は円筒であり、当該円筒は、1つまたは複数の格子ベースの特徴を囲み、当該円筒の各々の軸が、当該円筒の次元および囲まれた当該特徴の位置によって支配される表面の領域に対する法線に整列するように方向付けされた円筒である。

40

【0014】

一実施形態によれば、回転対称性を有する当該方向付けされた形状の各々は、1つまたは複数の格子ベースの特徴を囲む円筒であり、当該プロセッサは、当該円筒の次元および囲まれた当該特徴の位置によって支配される表面の領域に対する法線に当該円筒の軸を整列させるように当該円筒を方向付けるよう構成される。

【0015】

一実施形態によれば、当該補間はボリユメトリック補間であり、当該プロセッサは、当該円筒形表現の末端に対する点の近接性を測定する円筒補間係数を求めるよう構成される。一実施形態によれば、当該円筒補間係数は、(i)当該円筒の頂面に対する当該点の距

50

離と、当該円筒と当該円筒の対称軸に対する当該点の距離との体積における差とから計算される第1の係数と、(i i) 当該円筒の底面に対する当該点の距離と、当該円筒と当該円筒の当該対称軸に対する当該点の距離との体積における差とから計算される第2の係数と、(i i i) 当該円筒の残りから計算される第3の係数とを含む。

【 0 0 1 6 】

一実施形態によれば、当該ニューラルネットワークの当該エンコーダのトレーニング中に、当該プロセッサは、入力点群の複数のサンプリングされた点について、当該円筒補間係数を求めるよう構成される。

【 0 0 1 7 】

一実施形態によれば、当該プロセッサは、当該3Dオブジェクトの当該陰関数表現を使用して、表示装置上に当該3Dオブジェクトの画像をレンダリングするよう構成される。

【 0 0 1 8 】

別の局面によれば、当該AIシステムに、当該有線または無線通信チャネルを介して、動作するよう接続される画像処理システムであって、当該画像処理システムは、当該3Dオブジェクトの当該陰関数表現を使用して、表示装置上に当該3Dオブジェクトの画像をレンダリングするよう構成される。

【 0 0 1 9 】

一実施形態によれば、当該3Dオブジェクトの当該画像は、変動する視野角に対してレンダリングされる。

【 0 0 2 0 】

ある実施形態によれば、当該3Dオブジェクトの当該画像は、仮想現実またはゲームアプリケーション内で、変動する視野角に対してレンダリングされる。

【 0 0 2 1 】

別の局面によれば、当該AIシステムに、当該有線または無線通信チャネルを介して、動作するよう接続されるロボットシステムであって、当該ロボットシステムは、当該3Dオブジェクトの当該陰関数表現を使用してタスクを実行するよう構成される。

【 0 0 2 2 】

別の局面によれば、当該AIシステムに、当該有線または無線通信チャネルを介して、動作するよう接続される表示装置であって、当該プロセッサは、当該3Dオブジェクトの当該陰関数表現を使用して、当該3Dオブジェクトの画像をレンダリングするよう構成され、当該表示装置は、当該3Dオブジェクトのレンダリングされた当該画像を表示するよう構成される。

【 0 0 2 3 】

別の局面によれば、3次元(3D)オブジェクトの画像を、当該3Dオブジェクトの陰関数表現を使用してディスプレイ上にレンダリングするよう構成される画像処理システムであって、トレーニングされたニューラルネットワークを備え、当該トレーニングされたニューラルネットワークは、当該3Dオブジェクトの形状の3D点群を示すデータを、当該オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダと、当該格子ベースの特徴を、当該3Dオブジェクトを含む3Dシーン内の任意の点から当該オブジェクトまでの距離にデコードするよう構成されるデコーダとを含み、当該画像処理システムはさらに、少なくとも1つのプロセッサと、命令を格納したメモリとを備え、当該命令は、当該少なくとも1つのプロセッサに、3Dオブジェクトを含む3Dシーンの有向点群を示す入力データを受信させ、当該入力データは、当該3D点群の点の3D位置と、当該点の当該3D位置に近接する位置における当該3Dオブジェクトの表面に対する法線を定義する当該点の向きとを示し、当該命令は、さらに、当該少なくとも1つのプロセッサに、当該エンコーダで、当該点の当該位置および当該点の当該向きの両方を使用して、当該3Dオブジェクトの陰関数表現を生成させ、当該エンコーダで、当該3Dオブジェクトの当該陰関数表現を使用して、当該3Dオブジェクトの画像をレンダリングさせ、レンダリングされた当該画像をディスプレイに表示させる。

【 0 0 2 4 】

10

20

30

40

50

様々な実施形態のこれらおよび他の曲面は、以下に記載される実施形態から明らかになり、それらを参照して説明されるであろう。

【0025】

図面において、同様の参照符号は、異なる図を通して、概して、同じ部分を指す。様々な実施形態を実現する特徴および態様を示す図は、特許請求の範囲内に入る他の可能な実施形態に限定するものとして解釈されるべきではない。また、図面は、必ずしも縮尺通りではなく、代わりに、概して、種々の実施形態の原理を図示することに重点が置かれる。

【図面の簡単な説明】

【0026】

【図1】一実施形態による3Dオブジェクト表現生成方法およびシステムの概略図である 10

【0027】

【図2A】一実施形態による、3Dオブジェクト表現生成システムをトレーニングするための方法のフローチャート図である。

【0028】

【図2B】一実施形態による、3次元シーンの陰関数表現を生成し、その表現をデコードするための方法のフローチャート図である。

【0029】

【図3A】一実施形態による、3Dオブジェクト表現を生成するときに向きを判断するためのプロセスの概略図である。 20

【0030】

【図3B】一実施形態による、3Dオブジェクト表現を生成するときに向きを判断するためのプロセスの概略図である。

【0031】

【図3C】一実施形態による、3Dオブジェクト表現を生成するときに向きを判断するためのプロセスの概略図である。

【0032】

【図4A】一実施形態による、構造格子をオブジェクトの法線方向とともに使用して八分木表現を拡張する多分解能有向格子の概略図である。

【0033】

【図4B】一実施形態による、点の向きを含む3Dオブジェクトの有向点群を生成する概略図である。 30

【0034】

【図5】一実施形態による、3次元シーンの陰関数表現を生成するための方法のフローチャート図である。

【0035】

【図6】一実施形態による、円筒補間のための提案された方式を示す図である。

【0036】

【図7】一実施形態による、アブレーション研究の結果の表を示す図である。

【0037】

【図8】一実施形態による、アブレーション研究の結果の概略図である。 40

【0038】

【図9】一実施形態による、構造格子対有向格子を比較する表を示す図である。

【0039】

【図10】一実施形態による大規模シーンレンダリングの概略図である。

【0040】

【図11】一実施形態による、他のタイプの表現に対する本方法の実験結果を詳述する表を示す図である。

【発明を実施するための形態】

【0041】

本開示は、３次元（３Ｄ）シーンの陰関数表現を生成するよう構成されるシステムおよび方法の様々な実施形態を説明する。より一般的には、本出願人は、陰関数曲面表現を生成するときにオブジェクトの基底の幾何学的特徴のすべてを考慮する方法およびシステムを提供することが有益であろうことを認識し、理解した。したがって、ある陰関数表現生成システムは、ニューラルネットワークをトレーニングすることによって、３Ｄオブジェクトを含む３Ｄシーンの陰関数表現を生成する。ニューラルネットワークは、３Ｄオブジェクトの形状の３Ｄ点群を示すデータを、３Ｄオブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダを含む。ニューラルネットワークはまた、格子ベースの特徴を、３Ｄシーン内の任意の点からオブジェクトまでの距離にデコードするよう構成されるデコーダを含む。

10

【００４２】

３Ｄ陰関数表現パイプラインでは、単一の３Ｄ点が、幾何学的エンコーダ、位置エンコーダ、またはそれら両方を通して。次いで、特徴は、オブジェクトの表面をモデル化するデコーダに注入される。当該プロセスをすべての点群点について繰り返すことにより、モデル化された３Ｄ表面に関して、疎な出力表現が得られる。

【００４３】

図１を参照すると、一実施形態では、３Ｄオブジェクト表現生成方法およびシステムの概略図が示されている。図に関連して記載される方法は、例としてのみ提供され、本開示の範囲を限定しないと理解されるべきである。陰関数表現生成システム２００は、本明細書において記載されるかまたは別様に想定されるシステムのうちのいずれかであることができる。本陰関数表現生成システムは、単一のシステムまたは複数の異なるシステムであり得る。

20

【００４４】

３Ｄオブジェクト表現生成システム２００は、３Ｄオブジェクトを含む３Ｄシーンの陰関数表現を生成するための人工知能（ＡＩ）システムである。システム２００は、（１）オブジェクト１１０の形状の３Ｄ点群１２０を示すデータを、オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするよう構成されるエンコーダ２３３と、（２）格子ベースの特徴を３Ｄシーン内の任意の点からオブジェクトまでの距離にデコードするよう構成されるデコーダ２３４とを含むニューラルネットワーク２３２をトレーニングする。

30

【００４５】

システム２００は、３Ｄオブジェクト１１０を含む３Ｄシーンの有向点群１２０を示す入力データを受信し、入力データは、３Ｄ点群の点の３Ｄ位置と、点の３Ｄ位置に近接する位置における３Ｄオブジェクトの表面に対する法線を定義する点の向き１３０とを示す。システム２００は、３Ｄオブジェクトの陰関数表現を生成するために、点の位置および点の向きの両方を使用してエンコーダ２３３およびデコーダ２３４をトレーニングする。本システムは、次いで、通信インターフェース２５０などを介して、３Ｄオブジェクトの陰関数表現１４０（任意選択肢としてエンコーダおよびデコーダを含む）を、有線または無線通信チャネルを介して、送信することができる。

【００４６】

40

ある実施形態によれば、陰関数表現生成システム２００が提供される。図１に描写されるような陰関数表現生成システム２００の実施形態を参照すると、例えば、本システムは、１つまたは複数のシステムバス２６０を介して相互接続される、プロセッサ２２０、メモリ２３０、ユーザインターフェース２４０、および通信インターフェース２５０のうちの１つまたは複数を含む。図１は、いくつかの点において抽象概念を構成し、システム２００の構成要素の実際の編成は、図示されるものとは異なる、より複雑であり得ることが理解されるであろう。加えて、陰関数表現生成システム２００は、本明細書に記載されるかまたは別様に想定されるシステムのうちのいずれかであることができる。陰関数表現生成システム２００の他の要素および構成要素は、本明細書の他の箇所が開示および／または想定される。

50

【 0 0 4 7 】

一実施形態によれば、システム 2 0 0 は、メモリ 2 3 0 に記憶される命令を実行するか、または他の方法でデータを処理して、例えば、本方法の 1 つまたは複数のステップを実行することができるプロセッサ 2 2 0 を備える。プロセッサ 2 2 0 は、1 つまたは複数のモジュールから形成され得る。プロセッサ 2 2 0 は、マイクロプロセッサ、マイクロコントローラ、複数のマイクロコントローラ、回路系、フィールドプログラマブルゲートアレイ (F P G A)、特定用途向け集積回路 (A S I C)、単一のプロセッサ、または複数のプロセッサを含むがそれらに限定されはしない任意の好適な形態をとることができる。

【 0 0 4 8 】

メモリ 2 3 0 は、不揮発性メモリおよび / または R A M を含む、任意の好適な形態をとることができる。メモリ 2 3 0 は、例えば、L 1、L 2 もしくは L 3 キャッシュまたはシステムメモリなどの様々なメモリを含み得る。したがって、メモリ 2 3 0 は、スタティックランダムアクセスメモリ (S R A M)、ダイナミック R A M (D R A M)、フラッシュメモリ、読取り専用メモリ (R O M)、または他の同様のメモリデバイスを含み得る。メモリは、とりわけ、オペレーティングシステムを記憶することができる。R A M は、データの一時的な記憶のためにプロセッサによって使用される。一実施形態によれば、オペレーティングシステムは、プロセッサによって実行されると、システム 2 0 0 の 1 つまたは複数の構成要素の動作を制御するコードを含むことができる。プロセッサが本明細書で記載される機能のうちの 1 つまたは複数のハードウェアで実現する実施形態では、他の実施形態においてそのような機能に対応するものとして記載されるソフトウェアは省略され得ることは明らかであろう。メモリ 2 3 0 は、非一時的な機械可読媒体と見なされ得る。本明細書で使用するとき、非一時的という文言は、一時的な信号を除外するが、揮発性メモリおよび不揮発性メモリの両方を含むすべての形態のストレージを含むと理解される。一実施形態によれば、システム 2 0 0 のメモリ 2 3 0 は、1 つもしくは複数のアルゴリズム、モジュール、および / または命令を記憶して、本明細書で記載または別様に想定される方法の 1 つもしくは複数の機能またはステップを実行することができる。

【 0 0 4 9 】

ユーザインターフェース 2 4 0 は、ユーザとの通信を可能にするための 1 つまたは複数のデバイスを含んでもよい。ユーザインターフェースは、情報が伝達および / または受信されることを可能にする任意のデバイスまたはシステムであり得、ユーザコマンドを受信するためのディスプレイ、マウス、および / またはキーボードを含み得る。いくつかの実施形態では、ユーザインターフェース 2 4 0 は、通信インターフェース 2 5 0 を介して遠隔端末に提示され得る、コマンドラインインターフェースまたはグラフィカルユーザインターフェースを含んでもよい。ユーザインターフェースは、本システムの 1 つまたは複数の他の構成要素とともに位置してもよく、または本システムから遠隔に位置し、有線および / もしくは無線通信ネットワークを介して通信してもよい。

【 0 0 5 0 】

通信インターフェース 2 5 0 は、他のハードウェアデバイスとの通信を可能にするための 1 つまたは複数のデバイスを含み得る。たとえば、通信インターフェース 2 5 0 は、イーサネット (登録商標) プロトコルに従って通信するよう構成されるネットワークインターフェースカード (N I C) を含み得る。加えて、通信インターフェース 2 5 0 は、T C P / I P プロトコルに従う通信のために T C P / I P スタックを実現し得る。通信インターフェース 2 5 0 のための様々な代替もしくは追加のハードウェアまたは構成が明らかであろう。

【 0 0 5 1 】

システム 2 0 0 は、記載した各構成要素のうちの 1 つを含むものとして示されているが、様々な実施形態では、様々な構成要素が重複され得る。たとえば、プロセッサ 2 2 0 は、複数のプロセッサが協働して本明細書で記載する機能を達成するように、本明細書で記載する方法を独立して実行するよう構成されるか、または本明細書で記載する方法のステップもしくはサブルーチンを実行するよう構成される、複数のマイクロプロセッサを含み

10

20

30

40

50

得る。さらに、システム 200 の 1 つまたは複数の構成要素がクラウドコンピューティングシステムに実現される場合、様々なハードウェア構成要素は、別々の物理システムに属し得る。たとえば、プロセッサ 220 は、第 1 のサーバ内の第 1 のプロセッサと、第 2 のサーバ内の第 2 のプロセッサとを含み得る。多くの他の変形および構成が可能である。

【0052】

図 2 A を参照すると、一実施形態では、3D オブジェクト 110 を含む 3D シーンの陰関数表現を生成するように 3D オブジェクト表現生成システム 200 をトレーニングするための方法 250 のフローチャートである。トレーニング 260 中、本明細書に記載するように、本システムは、エンコーダ 233 およびデコーダ 234 を含むニューラルネットワークをトレーニングして、八分木表現を使用して、オブジェクトの形状の 3D 点群を示すデータを、オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードし、格子ベースの特徴を、3D シーン内の任意の点からオブジェクトまでの距離にデコードし、それにより、オブジェクトの表面をモデル化する。

10

【0053】

図 2 B は、一実施形態による、3次元シーンの陰関数表現を生成し、その表現をデコードするための方法のフローチャートである。本システムのニューラルネットワーク 232 は、オブジェクト 270 の形状の 3D 点群を示すデータを、オブジェクトの複数の分解能を捕捉する格子ベースの特徴にエンコードするようトレーニングされるエンコーダ 233 を含む。ニューラルネットワークはまた、格子ベースの特徴を 3D シーン内の任意の点からオブジェクトまでの距離にデコードするようトレーニングされるデコーダ 234 を含み、それによって、3D オブジェクトの陰関数表現 140 を介してオブジェクトの表面をモデル化する。3D オブジェクトのこの陰関数表現は、直ちに使用することができ、通信することができ、および / または将来の使用のために記憶することができる。実行 280 中、トレーニングされたニューラルネットワークは、トレーニングされたデコーダ 234 を使用して、3D オブジェクトのエンコードされた陰関数表現 140 をデコードするよう、使用される。

20

【0054】

図 3 A ~ 図 3 C を参照すると、一実施形態では、一実施形態による、3D オブジェクト表現を生成するときに向きを判断するためのプロセスの概略図が示されている。トレーニングデータを生成し、トレーニングされたエンコーダを使用して 3D オブジェクトをエンコードするために、本システムは、3D オブジェクトを含む 3D シーンの有向点群を示す入力データを生成する。その入力データは、3D 点群の点の 3D 位置と、点の 3D 位置に近接する位置における 3D オブジェクトの表面に対する法線を定義する点の向きとを示す。

30

【0055】

例えば、図 3 A を参照すると、幾何学的エンコーダは、異なる詳細レベルが明示的にモデル化されるときに、より良好に機能する。3D データ点は、310 において格子セルに示されるように、異なる空間分解能で配置される。各分解能に対して、異なるサイズを有する異なるセルが存在する。これらの異なる格子分解能は、それぞれのセル内にあるすべてのデータ点を集約するために使用される。数字 330 および 340 は、2 つの異なる分解能を示す。

40

【0056】

ある実施形態によれば、本システムは、5 つの異なる格子分解能を利用し、多分解能格子エンコーダを作成する。しかしながら、異なる格子分解能の数は、5 より多くても少なくともよい。例えば、異なる格子分解能の数は、オブジェクト / 表面の粗さに依存してもよく、したがって、それに応じて調整することができる。一実施形態によれば、単一のデコーダを、異なる多分解能格子エンコーダに使用することができる。

【0057】

図 3 B および図 3 C を参照すると、ツリー 380 は、単一の回転事例の玩具例について向きがどのように計算されるかを示す。実際には、3 つの回転があり、ツリー分岐係数を

50

6に増加させる。ツリーの深度の中に、より深く進むにつれて、向きは、380(図3B)および340(図3C)において異なるグレースケールとして図示される、正しい表面方向に、より近づくであろう。ツリーにおける深度の各レベルは、単一の分解能格子についての向きを設定する。

【0058】

図3A~図3Cを参照すると、例えば、セル340および330は、380において計算される向き(すなわち、ツリーにおけるそれぞれのグレースケール/深度)に従って回転され、回転されたセル360および370(図3C)をそれぞれ得る。この手順は、異なる分解能を含んで、各すべてのセルについて繰り返される。

【0059】

実施形態によれば、トレーニングのための入力3Dデータを生成するために、点の3D位置は、最初にオブジェクト表面により近くサンプリングし、次いで環境の周囲に、より分散して追加することによって、シーン全体にわたって射倅的に取得される。

【0060】

図4Aを参照すると、構造格子をオブジェクトの法線方向とともに使用して八分木表現を拡張する多分解能有向格子を生成するためのプロセスの概略図が示されている。セル410は、方向ツリーおよびそれぞれの深度レベル(LOD)レベルに回転され、有向格子420を形成する。

【0061】

図4Bを参照すると、オブジェクト430をエンコードする概略図が示されている。トレーニング中、例えば、(点群420内の)回転されたセル内のサンプリングされた点の特徴は、近傍のセル特徴が3DCNN疎カーネルで集約される、422で示される円筒補間方式に従って補間される。これらの特徴は、符号付き距離関数(SDF)および占有率といったオブジェクト表現のための現在の現状技術のデコードにおいて使用され得る。図4Bに関して円筒補間方式が記載されるが、軸周りに回転対称性を有する任意の形状が好適であることに留意することが重要である。例えば、それは、軸周りに回転対称性を有する球体または他の体積形状であり得る。

【0062】

図5を参照すると、一実施形態では、トレーニングされたニューラルネットワークのエンコーダを使用してオブジェクトをエンコードするための方法500が示されている。エンコーダは、入力510として受信された点群内の点の向きをエンコードする有向格子幾何学的エンコーダ520と、入力510として受信された点群データをエンコードする点エンコーダ530とを含む。各3D点は、幾何学的エンコーダ520、位置エンコーダ530、またはそれら両方を通して。ある実施形態によれば、入力510は、ある点、およびオブジェクトに最もよく適合する事前初期化されたツリーである。出力540は、3D表現を形成する深度レベル(LOD)ツリー特徴のセットである。

【0063】

点エンコーダ530は、位置エンコーダおよびアンカーの法線の両方を含む。有向格子幾何学的エンコーダ520は、550で有向格子構築から始まり、これは事前計算されたステップである。これは、有向格子の複数の分解能を含む。例えば、図5は、3つの分解能552を示すが、より少数の分解能およびより多数の分解能も可能である。各分解能について、(ツリーから抽出される)有向格子の特徴554が生成される。これらの特徴は、近傍情報を横断して局所的に集約される(すなわち、局所の特徴集約560)。集約された特徴562は、本明細書で記載されるかまたは別様に想定される円筒補間方式570において使用されて、最終的なLOD特徴580を生成する。LODツリー特徴580および点エンコーダ530からのエンコードされた点は、3D表現540をもたらす。図5の方法は円筒補間方式570を利用するが、軸周りに回転対称性を有する任意の形状が好適であることに留意されたい。例えば、それは、軸周りに回転対称性を有する球体または他の体積形状であり得る。

【0064】

10

20

30

40

50

ある実施形態によれば、有向格子が構築される。3D表現をモデル化するために八分木表現が使用される。具体的には、本システムは、格子ベースの3Dエンコーダをモデル化するために八分木表現を使用する。しかしながら、ある格子を後続の深度レベルにおいて8つのより小さい格子に分割するための標準的な8つのアクションに加えて、本システムは、セルの向きをモデル化するための回転アクションを含み、より高いレベルでは、より小さい（より密な）格子およびより微細な整列が、オブジェクトをより良好に表す。後続のLODごとに56の分岐因子（格子位置について8×向きについて7）をもたらすであろう、各アクションの個々のモデル化の代わりに、そして、格子サイズおよび向きは独立しているので、それらは2つのツリー、すなわち、(i) ツリー1：格子のサイズをモデル化するための構造化された八分木と、(ii) ツリー2：セルの向きをモデル化するための方向ツリーとに分割される。

10

【0065】

ツリー1における構造化された八分木について、その表現は、 $[-1, 1]$ 内に境界付けられる位置LOD（向きなし）からなる。典型的な八分木モデル化は、既存の研究から辿られる。

【0066】

方向ツリー2について、オブジェクトの表面点群から取られる正規化された点xについて、法線が、nとして示されて、この照会に関連付けられる。目標は、表面に沿ってセルを整列させることである。LOD内の整合性を維持するために、レベルごとに有限の可能な向きのセットを表す法線アンカーのセットを構築した。次いで、z軸が、照会法線nに最も近いアンカーであるアンカーに一致するように、セルを回転させた。この検索ツリーをモデル化するために、i) ノード状態、ii) アクション、iii) 状態遷移、およびiv) 初期状態を定義する必要がある。

20

【0067】

【数1】

一実施形態によれば、 δ として示されるノード状態は、3つの2タプル要素のセットによって定義される：

$$\delta = (r_{x-}, r_{x+}), (r_{y-}, r_{y+}), (r_{z-}, r_{z+}) \quad (\text{式1})$$

7つの可能なアクションがある：

30

$$\mathcal{A} = \{0, x_-, x_+, y_-, y_+, z_-, z_+\} \quad (\text{式2})$$

LOD1+1における状態は、LOD1における状態から、アクション x_- および x_+ について、以下の状態遷移を用いてアクション $a \in \mathcal{A}$ を適用した後に得られる：

$$\delta_{\alpha}^{l+1} = \mathcal{T}(\delta^l, \alpha): \quad (\text{式3})$$

$$\mathcal{T}(\delta^l, x_-) = \left(r_{x-}^l, \frac{r_{x-}^l + r_{x+}^l}{2} \right), (r_{y-}^l, r_{y+}^l), (r_{z-}^l, r_{z+}^l) \quad (\text{式4})$$

$$\mathcal{T}(\delta^l, x_+) = \left(\frac{r_{x-}^l + r_{x+}^l}{2}, r_{x+}^l \right), (r_{y-}^l, r_{y+}^l), (r_{z-}^l, r_{z+}^l) \quad (\text{式5})$$

40

アクション0は、回転なし、すなわち、 $\delta_{\alpha}^{l+1} = \delta^l$ を意味する。残りのアクションは、上記の式から導出することができる。初期状態は、 $(-\pi, \pi), (-\pi, \pi), (-\pi, \pi)$ として設定される。

【0068】

【数2】

50

方向ツリーモデル化は、回転を定義することによって終了する。回転は、状態 δ から LOD およびセルごとに得られる。回転アンカーは、状態のオイラー角範囲 (r_x, r_y, r_z) から定義され、

$$r_i = (r_{i+} + r_{i-})/2, \forall i \in \{x, y, z\} \quad (\text{式 6})$$

である。

【 0 0 6 9 】

各セルについて状態 を計算するために、回転アンカーは、コサイン類似度を使用して 10、ある回転自由度まで、セルの z 軸を表面法線と整列させることができる。

【 0 0 7 0 】

格子と照会点との関連付けに関して、ツリー 1 内の各セルは、ツリー 2 を検索することから計算される固定された向きを有する。トレーニングおよび評価中、照会点は、ツリー 1 (構造化されたツリー) 内の特定の LOD 上のセルに関連付けられる。次いで、対応する回転アンカーを用いてセルを回転させる。なお、点は、すべての八分木セルの外側にあってもよく；この場合、照会は破棄される。

【 0 0 7 1 】

トリリニア補間は、図 3 A に示すように、構造 (方向付けなし) 格子についての特徴を得る典型的な方法であったが、有向格子に対しては、構造格子と同じ手法を使用することは適用できない。したがって、本システムは、図 4 B に示されるように方向付けされた円筒 (これは、図 3 C における有向格子の 3 D 表現である) を使用し、それは、本明細書で議論されるように、セルの整列を活用し、格子の向き (法線方向周りの不変性回転) を規定する際に不変性の欠如を軽減することができる。この回転不変性は、格子内の点に明示的な滑らかさの制約を加える。 20

【 0 0 7 2 】

したがって、図 6 を参照すると、ある実施形態による、円筒補間のための提案された方式がある。図 6 の方法は円筒補間方式を利用するが、軸周りの回転対称性を有する任意の形状が好適であることに留意されたい。例えば、それは、軸周りの回転対称性を有する球体または他の体積形状であり得る。 30

【 0 0 7 3 】

入力セル格子は、(LOD ごとに) 有向格子から得られる対応するアンカー法線 n_{610} を有する。円筒 620 は、半径 R および高さ H を伴って、格子法線アンカー 610 と整列される。補間方式は、ボリュメトリック補間タイプのものである。これは、円筒の高さ境界までの照会点 x の距離 h_1 および h_2 、ならびに r として示される x と円筒対称軸との間の距離に依存する。

【 0 0 7 4 】

630 において、第 1 の係数 c_0 は、点の頂面までの距離 h_1 と、 R および点の対称軸までの距離 r を考慮した体積の差とから計算される。640 において、係数 c_2 は、底面までの距離 h_2 と、 R および点の対称軸までの距離 r を考慮した体積の差とから計算される。最後に、650 において、 c_1 は円筒の残りである。各係数は、 $k = \{0, 1, 2\}$ について、付随する学習可能な特徴 e_k を有する。補間された特徴 f は、 e_k の、 c_k 重みとの加重平均である。660 において、円筒補間特徴が、 c_0 、 c_1 、および c_2 から生成される。 40

【 0 0 7 5 】

照会点が与えられると、目的は、円筒内の点の相対位置を考慮して、特徴係数について相対空間体積を計算することである。円筒補間係数は、図 6 に示すように、円筒セル表現の末端に対する点の近接度を測定する。円筒の頂部および境界線により近い点では、より小さい境界体積 (図 6 の 630 における体積) を生成することになる。したがって、その反対面からの距離は大きくなり、したがって、より大きい体積係数 (図 6 の 640 に 50

ける体積)となる。この例における最も高い係数は、中心軸に従って反対になる(図6の650における体積)。

【0076】

【数3】

最後に、ボリュメトリック補間は、3つの係数 c_0 、 c_1 、および c_2 を与える(図6における導出)。これらの係数を用いて、学習可能な特徴 $e^l = \{e_0^l, e_1^l, e_2^l\}$ が、セルごと(単純化のため、セル添え字は省略される)およびLOD1ごとに、照会点特徴を補間するために使用される(線形結合)：

$$f^l = (c_0 e_0^l + c_1 e_1^l + c_2 e_2^l) / (c_0 + c_1 + c_2) \quad (式7)$$

10

提案される補間方式のみによって得られる特徴は、近傍およびLOD情報を欠いている。補間セルの角部からの特徴が近傍およびレベル間で共有されるトリリニア補間とは対照的に、円筒補間は、その近傍および他のLOD間の知識を本来組み込まない。したがって、局所的特徴集約のための共有されるレベル横断3DCNNの使用が提案され、ここでは $g_{\theta_l}(\cdot)$ として示される。特徴レベルごとの畳み込みカーネル \mathcal{K}_k 、ここで、 $k = 0, 1, 2$ は、ツリーにわたって共有される。各セルにおける各特徴 e_k^l に対して、関連付けられるカーネル \mathcal{K}_k が存在し：

$$\overline{e_k^l} = g_{\theta_l}(e_k^l, \mathcal{N}^l(x), \mathcal{K}_k), \forall l \quad (式8)$$

20

ここで、 $\mathcal{N}^l(\cdot)$ は近傍セルのセットである。3Dスパース畳み込みが利用され、すなわち、近傍は、存在しない場合、無視される。近傍は、現在のセルを中心として、カーネル内に近傍セルを含む。

【0077】

最後に、図6の幾何学的エンコーダからの補間された特徴は、係数および3DCNNから計算された局所的特徴の線形補間である：

【数4】

$$\bar{f} = \frac{c_0 \overline{e_0^l} + c_1 \overline{e_1^l} + c_2 \overline{e_2^l}}{c_0 + c_1 + c_2} \quad (式9)$$

30

【0078】

実現形態において、本方法は、新たな格子ベースのエンコーダに焦点を当てているため、本システムは、現状技術のデコーダアーキテクチャおよび出力表現を使用して、本方法を評価する。損失関数およびトレーニング手順についても以下で説明する。

【0079】

デコーダアーキテクチャでは、多層パーセプトロンが利用される。デコーダは、各レベルでトレーニングされ、すべてのLODにわたって共有される。入力される補間された特徴の他に、現状技術の位置エンコーダ $\phi_p(\cdot)$ が点上において L_p 個の周波数とともに追加され、 $\phi_n(\cdot)$ がアンカーの法線上において L_n 個の周波数とともに追加される。点および法線は各位置エンコーダに付され、サイズは、 $P = 3 \times 2 \times L_p + 3$ および $N = 3 \times 2 \times L_n + 3$ である。この方法は、SDFおよび占有率について出力表現として示される。

40

【0080】

トレーニング段階の間、照会 N_q が入力点群からサンプリングされ、それらがLODの各々についてどのボクセルにあるかが判断され、その特徴が選択されたボクセルに従って補間される。予測されるサンプルの二乗誤差の和またはクロスエントロピーは、それぞれSDFおよび占有率についてアクティブなLODから計算される。加えて、二重逆伝播を用いて法線を求める。次いで、計算された法線とアンカー法線との間でL2ノルムを正則化項として計算する。2つの項を加算(加重和)して最終損失を求める。

50

【 0 0 8 1 】

評価中、均一に分布した入力サンプルを、分解能 $Q = 512^3$ の単位立方体から得る。結果は、最後の LOD について示され、より細かい LOD に対応する。入力照会は、既存の八分木セルとマッチしない場合、破棄される。最後に、マーチングキューブを用いた出力からのメッシュが得られる。

【 0 0 8 2 】

実施例

【 0 0 8 3 】

以下は、本明細書で記載されるかまたは別様に想定される方法およびシステムを使用する、例示的実現形態および分析を説明する。これらは例としてのみ提供され、本発明の範囲を限定しないことが理解されるであろう。

【 0 0 8 4 】

ある実施形態によれば、3D再構築品質が、各オブジェクトについて、Chamfer Distance (CD)、Normal Consistency (NC)、およびIntersection over Union (IoU)を使用して評価された。CDは、照会点およびそのグラウンドトゥルースマッチの逆数最小距離として計算された。CDは5回計算され、その平均が示された。NCは、照会法線(CD計算中に得られる照会点に対応する)とその対応するグラウンドトゥルース法線との間のコサイン類似度から計算される、対応する法線である。NCは、両方の法線間のコサイン類似度の残差として報告される。IoUは、2つの格子セット間の重複を定量化する。メッシュは、IoUについて、分解能 $Q = 128^3$ の立方体を使用してレンダリングされた。

【 0 0 8 5 】

本方法を、3つのデータセット、すなわちABC、Thing10k、およびShapeNet上で評価した。合計32個のメッシュを、各々、Thing10kおよびABCからサンプリングし、150個のメッシュをShapeNetからサンプリングした。ShapeNetメッシュは水密にされ、作業はPyTorchで実施した。

【 0 0 8 6 】

デコーダアーキテクチャは、ReLUを有する次元128の1つの隠れ層を有する。各ボクセル特徴は、 $F = 32$ 次元特徴ベクトルとして表される。照会点および法線についての位置エンコードは、 $L_p = L_n = 6$ つの周波数で表わされる。スパース3D畳み込みは、局所的特徴集約について、カーネルサイズ K_k 、 V_l 、 $k = 5$ で、考慮される。円筒半径 R は、経験的に以下のように設定された：

【数5】

$$R = (h_1 + h_2) / \sqrt{2} \quad (\text{式10})$$

【 0 0 8 7 】

Adamオプティマイザを使用して、モデルを、0.001の学習率および $\eta = 0.1$ で、100エポックまでトレーニングした。 5×10^6 点の初期サンプルサイズが、512のバッチサイズで、考慮される。再サンプリングは、各すべてのエポック後に行われる。点は、表面およびその近傍付近から等比率でサンプリングされる。各ボクセルが表面サンプリングの前に少なくとも32個のサンプルを有することも保証される。LODs $\mathcal{L} = \{3, \dots, 7\}$ をすべてのデータセットについて考慮した。

【 0 0 8 8 】

ベースラインについて、この手法を、供給された設定でトレーニングされた現状技術の手法のBACON、SIREN、およびフーリエ特徴関数(FF)と比較した。構造格子直接手法もまた、有向格子に対する公平な比較のために、本方法に対して評価され、それはパイプラインにおいて、より小さい変更を必要とした。表面再構築設定は、上述のように、すべての方法について同じであった。

【 0 0 8 9 】

10

20

30

40

50

実験のために、アブレーションを下記のように利用した。構造格子対有向格子も比較され、本方法を、異なるエンコード戦略を使用する方法に対して評価した。これらは以下で論ずる。

【0090】

アブレーション

【0091】

各構成要素の関連性を分析するために、異なるパイプラインブロックに対して変更を徐々に加えた。ABCおよびThingy10kデータセットからの10個のメッシュを、トレーニングおよび試験のためにランダムにサンプリングする。図8は、エンコードになされた変更に基づいて、図7にリストされた異なるケースを示す。トリリニア補間を伴う有向格子の使用は、多くの穴をもたすことが注目された。セルはアンカー法線ごとに回転させるので、回転不変円筒表現を補間に使用した。依然として、より粗いが、これは、より適応した表現(CDの有意な改善)をもたす。

10

【0092】

したがって、図7を参照して、アブレーション研究の結果が示されている。この表は、最終エンコードに至る異なる段階を示す。トリリニア補間を用いた有向格子から開始して、提案される円筒補間、そして最後に、3DCNNを用いた局所的特徴集約である。CDに 10^{-5} を乗算し、NCに 10^{-4} を乗算する。

【0093】

図8を参照すると、レンダリングにおけるアブレーションの影響が示されており、図7の数値結果を反映している。パネル(a)は、トリリニア補間を用いた有向エンコードを表し、(b)は、円筒補間を加え、(c)および(d)は、特徴集約のために、 $3 \times 3 \times 3$ および $5 \times 5 \times 5$ の3DCNNカーネルをそれぞれ使用し、(e)は、(d)に法線正則化を加える。(f)はグラウンドトゥルースを示す。

20

【0094】

3DCNNの追加(図8(c)および(d))によるメッシュ平滑性の有意な改善(NCに反映される)が見られ、局所的特徴集約ステップに効果的に寄与する。実験は、 $5 \times 5 \times 5$ のカーネルが、より良好な性能を達成し、その後の実験に好ましいことを示す。提案される法線正則化は、平滑性を強化するが、精度を犠牲にする。

【0095】

30

構造格子対有向格子

【0096】

図9を参照して、構造格子対有向格子の比較が、SDFおよび占有率デコードについての結果とともに示されている。CDに 10^{-5} を乗算し、NCに 10^{-4} を乗算する。表は、本明細書に開示される方法の性能をSDFおよび占有率デコード上で構造格子と比較する。SDFおよび占有率デコードは、上述のようにトレーニングされる。法線正則化は、両方の場合について同じままである。本明細書に開示される方法は、すべてのフロント上でSDFデコードでの構造格子よりも性能が優れており、構造化された表面上で、より滑らかな結果をもたす。占有率フレームワークの性能不足にもかかわらず、メッシュ上に観察される穴およびくぼみは、より少ない(後者はIoUに有意な影響を有する)。これらの結果は、異なるデコード出力表現への本方法の適応性を示す。

40

【0097】

大規模シーン表現への本方法の可能な拡張を開くために、本方法のレンダリングが、図10に示されるように、Matterport3Dからのシーン上に示される。このシーンは、 4×4 クロップ(地面を含む)に分割され、あるモデルが、各クロップについて占有率デコードを用いてトレーニングされる。推論中、メッシュクロップは、図10に示されるように、マーチングキューブを使用してレンダリングされ、最終的に融合されてシーンを生成する。薄い表面のため、構造格子は、より粗くぼんやりした3D表現をもたす。本明細書に開示される方法は、薄い表面にうまく適応し、シーンを、より少ない粗さおよびより鮮明な品質で、レンダリングする。

50

【 0 0 9 8 】

有向格子の結果として、提案されるエンコーダは、より少ないトレーニングステップで、より効果的に、平面をレンダリングすることが観察される。特に、より構造的な規則的なオブジェクトでは、構造格子は苛性様効果（表面ノイズ）を生じる。有向格子の表面再構築において、最初のエポックからすぐに、ノイズ低減が観察される。

【 0 0 9 9 】

ベースライン

【 0 1 0 0 】

図 1 1 の表は、他のタイプの表現、すなわち、SIRENおよびBACONのような格子ベースのネットワークならびにBACONのようなマルチスケール表現に対する本方法（特徴ベースの幾何学的マルチスケール表現）の実験結果を詳述する。それはまた、非水密オブジェクトを考慮するべく、本方法を、符号なし出力表現を使用するNDFと比較する。NDFは、メッシュを得るためにball pivotingアルゴリズムを使用するが、多くのハンドチューニングを必要とする。それは計算的に高価であり、非常に不連続なメッシュ表現および不良な結果をもたらした。代わりに、このステップは、上記と同じ設定を使用して、点群をボクセル化し、マーチングキューブを使用して表面を得ることによって、置き換えられる。

【 0 1 0 1 】

したがって、格子ベースの方法は、すべてのフロントにおいて有意な改善を伴ってベースラインよりも性能が優れていることが示される。ABCのような、平面オブジェクトから構成される単純なデータセットでは、エンコーダは、有向格子の整列により、より滑らかな平面を再構築する。ホリスティックな詳細をレンダリングするものの、大部分のベースラインは、しばしば、過度に平滑化された表面を有する。本方法については、メッシュ内のより少ない穴および無視できるほどのスプラッティング（サンプリング領域の周囲の多数の小さいメッシュトレース）の結果として、より高いIoUも観察される。全体として、有向格子は、すべてのデータセットにわたって、より高い忠実度で、ロバストな3D表現を生成する。

【 0 1 0 2 】

メッシュをレンダリングするために必要なパラメータの数も提供される。多分解能格子表現の利点は、デコーダのサイズを、1つの隠れ層を有するMLPだけに低減できることである。この結果、この手法は、他の方法よりも速くメッシュを得る。

【 0 1 0 3 】

Thingi10kおよびShapeNetデータセットからの例を分析した。BACONおよびFFは、オブジェクトを妥当な精度でモデル化することができるが、多くのスプラッタを記録し、望ましくないノイズのある表面およびアーチファクトを生じさせる。SIRENおよびBACONは、過度に平滑化された表面を生成し、メッシュ上で複雑な詳細を失う。NDFは、多くの穴を生じさせるが、スプラッタなしにコンパクトなメッシュをどうにかして得る。BACON、SIREN、およびFFは、ShapeNetデータセット上で崩壊する。ShapeNetの水密版および非水密版の両方を試みたが、失敗したベースラインについて同様の結果が得られた。

【 0 1 0 4 】

小さな穴は、有向格子および構造格子の両方に影響を及ぼす非水密平面から生じ得る。しかしながら、有向格子は、構造格子の対応物よりも適切に穴を充填する。より一般的な多分解能格子表現の問題は、薄い表面をモデル化する困難さである。この制限にもかかわらず、本方法は、構造格子から実質的に改善する。

【 0 1 0 5 】

したがって、実施例および実験は、3D表現のための3D格子ベースのエンコーダのためのこの新規な手法が、現状技術の結果を、デコーダ表現変化に対してよりロバストかつ正確でありながら、もたらすことを実証する。エンコーダは、格子をオブジェクト表面法線と整列させ、整列によって引き起こされた問題を軽減する新たに開発された円筒補間技

10

20

30

40

50

法および局所的集約方式からセル特徴を集約することによって、オブジェクトにおける本来の構造的規則性を考慮する。

【0106】

加えて、実施例および実験は、本明細書に開示されるかまたは別様に想定されるシステムおよび方法が、改善されたコンピュータシステムをもたらすことを実証する。3Dオブジェクト表現生成システムは、先行技術のシステムよりも速く、より良好な3Dオブジェクト表現を生成することができる。したがって、トレーニングされたニューラルネットワークを含む本明細書に記載の3Dオブジェクト表現生成システムは、先行技術の3Dオブジェクト表現生成システムに対する改良である。

【0107】

ある実施形態によると、本明細書で開示されるかまたは別様に想定されるシステムおよび方法は、ニューラルネットワークのトレーニング中、およびトレーニングされたニューラルネットワークを使用する実行中に、何千または何百万ものデータ点を処理するよう構成される。例えば、トレーニングデータのコーパスから、機能的かつ熟練した、トレーニングされたニューラルネットワークを生成することは、入力データおよび生成された特徴からの何百万個ものデータ点の処理を要求する。これは、新規のトレーニングされたニューラルネットワークを生成するために、何百万個または何十億個もの計算を必要とし得る。その結果、各トレーニングされたニューラルネットワークは、入力データおよびアルゴリズムのパラメータに基づいて新規かつ相異なり、したがって、本システムの機能を改善する。機能的で熟練した、トレーニングされたニューラルネットワークを生成することは、人間の脳が生涯において達成不可能な量の計算および分析を有するプロセスを含む。

【0108】

本明細書で定義および使用されるすべての定義は、辞書の定義、参照により組み込まれる文書における定義、および/または定義された用語の通常の意味を支配すると理解されるべきである。

【0109】

本明細書および特許請求の範囲において使用される不定冠詞「a」および「an」は、明確に反対の指示がない限り、「少なくとも1つ」を意味すると理解されるべきである。

【0110】

本明細書および特許請求の範囲で使用される「および/または」という文言は、そのように結合された要素、すなわち、ある場合では接続的に存在し、他の場合では離接的に存在する要素、の「いずれかまたは両方」を意味すると理解されるべきである。「および/または」で列挙された複数の要素は、同じ態様で解釈されるべきであり、すなわち、そのように結合された要素の「1つまたは複数」である。「および/または」節によって具体的に識別される要素以外の他の要素が、具体的に識別されるそれらの要素に関連するかまたは関連しないかにかかわらず、選択肢として存在してもよい。

【0111】

本明細書および特許請求の範囲で使用されるように、「または」は、上記で定義された「および/または」と同じ意味を有すると理解されるべきである。例えば、リスト中の項目を分離する場合、「または」または「および/または」は、包括的である、すなわち、いくつかの要素または要素のリストのうちの少なくとも1つであるが、1つより多く、および選択肢として、追加のリストされていない項目も含む包含であるとして解釈されるべきである。「~のうちの1つのみ」もしくは「~のうちの厳密に1つ」、または特許請求の範囲で使用される場合、「~からなる」などの、反対に明確に示される文言のみが、いくつかの要素または要素のリストのうちの厳密に1つの要素の包含を指すことになる。一般に、本明細書で使用する「または」という文言は、「いずれか」、「~の1つ」、「~の1つのみ」、または「~の厳密に1つ」といった、排他性の文言が先行するときに排他的な代替物（すなわち、「一方または他方であるが、両方ではない」）を示すものとしてのみ解釈されるべきである。

【0112】

本明細書および特許請求の範囲において使用される場合、「少なくとも１つ」という文言は、１つまたは複数の要素のリストに関して、要素のリスト中の要素のうちの任意の１つまたは複数から選択される少なくとも１つの要素を意味すると理解されるべきであるが、必ずしも、要素のリスト内に具体的に列挙された各すべての要素の少なくとも１つを含むとは限らず、要素のリスト内の要素の任意の組合せを除外しない。この定義はまた、「少なくとも１つ」という文言が指す、要素のリスト内で具体的に識別される要素以外の要素が、具体的に識別される要素に関係するかまたは関係しないかにかかわらず、選択肢として存在し得ることを可能にする。

【０１１３】

反対のことが明確に示されない限り、複数のステップまたは行為を含む、ここで特許請求される任意の方法において、本方法のステップまたは行為の順序は、本方法のステップまたは行為が記載される順序に必ずしも限定されないことも理解されたい。

【０１１４】

特許請求の範囲ならびに上記明細書において、「備える」、「含む」、「持つ」、「有する」、「含有する」、「関与する」、「保持する」、「～から構成される」といったすべての移行句は、オープンエンドである、すなわち、含むが限定されないことを意味すると理解されるべきである。移行句「～からなる」および「本質的に～からなる」のみが、それぞれ、閉じた移行句または半ば閉じた移行句であるものとする。

【０１１５】

いくつかの発明の実施形態が本明細書に記載および示されたが、当業者は、本明細書に記載される機能を実行する、ならびに／もしくは結果および／もしくは利点の１つ以上を得るための種々の他の手段ならびに／または構造を容易に想定し、そのような変形例および／または修正例の各々は、本明細書に記載される発明の実施形態の範囲内であると見なされる。より一般的には、当業者は、本明細書に記載されるすべてのパラメータ、寸法、材料、および構成は例示的であることを意味し、実際のパラメータ、寸法、材料、および／または構成は、本発明の教示が使用される特定の用途に依存することになることを容易に理解するであろう。当業者は、本明細書に記載される特定の発明の実施形態に対する多くの均等物を認識するか、または日常的な実験のみを用いて確認することができるであろう。したがって、前述の実施形態は、例としてのみ提示されており、特許請求の範囲およびその均等物の範囲内で、発明の実施形態は、具体的に記載および特許請求されるものとは別の態様で実施され得ることを理解されたい。本開示の発明の実施形態は、本明細書に記載される各個々の特徴、システム、物品、材料、キット、および／または方法を対象とする。加えて、２つ以上のそのような特徴、システム、物品、材料、キット、および／または方法の任意の組み合わせは、そのような特徴、システム、物品、材料、キット、および／または方法が相互に矛盾しない場合、本開示の発明の範囲内に含まれる。

10

20

30

40

50

【図面】
【図 1】

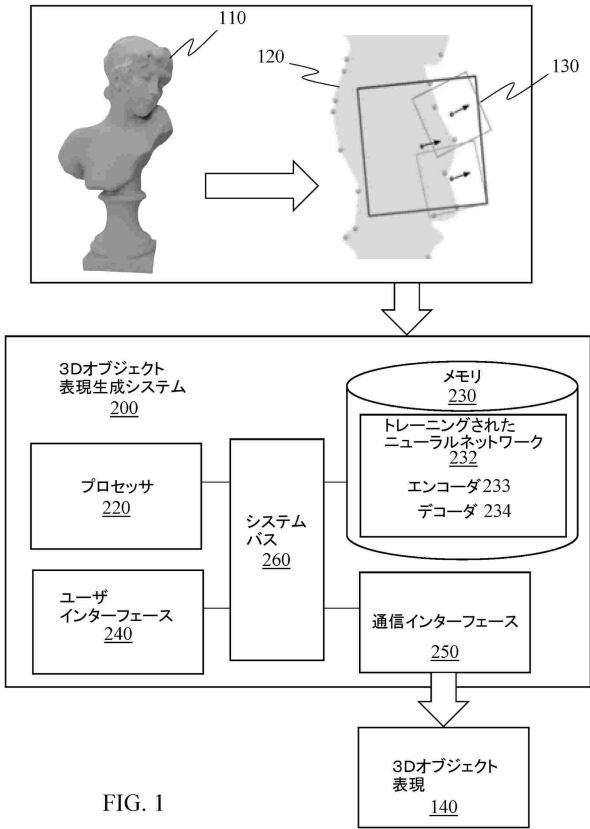


FIG. 1

【図 2 A】

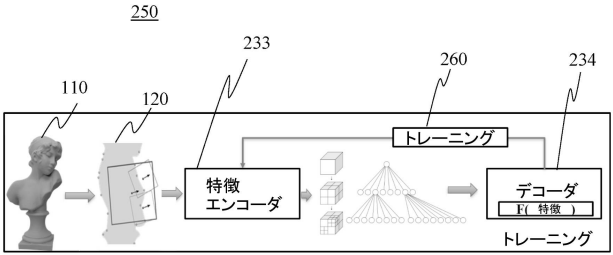


FIG. 2A

【図 2 B】

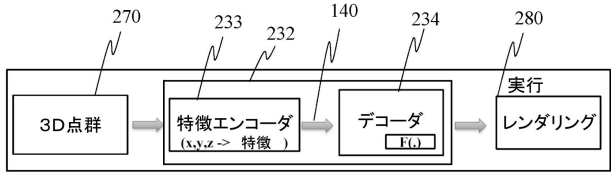


FIG. 2B

【図 3 A】

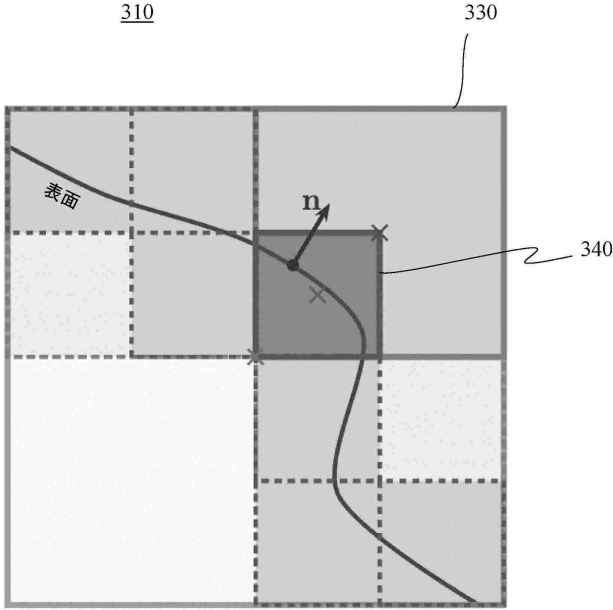


FIG. 3A

10

20

30

40

50

【 図 3 B 】

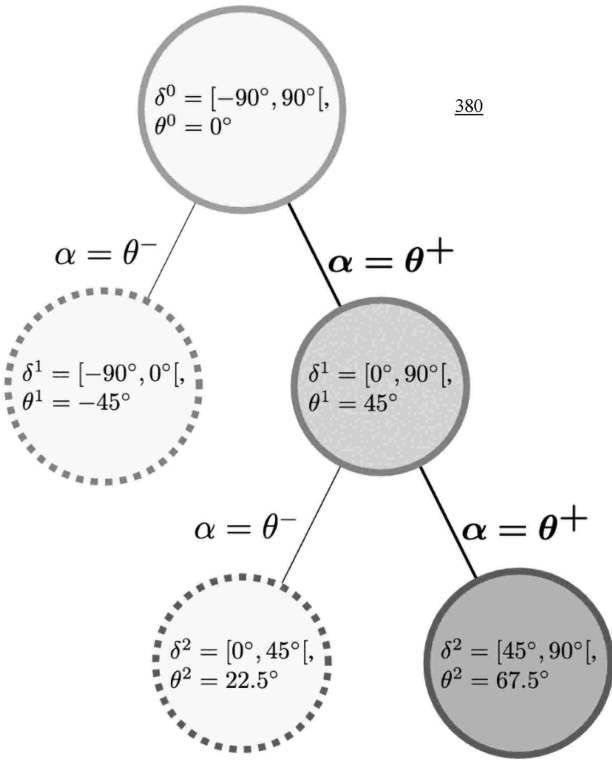


FIG. 3B

【 図 3 C 】

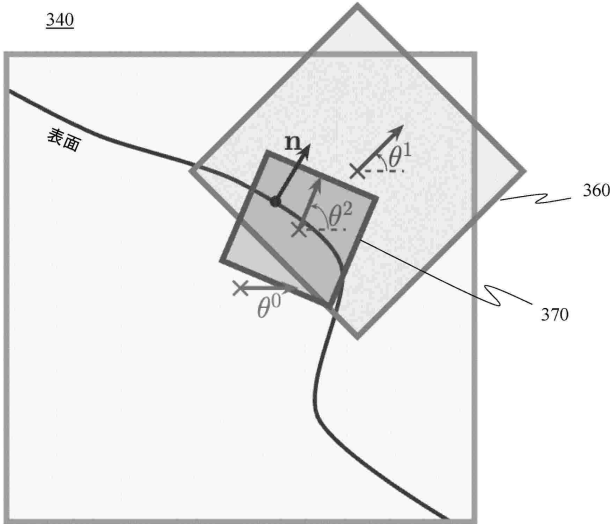
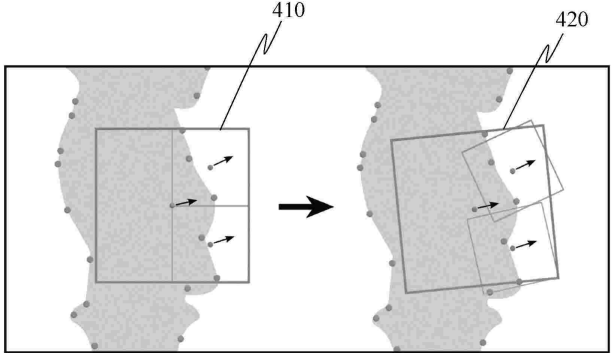


FIG. 3C

【 図 4 A 】



構造格子から有向格子へ

FIG. 4A

【 図 4 B 】

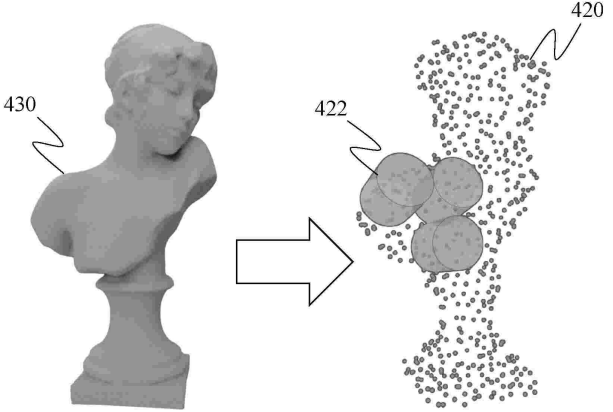


FIG. 4B

10

20

30

40

50

【 図 5 】

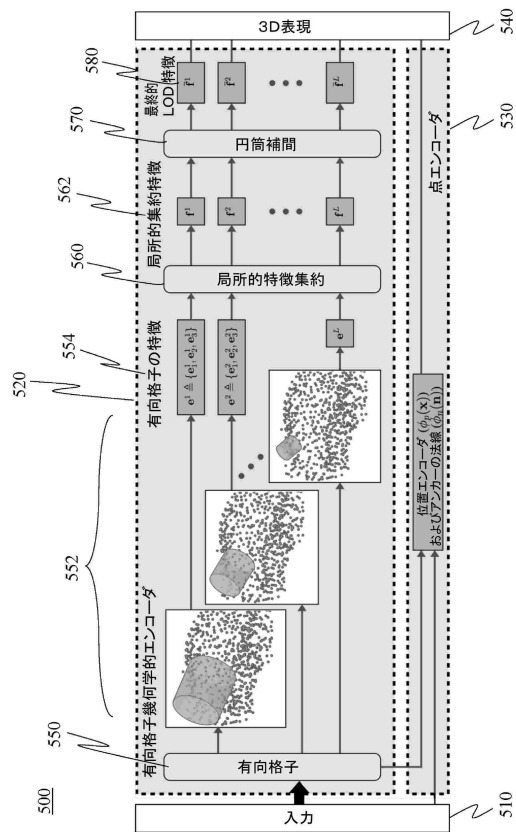


FIG. 5

【 図 6 】

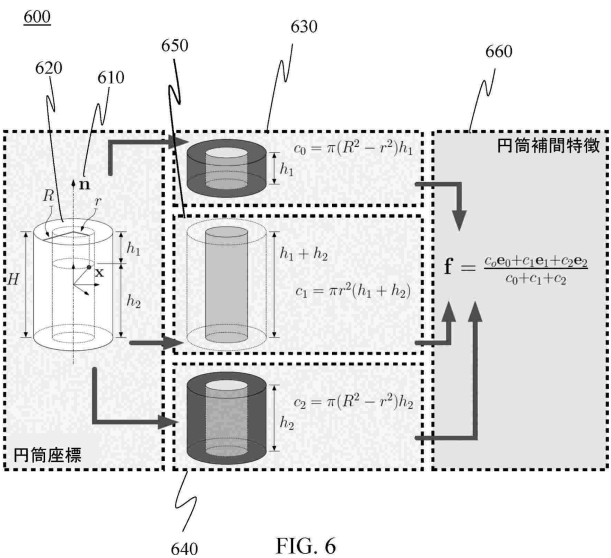


FIG. 6

【 図 7 】

ネットワークオプション					結果		
トリリニア補間	円筒補間	3DCNN 3x3x3	3DCNN 5x5x5	法線正則化	CD↓	NC↓	IoU↑
✓					1.278	4.410	0.881
	✓				0.570	4.348	0.990
	✓	✓			0.434	4.094	0.998
	✓		✓		0.431	4.093	0.998
	✓		✓	✓	0.443	4.058	0.998

FIG. 7

【 図 8 】

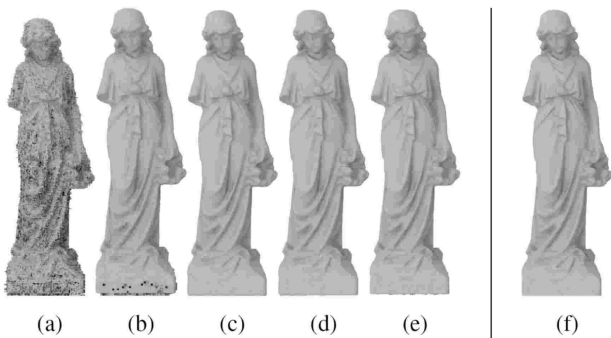


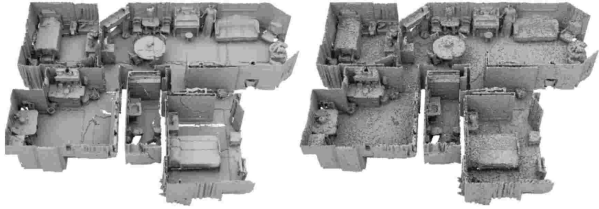
FIG. 8

【 図 9 】

ネットワークオプション				結果		
構造格子	有向格子	SDF デコーダ	占有率 デコーダ	CD↓	NC↓	IoU↑
✓		✓		0.445	4.257	0.997
✓			✓	0.687	4.117	0.989
	✓		✓	0.788	4.140	0.995
	✓	✓		0.443	4.058	0.998

FIG. 9

【 図 1 0 】



(a) 有向格子 (b) 構造格子

FIG. 10

10

【 図 1 1 】

	SIREN	BACON	FF	NDF	本発明
ABCデータセットについての結果					
CD↓	5.837	2.229	18.98	5.020	0.603
NC↓	5.150	4.658	5.170	4.732	3.987
IoU ↑	0.879	0.987	0.916	0.950	0.998
Thing10kデータセットについての結果					
CD↓	62.85	60.72	67.93	4.421	0.608
NC↓	49.11	5.052	58.31	4.636	4.413
IoU ↑	0.716	0.928	0.877	0.920	0.995
ShapeNETデータセットについての結果					
CD↓	—	—	—	4.125	0.392
NC↓	—	—	—	6.186	5.425
IoU ↑	—	—	—	0.984	0.999
推論パラメータの数					
#パラメータ	199K	537K	527K	4.62M	388K

FIG. 11

20

【 外国語明細書 】

2025071766000022.pdf

30

40

50

フロントページの続き

・リサーチ・ラボラトリーズ・インコーポレイテッド

(72)発明者 ガウアー, アリハント

アメリカ合衆国、マサチューセッツ州、ケンブリッジ、ブロードウェイ 201、ケアオブ・ミッ
シシ・エレクトリック・リサーチ・ラボラトリーズ・インコーポレイテッド

Fターム(参考) 5C159 LB11 PP03 PP13 TD16 TD17 UA02 UA05