

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2022年12月1日(01.12.2022)



(10) 国際公開番号

WO 2022/249232 A1

(51) 国際特許分類:
G06T 7/00 (2017.01)

(21) 国際出願番号: PCT/JP2021/019579

(22) 国際出願日: 2021年5月24日(24.05.2021)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(71) 出願人: 日本電信電話株式会社 (NIPPON TELEGRAPH AND TELEPHONE CORPORATION) [JP/JP]; 〒1008116 東京都千代田区大手町一丁目5番1号 Tokyo (JP).

(72) 発明者: 金子 卓弘 (KANEKO Takuhiro); 〒1808585 東京都武蔵野市緑町3丁目9-1 NTT知的財産センタ内 Tokyo (JP).

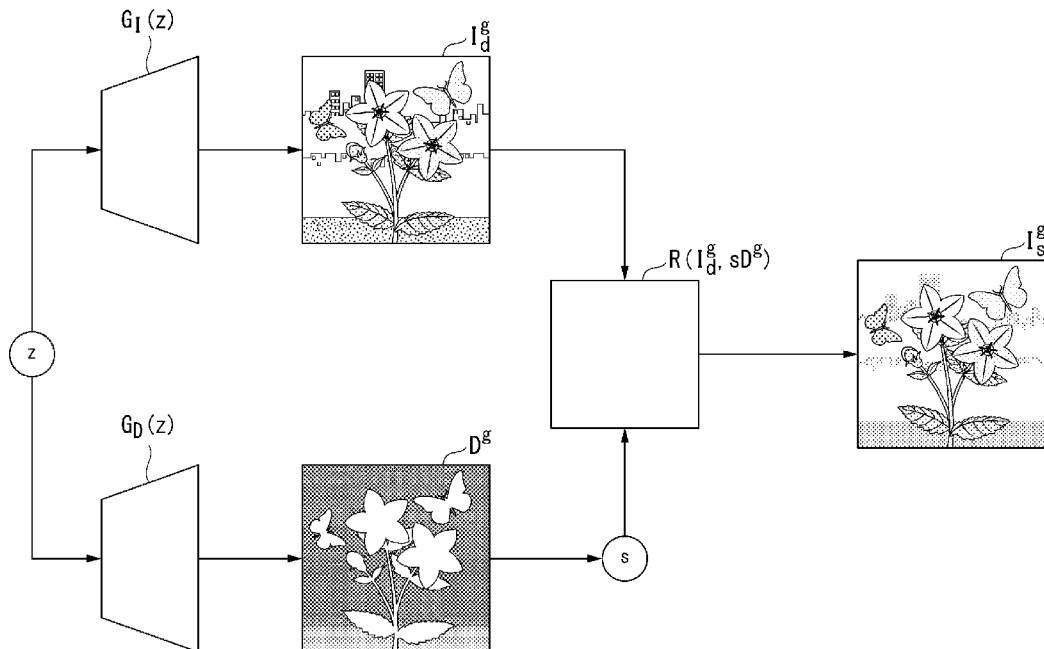
(74) 代理人: 特許業務法人 志賀国際特許事務所 (SHIGA INTERNATIONAL PATENT OFFICE); 〒1006620 東京都千代田区丸の内一丁目9番2号 Tokyo (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, IT, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY,

(54) Title: LEARNING DEVICE, LEARNED MODEL GENERATION METHOD, DATA GENERATION DEVICE, DATA GENERATION METHOD, AND PROGRAM

(54) 発明の名称: 学習装置、学習済みモデルの生成方法、データ生成装置、データ生成方法およびプログラム

[図4]



(57) Abstract: In the present invention, an image generation unit generates first image data simulating a captured image with a first depth of field by inputting latent variables to an image generation model, which is a machine learning model. A three-dimensional data generation unit generates three-dimensional data corresponding to the first image data by inputting the latent variables into a three-dimensional generation model, which is a machine learning model. A conversion unit



MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ,
NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT,
QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,
ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG,
US, UZ, VC, VN, WS, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類：

一 国際調査報告(条約第21条(3))

generates second image data simulating a captured image with a second depth of field on the basis of the first image data and the three-dimensional data. A calculation unit calculates a learning reference value indicating the degree of closeness between the distribution of the first image data and the second image data and the empirical distribution of the actual captured images. An updating unit updates the parameters of the image generation model and the three-dimensional generation model on the basis of the learning reference value.

(57) 要約：画像生成部は、潜在変数を機械学習モデルである画像生成モデルに入力することで、第1被写界深度に係る撮像画像を模擬した第1画像データを生成する。三次元データ生成部は、潜在変数を機械学習モデルである三次元生成モデルに入力することで、第1画像データに対応する三次元データを生成する。変換部は、第1画像データと三次元データとに基づいて、第2被写界深度に係る撮像画像を模擬した第2画像データを生成する。算出部は、第1画像データおよび第2画像データの分布と、実際の撮像画像の経験分布との近さの度合いを示す学習基準値を算出する。更新部は、学習基準値に基づいて画像生成モデルおよび三次元生成モデルのパラメータを更新する。

明 細 書

発明の名称：

学習装置、学習済みモデルの生成方法、データ生成装置、データ生成方法
およびプログラム

技術分野

[0001] 本発明は、学習装置、学習済みモデルの生成方法、データ生成装置、データ生成方法およびプログラムに関する。

背景技術

[0002] 二次元の画像データから、その画像データに対応する三次元データ（深度情報など）を推定することは、長く関心を集めている問題の一つである。この問題を解くための方法として、二次元画像と三次元データのペアデータを教師データとして、二次元画像から三次元データを得る変換器を学習する方法が知られている。しかし、ペアデータを集めるためには専用の機器が必要であり、さらにデータ取得後もデータ間のアライメントを正確に取ることが必要であるため、データ収集コストが高いという問題がある。

[0003] 非特許文献1には、生成モデルによって視点の異なる画像データを、生成される画像データの経験分布が実画像の経験分布と一致するように学習する技術が開示されている。

先行技術文献

非特許文献

[0004] 非特許文献1：A. Noguchi, T. Harada, “RGBD-GAN: Unsupervised 3D Representation Learning From Natural Image Datasets via RGBD Image Synthesis,” in Proc. ICLR 2020.

発明の概要

発明が解決しようとする課題

[0005] しかしながら、非特許文献1に記載の技術を実現するためには、学習に用

いるデータセットに、複数の異なる視点から撮像された画像データが十分な量含まれている必要がある。そのため、異なる視点に係る画像データを収集することが困難な場合に、適切な学習を行うことができない可能性がある。

本発明の目的は、三次元データに関する補足データなしに、二次元画像と、当該二次元画像に対応する三次元データとを生成することができる学習装置、学習済みモデルの生成方法、データ生成装置、データ生成方法およびプログラムを提供することにある。

課題を解決するための手段

[0006] 本発明の一態様は、潜在変数を機械学習モデルである画像生成モデルに入力することで、第1被写界深度に係る撮像画像を模擬した第1画像データを生成する画像生成部と、前記潜在変数を機械学習モデルである三次元生成モデルに入力することで、前記第1画像データに対応する三次元データを生成する三次元データ生成部と、前記第1画像データと前記三次元データとに基づいて、第2被写界深度に係る撮像画像を模擬した第2画像データを生成する変換部と、前記第1画像データおよび前記第2画像データの分布と、実際の撮像画像の経験分布との近さの度合いを示す学習基準値を算出する算出部と、前記学習基準値に基づいて前記画像生成モデルおよび前記三次元生成モデルのパラメータを更新する更新部とを備える学習装置である。

[0007] 本発明の一態様は、潜在変数を機械学習モデルである画像生成モデルに入力することで、第1被写界深度に係る撮像画像を模擬した第1画像データを生成するステップと、前記潜在変数を機械学習モデルである三次元生成モデルに入力することで、前記第1画像データに対応する三次元データを生成するステップと、前記第1画像データと前記三次元データとに基づいて、第2被写界深度に係る撮像画像を模擬した第2画像データを生成するステップと、前記第1画像データおよび前記第2画像データの分布と、実際の撮像画像の経験分布との近さの度合いを示す学習基準値を算出するステップと、前記学習基準値に基づいて前記画像生成モデルおよび前記三次元生成モデルのパラメータを更新するステップと、学習済みの前記画像生成モデルおよび前記三

次元生成モデルを出力するステップとを有する学習済みモデルの生成方法である。

[0008] 本発明の一態様は、上記態様に係る学習済みモデルの生成方法によって出力された画像生成モデルに潜在変数を入力することで、撮像画像を模擬した画像データを生成する画像生成部と、上記態様に係る学習済みモデルの生成方法によって出力された三次元生成モデルに前記潜在変数を入力することで、三次元データを生成する三次元データ生成部と、を備えるデータ生成装置である。

[0009] 本発明の一態様は、上記態様に係る学習済みモデルの生成方法によって出力された画像生成モデルに潜在変数を入力することで、撮像画像を模擬した画像データを生成するステップと、上記態様に係る学習済みモデルの生成方法によって出力された三次元生成モデルに前記潜在変数を入力することで、三次元データを生成するステップと、を備えるデータ生成方法である。

[0010] 本発明の一態様は、コンピュータを、上記態様に係る学習装置として機能させるためのプログラムである。

発明の効果

[0011] 上記少なくとも1つの態様によれば、三次元データに関する補足データなしに、二次元画像と、当該二次元画像に対応する三次元データとを生成することができる。

図面の簡単な説明

[0012] [図1]第1の実施形態に係るデータ生成システムの構成を示す図である。

[図2]第1の実施形態に係るモデル学習装置の構成を示す概略ブロック図である。

[図3]第1の実施形態に係るモデル学習装置の動作を示すフローチャートである。

[図4]第1の実施形態に係る学習処理におけるデータの変遷を示す図である。

[図5]第1の実施形態に係るデータ生成装置の構成を示す概略ブロック図である。

[図6]少なくとも1つの実施形態に係るコンピュータの構成を示す概略ブロック図である。

発明を実施するための形態

[0013] 〈第1の実施形態〉

《データ生成システム1の構成》

図1は、第1の実施形態に係るデータ生成システム1の構成を示す図である。データ生成システム1は、同一の被写体に係る被写界深度の深い撮像画像を模擬した画像データ（ディープデプスオブフィールドイメージ）、被写界深度の浅い撮像画像を模擬した画像データ（シャローデプスオブフィールドイメージ）、および被写体の深度を示す深度データを生成する。ディープデプスオブフィールドイメージは、ピントが合っているように見える範囲が広い画像データである。シャローデプスオブフィールドイメージは、ピントが合っているように見える範囲が狭い画像データである。

[0014] データ生成システム1は、データ生成装置11とモデル学習装置13とを備える。

データ生成装置11は、機械学習モデルである画像生成モデルおよび三次元生成モデルを用いて、ディープデプスオブフィールドイメージ、シャローデプスオブフィールドイメージおよび深度データの組を生成する。

モデル学習装置13は、実際の撮像画像を学習用データとして用いて画像生成モデルおよび三次元生成モデルの学習を行う。なお、学習用データに係る撮像画像は、三次元データに関する補足データを有する必要はない。つまり、学習用データは深度データなどの三次元データや、視点を異ならせた撮像画像のペアなどを含まなくてもよい。

[0015] 《モデル学習装置13の構成》

図2は、第1の実施形態に係るモデル学習装置13の構成を示す概略ブロック図である。第1の実施形態に係るモデル学習装置13は、学習用データ記憶部131、モデル記憶部132、潜在変数生成部133、画像生成部134、三次元データ生成部135、変換部136、識別部137、算出部1

38、更新部139を備える。

[0016] 学習用データ記憶部131は、複数の画像データを記憶する。各画像データは、撮像装置によって撮像された画像である。学習用データ記憶部131は、様々な被写界深度に係る画像データを記憶する。

[0017] モデル記憶部132は、画像生成モデル G_I 、深度生成モデル G_D 、識別モデル C 、および被写界深度効果レンダラ R を記憶する。画像生成モデル G_I 、深度生成モデル G_D および識別モデル C は、いずれもニューラルネットワーク（例えば、畳み込みニューラルネットワーク、全結合ニューラルネットワーク、再帰ニューラルネットワークなど）によって構成される。被写界深度効果レンダラ R は、撮像装置の光学系を模擬する数理モデルによって構成される。

[0018] 画像生成モデル G_I は、潜在変数 z を入力とし、被写界深度の深い撮像画像を模擬した画像データであるディープデプスオブフィールドイメージ I_d を出力とする。潜在変数は、画像データを生成するシードとなる任意の数値である。

[0019] 深度生成モデル G_D は、潜在変数 z を入力とし、ディープデプスオブフィールドイメージ I_d の被写体の深度を表す深度データ D_d を出力とする。深度データ D_d の要素数はディープデプスオブフィールドイメージ I_d の要素数と等しい。画像生成モデル G_I と深度生成モデル G_D とは、一部の層（例えば、入力層および中間層）を共通とするものであってよい。

[0020] 被写界深度効果レンダラ R は、ディープデプスオブフィールドイメージ I_d と深度データ D_d の組を入力とし、ディープデプスオブフィールドイメージ I_d と同じ被写体に係る被写界深度の浅い撮像画像を模擬した画像データであるシャローデプスオブフィールドイメージ I_s を出力とする。被写界深度効果レンダラ R の詳細については後述する。

[0021] 識別モデル C は、画像データを入力とし、入力された画像データが実際の撮像画像である確率または実際の撮像画像である度合を示す評価値を出力とする。例えば、識別モデル C は、実際の撮像画像である確率を出力する場合

、入力された画像データがディープデプスオブフィールドイメージ I_d^g またはシャローデプスオブフィールドイメージ I_s^g である確率が高いほど0に近い値を出力し、実際の撮像データである確率が高いほど1に近い値を出力する。

[0022] 画像生成モデル G_I 、深度生成モデル G_D 、識別モデル C および被写界深度効果レンダラ R は、GANs (Generative Adversarial Networks) を構成する。画像生成モデル G_I 、深度生成モデル G_D および被写界深度効果レンダラ R の組み合わせは、Generatorである。識別モデル C は、Discriminatorである。

[0023] 潜在変数生成部133は、乱数に基づいて潜在変数 z を生成する。例えば、潜在変数生成部133は、ガウシアン分布や一様分布などの任意の分布において、ランダムに潜在変数 z を抽出する。なお、潜在変数生成部133は、学習用データ記憶部131が記憶する画像データから潜在変数 z を抽出してもよい。

[0024] 画像生成部134は、潜在変数生成部133が生成した潜在変数 z をモデル記憶部132が記憶する画像生成モデル G_I に入力することで、ディープデプスオブフィールドイメージ I_d^g を生成する。つまり、画像生成部134は、以下の式(1)によりディープデプスオブフィールドイメージ I_d^g を算出する。

[0025] [数1]

$$I_d^g = G_I(z) \quad \dots (1)$$

[0026] 三次元データ生成部135は、潜在変数生成部133が生成した潜在変数 z をモデル記憶部132が記憶する深度生成モデル G_D に入力することで、三次元データである深度データ D^g を生成する。つまり、三次元データ生成部135は、以下の式(2)により深度データ D^g を算出する。

[0027] [数2]

$$D^g = G_D(z) \quad \dots (2)$$

[0028] 変換部136は、画像生成部134が生成したディープデプスオブフィールドイメー ジ I_d^g と三次元データ生成部135が生成した深度データ D^g とをモデル記憶部132が記憶する被写界深度効果レンダラRに入力することで、シャローデプスオブフィールドイメー ジ I_s を生成する。つまり、変換部136は、ディープデプスオブフィールドイメー ジ I_d^g をシャローデプスオブフィールドイメー ジ I_s に変換する。つまり、変換部136は、以下の式(3)によりシャローデプスオブフィールドイメー ジ I_s を生成する。式(3)においてsは、ディープデプスオブフィールドイメー ジ I_d^g とシャローデプスオブフィールドイメー ジ I_s の混合度、すなわち被写界深度の深さの度合いを表す。なお、混合度sが0である場合、計算結果はディープデプスオブフィールドイメー ジ I_d^g と等しくなる。

[0029] [数3]

$$I_s^g = R(I_d^g, sD^g) \quad \dots (3)$$

[0030] ここで、被写界深度効果レンダラRについて説明する。第1の実施形態に係る被写界深度効果レンダラRは、仮想の光学系における光線の経路を計算し、仮想の光学系における開口面積を広げることで被写界深度の変換を行う。具体的には、被写界深度効果レンダラRは、深度データ D^g を変形関数Tを用いて変形することで、式(4)に示すように画像面上の位置座標xおよび開口面上の角度座標uと被写体の深度の関係を示す深度マップ $M^g(x, u)$ を演算する。次に被写界深度効果レンダラRは、深度マップ $M^g(x, u)$ に基づいて、式(5)に示すようにディープデプスオブフィールドイメー ジ I_d^g から、開口面上の視線方向ごとに入射する光によって結像される画像 $L^g(x, u)$ を演算する。そして、被写界深度効果レンダラRは、式(6)に示すように光学系の開口を模擬するインディケータ $A(u)$ を用いて画像 $L^g(x, u)$ を統合することで、シャローデプスオブフィールドイメー ジ I_s を生成する。

[0031] [数4]

$$M^g(x, u) = T(D^g(x), u) \quad \dots (4)$$

[0032] [数5]

$$L^g(\mathbf{x}, \mathbf{u}) = I_d^g(\mathbf{x} + \mathbf{u}(M^g(\mathbf{x}, \mathbf{u}) + \hat{m})) \quad \dots (5)$$

[0033] [数6]

$$I_s^g(\mathbf{x}) = \sum_{\mathbf{u}} A(\mathbf{u}) L^g(\mathbf{x} + \mathbf{u}, \mathbf{u}) \quad \dots (6)$$

[0034] 式(5)における $\hat{m}$ は、ピント位置までの距離(過焦点距離)を示す。 $\hat{m}$ は、学習によって獲得される値であってもよいし、所定の分布から抽出される乱数であってもよいし、定数(例えばゼロ)であってもよい。

インディケータ $A(\mathbf{u})$ は、開口部に相当する要素の値が正值で、開口部以外の要素の値が0であり、すべての要素の値の和が1となる行列である。インディケータ $A(\mathbf{u})$ が模擬する開口部の形状は例えば円形である。このように、被写界深度効果レンダラ R には、光線空間を考慮した光学的な制約が与えられている。これにより、被写界深度効果レンダラ R は、光学的に整合性のとれた画像変換を実現することができる。

[0035] なお、被写界深度効果レンダラ R は、深度データ $D_g(\mathbf{x})$ に基づくワーピング関数(変形関数)によって構成されてもよいし、ニューラルネットワークモデルによって構成されてもよい。被写界深度効果レンダラ R がニューラルネットワークモデルを構成に含む場合、ワーピング関数による変形結果に基づく制約を与えられてもよい。具体的には、被写界深度効果レンダラ R は、ワーピング関数によって変形された深度データ D_g をニューラルネットワークモデルに入力するように構成されてもよい。

[0036] 識別部137は、ディープデプスオブフィールドイメージ I_d 、シャローデプスオブフィールドイメージ I_s および学習用データ記憶部131が記憶する撮像画像を識別モデル C に入力することで、入力された画像データが実際の撮像画像である度合を示す評価値を算出する。

[0037] 算出部138は、画像生成モデル G_I 、深度生成モデル G_D および識別モデル C の学習に用いる学習基準(損失関数)を算出する。具体的には、算出部

138は、敵対的学習基準に基づいて学習基準を算出する。

敵対的学習基準 $\mathcal{L}_{AR-GAN}$ とは、画像データが実際の撮像画像であるか撮像画像を模擬した画像データであるかの判断の正確さを示す指標である。算出部138は、以下の式(7)に示すように、敵対的学習基準 $\mathcal{L}_{AR-GAN}$ を求める。

[0038] [数7]

$$\mathcal{L}_{AR-GAN} = \mathbb{E}_{I^r \sim P^r(I)} [\log C(I^r)] + \mathbb{E}_{z \sim P_z(z), s \sim P_s(s)} [\log(1 - C(R(G_I(z), sG_D(z))))] \quad \dots (7)$$

[0039] 式(7)において、 $s \sim P_s(s)$ は、ディープデプスオブフィールドイメージ I_d^s とシャローデプスオブフィールドイメージ I_s^s の混合度、すなわち被写界深度の深さの度合いを表す。分布 $P_s(s)$ は、0以上1以下の値域に係る分布、例えば二項分布や一様分布などを用いることができる。また、 $z \sim P_z(z)$ は、潜在変数 z を分布 $P_z(z)$ から抽出する処理を示す。なお、式(7)では学習基準としてクロスエントロピーを用いるが、これに限らず、L1距離やL2距離、ワッサーズタイン距離などの任意の距離基準に基づく学習基準を用いてもよい。

[0040] 更新部139は、算出部138が算出した敵対的学習基準 $\mathcal{L}_{AR-GAN}$ に基づいて画像生成モデル G_I 、深度生成モデル G_D および識別モデル C のパラメータを更新する。具体的には、更新部139は、識別モデル C について、敵対的学習基準 $\mathcal{L}_{AR-GAN}$ が大きくなるようにパラメータを更新する。また更新部139は、画像生成モデル G_I および深度生成モデル G_D について、敵対的学習基準 $\mathcal{L}_{AR-GAN}$ が小さくなるようにパラメータを更新する。また、被写界深度効果レンダラ R が学習可能なパラメータを持つ場合、被写界深度効果レンダラ R について、敵対的学習基準 $\mathcal{L}_{AR-GAN}$ が小さくなるようにパラメータを更新する。

[0041] シャローデプスオブフィールドイメージ I_s^s は、光学的制約を有する被写界深度効果レンダラ R を用いて、深度データ D^s から生成される。そのため、識別モデル C がシャローデプスオブフィールドイメージ I_s^s を実際の撮像画

像であると誤判定させるためには、上記光学的制約の下、適切な深度データ D^g を生成する必要がある。したがって、第1の実施形態に係るモデル学習装置13は、上記の学習基準に従ってパラメータを更新することで、ディープデプスオブフィールドイメー ジ I^g_d 、シャローデプスオブフィールドイメー ジ I^g_s および深度データ D^g の組が適切に生成できるように画像生成モデル G_I および深度生成モデル G_D のパラメータを更新することができる。また、被写界深度効果レンダラ R が学習可能なパラメータを持つ場合、被写界深度効果レンダラ R のパラメータを更新することができる。

[0042] 《モデル学習装置13の動作》

図3は、第1の実施形態に係るモデル学習装置13の動作を示すフローチャートである。図4は、第1の実施形態に係る学習処理におけるデータの変遷を示す図である。

モデル学習装置13が学習処理を開始すると、以下に示すステップS1からステップS6の処理を、所定回数繰り返し実行する。まず潜在変数生成部133は、乱数と所定の分布とに基づいて潜在変数 z を生成する（ステップS1）。次に、画像生成部134は、ステップS1で生成した潜在変数 z をモデル記憶部132が記憶する画像生成モデル G_I に入力することで、ディープデプスオブフィールドイメー ジ I^g_d を生成する（ステップS2）。

[0043] また三次元データ生成部135は、ステップS1で生成した潜在変数 z をモデル記憶部132が記憶する深度生成モデル G_D に入力することで、深度データ D^g を生成する（ステップS3）。次に、変換部136は、0以上1以下の混合度 s を所定の分布に従って決定する（ステップS4）。変換部136は、ステップS2で生成したディープデプスオブフィールドイメー ジ I^g_d と、ステップS3で生成した深度データ D^g に混合度 s を乗算したものとを、モデル記憶部132が記憶する被写界深度効果レンダラ R に入力することで、シャローデプスオブフィールドイメー ジ I^g_s を生成する（ステップS5）。なお、混合度 s がゼロである場合、生成されるシャローデプスオブフィールドイメー ジ I^g_s はディープデプスオブフィールドイメー ジ I^g_d と一致する。

次に、識別部137は、シャローデプスオブフィールドイメージ I_s を識別モデルCに入力することで、入力されたシャローデプスオブフィールドイメージ I_s が実際の撮像画像である度合を示す評価値を算出する（ステップS6）。

[0044] モデル学習装置13は、所定数の潜在変数 z から生成されたシャローデプスオブフィールドイメージ I_s についての評価値を算出すると、以下に示すステップS7からステップS8の処理を所定回数繰り返し実行する。まず識別部137は、学習用データ記憶部131から任意の撮像画像を読み出す（ステップS7）。識別部137は、読み出した撮像画像を識別モデルCに入力することで、入力された撮像画像が実際の撮像画像である度合を示す評価値を算出する（ステップS8）。

[0045] 算出部138は、ステップS6で算出した評価値およびステップS8で算出した評価値を用いて、上述の式（7）に基づいて敵対的学習基準 L_{AR-GAN} を算出する（ステップS9）。更新部139は、ステップS9で算出した敵対的学習基準 L_{AR-GAN} に基づいて画像生成モデル G_I 、深度生成モデル G_D および識別モデルCのパラメータを更新する（ステップS10）。また、被写界深度効果レンダラRが学習可能なパラメータを持つ場合、被写界深度効果レンダラRのパラメータを更新する。

[0046] 更新部139は、ステップS1からステップS10によるパラメータの更新を、所定のエポック数だけ繰り返し実行したか否かを判定する（ステップS11）。繰り返しが所定のエポック数に満たない場合（ステップS11：NO）、モデル学習装置13はステップS1に処理を戻し、学習処理を繰り返し実行する。

[0047] 他方、繰り返しが所定のエポック数に達した場合（ステップS11：YES）、モデル学習装置13は学習処理を終了する。これにより、モデル学習装置13は、学習済みモデルである画像生成モデル G_I および深度生成モデル G_D を生成することができる。また、被写界深度効果レンダラRが学習可能なパラメータを持つ場合、学習済みモデルである被写界深度効果レンダラRを

生成することができる。

[0048] 《データ生成装置 11 の構成》

図 5 は、第 1 の実施形態に係るデータ生成装置 11 の構成を示す概略ブロック図である。

第 1 の実施形態に係るデータ生成装置 11 は、モデル記憶部 111、潜在変数生成部 112、画像生成部 113、三次元データ生成部 114、変換部 115、出力部 116 を備える。

[0049] モデル記憶部 111 は、モデル学習装置 13 による学習済みの画像生成モデル G_I および深度生成モデル G_D 、およびモデル学習装置 13 と同じ被写界深度効果レンダラ R を記憶する。

[0050] 潜在変数生成部 112、画像生成部 113、三次元データ生成部 114 および変換部 115 は、モデル学習装置 13 が備える、潜在変数生成部 133、画像生成部 134、三次元データ生成部 135 および変換部 136 と同様の処理を実行する。なお、変換部 115 は、シャローデプスオブフィールドイメージ I_s を生成する際、混合度 s を 1 として計算する。

[0051] 出力部 116 は、画像生成部 113 が生成したディープデプスオブフィールドイメージ I_d 、三次元データ生成部 114 が生成した深度データ D_s および変換部 115 が生成したシャローデプスオブフィールドイメージ I_s を出力する。

[0052] 《作用・効果》

これにより、データ生成装置 11 は、二次元画像であるディープデプスオブフィールドイメージ I_d およびシャローデプスオブフィールドイメージ I_s と、三次元データである深度データ D_s との組を生成することができる。データ生成装置 11 がこれらのデータの組を生成するために用いる画像生成モデル G_I および深度生成モデル G_D は、三次元データに関する補足データなしに学習されたものである。つまり、第 1 の実施形態に係るデータ生成システム 1 によれば、三次元データに関する補足データなしに、二次元画像と、当該二次元画像に対応する三次元データとを生成することができる。

[0053] これは、モデル学習装置 13 が、ディープデプスオブフィールドイメー ジ I_d の経験分布およびディープデプスオブフィールドイメー ジ I_s と深度データ D_s から生成されたシャローデプスオブフィールドイメー ジ I_s の経験分布が、いずれも実際の撮像画像の経験分布に近くなるようにモデルを学習させるためである。つまり、ディープデプスオブフィールドイメー ジ I_d と深度データ D_s から生成されたシャローデプスオブフィールドイメー ジ I_s が、実際の撮像画像の経験分布に近くなるためには、適切な深度データ D_s が得られている必要があり、シャローデプスオブフィールドイメー ジ I_s が、実際の撮像画像の経験分布に近くなったということは、モデルが適切な深度データ D_s を得ることができるように学習されたことを示すためである。

[0054] 《実験結果》

第 1 の実施形態に係るデータ生成システム 1 を用いたデータペアの生成の実験結果の一例を説明する。実験では、学習用データとして花画像、鳥画像、顔画像に係る撮像画像が用いられた。

[0055] 実験では、データ生成システム 1 は正規分布 $N(0, 1)$ に基づいてランダムに潜在変数 z を抽出した。画像生成モデル G_I 、深度生成モデル G_D および識別モデル C は、CNN によって構成した。被写体深度効果レンダラ R の変形関数 T は、 $D_s(x)$ に基づいてワーピングを行った後、CNN を適用する構成を用いた。

[0056] 実験において、非特許文献 1 に記載の RGBD-GAN を比較例とした。

評価方法として、第 1 の実施形態に係るデータ生成システム 1 および非特許文献 1 の手法のそれぞれで生成されたデータペアを教師データとして用いた画像データから深度データへの変換器の精度を比較する方法を採用した。具体的には、第 1 の実施形態に係るデータ生成システム 1 および非特許文献 1 の手法のそれぞれで生成されたデータペアを教師データとして用いた変換器による計算結果と、実際の撮像画像と実際の深度データのペア（実ペアデータ）とを教師データとして用いた変換器による計算結果との一致度を比較した。評価尺度は、Scale-Invariant Depth Error (SIDE) を用いた。S

I D E は、値が小さいほど性能がよいことを示す。

[0057] その結果、第 1 の実施形態に係るデータ生成システム 1 が非特許文献 1 の手法よりも S I D E の値が小さいことを確認した。すなわち、第 1 の実施形態に係るデータ生成システム 1 が非特許文献 1 の手法よりも実ペアデータを用いて学習した変換結果との一致度が高いことを確認した。なお、非特許文献 1 の手法は、視点に関する手がかりを予め取得しておく必要があるのに対し、第 1 の実施形態に係るデータ生成システム 1 ではこのような情報が不要である。また、非特許文献 1 の手法は、画像データと深度データのペアを生成するのに対し、第 1 の実施形態に係るデータ生成システム 1 は、ディープデプスオブフィールドイメージ I_d 、深度データ D_d およびシャローデプスオブフィールドイメージ I_s の組を得ることができる。

[0058] 〈第 2 の実施形態〉

第 1 の実施形態に係るデータ生成システム 1 は、光線空間に基づく光学的な制約を有する被写界深度効果レンダラ R を用いてシャローデプスオブフィールドイメージ I_s を生成する。これに対し、第 2 の実施形態に係るデータ生成システム 1 は、深度とボケ効果の大きさとの関係に基づく光学的な制約を有する被写界深度効果レンダラ R を用いてシャローデプスオブフィールドイメージ I_s を生成する。

[0059] 第 2 の実施形態に係る被写界深度効果レンダラ R は、以下の式 (8)、(9) で表されるカーネル k を用いて計算を行う。

[0060] [数 8]

$$\hat{k}(x, m) = [\|x\|_2^2 \leq m^2] \quad \dots (8)$$

[0061] [数 9]

$$k(x, m) = \hat{k}(x, m) / \sum \hat{k}(x, m) \quad \dots (9)$$

[0062] 式 (8) において、 $[\cdot]$ はアイバーソンの記法を表し、括弧内の条件が真ならば 1、偽ならば 0 をとる。m は、被写体の位置を基準としたピント位置の深度を示す。深度 m がゼロであることは、被写体にフォーカスがあつて

いることを示す。深度 m が正の値であることは、ピント位置が被写体より手前側に位置することを示す。深度 m が負の値であることは、ピント位置が被写体より奥側に位置することを示す。この定義により、 $\hat{k}(x, m)$ は、半径 $|m|$ 以内の中央部のみ値 1 を持ち、それ以外は値 0 を持つ。つまり、深度 $|m|$ が大きければ大きいほど、中央の円形部の大きさは大きくなり、このカーネルを画像に畳み込んだ時のボケ効果は大きくなる。また式 (9) は、 $k(x, m)$ の要素の値の合計が 1 となるように $\hat{k}(x, m)$ を正規化する処理を表す式である。

[0063] また、第 2 の実施形態では、三次元データ生成部 135 は、 D^g を算出する前段階として、三次元データ生成部 135 は、深度情報の確率分布 $P(x, m)$ を生成する。ここで、 $P(x, m)$ は位置座標 x の深度 m における深度の存在確率を表す。つまり、各座標 x について存在確率 $P(x, m)$ の総和は 1 となる。 $P(x, m)$ が得られたとき、 $D^g(x)$ は、例えば式 (10) に示すように、 $P(x, m)$ に対して、各座標 x ごとに最大となる深度 m を算出することによって求めることが可能である。

[0064] [数10]

$$D^g(x) = \underset{m}{\operatorname{argmax}} P(x, m) \quad \dots (10)$$

[0065] 三次元データ生成部 135 は、式 (10) の演算を行う際、前処理として、 $P(x, m)$ に対してスムージングなどを行なってもよい。そして、変換部 136 は、以下の式 (11) により、カーネル $k(x, m)$ とディープデプスオブフィールドイメーজ I_d^g とを畳み込んだ結果と、深度データの確率分布 $P(x, m)$ とを乗算し、その和をとることによりシャローデプスオブフィールドイメージ I_s^g を合成することができる。

[0066] [数11]

$$I_s^g = \sum_m P(x, m - \hat{m})(I_d^g(x) * k(x, m)) \quad \dots (11)$$

[0067] 式 (11) において演算子 $*$ は畳み込み演算を表す。このように、第 2 の実施形態に係る被写界深度効果レンダラ R を用いた演算は、形状が事前に定

められたカーネル k との畳み込みと、その重み付き和とに制約される。これにより、変換部136は、第1の実施形態と同様に、光学的に整合性のとれた変換を実現することができ、変換の過程において画像の内容を大きく棄損することなくボケ度合だけを変えることができる。

また、第1の実施形態に係るデータ生成システム1では、式(3)において、深度データに混合度 s を乗算することによって被写界深度の度合いを調整するが、第2の実施形態に係るデータ生成システム2においては、式(11)において、 $k(x, m)$ の m に混合度 s を乗算する(つまり、 $k(x, s \cdot m)$ とする)ことによって、被写界深度の度合いを調整することが可能である。例えば、 s が0の場合、 $k(x, m)$ は中心のみが1で、それ以外は0のカーネルとなるため、式(11)の出力は、被写界深度の深い画像になる。

[0068] 〈第3の実施形態〉

画像におけるボケは、被写体がピント位置より手前側にある場合と、奥側にあるために発生している場合とのそれぞれにおいて発生する。一方で、画像データから、被写体が手前側に存在するためにボケが生じているか、奥側に存在するためにボケが生じているかを判断することは困難である。そこで、第3の実施形態に係るデータ生成システム1では、ピントが合う被写体(対象物)が画像の中心近傍に存在する可能性が高く、また対象物の近傍に写る被写体は背景であることが多いというヒューリスティックスに基づいて、モデルの更新に用いる目的関数を算出する。

[0069] 第3の実施形態に係る算出部138は、式(7)に示す敵対的学習基準 L_A 、 $R-GAN$ に加え、以下の式(12)、(13)で表される深度学習基準 L_p を求める。

[0070] [数12]

$$D_P = \begin{cases} 0 & (r \leq r_{th}) \\ -g(r - r_{th}) & (r > r_{th}) \end{cases} \quad \dots (12)$$

[0071]

[数13]

$$\mathcal{L}_p = \lambda_p \|D^g - D_p\|_2^2 \quad \dots (13)$$

[0072] 式(12)において r は画像中心からの距離を示し、 r_{th} と g はそれぞれ対象物の大きさと深さを定めるハイパーパラメータである。式(12)に示す事前深度データ D_p は、半径 r_{th} 以内はピント位置にあることを示し、中心からの距離が r_{th} より遠いほど、ピント位置より奥側に位置することを示す。また式(13)において λ_p は深度学習基準 $\mathcal{L}_p$ の重みを表すハイパーパラメータである。深度学習基準 $\mathcal{L}_p$ は、深度データ D_g と事前深度データ D_p との距離がゼロに近いほど低くなる。つまり、深度学習基準 $\mathcal{L}_p$ は、深度データ D_g のうち少なくともフォーカスされる可能性が高い部分と予め定めたピント位置に係る深度との距離がゼロに近いほど低くなる。なお、式(13)では学習基準としてL2距離に基づくものを用いるが、これに限られず、L1距離やワッサースタイン距離などの任意の距離基準に基づく学習基準を用いてもよい。

[0073] 上記の例では、予め定めたピント位置での深度をゼロ、当該位置以外での深度を、ピント位置より奥側となるように予め定めた値としているが、これに限られない。例えば、ヒューリスティックスに基づいて、画像の一部または全体領域における各位置での深度を予め定めておくことができる場合、深度学習基準 $\mathcal{L}_p$ は、各位置において深度データ D_g と当該予め定めた深度との距離がゼロに近いほど低くなるものであってよい。また、式(13)に示す深度学習基準 $\mathcal{L}_p$ は、画像の全体領域について深度データ D_g が示す深度と予め定めた深度との距離を求めるが、これに限られない。例えば深度学習基準 $\mathcal{L}_p$ は、深度データ D_g のうち半径 r_{th} 以内の領域や、ヒューリスティックスに基づいてフォーカスされる可能性が高いと判定された領域など、深度データ D_g の一部領域のみにおける距離に基づいて計算されるものであってもよい。また例えば深度学習基準 $\mathcal{L}_p$ は、深度データ D_g のうち半径 r_{th} より外側の領域や、ヒューリスティックスに基づいてフォーカスされる可能性が低いと

判定された領域などにおける深度が、遠景または近景に係る予め定めた深度に近いほど低くなるものであってもよい。

[0074] 更新部139は、敵対的学習基準 L_{AR-GAN} と深度学習基準 L_p の和に基づいて、深度生成モデル G_D のパラメータを更新する。これにより、中心にフォーカスがあい、周囲が背景となる深度データ D_p が生成されることが促進される。なお、更新部139は、深度学習基準 L_p を学習の全工程において用いてもよいし、学習の途中段階まで用いてもよい。更新部139は、深度学習基準 L_p を所定のエポック数に至るまで用い、以降用いないようにすることで、実際の深度データと事前深度データ D_p とにギャップがあった場合のネガティブな効果を抑制することができる。

[0075] 《変形例》

第3の実施形態に係るデータ生成システム1は、対象物が画像の中心近傍に存在する可能性が高く、また対象物の近傍に写る被写体は背景であることが多いというヒューリスティックスに基づいて、式(12)に基づいて事前深度データ D_p を算出するが、これに限られない。例えば、他の実施形態においては、対象物の近傍に写る被写体は前景であることが多いというヒューリスティックスに基づいて計算してもよいし、人の顔がフォーカスされる可能性が高いというヒューリスティックスに基づいて、パターンマッチング処理により検出された顔の位置に基づいて事前深度データ D_p を算出してもよい。また、 r_{th} および g をハイパーパラメータではなく学習により更新するパラメータとしてもよい。

[0076] 〈他の実施形態〉

以上、図面を参照して一実施形態について詳しく説明してきたが、具体的な構成は上述のものに限られることはなく、様々な設計変更等を行うことが可能である。すなわち、他の実施形態においては、上述の処理の順序が適宜変更されてもよい。また、一部の処理が並列に実行されてもよい。

上述した実施形態に係るデータ生成システム1は、データ生成装置11とモデル学習装置13とを備えるが、単独のコンピュータによって構成される

ものであってもよい。

[0077] 上述した実施形態に係るデータ生成システム 1 は、GANs によって画像生成モデル G_I および深度生成モデル G_D のパラメータを更新するが、これに限られない。例えば、他の実施形態に係るデータ生成システム 1 は、Variational Autoencoder、Flow Model、Denoising Diffusion Probabilistic Model などの任意の生成モデルにおける学習基準によって画像生成モデル G_I および深度生成モデル G_D のパラメータを更新してもよい。

[0078] また、上述の実施形態に係る被写界深度効果レンダラ R は、光学系の制約を模擬する数理モデルであるが、これに限られない。例えば被写界深度効果レンダラ R は、ニューラルネットワークによって構成された学習済みモデルであってもよい。また、光学系に関わるパラメータなど被写界深度効果レンダラ R の一部のパラメータを学習パラメータとして持ってもよい。また、被写界深度効果レンダラ R が学習可能なパラメータを持つ場合、画像生成モデル G_I および深度生成モデル G_D と同時に学習してもよい。

[0079] また、上述の実施形態に係るデータ生成装置 11 は、ディープデプスオブフィールドイメージ I_d 、シャローデプスオブフィールドイメージ I_s および深度データ D の組を出力するが、これに限られない。例えば、他の実施形態に係るデータ生成装置 11 は、ディープデプスオブフィールドイメージ I_d と深度データ D の組を出力し、シャローデプスオブフィールドイメージ I_s を出力しなくてもよい。また、他の実施形態に係るデータ生成装置 11 は、シャローデプスオブフィールドイメージ I_s と深度データ D の組を出力し、ディープデプスオブフィールドイメージ I_d を出力しなくてもよい。また、他の実施形態に係るデータ生成装置 11 は、ディープデプスオブフィールドイメージ I_d とシャローデプスオブフィールドイメージ I_s の組を出力し、深度データ D を出力しなくてもよい。また、他の実施形態に係るデータ生成装置 11 は、ディープデプスオブフィールドイメージ I_d 、シャローデプスオブフィールドイメージ I_s および深度データ D の組の少なくとも一部を統合したデータを出力してもよい。例えばデータ生成装置 11 は、ディープデ

プスオブフィールドイメージ I_d と深度データ D_d を、深度情報を含む画像データとして出力してもよい。

[0080] 〈コンピュータ構成〉

図6は、少なくとも1つの実施形態に係るコンピュータの構成を示す概略ブロック図である。

コンピュータ20は、プロセッサ21、メインメモリ23、ストレージ25、インタフェース27を備える。

上述のデータ生成装置11およびモデル学習装置13は、コンピュータ20に実装される。そして、上述した各処理部の動作は、プログラムの形式でストレージ25に記憶されている。プロセッサ21は、プログラムをストレージ25から読み出してメインメモリ23に展開し、当該プログラムに従って上記処理を実行する。また、プロセッサ21は、プログラムに従って、上述した各記憶部に対応する記憶領域をメインメモリ23に確保する。プロセッサ21の例としては、CPU (Central Processing Unit)、GPU (Graphic Processing Unit)、マイクロプロセッサなどが挙げられる。

[0081] プログラムは、コンピュータ20に発揮させる機能の一部を実現するためのものであってもよい。例えば、プログラムは、ストレージに既に記憶されている他のプログラムとの組み合わせ、または他の装置に実装された他のプログラムとの組み合わせによって機能を発揮させるものであってもよい。なお、他の実施形態においては、コンピュータ20は、上記構成に加えて、または上記構成に代えてPLD (Programmable Logic Device) などのカスタムLSI (Large Scale Integrated Circuit) を備えてもよい。PLDの例としては、PAL (Programmable Array Logic)、GAL (Generic Array Logic)、CPLD (Complex Programmable Logic Device)、FPGA (Field Programmable Gate Array) が挙げられる。この場合、プロセッサ21によって実現される機能の一部または全部が当該集積回路によって実現されてよい。このような集積回路も、プロセッサの一例に含まれる。

[0082] ストレージ25の例としては、磁気ディスク、光磁気ディスク、光ディス

ク、半導体メモリ等が挙げられる。ストレージ25は、コンピュータ20のバスに直接接続された内部メディアであってもよいし、インタフェース27または通信回線を介してコンピュータ20に接続される外部メディアであってもよい。また、このプログラムが通信回線によってコンピュータ20に配信される場合、配信を受けたコンピュータ20が当該プログラムをメインメモリ23に展開し、上記処理を実行してもよい。少なくとも1つの実施形態において、ストレージ25は、一時的でない有形の記憶媒体である。

[0083] また、当該プログラムは、前述した機能の一部を実現するためのものであってもよい。さらに、当該プログラムは、前述した機能をストレージ25に既に記憶されている他のプログラムとの組み合わせで実現するもの、いわゆる差分ファイル（差分プログラム）であってもよい。

符号の説明

[0084] 1…データ生成システム 11…データ生成装置 111…モデル記憶部
112…潜在変数生成部 113…画像生成部 114…三次元データ生成部
115…変換部 116…出力部 13…モデル学習装置 131…学習用データ記憶部
132…モデル記憶部 133…潜在変数生成部 134…画像生成部
135…三次元データ生成部 136…変換部 137…識別部 138…算出部 139…更新部

請求の範囲

- [請求項1] 潜在変数を機械学習モデルである画像生成モデルに入力することで、第1被写界深度に係る撮像画像を模擬した第1画像データを生成する画像生成部と、
- 前記潜在変数を機械学習モデルである三次元生成モデルに入力することで、前記第1画像データに対応する三次元データを生成する三次元データ生成部と、
- 前記第1画像データと前記三次元データとに基づいて、第2被写界深度に係る撮像画像を模擬した第2画像データを生成する変換部と、
- 前記第1画像データおよび前記第2画像データの分布と実際の撮像画像の経験分布との近さの度合いを示す学習基準値を算出する算出部と、
- 前記学習基準値に基づいて前記画像生成モデルおよび前記三次元生成モデルのパラメータを更新する更新部と
- を備える学習装置。
- [請求項2] 前記第2被写界深度は、前記第1被写界深度より浅い
- 請求項1に記載の学習装置。
- [請求項3] 前記変換部は、前記第1画像データと前記三次元データとを光学系の制約を模擬する数理モデルに入力することで、前記第2画像データを生成する
- 請求項1または請求項2に記載の学習装置。
- [請求項4] 画像データを入力とし、前記画像データが実際の撮像画像であるか撮像画像を模擬した画像データであるかを判定する機械学習モデルである識別モデルに、前記第1画像データおよび前記第2画像データの少なくとも一方、並びに前記実際の撮像画像を入力し、入力された画像データを識別する識別部を備え、
- 前記学習基準値は、前記識別部による識別しにくさの度合いを示し、

前記更新部は、前記学習基準値に基づいてさらに前記識別モデルのパラメータを更新する

請求項 1 から請求項 3 の何れか 1 項に記載の学習装置。

[請求項5]

前記学習基準値は、画像データの分布と実際の撮像画像の経験分布との近さの度合いを示す第 1 基準値と、前記三次元データのうち少なくとも一部に係る各位置と予め定めた深度との距離に係る第 2 基準値の和である

請求項 1 から請求項 4 の何れか 1 項に記載の学習装置。

[請求項6]

前記三次元データは、過焦点距離を基準とした深度を表す深度データであって、

前記変換部は、前記三次元データまたは深度の大きさを表すパラメータに係数を乗算することで、前記第 2 被写界深度の深さを異ならせる

請求項 1 から請求項 5 の何れか 1 項に記載の学習装置。

[請求項7]

潜在変数を機械学習モデルである画像生成モデルに入力することで、第 1 被写界深度に係る撮像画像を模擬した第 1 画像データを生成するステップと、

前記潜在変数を機械学習モデルである三次元生成モデルに入力することで、前記第 1 画像データに対応する三次元データを生成するステップと、

前記第 1 画像データと前記三次元データとに基づいて、第 2 被写界深度に係る撮像画像を模擬した第 2 画像データを生成するステップと、

前記第 1 画像データおよび前記第 2 画像データの分布と、実際の撮像画像の経験分布との近さの度合いを示す学習基準値を算出するステップと、

前記学習基準値に基づいて前記画像生成モデルおよび前記三次元生成モデルのパラメータを更新するステップと、

学習済みの前記画像生成モデルおよび前記三次元生成モデルを出力するステップと

を有する学習済みモデルの生成方法。

[請求項8] 請求項7に記載の学習済みモデルの生成方法によって出力された画像生成モデルに潜在変数を入力することで、撮像画像を模擬した画像データを生成する画像生成部と、

請求項7に記載の学習済みモデルの生成方法によって出力された三次元生成モデルに前記潜在変数を入力することで、三次元データを生成する三次元データ生成部と、

を備えるデータ生成装置。

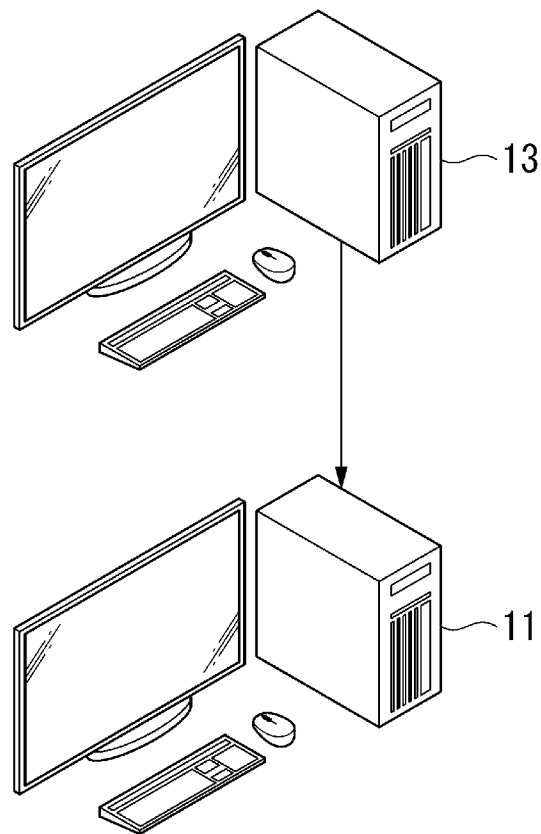
[請求項9] 請求項7に記載の学習済みモデルの生成方法によって出力された画像生成モデルに潜在変数を入力することで、撮像画像を模擬した画像データを生成するステップと、

請求項7に記載の学習済みモデルの生成方法によって出力された三次元生成モデルに前記潜在変数を入力することで、三次元データを生成するステップと、

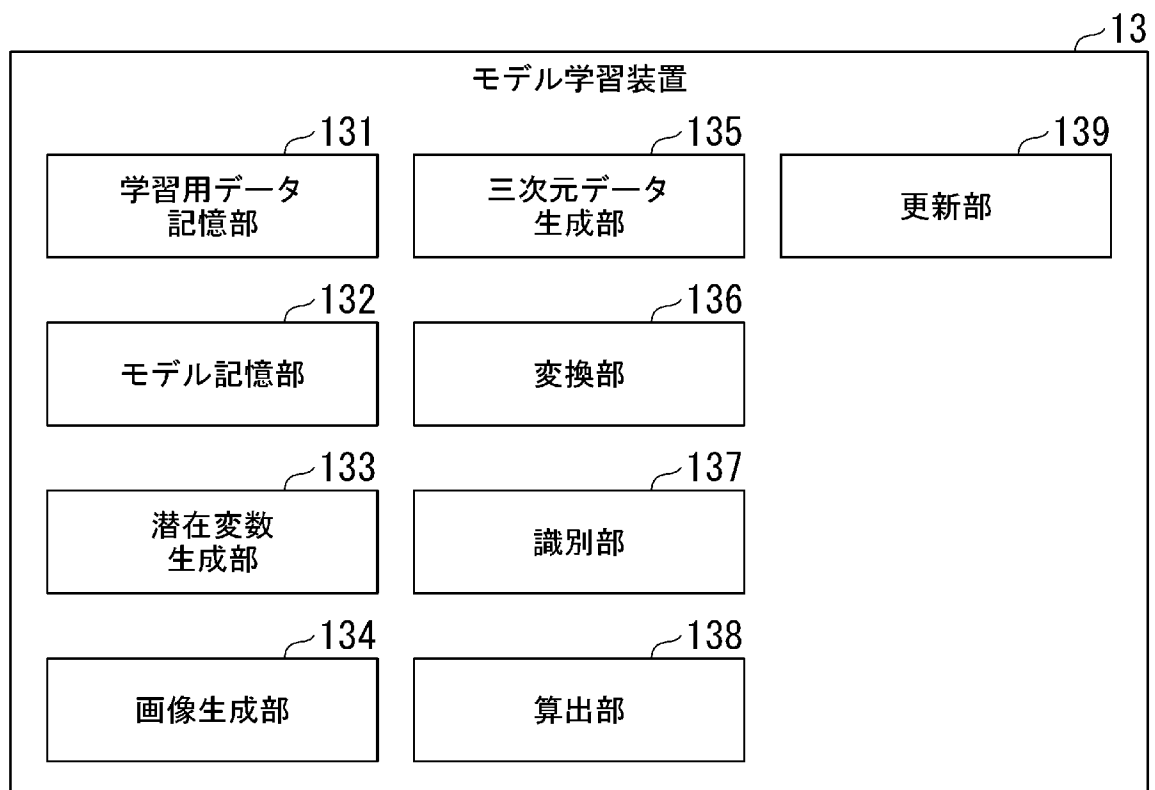
を備えるデータ生成方法。

[請求項10] コンピュータを、請求項1から請求項6の何れか1項に記載の学習装置として機能させるためのプログラム。

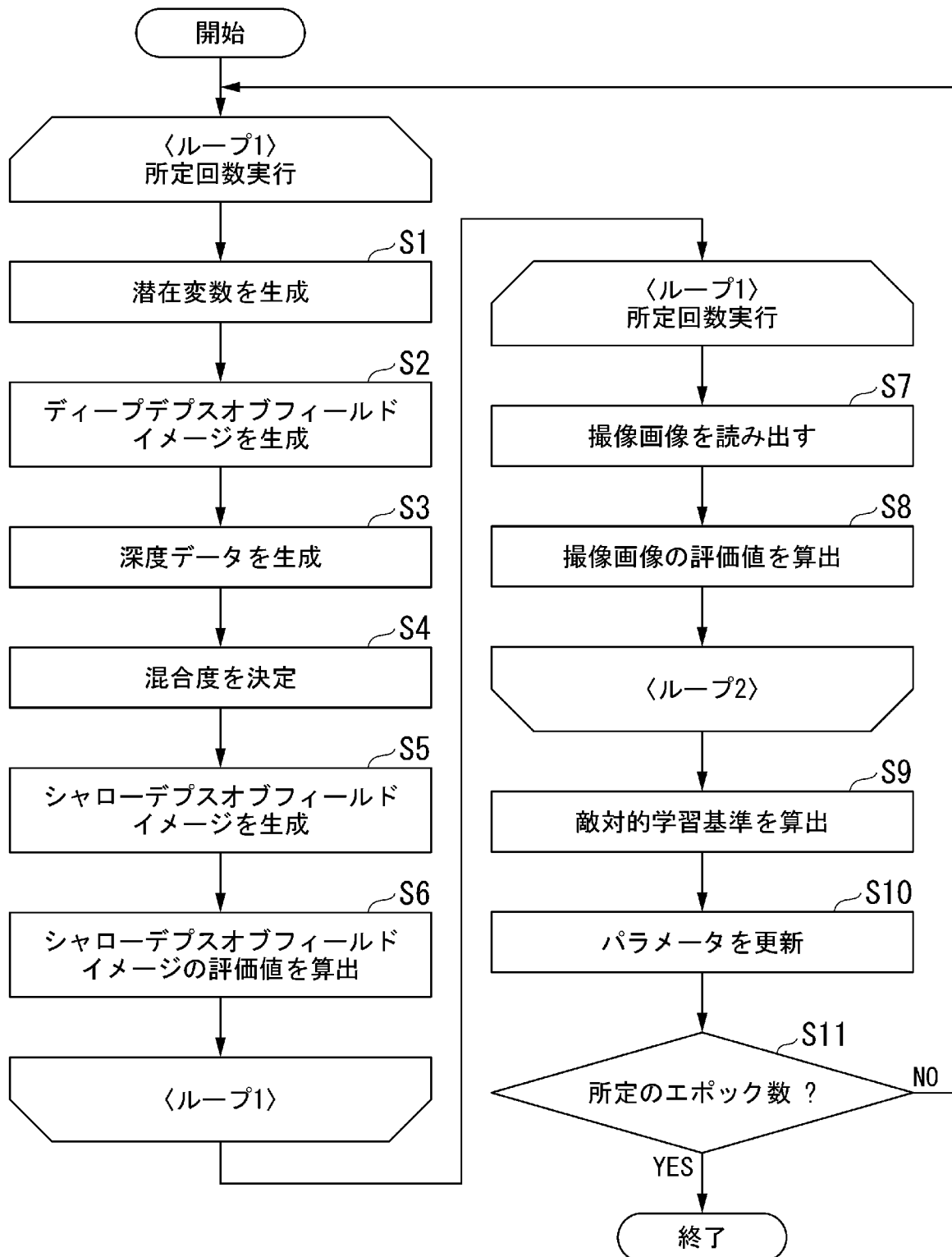
[図1]



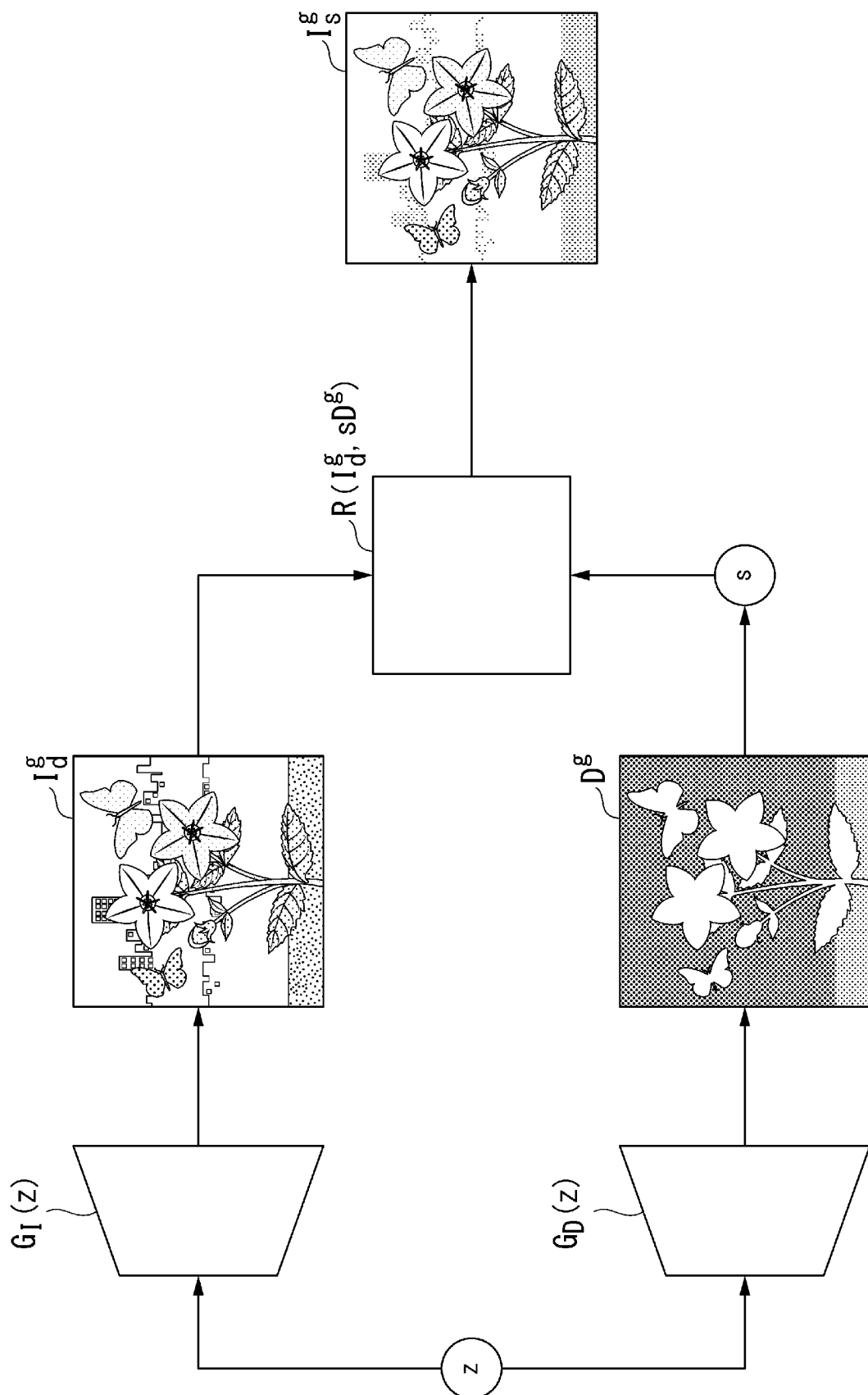
[図2]



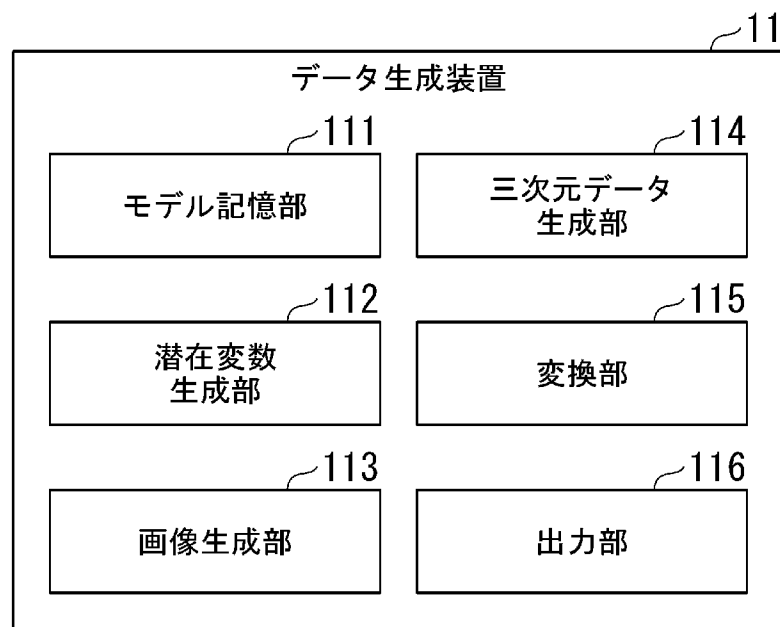
[図3]



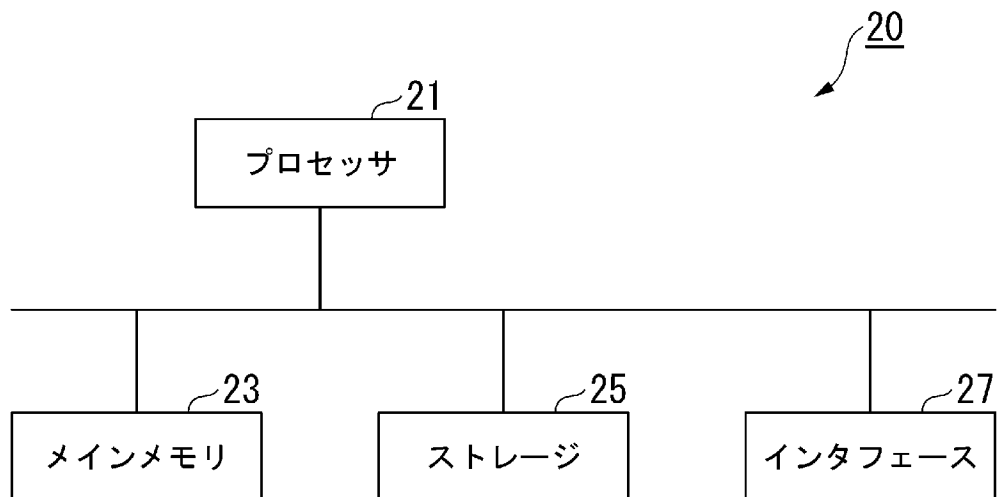
[図4]



[図5]



[図6]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2021/019579

A. CLASSIFICATION OF SUBJECT MATTER

G06T 7/00(2017.01)i

FI: G06T7/00 350C

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

G06T7/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2021

Registered utility model specifications of Japan 1996-2021

Published registered utility model applications of Japan 1994-2021

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2020-042818 A (BAIDU ONLINE NETWORK TECHNOLOGY BEIJING CO LTD) 19 March 2020 (2020-03-19) entire text, all drawings	1-10
A	NOGUCHI, Atsuhiko et al. RGBD-GAN: Unsupervised 3D Representation Learning From Natural Image Datasets via RGBD Image Synthesis. ICLR 2020., 27 September 2019 entire text, all drawings	1-10

☐ Further documents are listed in the continuation of Box C.

☒ See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search

05 July 2021 (05.07.2021)

Date of mailing of the international search report

10 August 2021 (10.08.2021)

Name and mailing address of the ISA/

Japan Patent Office

3-4-3, Kasumigaseki, Chiyoda-ku,

Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

Information on patent family members

International application No.

PCT/JP2021/019579

Patent Documents referred in the Report	Publication Date	Patent Family	Publication Date
---	---------------------	---------------	---------------------

JP 2020-042818 A

19 Mar. 2020

US 2020/0082553 A1
entire text, all
drawings
EP 3621036 A1
CN 109242903 A

A. 発明の属する分野の分類（国際特許分類（IPC）） G06T 7/00(2017.01)i FI: G06T7/00 350C														
B. 調査を行った分野														
調査を行った最小限資料（国際特許分類（IPC）） G06T7/00														
最小限資料以外の資料で調査を行った分野に含まれるもの <table border="0"> <tr> <td>日本国実用新案公報</td> <td>1922 - 1996年</td> </tr> <tr> <td>日本国公開実用新案公報</td> <td>1971 - 2021年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996 - 2021年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994 - 2021年</td> </tr> </table>			日本国実用新案公報	1922 - 1996年	日本国公開実用新案公報	1971 - 2021年	日本国実用新案登録公報	1996 - 2021年	日本国登録実用新案公報	1994 - 2021年				
日本国実用新案公報	1922 - 1996年													
日本国公開実用新案公報	1971 - 2021年													
日本国実用新案登録公報	1996 - 2021年													
日本国登録実用新案公報	1994 - 2021年													
国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）														
C. 関連すると認められる文献														
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号												
A	JP 2020-042818 A (パイドゥ オンライン ネットワーク テクノロジー (ペイジン) カンパニー リミテッド) 19.03.2020 (2020-03-19) 全文, 全図	1-10												
A	Atsuhiko Noguchi et al., RGBD-GAN: Unsupervised 3D Representation Learning From Natural Image Datasets via RGBD Image Synthesis, ICLR 2020, 2019.09.27 全文, 全図	1-10												
☐ C欄の続きにも文献が列挙されている。 ☒ パテントファミリーに関する別紙を参照。														
<table border="0"> <tr> <td>* 引用文献のカテゴリー</td> <td>“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの</td> </tr> <tr> <td>“A” 特に関連のある文献ではなく、一般的技術水準を示すもの</td> <td>“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの</td> </tr> <tr> <td>“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの</td> <td>“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの</td> </tr> <tr> <td>“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）</td> <td>“&” 同一パテントファミリー文献</td> </tr> <tr> <td>“O” 口頭による開示、使用、展示等に言及する文献</td> <td></td> </tr> <tr> <td>“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献</td> <td></td> </tr> </table>			* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの	“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの	“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの	“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献	“O” 口頭による開示、使用、展示等に言及する文献		“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献	
* 引用文献のカテゴリー	“T” 国際出願日又は優先日後に公表された文献であって出願と抵触するものではなく、発明の原理又は理論の理解のために引用するもの													
“A” 特に関連のある文献ではなく、一般的技術水準を示すもの	“X” 特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの													
“E” 国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの	“Y” 特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの													
“L” 優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）	“&” 同一パテントファミリー文献													
“O” 口頭による開示、使用、展示等に言及する文献														
“P” 国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献														
国際調査を完了した日 05.07.2021	国際調査報告の発送日 10.08.2021													
名称及びあて先 日本国特許庁(ISA/JP) 〒100-8915 日本国 東京都千代田区霞が関三丁目4番3号	権限のある職員（特許庁審査官） 小池 正彦 5H 8726 電話番号 03-3581-1101 内線 3531													

PCT/JP2021/019579