

(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9月29日(29.09.2022)



(10) 国際公開番号

WO 2022/201803 A1

(51) 国際特許分類:
G06T 7/00 (2017.01) *G01S 17/89* (2020.01)
G01S 17/86 (2020.01)

(21) 国際出願番号: PCT/JP2022/001918

(22) 国際出願日: 2022年1月20日(20.01.2022)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2021-051097 2021年3月25日(25.03.2021) JP

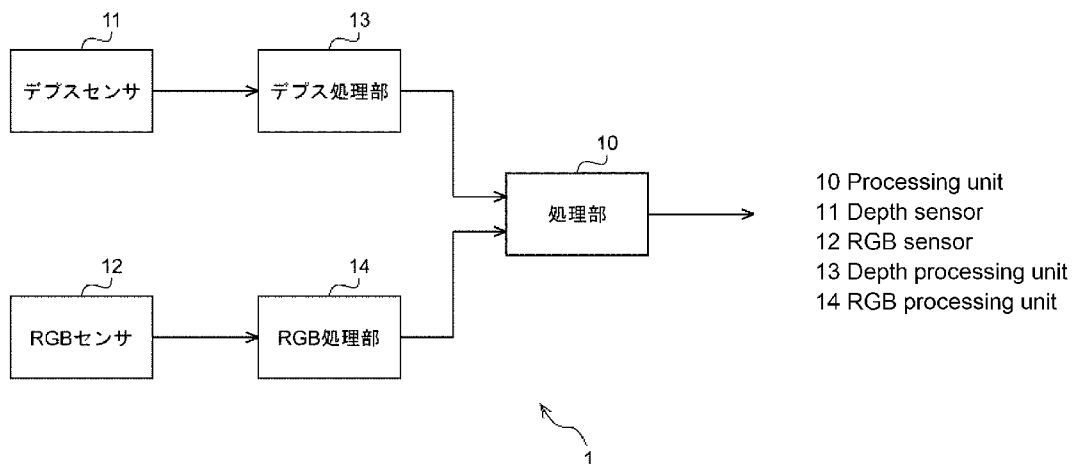
(71) 出願人: ソニーセミコンダクタソリューションズ株式会社(SONY SEMICONDUCTOR SOLUTIONS CORPORATION) [JP/JP]; 〒2430014

神奈川県厚木市旭町四丁目14番1号 Kanagawa (JP).

(72) 発明者: 柴山 憲文 (SHIBAYAMA Noribumi); 〒2430014 神奈川県厚木市旭町四丁目14番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 吉田 隆彦(YOSHIDA Takahiko); 〒2430014 神奈川県厚木市旭町四丁目14番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP). 山田 英史(YAMADA Hideshi); 〒2430014 神奈川県厚木市旭町四丁目14番1号 ソニーセミコンダクタソリューションズ株式会社内 Kanagawa (JP).

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置、情報処理方法、及びプログラム



(57) Abstract: The present invention relates to an information processing device, information processing method and a program which enable more suitably processing correction target pixels in order to use sensor fusion. An information processing device is provided with a processing unit which performs processing using a trained model trained with machine learning on at least a portion of a first image, in which a target acquired by a first sensor is indicated with depth information, a second image, in which an image of the target acquired by a second sensor is indicated with surface information, and a third image obtained from the first image and the second image, and which specifies correction target pixels contained in

(74) 代理人: 西川 孝, 外 (NISHIKAWA Takashi et al.); 〒1700013 東京都豊島区東池袋 3 丁目 9 番 1 0 号 池袋 F N ビル 4 階 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, 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))

the first image. This invention can be applied, for example, to machines having multiple sensors.

(57) 要約: 本開示は、センサフュージョンを用いるに際して、より適切に補正対象画素を処理することができるようにする情報処理装置、情報処理方法、及びプログラムに関する。第1のセンサにより取得された対象物を深度情報で示した第1の画像、第2のセンサにより取得された対象物の像を面情報で示した第2の画像、及び第1の画像と第2の画像から得られる第3の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理を行い、第1の画像に含まれる補正対象画素を特定する処理部を備える情報処理装置が提供される。本開示は、例えば、複数のセンサを有する機器に適用することができる。

明 細 書

発明の名称： 情報処理装置、情報処理方法、及びプログラム

技術分野

[0001] 本開示は、情報処理装置、情報処理方法、及びプログラムに関し、特に、センサフュージョンを用いるに際して、より適切に補正対象画素を処理することができるようにした情報処理装置、情報処理方法、及びプログラムに関する。

背景技術

[0002] 近年、検出原理の異なる複数のセンサを組み合わせ、それらの測定結果を融合するセンサフュージョンに関する研究開発が盛んに行われている。

[0003] 特許文献1には、深度マップの品質を向上させるために、深度測定データ内の欠陥画素を検出し、検出した欠陥画素の深度修正を定義し、検出した欠陥画素の深度測定データに深度修正を適用する技術が開示されている。

先行技術文献

特許文献

[0004] 特許文献1：特表2014-524016号公報

発明の概要

発明が解決しようとする課題

[0005] センサフュージョンを用いるに際しては、処理対象となる画像に、欠陥画素等の補正対象画素が含まれる場合があり、より適切に補正対象画素を処理することが求められる。

[0006] 本開示はこのような状況に鑑みてなされたものであり、センサフュージョンを用いるに際して、より適切に補正対象画素を処理することができるようにするものである。

課題を解決するための手段

[0007] 本開示の第1の側面の情報処理装置は、第1のセンサにより取得された対象物を深度情報で示した第1の画像、第2のセンサにより取得された前記対

象物の像を面情報で示した第2の画像、及び前記第1の画像と前記第2の画像から得られる第3の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記第1の画像に含まれる補正対象画素を特定する処理部を備える情報処理装置である。

[0008] 本開示の第1の側面の情報処理方法、及びプログラムは、上述した本開示の第1の側面の情報処理装置に対応する情報処理方法、及びプログラムである。

[0009] 本開示の第1の側面の情報処理装置、情報処理方法、及びプログラムにおいては、第1のセンサにより取得された対象物を深度情報で示した第1の画像、第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像、及び前記第1の画像と前記第2の画像から得られる第3の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理が行われ、前記第1の画像に含まれる補正対象画素が特定される。

[0010] 本開示の第2の側面の情報処理装置は、第1のセンサにより取得された対象物を深度情報で示した第1の画像、及び第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像を取得し、前記第1の画像と対になる前記第2の画像に基づいて、第3の画像として前記第1の画像を擬似的に生成し、前記第1の画像と前記第3の画像とを比較し、比較結果に基づいて、前記第1の画像に含まれる補正対象画素を特定する処理部を備える情報処理装置である。

[0011] 本開示の第2の側面の情報処理方法、及びプログラムは、上述した本開示の第2の側面の情報処理装置に対応する情報処理方法、及びプログラムである。

[0012] 本開示の第2の側面の情報処理装置、情報処理方法、及びプログラムにおいては、第1のセンサにより取得された対象物を深度情報で示した第1の画像、及び第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像が取得され、前記第1の画像と対になる前記第2の画像に基づいて、第3の画像として前記第1の画像が擬似的に生成され、前記第1の画像と前

記第3の画像とが比較され、比較結果に基づいて、前記第1の画像に含まれる補正対象画素が特定される。

[0013] 本開示の第3の側面の情報処理装置は、第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成する処理部を備え、前記処理部は、前記第1の画像の各画素に応じた第1の位置の深度情報に基づいて、前記第1の位置を前記第2の画像の画像面に写像し、前記第2の画像の各画素に応じた第2の位置のうち、前記第1の位置の深度情報が割り当てられていない第2の位置を、画素補正位置として特定し、機械学習により学習された学習済みモデルを用いて、前記第2の画像における前記画素補正位置の深度情報を推論する情報処理装置である。

[0014] 本開示の第3の側面の情報処理方法、及びプログラムは、上述した本開示の第3の側面の情報処理装置に対応する情報処理方法、及びプログラムである。

[0015] 本開示の第3の側面の情報処理装置、情報処理方法、及びプログラムにおいては、第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成するに際して、前記第1の画像の各画素に応じた第1の位置の深度情報に基づいて、前記第1の位置が前記第2の画像の画像面に写像され、前記第2の画像の各画素に応じた第2の位置のうち、前記第1の位置の深度情報が割り当てられていない第2の位置が、画素補正位置として特定され、機械学習により学習された学習済みモデルを用いて、前記第2の画像における前記画素補正位置の深度情報が推論される。

[0016] 本開示の第4の側面の情報処理装置は、第2のセンサにより取得された対象物の像を色情報で示した第2の画像を、第1のセンサにより取得された前記対象物を深度情報で示した第1の画像の画像面に写像して第3の画像を生成する処理部を備え、前記処理部は、前記第1の画像の各画素に応じた第1

の位置のうち、有効な深度情報が割り当てられていない第1の位置を、画素補正位置として特定し、機械学習により学習された学習済みモデルを用いて、前記第1の画像における前記画素補正位置の深度情報を推論し、前記第1の位置に割り当てられた深度情報に基づいて、前記第2の画像における第2の位置から色情報をサンプリングして、前記第2の位置を前記第1の画像の画像面に写像する情報処理装置である。

[0017] 本開示の第4の側面の情報処理方法、及びプログラムは、上述した本開示の第4の側面の情報処理装置に対応する情報処理方法、及びプログラムである。

[0018] 本開示の第4の側面の情報処理装置、情報処理方法、及びプログラムにおいては、第2のセンサにより取得された対象物の像を色情報で示した第2の画像を、第1のセンサにより取得された前記対象物を深度情報で示した第1の画像の画像面に写像して第3の画像を生成するに際して、前記第1の画像の各画素に応じた第1の位置のうち、有効な深度情報が割り当てられていない第1の位置が、画素補正位置として特定され、機械学習により学習された学習済みモデルを用いて、前記第1の画像における前記画素補正位置の深度情報が推論され、前記第1の位置に割り当てられた深度情報に基づいて、前記第2の画像における第2の位置から色情報をサンプリングして、前記第2の位置が前記第1の画像の画像面に写像される。

[0019] なお、本開示の第1乃至第4の側面の情報処理装置は、独立した装置であってもよいし、1つの装置を構成している内部ブロックであってもよい。

図面の簡単な説明

[0020] [図1]本技術を適用した情報処理装置の構成例を示す図である。

[図2]教師あり学習を用いた場合における学習時の処理を行う学習装置の構成例を示す図である。

[図3]センサフュージョン用のDNNの構造と出力の第1の例を示す図である。

[図4]センサフュージョン用のDNNの構造と出力の第2の例を示す図である。

[図5]教師あり学習を用いた場合における推論時の処理を行う処理部の構成例

を示す図である。

[図6]教師なし学習を用いた場合における学習時の処理を行う学習装置の構成例を示す図である。

[図7]教師なし学習を用いた場合における推論時の処理を行う処理部の構成例を示す図である。

[図8]推論時の処理を行う処理部の構成例を示す図である。

[図9]処理部における特定部の詳細な構成例を示す図である。

[図10]GANを利用したデプス画像の生成例を示す図である。

[図11]特定処理の流れを説明するフローチャートである。

[図12]補正処理の流れを説明するフローチャートである。

[図13]推論時の処理を行う処理部の構成例を示す図である。

[図14]RGB画像とデプス画像の例を示す図である。

[図15]教師あり学習を用いた場合における学習装置と推論部の構成例を示す図である。

[図16]教師なし学習を用いた場合における学習装置と推論部の構成例を示す図である。

[図17]画像生成処理の第1の例の流れを説明するフローチャートである。

[図18]画像生成処理の第2の例の流れを説明するフローチャートである。

[図19]本開示を適用可能なユースケースの第1の例を示す図である。

[図20]本開示を適用可能なユースケースの第2の例を示す図である。

[図21]本開示を適用可能なユースケースの第3の例を示す図である。

[図22]A I 処理を行う装置を含むシステムの構成例を示す図である。

[図23]電子機器の構成例を示すブロック図である。

[図24]エッジサーバ又はクラウドサーバの構成例を示すブロック図である。

[図25]光センサの構成例を示すブロック図である。

[図26]処理部の構成例を示すブロック図である。

[図27]複数の装置間でのデータの流れを示す図である。

発明を実施するための形態

[0021] (装置の構成例)

図1は、本技術を適用した情報処理装置の構成例を示す図である。

[0022] 情報処理装置1は、複数のセンサを組み合わせることでそれらの測定結果を融合するセンサフュージョンに関する機能を有している。図1において、情報処理装置1は、処理部10、デプスセンサ11、RGBセンサ12、デプス処理部13、及びRGB処理部14を含んで構成される。

[0023] デプスセンサ11は、ToF(Time of Flight)センサなどの測距センサである。ToFセンサは、dToF(direct Time of Flight)方式とiToF(indirect Time of Flight)方式のいずれの方式であってもよい。デプスセンサ11は、対象物までの距離を測定し、その結果得られる測距信号をデプス処理部13に供給する。なお、デプスセンサ11は、ストラクチャライト方式のセンサ、LiDAR(Light Detection and Ranging)方式のセンサ、ステレオカメラなどであってもよい。

[0024] デプス処理部13は、DSP等の信号処理回路である。デプス処理部13は、デプスセンサ11から供給される測距信号に対し、デプス現像処理やデプス前処理（例えばリサイズ処理等）などの信号処理を行い、その結果得られるデプス画像データを処理部10に供給する。デプス画像は、対象物を深度情報で示した画像である。なお、デプス処理部13は、デプスセンサ11内に含まれてもよい。

[0025] RGBセンサ12は、CMOS(Complementary Metal Oxide Semiconductor)イメージセンサやCCD(Charge Coupled Device)イメージセンサ等のイメージセンサである。RGBセンサ12は、対象物の像を撮像し、その結果得られる撮像信号をRGB処理部14に供給する。なお、RGBセンサ12は、RGBカメラに限らず、モノクロカメラや赤外線カメラなどであってもよい。

[0026] RGB処理部14は、DSP(Digital Signal Processor)等の信号処理回路である。RGB処理部14は、RGBセンサ12から供給される撮像信号に対し、RGB現像処理やRGB前処理（例えばリサイズ処理等）などの信号処理を行い、その結果得られるRGB画像データを処理部10に供給する。RGB画像は、対象物の像

を色情報（面情報）で示した画像である。なお、RGB処理部 1 4 は、RGBセンサ 1 2 内に含まれてもよい。

[0027] 処理部 1 0 は、CPU(Central Processing Unit)等のプロセッサである。処理部 1 0 には、デプス処理部 1 3 からのデプス画像データと、RGB処理部 1 4 からのRGB画像データとが供給される。

[0028] 処理部 1 0 は、デプス画像データ、RGB画像データ、及びデプス画像データとRGB画像データから得られる画像データの少なくとも一部に機械学習により学習された学習済みモデル（学習モデル）を用いた処理を行う。以下、処理部 1 0 で行われる学習モデルを用いた処理の詳細を説明する。

[0029] < 1. 第 1 の実施の形態 >

[0030] デプスセンサ 1 1 として、ToFセンサを用いてデプス画像を生成した場合に、フライングピクセルと呼ばれる異常画素が含まれることがある。この異常画素のために、デプス画像を用いた認識処理の精度が低下する恐れがある。そこで、以下、機械学習により学習された学習済みモデルを利用して、デプス画像に含まれるフライングピクセルや欠陥画素などの補正対象画素を特定する手法について説明する。ここでは、機械学習を利用するに際して、教師あり学習の場合と教師なし学習の場合をそれぞれ説明する。

[0031] (A) 教師あり学習

図 2 は、教師あり学習を用いた場合における学習時の処理を行う学習装置の構成例を示す図である。

[0032] 図 2 において、学習装置 2 は、視点変換部 1 1 1、欠陥領域指定部 1 1 2、学習モデル 1 1 3、及び減算部 1 1 4 を有する。

[0033] 学習装置 2 には、学習データとしてデプス画像とRGB画像が入力され、デプス画像が視点変換部 1 1 1 に、RGB画像が学習モデル 1 1 3 にそれぞれ供給される。ただし、ここで入力されたデプス画像には、欠陥領域（欠陥画素）が含まれる。

[0034] 視点変換部 1 1 1 は、入力されたデプス画像に対し視点変換処理を行い、その結果得られる視点が変換されたデプス画像である視点変換デプス画像を

、欠陥領域指定部 1 1 2 及び学習モデル 1 1 3 に供給する。

[0035] 視点変換処理としては、撮影パラメータを用いて、デプスセンサ 1 1 からの測距信号から得られたデプス画像を RGB センサ 1 2 の視点に変換する処理が行われ、RGB センサ 1 2 の視点となった視点変換デプス画像が生成される。撮影パラメータとしては、例えば、デプスセンサ 1 1 と RGB センサ 1 2 との相対位置や姿勢に関する情報などが用いられる。

[0036] 欠陥領域指定部 1 1 2 は、視点変換部 1 1 1 から供給された視点変換デプス画像における欠陥領域を指定することで、欠陥領域教師データを生成し、減算部 1 1 4 に供給する。

[0037] 例えば、アノテーション作業としてユーザが目視などで欠陥領域（例えば欠陥画素の領域）を指定することで、欠陥領域を塗り潰した画像、又は視点変換デプス画像における欠陥領域（欠陥画素）の座標などが、欠陥領域教師データとして生成される。欠陥領域や欠陥画素の座標としては、例えば矩形や点を表した座標を用いることができる。

[0038] 学習モデル 1 1 3 は、RGB 画像と視点変換デプス画像を入力とし、欠陥領域を出力としたディープニューラルネットワーク（DNN : Deep Neural Network）による機械学習を行うモデルである。DNN は、対象の全体像から細部までの各々の粒度の概念を階層構造として関連させて学習するディープラーニング（Deep Learning）のうち、多層の人工ニューラルネットワークによる機械学習の手法である。

[0039] また、減算部 1 1 4 では、学習モデル 1 1 3 の出力である欠陥領域と、欠陥領域指定部 1 1 2 からの欠陥領域教師データとの差分（ずれ）が計算され、それらの欠陥領域の誤差として、学習モデル 1 1 3 にフィードバックされる。学習モデル 1 1 3 では、バックプロパゲーション（誤差逆伝播法）を利用して、減算部 1 1 4 からの誤差を減らすように、DNN の各ニューロンの重みなどが調整される。

[0040] すなわち、学習モデル 1 1 3 では、RGB 画像と視点変換デプス画像を入力したときに、その出力として欠陥領域を出力することを期待しているが、学習

の初期段階では、欠陥領域とは異なる領域を欠陥領域として出力してしまう。ここでは、学習モデル 1 1 3 から出力される欠陥領域と、欠陥領域教師データとの差分（ずれ）をフィードバックすることを繰り返すことで、学習が進んで行くにつれて、学習モデル 1 1 3 の出力である欠陥領域が、徐々に欠陥領域教師データとして出力されるようになり、学習モデル 1 1 3 の学習が収束することになる。

[0041] 学習モデル 1 1 3 における DNN の基本構造としては、例えば、FuseNet（下記の文献 1 に記載）等のセマンティックセグメンテーション (Semantic Segmentation) 用 DNN、又は SSD (Single Shot Multibox Detector) や YOLO (You Only Look Once) 等の物体検知の DNN などを用いることができる。

[0042] 図 3 は、センサフュージョン用の DNN の構造と出力の例として、セマンティックセグメンテーションによる二値分類画像を出力する例を示している。

[0043] 図 3 においては、図中の左側から RGB 画像と視点変換デプス画像を入力したときに、RGB 画像に対する畳み込み演算を行うことで段階的に得られる特徴量に、視点変換デプス画像に対する畳み込み演算を行うことで段階的に得られる特徴量を加算している。すなわち、RGB 画像と視点変換デプス画像について、畳み込み演算により特徴量（マトリックス）が段階的に求められ、フュージョン要素ごとに加算が行われる。

[0044] これにより、デプスセンサ 1 1 と RGB センサ 1 2 の 2 つのセンサ出力であるデプス画像（視点変換デプス画像）と RGB 画像とが合成され、セマンティックセグメンテーションの出力として二値分類画像が出力される。二値分類画像は、欠陥領域（欠陥画素の領域）とそれ以外の領域とを塗り分けたような画像とされる。例えば、二値分類画像においては、欠陥画素であるか否かに応じて欠陥画素を塗り潰すことができる。

[0045] なお、センサフュージョンにおけるセマンティックセグメンテーションに関する技術としては、例えば、下記の文献 1 に開示された技術がある。

[0046] 文献 1 : "FuseNet: Incorporating Depth into Semantic Segmentation via a Fusion-based CNN Architecture", Caner Hazirbas, Lingni Ma, Csaba Do

mokos, and Daniel Cremer <URL : <https://hazirbas.com/projects/fusenet/>>

[0047] 図4は、センサフュージョン用のDNNの構造と出力の例として、欠陥領域の座標などの数値データを出力する例を示している。

[0048] 図4においては、図3と同様に、図中の左側からRGB画像と視点変換デプス画像が入力され、各画像に対する畳み込み演算により特徴量（マトリックス）が段階的に求められ、フュージョン要素ごとに加算される。また、その後段には、SSD(Single Shot Multibox Detector)の構造が示されており、フュージョン要素ごとに加算するなどして得られた特徴量を入力することで、その出力として欠陥領域（欠陥画素）の座標が出力される。例えば、欠陥領域や欠陥画素の座標としては、矩形や点を表した座標（xy座標）が出力される。

[0049] なお、SSDを用いた物体検知に関する技術としては、例えば、下記の文献2に開示された技術がある。

[0050] 文献2：“SSD: single shot multibox detector”, W. Liu, D. Anguelov, D. Erhan, S. Christian, S. Reed, C.-Y. Fu, and A. C. Berg

[0051] このようにして学習時にDNNにより学習した学習モデル113を、学習済みモデルとして、推論時に用いることができる。図5は、教師あり学習を用いた場合における推論時の処理を行う処理部の構成例を示す図である。

[0052] 図5において、処理部10は、図1の処理部10に対応している。処理部10は、視点変換部121、及び学習モデル122を有する。ただし、学習モデル122は、学習時にDNNにより学習することで学習済みとなった学習モデル113（図2）に対応している。

[0053] 処理部10には、測定データとしてデプス画像とRGB画像が入力され、デプス画像が視点変換部121に、RGB画像が学習モデル122にそれぞれ供給される。

[0054] 視点変換部121は、撮影パラメータを用いて、入力されたデプス画像に対する視点変換処理を行い、その結果得られるRGBセンサ12の視点に対応し

た視点変換デプス画像を、学習モデル 1 2 2 に供給する。

[0055] 学習モデル 1 2 2 は、測定データとしての RGB 画像と、視点変換部 1 2 1 から供給される視点変換デプス画像を入力とした推論を行うことで、欠陥領域を出力する。すなわち、学習モデル 1 2 2 は、学習時に DNN により学習済みとなった学習モデル 1 1 3（図 2）に対応しており、RGB 画像と視点変換デプス画像を入力したとき、その出力である欠陥領域として、欠陥画素を塗り潰した二値分類画像や、欠陥領域や欠陥画素の座標（矩形や点を表した xy 座標）が出力される。

[0056] （B）教師なし学習

図 6 は、教師なし学習を用いた場合における学習時の処理を行う学習装置の構成例を示す図である。

[0057] 図 6 において、学習装置 2 は、視点変換部 1 3 1、学習モデル 1 3 2、及び減算部 1 3 3 を有する。

[0058] 学習装置 2 には、学習データとしてデプス画像と RGB 画像が入力され、デプス画像が視点変換部 1 3 1 に、RGB 画像が学習モデル 1 3 2 にそれぞれ供給される。ただし、ここで入力されるデプス画像は、欠陥がないデプス画像とされる。

[0059] 視点変換部 1 3 1 は、撮影パラメータを用いて、デプス画像に対する視点変換処理を行い、その結果得られる RGB センサ 1 2 の視点に対応した視点変換デプス画像を、学習モデル 1 3 2 及び減算部 1 3 3 に供給する。

[0060] 学習モデル 1 3 2 は、RGB 画像と視点変換デプス画像を入力とし、視点変換デプス画像を出力としたオートエンコーダ（Autoencoder）による機械学習を行うモデルである。オートエンコーダは、ニューラルネットワークの 1 つであって入力と出力の差分をとることで異常検知などに利用されるが、学習モデル 1 3 2 では、入力された視点変換デプス画像と同じフォーマットのデータである視点変換デプス画像が出力されるように調整されている。

[0061] また、減算部 1 3 3 では、学習モデル 1 3 2 の出力である視点変換デプス画像と、視点変換部 1 3 1 からの視点変換デプス画像との差分が計算され、

それらの視点変換デプス画像の誤差として、学習モデル 1 3 2 にフィードバックされる。例えば、視点変換デプス画像の差分としては、画像内の各画素についてz座標の値の差を用いることができる。学習モデル 1 3 2 では、バックプロパゲーションを利用して、減算部 1 3 3 からの誤差を減らすように、N Nの各ニューロンの重みなどが調整される。

[0062] すなわち、学習モデル 1 3 2 では、欠陥がない視点変換デプス画像を入力として、視点変換デプス画像を出力するとともに、入力した視点変換デプス画像と出力した視点変換デプス画像との差分がフィードバックされることが繰り返して行われる。これにより、学習モデル 1 3 2 では、欠陥があるデプス画像を知らずにオートエンコーダにより学習が行われるため、その出力として、欠陥が消えている視点変換デプス画像を出力するようになる。

[0063] 学習モデル 1 3 2 におけるオートエンコーダの基本構造としては、例えば、上記の文献 1 に記載のFuseNetを用いることができる。具体的には、上述した教師あり学習の場合、セマンティックセグメンテーションの出力として二値分類画像が出力されたとしたが、教師なし学習の場合には、セマンティックセグメンテーションの出力としてデプス画像（視点変換デプス画像）が出力されるようにすればよい。

[0064] このようにして学習時にオートエンコーダにより学習した学習モデル 1 3 2 を、学習済みモデルとして、推論時に用いることができる。図 7 は、教師なし学習を用いた場合における推論時の処理を行う処理部の構成例を示す図である。

[0065] 図 7 において、処理部 1 0 は、図 1 の処理部 1 0 に対応している。処理部 1 0 は、視点変換部 1 4 1、学習モデル 1 4 2、及び比較部 1 4 3 を有する。ただし、学習モデル 1 4 2 は、学習時にオートエンコーダにより学習することで学習済みとなった学習モデル 1 3 2（図 6）に対応している。

[0066] 処理部 1 0 には、測定データとしてデプス画像とRGB画像が入力され、デプス画像が視点変換部 1 4 1 に、RGB画像が学習モデル 1 4 2 にそれぞれ供給される。ただし、ここで入力されるデプス画像は、欠陥がある（欠陥がある可

能性がある) デプス画像とされる。

- [0067] 視点変換部 1 4 1 は、撮影パラメータを用いて、デプス画像に対する視点変換処理を行い、その結果得られるRGBセンサ 1 2 の視点に対応した視点変換デプス画像を、学習モデル 1 4 2 及び比較部 1 4 3 に供給する。
- [0068] 学習モデル 1 4 2 は、測定データとしてのRGB画像と、視点変換部 1 4 1 から供給される視点変換デプス画像を入力とした推論を行うことで、その出力としての視点変換デプス画像を比較部 1 4 3 に供給する。すなわち、学習モデル 1 4 2 は、学習時にオートエンコーダにより学習済みとなった学習モデル 1 3 2 (図 6) に対応しており、RGB画像と視点変換デプス画像を入力したとき、その出力である視点変換デプス画像として、欠陥が消えている視点変換デプス画像が出力される。
- [0069] 比較部 1 4 3 は、視点変換部 1 4 1 から供給される視点変換デプス画像と、学習モデル 1 4 2 から供給される視点変換デプス画像とを比較し、その比較結果を欠陥領域として出力する。すなわち、視点変換部 1 4 1 からの視点変換デプス画像は、欠陥がある可能性があるが、学習モデル 1 4 2 から出力された視点変換デプス画像は、欠陥がない(欠陥が消えている)ので、比較部 1 4 3 では、それらの視点変換デプス画像を比較することで、欠陥領域が得られる。
- [0070] 具体的には、例えば、比較対象となる 2 つの視点変換デプス画像内の各画素について、両者のZ座標の値(距離値)の比を計算し、計算した比が所定の閾値以上(又は閾値未満)となる画素を欠陥画素であるとみなすことができる。比較部 1 4 3 は、欠陥画素であるとみなした画素について、画像内のXY座標を、欠陥領域として出力することができる。
- [0071] 以上のように、第 1 の実施の形態では、機械学習により学習された学習済みモデルを用いて、デプス画像に含まれる欠陥領域(欠陥画素)を出力することができる。この欠陥領域に含まれる欠陥画素は、補正対象画素として、後段の処理で補正することができる。例えば、欠陥画素等の補正対象画素に対して、後述する補正処理(図 1 2)を適用することが可能である。あるいは

は、後段の処理では、欠陥画素等の補正対象画素を補正せずに、無効として無視しても構わない。なお、学習済みモデルの出力として、欠陥画素等の補正対象画素が補正されたデプス画像が出力されるようにしてもよい。

[0072] このように、欠陥画素等の補正対象画素を特定することで、補正対象画素を補正したり、無効として無視したりするなどの処理が可能となり、例えば、後続するデプス画像を用いた認識処理では、認識処理の精度を向上させることができる。

[0073] <2. 第2の実施の形態>

[0074] デプス画像に含まれる欠陥画素等の補正対象画素を特定する際には、GAN(Generative Adversarial Networks)等を用いて擬似的に生成したデプス画像を利用することができる。以下、GANにより生成したデプス画像を利用して、欠陥画素等の補正対象画素を特定する手法と、特定した補正対象画素を補正する手法について説明する。

[0075] (処理部の構成例)

図8は、推論時の処理を行う処理部の構成例を示す図である。

[0076] 図8において、処理部10は、図1の処理部10に対応している。処理部10は、特定部201、及び補正部202を有する。

[0077] 処理部10には、測定データとしてRGB画像とデプス画像が入力され、RGB画像とデプス画像が特定部201に、デプス画像が補正部202にそれぞれ供給される。

[0078] 特定部201は、入力されたRGB画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた推論を行い、入力されたデプス画像に含まれる欠陥領域(欠陥画素)を特定する。特定部201は、欠陥領域(欠陥画素)の特定結果を、補正部202に供給する。

[0079] 補正部202は、特定部201から供給される欠陥領域(欠陥画素)の特定結果に基づいて、入力されたデプス画像に含まれる欠陥領域(欠陥画素)を補正する。補正部202は、補正したデプス画像を出力する。

[0080] ここで、図9を参照して、特定部201の詳細な構成を説明する。図9に

において、特定部201は、学習モデル211、視点変換部212、及び比較部213を有する。

[0081] 特定部201には、測定データとしてRGB画像とデプス画像が入力され、RGB画像が学習モデル211に、デプス画像が視点変換部212にそれぞれ供給される。

[0082] 学習モデル211は、デプス画像と、デプス画像と対となるRGB画像との対応関係を、GAN等の機械学習により学習した学習済みモデルである。学習モデル211は、入力されたRGB画像からデプス画像を生成し、その出力として生成デプス画像を比較部213に供給する。ここでは、学習モデル211を利用して生成されたデプス画像を生成デプス画像と呼び、デプスセンサ11により取得されたデプス画像と区別する。

[0083] 視点変換部212は、撮影パラメータを用いて、デプス画像をRGBセンサ12の視点に変換する処理を行い、その結果得られる視点変換デプス画像を、比較部213に供給する。撮影パラメータとしては、例えば、デプスセンサ11とRGBセンサ12との相対位置や姿勢に関する情報などが用いられる。

[0084] 比較部213には、学習モデル211からの生成デプス画像と、視点変換部212からの視点変換デプス画像が供給される。比較部213は、生成デプス画像と視点変換デプス画像とを比較し、その比較結果が所定の条件を満たすとき、欠陥画素が検出されたとして、比較結果を出力する。

[0085] 例えば、比較部213では、生成デプス画像と視点変換デプス画像に対し、対応する画素ごとの輝度の差を求めて、その輝度差の絶対値が所定の閾値以上となるかどうかを判定し、輝度差の絶対値が所定の閾値以上となる画素を、欠陥候補の画素（欠陥画素）とすることができる。

[0086] また、上述した説明では、比較部213により生成デプス画像と視点変換デプス画像との比較をする際に、画素ごとに輝度の差をとって閾値判定を行う場合を示したが、輝度の差に限らず、例えば、画素ごとの輝度の比など、他の演算値を用いても構わない。

[0087] ここで、輝度差や輝度比が所定の閾値以上となる画素を、欠陥候補の画素

とする理由は次の通りである。すなわち、GAN等により学習した学習済みモデルを利用して生成された生成デプス画像が、期待通りに生成されている場合、生成デプス画像はデプス画像に似るため、輝度差や輝度比が大きい画素が欠陥のある画素であるとしている。

[0088] (GANによる画像生成)

図10は、GANにより学習することで学習済みとなった学習モデル211を利用して、RGB画像から生成デプス画像を生成する例を示している。

[0089] GANは、生成ネットワーク(Generator)と識別ネットワーク(Discriminator)と呼ばれる2つのネットワークを利用し、これらを互いに競わせることで、高精度な生成モデルを学習させている。

[0090] 例えば、生成ネットワークは、識別ネットワークを騙すような本物そっくりの学習サンプル(生成デプス画像)を適当なデータ(RGB画像)から生成する一方で、識別ネットワークは、与えられたサンプルが生成ネットワークにより生成されたものがあるのか本物そのものであるのかその正否を判定する。これらの2つのモデルを学習させることで、最終的に、生成ネットワークは、限りなく本物に類似したサンプル(生成デプス画像)を、適当なデータ(RGB画像)から生成できるようになる。

[0091] 学習モデル211においても、学習時に生成ネットワークと識別ネットワークの2つのネットワークを利用した機械学習を実施済みであり、推論時には、図10に示すように、RGB画像を入力することで、生成デプス画像を出力することができる。

[0092] なお、RGB画像からデプス画像を生成する技術としては、例えば、下記の文献3に開示された技術がある。

[0093] 文献3: "Depth Map Prediction from a Single Image using a Multi-Scale Deep Network", David Eigen, Christian Puhrsch, Rob Fergus

[0094] また、学習モデル211は、GANに限らず、VAE(Variational Autoencoder)等の他のニューラルネットワークにより機械学習を実施して、推論時に、入力されたRGB画像から生成デプス画像を生成するようにしても構わない。

[0095] (特定処理)

図 11 のフローチャートを参照して、特定部 201 による特定処理の流れを説明する。

[0096] ステップ S201 において、比較部 213 は、欠陥候補の画素の判定に用いる閾値 Th を設定する。

[0097] ステップ S202 において、比較部 213 は、学習モデル 211 から出力された生成デプス画像の画素 (i, j) での輝度値 p を取得する。また、ステップ S203 において、比較部 213 は、視点変換部 212 からの視点変換デプス画像の画素 (i, j) での輝度値 q を取得する。

[0098] ここでは、各画像内の画素の i 行 j 列を、画素 (i, j) と表記しており、生成デプス画像の画素 (i, j) と視点変換デプス画像の画素 (i, j) とは、それらの画像において対応する位置（同じ座標）に存在する画素を示している。

[0099] ステップ S204 において、比較部 213 は、輝度値 p と輝度値 q の差分の絶対値が閾値 Th 以上であるかどうかを判定する。つまり、下記の式 (1) の関係を満たすかどうか判定される。

$$[0100] \quad |p - q| \geq Th \cdots (1)$$

[0101] ステップ S204 において、比較部 213 は、輝度値 p と輝度値 q の差分の絶対値が閾値 Th 以上であると判定された場合、処理はステップ S205 に進められる。ステップ S205 において、比較部 213 は、比較対象となっている画素 (i, j) を欠陥候補として格納する。例えば、欠陥候補の画素に関する情報（例えば座標）は、画素補正位置情報として、メモリに保持しておくことができる。

[0102] 一方で、ステップ S204 において、輝度値 p と輝度値 q の差分の絶対値が閾値 Th 未満であると判定された場合、ステップ S205 の処理はスキップされる。

[0103] ステップ S206 においては、画像中の画素を全て探索したかどうか判定される。ステップ S206 において、画像中の画素を全て探索していないと判定された場合、処理はステップ S202 に戻り、それ以降の処理が繰り返される。

返される。

[0104] ステップS 2 0 2乃至S 2 0 6の処理が繰り返されることで、生成デプス画像と視点変換デプス画像における対応する画素の全てについて、輝度値の差分の閾値判定が行われ、画像内に含まれる欠陥候補の画素の全てが特定されてその情報が保持される。

[0105] ステップS 2 0 6において、画像中の画素を全て探索したと判定された場合に、一連の処理は終了する。

[0106] 以上、特定処理の流れを説明した。この特定処理では、デプス画像に含まれる画素から、欠陥候補の画素となる画素が全て特定される。

[0107] (補正処理)

図 1 2 のフローチャートを参照して、補正部 2 0 2 による補正処理の流れを説明する。なお、補正部 2 0 2 では、入力されたデプス画像から視点変換デプス画像が生成され、視点変換デプス画像に対して処理が実施される。視点変換デプス画像は、特定部 2 0 1 から供給されてもよい。

[0108] ステップS 2 3 1 において、補正部 2 0 2 は、欠陥画素を設定する。ここでは、図 1 1 のステップS 2 0 5 の処理で格納された欠陥候補の画素が、欠陥画素として設定される。例えば、欠陥画素を設定するに際しては、メモリに保持された画素補正位置情報を用いることができる。

[0109] ステップS 2 3 2 において、補正部 2 0 2 は、視点変換デプス画像における欠陥画素の周辺領域を設定する。例えば、欠陥画素を含む $N \times N$ の正方領域を周辺領域とすることができる。 N の値は、画素を単位として任意の値を設定可能であるが、例えば、 $N = 5$ とすることができる。なお、正方領域に限らず、矩形等の他の形状の領域を周辺領域としても構わない。

[0110] ステップS 2 3 3 において、補正部 2 0 2 は、視点変換デプス画像における欠陥画素の周辺領域の輝度を置き換える。周辺領域の輝度の置き換え方法としては、例えば、次の 2 つの方法のいずれかを用いることができる。

[0111] 第 1 に、測定データとしての視点変換デプス画像における周辺領域に含まれる画素のうち、欠陥画素を除いた画素の輝度値の中央値を算出し、算出し

た輝度値の中央値で、周辺領域の輝度値を置き換える方法である。ここでは、輝度値の中央値を用いることで、輝度の置き換えに際してノイズの影響を抑制することができるが、平均値などの他の統計量を用いても構わない。

[0112] 第2に、学習モデル211からの出力である生成デプス画像における周辺領域に対応する領域の輝度値で、周辺領域の輝度値を置き換える方法である。すなわち、生成デプス画像は、GAN等により学習した学習モデル211を利用して生成された擬似的なデプス画像であるため、欠陥のような不自然な領域がなく、そのため、周辺領域の輝度の置き換えに利用することができる。

[0113] ステップS234においては、欠陥画素を全て置き換えたかどうか判定される。ステップS234において、欠陥画素を全て置き換えていないと判定された場合、処理はステップS231に戻り、それ以降の処理が繰り返される。

[0114] ステップS231乃至234の処理が繰り返されることで、視点変換デプス画像に含まれる欠陥画素（を含む周辺領域）の全てが補正される。

[0115] ステップS234において、欠陥画素を全て置き換えた判定された場合に、一連の処理は終了する。

[0116] 以上、補正処理の流れを説明した。この補正処理では、補正対象画素としての欠陥画素が設定され、その周辺領域の輝度を置き換えることで、補正対象画素（を含む領域）が補正される。そして、補正対象画素が補正されたデプス画像（視点変換デプス画像）が出力される。

[0117] 以上のように、第2の実施の形態では、GAN等を利用して擬似的に生成したデプス画素を利用して、欠陥画素等の補正対象画素を特定して補正することができる。そのため、例えば、後続するデプス画像を用いた認識処理では、認識処理の精度を向上させることができる。

[0118] <3. 第3の実施の形態>

[0119] RGB画像とデプス画像からRGBD画像を生成する際に、デプス値（距離値）が割り当てられない場合や、デプス値が割り当てられても正しいデプス値が割り当てられていない場合がある。デプス値が割り当てられない要因としては

、視差による遮蔽や飽和、低反射率物体や透明物体が対象物となることなどが挙げられる。正しいデプス値が割り当てられていない要因としては、対象物がマルチパスである場合や、鏡面、半透明物体、高コントラスト模様などであることが挙げられる。

[0120] そのため、RGB画像とデプス画像から、欠陥がないRGBD画像を生成する手法が求められていた。以下、機械学習により学習された学習済みモデルを利用して、RGB画像とデプス画像から、欠陥がないRGBD画像を生成する手法について説明する。

[0121] (処理部の構成例)

図13は、推論時の処理を行う処理部の構成例を示す図である。

[0122] 図13において、処理部10は、図1の処理部10に対応している。処理部10は、画像生成部301を有する。

[0123] 処理部10には、測定データとしてRGB画像とデプス画像が入力され、画像生成部301に供給される。

[0124] 画像生成部301は、入力されたRGB画像とデプス画像から、RGBの色情報とデプス値(D値)による深度情報を持ったRGBD画像を生成する。RGBD画像の生成に際しては、デプス画像をRGB画像の画像面へ写像するか、あるいはRGB画像をデプス画像の画像面へ写像することでRGBD画像を生成することができる。例えば、図14に示すようなRGB画像とデプス画像とが合成され、RGBD画像が生成される。

[0125] 画像生成部301は、推論部311を有する。推論部311は、学習済みの学習モデルを用いて、デプス値に欠陥があるRGBD画像などを入力として推論を行い、欠陥を補正済みのRGBD画像などを出力する。以下、推論部311で用いられる学習モデルとして、学習時に、教師あり学習により学習した場合と、教師なし学習により学習した場合について説明する。

[0126] (A) 教師あり学習

図15は、教師あり学習を用いた場合における学習時の処理を行う学習装置と、推論時の処理を行う推論部の構成例を示す図である。

- [0127] 図15においては、上段に、学習時の処理を行う学習装置2を示し、下段に、推論時の処理を行う推論部311が示されている。推論部311は、図13の推論部311に対応している。
- [0128] 図15において、学習装置2は、学習モデル321を有する。学習モデル321は、デプス値に欠陥があるRGBD画像と欠陥画素の位置を示した画素位置情報（欠陥画素位置情報）を入力とし、RGBD画像を出力するニューラルネットワークによる機械学習を行うモデルである。例えば、学習モデル321では、デプス値に欠陥があるRGBD画像と欠陥画素位置情報を学習データとし、欠陥画素位置（を含む領域）の補正に関する情報を教師データとした学習を繰り返すことで、その出力として欠陥を補正済みのRGBD画像を出力することができるようになる。ニューラルネットワークとしては、例えば、オートエンコーダやDNNなどを用いることができる。
- [0129] このようにして学習時に機械学習により学習した学習モデル321を、学習済みモデルとして推論時に用いることができる。
- [0130] 図15において、推論部311は、学習モデル331を有する。学習モデル331は、学習時に機械学習により学習することで学習済みとなった学習モデル321に対応している。
- [0131] 学習モデル331は、デプス値に欠陥があるRGBD画像と欠陥画素位置情報を入力とした推論を行うことで、欠陥を補正済みのRGBD画像を出力する。ここで、デプス値に欠陥があるRGBD画像は、測定データとしてのRGB画像とデプス画像から生成されたRGBD画像である。欠陥画素位置情報は、測定データとしてのRGB画像とデプス画像から特定される欠陥画素の位置に関する情報である。
- [0132] なお、教師あり学習として、他の機械学習を行ってもよい。例えば、学習時において、学習モデル321の出力として欠陥を補正済みの画素位置に関する情報が出力されるように学習することで、推論時において、学習モデル331では、デプス値に欠陥があるRGBD画像と欠陥画素位置情報を入力とした推論を行い、欠陥を補正済みの画素位置に関する情報を出力するようにし

てもよい。

[0133] (B) 教師なし学習

図 1 6 は、教師なし学習を用いた場合における学習時の処理を行う学習装置と、推論時の処理を行う推論部の構成例を示す図である。

[0134] 図 1 6 においては、上段に、学習時の処理を行う学習装置 2 を示し、下段に、推論時の処理を行う推論部 3 1 1 が示されている。推論部 3 1 1 は、図 1 3 の推論部 3 1 1 に対応している。

[0135] 図 1 6 において、学習装置 2 は、学習モデル 3 4 1 を有する。学習モデル 3 4 1 は、欠陥がない RGBD 画像を入力としてニューラルネットワークにより機械学習を行うモデルである。すなわち、学習モデル 3 4 1 は、欠陥がある RGBD 画像を知らずにニューラルネットワークにより教師なしの学習を繰り返すため、その出力として、欠陥が消えている RGBD 画像を出力するようになる。

[0136] このようにして学習時に機械学習により教師なしの学習をした学習モデル 3 4 1 を、学習済みモデルとして推論時に用いることができる。

[0137] 図 1 6 において、推論部 3 1 1 は、学習モデル 3 5 1 を有する。学習モデル 3 5 1 は、学習時に機械学習により教師なしの学習をすることで学習済みとなった学習モデル 3 4 1 に対応している。

[0138] 学習モデル 3 5 1 は、デプス値に欠陥がある RGBD 画像を入力とした推論を行うことで、欠陥を補正済みの RGBD 画像を出力する。ここで、デプス値に欠陥がある RGBD 画像は、測定データとしての RGB 画像とデプス画像から生成された RGBD 画像である。

[0139] (画像生成処理)

次に、図 1 7 のフローチャートを参照して、画像生成部 3 0 1 による画像生成処理の第 1 の例の流れを説明する。第 1 の例では、デプス画像を RGB 画像の画像面へ写像することで RGBD 画像を生成する場合における画像生成処理の流れを示している。

[0140] ステップ S 3 0 1 において、画像生成部 3 0 1 は、デプス画像に含まれる D 画素を全て処理したかどうかを判定する。ここでは、デプス画像に含まれる

画素をD画素と呼んでいる。

- [0141] ステップS301において、D画素を全て処理していないと判定された場合、処理はステップS302に進められる。ステップS302において、画像生成部301は、処理対象のD画素についてデプス値と画素位置(x, y)を取得する。
- [0142] ステップS303において、画像生成部301は、取得した処理対象のD画素のデプス値が有効なデプス値であるかどうかを判定する。
- [0143] ステップS303において、処理対象のD画素のデプス値が有効なデプス値であると判定された場合、処理はステップS304に進められる。ステップS304において、画像生成部301は、画素位置(x, y)とデプス値に基づいて、RGB画像における写像先位置(x', y')を取得する。
- [0144] ステップS305において、画像生成部301は、写像先位置(x', y')にデプス値がまだ割り当てられていないかどうかを判定する。ここで、1つの写像先位置(x', y')に対し、複数のデプス値が割り当てられる場合があるので、ステップS305では、写像先位置(x', y')にデプス値が既に割り当てられている場合には、割り当てようとしているデプス値が、既に割り当て済みのデプス値よりも小さいかどうかをさらに判定する。
- [0145] ステップS305において、デプス値がまだ割り当てられていないと判定された場合、又はデプス値が既に割り当てられている場合に割り当てようとしているデプス値が既に割り当て済みのデプス値よりも小さいとき、処理はステップS306に進められる。ステップS306において、画像生成部301は、写像先位置(x', y')にデプス値を割り当てる。
- [0146] ステップS306の処理が終了すると、処理はステップS301に戻る。また、ステップS303において、処理対象のD画素のデプス値が有効なデプス値ではないと判定された場合、又は、ステップS305において、デプス値が既に割り当てられているが、割り当てようとしているデプス値が既に割り当て済みのデプス値よりも大きい場合、処理はステップS301に戻る。
- [0147] デプス画像に含まれるD画素を順次、処理対象のD画素として、当該D画素の

画素位置(x, y)のデプス値が有効であって、対応する写像先位置(x', y')にデプス値が割り当てられていない場合、又はデプス値が既に割り当てられている場合に割り当てようとしているデプス値が既に割り当て済みのデプス値よりも小さいときには、写像先位置(x', y')にデプス値が割り当てられる。

[0148] 上述した処理が繰り返されて、ステップS 3 0 1において、D画素を全て処理したと判定された場合、処理はステップS 3 0 7に進められる。すなわち、D画素を全て処理したときに、デプス画像をRGB画像の画像面へ写像することが完了してRGBD画像が生成されるが、このRGBD画像は、欠陥があるRGBD画像（不完全なRGBD画像）である可能性があるため、ステップS 3 0 7以降の処理が行われる。

[0149] ステップS 3 0 7において、画像生成部3 0 1は、デプス値が割り当てられていないRGB画素があるかどうかを判定する。ここでは、RGB画像に含まれる画素をRGB画素と呼んでいる。

[0150] ステップS 3 0 7において、デプス値が割り当てられていないRGB画素があると判定された場合、処理はステップS 3 0 8に進められる。

[0151] ステップS 3 0 8において、画像生成部3 0 1は、デプス値が割り当てられていないRGB画素の位置に関する情報に基づいて、画素補正位置情報を生成する。この画素補正位置情報は、デプス値が割り当てられていないRGB画素を、補正する必要がある画素（欠陥画素）であるとして、その画素位置を特定する情報（例えば欠陥画素の座標）を含む。

[0152] ステップS 3 0 9において、推論部3 1 1は、学習モデル3 3 1（図1 5）を用いて、欠陥があるRGBD画像と画素補正位置情報を入力として推論を行い、欠陥を補正済みのRGBD画像を生成する。学習モデル3 3 1は、学習時に、デプス値に欠陥があるRGBD画像と欠陥画素位置情報を入力としてニューラルネットワークによる学習を行った学習済みモデルであって、欠陥を補正済みのRGBD画像を出力することができる。つまり、欠陥を補正済みのRGBD画像では、RGB画像における画素補正位置のデプス値が推論されたことで、欠陥が補正されている。

- [0153] なお、ここでは、学習モデル 331 を用いた場合を示したが、デプス値に欠陥があるRGBD画像を入力とした推論を行うことで欠陥を補正済みのRGBD画像を出力する学習モデル 351（図 16）などの他の学習済みモデルを用いても構わない。
- [0154] ステップ S309 の処理が終了すると、一連の処理は終了する。また、ステップ S307 において、デプス値が割り当てられていないRGB画素がないと判定された場合には、欠陥がないRGBD画像（完全なRGBD画像）が生成されて補正する必要がないため、ステップ S308，S309 の処理がスキップされ、一連の処理は終了する。
- [0155] 以上、画像生成処理の第 1 の例の流れを説明した。この画像生成処理では、デプスセンサ 11 により取得されたデプス画像を、RGBセンサ 12 により取得されたRGB画像の画像面に写像してRGBD画像を生成するに際して、次のような処理が行われる。すなわち、デプス画像の各画素に応じた画素位置(x, y)のデプス値に基づいて、位置(x, y)をRGB画像の画像面に写像し、RGB画像の各画素に応じた写像先位置(x', y')のうち、画素位置(x, y)のデプス値が割り当てられていない写像先位置(x', y')を画素補正位置として特定し、学習モデルを用いて、RGB画像における画素補正位置のデプス値を推論することで、補正済みのRGBD画像を生成している。
- [0156] 次に、図 18 のフローチャートを参照して、画像生成部 301 による画像生成処理の第 2 の例の流れを説明する。第 2 の例では、RGB画像をデプス画像の画像面へ写像することでRGBD画像を生成する場合における画像生成処理の流れを示している。
- [0157] ステップ S331 において、画像生成部 301 は、デプス画像に含まれるD画素を全て処理したかどうかを判定する。
- [0158] ステップ S331 において、D画素を全て処理していないと判定された場合、処理はステップ S332 に進められる。ステップ S332 において、画像生成部 301 は、処理対象のD画素についてデプス値と画素位置(x, y)を取得する。

- [0159] ステップS 3 3 3において、画像生成部3 0 1は、取得した処理対象のD画素のデプス値が有効なデプス値であるかどうかを判定する。
- [0160] ステップS 3 3 3において、処理対象のD画素のデプス値が有効なデプス値ではないと判定された場合、処理はステップS 3 3 4に進められる。
- [0161] ステップS 3 3 4において、推論部3 1 1は、学習モデルを用いて、欠陥があるデプス画像と画素補正位置情報を入力として推論を行い、補正済みのデプス値を生成する。ここで用いられる学習モデルは、学習時に、デプス値に欠陥があるデプス画像と画素補正位置情報を入力としてニューラルネットワークによる学習を行った学習済みモデルであって、補正済みのデプス値を出力することができる。なお、補正済みのデプス値を生成可能であれば、他のニューラルネットワークにより学習した学習済みモデルを用いても構わない。
- [0162] ステップS 3 3 4の処理が終了すると、処理はステップS 3 3 5に進められる。また、ステップS 3 3 3において、処理対象のD画素のデプス値が有効なデプス値であると判定された場合、ステップS 3 3 4の処理がスキップされ、処理はステップS 3 3 5に進められる。
- [0163] ステップS 3 3 5において、画像生成部3 0 1は、デプス値と撮影パラメータに基づいて、RGB画像におけるサンプリング位置(x' , y')を計算する。撮影パラメータとしては、例えば、デプスセンサ1 1とRGBセンサ1 2との相対位置や姿勢に関する情報などが用いられる。
- [0164] ステップS 3 3 6において、画像生成部3 0 1は、RGB画像のサンプリング位置(x' , y')からRGB値をサンプリングする。
- [0165] ステップS 3 3 6の処理が終了すると、処理はステップS 3 3 1に戻り、上述した処理が繰り返される。すなわち、デプス画像に含まれるD画素を順次、処理対象のD画素として、当該D画素の画素位置(x , y)のデプス値が有効でない場合には、学習モデルを利用して補正済みのデプス値を生成することで、処理対象のD画素のデプス値に対応したサンプリング位置(x' , y')が計算され、RGB画像からRGB値がサンプリングされる。

[0166] 上述した処理が繰り返されて、ステップS 3 3 1において、D画素を全て処理したと判定された場合、RGB画像をデプス画像の画像面へ写像することが完了してRGBD画像が生成されるため、一連の処理は終了する。

[0167] 以上、画像生成処理の第2の例の流れを説明した。この画像生成処理では、RGBセンサ1 2により取得されたRGB画像を、デプスセンサ1 1により取得されたデプス画像の画像面に写像してRGBD画像を生成するに際して、次のような処理が行われる。すなわち、デプス画像の各画素に応じた画素位置(x, y)のうち、有効なデプス値が割り当てられていない画素位置(x, y)を画素補正位置として特定し、学習モデルを用いて、デプス画像における画素補正位置のデプス値を推論し、画素位置(x, y)に割り当てられたデプス値に基づいて、RGB画像におけるサンプリング位置(x', y')からRGB値をサンプリングして、サンプリング位置(x', y')をデプス画像の画像面に写像することで、補正済みのRGBD画像を生成している。

[0168] (ユースケースの例)

図1 9乃至図2 1は、本開示を適用可能なユースケースの例を示している。

[0169] 図1 9は、ユースケースの第1の例を示す図である。図1 9において、人物を被写体としたポートレートやテレビ会議などのRGBD画像3 6 1に遮蔽領域3 6 2が含まれる場合、遮蔽領域3 6 2ではデプス値を求めることが困難となるため、人物の部分を残して背景を除去するときに、遮蔽領域3 6 2が背景に映り込んでしまう恐れがある。

[0170] 本開示に係る技術では、RGBD画像を生成する画像生成部3 0 1において、推論部3 1 1が学習済みモデル（学習モデル）を用いて、デプス値に欠陥があるRGBD画像を入力として、欠陥（遮蔽領域3 6 2の部分）を補正済みのRGBD画像を出力するため、このような事象を回避することができる。

[0171] 図2 0は、ユースケースの第2の例を示す図である。図2 0において、工事現場の作業員をセンシングして得られたRGBD画像3 7 1に、当該作業員が身に着けている反射ベスト3 7 2が含まれる場合、反射ベスト3 7 2は再帰

性反射素材からなり、光を強く反射するため、光源からの光を出射するデプスセンサ 11 では飽和してしまい、測距を行うことが難しい。また、自動運転の車両でセンシングを行う場合に、強い反射率を有する再帰性反射素材からなる道路標識 373 などがRGBD画像 371 に含まれるときにも、デプスセンサ 11 で測距を行うことは困難である。

[0172] 本開示に係る技術では、RGBD画像を生成する画像生成部 301 において、推論部 311 が学習済みモデルを用いて、デプス値に欠陥があるRGBD画像を入力として、欠陥（反射ベスト 372 や道路標識 373 の部分）を補正済みのRGBD画像を出力するため、このような事象を回避することができる。

[0173] 図 21 は、ユースケースの第 3 の例を示す図である。例えば、建物測量や 3D のAR(Augmented Reality)ゲーム等のアプリケーションにおいては、部屋の中を 3D スキャンしたいケースがある。図 21 において、部屋の中をセンシングして得られたRGBD画像 381 に、透明な窓 382、高周波の模様 383、鏡や鏡面 384、壁の角 385 などが含まれる場合、デプス値がとれないことや、デプス値が誤ってしまう恐れがある。

[0174] 本開示に係る技術では、RGBD画像を生成する画像生成部 301 において、推論部 311 が学習済みモデルを用いて、デプス値に欠陥があるRGBD画像を入力として、欠陥（透明な窓 382、高周波の模様 383、鏡や鏡面 384、壁の角 385 などの部分）を補正済みのRGBD画像を出力することができる。そのため、建物測量や 3D のARゲーム等のアプリケーションにおいては、本開示に係る技術を適用して部屋の中を 3D スキャンすることで、それらのアプリケーションが期待する動作が行われるようになる。

[0175] <4. 変形例>

[0176] 図 22 は、AI 処理を行う装置を含むシステムの構成例を示している。

[0177] 電子機器 20001 は、スマートフォン、タブレット型端末、携帯電話機等のモバイル端末である。電子機器 20001 は、例えば、図 1 の情報処理装置 1 に対応しており、デプスセンサ 11（図 1）に対応した光センサ 20011 を有する。光センサは、光を電気信号に変換するセンサ（画像センサ

）である。電子機器 20001 は、所定の通信方式に対応した無線通信によって所定の場所に設置された基地局 20020 に接続することで、コアネットワーク 20030 を介して、インターネット等のネットワーク 20040 に接続することができる。

[0178] 基地局 20020 とコアネットワーク 20030 の間などのモバイル端末により近い位置には、モバイルエッジコンピューティング（MEC : Mobile Edge Computing）を実現するためのエッジサーバ 20002 が設けられる。ネットワーク 20040 には、クラウドサーバ 20003 が接続される。エッジサーバ 20002 とクラウドサーバ 20003 は、用途に応じた各種の処理を行うことができる。なお、エッジサーバ 20002 は、コアネットワーク 20030 内に設けられてもよい。

[0179] 電子機器 20001、エッジサーバ 20002、クラウドサーバ 20003、又は光センサ 20011 により、AI 処理が行われる。AI 処理は、本開示に係る技術を、機械学習等の AI を利用して処理するものである。AI 処理は、学習処理と推論処理を含む。学習処理は、学習モデルを生成する処理である。また、学習処理には、後述する再学習処理も含まれる。推論処理は、学習モデルを用いた推論を行う処理である。

[0180] 電子機器 20001、エッジサーバ 20002、クラウドサーバ 20003、又は光センサ 20011 においては、CPU (Central Processing Unit) 等のプロセッサがプログラムを実行したり、あるいは特定用途に特化したプロセッサ等の専用のハードウェアを用いたりすることで、AI 処理が実現される。例えば、特定用途に特化したプロセッサとしては、GPU (Graphics Processing Unit) を用いることができる。

[0181] 図 23 は、電子機器 20001 の構成例を示している。電子機器 20001 は、各部の動作の制御や各種の処理を行う CPU 20101 と、画像処理や並列処理に特化した GPU 20102 と、DRAM (Dynamic Random Access Memory) 等のメインメモリ 20103 と、フラッシュメモリ等の補助メモリ 20104 を有する。

- [0182] 補助メモリ20104は、AI処理用のプログラムや各種パラメータ等のデータを記録している。CPU20101は、補助メモリ20104に記録されたプログラムやパラメータをメインメモリ20103に展開してプログラムを実行する。あるいは、CPU20101とGPU20102は、補助メモリ20104に記録されたプログラムやパラメータをメインメモリ20103に展開してプログラムを実行する。これにより、GPU20102を、GPGPU(General-Purpose computing on Graphics Processing Units)として用いることができる。
- [0183] なお、CPU20101やGPU20102は、SoC(System on a Chip)として構成されてもよい。CPU20101がAI処理用のプログラムを実行する場合には、GPU20102を設けなくてもよい。
- [0184] 電子機器20001はまた、本開示に係る技術を適用した光センサ20011と、物理的なボタンやタッチパネル等の操作部20105と、少なくとも1以上のセンサを含むセンサ20106と、画像やテキスト等の情報を表示するディスプレイ20107と、音を出力するスピーカ20108と、所定の通信方式に対応した通信モジュール等の通信I/F20109と、それらを接続するバス20110を有する。
- [0185] センサ20106は、光センサ(画像センサ)、音センサ(マイクロフォン)、振動センサ、加速度センサ、角速度センサ、圧力センサ、匂いセンサ、生体センサ等の各種のセンサを少なくとも1以上有している。AI処理では、光センサ20011から取得したデータ(画像データ)とともに、センサ20106の少なくとも1以上のセンサから取得したデータを用いることができる。すなわち、光センサ20011は、デプスセンサ11(図1)に対応し、センサ20106は、RGBセンサ12(図1)に対応している。
- [0186] なお、センサフュージョンの技術によって2以上の光センサから取得したデータやそれらを統合的に処理して得られるデータが、AI処理で用いられてもよい。2以上の光センサとしては、光センサ20011とセンサ20106内の光センサの組み合わせでもよいし、あるいは光センサ20011内

に複数の光センサが含まれていてもよい。例えば、光センサには、RGBの可視光センサ、ToF (Time of Flight) 等の測距センサ、偏光センサ、イベントベースのセンサ、IR像を取得するセンサ、多波長取得可能なセンサなどが含まれる。

[0187] 電子機器20001においては、CPU20101やGPU20102等のプロセッサによってAI処理を行うことができる。電子機器20001のプロセッサが推論処理を行う場合には、光センサ20011で画像データを取得した後に時間を要せずに処理を開始することができるため、高速に処理を行うことができる。そのため、電子機器20001では、短い遅延時間で情報を伝達することが求められるアプリケーションなどの用途に推論処理が用いられた際に、ユーザは遅延による違和感なく操作を行うことができる。また、電子機器20001のプロセッサがAI処理を行う場合、クラウドサーバ20003等のサーバを利用する場合と比べて、通信回線やサーバ用のコンピュータ機器などを利用する必要がなく、低コストで処理を実現することができる。

[0188] 図24は、エッジサーバ20002の構成例を示している。エッジサーバ20002は、各部の動作の制御や各種の処理を行うCPU20201と、画像処理や並列処理に特化したGPU20202を有する。エッジサーバ20002はさらに、DRAM等のメインメモリ20203と、HDD(Hard Disk Drive)やSSD(Solid State Drive)等の補助メモリ20204と、NIC(Network Interface Card)等の通信I/F20205を有し、それらがバス20206に接続される。

[0189] 補助メモリ20204は、AI処理用のプログラムや各種パラメータ等のデータを記録している。CPU20201は、補助メモリ20204に記録されたプログラムやパラメータをメインメモリ20203に展開してプログラムを実行する。あるいは、CPU20201とGPU20202は、補助メモリ20204に記録されたプログラムやパラメータをメインメモリ20203に展開してプログラムを実行することで、GPU20202をGPG

PUとして用いることができる。なお、CPU 20201がAI処理用のプログラムを実行する場合には、GPU 20202を設けなくてもよい。

[0190] エッジサーバ20002においては、CPU 20201やGPU 20202等のプロセッサによってAI処理を行うことができる。エッジサーバ20002のプロセッサがAI処理を行う場合、エッジサーバ20002はクラウドサーバ20003と比べて、電子機器20001に近い位置に設けられるため、処理の低遅延化を実現することができる。また、エッジサーバ20002は、電子機器20001や光センサ20011に比べて、演算速度などの処理能力が高いため、汎用的に構成することができる。そのため、エッジサーバ20002のプロセッサがAI処理を行う場合、電子機器20001や光センサ20011の仕様や性能の違いに依らず、データを受信できればAI処理を行うことができる。エッジサーバ20002でAI処理を行う場合には、電子機器20001や光センサ20011における処理の負荷を軽減することができる。

[0191] クラウドサーバ20003の構成は、エッジサーバ20002の構成と同様であるため、説明は省略する。

[0192] クラウドサーバ20003においては、CPU 20201やGPU 20202等のプロセッサによってAI処理を行うことができる。クラウドサーバ20003は、電子機器20001や光センサ20011に比べて、演算速度などの処理能力が高いため、汎用的に構成することができる。そのため、クラウドサーバ20003のプロセッサがAI処理を行う場合、電子機器20001や光センサ20011の仕様や性能の違いに依らず、AI処理を行うことができる。また、電子機器20001又は光センサ20011のプロセッサで負荷の高いAI処理を行うことが困難である場合には、その負荷の高いAI処理をクラウドサーバ20003のプロセッサが行い、その処理結果を電子機器20001又は光センサ20011のプロセッサにフィードバックすることができる。

[0193] 図25は、光センサ20011の構成例を示している。光センサ2001

1 は、例えば複数の基板が積層された積層構造を有する 1 チップの半導体装置として構成することができる。光センサ 20011 は、基板 20301 と基板 20302 の 2 枚の基板が積層されて構成される。なお、光センサ 20011 の構成としては積層構造に限らず、例えば、撮像部を含む基板が、CPU や DSP (Digital Signal Processor) 等の AI 処理を行うプロセッサを含んでいてもよい。

[0194] 上層の基板 20301 には、複数の画素が 2 次元に並んで構成される撮像部 20321 が搭載されている。下層の基板 20302 には、撮像部 20321 での画像の撮像に関する処理を行う撮像処理部 20322 と、撮像画像や信号処理結果を外部に出力する出力 I/F 20323 と、撮像部 20321 での画像の撮像を制御する撮像制御部 20324 が搭載されている。撮像部 20321、撮像処理部 20322、出力 I/F 20323、及び撮像制御部 20324 により撮像ブロック 20311 が構成される。

[0195] 下層の基板 20302 には、各部の制御や各種の処理を行う CPU 20331 と、撮像画像や外部からの情報等を用いた信号処理を行う DSP 20332 と、SRAM (Static Random Access Memory) や DRAM (Dynamic Random Access Memory) 等のメモリ 20333 と、外部と必要な情報のやり取りを行う通信 I/F 20334 が搭載されている。CPU 20331、DSP 20332、メモリ 20333、及び通信 I/F 20334 により信号処理ブロック 20312 が構成される。CPU 20331 及び DSP 20332 の少なくとも 1 つのプロセッサによって AI 処理を行うことができる。

[0196] このように、複数の基板が積層された積層構造における下層の基板 20302 に、AI 処理用の信号処理ブロック 20312 を搭載することができる。これにより、上層の基板 20301 に搭載される撮像用の撮像ブロック 20311 で取得される画像データが、下層の基板 20302 に搭載された AI 処理用の信号処理ブロック 20312 で処理されるため、1 チップの半導体装置内で一連の処理を行うことができる。

[0197] 光センサ 20011 においては、CPU 20331 等のプロセッサによっ

てA I処理を行うことができる。光センサ20011のプロセッサが推論処理等のA I処理を行う場合、1チップの半導体装置内で一連の処理が行われるため、センサ外部に情報が漏れないことから情報の秘匿性を高めることができる。また、画像データ等のデータを他の装置に送信する必要がないため、光センサ20011のプロセッサでは、画像データを用いた推論処理等のA I処理を高速に行うことができる。例えば、リアルタイム性が求められるアプリケーションなどの用途に推論処理が用いられた際に、リアルタイム性を十分に確保することができる。ここで、リアルタイム性を確保するということは、短い遅延時間で情報を伝達できることを指す。さらに、光センサ20011のプロセッサがA I処理を行うに際して、電子機器20001のプロセッサにより各種のメタデータを渡すことで、処理を削減して低消費電力化を図ることができる。

[0198] 図26は、処理部20401の構成例を示している。処理部20401は、図1の処理部10に対応している。電子機器20001、エッジサーバ20002、クラウドサーバ20003、又は光センサ20011のプロセッサがプログラムに従った各種の処理を実行することで、処理部20401として機能する。なお、同一の又は異なる装置が有する複数のプロセッサを処理部20401として機能させてもよい。

[0199] 処理部20401は、A I処理部20411を有する。A I処理部20411は、A I処理を行う。A I処理部20411は、学習部20421と推論部20422を有する。

[0200] 学習部20421は、学習モデルを生成する学習処理を行う。学習処理では、画像データに含まれる補正対象画素を補正するための機械学習を行った機械学習済みの学習モデルが生成される。また、学習部20421は、生成済みの学習モデルを更新する再学習処理を行ってもよい。以下の説明では、学習モデルの生成と更新を区別して説明するが、学習モデルを更新することで、学習モデルを生成しているとも言えるため、学習モデルの生成には、学習モデルの更新の意味が含まれるものとする。

[0201] また、生成された学習モデルは、電子機器20001、エッジサーバ20002、クラウドサーバ20003、又は光センサ20011などが有するメインメモリ又は補助メモリなどの記憶媒体に記録されることで、推論部20422が行う推論処理において新たに利用可能となる。これにより、当該学習モデルに基づく推論処理を行う電子機器20001、エッジサーバ20002、クラウドサーバ20003、又は光センサ20011などを生成することができる。さらに、生成された学習モデルは、電子機器20001、エッジサーバ20002、クラウドサーバ20003、又は光センサ20011などとは独立した記憶媒体又は電子機器に記録され、他の装置で使用するために提供されてもよい。なお、これらの電子機器20001、エッジサーバ20002、クラウドサーバ20003、又は光センサ20011などの生成とは、製造時において、それらの記憶媒体に新たに学習モデルを記録することだけでなく、既に記録されている生成済学習モデルを更新することにも含まれるものとする。

[0202] 推論部20422は、学習モデルを用いた推論処理を行う。推論処理では、学習モデルを用いて、画像データに含まれる補正対象画素を特定したり、特定した補正対象画素を補正したりする処理が行われる。補正対象画素は、画像データに応じた画像内の複数個の画素のうち、所定の条件を満たした補正対象となる画素である。

[0203] 機械学習の手法としては、ニューラルネットワークやディープラーニングなどを用いることができる。ニューラルネットワークとは、人間の脳神経回路を模倣したモデルであって、入力層、中間層（隠れ層）、出力層の3種類の層からなる。ディープラーニングとは、多層構造のニューラルネットワークを用いたモデルであって、各層で特徴的な学習を繰り返し、大量データの中に潜んでいる複雑なパターンを学習することができる。

[0204] 機械学習の問題設定としては、教師あり学習を用いることができる。例えば、教師あり学習は、与えられたラベル付きの教師データに基づいて特徴量を学習する。これにより、未知のデータのラベルを導くことが可能となる。

学習データは、実際に光センサにより取得された画像データや、集約して管理されている取得済みの画像データ、シミュレータにより生成されたデータセットなどを用いることができる。

[0205] なお、教師あり学習に限らず、教師なし学習、半教師あり学習、強化学習などを用いてもよい。教師なし学習は、ラベルが付いていない学習データを大量に分析して特徴量を抽出し、抽出した特徴量に基づいてクラスタリング等を行う。これにより、膨大な未知のデータに基づいて傾向の分析や予測を行うことが可能となる。半教師あり学習は、教師あり学習と教師なし学習を混在させたものであって、教師あり学習で特徴量を学ばせた後、教師なし学習で膨大な学習データを与え、自動的に特徴量を算出させながら繰り返し学習を行う方法である。強化学習は、ある環境内におけるエージェントが現在の状態を観測して取るべき行動を決定する問題を扱うものである。

[0206] このように、電子機器 20001、エッジサーバ 20002、クラウドサーバ 20003、又は光センサ 20011 のプロセッサが A1 処理部 20411 として機能することで、それらの装置のいずれか又は複数の装置で A1 処理が行われる。

[0207] A1 処理部 20411 は、学習部 20421 及び推論部 20422 のうち少なくとも一方を有していればよい。すなわち、各装置のプロセッサは、学習処理と推論処理の両方の処理を実行することは勿論、学習処理と推論処理のうちの一方の処理を実行するようにしてもよい。例えば、電子機器 20001 のプロセッサが推論処理と学習処理の両方を行う場合には、学習部 20421 と推論部 20422 を有するが、推論処理のみを行う場合には、推論部 20422 のみを有していればよい。

[0208] 各装置のプロセッサは、学習処理又は推論処理に関する全ての処理を実行してもよいし、一部の処理を各装置のプロセッサで実行した後に、残りの処理を他の装置のプロセッサで実行してもよい。また、各装置においては、学習処理や推論処理などの A1 処理の各々の機能を実行するための共通のプロセッサを有してもよいし、機能ごとに個別にプロセッサを有してもよい。

- [0209] なお、上述した装置以外の他の装置によりA I処理が行われてもよい。例えば、電子機器20001が無線通信などにより接続可能な他の電子機器によって、A I処理を行うことができる。具体的には、電子機器20001がスマートフォンである場合に、A I処理を行う他の電子機器としては、他のスマートフォン、タブレット型端末、携帯電話機、P C(Personal Computer)、ゲーム機、テレビ受像機、ウェアラブル端末、デジタルスチルカメラ、デジタルビデオカメラなどの装置とすることができる。
- [0210] また、自動車等の移動体に搭載されるセンサや、遠隔医療機器に用いられるセンサなどを用いた構成においても、推論処理等のA I処理を適用可能であるが、それらの環境では遅延時間が短いことが求められる。このような環境においては、ネットワーク20040を介してクラウドサーバ20003のプロセッサでA I処理を行うのではなく、ローカル側の装置（例えば車載機器や医療機器としての電子機器20001）のプロセッサでA I処理を行うことで遅延時間を短くすることができる。さらに、インターネット等のネットワーク20040に接続する環境がない場合や、高速な接続を行うことができない環境で利用する装置の場合にも、例えば電子機器20001や光センサ20011等のローカル側の装置のプロセッサでA I処理を行うことで、より適切な環境でA I処理を行うことができる。
- [0211] なお、上述した構成は一例であって、他の構成を採用しても構わない。例えば、電子機器20001は、スマートフォン等のモバイル端末に限らず、P C、ゲーム機、テレビ受像機、ウェアラブル端末、デジタルスチルカメラ、デジタルビデオカメラなどの電子機器、車載機器、医療機器であってもよい。また、電子機器20001は、無線LAN(Local Area Network)や有線LANなどの所定の通信方式に対応した無線通信又は有線通信によってネットワーク20040に接続してもよい。A I処理は、各装置のC P UやG P U等のプロセッサに限らず、量子コンピュータやニューロモーフィック・コンピュータなどを利用して構わない。
- [0212] ところで、学習モデルや画像データ、補正済みデータ等のデータは、単一

の装置内で用いられることは勿論、複数の装置の間でやり取りされ、それらの装置内で用いられてもよい。図27は、複数の装置間でのデータの流れを示している。

[0213] 電子機器20001-1乃至20001-N (Nは1以上の整数)は、例えばユーザごとに所持され、それぞれ基地局(不図示)等を介してインターネット等のネットワーク20040に接続可能である。製造時において、電子機器20001-1には、学習装置20501が接続され、学習装置20501により提供される学習モデルを補助メモリ20104に記録することができる。学習装置20501は、シミュレータ20502により生成されたデータセットを学習データとして用いて学習モデルを生成し、電子機器20001-1に提供する。なお、学習データは、シミュレータ20502から提供されるデータセットに限らず、実際に光センサにより取得された画像データや、集約して管理されている取得済みの画像データなどを用いても構わない。

[0214] 図示は省略しているが、電子機器20001-2乃至20001-Nについても、電子機器20001-1と同様に、製造時の段階で学習モデルを記録することができる。以下、電子機器20001-1乃至20001-Nをそれぞれ区別する必要がある場合には、電子機器20001と呼ぶ。

[0215] ネットワーク20040には、電子機器20001のほかに、学習モデル生成サーバ20503、学習モデル提供サーバ20504、データ提供サーバ20505、及びアプリサーバ20506が接続され、相互にデータをやり取りすることができる。各サーバは、クラウドサーバとして設けることができる。

[0216] 学習モデル生成サーバ20503は、クラウドサーバ20003と同様の構成を有し、CPU等のプロセッサによって学習処理を行うことができる。学習モデル生成サーバ20503は、学習データを用いて学習モデルを生成する。図示した構成では、製造時に電子機器20001が学習モデルを記録する場合を例示しているが、学習モデルは、学習モデル生成サーバ2050

3から提供されてもよい。学習モデル生成サーバ20503は、生成した学習モデルを、ネットワーク20040を介して電子機器20001に送信する。電子機器20001は、学習モデル生成サーバ20503から送信されてくる学習モデルを受信し、補助メモリ20104に記録する。これにより、その学習モデルを備える電子機器20001が生成される。

[0217] すなわち、電子機器20001では、製造時の段階で学習モデルを記録していない場合には、学習モデル生成サーバ20503からの学習モデルを新規で記録することで、新たな学習モデルを記録した電子機器20001が生成される。また、電子機器20001では、製造時の段階で学習モデルを既に記録している場合、記録済みの学習モデルを、学習モデル生成サーバ20503からの学習モデルに更新することで、更新済みの学習モデルを記録した電子機器20001が生成される。電子機器20001では、適宜更新される学習モデルを用いて推論処理を行うことができる。

[0218] 学習モデルは、学習モデル生成サーバ20503から電子機器20001に直接提供するに限らず、各種の学習モデルを集約して管理する学習モデル提供サーバ20504がネットワーク20040を介して提供してもよい。学習モデル提供サーバ20504は、電子機器20001に限らず、他の装置に学習モデルを提供することで、その学習モデルを備える他の装置を生成しても構わない。また、学習モデルは、フラッシュメモリ等の着脱可能なメモリカードに記録して提供しても構わない。電子機器20001では、スロットに装着されたメモリカードから学習モデルを読み出して記録することができる。これにより、電子機器20001では、過酷環境下で使用される場合や、通信機能を有していない場合、通信機能を有しているが伝送可能な情報量が少ない場合などであっても、学習モデルを取得することができる。

[0219] 電子機器20001は、画像データや補正済みデータ、メタデータなどのデータを、ネットワーク20040を介して他の装置に提供することができる。例えば、電子機器20001は、画像データや補正済みデータ等のデータを、ネットワーク20040を介して学習モデル生成サーバ20503に

送信する。これにより、学習モデル生成サーバ20503では、1又は複数の電子機器20001から収集された画像データや補正済みデータ等のデータを学習データとして用い、学習モデルを生成することができる。より多くの学習データを用いることで、学習処理の精度を上げることができる。

[0220] 画像データや補正済みデータ等のデータは、電子機器20001から学習モデル生成サーバ20503に直接提供するに限らず、各種のデータを集約して管理するデータ提供サーバ20505が提供してもよい。データ提供サーバ20505は、電子機器20001に限らず他の装置からデータを収集してもよいし、学習モデル生成サーバ20503に限らず他の装置にデータを提供しても構わない。

[0221] 学習モデル生成サーバ20503は、既に生成された学習モデルに対し、電子機器20001又はデータ提供サーバ20505から提供された画像データや補正済みデータ等のデータを学習データに追加した再学習処理を行い、学習モデルを更新してもよい。更新された学習モデルは、電子機器20001に提供することができる。学習モデル生成サーバ20503において、学習処理又は再学習処理を行う場合、電子機器20001の仕様や性能の違いに依らず、処理を行うことができる。

[0222] また、電子機器20001において、補正済みデータやメタデータに対してユーザが修正の操作を行った場合（例えばユーザが正しい情報を入力した場合）に、その修正処理に関するフィードバックデータが、再学習処理に用いられてもよい。例えば、電子機器20001からのフィードバックデータを学習モデル生成サーバ20503に送信することで、学習モデル生成サーバ20503では、電子機器20001からのフィードバックデータを用いた再学習処理を行い、学習モデルを更新することができる。なお、電子機器20001では、ユーザによる修正の操作が行われる際に、アプリサーバ20506により提供されるアプリケーションが利用されてもよい。

[0223] 再学習処理は、電子機器20001が行ってもよい。電子機器20001において、画像データやフィードバックデータを用いた再学習処理を行って

学習モデルを更新する場合、装置内で学習モデルの改善を行うことができる。これにより、その更新された学習モデルを備える電子機器 20001 が生成される。また、電子機器 20001 は、再学習処理で得られる更新後の学習モデルを学習モデル提供サーバ 20504 に送信して、他の電子機器 20001 に提供されるようにしてもよい。これにより、複数の電子機器 20001 の間で、更新後の学習モデルを共有することができる。

[0224] あるいは、電子機器 20001 は、再学習された学習モデルの差分情報（更新前の学習モデルと更新後の学習モデルに関する差分情報）を、アップデート情報として、学習モデル生成サーバ 20503 に送信してもよい。学習モデル生成サーバ 20503 では、電子機器 20001 からのアップデート情報に基づき改善された学習モデルを生成して、他の電子機器 20001 に提供することができる。このような差分情報をやり取りすることで、全ての情報をやり取りする場合と比べてプライバシーを保護することができ、また通信コストを削減することができる。なお、電子機器 20001 と同様に、電子機器 20001 に搭載された光センサ 20011 が再学習処理を行ってもよい。

[0225] アプリサーバ 20506 は、ネットワーク 20040 を介して各種のアプリケーションを提供可能なサーバである。アプリケーションは、学習モデルや補正済みデータ、メタデータ等のデータを用いた所定の機能を提供する。電子機器 20001 は、ネットワーク 20040 を介してアプリサーバ 20506 からダウンロードしたアプリケーションを実行することで、所定の機能を実現することができる。あるいは、アプリサーバ 20506 は、例えば API（Application Programming Interface）などを介して電子機器 20001 からデータを取得し、アプリサーバ 20506 上でアプリケーションを実行することで、所定の機能を実現することもできる。

[0226] このように、本技術を適用した装置を含むシステムでは、各装置の間で、学習モデル、画像データ、補正済みデータ等のデータがやり取りされて流通し、それらのデータを用いた様々なサービスを提供することが可能となる。

例えば、学習モデル提供サーバ20504を介した学習モデルを提供するサービスや、データ提供サーバ20505を介した画像データや補正済みデータ等のデータを提供するサービスを提供することができる。また、アプリケーションサーバ20506を介したアプリケーションを提供するサービスを提供することができる。

[0227] あるいは、学習モデル提供サーバ20504により提供される学習モデルに、電子機器20001の光センサ20011から取得した画像データを入力して、その出力として得られる補正済みデータが提供されてもよい。また、学習モデル提供サーバ20504により提供される学習モデルを実装した電子機器などの装置を生成して提供してもよい。さらに、学習モデルや補正済みデータ、メタデータ等のデータを読み出し可能な記憶媒体に記録することで、それらのデータが記録された記憶媒体や、当該記憶媒体を搭載した電子機器などの装置を生成して提供してもよい。当該記憶媒体は、磁気ディスク、光ディスク、光磁気ディスク、半導体メモリなどの不揮発性メモリでもよいし、SRAMやDRAMなどの揮発性メモリでもよい。

[0228] なお、本開示の実施の形態は、上述した実施の形態に限定されるものではなく、本開示の要旨を逸脱しない範囲において種々の変更が可能である。また、本明細書に記載された効果はあくまで例示であって限定されるものではなく、他の効果があってもよい。

[0229] また、本開示は、以下のような構成をとることができる。

[0230] (1)

第1のセンサにより取得された対象物を深度情報で示した第1の画像、第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像、及び前記第1の画像と前記第2の画像から得られる第3の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記第1の画像に含まれる補正対象画素を特定する

処理部を備える

情報処理装置。

(2)

前記学習済みモデルは、前記第1の画像及び前記第2の画像を入力とし、前記第1の画像に対して指定された補正対象画素を含む第1の領域を教師データとして学習したディープニューラルネットワークである

前記(1)に記載の情報処理装置。

(3)

前記学習済みモデルは、特定した前記補正対象画素を含む第2の領域として、セマンティックセグメンテーションによる二値分類画像、又は物体検知アルゴリズムによる座標情報を出力する

前記(1)又は(2)に記載の情報処理装置。

(4)

前記第1の画像は、前記第2のセンサの視点に変換されて処理される

前記(2)又は(3)に記載の情報処理装置。

(5)

前記学習済みモデルは、欠陥がない前記第1の画像及び前記第2の画像を入力として教師なし学習をしたオートエンコーダであり、

前記処理部は、

欠陥の可能性がある前記第1の画像と、前記学習済みモデルから出力される前記第1の画像とを比較し、

比較結果に基づいて、前記補正対象画素を特定する

前記(1)に記載の情報処理装置。

(6)

前記処理部は、

比較対象となる2つの前記第1の画像の各画素の距離値の比を算出し、

算出した前記比が所定の閾値以上となる画素を、前記補正対象画素として特定する

前記(5)に記載の情報処理装置。

(7)

前記第 1 の画像は、前記第 2 のセンサの視点に変換されて処理される
前記（５）又は（６）に記載の情報処理装置。

（８）

情報処理装置が、

第 1 のセンサにより取得された対象物を深度情報で示した第 1 の画像、第
2 のセンサにより取得された前記対象物の像を面情報で示した第 2 の画像、
及び前記第 1 の画像と前記第 2 の画像から得られる第 3 の画像の少なくとも
一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記
第 1 の画像に含まれる補正対象画素を特定する

情報処理方法。

（９）

コンピュータを、

第 1 のセンサにより取得された対象物を深度情報で示した第 1 の画像、第
2 のセンサにより取得された前記対象物の像を面情報で示した第 2 の画像、
及び前記第 1 の画像と前記第 2 の画像から得られる第 3 の画像の少なくとも
一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記
第 1 の画像に含まれる補正対象画素を特定する

処理部を備える

情報処理装置として機能させるプログラム。

（１０）

第 1 のセンサにより取得された対象物を深度情報で示した第 1 の画像、及
び第 2 のセンサにより取得された前記対象物の像を面情報で示した第 2 の画
像を取得し、

前記第 1 の画像と対になる前記第 2 の画像に基づいて、第 3 の画像として
前記 1 の画像を擬似的に生成し、

前記第 1 の画像と前記第 3 の画像とを比較し、

比較結果に基づいて、前記第 1 の画像に含まれる補正対象画素を特定する
処理部を備える

情報処理装置。

(11)

前記処理部は、GANを利用して、前記第2の画像から前記第3の画像を生成する

前記(10)に記載の情報処理装置。

(12)

前記処理部は、前記第1の画像と対になる前記第2の画像の対応関係をGANにより学習した学習済みモデルを利用する

前記(11)に記載の情報処理装置。

(13)

前記処理部は、

撮影パラメータに基づいて、前記第1の画像を、前記第2のセンサの視点に変換した第4の画像を生成し、

前記第4の画像を、前記第3の画像と比較する

前記(10)乃至(12)のいずれかに記載の情報処理装置。

(14)

前記処理部は、前記第1の画像と前記第3の画像に対し、対応する画素ごとの輝度の差又は比をとることで比較を行う

前記(10)乃至(13)のいずれかに記載の情報処理装置。

(15)

前記処理部は、

予め定められた閾値を設定し、

画素ごとの輝度の差又は比の絶対値が前記閾値以上となる画素を、前記補正対象画素として特定する

前記(14)に記載の情報処理装置。

(16)

前記処理部は、前記第1の画像における前記補正対象画素を含む周辺領域の輝度を置き換えることで、前記補正対象画素を補正する

前記（１０）乃至（１５）のいずれかに記載の情報処理装置。

（１７）

前記処理部は、前記周辺領域に含まれる画素のうち、前記補正対象画素を除いた画素の輝度値の統計量を算出して前記周辺領域の輝度値に置き換えるか、又は、前記第３の画像における前記周辺領域に対応する領域の輝度値で前記周辺領域の輝度値を置き換える

前記（１６）に記載の情報処理装置。

（１８）

情報処理装置が、

第１のセンサにより取得された対象物を深度情報で示した第１の画像、及び第２のセンサにより取得された前記対象物の像を面情報で示した第２の画像を取得し、

前記第１の画像と対になる前記第２の画像に基づいて、第３の画像として前記第１の画像を擬似的に生成し、

前記第１の画像と前記第３の画像とを比較し、

比較結果に基づいて、前記第１の画像に含まれる補正対象画素を特定する情報処理方法。

（１９）

コンピュータを、

第１のセンサにより取得された対象物を深度情報で示した第１の画像、及び第２のセンサにより取得された前記対象物の像を面情報で示した第２の画像を取得し、

前記第１の画像と対になる前記第２の画像に基づいて、第３の画像として前記第１の画像を擬似的に生成し、

前記第１の画像と前記第３の画像とを比較し、

比較結果に基づいて、前記第１の画像に含まれる補正対象画素を特定する処理部を備える

情報処理装置として機能させるプログラム。

(20)

第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成する処理部を備え、

前記処理部は、

前記第1の画像の各画素に応じた第1の位置の深度情報に基づいて、前記第1の位置を前記第2の画像の画像面に写像し、

前記第2の画像の各画素に応じた第2の位置のうち、前記第1の位置の深度情報が割り当てられていない第2の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第2の画像における前記画素補正位置の深度情報を推論する

情報処理装置。

(21)

前記学習済みモデルは、深度情報に欠陥がある前記第3の画像と前記画素補正位置を入力とした学習によって、補正済みの前記第3の画像を出力するようになったニューラルネットワークである

前記(20)に記載の情報処理装置。

(22)

前記学習済みモデルは、欠陥がない前記第3の画像を入力とした教師なし学習によって、補正済みの前記第3の画像を出力するようになったニューラルネットワークである

前記(20)に記載の情報処理装置。

(23)

情報処理装置が、

第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成するに際して、

前記第 1 の画像の各画素に応じた第 1 の位置の深度情報に基づいて、前記第 1 の位置を前記第 2 の画像の画像面に写像し、

前記第 2 の画像の各画素に応じた第 2 の位置のうち、前記第 1 の位置の深度情報が割り当てられていない第 2 の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第 2 の画像における前記画素補正位置の深度情報を推論する

情報処理方法。

(24)

コンピュータを、

第 1 のセンサにより取得された対象物を深度情報で示した第 1 の画像を、第 2 のセンサにより取得された前記対象物の像を色情報で示した第 2 の画像の画像面に写像して第 3 の画像を生成する処理部を備え、

前記処理部は、

前記第 1 の画像の各画素に応じた第 1 の位置の深度情報に基づいて、前記第 1 の位置を前記第 2 の画像の画像面に写像し、

前記第 2 の画像の各画素に応じた第 2 の位置のうち、前記第 1 の位置の深度情報が割り当てられていない第 2 の位置を、画素補正位置として特定し

、

機械学習により学習された学習済みモデルを用いて、前記第 2 の画像における前記画素補正位置の深度情報を推論する

情報処理装置として機能させるプログラム。

(25)

第 2 のセンサにより取得された対象物の像を色情報で示した第 2 の画像を、第 1 のセンサにより取得された前記対象物を深度情報で示した第 1 の画像の画像面に写像して第 3 の画像を生成する処理部を備え、

前記処理部は、

前記第 1 の画像の各画素に応じた第 1 の位置のうち、有効な深度情報が割り当てられていない第 1 の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第 1 の画像における前記画素補正位置の深度情報を推論し、

前記第 1 の位置に割り当てられた深度情報に基づいて、前記第 2 の画像における第 2 の位置から色情報をサンプリングして、前記第 2 の位置を前記第 1 の画像の画像面に写像する

情報処理装置。

(26)

前記学習済みモデルは、欠陥がある前記第 1 の画像と前記画素補正位置を入力とした学習によって、補正後の深度情報を出力するようになったニューラルネットワークである

前記(25)に記載の情報処理装置。

(27)

情報処理装置が、

第 2 のセンサにより取得された対象物の像を色情報で示した第 2 の画像を、第 1 のセンサにより取得された前記対象物を深度情報で示した第 1 の画像の画像面に写像して第 3 の画像を生成するに際して、

前記第 1 の画像の各画素に応じた第 1 の位置のうち、有効な深度情報が割り当てられていない第 1 の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第 1 の画像における前記画素補正位置の深度情報を推論し、

前記第 1 の位置に割り当てられた深度情報に基づいて、前記第 2 の画像における第 2 の位置から色情報をサンプリングして、前記第 2 の位置を前記第 1 の画像の画像面に写像する

情報処理方法。

(28)

コンピュータを、

第 2 のセンサにより取得された対象物の像を色情報で示した第 2 の画像を、第 1 のセンサにより取得された前記対象物を深度情報で示した第 1 の画像

の画像面に写像して第3の画像を生成する処理部を備え、

前記処理部は、

前記第1の画像の各画素に応じた第1の位置のうち、有効な深度情報が割り当てられていない第1の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第1の画像における前記画素補正位置の深度情報を推論し、

前記第1の位置に割り当てられた深度情報に基づいて、前記第2の画像における第2の位置から色情報をサンプリングして、前記第2の位置を前記第1の画像の画像面に写像する

情報処理装置として機能させるプログラム。

符号の説明

[0231] 1 情報処理装置, 2 学習装置, 10 処理部, 11 デプスセンサ, 12 RGBセンサ, 13 デプス処理部, 14 RGB処理部, 111 視点変換部, 112 欠陥領域指定部, 113 学習モデル, 114 減算部, 121 視点変換部, 122 学習モデル, 131 視点変換部, 132 学習モデル, 133 減算部, 141 視点変換部, 142 学習モデル, 143 比較部, 201 特定部, 202 補正部, 211 学習モデル, 212 視点変換部, 213 比較部, 301 画像生成部, 311 推論部, 321 学習モデル, 331 学習モデル, 341 学習モデル, 351 学習モデル

請求の範囲

- [請求項1] 第1のセンサにより取得された対象物を深度情報で示した第1の画像、第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像、及び前記第1の画像と前記第2の画像から得られる第3の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記第1の画像に含まれる補正対象画素を特定する
- 処理部を備える
- 情報処理装置。
- [請求項2] 前記学習済みモデルは、前記第1の画像及び前記第2の画像を入力とし、前記第1の画像に対して指定された補正対象画素を含む第1の領域を教師データとして学習したディープニューラルネットワークである
- 請求項1に記載の情報処理装置。
- [請求項3] 前記学習済みモデルは、特定した前記補正対象画素を含む第2の領域として、セマンティックセグメンテーションによる二値分類画像、又は物体検知アルゴリズムによる座標情報を出力する
- 請求項2に記載の情報処理装置。
- [請求項4] 前記第1の画像は、前記第2のセンサの視点に変換されて処理される
- 請求項2に記載の情報処理装置。
- [請求項5] 前記学習済みモデルは、欠陥がない前記第1の画像及び前記第2の画像を入力として教師なし学習をしたオートエンコーダであり、
- 前記処理部は、
- 欠陥の可能性がある前記第1の画像と、前記学習済みモデルから出力される前記第1の画像とを比較し、
- 比較結果に基づいて、前記補正対象画素を特定する
- 請求項1に記載の情報処理装置。

- [請求項6] 前記処理部は、
比較対象となる２つの前記第１の画像の各画素の距離値の比を算出し、
算出した前記比が所定の閾値以上となる画素を、前記補正対象画素として特定する
請求項５に記載の情報処理装置。
- [請求項7] 前記第１の画像は、前記第２のセンサの視点に変換されて処理される
請求項５に記載の情報処理装置。
- [請求項8] 情報処理装置が、
第１のセンサにより取得された対象物を深度情報で示した第１の画像、第２のセンサにより取得された前記対象物の像を面情報で示した第２の画像、及び前記第１の画像と前記第２の画像から得られる第３の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記第１の画像に含まれる補正対象画素を特定する
情報処理方法。
- [請求項9] コンピュータを、
第１のセンサにより取得された対象物を深度情報で示した第１の画像、第２のセンサにより取得された前記対象物の像を面情報で示した第２の画像、及び前記第１の画像と前記第２の画像から得られる第３の画像の少なくとも一部に機械学習により学習された学習済みモデルを用いた処理を行い、前記第１の画像に含まれる補正対象画素を特定する
処理部を備える
情報処理装置として機能させるプログラム。
- [請求項10] 第１のセンサにより取得された対象物を深度情報で示した第１の画像、及び第２のセンサにより取得された前記対象物の像を面情報で示

した第2の画像を取得し、

前記第1の画像と対になる前記第2の画像に基づいて、第3の画像として前記1の画像を擬似的に生成し、

前記第1の画像と前記第3の画像とを比較し、

比較結果に基づいて、前記第1の画像に含まれる補正対象画素を特定する

処理部を備える

情報処理装置。

[請求項11] 前記処理部は、GANを利用して、前記第2の画像から前記第3の画像を生成する

請求項10に記載の情報処理装置。

[請求項12] 前記処理部は、前記第1の画像と対になる前記第2の画像の対応関係をGANにより学習した学習済みモデルを利用する

請求項11に記載の情報処理装置。

[請求項13] 前記処理部は、

撮影パラメータに基づいて、前記第1の画像を、前記第2のセンサの視点に変換した第4の画像を生成し、

前記第4の画像を、前記第3の画像と比較する

請求項10に記載の情報処理装置。

[請求項14] 前記処理部は、前記第1の画像と前記第3の画像に対し、対応する画素ごとの輝度の差又は比をとることで比較を行う

請求項10に記載の情報処理装置。

[請求項15] 前記処理部は、

予め定められた閾値を設定し、

画素ごとの輝度の差又は比の絶対値が前記閾値以上となる画素を、前記補正対象画素として特定する

請求項14に記載の情報処理装置。

[請求項16] 前記処理部は、前記第1の画像における前記補正対象画素を含む周

辺領域の輝度を置き換えることで、前記補正対象画素を補正する
請求項 10 に記載の情報処理装置。

[請求項17] 前記処理部は、前記周辺領域に含まれる画素のうち、前記補正対象画素を除いた画素の輝度値の統計量を算出して前記周辺領域の輝度値に置き換えるか、又は、前記第3の画像における前記周辺領域に対応する領域の輝度値で前記周辺領域の輝度値を置き換える
請求項 16 に記載の情報処理装置。

[請求項18] 情報処理装置が、
第1のセンサにより取得された対象物を深度情報で示した第1の画像、及び第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像を取得し、
前記第1の画像と対になる前記第2の画像に基づいて、第3の画像として前記1の画像を擬似的に生成し、
前記第1の画像と前記第3の画像とを比較し、
比較結果に基づいて、前記第1の画像に含まれる補正対象画素を特定する
情報処理方法。

[請求項19] コンピュータを、
第1のセンサにより取得された対象物を深度情報で示した第1の画像、及び第2のセンサにより取得された前記対象物の像を面情報で示した第2の画像を取得し、
前記第1の画像と対になる前記第2の画像に基づいて、第3の画像として前記1の画像を擬似的に生成し、
前記第1の画像と前記第3の画像とを比較し、
比較結果に基づいて、前記第1の画像に含まれる補正対象画素を特定する
処理部を備える
情報処理装置として機能させるプログラム。

[請求項20] 第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成する処理部を備え、
前記処理部は、

前記第1の画像の各画素に応じた第1の位置の深度情報に基づいて、前記第1の位置を前記第2の画像の画像面に写像し、

前記第2の画像の各画素に応じた第2の位置のうち、前記第1の位置の深度情報が割り当てられていない第2の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第2の画像における前記画素補正位置の深度情報を推論する
情報処理装置。

[請求項21] 前記学習済みモデルは、深度情報に欠陥がある前記第3の画像と前記画素補正位置を入力とした学習によって、補正済みの前記第3の画像を出力するようになったニューラルネットワークである
請求項20に記載の情報処理装置。

[請求項22] 前記学習済みモデルは、欠陥がない前記第3の画像を入力とした教師なし学習によって、補正済みの前記第3の画像を出力するようになったニューラルネットワークである
請求項20に記載の情報処理装置。

[請求項23] 情報処理装置が、
第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成するに際して、
前記第1の画像の各画素に応じた第1の位置の深度情報に基づいて、前記第1の位置を前記第2の画像の画像面に写像し、
前記第2の画像の各画素に応じた第2の位置のうち、前記第1の位

置の深度情報が割り当てられていない第2の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第2の画像における前記画素補正位置の深度情報を推論する
情報処理方法。

[請求項24]

コンピュータを、

第1のセンサにより取得された対象物を深度情報で示した第1の画像を、第2のセンサにより取得された前記対象物の像を色情報で示した第2の画像の画像面に写像して第3の画像を生成する処理部を備え、

前記処理部は、

前記第1の画像の各画素に応じた第1の位置の深度情報に基づいて、前記第1の位置を前記第2の画像の画像面に写像し、

前記第2の画像の各画素に応じた第2の位置のうち、前記第1の位置の深度情報が割り当てられていない第2の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第2の画像における前記画素補正位置の深度情報を推論する
情報処理装置として機能させるプログラム。

[請求項25]

第2のセンサにより取得された対象物の像を色情報で示した第2の画像を、第1のセンサにより取得された前記対象物を深度情報で示した第1の画像の画像面に写像して第3の画像を生成する処理部を備え、

前記処理部は、

前記第1の画像の各画素に応じた第1の位置のうち、有効な深度情報が割り当てられていない第1の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第1の

画像における前記画素補正位置の深度情報を推論し、

前記第 1 の位置に割り当てられた深度情報に基づいて、前記第 2 の画像における第 2 の位置から色情報をサンプリングして、前記第 2 の位置を前記第 1 の画像の画像面に写像する
情報処理装置。

[請求項26] 前記学習済みモデルは、欠陥がある前記第 1 の画像と前記画素補正位置を入力とした学習によって、補正後の深度情報を出力するようになったニューラルネットワークである

請求項 25 に記載の情報処理装置。

[請求項27] 情報処理装置が、

第 2 のセンサにより取得された対象物の像を色情報で示した第 2 の画像を、第 1 のセンサにより取得された前記対象物を深度情報で示した第 1 の画像の画像面に写像して第 3 の画像を生成するに際して、

前記第 1 の画像の各画素に応じた第 1 の位置のうち、有効な深度情報が割り当てられていない第 1 の位置を、画素補正位置として特定し、

機械学習により学習された学習済みモデルを用いて、前記第 1 の画像における前記画素補正位置の深度情報を推論し、

前記第 1 の位置に割り当てられた深度情報に基づいて、前記第 2 の画像における第 2 の位置から色情報をサンプリングして、前記第 2 の位置を前記第 1 の画像の画像面に写像する

情報処理方法。

[請求項28] コンピュータを、

第 2 のセンサにより取得された対象物の像を色情報で示した第 2 の画像を、第 1 のセンサにより取得された前記対象物を深度情報で示した第 1 の画像の画像面に写像して第 3 の画像を生成する処理部を備え、

前記処理部は、

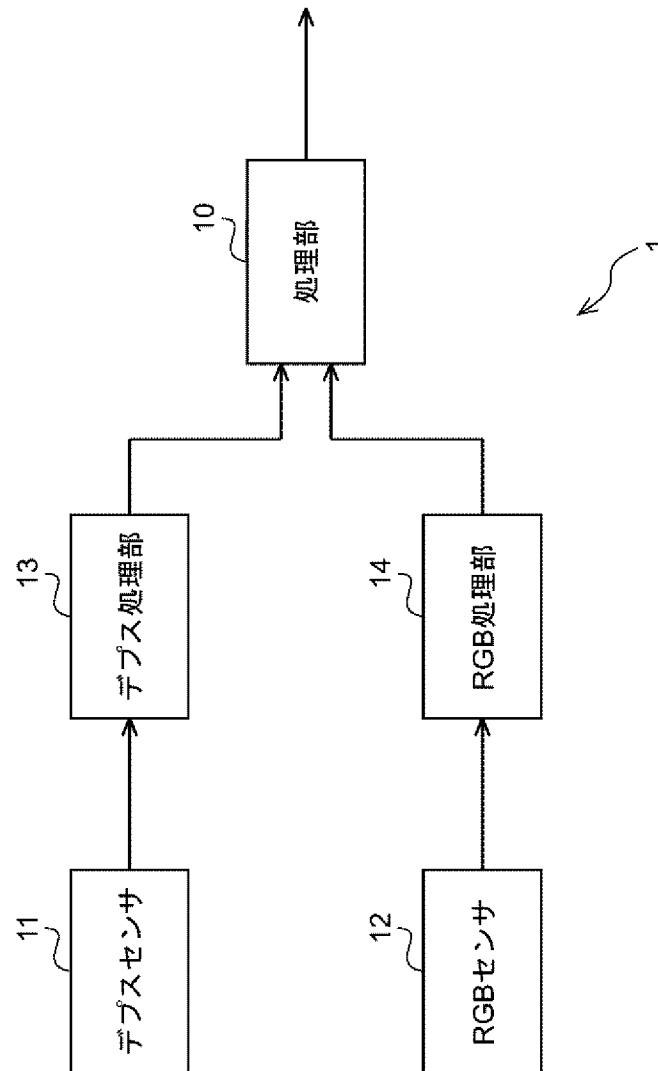
前記第 1 の画像の各画素に応じた第 1 の位置のうち、有効な深度情報が割り当てられていない第 1 の位置を、画素補正位置として特定し、

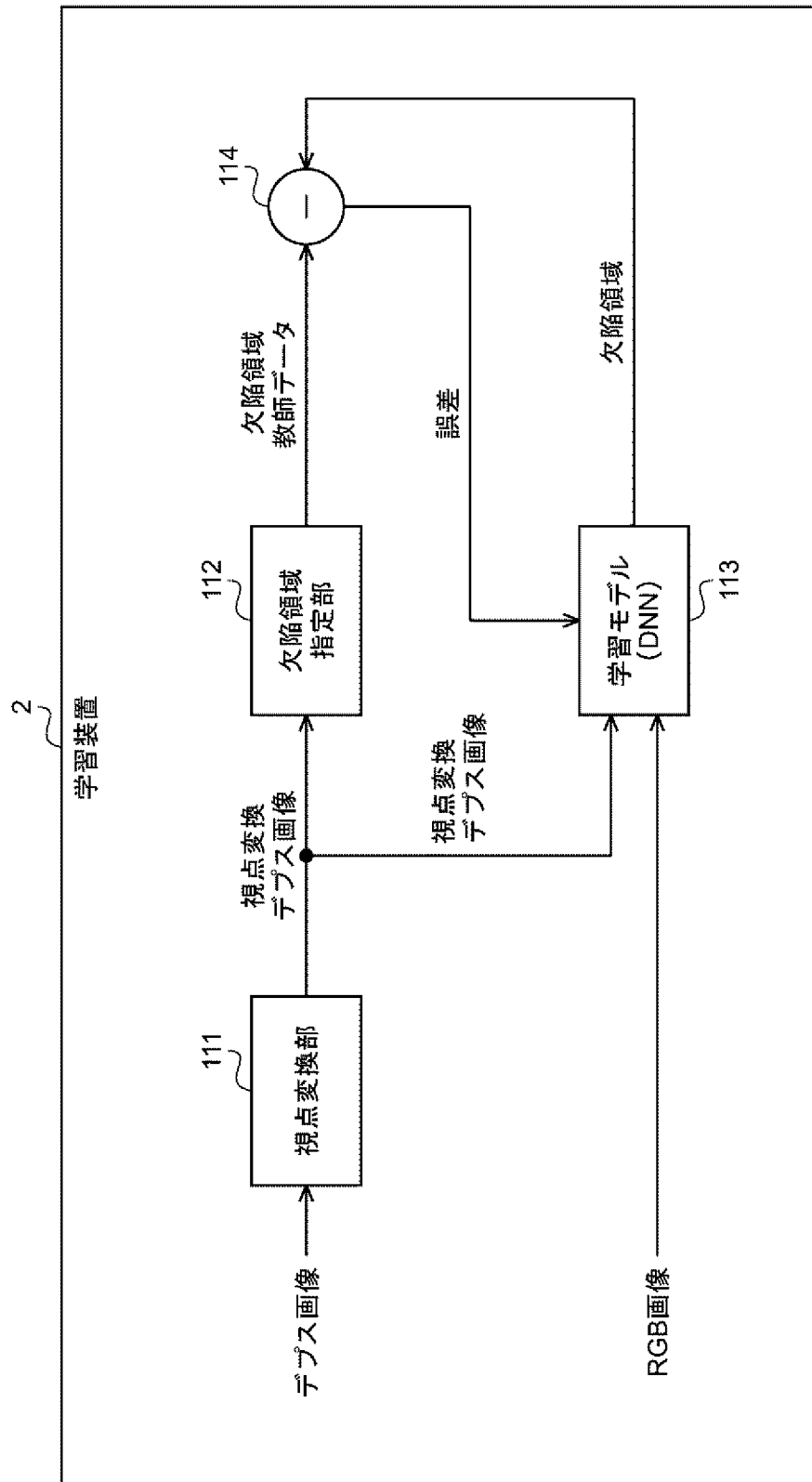
機械学習により学習された学習済みモデルを用いて、前記第 1 の画像における前記画素補正位置の深度情報を推論し、

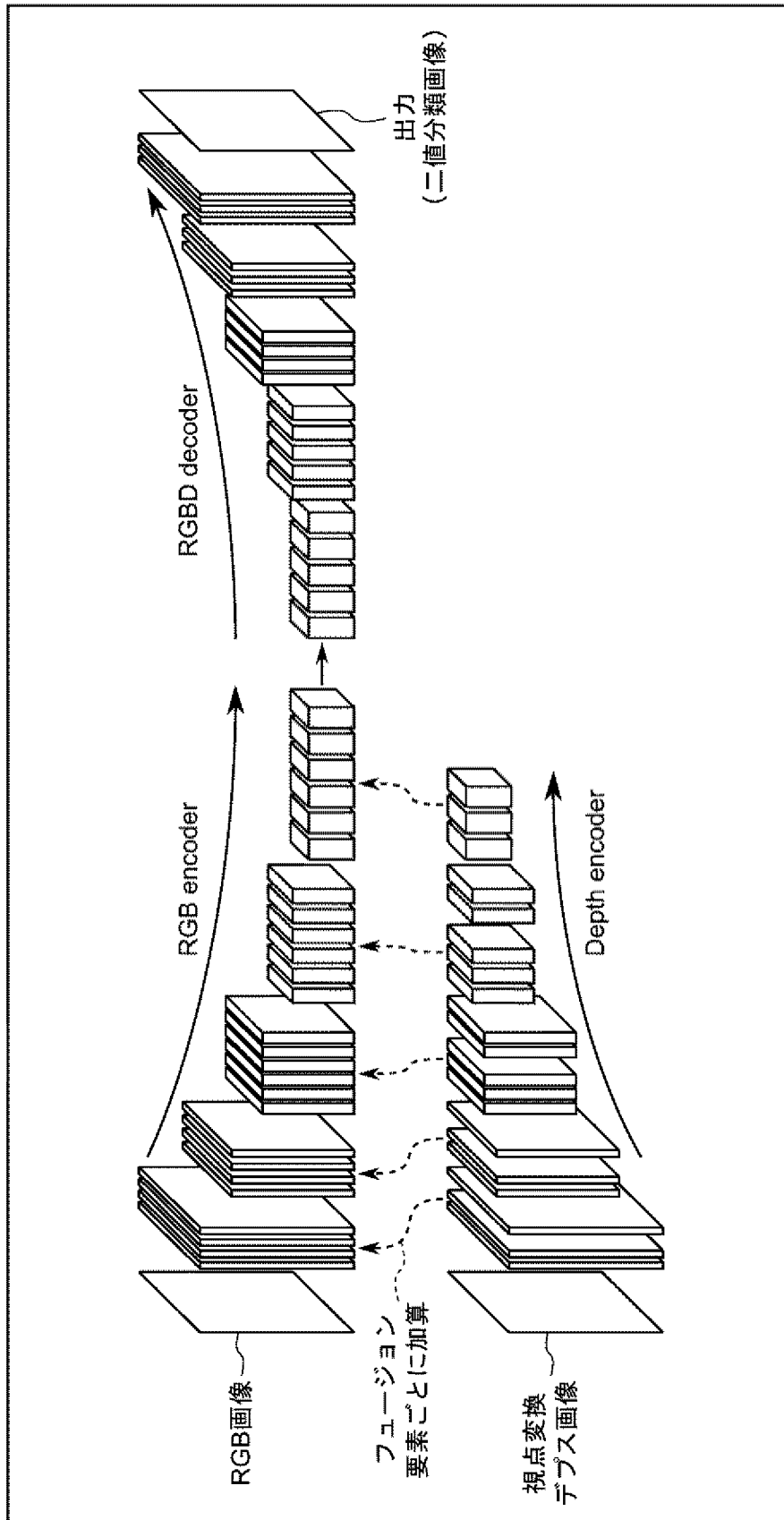
前記第 1 の位置に割り当てられた深度情報に基づいて、前記第 2 の画像における第 2 の位置から色情報をサンプリングして、前記第 2 の位置を前記第 1 の画像の画像面に写像する

情報処理装置として機能させるプログラム。

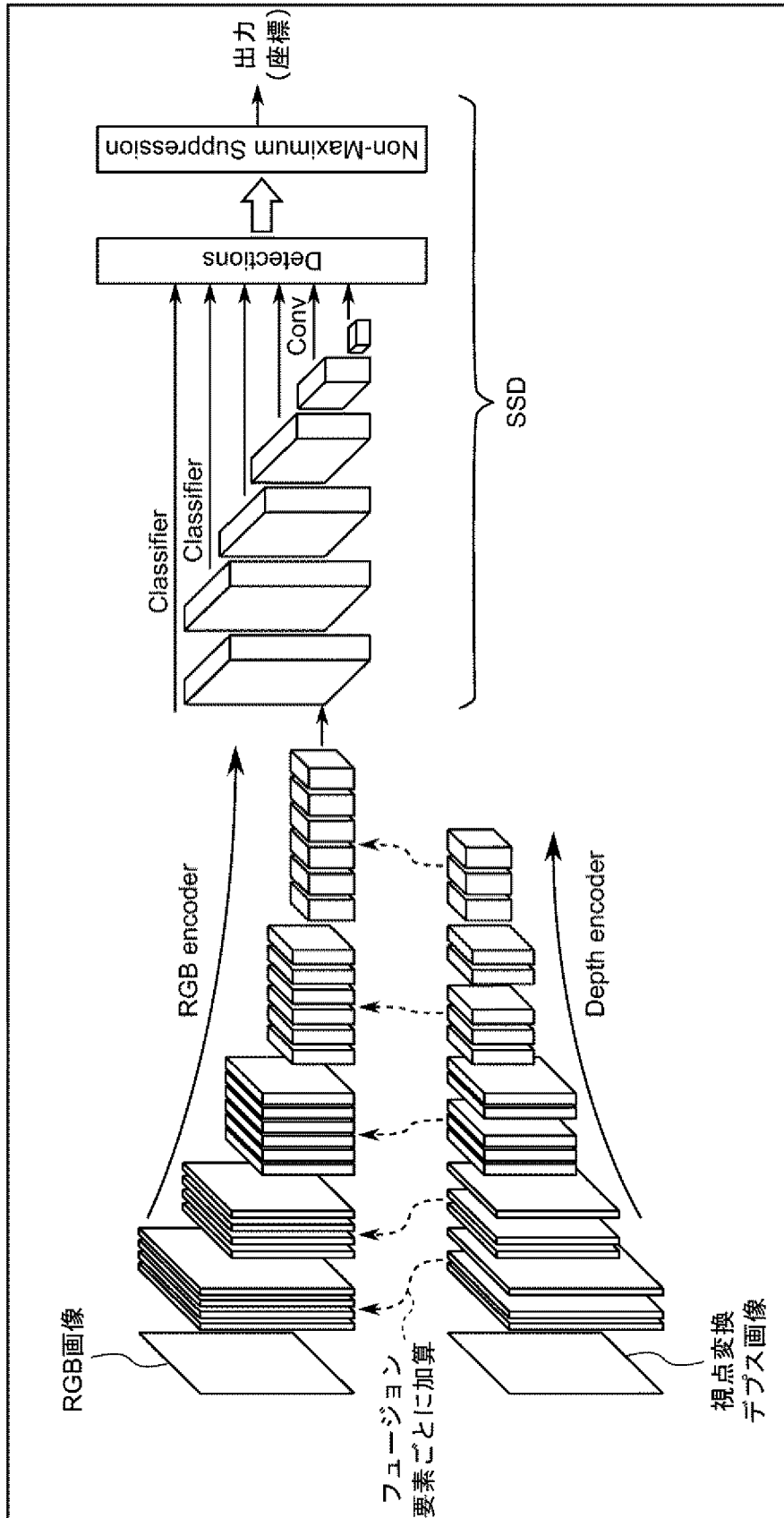
[図1]
FIG.1

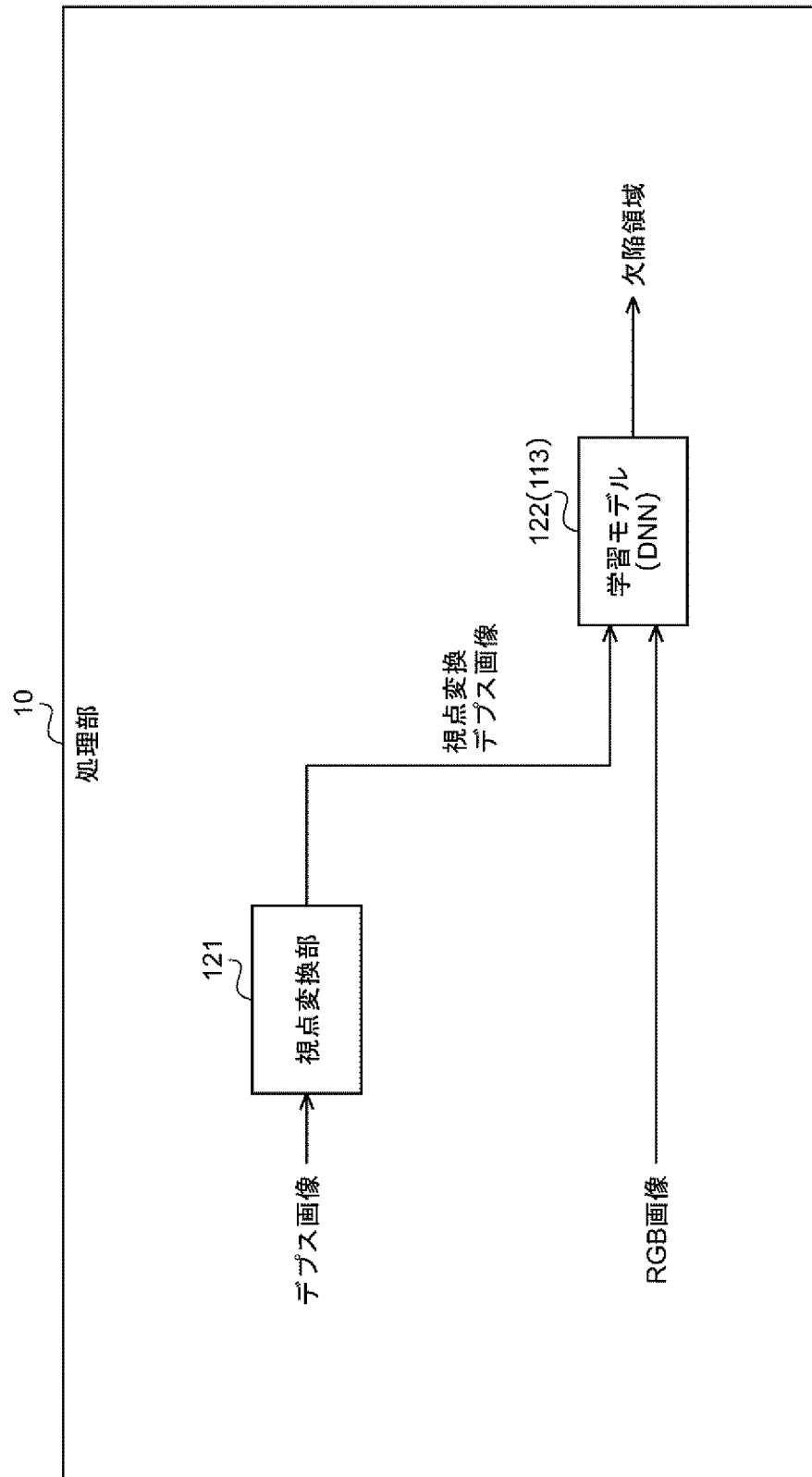


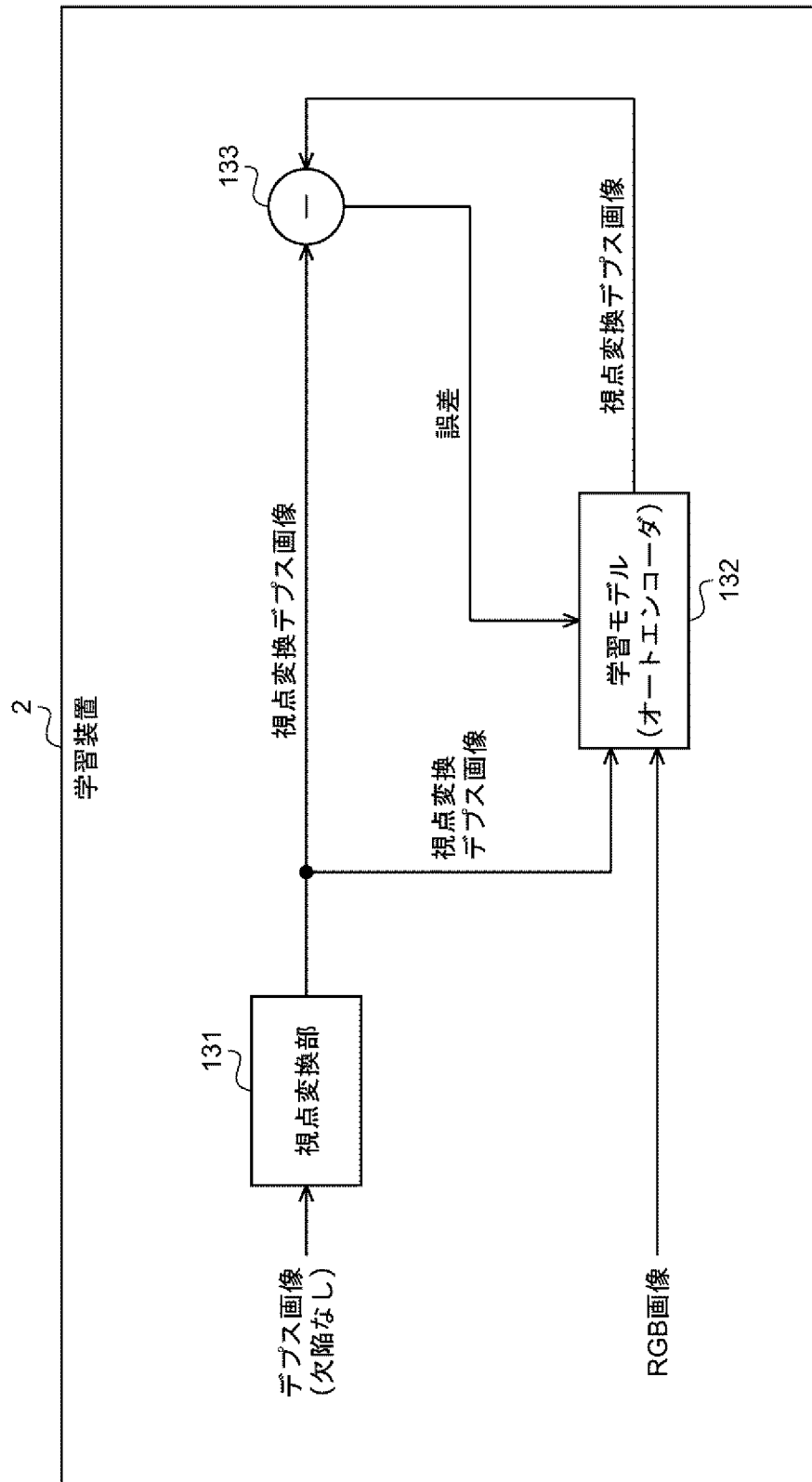
[図2]
FIG.2

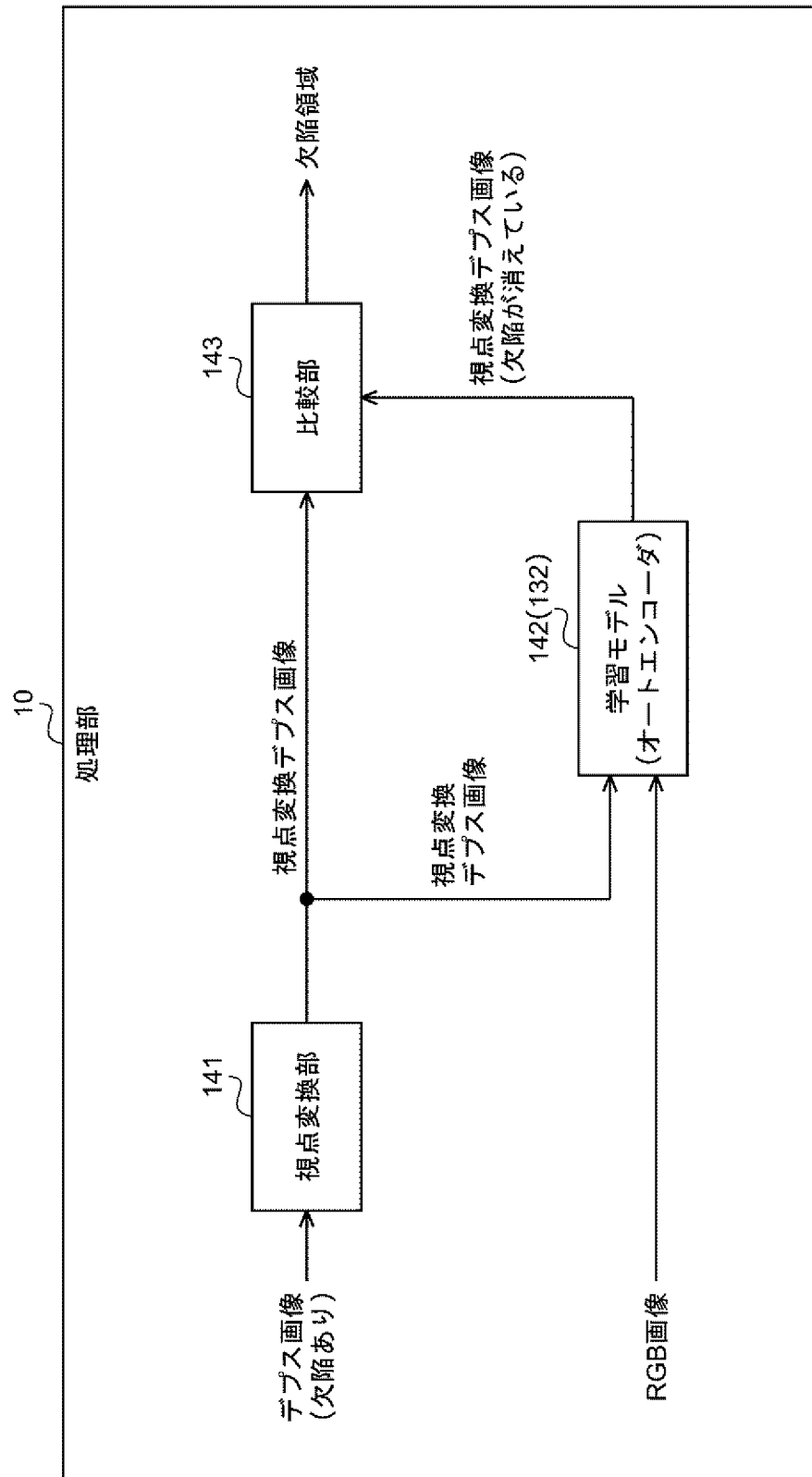
[図3]
FIG.3

[図4]
FIG.4

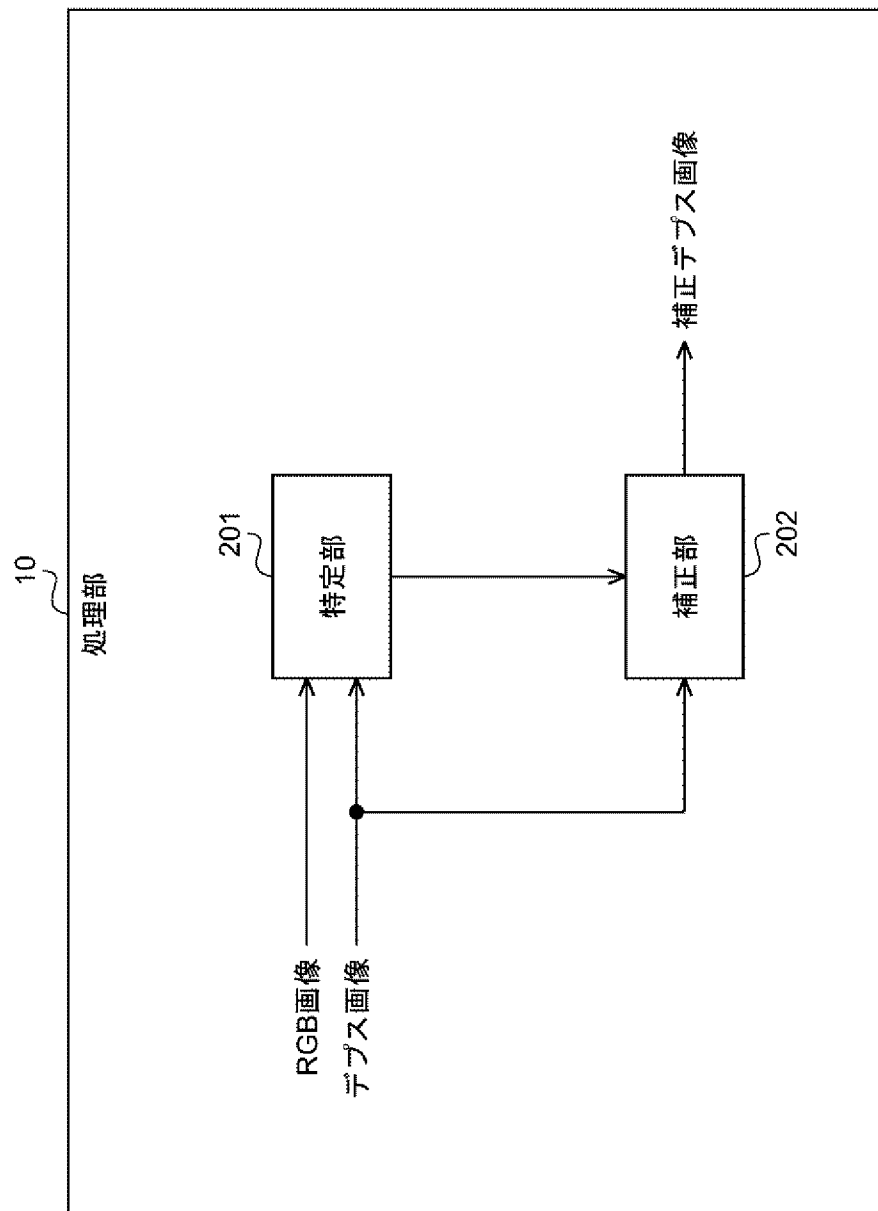


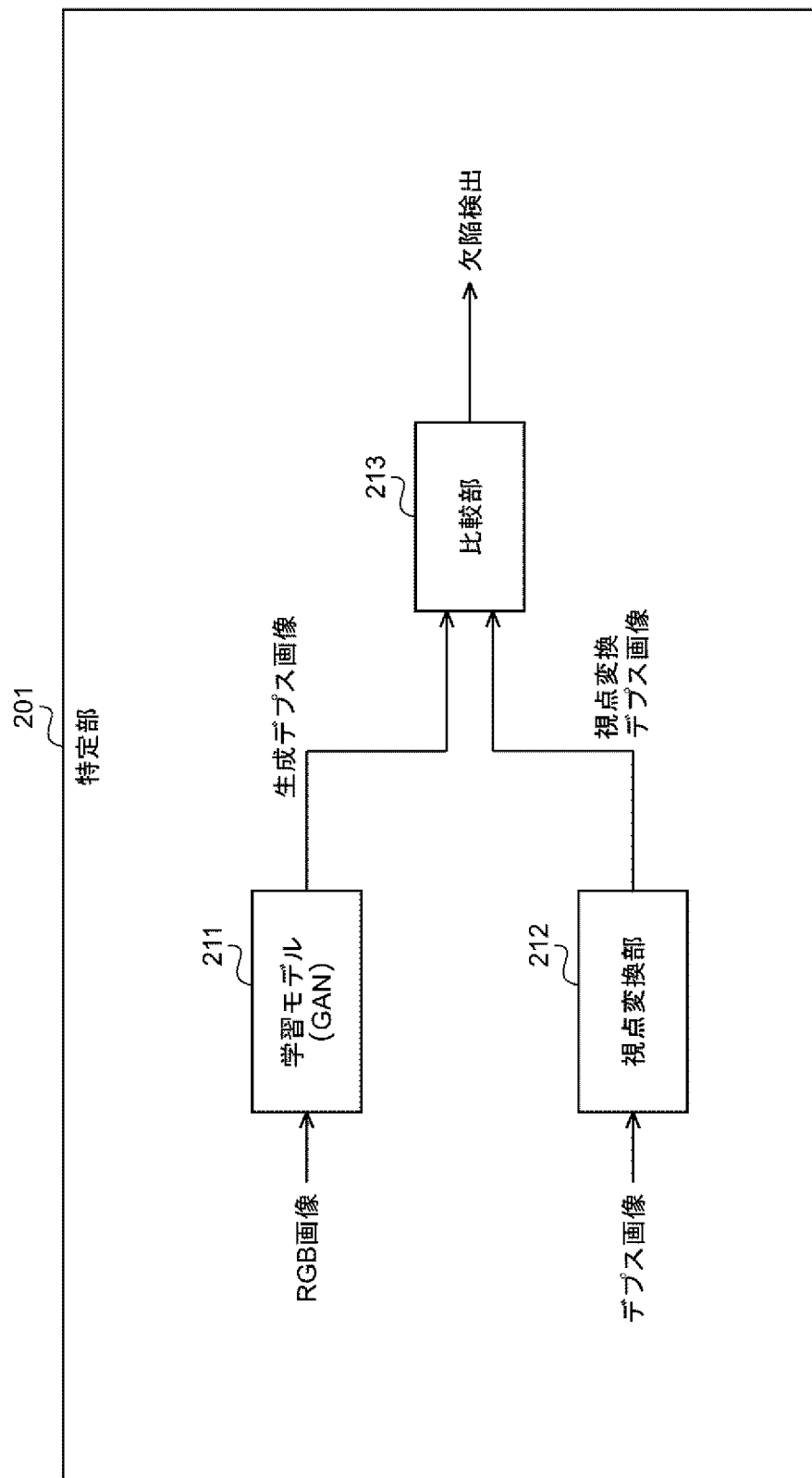
[図5]
FIG.5

[図6]
FIG.6

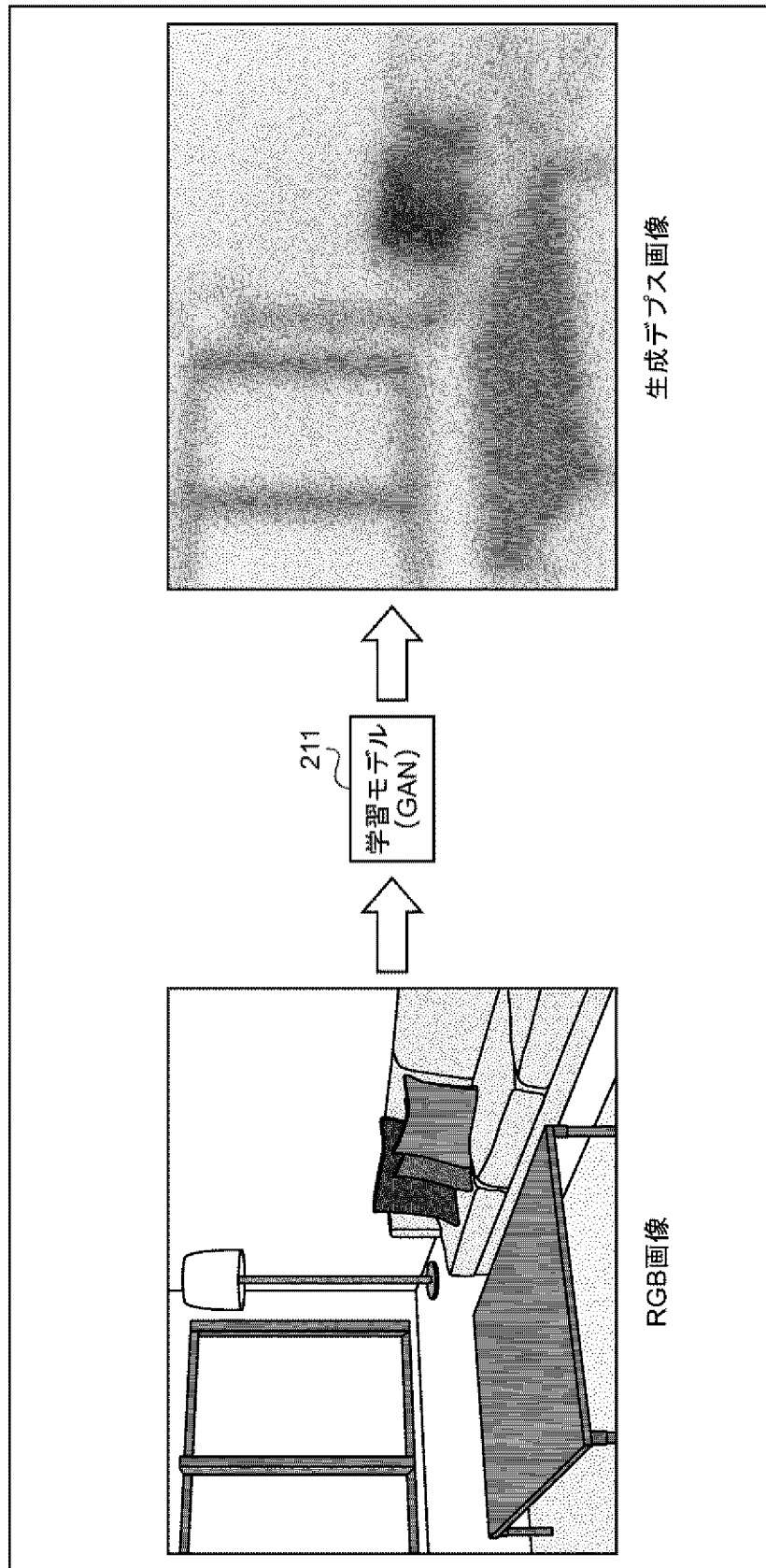
[図7]
FIG.7

[図8]
FIG.8

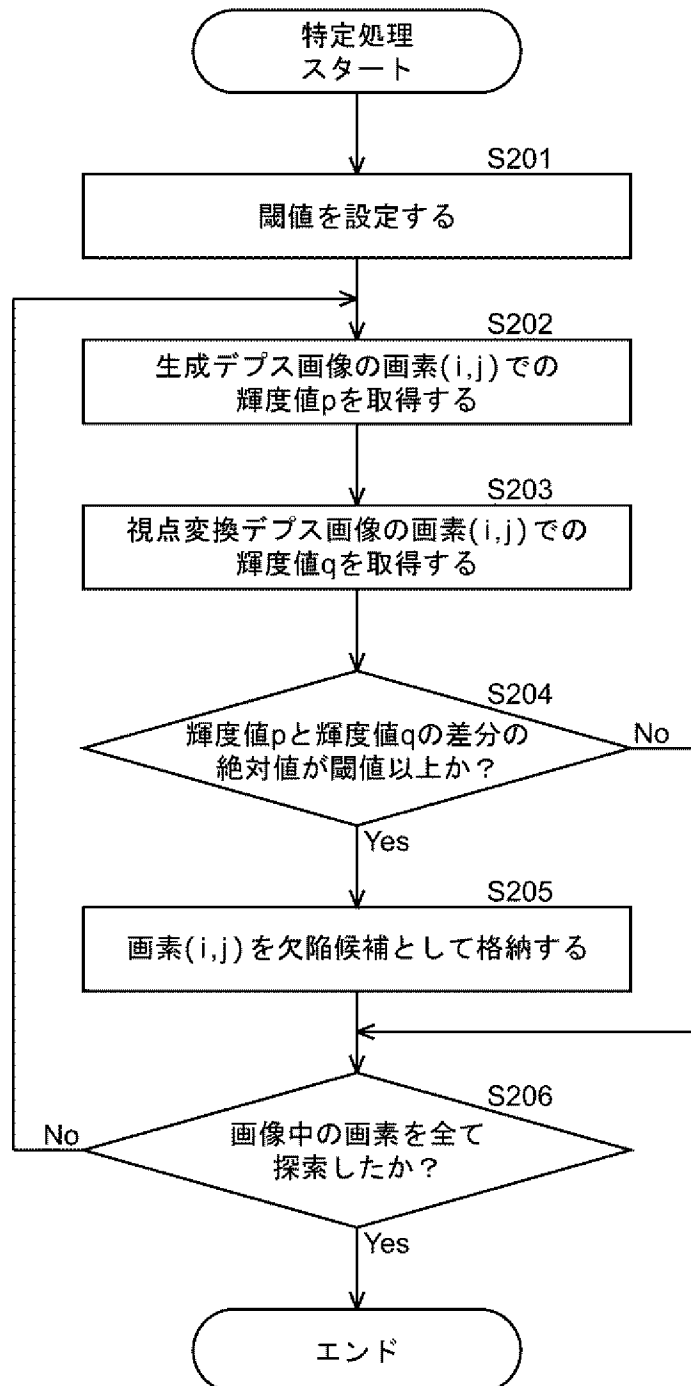


[図9]
FIG.9

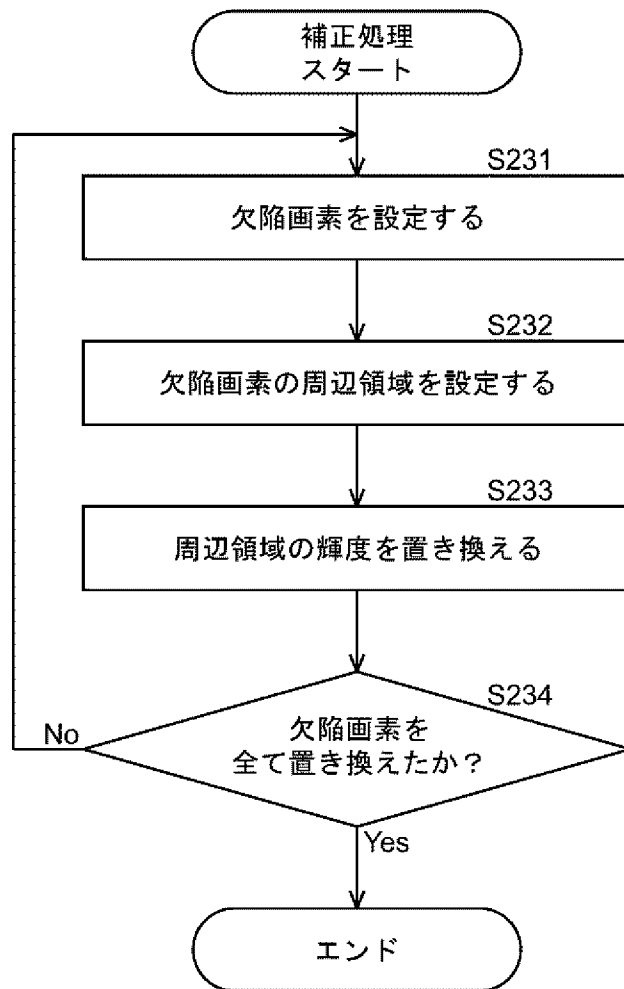
[図10]
FIG.10



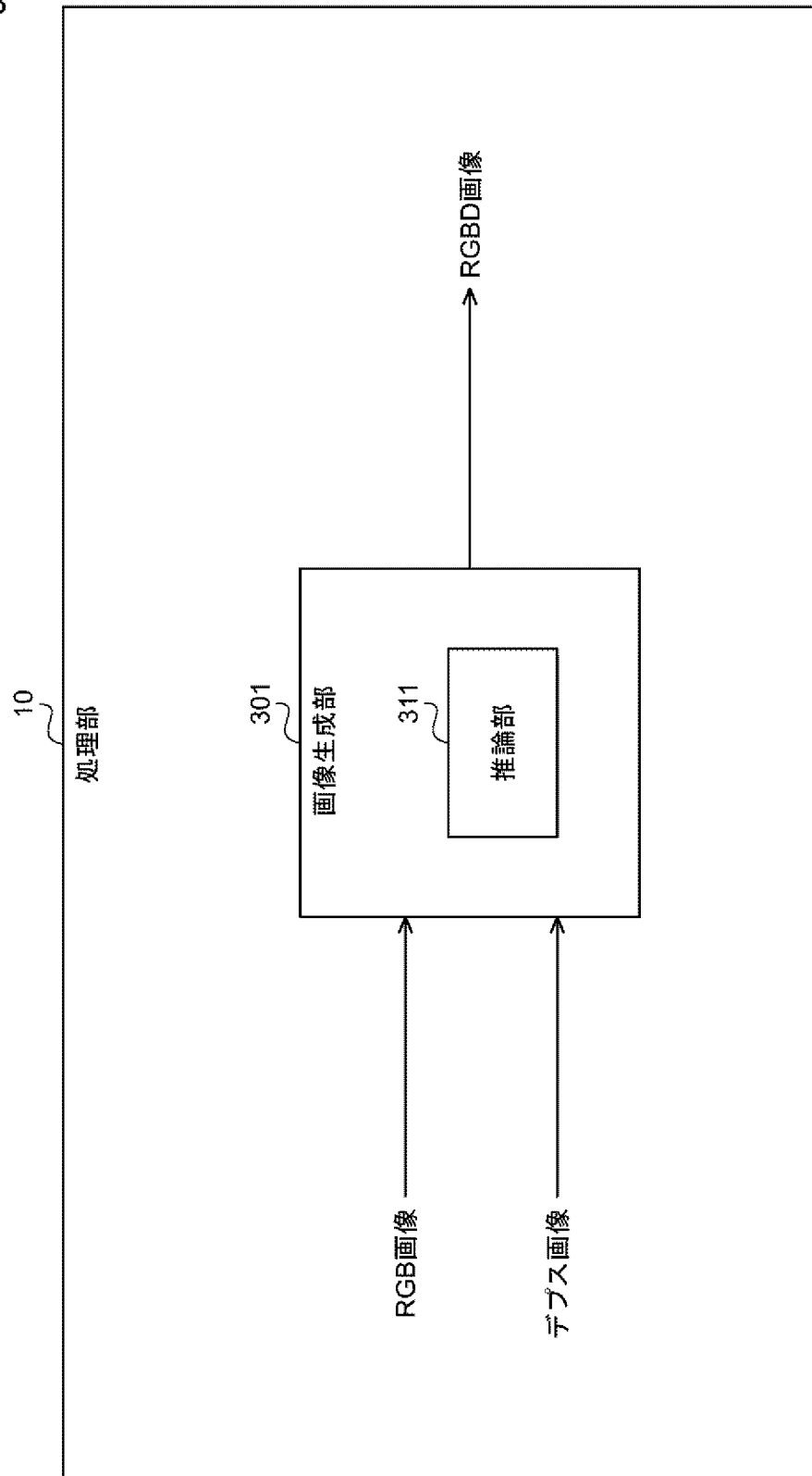
[図11]
FIG.11



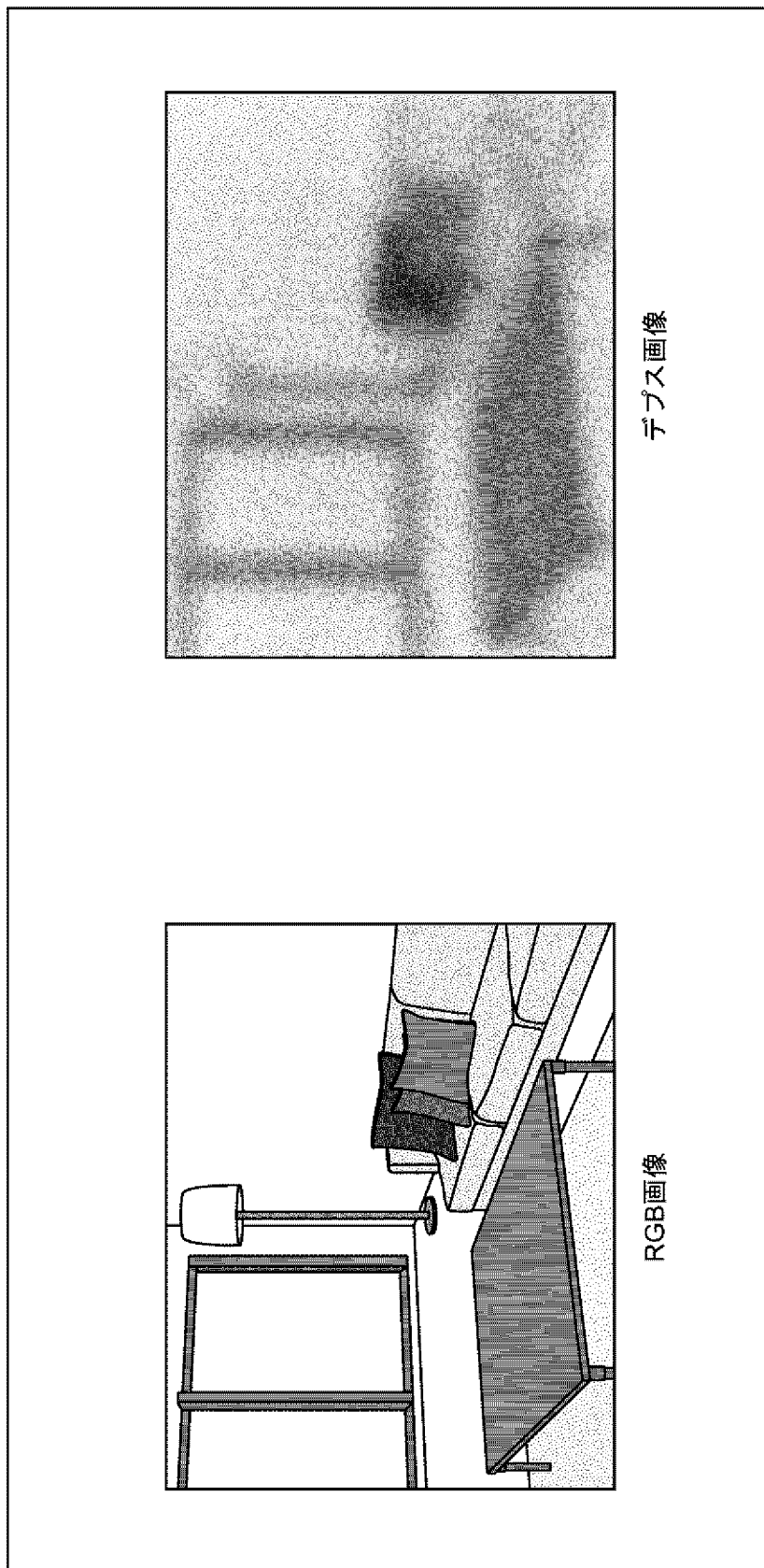
[図12]
FIG.12



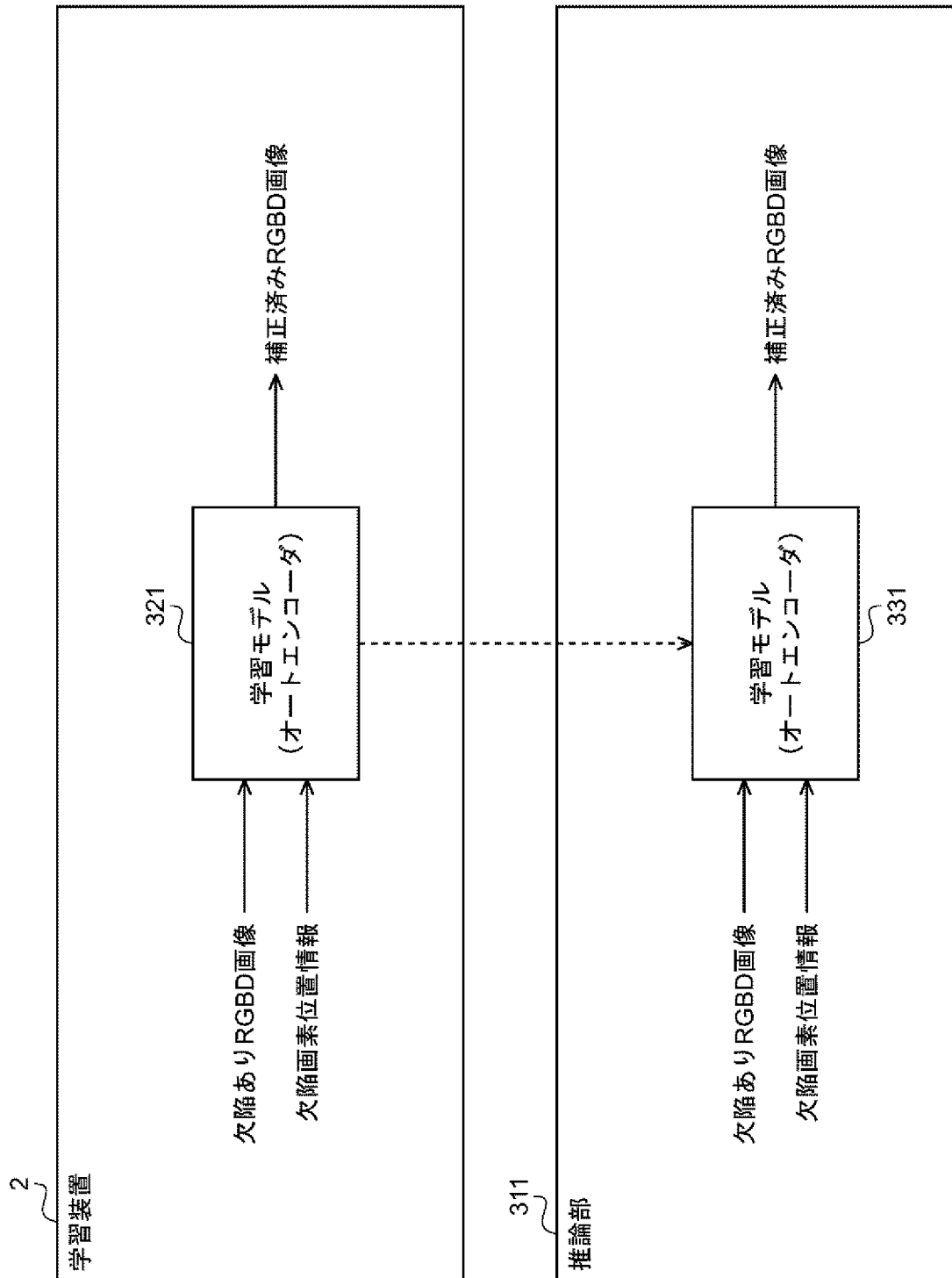
[図13]
FIG.13



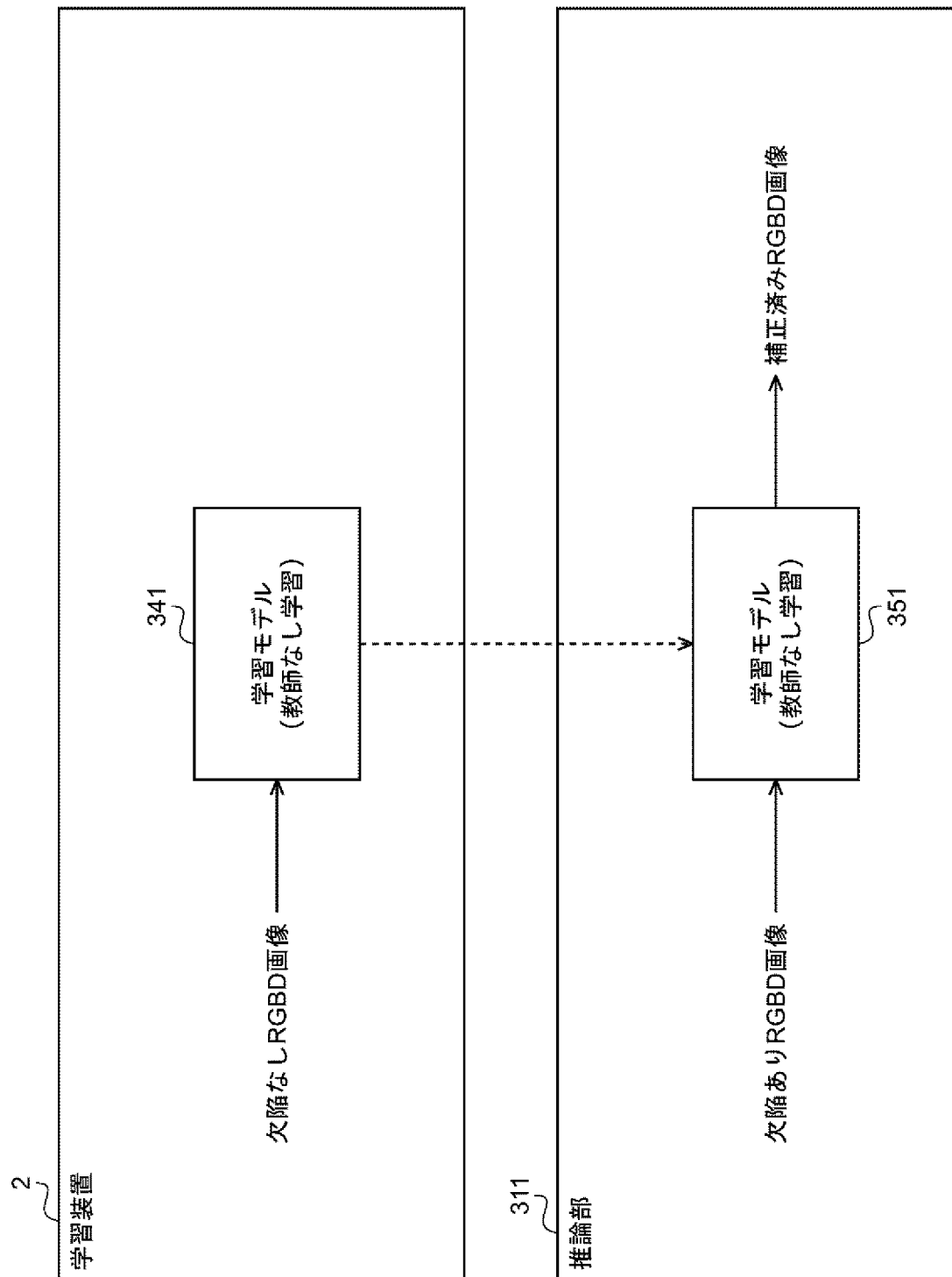
[図14]
FIG.14



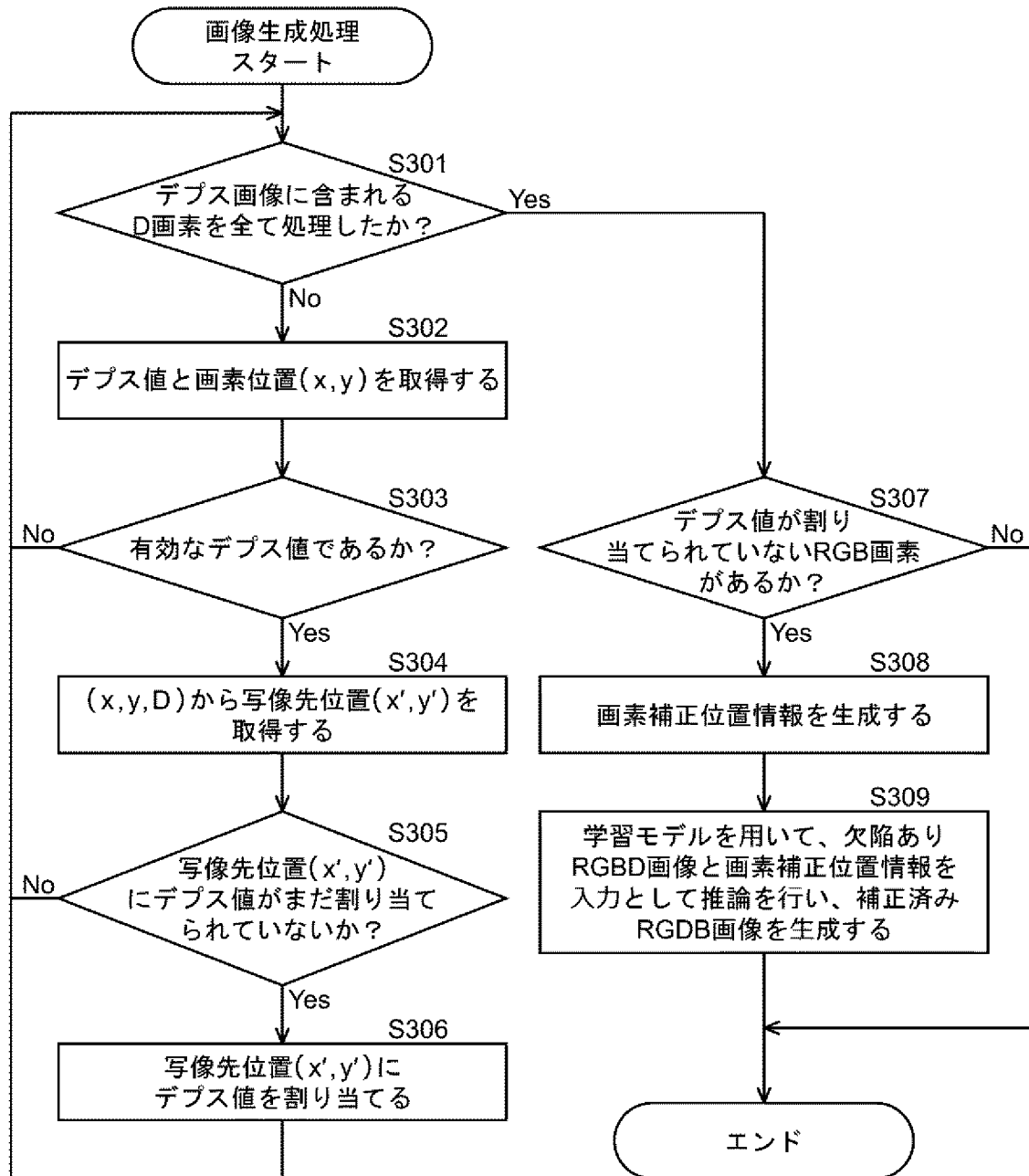
[図15]
FIG.15



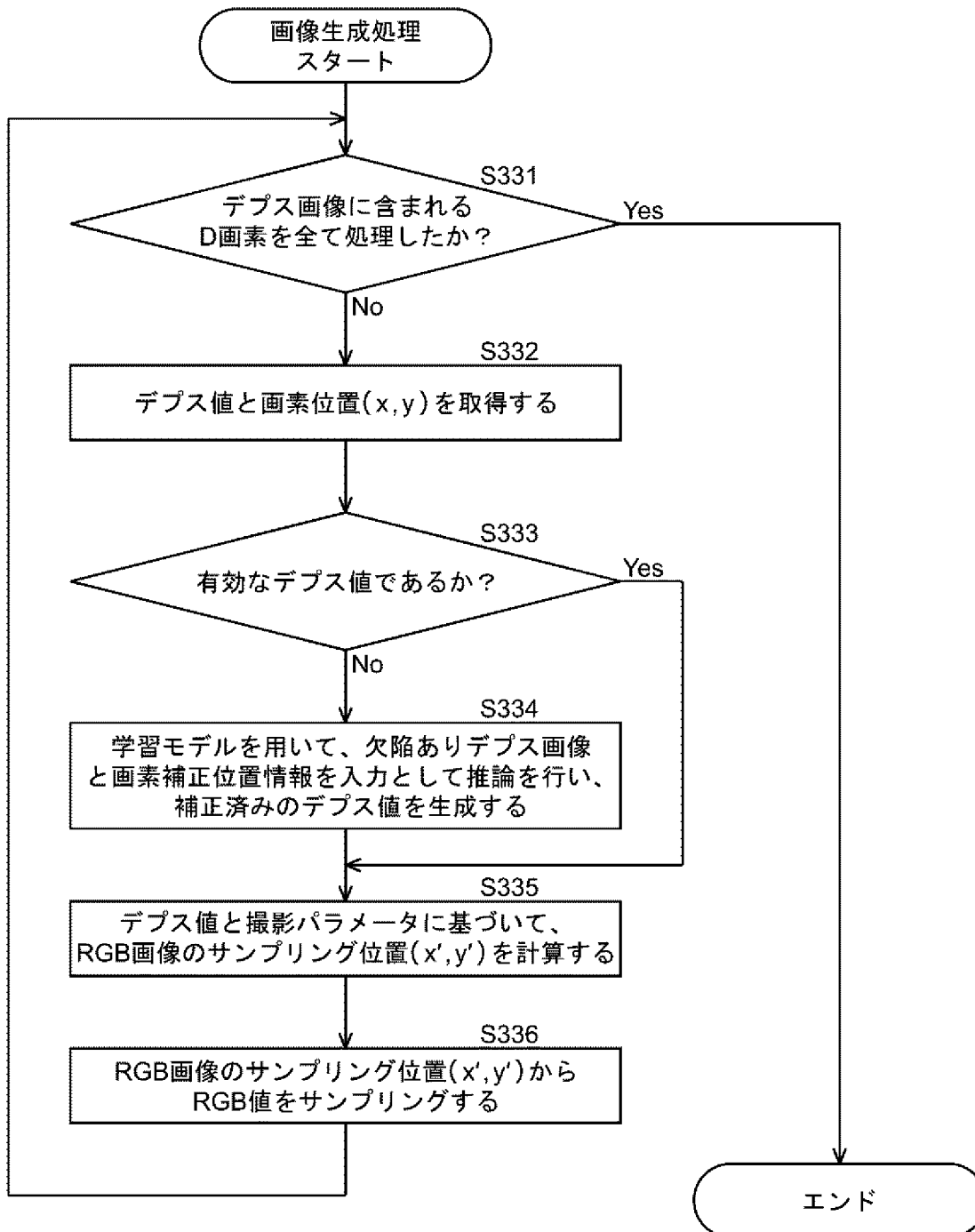
[図16]
FIG.16



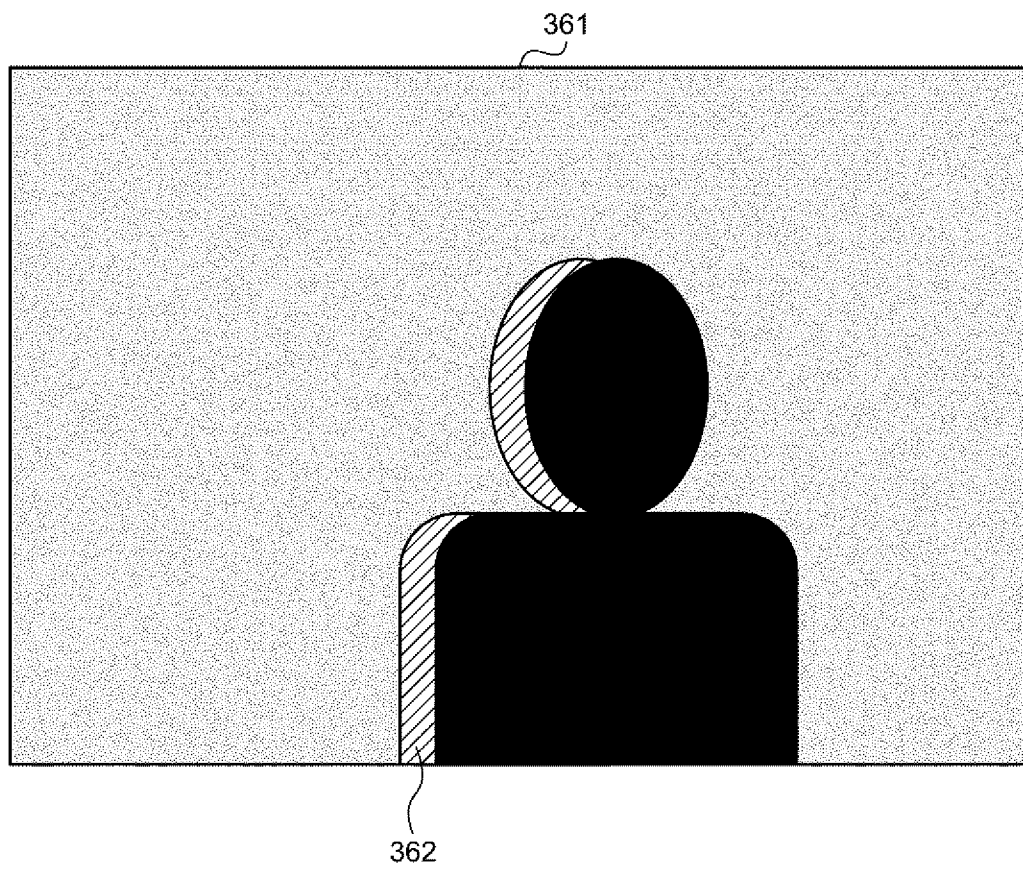
[図17]
FIG.17



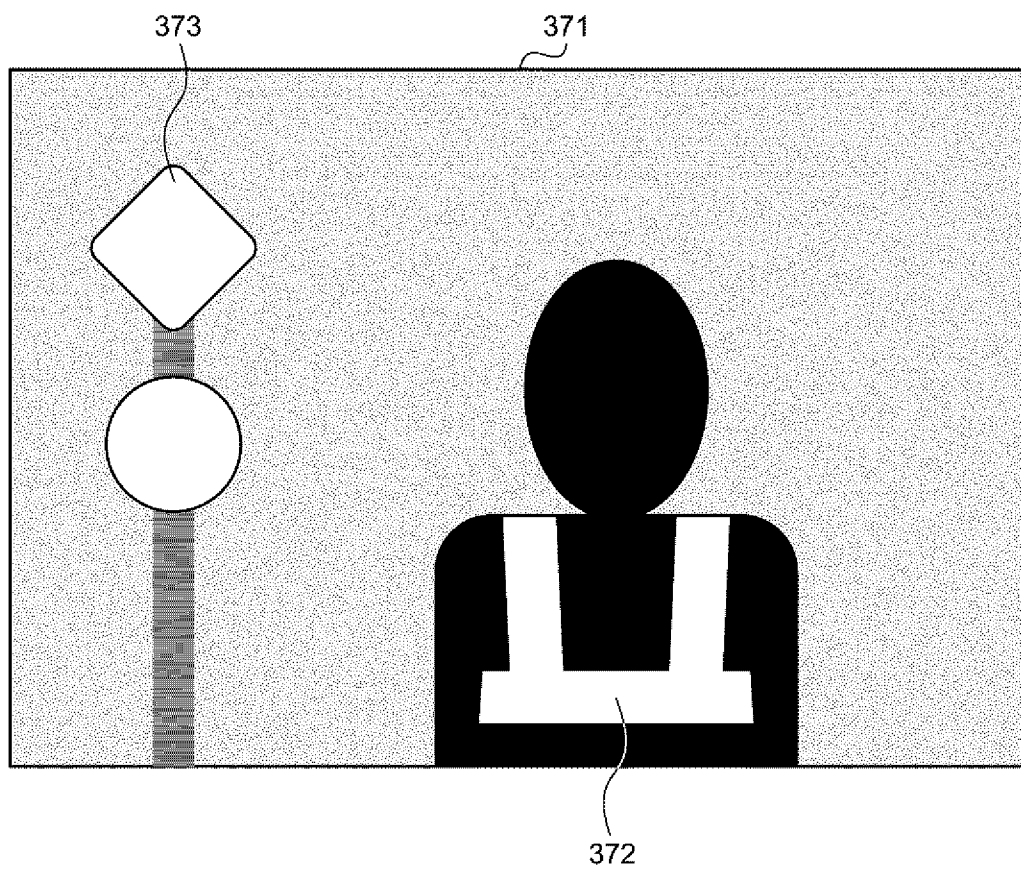
[図18]
FIG.18



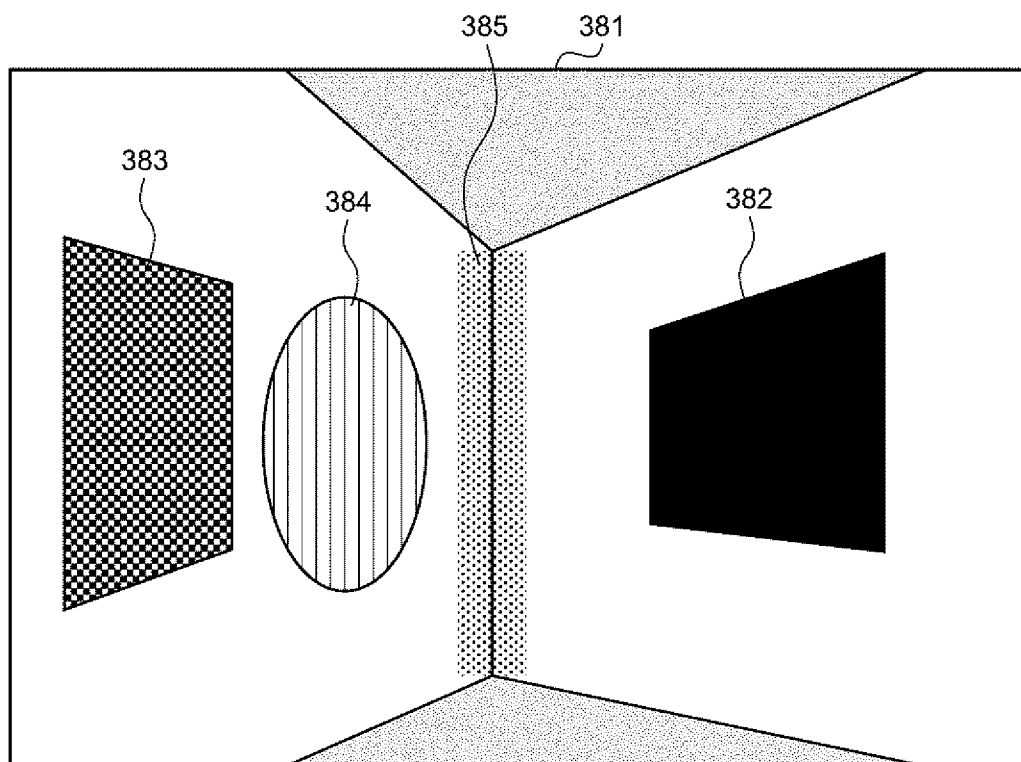
[図19]
FIG.19



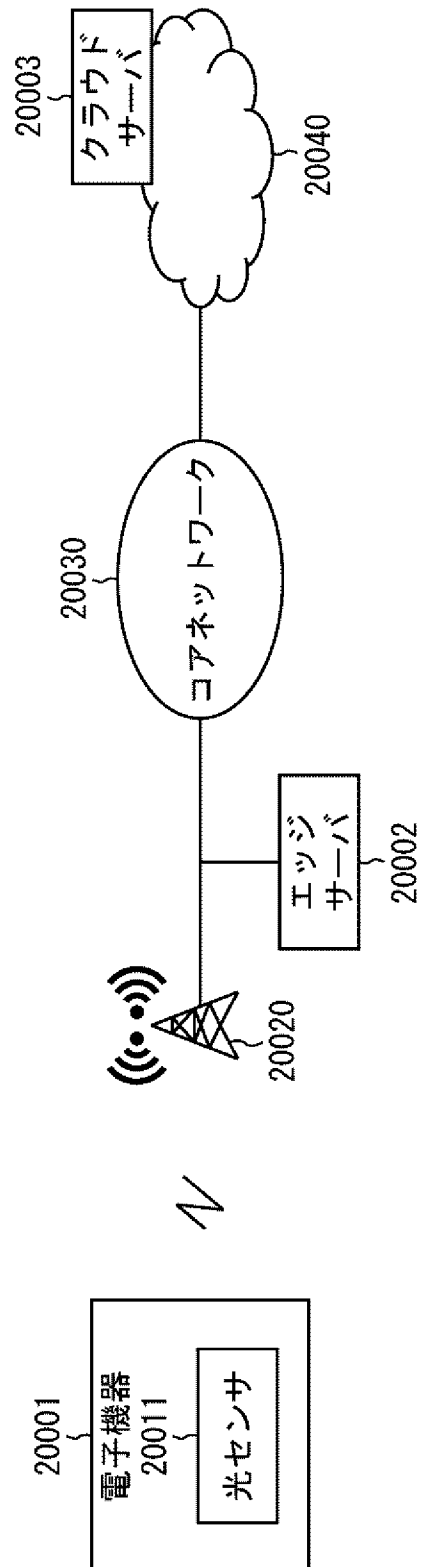
[図20]
FIG.20



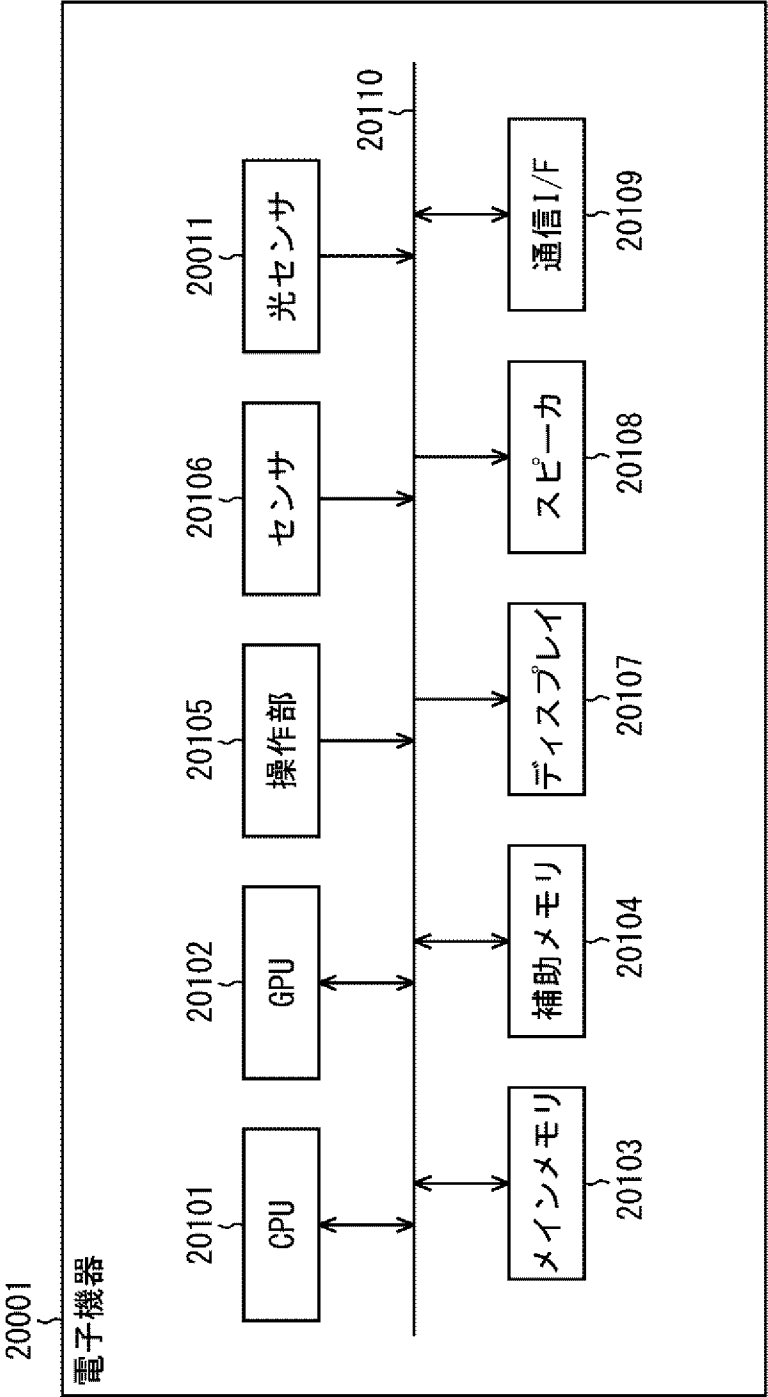
[図21]
FIG.21



[図22]

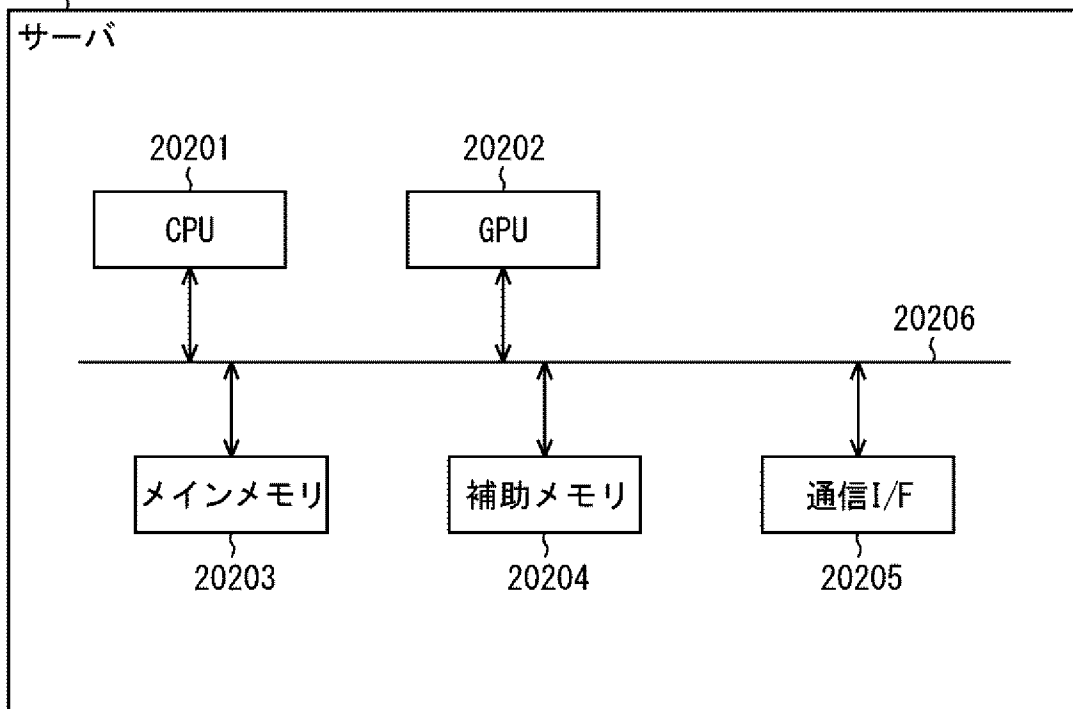


[図23]

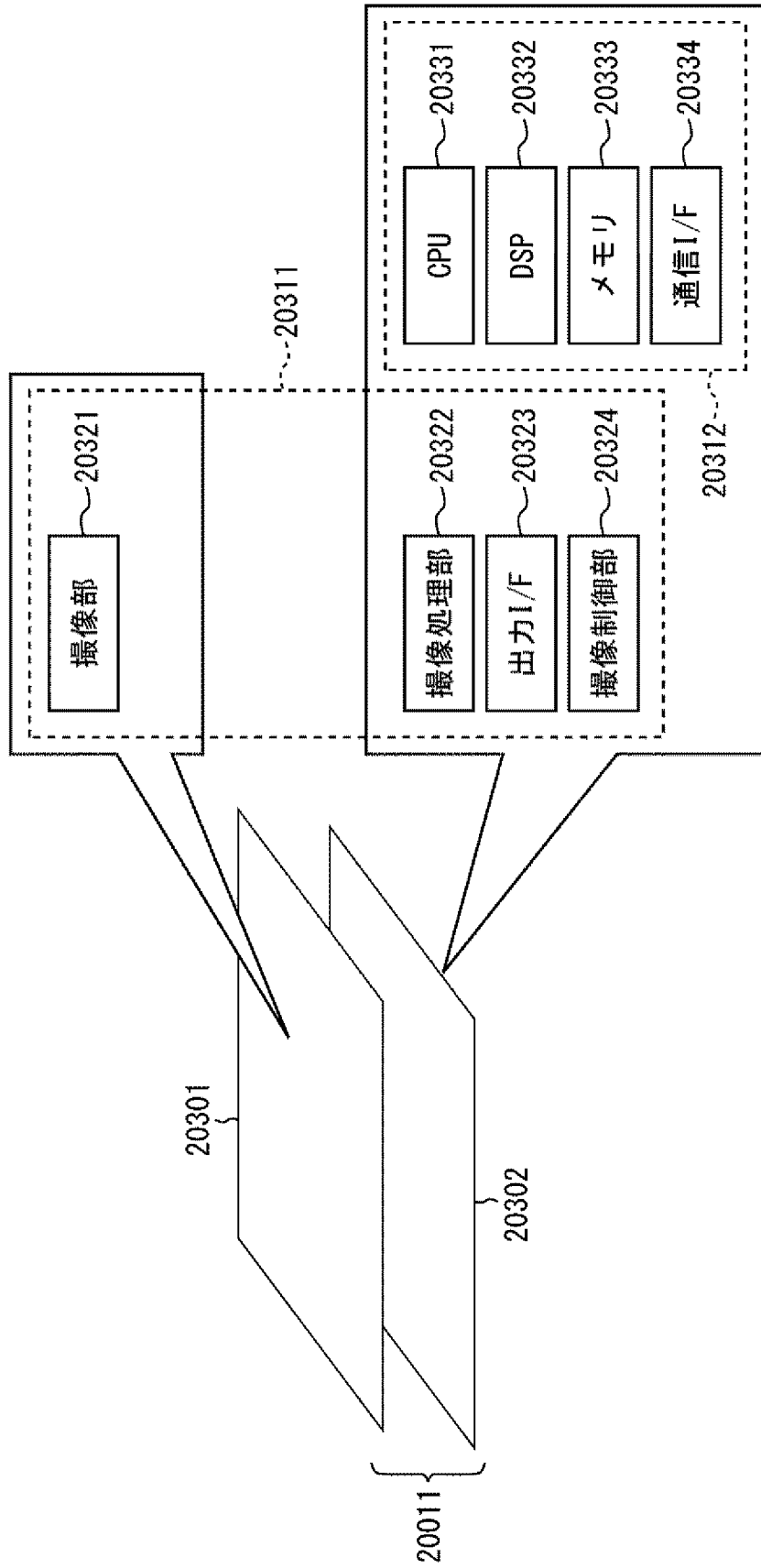


[図24]

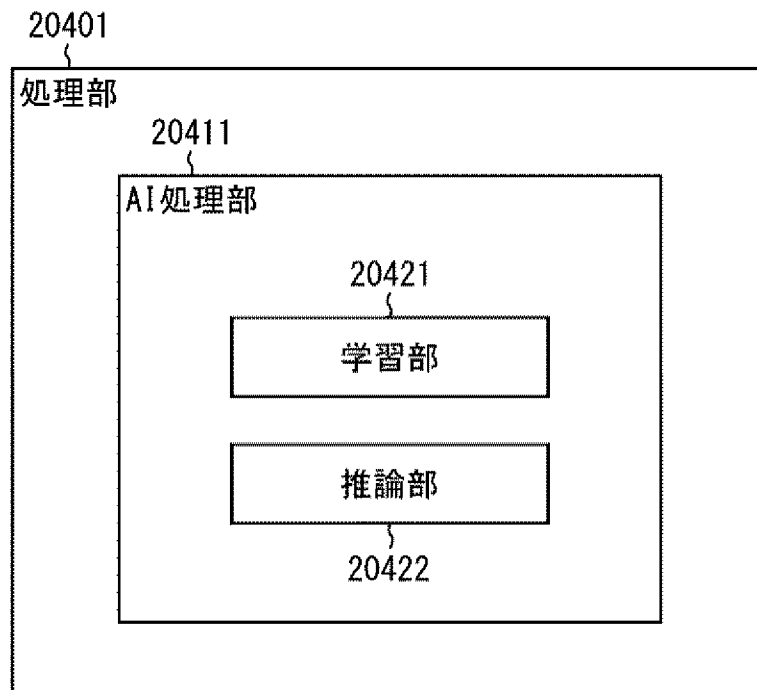
20002 (20003)



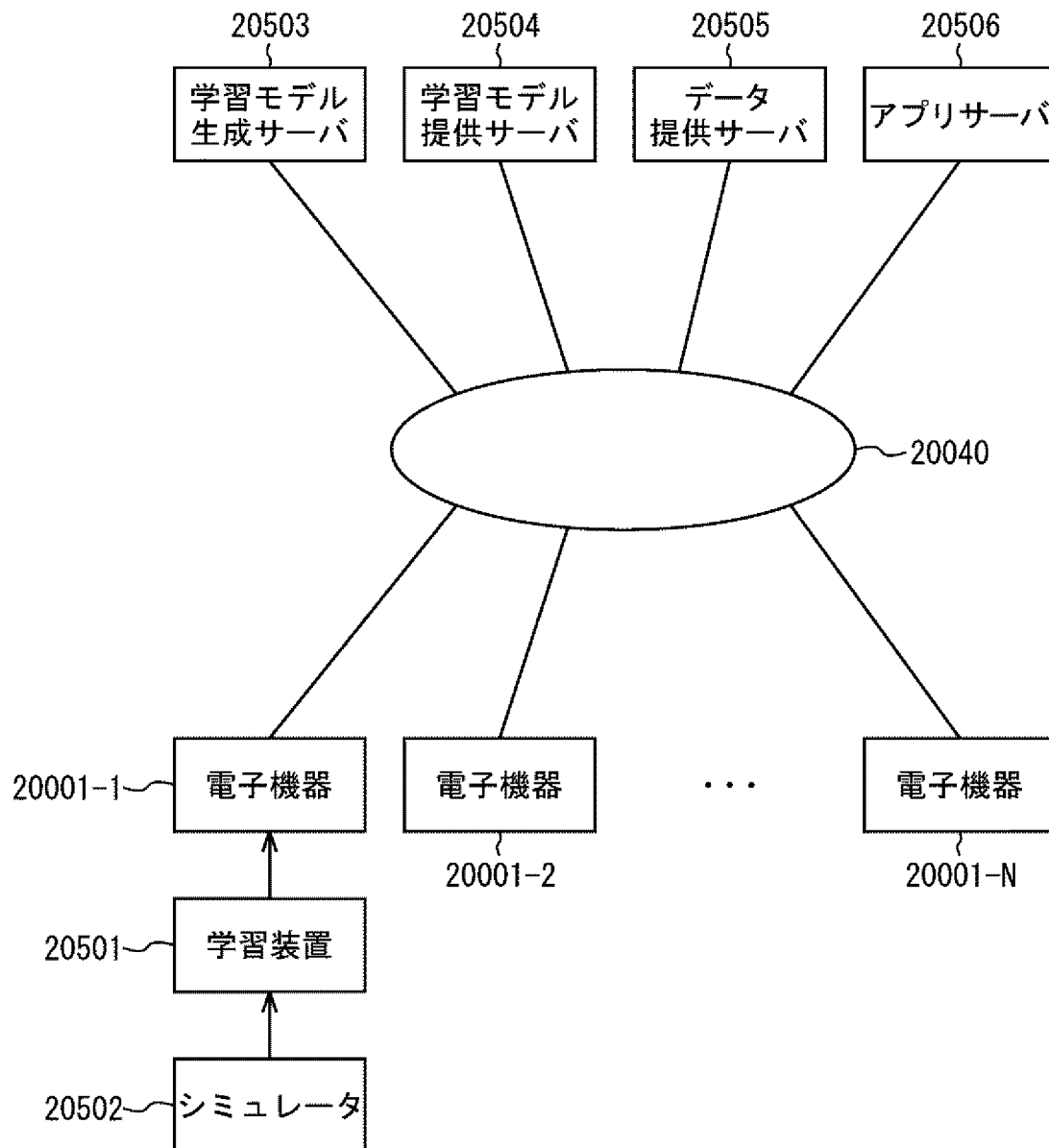
[図25]



[図26]



[図27]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2022/001918

A. CLASSIFICATION OF SUBJECT MATTER**G06T 7/00**(2017.01)i; **G01S 17/86**(2020.01)i; **G01S 17/89**(2020.01)i

FI: G06T7/00 350C; G01S17/89; G01S17/86

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; G01S17/86; G01S17/89

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 1971-2022
 Registered utility model specifications of Japan 1996-2022
 Published registered utility model applications of Japan 1994-2022

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.
X	CN 111626086 A (BEIJING SENSETIME SCIENCE TECH LTD.) 04 September 2020 (2020-09-04) paragraphs [0087], [0099]-[0100], [0113], [0148]-[0153]	1-19
A	paragraphs [0087], [0099]-[0100], [0113], [0148]-[0153]	20-28
A	JP 2020-149162 A (FUJITSU LTD.) 17 September 2020 (2020-09-17) abstract	1-28
A	JP 2019-016275 A (CANON INC.) 31 January 2019 (2019-01-31) abstract	1-28
A	US 10907960 B1 (OUTSIGHT SA) 02 February 2021 (2021-02-02) abstract	1-28
A	JP 2020-089947 A (SONY CORP.) 11 June 2020 (2020-06-11) paragraphs [0068]-[0069]	1-28
A	JP 2019-510325 A (MITSUBISHI ELECTRIC CORP.) 11 April 2019 (2019-04-11) paragraph [0056]	1-28

☐ 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

28 March 2022

Date of mailing of the international search report

05 April 2022

Name and mailing address of the ISA/JP

Japan Patent Office (ISA/JP)
 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/JP2022/001918

Patent document cited in search report			Publication date (day/month/year)	Patent family member(s)	Publication date (day/month/year)
CN	111626086	A	04 September 2020	(Family: none)	
JP	2020-149162	A	17 September 2020	(Family: none)	
JP	2019-016275	A	31 January 2019	(Family: none)	
US	10907960	B1	02 February 2021	(Family: none)	
JP	2020-089947	A	11 June 2020	WO 2020/115972 A1 paragraphs [0068]-[0069]	
JP	2019-510325	A	11 April 2019	WO 2017/209307 A1 paragraph [0056]	
				US 2017/0351935 A1	

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; G01S 17/86(2020.01)i; G01S 17/89(2020.01)i FI: G06T7/00 350C; G01S17/89; G01S17/86														
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; G01S17/86; G01S17/89														
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>1971-2022年</td> </tr> <tr> <td>日本国実用新案登録公報</td> <td>1996-2022年</td> </tr> <tr> <td>日本国登録実用新案公報</td> <td>1994-2022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	1971-2022年	日本国実用新案登録公報	1996-2022年	日本国登録実用新案公報	1994-2022年				
日本国実用新案公報	1922-1996年													
日本国公開実用新案公報	1971-2022年													
日本国実用新案登録公報	1996-2022年													
日本国登録実用新案公報	1994-2022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												
X	CN 111626086 A (BEIJING SENSETIME SCIENCE TECH LTD.) 04.09.2020 (2020-09-04) 段落0087, 0099-0100, 0113, 0148-0153	1-19												
A	段落0087, 0099-0100, 0113, 0148-0153	20-28												
A	JP 2020-149162 A (富士通株式会社) 17.09.2020 (2020-09-17) 要約	1-28												
A	JP 2019-016275 A (キャノン株式会社) 31.01.2019 (2019-01-31) 要約	1-28												
A	US 10907960 B1 (Outsight SA) 02.02.2021 (2021-02-02) 要約	1-28												
A	JP 2020-089947 A (ソニー株式会社) 11.06.2020 (2020-06-11) 段落0068-0069	1-28												
A	JP 2019-510325 A (三菱電機株式会社) 11.04.2019 (2019-04-11) 段落0056	1-28												
☐ 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、一般的な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、一般的な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、一般的な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 28.03.2022	国際調査報告の発送日 05.04.2022													
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													

国際調査報告
 パテントファミリーに関する情報

国際出願番号

PCT/JP2022/001918

引用文献			公表日	パテントファミリー文献	公表日
CN	111626086	A	04.09.2020	(ファミリーなし)	
JP	2020-149162	A	17.09.2020	(ファミリーなし)	
JP	2019-016275	A	31.01.2019	(ファミリーなし)	
US	10907960	B1	02.02.2021	(ファミリーなし)	
JP	2020-089947	A	11.06.2020	W0 2020/115972 A1	
				段落0068-0069	
JP	2019-510325	A	11.04.2019	W0 2017/209307 A1	
				段落0056	
				US 2017/0351935 A1	