

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7242185号
(P7242185)

(45)発行日 令和5年3月20日(2023.3.20)

(24)登録日 令和5年3月10日(2023.3.10)

(51)国際特許分類 F I
G 0 6 T 1/00 (2006.01) G 0 6 T 1/00 5 1 0
G 0 6 T 1/40 (2006.01) G 0 6 T 1/40

請求項の数 12 (全17頁)

(21)出願番号	特願2018-1552(P2018-1552)	(73)特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22)出願日	平成30年1月10日(2018.1.10)	(74)代理人	100110412 弁理士 藤元 亮輔
(65)公開番号	特開2019-121252(P2019-121252 A)	(74)代理人	100104628 弁理士 水本 敦也
(43)公開日	令和1年7月22日(2019.7.22)	(74)代理人	100121614 弁理士 平山 倫也
審査請求日	令和2年12月22日(2020.12.22)	(72)発明者	日浅 法人 東京都大田区下丸子3丁目30番2号 キヤノン株式会社内
		審査官	山口 大志

最終頁に続く

(54)【発明の名称】 画像処理方法、画像処理装置、画像処理プログラム、および、記憶媒体

(57)【特許請求の範囲】

【請求項1】

訓練画像と正解画像を取得する工程と、
前記訓練画像を多層のニューラルネットワークへ入力して出力画像を生成する工程と、
前記正解画像と前記出力画像のそれぞれにガンマ補正を実行し、ガンマ補正された前記正解画像とガンマ補正された前記出力画像との誤差を算出する工程と、
前記誤差に基づいて、前記ニューラルネットワークのネットワークパラメータを更新する工程と、を有することを特徴とする画像処理方法。

【請求項2】

前記訓練画像および前記正解画像はそれぞれ、周期的に配列された複数の色成分を有し、
前記画像処理方法は、前記訓練画像または前記正解画像の各色成分のみで構成される色成分画像を生成する工程を有し、
前記色成分画像を生成する工程は、前記訓練画像に対して前記ニューラルネットワークへの入力前に実行され、前記正解画像に対して前記誤差の算出前に実行されることを特徴とする請求項1に記載の画像処理方法。

10

【請求項3】

前記ガンマ補正の実行前に、前記出力画像の信号値を所定の範囲でクリッピングする工程を更に有することを特徴とする請求項1または2に記載の画像処理方法。

【請求項4】

前記正解画像は、前記訓練画像よりも高い解像度またはコントラストを有する画像であ

20

ることを特徴とする請求項 1 乃至 3 のいずれか 1 項に記載の画像処理方法。

【請求項 5】

前記訓練画像と前記正解画像のそれぞれの信号値が取り得る下限値と上限値を取得する工程と、

前記下限値と前記上限値に基づいて、前記訓練画像と前記正解画像のそれぞれの前記信号値を規格化する工程と、を更に有することを特徴とする請求項 1 乃至 4 のいずれか 1 項に記載の画像処理方法。

【請求項 6】

ガンマ補正を指定する工程を更に有し、

前記誤差を算出する工程において、前記ガンマ補正を指定する工程で指定されたガンマ補正を前記出力画像と前記正解画像のそれぞれに実行することを特徴とする請求項 1 乃至 5 のいずれか 1 項に記載の画像処理方法。

10

【請求項 7】

前記ガンマ補正を指定する工程で指定された前記ガンマ補正に関する情報と、前記ネットワークパラメータを更新する工程で更新された前記ネットワークパラメータを記憶する工程を更に有することを特徴とする請求項 6 に記載の画像処理方法。

【請求項 8】

前記ネットワークパラメータを更新する工程において、前記正解画像と前記出力画像に複数のガンマ補正をそれぞれ実行して算出された複数の前記誤差に基づいて、前記ネットワークパラメータを更新することを特徴とする請求項 1 乃至 7 のいずれか 1 項に記載の画像処理方法。

20

【請求項 9】

RAW 画像である訓練画像と正解画像を取得する取得部と、

前記訓練画像を多層のニューラルネットワークへ入力して出力画像を生成する生成部と、

前記正解画像と前記出力画像のそれぞれにガンマ補正を実行し、ガンマ補正された前記正解画像とガンマ補正された前記出力画像との誤差を算出する算出部と、

前記誤差に基づいて、前記ニューラルネットワークのネットワークパラメータを更新する更新部と、を有することを特徴とする画像処理装置。

【請求項 10】

前記ネットワークパラメータを記憶する記憶部を更に有することを特徴とする請求項 9 に記載の画像処理装置。

30

【請求項 11】

請求項 1 乃至 8 のいずれか 1 項に記載の画像処理方法をコンピュータに実行させることを特徴とする画像処理プログラム。

【請求項 12】

請求項 11 に記載の画像処理プログラムを記憶していることを特徴とする記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ディープラーニングを用いた画像処理方法に関する。

40

【背景技術】

【0002】

特許文献 1 には、ガラス越しに被写体を撮影することで得られる入力画像から、畳み込みニューラルネットワークを用いてガラスの汚れや水滴の像が除去された画像を推定する方法が開示されている。特許文献 1 に開示された方法では、畳み込みニューラルネットワークにより、汚れ及び水滴の像の検出と、検出された領域のインペインティングが行われる。非特許文献 1 には、様々な回帰問題に対して汎用的に適用可能なネットワーク構成が開示されている。また非特許文献 1 には、ネットワークを用いて、入力画像のアップサンプリング、JPEG デブロッキング（圧縮ノイズの除去）、デノイジング、ノンブライントなデブラー、または、インペインティングを実行することが開示されている。

50

【先行技術文献】

【特許文献】

【0003】

【文献】米国特許第9373160号

X. Mao, C. Shen, Y. Yang, "Image Restoration Using Convolutional Auto-encoders with Symmetric Skip Connections", <https://arxiv.org/abs/1606.08921>.

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかし、特許文献1および非特許文献1に開示された方法では、入力画像がRAW画像である際に適切な推定を行うことができない。

【0005】

ところで、多層のニューラルネットワークにRAW画像を入力し、所望の出力画像を推定しようとする場合、ニューラルネットワークの学習もRAW画像を用いて行う必要がある。学習では、訓練用のRAW画像をネットワークへ入力して得られた出力と正解のRAW画像の誤差を最小化することで、ネットワークパラメータを最適化する。このため、学習時の誤差は、RAW画像の状態（光量と信号値が略線型の状態）で最小化される。しかし、実際にユーザが現像された出力画像を鑑賞する際は、ガンマ補正が実行されるため、信号値の大きさによって誤差が変化する。具体的には、推定画像の暗部において誤差が拡大され、推定精度が低下する。

【0006】

そこで本発明は、現像された画像の信号の大きさに対して、推定精度が影響されにくい多層のニューラルネットワークを取得可能な画像処理方法、画像処理装置、画像処理プログラム、および、記憶媒体を提供することを目的とする。

【課題を解決するための手段】

【0007】

本発明の一側面としての画像処理方法は、RAW画像である訓練画像と正解画像を取得する工程と、前記訓練画像を多層のニューラルネットワークへ入力して出力画像を生成する工程と、前記正解画像と前記出力画像のそれぞれにガンマ補正を実行し、ガンマ補正された前記正解画像とガンマ補正された前記出力画像との誤差を算出する工程と、前記誤差に基づいて、前記ニューラルネットワークのネットワークパラメータを更新する工程とを有する。

【0008】

本発明の他の側面としての画像処理装置は、RAW画像である訓練画像と正解画像を取得する取得部と、前記訓練画像を多層のニューラルネットワークへ入力して出力画像を生成する生成部と、前記正解画像と前記出力画像のそれぞれにガンマ補正を実行し、ガンマ補正された前記正解画像とガンマ補正された前記出力画像との誤差を算出する算出部と、前記誤差に基づいて、前記ニューラルネットワークのネットワークパラメータを更新する更新部とを有する。

【0009】

本発明の他の側面としての画像処理プログラムは、前記画像処理方法をコンピュータに実行させる。

【0010】

本発明の他の側面としての記憶媒体は、前記画像処理プログラムを記憶している。

【0011】

本発明の目的及び特徴は、以下の実施例において説明される。

【発明の効果】

【0012】

10

20

30

40

50

本発明によれば、現像された画像の信号の大きさに対して、推定精度が影響されにくい多層のニューラルネットワークを取得可能な画像処理方法、画像処理装置、画像処理プログラム、および、記憶媒体を提供することができる。

【図面の簡単な説明】

【0013】

【図1】各実施例における畳み込みニューラルネットワークを示す図である。

【図2】ガンマ補正の説明図である。

【図3】実施例1における画像処理システムのブロック図である。

【図4】実施例1における画像処理システムの外觀図である。

【図5】各実施例における学習工程のフローチャートである。

【図6】各実施例における画像の色成分に関する説明図である。

【図7】各実施例における推定工程のフローチャートである。

【図8】実施例2における画像処理システムのブロック図である。

【図9】実施例2における画像処理システムの外觀図である。

【発明を実施するための形態】

【0014】

以下、本発明の実施例について、図面を参照しながら詳細に説明する。各図において、同一の部材については同一の参照符号を付し、重複する説明は省略する。

【0015】

まず、以下に用語を定義する。本発明は、ディープラーニングによって回帰問題を解き、入力画像から様々な出力画像を推定する方法に関する。ディープラーニングとは、多層のニューラルネットワークを用いた機械学習である。大量の訓練画像とそれに対応する正解画像（得たい出力）のペアから、ネットワークパラメータ（重みとバイアス）を学習することで、未知の入力画像に対しても高精度な推定が可能となる。

【0016】

多層のニューラルネットワークを用いた画像処理には、ネットワークパラメータ（重みとバイアス）を更新するための処理工程と、更新されたパラメータを用いて未知の入力に対して推定を行う処理工程の二つが存在する。以下、前者を学習工程と呼び、後者を推定工程と呼ぶ。本発明は、学習工程において特徴を有する。

【0017】

次に、学習工程と推定工程における画像の名称を定める。ネットワークへ入力する画像を入力画像とし、特に学習工程の際に用いる、正解画像が既知の入力画像を訓練画像と呼称する。ネットワークから出力された画像を出力画像とし、特に推定工程の際の出力画像を推定画像と呼称する。ネットワークの入力画像と、正解画像はRAW画像である。ここでRAW画像とは、撮像素子から出力された未現像の画像データであり、各画素の光量と信号値とが略線型の関係にある。ユーザが画像を鑑賞する前にRAW画像は現像されるが、その際にガンマ補正が実行される。ガンマ補正は、例えば入力の信号値を冪乗する処理であり、その冪指数として $1/2$ 、 2 などが用いられる。

【0018】

また出力画像も、推定によって正解画像に準ずる画像として生成されるため、RAW画像の性質を有する。推定工程には、様々な処理が含まれる。例えば、アップサンプリング、デノイジング、圧縮ノイズの除去、デプラー（ぼけ補正）、インペインティング、デモザイキング、ディヘイズ（Dehaze）、高階調化、色成分の推定（モノクロのカラー化など）、リライティング（照明環境の変更）がある。

【0019】

次に、本発明の要旨を簡単に述べる。本発明は、RAW画像を入力とする多層のニューラルネットワークの学習工程において、ガンマ補正の影響を考慮する。これにより、現像後の推定画像が信号値の大きさに依らず、おおよそ一定の推定精度を保つことができる。特に、高解像化（アップサンプリングやデプラー）や高コントラスト化（ディヘイズ）で出やすいアンダーシュートやリングングが抑制される。

10

20

30

40

50

【 0 0 2 0 】

学習時にガンマ補正を考慮する方法として、本発明では最小化する誤差を算出する前に、ネットワークの出力画像と正解画像に対してガンマ補正を実行する。この方法以外に、正解画像とネットワークに入力する前の入力画像をガンマ補正する方法、又はネットワーク内でガンマ補正相当の処理を行う方法が考えられる。しかし、本発明の手法では、出力画像（推定画像を含む）において光量と信号値が略線型の関係となるのに対し、後者の二つの手法は出力画像における光量と信号値の関係がガンマ補正後のものとなる。このため、後者の二つの手法では、出力画像に対して通常の現像処理を行う場合、その前に逆ガンマ補正を実行しなければならない。しかし、ガンマ補正には様々な形があり、高輝度部の階調潰れが問題になることがある。さらに、逆ガンマ補正の解が一意的に定まらない場合もある。その例を図 2 に示す。

10

【 0 0 2 1 】

図 2 は、ガンマ補正の説明図であり、ガンマ補正に用いられるガンマカーブの一例を表す。図 2 において、横軸はガンマ補正前の信号値、縦軸はガンマ補正後の信号値をそれぞれ示し、それぞれの信号値は規格化されている。図 2 のガンマカーブでは、高輝度部においてガンマ補正後の信号が同一値になっている。このため、逆ガンマ補正時にはこの部分の信号値が一意的に定まらない。このような高輝度部での階調潰れは、ユーザが現像時に露出の調整などの編集作業を行った際に正しい画像が得られない要因となる。この対策として、例えばユーザが露出調整を行った場合、露出調整した状態の RAW 画像を入力画像として、ネットワークへ入力し直すことが考えられる。しかしこの方法では、ユーザが露出調整する度にネットワーク内の演算を再実行する必要があり、計算負荷が増大する。そこで、以下の各実施例の手法を用いることで、高輝度の階調潰れや演算負荷の増大を抑制することができる。

20

【 実施例 1 】

【 0 0 2 2 】

まず、本発明の実施例 1 における画像処理システムについて説明する。本実施例の画像処理システムでは、学習工程を実行する装置（パラメータ学習装置 1 0 1）と、推定工程を行う装置（画像編集装置 1 1 1）とが個別に存在する。本実施例における推定工程は、収差・回折によるぼけの補正（デブラーの一種）であるが、これに限定されるものではない。

30

【 0 0 2 3 】

図 3 および図 4 を参照して、本実施例における画像処理システムについて説明する。図 3 は、画像処理システム 1 0 0 のブロック図である。図 4 は、画像処理システム 1 0 0 の外観図である。図 3 および図 4 に示されるように、画像処理システム 1 0 0 は、パラメータ学習装置 1 0 1、画像編集装置 1 1 1、撮像装置 1 2 1、表示装置 1 2 3、記録媒体 1 2 2、および、出力装置 1 2 4 を備えて構成される。

【 0 0 2 4 】

パラメータ学習装置 1 0 1 は、学習工程を実行する画像処理装置であり、記憶部 1 0 2 および学習部 1 0 3 を有する。学習部 1 0 3 は、推定工程に必要なネットワークパラメータ（重みとバイアス）を算出する学習工程を実行する。この際、学習部 1 0 3 は、記憶部 1 0 2 に記憶された複数の訓練画像と正解画像を用いる。訓練画像と正解画像は、収差・回折によるぼけの有無が異なる画像である。

40

【 0 0 2 5 】

本実施例において、学習部 1 0 3 は、取得部 1 0 3 a、生成部 1 0 3 b、算出部 1 0 3 c、および、更新部 1 0 3 d を有する。取得部 1 0 3 a は、訓練画像と正解画像を取得する。本実施例において、訓練画像と正解画像はそれぞれ RAW 画像である。生成部 1 0 3 b は、訓練画像を多層のニューラルネットワークへ入力して出力画像を生成する。算出部 1 0 3 c は、正解画像と出力画像のそれぞれにガンマ補正を実行し、ガンマ補正された正解画像と出力画像の誤差を算出する。更新部 1 0 3 d は、算出部 1 0 3 c により算出された誤差に基づいて、ニューラルネットワークのネットワークパラメータを更新する。なお

50

、学習工程に関する詳細は、フローチャートを用いて後述する。学習されたネットワークパラメータは、記憶部 102 に記憶される。

【0026】

画像編集装置 111 は、推定工程を実行する装置であり、記憶部 112 および推定部 113 を有する。撮像装置 121 は、CMOS センサや CCD センサなどの撮像素子 125 を有し、光学系（撮像光学系）により形成された光学像を光電変換して撮像画像を出力する。画像編集装置 111 は、有線または無線で接続された撮像装置 121、または半導体メモリなどの記録媒体 122 から、RAW 画像である撮像画像を取得し、記憶部 112 に記憶する。推定部 113 は、撮像画像から入力画像を取得して多層のニューラルネットワークへ入力し、収差・回折によるぼけが補正された推定画像を生成する。画像編集装置 111 とパラメータ学習装置 101 は、有線または無線で互いに接続されており、画像編集装置 111 は、記憶部 102 に記憶されたネットワークパラメータを読み出し、推定工程に用いる。また画像編集装置 111 は、推定画像から収差・回折によるぼけが補正された撮像画像を生成し、補正された撮像画像は記録媒体 122、表示装置 123、または、出力装置 124 の少なくとも一つに出力される。表示装置 123 は、例えば液晶ディスプレイやプロジェクタなどである。ユーザは、表示装置 123 を介して、編集途中の画像を確認しながら作業を行うことができる。出力装置 124 は、プリンタなどである。画像編集装置 111 は、その他に現像処理などを行う機能を有していてもよい。

10

【0027】

次に、図 5 を参照して、パラメータ学習装置 101 の学習部 103 で実行される学習工程に関して説明する。図 5 は、学習工程のフローチャートである。図 5 の各ステップは、学習部 103 の各部（取得部 103a、生成部 103b、算出部 103c、および、更新部 103d）により実行される。

20

【0028】

まず、ステップ S101 において、学習部 103（取得部 103a）は、記憶部 102 から複数の訓練画像と正解画像を取得する。訓練画像と正解画像はそれぞれ、未現像の RAW 画像である。訓練画像とは、収差・回折によるぼけが作用した画像である。正解画像は、訓練画像と同一の被写体が写っており、且つぼけのない（または非常に小さい）画像である。訓練画像は、収差・回折の影響が少ない RAW 画像（正解画像）に対して、シミュレーションで収差・回折によるぼけを付与することで生成できる。また訓練画像と正解画像は、同一の被写体を異なる F 値で撮影することでも生成可能である。この場合、例えば、絞りが開放に近く、収差の影響が大きい画像が訓練画像であり、絞りをある程度絞って収差と回折による劣化が抑えられた画像が正解画像となる。学習工程において含まれていない収差・回折によるぼけ、または被写体の構造（エッジなど）は、後の推定工程で精度良く推定することができない。このため学習部 103 は、様々な種類の被写体の構造や補正しようとする収差・回折のぼけが含まれるように、ステップ S101 では複数の訓練画像と正解画像を取得する。

30

【0029】

なお、収差・回折によるぼけ補正以外の処理に関しても、同様にシミュレーションで訓練画像と正解画像のペアを用意することで、学習工程を実行することができる。デノイジングに関しては、低ノイズの正解画像に想定されるノイズを付与することで、訓練画像を生成することができる。アップサンプリングに関しては、実施例 2 にて後述する。圧縮ノイズの除去に関しては、無圧縮または圧縮率の小さい正解画像を圧縮することで、訓練画像を生成することができる。収差・回折以外（デフォーカスぼけなど）のデブラーに関しては、ぼけの少ない正解画像に想定されるぼけを畳み込むことで、訓練画像を生成することができる。デフォーカスぼけの場合、距離に依存するため、複数の訓練画像と正解画像に異なる距離のデフォーカスぼけを畳み込むようにする。インペインティングに関しては、欠損のない正解画像に欠損を与えることで、訓練画像を生成することができる。デモザイキングに関しては、三板式の撮像素子などで撮像された正解画像を Bayer 配列などで再サンプリングすることで、訓練画像を生成することができる。ディヘイズに関しては

40

50

、霧や霽のない正解画像に対して散乱光を付与することで、訓練画像を生成することができる。霧や霽は、濃度や距離によって散乱光の強さが変化するため、異なる濃度や距離の散乱光に対して複数の訓練画像を生成する。高階調化に関しては、高階調な正解画像を低階調化することで訓練画像を得ることができる。色成分の推定に関しては、複数の色成分を有する正解画像の色成分を減らすことで、訓練画像を生成することができる。リライティングに関しては、正解画像の被写体における法線、形状、反射率の分布が既知であれば、シミュレーションで異なる光源環境の訓練画像を生成することができる。ただし、この場合、計測の負荷が大きいいため、実際に異なる照明環境で被写体を撮影して正解画像と訓練画像のペアを生成してもよい。

【 0 0 3 0 】

続いてステップ S 1 0 2 において、学習部 1 0 3 (取得部 1 0 3 a) は、学習工程で用いられるガンマ補正に関する情報を取得する。学習工程で用いられるガンマ補正を、推定画像の現像に用いるガンマ補正と合わせることで、補正後の信号値の大きさに依らないより安定した精度の推定が可能となる。本実施例では、様々なガンマ補正に対応するため、学習部 1 0 3 は複数のガンマ補正に対してそれぞれ学習工程を実行し、各ガンマ補正で最適化されたネットワークパラメータを記憶部 1 0 2 に記憶しておく。ガンマ補正に関する情報の取得は、記憶部 1 0 2 に記憶された複数のガンマ補正に関する情報から選択することができる。または、ユーザがガンマ補正の数式やルックアップテーブルなどを入力してもよい。なおガンマ補正とは、図 2 に示されるように、補正前後における信号値の関係を示すカーブ (ガンマカーブ) が傾き 1 の直線 (図 2 中の一点鎖線) 以上の位置に存在する処理である。ガンマ補正を表す関数 g の具体的な例としては、以下の式 (1) 乃至 (3) などが挙げられる。

【 0 0 3 1 】

【数 1】

$$g(s) = s^{1/\gamma} \quad (1)$$

【 0 0 3 2 】

【数 2】

$$g(s) = \log_{\alpha}(\beta \cdot s + 1) \quad (2)$$

【 0 0 3 3 】

【数 3】

$$g(s) = 1 - \exp(-\delta \cdot s) \quad (3)$$

【 0 0 3 4 】

式 (1) 乃至 (3) において、 s はガンマ補正前の規格化された信号値であり、 $\gamma > 1$ 、 $\alpha > 0$ 、 $\beta > 1$ 、 $\delta > 0$ は定数である。ただし、ガンマ補正は式 (1) 乃至 (3) の形に限定されず、区分線形関数やルックアップテーブルでもよい。なお、ステップ S 1 0 2 は、ステップ S 1 0 4 の前であれば、いつ実行してもよい。

【 0 0 3 5 】

続いてステップ S 1 0 3 において、学習部 1 0 3 (生成部 1 0 3 b) は、ステップ S 1 0 1 にて取得された複数の訓練画像のうち少なくとも一の訓練画像を選択し、選択された訓練画像をネットワークへ入力して出力画像を算出 (生成) する。複数の訓練画像の全てを選択する (訓練画像の全てをネットワークへ入力し、それら全ての出力を用いてネットワークパラメータを更新する) 場合をバッチ学習と呼ぶ。この方法は、訓練画像の数が増えるにつれて、演算負荷が膨大になる。一枚の訓練画像のみを選択する場合 (ネットワー

10

20

30

40

50

クパラメータの更新に一枚の訓練画像のみを用いて、更新ごとに異なる訓練画像を用いる)場合をオンライン学習と呼ぶ。この手法は、訓練画像の総数が増えても演算量が増大しないが、一枚の訓練画像に存在するノイズの影響を受けやすい。このため、複数の訓練画像から少数(ミニバッチ)を選択し、それらを用いてネットワークパラメータの更新を行なうミニバッチ法を用いることが好ましい。次の更新では、異なる小数の訓練画像を選択して用いる。この処理を繰り返すことにより、バッチ学習とオンライン学習の欠点を小さくすることができる。

【0036】

ここで、図1を参照して、多層のニューラルネットワークで行われる処理に関して説明する。図1は、畳み込みニューラルネットワーク(CNN)を示す図である。ただし本実施例は、これに限定されるものではなく、例えばCNNに残差ネットワークを採用することができ、または、GAN(Generative Adversarial Network)などを用いてもよい。なお図1では、簡単のため、入力する訓練画像201を一枚だけ描画しているが、実際には選択された複数の訓練画像それぞれに対して、出力画像が生成される。訓練画像201は、RAW画像を色成分ごとに三次元方向に配列した画像である。

10

【0037】

図6は、画像の色成分に関する説明図である。本実施例において、訓練画像は図6(A)に示されるようなBayer配列の画像である。ここでRGBは、それぞれ赤、緑、青を表す。図6(A)のBayer配列から、各色の成分だけを配列し直した構成が図6(B)である。Gは、G₁とG₂の2種類があるため、それぞれを抽出して配列する。図6(B)の四枚の画像を三次元方向に配列した4チャンネルの画像が、図1における訓練画像201である。この作業は必ずしも必要ではないが、収差・回折は波長によって変化するため、同一のぼけを持つ色成分を配列させた方が補正しやすい。また、RGBが同一次元内に配列されていると、局所的に異なる明るさを有する画素が混合されるため、推定精度が低下しやすい。このため、訓練画像を色成分ごとに分離することが好ましい。なお、ここではBayer配列の場合を示しているが、その他の配列(ハニカム構造など)に関しても同様である。なお、モノクロの場合、色成分の再配列は実行しなくてよい。また本実施例では、複数の色成分を一括で学習、推定する例を示すが、各色で個別に学習、推定してもよい。図1では描画を簡略化するため、訓練画像201を4×4の4チャンネル画像としているが、縦横の画像サイズはこれに限定されるものではない。

20

30

【0038】

本実施例において、訓練画像および正解画像はそれぞれ、周期的に配列された複数の色成分を有し、訓練画像または正解画像の各色成分のみで構成される色成分画像を生成する工程を設けてもよい。ここで、色成分画像を生成する工程は、訓練画像に対してニューラルネットワークへの入力前に実行され、正解画像に対して誤差の算出前に実行される。

【0039】

CNNは複数の層構造になっており、各層で線型変換と非線型変換が実行される。線型変換は、入力された画像(または特徴マップ)とフィルタの畳み込み、およびバイアス(図1中のbias)との和で表現される。各層におけるフィルタの重みとバイアスをネットワークパラメータと呼び、その値を学習工程によって更新する。非線形変換は、活性化関数(Activation Function)と呼ばれる非線型関数による変換である(図1中のAF)。活性化関数の例としては、シグモイド関数やハイパボリックタンジェント関数などがあり、本実施例では以下の式(4)で表されるReLU(Rectified Linear Unit)が用いられる。

40

【0040】

【数4】

$$f(x) = \max(x, 0) \quad (4)$$

50

【 0 0 4 1 】

式 (4) において、 \max は、引数のうち最大値を出力する MAX 関数を表す。

【 0 0 4 2 】

入力層に入力された訓練画像 2 0 1 は、第 1 畳み込み層で複数のフィルタ 2 0 2 それぞれとのコンボリューションと、バイアスとの和を取られる。フィルタ 2 0 2 それぞれのチャンネル数は、訓練画像 2 0 1 と一致し、訓練画像 2 0 1 のチャンネル数が 2 以上の場合、3 次元フィルタとなる (3 次元目がチャンネル数を表す)。なお、フィルタの縦横の大きさは任意である。コンボリューションと和の結果は、活性化関数によって非線形変換が施され、第 1 特徴マップ 2 0 3 が第 1 中間層に出力される。ここで、第 1 特徴マップ 2 0 3 のチャンネル数 (3 次元方向の配列数) は、フィルタ 2 0 2 の数と同じである。次に、第 2 畳み込み層へ第 1 特徴マップ 2 0 3 が入力され、前述と同様に複数のフィルタ 2 0 4 のそれぞれとのコンボリューションと、バイアスとの和が取られる。その結果を非線形変換し、以下同様に畳み込み層の数だけ繰り返す。一般に、畳み込み層が 3 層以上ある CNN が、ディープラーニングに該当する。最後の畳み込み層から出力された結果が、CNN の出力画像 2 1 1 である。なお、最後の畳み込み層では、活性化関数による非線形変換を実行しなくてもよい。

10

【 0 0 4 3 】

続いてステップ S 1 0 4 において、学習部 1 0 3 (算出部 1 0 3 c) は、出力画像 2 1 1 および正解画像 2 2 1 のそれぞれに対してガンマ補正を実行し、ガンマ補正された出力画像と正解画像との誤差を算出する。正解画像 2 2 1 も訓練画像 2 0 1 と同様に、色成分ごとに配列してチャンネル方向にスタックされている。本実施例において、学習部 1 0 3 は、以下の式 (5) を用いて誤差 L を算出する。

20

【 0 0 4 4 】

【 数 5 】

$$L = \frac{1}{2N} \sum_j^N |g(t_j) - g(y_j)|^2 \quad (5)$$

【 0 0 4 5 】

式 (5) において、 t は正解画像 2 2 1 の信号値、 y は出力画像 2 1 1 の信号値、 j は画素の番号、 N は総画素数、 g はガンマ補正を示す。式 (5) ではユークリッドノルムを用いているが、正解画像と出力画像の差異を表す値であれば、他の指標を用いてもよい。

30

【 0 0 4 6 】

続いてステップ S 1 0 5 において、学習部 1 0 3 (更新部 1 0 3 d) は、ステップ S 1 0 4 にて算出された誤差からネットワークパラメータの更新量を算出し、ネットワークパラメータを更新する。ここでは、誤差逆伝搬法 (Backpropagation) が用いられる。誤差逆伝搬法では、誤差の微分に基づいて更新量を算出する。ただし、本実施例はこれに限定されるものではない。

【 0 0 4 7 】

続いてステップ S 1 0 6 において、学習部 1 0 3 は、所定の終了条件を満たすか否か、すなわち、ネットワークパラメータの最適化が終了したか否かを判定する。ここで所定の終了条件とは、例えば、学習工程が既定の時間に達した場合、パラメータの更新回数が既定の回数に達した場合、パラメータ更新には用いない訓練画像と正解画像を用意しておき、その出力画像と正解画像の誤差が所定の値以下になった場合などである。または、ユーザが最適化終了を指示してもよい。所定の終了条件を満たさない場合、ステップ S 1 0 3 に戻り、学習部 1 0 3 は新たなミニバッチを取得してネットワークパラメータを更新する。一方、所定の終了条件を満たす場合、ステップ S 1 0 7 へ進む。

40

【 0 0 4 8 】

ステップ S 1 0 7 において、学習部 1 0 3 は、更新されたネットワークパラメータを記憶部 1 0 2 に出力する。記憶部 1 0 2 は、そのネットワークパラメータを記憶する。本実

50

施例において、学習部 103 は異なるガンマ補正ごとにネットワークパラメータを学習するため、記憶部 102 はネットワークパラメータとそれに対応するガンマ補正の情報とを合わせて記憶する。以上の学習工程により、現像された画像の信号値の大きさに対して、推定精度が影響されにくい多層のニューラルネットワークを得ることができる。

【0049】

次に、図7を参照して、画像編集装置111の推定部113で実行される推定工程に関して説明する。図7は、推定工程のフローチャートである。

【0050】

まず、ステップS201において、推定部113は、撮像装置121または記録媒体122から、撮像画像を取得する。撮像画像は、未現像のRAW画像である。RAW画像の信号値が符号化されている場合、推定部113は復号処理を実行する。また推定部113は、撮像画像を後に現像する際に用いられるガンマ補正に関する情報を取得する。

10

【0051】

続いてステップS202において、推定部113は、ステップS201にて取得したガンマ補正に関する情報から、その情報に対応するネットワークパラメータを取得する。ネットワークパラメータは、パラメータ学習装置101の記憶部102から読み出される。または、予め画像編集装置111の記憶部112に複数のネットワークパラメータを保存しておき、記憶部112から読み出してもよい。取得するネットワークパラメータは、ステップS201にて取得したガンマ補正に関する情報と学習工程で用いられたガンマ補正に関する情報とが互いに一致するもの、または、最も近いものである。

20

【0052】

続いてステップS203において、推定部113は、撮像画像からCNNへ入力する入力画像を取得する。入力画像は、訓練画像と同様に、色成分ごとに配列して三次元方向にスタックされる。なお、推定工程の入力画像のサイズは、学習工程における訓練画像のサイズと、必ずしも一致する必要はない。続いてステップS204において、推定部113は、入力画像とネットワークパラメータに基づいて、推定画像を生成する。推定画像の生成には、学習工程と同様に、図1に示されるCNNが用いられる。ただし、図1中の出力画像211が推定画像となり、それ以降のガンマ補正および正解画像との誤差算出は行わない。

【0053】

続いてステップS205において、推定部113は、撮像画像の所定の領域に対して推定（本実施例では、収差・回折のぼけ補正）が完了したか否かを判定する。推定が完了していない場合、ステップS203へ戻り、推定部113は、撮像画像の所定の領域から新たな入力画像を取得する。推定に用いられるCNNにおいて、出力画像のサイズが入力画像よりも小さくなる場合、所定の領域からオーバーラップして入力画像を取得する必要がある。所定の領域は、撮像画像の全体または一部である。撮像画像はRAW画像であるため、受光して得られた画像の他に、ヘッダ（前記画像の画素数や撮影時刻などの情報）や撮像素子のオプティカルブラックの情報が含まれていることがある。ヘッダやオプティカルブラックは、収差・回折のぼけと無関係であるため、所定の領域からそれらを除いてもよい。

30

40

【0054】

続いてステップS206において、推定部113は、生成された複数の推定画像を合成して、収差・回折によるぼけが補正された撮像画像を出力する。必要に応じて、推定部113は、ヘッダやオプティカルブラックの情報を含めて出力する。以上の推定工程により、現像後の信号値の大きさに対する影響を抑えて、収差・回折によるぼけを補正することができる。推定工程後、ユーザが任意で露出補正などの編集を行い、現像処理により最終的な現像画像を得る。

【0055】

次に、本発明の効果を増すための好ましい条件に関して説明する。まず、CNNへ入力する入力画像と正解画像の信号値を規格化する工程を設けることが好ましい。撮像画像は

50

、撮像装置の構成によって、画像の信号値の取り得る範囲が異なる。学習工程と推定工程とで入力画像における信号値の範囲が異なる場合、正しい推定結果が得られない。このため、信号値の規格化を行うことが好ましい。信号値が取り得る範囲は、下限値（オプティカルブラックの信号値）と上限値（輝度飽和値）とで規定される。信号値の範囲に関する情報は、撮像画像のヘッダやオプティカルブラック領域から取得することができる。具体的には、例えば以下の式（6）に従って規格化するとよい。

【0056】

【数6】

$$S_{\text{nor}} = \frac{S - S_{\text{OB}}}{S_{\text{satu}} - S_{\text{OB}}} \quad (6)$$

10

【0057】

式（6）において、 s はRAW画像の信号値、 s_{OB} はオプティカルブラックの信号値、 s_{satu} は信号の輝度飽和値、 s_{nor} は規格化された信号値を示す。推定工程では、出力された推定画像に対して、式（6）の逆を行って元のスケールに戻す。

【0058】

また、式（5）の誤差に対し、式（1）または式（2）で表されるようなガンマ補正を用いる場合、出力画像の信号値がゼロになる近傍で微分が発散することがある。誤差の微分が発散すると、誤差逆伝搬法が正しく動作しない。また、ガンマ補正の式によっては、補正される信号値がある範囲に収まっていないと、正しい補正值が出力されない。これらを回避するため、ガンマ補正の実行前に、出力画像の信号値を所定の範囲でクリッピングする工程を設けることが好ましい。所定の範囲は、ガンマ補正が定義されて、かつ微分が発散しない範囲である。例えば式（1）では、クリッピングの下限を正の微小な値とし、上限を1とする。

20

【0059】

同様に、推定画像の信号値も画像信号が取り得る範囲を超過することがある。このため、ステップS204では、推定画像の信号値をクリッピングすることが好ましい。規格化されている場合には0～1で、規格化されていない場合にはオプティカルブラックの信号値と輝度飽和値でクリッピングする。

30

【0060】

また、本実施例の学習工程では、一つのガンマ補正に関する情報ごとにネットワークパラメータを学習する。このため、取り得るガンマ補正の種類が多い場合、ネットワークパラメータのデータ容量が大きくなる。この対策として、複数のガンマ補正に対して一括でネットワークパラメータを学習してもよい。例えば、以下の式（7）で表される誤差 L_{sum} を用いて学習することができる。

【0061】

【数7】

$$L_{\text{sum}} = \frac{1}{2NM} \sum_k^M \sum_j^N |g_k(t_j) - g_k(y_j)|^2 \quad (7)$$

40

【0062】

式（7）において、 g_k は k で指定されるガンマ補正であり、 k によって互いに異なる。Mはガンマ補正の種類総数である。正解画像と出力画像に複数のガンマ補正それぞれを実行し、算出された複数の誤差を合わせた誤差 L_{sum} を用いて学習することで、ネットワークパラメータの容量を削減することができる。

【0063】

本実施例によれば、現像された画像の信号値の大きさに対して、推定精度が影響されに

50

くい多層のニューラルネットワークを取得し、そのニューラルネットワークを用いて推定を行うことが可能な画像処理システムを提供することができる。

【実施例 2】

【0064】

次に、本発明の実施例 2 における画像処理システムについて説明する。本実施例の画像処理システムでは、学習工程および推定工程を行う装置（サーバ 301）と、複数の撮像装置（第 1 撮像装置 311、第 2 撮像装置 312、第 3 撮像装置 313）とが互いに接続されている。本実施例における推定工程は、アップサンプリングであるが、これに限定されるものではない。

【0065】

図 8 および図 9 を参照して、本実施例における画像処理システムについて説明する。図 8 は、画像処理システム 300 のブロック図である。図 9 は、画像処理システム 300 の外観図である。図 8 および図 9 に示されるように、画像処理システム 300 は、サーバ（画像処理装置）301、第 1 撮像装置 311、第 2 撮像装置 312、および、第 3 撮像装置 313 を有する。サーバ 301 は、記憶部 302、学習部 303、推定部 304、および、現像部 305 を有する。学習部 303 は、記憶部 302 に記憶された複数の訓練画像と正解画像を読み出して学習工程を実行し、ニューラルネットワークのネットワークパラメータ（重みとバイアス）を学習する。学習部 303 は、図 3 に示される取得部 103a、生成部 103b、算出部 103c、および、更新部 103d と同様の機能を有する。

【0066】

本実施例において、訓練画像は低解像度の画像（第 1 の解像度を有する画像）であり、正解画像は高解像度の画像（第 1 の解像度よりも高い第 2 の解像度を有する画像）である。訓練画像は、正解画像をダウンサンプリングすることで用意することができる。または、本実施例において、訓練画像は低コントラストの画像（第 1 のコントラストを有する画像）であり、正解画像は高コントラストの画像（第 1 のコントラストよりも高い第 2 のコントラストを有する画像）であってもよい。

【0067】

なお、訓練画像と正解画像は、サイズ合わせを行ってもよいし、行わなくてもよい。正解画像を $1/n$ 倍にダウンサンプリングして訓練画像を生成した場合、訓練画像をバイキュービックなどの補間で n 倍に引き伸ばすことにより、サイズ合わせを行うことができる。サイズ合わせを行わない場合、ネットワーク内に逆畳み込み（Deconvolution）層や逆プーリング（Depooling）層などを挿入することが好ましい。

【0068】

学習部 303 は、図 1 に示されるネットワークを用いて、ネットワークパラメータを学習する。本実施例において、学習部 303 は、式（7）のように複数のガンマ補正が混合された誤差を最小化するように、ネットワークパラメータを学習する。学習されたネットワークパラメータは、記憶部 302 に記憶される。サーバ 301 は、複数の撮像装置（第 1 撮像装置 311、第 2 撮像装置 312、および、第 3 撮像装置 313）と無線接続されており、各撮像装置から撮像画像を受信する。撮像画像は RAW 画像である。推定部 304 は、記憶部 302 からネットワークパラメータを読み出し、撮像画像に対して推定工程を実行する。これにより、アップサンプリングされた撮像画像が生成される。アップサンプリングされた撮像画像は、現像部 305 により現像され、元の撮像装置へ送信される。

【0069】

学習部 303 で実行される学習工程は、図 5 を参照して説明した学習工程と同様である。ただし、本実施例ではステップ S102 が存在しない。推定部 304 で実行される推定工程は、図 7 を参照して説明した推定工程と同様である。

【0070】

本実施例によれば、現像された画像の信号値の大きさに対して、推定精度が影響されにくい多層のニューラルネットワークを取得し、そのニューラルネットワークにより推定が可能な画像処理システムを提供することができる。

10

20

30

40

50

【 0 0 7 1 】

(その他の実施例)

本発明は、上述の実施例の 1 以上の機能を実現するプログラムを、ネットワーク又は記憶媒体を介してシステム又は装置に供給し、そのシステム又は装置のコンピュータにおける 1 つ以上のプロセッサがプログラムを読み出し実行する処理でも実現可能である。また、1 以上の機能を実現する回路 (例えば、A S I C) によっても実現可能である。

【 0 0 7 2 】

各実施例によれば、現像された画像の信号の大きさに対して、推定精度が影響されにくい多層のニューラルネットワークを取得可能な画像処理方法、画像処理装置、画像処理プログラム、および、記憶媒体を提供することが可能である。

10

【 0 0 7 3 】

以上、本発明の好ましい実施例について説明したが、本発明はこれらの実施例に限定されず、その要旨の範囲内で種々の変形及び変更が可能である。

【 符号の説明 】

【 0 0 7 4 】

- 1 0 1 パラメータ学習装置 (画像処理装置)
- 1 0 3 a 取得部
- 1 0 3 b 生成部
- 1 0 3 c 算出部
- 1 0 3 d 更新部

20

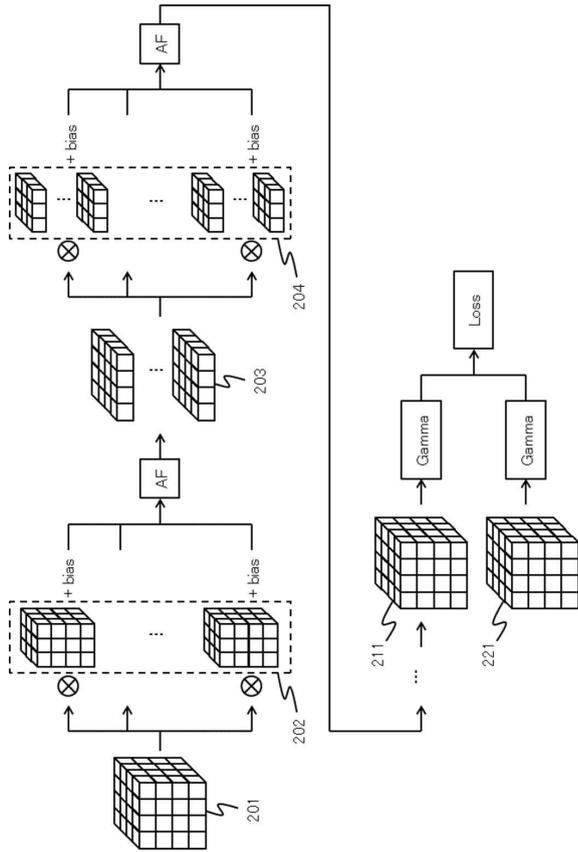
30

40

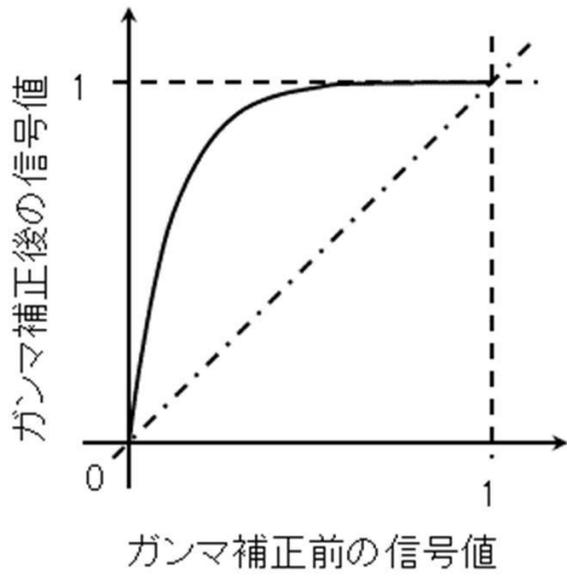
50

【図面】

【図 1】



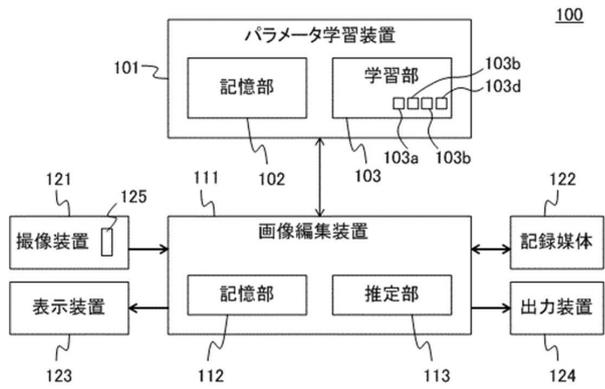
【図 2】



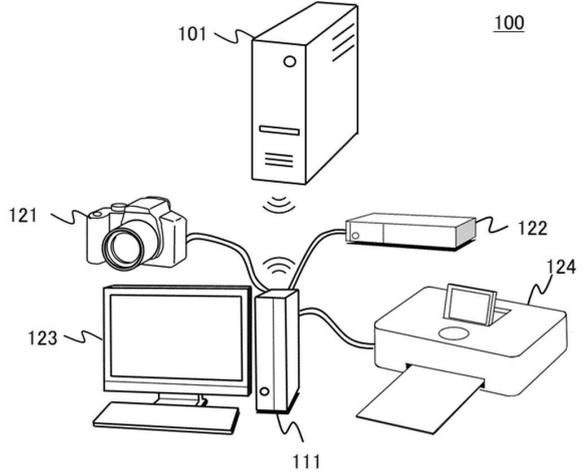
10

20

【図 3】



【図 4】

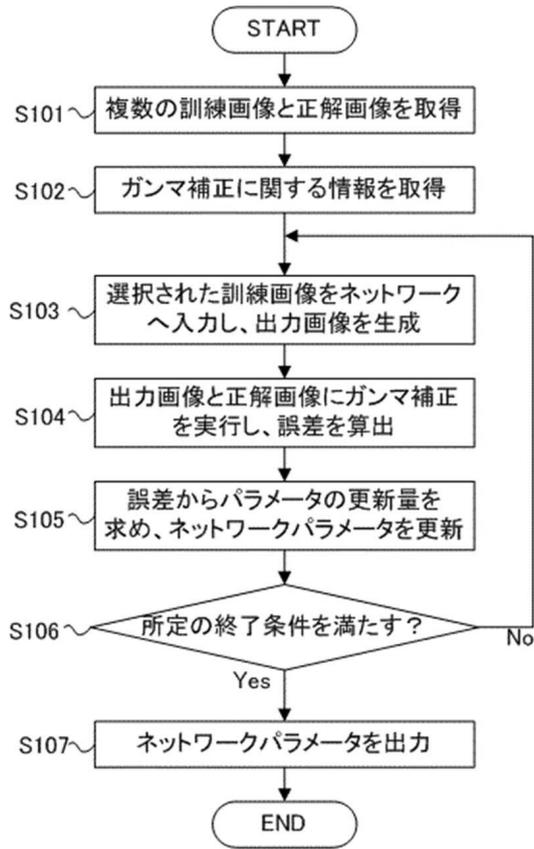


30

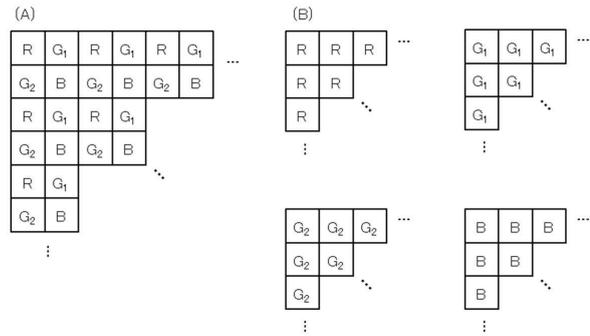
40

50

【図5】



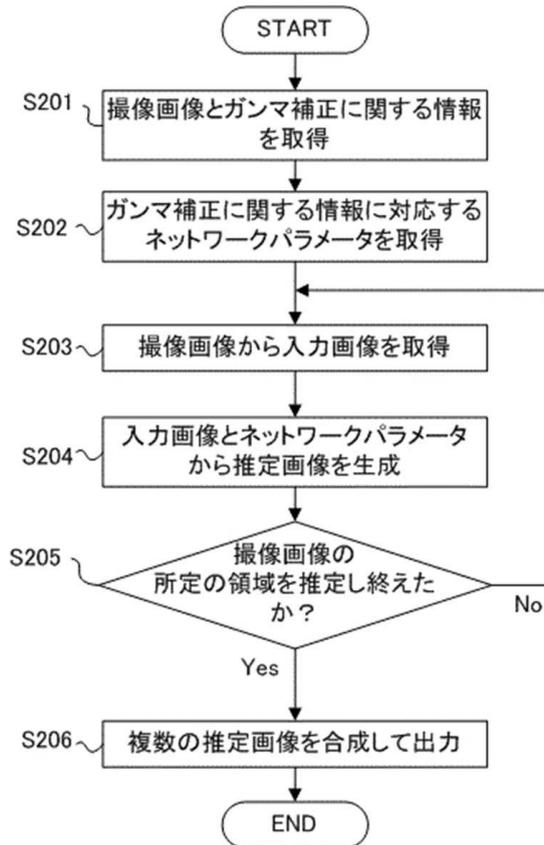
【図6】



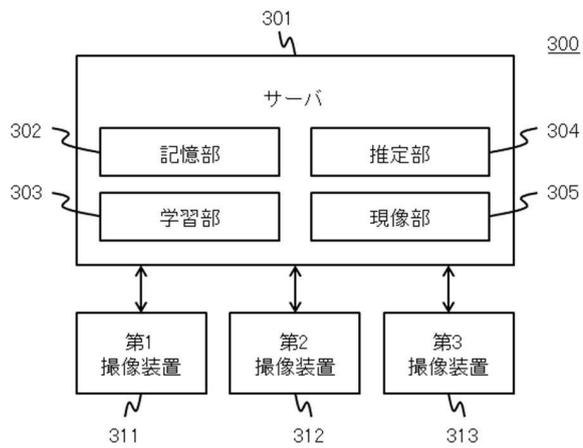
10

20

【図7】



【図8】

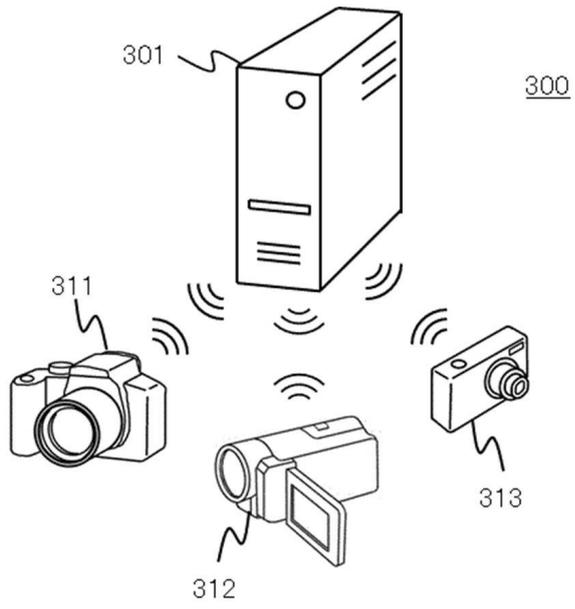


30

40

50

【 9】



10

20

30

40

50

フロントページの続き

(56)参考文献 米国特許出願公開第2017/0185871(US, A1)

特開2009-032021(JP, A)

特開2005-217592(JP, A)

特開2006-140952(JP, A)

(58)調査した分野 (Int.Cl., DB名)

G06T 1/00

G06T 1/40