

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-16114

(P2019-16114A)

(43) 公開日 平成31年1月31日(2019.1.31)

(51) Int.Cl.			F I			テーマコード (参考)	
<b>G06T</b>	<b>7/00</b>	<b>(2017.01)</b>	G06T	7/00	350B	5C122	
<b>H04N</b>	<b>5/232</b>	<b>(2006.01)</b>	H04N	5/232	120	5L096	
<b>H04N</b>	<b>5/235</b>	<b>(2006.01)</b>	H04N	5/235	100		

審査請求 未請求 請求項の数 20 O L (全 27 頁)

(21) 出願番号	特願2017-132259 (P2017-132259)	(71) 出願人	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29年7月5日(2017.7.5)	(74) 代理人	100076428 弁理士 大塚 康德
		(74) 代理人	100115071 弁理士 大塚 康弘
		(74) 代理人	100112508 弁理士 高柳 司郎
		(74) 代理人	100116894 弁理士 木村 秀二
		(74) 代理人	100130409 弁理士 下山 治
		(74) 代理人	100134175 弁理士 永川 行光

最終頁に続く

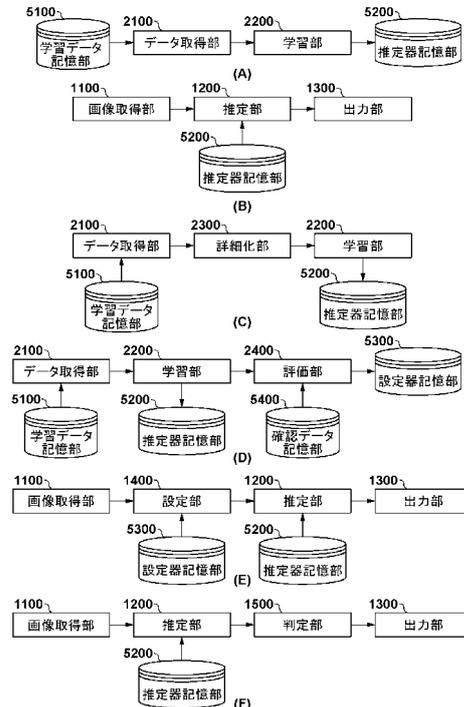
(54) 【発明の名称】 画像処理装置、学習装置、フォーカス制御装置、露出制御装置、画像処理方法、学習方法、及びプログラム

(57) 【要約】 (修正有)

【課題】 分類結果を用いた処理の精度を向上できるように、画像の各領域のクラス分類を行うことができる画像処理装置、学習装置、フォーカス制御装置、露出制御装置、画像処理方法、学習方法、及びプログラムを提供する。

【解決手段】 画像処理装置は、対象画像から特徴量を抽出する抽出手段と、特徴量に基づいて、対象画像における互いに異なる属性を持つ領域の混合状態を推定する推定手段と、を備える。

【選択図】 図1



**【特許請求の範囲】****【請求項 1】**

対象画像から特徴量を抽出する抽出手段と、  
前記特徴量に基づいて、前記対象画像における互いに異なる属性を持つ領域の混合状態を推定する推定手段と、  
を備えることを特徴とする画像処理装置。

**【請求項 2】**

前記領域の 1 つは、特定のクラスに属する被写体の領域であり、前記領域の別の 1 つは、前記特定のクラスとは別のクラスに属する被写体の領域であることを特徴とする、請求項 1 に記載の画像処理装置。

10

**【請求項 3】**

前記混合状態を示す情報を出力する出力手段をさらに備えることを特徴とする、請求項 1 又は 2 に記載の画像処理装置。

**【請求項 4】**

前記混合状態を示す情報は、前記対象画像における前記領域の分布に依存して定まる情報であることを特徴とする、請求項 3 に記載の画像処理装置。

**【請求項 5】**

前記混合状態を示す情報は、前記対象画像における前記領域のそれぞれの比率を表す情報であることを特徴とする、請求項 3 に記載の画像処理装置。

**【請求項 6】**

前記混合状態を示す情報は、前記対象画像における前記領域の配置を表す情報であることを特徴とする、請求項 3 に記載の画像処理装置。

20

**【請求項 7】**

前記混合状態を示す情報は、前記対象画像における領域間の境界に係る情報であることを特徴とする、請求項 3 に記載の画像処理装置。

**【請求項 8】**

入力画像を取得する取得手段と、  
領域設定パターンに従って、前記入力画像中に複数の領域を設定する設定手段と、をさらに備え、

前記対象画像は、前記設定された領域のそれぞれに含まれる前記入力画像の部分画像であることを特徴とする、請求項 1 から 7 のいずれか 1 項に記載の画像処理装置。

30

**【請求項 9】**

前記取得手段は、前記入力画像の撮影情報をさらに取得し、  
前記設定手段は、前記撮影情報に応じて前記領域設定パターンを選択することを特徴とする、請求項 8 に記載の画像処理装置。

**【請求項 10】**

前記入力画像の第 1 の部分にある、第 1 の領域設定パターンに従う大きさの第 1 の対象画像における混合状態に従って、前記第 1 の部分の混合状態を再判定するか否かを判定する判定手段をさらに備え、

再判定を行うとの判定に応じて、前記推定手段は、前記第 1 の部分にある、第 2 の領域設定パターンに従う大きさの第 2 の対象画像における混合状態を推定し、

前記第 2 の対象画像は前記第 1 の対象画像よりも小さいことを特徴とする、請求項 8 に記載の画像処理装置。

40

**【請求項 11】**

前記推定手段は、前記領域設定パターンに従って複数の推定器から選択された推定器を推定に用いることを特徴とする、請求項 8 から 10 のいずれか 1 項に記載の画像処理装置。

**【請求項 12】**

推定器の学習に用いる識別画像の特徴量を抽出する抽出手段と、  
前記識別画像における互いに異なる属性を持つ領域の混合状態を示す情報を教師情報と

50

して取得する取得手段と、

前記特徴量から前記混合状態を推定する推定器の学習を、前記識別画像の特徴量と前記教師情報との組み合わせを用いて行う学習手段と、

を備えることを特徴とする学習装置。

【請求項 13】

前記取得手段は、前記識別画像の各画素の属性を示す情報を取得し、前記各画素の属性を示す情報を用いて前記混合状態を示す情報を生成することにより、前記教師情報を取得することを特徴とする、請求項 12 に記載の学習装置。

【請求項 14】

識別画像と、前記識別画像における第 1 の属性の領域、第 2 の属性の領域、及び第 1 の属性の領域と第 2 の属性の領域とが混在している混在領域を示す情報と、を取得する取得手段と、

10

前記第 1 の属性の領域に含まれる画素の画素値、及び前記第 2 の属性の領域に含まれる画素の画素値に基づいて、前記混在領域の各画素の属性を判定する判定手段と、

を備えることを特徴とする画像処理装置。

【請求項 15】

対象画像と、前記対象画像における互いに異なる属性を持つ領域の混合状態を示す情報と、を取得する取得手段と、

前記対象画像の各画素の属性を判定する判定手段と、を備える画像処理装置であって、

前記判定手段は、評価値に基づいて各画素の属性を決定し、

20

前記評価値は、前記各画素の属性に基づいて計算される混合状態を示す情報と、前記取得手段が取得した前記混合状態を示す情報と、の類似度が大きいほど高くなることを特徴とする、画像処理装置。

【請求項 16】

複数の測距点を備える撮像装置のためのフォーカス制御装置であって、

前記撮像装置により得られた画像のうち前記複数の測距点のそれぞれに対応する領域について、前記領域に占める特定の属性の領域の面積比を示す情報を取得する取得手段と、

前記面積比に応じて前記複数の測距点を重み付けし、前記撮像装置のフォーカス制御を行う制御手段と、

を備えることを特徴とするフォーカス制御装置。

30

【請求項 17】

撮像装置により得られた画像と、前記画像の各領域について、前記領域に占める特定の属性の領域の面積比を示す情報と、を取得する取得手段と、

前記画像の全体に占める前記特定の属性の領域の面積比を算出する算出手段と、

前記算出手段が算出した面積比に応じて、露出制御アルゴリズムを選択する選択手段と

、

選択された前記露出制御アルゴリズムを用いて、前記撮像装置の露出制御を行う制御手段と、

を備えることを特徴とする露出制御装置。

【請求項 18】

40

画像処理装置が行う画像処理方法であって、

対象画像から特徴量を抽出する抽出工程と、

前記特徴量に基づいて、前記対象画像における互いに異なる属性を持つ領域の混合状態を推定する推定工程と、

を有することを特徴とする画像処理方法。

【請求項 19】

学習装置が行う学習方法であって、

推定器の学習に用いる識別画像の特徴量を抽出する抽出工程と、

前記識別画像における互いに異なる属性を持つ領域の混合状態を示す情報を教師情報として取得する取得工程と、

50

前記特徴量から前記混合状態を推定する推定器の学習を、前記識別画像の特徴量と前記教師情報との組み合わせを用いて行う学習工程と、

を有することを特徴とする学習方法。

【請求項 20】

コンピュータを、請求項 1 から 11、14、及び 15 のいずれか 1 項に記載の画像処理装置、請求項 12 若しくは 13 に記載の学習装置、請求項 16 に記載のフォーカス制御装置、又は請求項 17 に記載の露出制御装置が有する各手段として動作させるためのプログラム。

【発明の詳細な説明】

【技術分野】

10

【0001】

本発明は、画像処理装置、学習装置、フォーカス制御装置、露出制御装置、画像処理方法、学習方法、及びプログラムに関する。

【背景技術】

【0002】

近年、画像を領域分割する研究も広く行われている。例えば、画像から人物の領域、自動車の領域、道路の領域、建物の領域、又は空の領域等を切り出すことができる。これは意味的領域分割 (Semantic Segmentation) と呼ばれ、分割結果は被写体の種類に対応した画像補正又はシーン解釈等に応用可能である。

【0003】

20

意味的領域分割方法としては、画像を事前にいくつかの領域に分割し、分割された各領域をクラス分類する方法がある。例えば、画像を複数の矩形ブロックに分割し、それぞれのブロックをクラス分類することができる。画像をクラス分類する方法としては、非特許文献 1 に記載のように、深層学習を用いた分類が広く研究されている。また、例えば非特許文献 2 に記載の方法を用いて画像を不定形の小領域 (superpixel) に分割し、その領域の特徴量及び領域周辺のコンテキスト特徴量を用いて領域をクラス分類することもできる。クラス分類には、学習画像を用いて学習が行われた推定器を用いることができる。

【0004】

近年では、深層学習を利用した領域分割も研究されている。非特許文献 3 は、CNN (Convolutional Neural Network) の中間層出力を特徴量として利用し、複数の中間層特徴による画素ごとのクラス判定結果を統合する。この方法では、小領域分割結果を利用することなく、画素ごとに直接クラス判定を行うことができる。

30

【先行技術文献】

【非特許文献】

【0005】

【非特許文献 1】A. Krizhevsky et al. "ImageNet Classification with Deep Convolutional Neural Networks", Proc. Advances in Neural Information Processing Systems 25 (NIPS 2012).

40

【非特許文献 2】R. Achanta et al. "SLIC Superpixels", EPFL Technical Report 1493 00, 2010.

【非特許文献 3】J. Long et al. "Fully Convolutional Networks for Semantic Segmentation", The IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015.

【発明の概要】

【発明が解決しようとする課題】

【0006】

従来の方法によれば、画像上のそれぞれの小領域を、被写体の種類に基づいてクラス分類することができた。例えば、各領域の特徴量に基づいて、領域が空を表す領域であるか

50

前景（空以外）の領域であるかを判定することができた。一方、異なる種類の被写体が混在している領域については、適切な判定を行うことが困難であった。例えば、たくさんの木の枝の隙間に空が見えるような複数の領域について判定を行うと、テクスチャが類似しているために、全ての領域について前景であると判定するか、全ての領域について空であると判定してしまう可能性が高かった。

【0007】

本発明は、分類結果を用いた処理の精度を向上できるように、画像の各領域のクラス分類を行うことを目的とする。

【課題を解決するための手段】

【0008】

本発明の目的を達成するために、例えば、本発明の画像処理装置は以下の構成を備える。すなわち、

対象画像から特徴量を抽出する抽出手段と、

前記特徴量に基づいて、前記対象画像における互いに異なる属性を持つ領域の混合状態を推定する推定手段と、

を備えることを特徴とする。

【発明の効果】

【0009】

本発明によれば、分類結果を用いた処理の精度を向上できるように、画像の各領域のクラス分類を行うことができる。

【図面の簡単な説明】

【0010】

【図1】各実施形態に係る装置の構成例を示すブロック図。

【図2】各実施形態に係る処理を説明するフローチャート。

【図3】学習画像及びクラスラベルを説明する図。

【図4】2クラス識別時のクラスラベル及び混合状態を説明する図。

【図5】3クラス識別時のクラスラベル及び混合状態を説明する図。

【図6】3クラス識別時の混合状態を説明する図。

【図7】混合状態のマッピングの説明図。

【図8】各実施形態に係る学習部又は推定部の構成例を示す図。

【図9】混在領域におけるクラスラベル詳細化を説明する図。

【図10】各実施形態を実現可能なコンピュータの構成例を示す図。

【発明を実施するための形態】

【0011】

本発明の一実施形態によれば、入力画像上の識別単位となる所定領域における、複数のクラスの混ざり方（以下、混合状態と呼ぶ）を推定することができる。以下では、推定対象となる領域内の画像のことを対象画像と呼ぶことがある。より具体的には、本発明の一実施形態によれば、対象画像における、互いに異なる属性を持つ領域の混合状態が判定される。それぞれの属性の領域は、同じクラスに属する被写体が占める領域である。すなわち、この属性の領域の1つは、特定のクラスに属する被写体の領域であり、この属性の領域の別の1つは、特定のクラスとは別のクラスに属する被写体の領域である。

【0012】

一実施形態によれば、例えばたくさんの木の枝（前景）の隙間に空が見えるような領域について、前景部分と空部分との混合状態（例えば面積比、エッジ面積、又は配置パターン等）を推定することができる。従来の方法で得られるような各領域のクラス情報（例えば前景領域であるか空領域であるかを示す情報）だけでなく、このような混合状態を示す情報を用いることにより、後に画像に対して行う処理の精度を向上させることができる。具体例については各実施形態において詳しく説明する。

【0013】

以下、本発明の実施形態を図面に基づいて説明する。ただし、本発明の範囲は以下の実

10

20

30

40

50

形態に限定されるものではない。以下の実施形態においては、図 1 等に示される各処理部は、コンピュータにより実現してもよいし、専用のハードウェアによって実現してもよい。

#### 【0014】

図 10 は、各実施形態を実現可能なコンピュータの基本構成を示す図である。図 10 においてプロセッサ 101 は、例えば CPU であり、コンピュータ全体の動作をコントロールする。メモリ 102 は、例えば RAM であり、プログラム及びデータ等を一時的に記憶する。コンピュータが読み取り可能な記憶媒体 103 は、例えばハードディスク又は CD-ROM 等であり、プログラム及びデータ等を長期的に記憶する。本実施形態においては、記憶媒体 103 が格納している、各部の機能を実現するプログラムが、メモリ 102 へと読み出される。そして、プロセッサ 101 が、メモリ 102 上のプログラムに従って動作することにより、各部の機能が実現される。

10

#### 【0015】

図 10 において、入力インタフェース 104 は外部の装置から情報を取得するためのインタフェースである。また、出力インタフェース 105 は外部の装置へと情報を出力するためのインタフェースである。バス 106 は、上述の各部を接続し、データのやりとりを可能とする。

#### 【0016】

##### [実施形態 1]

図 1 (A) 及び (B) に沿って、実施形態 1 に係る画像処理装置及び学習装置の基本的な構成を説明する。

20

#### 【0017】

最初に、図 1 (A) に従って、学習装置の装置構成の概要を説明する。本実施形態において学習装置は、後述する画像処理装置が混合状態を認識する処理を行う際に利用する推定器を、事前に用意された学習画像から生成する。学習処理の詳細については後述する。学習データ記憶部 5100 には、あらかじめ用意された学習データを記憶している。学習データは、学習画像と教師情報とを含む。データ取得部 2100 は、学習データ記憶部 5100 から、学習画像と教師情報とを取得する。学習部 2200 は、特徴抽出部 610 を用いて、学習画像の所定領域にある、推定器の学習に用いる識別画像の特徴量を抽出する。また、学習部 2200 は、特徴量から混合状態を推定する推定器の学習を、識別画像の特徴量と教師情報との組み合わせを用いて行う。例えば、学習部 2200 は、特徴量を入力されると混合状態を示す情報を出力する推定器の学習を行うことができる。ここで、教師情報は、識別画像における、互いに異なる属性の領域間の混合状態を示す情報である。学習により得られた推定器は、推定器記憶部 5200 に記憶される。具体的には、推定器記憶部 5200 は、学習により決定された推定器のパラメータを記憶することができる。

30

#### 【0018】

次に、図 1 (B) に従って、画像処理装置の装置構成の概要を説明する。本実施形態において画像処理装置は、未知の入力画像における混合状態を推定する処理を行う。処理内容の詳細は後述する。画像取得部 1100 は、入力画像を取得する。推定部 1200 は、特徴抽出部 610 を用いて、入力画像の所定領域にある、混合状態の識別対象となる対象画像から特徴量を抽出する。推定部 1200 が用いる特徴抽出部 610 は、学習部 2200 が用いる特徴抽出部 610 と同様に動作できる。また、推定部 1200 は、特徴量に基づいて、対象画像における互いに異なる属性を持つ領域の混合状態を推定する。例えば、推定部 1200 は、あらかじめ学習が行われた推定器 620 を推定器記憶部 5200 から読み込み、推定器に特徴量を入力することにより得られた、対象画像における互いに異なる属性の領域間の混合状態を示す情報を、出力部 1300 に出力する。推定器 620 は、学習部 2200 による学習により得られたものでありうる。出力部 1300 は、推定部 1200 による推定結果を出力する。

40

#### 【0019】

学習装置のデータ取得部 2100 及び学習部 2200 は、同じ計算機上で実現されても

50

よいし、それぞれ独立したモジュールとして構成されていてもよいし、計算機上で動作するプログラムとして実装されてもよい。学習装置の学習データ記憶部 5 1 0 0 及び推定器記憶部 5 2 0 0 は、計算機の内部又は外部にあるストレージを用いて実現することができる。

#### 【 0 0 2 0 】

画像処理装置の画像取得部 1 1 0 0 及び推定部 1 2 0 0 は、同じ計算機上で実現されてもよいし、それぞれ独立したモジュールとして構成されていてもよいし、計算機上で動作するプログラムとして実装されてもよい。また、これらはカメラ等の撮影装置内部に回路又はプログラムとして実装されてもよい。

#### 【 0 0 2 1 】

画像処理装置は、学習装置と同じ計算機上で実現されてもよいし、別々の計算機上で実現されてもよい。学習装置及び画像処理装置が備える推定器記憶部 5 2 0 0 は、同じストレージであってもよいし、異なるストレージであってもよい。異なるストレージを用いる場合、学習装置により推定器記憶部 5 2 0 0 に格納された推定器を、画像処理装置が備える推定器記憶部 5 2 0 0 にコピー又は移動することができる。

#### 【 0 0 2 2 】

以下に、本実施形態に係る処理を詳細に説明する。まず、学習装置が行う学習時の処理に関して図 2 ( A ) のフローに従って説明する。S 2 1 0 0 においてデータ取得部 2 1 0 0 は、学習データ記憶部 5 1 0 0 から、学習画像と、混合状態の教師情報と、を学習データとして取得する。

#### 【 0 0 2 3 】

学習データ記憶部 5 1 0 0 には、あらかじめ複数の学習画像と混合状態の教師情報が記憶されている。学習画像とは、推定器の学習に用いられる画像のことを指す。学習画像は、例えば、デジタルカメラ等で撮影された画像データでありうる。画像データの形式は特に限定されず、例えば J P E G、P N G、又は B M P 等でありうる。以下では、用意された学習画像の枚数を N 枚とし、n 番目の学習画像を  $I_n$  ( $n = 1 \cdots N$ ) と表す。

#### 【 0 0 2 4 】

混合状態の教師情報は、学習画像の所定領域における混合状態を示す。この教師情報は予め用意されており、例えば人間が学習画像を見ながら作成することができる。本実施形態においては、学習画像には識別単位となる複数の領域が設定されており、それぞれの領域について教師情報が用意されている。以下、1 つの識別単位となる、学習画像中の所定領域の画像を、識別画像と呼ぶ。

#### 【 0 0 2 5 】

領域の設定方法は特に限定されない。例えば、所定の領域設定パターンに従って、入力画像中に複数の領域を設定することができる。具体例としては、学習画像を所定サイズ (例えば  $16 \times 16$  ピクセル) の複数の矩形領域に分割し、それぞれの矩形領域を識別単位として扱うことができる。また、非特許文献 2 に記載の手法で得られた小領域を識別単位の領域として扱うことができる。一方、学習画像の一部にのみ所定サイズの矩形領域を設定してもよい。なお、学習データ記憶部 5 1 0 0 には、所定サイズの識別画像が学習データとして記憶されていてもよい。

#### 【 0 0 2 6 】

以下に、教師情報が示す混合状態について説明する。画像上の被写体は、複数のクラスに分類することができる。図 3 は、このようなクラス分類の例を示す。図 3 ( A ) には、学習画像 5 0 0 の例を示す。学習画像 5 0 0 には、空、人物、及び植物が写っており、それぞれを異なるクラスに分類することができる。すなわち、図 3 ( B ) に示すように、領域 5 4 1 に含まれる画素には「空」のクラスラベルを、領域 5 4 2 に含まれる画素には「人物」のクラスラベルを、領域 5 4 3 に含まれる画素には「植物」のクラスラベルを、それぞれ与えることができる。

#### 【 0 0 2 7 】

クラス及びクラスラベルの定義は様々であり、クラス分類の方法は特に限定されない。

10

20

30

40

50

図3の例では、被写体の種類に応じてクラス分類を行った。別のクラスラベルの例としては、肌領域又は髪領域、犬又は猫等の動物、及び自動車又は建物等の人工物、等が挙げられる。工場で用いられる部品A又は部品Bといったような、特定物体を示すクラスラベルを用いることもできる。一方、各画素を主被写体領域と背景領域にクラス分類してもよい。また、光沢面又はマット面といった表面性状の違いや、金属面又はプラスチック面のような素材の違いに応じて、クラス分類を行ってもよい。以下では、クラスは全部でM種類あるものとする。

#### 【0028】

クラスの混合状態とは、対象画像における互いに異なる属性の領域間の混合状態のことである。それぞれの属性の領域は、同じ特定のクラスに属する被写体が占める領域である。互いに異なる属性の領域の一方は、特定のクラスに属する被写体の領域であり、他方は、特定のクラスとは別のクラスに属する被写体の領域である。以下では、あるクラスに属する被写体が占める属性の領域のことを、単にそのクラスに属する領域と呼ぶことがある。また、各画素に写っている被写体のクラスのことを、以下では画素の属性又はクラスと呼ぶことがある。

10

#### 【0029】

混合状態の定義はさまざまなものが考えられる。本実施形態では、以下のように混合状態を数値で表すものとする。一実施形態において、混合状態を示す情報は、対象画像における属性の領域の分布に依存して定まる情報である。例えば、混合状態を示す情報は、対象画像における属性の領域のそれぞれの比率を表す情報である。具体例として、混合状態を示す情報は、対象画像において各クラスに属する領域の面積比でありうる。クラスが「空」と「非空」の2クラスである場合について、図4の例を参照して説明する。図4には、学習画像510と、そのクラスラベル520が表示されている。クラスラベル520は、クラスが「空」である画素を白で、クラスが「非空」である画素を黒で表す。図4には、学習画像510中の識別画像515に対応するクラスラベル520上の領域525を拡大した、拡大図(526)が示されている。拡大図(526)には、非空領域511と、空領域522とが示されている。このとき、識別画像515の混合状態を、対応する領域525における空領域と非空領域との面積比 $r$ で表すことができる。例えば、 $16 \times 16 \text{ pixel}$ の矩形領域において、空領域画素が $192 \text{ pixel}$ 、非空領域画素が $64 \text{ pixel}$ であった場合には、 $r = 192 / 256 = 0.75$ となる。

20

30

#### 【0030】

上の例では2クラスの面積比について説明したが、3つ以上のクラスの面積比を表すこともできる。図5には、学習画像530と、学習画像530中の識別画像535と、識別画像535についてのクラスラベル536が表示されている。この例では、学習画像530の各画素は「空」、「植物」、及び「人工物」の3クラスに分類されている。この場合の混合状態は、植物領域531、空領域532、及び人工物領域533の面積比に応じて定めることができる。一例として、この場合に面積比を示す点545を、それぞれのクラスの面積比に応じて、図6(A)に示す座標空間中の単体540(図5の場合三角形)上にプロットすることができる。この点545は、単体の二つの辺を内分する内分比 $t_1$ 及び $t_2$ を用いて一意に表すことができるため、このときの面積比を、 $r = (t_1, t_2)$ というベクトルで表すことができる。これは一般M次元においても同様であるため、クラス数Mのときの面積比は、 $r = (t_1, t_2, \dots, t_{M-1})$ というM-1次元のベクトルで一意に表すことができる。なお、前述のような2クラスの場合の面積比は、この一般化された形式において $M = 2$ と設定した場合と同値である。

40

#### 【0031】

また、上記のように表されるMクラスの面積比を、低次の空間にマッピングして扱ってもよい。例えば、識別画像における面積比をM次元空間にプロットし、SOM(Self-Organizing Map)やLLE(Locally Linear Embedding)を用いて低次の空間にマッピングすることができる。図6(B)は、上記の3クラスの混合比の空間を1次元のSOM550で量子化した例を示す。551はSOM

50

550の始端ノードを表し、552はSOM550の終端ノードを表す。図6(C)は、これらを1次元空間にマッピングすることにより得られる。始端ノードの位置を0、終端ノードの位置を1とし、ノードを均等に配置すると、マップ上の位置を表すスカラ値により面積比を示す点545を表現することができる。例えば、図6(B)に示す面積比を示す点545は、図6(C)上の点546として近似することができる。そして点546のマップ上の位置(この例では $r = 0.37$ )を用いて面積比を表すことができる。マッピング後の次元数は、図6(C)に示すように1次元には限られず、一般的にMの値が大きい場合には1以上の次元RのSOMで近似することができる。その場合rはR次元のベクトルで表すことができる。例えば、クラス数 $M = 5$ のとき、2次元のSOMで面積比空間を量子化した場合には、rは2次元ベクトルで表すことができる。

10

#### 【0032】

また、Mクラスの面積比を、複数の基底ベクトルの合成ベクトルとして表現してもよい。例えば、さまざまな識別画像から得られたクラス面積比を、主成分分析又はスパースコーディング等を用いて複数の基底ベクトルに分解し、寄与度の大きい少数のベクトルでこれを近似することができる。この場合、面積比空間における面積比は、これら基底ベクトルの合成ベクトルとして表現することができ、そのときの各基底ベクトルに対する重み係数を用いて面積比を表すことができる。

#### 【0033】

別の例として、混合状態を示す情報は、対象画像における互いに異なる属性の領域間の境界に係る情報、例えば対象画像中におけるこの境界を表す画素の比率を表す情報でありうる。一例として、各画素のクラス(例えば空領域又は非空領域)を示す二値画像に対してエッジ検出を行い、得られたエッジ画素数をカウントし、所定領域の画素数とエッジ画素数との比 $e$ を用いて混合状態を表すことができる。図4には、クラスラベル526に対するエッジ検出結果527が示されており、検出されたエッジ画素523が表されている。16×16 pixelの矩形領域におけるエッジ画素のカウント結果が64である場合、エッジ画素率は $e = 64 / 256 = 0.25$ と表すことができる。

20

#### 【0034】

さらなる別の例として、混合状態を示す情報は、対象画像における属性の領域の配置を表す情報でありうる。例えば、所定領域内における各クラスの画素の配置パターンに従って混合状態を表すことができる。クラス数がM、所定領域の画素数がKである場合、所定領域内の各画素のクラスを、 $M \times K$ 次元の二値ベクトルで表すことができる。例えば、「空」「非空」の2クラスが定義されており、所定領域のサイズが16×16 pixelである場合、所定領域内のクラスラベル配置パターンを、 $2 \times 16 \times 16 = 512$ 次元の二値ベクトルとして表現することができる。このように識別画像から得られたさまざまな二値ベクトルをベクトル空間上にプロットし、SOM若しくはLLE等を用いて量子化することにより、所定領域におけるクラスラベル配置パターンをベクトルpとして表現することができる。また、識別画像から得られたさまざまな二値ベクトルを主成分分析若しくはスパースコーディング等を用いて基底ベクトルで表現する方法を用いることもできる。

30

#### 【0035】

図7は、所定領域におけるクラスラベル配置パターンを2次元のSOMを用いてマッピングすることにより得られたマップ900を示す。マップ900において、各四角形はSOMのノードであって、量子化されたクラスラベル配置パターンをそれぞれ表す。SOMの特性上、類似したパターンがマップ上で近い位置に配置される。識別画像におけるクラスラベル配置パターンを、各ノードのパターンとの近さに基づいて、マップ上の位置座標pで表すことができる。例えば図7の2次元SOMの例では、マップ上の位置を $p = (p_1, p_2)$ の2次元ベクトルで表すことができる。

40

#### 【0036】

このように、混合状態はさまざまな方法で表現することができる。混合状態は、これらの表現のうちいずれか一つを用いて表してもよい。例えば、混合状態を面積比だけで表すのであれば、 $C = r$ と定義すればよく、エッジ画素率だけで表すのであれば $C = e$ と表

50

わせばよく、クラスラベル配置パターンだけで表すのであれば  $C = p$  と定義すればよい。また、混合状態は複数の表現を組み合わせて表してもよい。例えば、混合状態を面積比とエッジ画素率との組み合わせ  $C = (r, e)$  と定義してもよいし、面積比、エッジ画素率、及びクラスラベル配置パターンの組み合わせ  $C = (r, e, p)$  と定義してもよい。本発明において、混合状態の表現方法は特に限定されない。

#### 【0037】

上記のように、混合状態  $C$  は1つ以上の数値で表されるベクトルとして表現することができる。すなわち、一実施形態において得られる混合状態を表す情報は、所定領域における混合状態を表す特徴量であるといえる。混合状態  $C$  を表すベクトルの次元数を  $L$  とする。以下では、画像  $I_n$  上の所定領域  $i$  における混合状態ベクトルを  $C_{n,i}$  と表し、混合状態ベクトル  $C_{n,i}$  の  $l$  ( $l = 1, \dots, L$ ) 番目の要素を  $c(n, i, l)$  と表す。なお、混合状態は、所定領域において各クラスの画素がどのように混合されているかだけでなく、所定領域が特定の1つのクラスの画素で構成されることを示してもよい。

10

#### 【0038】

本実施形態においては、各学習画像の各画素について図3(B)に示すようにクラスラベルが与えられているものとする。そして、このクラスラベルに基づいて、学習画像から得られるそれぞれの識別画像について、上記のようにスカラー値又はベクトル値として表される混合状態  $C$  が教師情報として予め算出され、学習データ記憶部5100に予め格納されているものとする。しかしながら、データ取得部2100は、識別画像の各画素の属性を示す情報を取得し、各画素の属性を示す情報を用いて混合状態を示す情報を生成することにより、教師情報を取得してもよい。例えば、データ取得部2100は、学習データ記憶部5100に格納されている学習画像の各画素のクラスラベルに基づいて、上記のようにそれぞれの識別画像の混合状態  $C$  を算出することができる。さらに、各学習画像の各画素について図3(B)に示すようにクラスラベルが与えられていることは必須ではない。例えば、学習画像から得られる識別画像を見ながら作業者が入力したこの識別画像の混合状態、又は作業者が入力した情報(エッジ情報など)に基づいて自動的に算出されたこの識別画像の混合状態が、学習データ記憶部5100に予め格納されていてもよい。

20

#### 【0039】

ステップS2200で、学習部2200は、データ取得部2100から識別画像及び混合状態の教師情報を取得して、混合状態を推定する推定器の学習を行う。以下では、推定器としてCNN(Convolutional Neural Network)を利用する場合について説明する。CNNの構成としては従来既知のものを用いることができる。典型的には、CNNは、畳み込み層とプーリング層とを繰り返すことにより入力信号の局所的な特徴を次第にまとめていき、変形や位置ずれに対してロバストな特徴を得ることにより、認識タスクを行うニューラルネットワークである。

30

#### 【0040】

CNNを用いた推定処理の例を、図8(A)を参照しながら説明する。学習部2200は、特徴抽出部610を用いて、推定器の学習に用いる識別画像の特徴量を抽出する。また、学習部2200は、特徴量を入力されると混合状態を示す情報を出力する推定器の学習を、識別画像の特徴量と教師情報との組み合わせを用いて行う。図8(A)は、学習部2200が処理のために用いることができるCNNの一例を示す。図8(A)には、特徴抽出部610が行う処理に相当する部分が示されており、これは特徴抽出を行うCNNの畳み込み層に相当する。また、図8(A)には、学習を行う推定器620に相当する部分が示されており、これはパターン推定を行うCNNの完全結合層に相当する。

40

#### 【0041】

畳み込み層は、学習画像の部分画像である識別画像630の各位置における畳み込み演算結果を信号として受け取る入力層611を有する。入力層611からの信号は、畳み込み層とプーリング層とが配置され、畳み込み演算とプーリングによる信号の選択とが繰り返される複数の中間層612, 613を介して、最終層615へと送られる。特徴抽出部610の最終層615からの出力信号は、推定器620へと送られる。以下では、特徴抽

50

出部 6 1 0 の出力信号を  $X$  とする。完全結合層では、各層の素子が前後の層と全結合しており、特徴抽出部 6 1 0 から入力された信号は、重み係数を用いた積和演算を介して出力層 6 4 0 へと送られる。出力層 6 4 0 は、混合状態ベクトル  $C$  の次元数  $L$  と同数の出力素子を有している。

#### 【 0 0 4 2 】

学習部 2 2 0 0 は、推定器の学習を行う際に、学習画像  $I_n$  の所定領域  $i$  から得られた識別画像を CNN に入力した際に、出力層 6 4 0 で得られる出力信号の値を、教師情報と比較する。ここで、学習画像  $I_n$  の所定領域  $i$  を特徴抽出部 6 1 0 に入力して得られた特徴量を  $X^n_i$ 、これを推定器 6 2 0 に入力した結果得られた出力層 6 4 0 における  $l$  番目の素子の出力信号を  $y_l(X^n_i)$  とする。また、出力層 6 4 0 のうち  $l$  番目の出力素子における教師信号は、混合状態  $C_{n,i}$  の  $l$  番目の要素  $c(n, i, l)$  で表される。この場合、出力信号と教師情報との誤差は下記のように計算される。

10

#### 【 数 1 】

$$E_{(n,i,l)} = (y_l(X^n_i) - c(n, i, l))^2$$

#### 【 0 0 4 3 】

誤差逆伝搬法を用いて、このように得られた誤差を出力層から入力層へと順次逆伝搬することにより、CNN の学習を行うことができる。例えば、確率的勾配降下法等を用いて CNN における各層の重み係数を更新することができる。CNN の重み係数の初期値としては、ランダムな値を用いることもできるし、何らかのタスクに関する学習により得られた重み係数を用いてもよい。例えば、画像分類タスクにおいては画像ごとにクラスラベルが与えられた学習画像を用いるが、領域分割タスクにおいては画素ごとにクラスラベルが与えられた学習画像を用いるため、領域分割タスク用の学習画像を人間が用意するための負荷は大きい。一方、画像分類タスク用の学習画像は一般に公開されており、簡単に入手することができる。例えば ILSVRC (ImageNet Large-scale Visual Recognition Challenge) では 120 万枚の画像分類タスク用の学習画像が公開されている。よって、このような画像分類タスクのために CNN の学習を行い、この学習により得られた重み係数を初期値として用いて、本実施形態のような混合状態推定タスクのための学習を行ってもよい。

20

30

#### 【 0 0 4 4 】

ここでは CNN を用いた推定器について説明したが、推定器の構成は特に限定されない。図 8 (B) には、推定部 1 2 0 0 が処理のために用いることができる構成の別の一例を示す。図 8 (B) には、特徴抽出部 6 1 0 が行う処理に相当する部分、及び推定器 6 5 0 が行う処理に相当する部分が示されている。推定器 6 5 0 は、特徴抽出部 6 1 0 における各層の出力信号を連結して得られた一つの特徴量に対する回帰値を与える。推定器 6 5 0 が用いる手法としては、例えば SVR (Support Vector Regression) やロジスティック回帰等が挙げられるが、特に限定されない。そして、学習画像を用いて、この推定器 6 5 0 が用いる回帰関数の学習を行うことができる。例えば、上記の出力信号と教師情報との誤差に基づく誤差関数を最小化するように、回帰関数の学習を行うことができる。また、図 8 (A) のような構成を用いて CNN の学習をあらかじめ行っておき、その後、その後に CNN の各層の出力信号に基づく特徴量を使って推定器 6 5 0 のみの学習を行ってもよい。ここで、推定器 6 5 0 を完全結合の多層ニューラルネットワークで構成すれば、図 8 (A) の構成と同様に、誤差逆伝搬法を用いて CNN と推定器 6 5 0 との学習を同時に行うこともできる。

40

#### 【 0 0 4 5 】

また、特徴抽出部 6 1 0 は、HOG 又は SIFT のような別の特徴抽出手法を用いて特徴量を抽出することができる。また、推定器は、SVR、ロジスティック回帰、又は多層ニューラルネットワーク等の識別関数を用いて、混合状態の推定を行うことができる。このように、一実施形態においては、特徴抽出手法と推定手法との任意の組み合わせを用い

50

ることができる。このような場合にも、従来の方法に従って推定器の学習を行うことができる。ステップS 2 2 0 0における学習により得られた推定器のパラメータは、推定器記憶部5 2 0 0に記憶される。

【0 0 4 6】

このようにして学習が行われた推定器を用いて、入力画像の混合状態を識別する方法について、図2 ( B )のフローチャートを参照して説明する。S 1 1 0 0において、画像取得部1 1 0 0は、混合状態の識別対象となる入力画像を取得する。画像取得部1 1 0 0は、撮像装置から得られた現像前の画像データを取得することもできる。

【0 0 4 7】

以下で、入力画像の所定領域にある混合状態の推定対象となる画像を対象画像と呼ぶ。画像取得部1 1 0 0は、所定の領域設定パターンに従って、入力画像中に複数の領域を設定することができる。設定された領域のそれぞれに含まれる入力画像の部分画像が、対象画像となる。対象画像は、識別単位に従う所定サイズの部分画像であり、その設定方法は特に限定されない。例えば、学習時と同様、入力画像を所定サイズ(例えば1 6 × 1 6ピクセル)の複数の矩形領域に分割し、それぞれの矩形領域にある複数の対象画像について判定を行うことができる。一方、入力画像の一部領域にある対象画像に対して判定を行ってもよい。

10

【0 0 4 8】

ステップS 1 2 0 0において、推定部1 2 0 0は、特徴抽出部6 1 0を用いて、S 1 1 0 0で得られた入力画像の所定領域にある対象画像から特徴量を抽出する。また、推定部1 2 0 0は、学習された推定器6 2 0を推定器記憶部5 2 0 0から読み込み、推定器6 2 0に特徴量を入力することにより、対象画像における互いに異なる属性の領域間の混合状態を示す情報を生成する。こうして、推定部1 2 0 0は、ステップS 1 1 0 0で取得した入力画像中の対象画像について混合状態を推定する。図8 ( A )は、推定部1 2 0 0が処理のために用いることができるCNNの一例を示す。図8 ( A )には、特徴抽出部6 1 0が行う処理に相当する部分が示されており、ここでは入力画像中の所定領域における信号が各層に順伝搬され、対象画像の特徴量 $X_i$ が抽出される。また、図8 ( A )には、推定器6 2 0に相当する部分が示されており、ここでは得られた特徴量 $X_i$ から、混合状態ベクトルの各要素に割り当てられた出力素子6 2 1における出力信号が生成される。各素子1の出力信号の値は、混合状態ベクトルの各要素 $y_1 ( X_i )$ の値となる。

20

30

【0 0 4 9】

ステップS 1 3 0 0において、出力部1 3 0 0は、ステップS 1 2 0 0で得られた推定結果を出力する。出力部1 3 0 0が行う処理は、識別結果の利用方法に依存し、特に限定されない。混合状態を示す情報を用いた処理例を以下に挙げる。

【0 0 5 0】

例えば、入力画像の各領域に対する画像処理を、その領域における混合状態に応じて変更することができる。この場合、出力部1 3 0 0は画像補正アプリケーションに対して各領域の混合状態を出力することができる。

【0 0 5 1】

また、別の例として、混合状態に応じたカメラのフォーカス制御を行うこともできる。例えば、複数の測距点を備える撮像装置のためのフォーカス制御装置は、取得部と、制御部とを備えることができる。取得部は、撮像装置により得られた画像のうち複数の測距点のそれぞれに対応する領域について、領域に占める特定の属性の領域の面積比を示す情報を取得する。そして、制御部は、面積比に応じて複数の測距点を重み付けし、撮像装置のフォーカス制御を行う。より具体的には、多点測距AFを行う場合に、フォーカスを合わせる対象となる被写体成分がより多い測距点の重みを大きくすることができる。例えば、前景に重点を置くフォーカス制御を行う場合、前景成分がより多い測距点の重みを大きくすることができ、特定の被写体に重点を置くフォーカス制御を行う場合、特定の被写体成分がより多い測距点の重みを大きくすることができる。このようなフォーカス制御装置は、上記の情報処理装置から混合状態を示す情報を取得してもよいし、上記の情報処理装置

40

50

が備える上記の各構成を有していてもよいし、本実施形態とは異なる方法で生成された混合状態を示す情報を取得してもよい。

#### 【0052】

さらなる別の例として、混合状態に応じたカメラの露出制御を行うこともできる。例えば、撮像装置のための露出制御装置は、取得部と、算出部と、選択部と、制御部とを備えることができる。取得部は、撮像装置により得られた画像と、画像の各領域について、領域に占める特定の属性の領域の面積比を示す情報を取得することができる。算出部は、画像全体に占める特定の属性の領域の面積比を算出することができる。選択部は、算出された面積比に応じて、露出制御アルゴリズムを選択することができる。制御部は、選択された露出制御アルゴリズムを用いて、撮像装置の露出制御を行うことができる。より具体的には、視野における空の面積に応じて異なる露出制御を行う場合に、混合状態に基づいて空の面積を算出することができる。この場合、従来技術のように、空と枝が混ざっている領域について、ほとんどの領域を前景であると判定したり、ほとんどの領域について空であると判定したりすることにより、空の面積が実際の値と大きく異なってしまう可能性を減らせることが期待できる。

10

#### 【0053】

ここでは静止画像を学習画像及び入力画像として用いる場合について説明したが、動画を学習画像及び入力画像として用いることもできる。この場合、混合状態の定義は時間方向に拡張される。例えば、 $16 \times 16 \text{ pixel}$ の所定領域及び5フレームを識別単位とする場合、 $16 \times 16 \times 5$ のボクセルに関して混合状態を定義することができる。例えば、面積比を用いて混合状態を表す上記の例を拡張することにより、体積比を用いて混合状態を表すことが可能である。

20

#### 【0054】

本実施形態では、それぞれが複数の画素を含む複数の領域へと入力画像（及び学習画像）が分割され、この領域内の混合状態が推定された。このような処理によれば、全ての画素のそれぞれについてクラスを推定する場合と比較して、推定処理の回数が少なくなるため、処理の高速化が期待できる。一方、入力画像のそれぞれの画素について混合状態を推定することもできる。すなわち、1つの画素に異なるクラスに属する複数の被写体が写っていることがあり、この1つの画素に対応する被写体領域における、それぞれのクラスの被写体の混合状態を推定することもできる。

30

#### 【0055】

本実施形態では、混合状態を示す情報はスカラ値又は複数のスカラ値で構成されるベクトルとして得られた。一方で、混合状態を示す情報は、3つ以上の値から選択される情報でありうる。例えば、所定領域におけるクラス「空」及び「非空」の混合状態を示す情報は、所定領域が「空」で構成されることを示す値、所定領域が「非空」で構成されることを示す値、又は所定領域において「空」及び「非空」が混合されていることを示す値でありうる。このような混合状態を示す情報も、上述の処理例及び後述する実施形態4、5において利用可能である。

#### 【0056】

##### [実施形態2]

実施形態1では、学習画像の各画素に対してクラスラベルが設定されていることを前提にして説明した。しかしながら、画素毎にクラスラベルを設定するには時間がかかる。実施形態2では、学習画像に対してクラスラベルを入力するユーザ作業を軽減する方法を説明する。本実施形態においては、学習画像の各領域に対して入力されたクラスラベルに基づいて、データ取得部2100は、画素ごとのクラスラベルを自動的に算出する。

40

#### 【0057】

以下、図1(C)を参照して、本実施形態における学習装置の基本的な構成を説明する。本実施形態における画像処理装置の構成は実施形態1と同様であり、説明を省略する。本実施形態において学習データ記憶部5100には、識別画像の他に、識別画像における第1の属性の領域、第2の属性の領域、及び第1の属性の領域と第2の属性の領域とが混

50

在している混在領域を示す情報が格納されている。例えば、学習データ記憶部 5 1 0 0 は、学習画像と、学習画像上の各領域に対して付与されたクラスラベルと、を含む学習データが記憶する。ここで、複数のクラスが混在している領域には、混在領域であることを示すクラスラベルが与えられている。

【 0 0 5 8 】

データ取得部 2 1 0 0 は、学習データ記憶部 5 1 0 0 から学習データを読み込む。すなわち、データ取得部 2 1 0 0 は、識別画像の他に、識別画像における第 1 の属性の領域、第 2 の属性の領域、及び第 1 の属性の領域と第 2 の属性の領域とが混在している混在領域を示す情報を取得する。

【 0 0 5 9 】

詳細化部 2 3 0 0 は、第 1 の属性の領域に含まれる画素の画素値、及び前記第 2 の属性の領域に含まれる画素の画素値に基づいて、混在領域の各画素の属性を判定する。例えば、詳細化部 2 3 0 0 は、混在領域であることを示すクラスラベルが与えられている領域について、混合状態を示す教師情報を算出する。詳細については後述する。学習部 2 2 0 0 は、学習画像と混合状態の教師情報とを用いて、実施形態 1 と同様に推定器の学習を行う。

【 0 0 6 0 】

本実施形態において学習装置が行う処理のフローを、図 2 ( C ) に従って説明する。ステップ S 2 1 0 0 においてデータ取得部 2 1 0 0 は、学習データ記憶部 5 1 0 0 から、学習画像とクラスラベルデータとを学習データとして読み込む。学習データ記憶部 5 1 0 0 には、あらかじめ複数の学習画像とそれぞれについてのクラスラベルデータとが用意されている。

【 0 0 6 1 】

ここで、本実施形態におけるクラスラベルデータについて説明する。図 9 ( A ) には学習画像 5 0 0 が示されており、図 9 ( B ) には学習画像 5 0 0 についてのクラスラベルデータ 4 0 0 が示されている。この例では、学習画像 5 0 0 は空領域 4 1 0、非空領域 4 2 0、及び混在領域 4 3 0 から構成されており、それぞれの領域の画素には「空」、「非空」、及び「混在」がそれぞれクラスラベルとして付されている。このように、学習画像 5 0 0 には、単一クラスの領域と、複数クラスが混在している領域と、が設定されている。

【 0 0 6 2 】

これらのクラスラベルは、ツール等を介して予め人間が入力することができる。例えば作業者は、学習画像の空領域及び非空領域を決定することができる。その際、前景の木の枝が細かく複雑になっている箇所においては、空領域と非空領域とを正確に切り分けることは、作業者に対する大きな作業負荷を要求する。そこで、作業者は、このように複数のクラスが混在している領域に対しては、「混在」というクラスラベルを与えることができる。

【 0 0 6 3 】

ここでは、「空」と「非空」が混在している領域について説明したが、実施形態 1 で説明したように、クラス定義はこのようなものに限定されない。また、クラスが 3 クラス以上ある場合には、クラスの組み合わせの数だけ混在領域の種類を定義することができる。例えば、図 5 に示すように「空」、「植物」、「人工物」の 3 クラスが定義されている場合には、「空と植物の混在領域」、「空と人工物の混在領域」、「植物と人工物の混在領域」、「空と植物と人工物の混在領域」、の 4 種類の混在領域クラスを定義できる。以下では、「空」と「非空」の 2 クラスが定義されている場合を例にして説明する。

【 0 0 6 4 】

ステップ S 2 3 0 0 において、詳細化部 2 3 0 0 は、混在領域に関してクラスラベルの詳細化を行う。具体的には、詳細化部 2 3 0 0 は、混在領域の各画素についてクラスラベルを設定する。ここで、詳細化部 2 3 0 0 は、第 1 の属性の領域に含まれる画素の画素値、及び第 2 の属性の領域に含まれる画素の画素値に基づいて、混在領域の各画素の属性を判定する。例えば、詳細化部 2 3 0 0 は、各クラスの色情報を参考に、混在領域のクラス

10

20

30

40

50

ラベルを判定することができる。具体例として、詳細化部2300は、学習画像 $I_n$ における、空領域と非空領域とのそれぞれについて、各画素のRGB値を抽出してRGB色空間にプロットする。混在領域以外の空領域及び非空領域は、学習データに示されている。そして、詳細化部2300は、空領域と非空領域とのそれぞれについて、混合ガウス分布を推定する。すると、混合領域の各画素について、そのRGB値及び空領域の混合ガウス分布に基づいて空領域の尤度を求めることができ、またそのRGB値及び非空領域の混合ガウス分布に基づいて非空領域にある尤度を推定することができる。詳細化部2300は、そして、「空」「非空」のうち尤度が高い方のクラスラベルを画素に割り当てることができる。こうして、詳細化部2300は、混在領域内のクラスラベルを詳細化することができる。図9(C)は、このようにして詳細化されたクラスラベルデータ450を示し、ここには詳細化された空領域460及び非空領域470が表されている。

10

#### 【0065】

このようにして詳細化されたクラスラベルデータを基に、詳細化部2300は、識別単位となる識別領域について混合状態を表す教師情報を算出する。識別領域、並びに混合状態を示す教師情報の定義及び算出方法に関しては、実施形態1で詳しく説明したとおりであるため、ここでは詳細な説明は省く。なお、詳細化部2300がクラスラベルの詳細化を行うことは必須ではない。例えば、識別領域内の混在領域にある画素のRGB値分布と、空領域及び非空領域の混合ガウス分布とに基づいて、混在領域における混合状態を推定することが可能であり、これに基づいて識別領域における混合状態を表す教師情報を算出してもよい。

20

#### 【0066】

変形例として、学習データにおいて、複数のクラスが混在している領域に対しては混合状態が設定されていてもよい。例えば、作業者は、特定の領域について、「非空領域の割合は30%」というような、クラスの面積比を示す情報を入力することができる。この場合、詳細化部2300は、各画素についてのクラスラベルを推定することなく、識別単位となる識別領域について混合状態を表す教師情報を算出することができる。一方、詳細化部2300は、混合状態を参照して、入力画像の各画素のクラスラベルを推定することもできる。この場合には、後述する実施形態5と同様に、学習データから計算可能な混合状態を表す情報と、推定された各画素の属性に基づいて計算される混合状態を示す情報と、の類似度が大きいほど高くなる評価値を用いて、推定を行うことができる。

30

#### 【0067】

##### [実施形態3]

実施形態1, 2では、識別単位となる識別領域は、あらかじめ矩形領域又は小領域として設定されているという前提で説明を行った。一方で、識別領域の大きさや切り方を、さまざまな撮影情報に基づいて変えることができる。例えば、ポケの強い領域では、細かいテクスチャが情報として失われるため、より広い識別領域に対して推定を行うことにより、混合状態の推定精度を向上できる可能性がある。

#### 【0068】

撮影情報は、撮像装置固有の情報と、撮影された画像固有の情報を含む。撮像装置固有の情報としては、センサのサイズ若しくは許容錯乱円径、及び光学系の明るさ若しくは焦点距離等が挙げられる。撮影された画像固有の情報としては、絞り値、合焦距離、BV値、RAW画像、露出時間、ゲイン(ISO感度)、ホワイトバランス係数、距離情報、GPS等による位置情報、日時等の時間情報、等が挙げられる。他にも、撮影された画像固有の情報としては、撮影時における、重力センサ値、加速度、地磁気方向、温度、湿度、気圧、又は高度等が挙げられる。また、可視光以外に、赤外光や紫外光の情報を得ることができる撮像系もある。撮像装置の仕様により、得られる撮影情報はそれぞれ異なる。撮影情報は、入力画像の撮影時に入力画像に関連付けて付された情報、入力画像の撮影時における撮像装置の状態を示す情報、又は入力画像の撮影時に撮像装置によって測定された情報でありうる。また、撮影情報は、入力画像の撮影時に撮像装置によって検知された入力画像の特性を表す情報でありうる。また、撮影情報は、入力画像自体のデータとは異なる

40

50

る情報である。

【0069】

図1(D)に沿って、実施形態3に係る学習装置の基本的な構成を説明する。学習データ記憶部5100には、あらかじめ学習データが記憶されている。本実施形態において学習データは、学習画像と、各学習画像に対応する撮影情報と、学習画像上のさまざまな大きさの領域に対して付与された混合状態の教師情報と、を含む。データ取得部2100は、学習データ記憶部5100から、学習画像、撮影情報、及び教師情報を読み込む。学習部2200は、学習画像と混合状態の教師情報を用いて、混合状態を推定する推定器の学習を行い、得られた推定器を推定器記憶部5200に記憶する。ここで、学習部2200は、第1の領域設定パターンに従って設定された所定領域にある識別画像を用いた学習により第1の推定器を生成し、第2の領域設定パターンに従って設定された所定領域にある識別画像を用いた学習により第2の推定器を生成する。評価部2400は、確認データ記憶部5400から読み込んだ確認データを使って、学習により得られたそれぞれの推定器の推定精度を評価する。そして、評価部2400は、撮影情報と推定精度とに基づいて領域設定器を生成し、設定器記憶部5300に記憶する。

10

【0070】

次に図1(E)に沿って、画像処理装置の装置構成の概要を説明する。画像取得部1100は、入力画像と撮影情報とを取得する。領域設定部1400は、撮影情報に応じて、複数の領域設定パターンの中から、対象画像の設定に用いる領域設定パターンを選択する。本実施形態において領域設定部1400は、領域設定器を設定器記憶部5300から読み込み、撮影情報に従って識別単位となる領域を設定する。推定部1200は、推定器を推定器記憶部5200から読み込み、設定された識別単位に従って設定された所定領域にある対象画像について、推定器を使って混合状態を推定する。

20

【0071】

本実施形態における処理の詳細な説明を以下に記す。まず、学習時の処理に関して図2D)のフローチャートを参照して説明する。ステップS2100において、データ取得部2100は、学習データ記憶部5100から、学習画像、撮影情報、及び混合状態の教師情報を、学習データとして読み込む。

【0072】

ステップS2200において、学習部2200は、データ取得部2100が取得した学習画像と混合状態の教師情報とを用いて、混合状態を推定する推定器の学習を行う。上述のように、本実施形態においては複数種類の領域設定パターンのそれぞれに従って識別単位が設定される。すなわち、識別単位となる領域としては、さまざまなものが用意されている。例えば、 $3 \times 3$ 、 $9 \times 9$ 、及び $15 \times 15$ の矩形領域など、サイズの異なる複数パターンの識別単位を用意することができる。実施形態1でも説明したように、識別単位は矩形領域には限られない。例えば、実施形態1で説明したように、複数の領域設定パターンとして、領域分割により小領域を設定する際に用いるパラメータを複数用意することができる。

30

【0073】

領域設定パターンの違いにより、画像上の同じ位置であっても、混合状態の教師情報は変化しうる。図3(C)には、学習画像の同じ位置にあるさまざまなサイズの矩形領域551、552、及び553が示されている。最も小さい矩形領域551においては、空：非空の面積比は $r = 1$ である。一方、矩形領域552及び553は、それぞれ非空領域を含むため、面積比はそれぞれ $r = 0.9$ 及び $r = 0.8$ となる。

40

【0074】

学習部2200は、それぞれの領域設定パターンに対応する推定器の学習を行う。すなわち、学習部2200は、着目領域設定パターンに従って設定された識別領域と、この識別領域について与えられた教師情報とに基づき、着目領域設定パターンに対応する推定器の学習を行う。この結果、学習部2200は、複数の領域設定パターンのそれぞれに対応する推定器を生成する。例えば、領域設定パターンのインデックスを $q$ とし、領域設定パ

50

ターンの総数を  $Q$  とすると、学習によって  $Q$  種類の推定器  $y_q$  を得ることができる。推定器の学習は実施形態 1 と同様に行うことができる。一例として、それぞれの推定器  $y_q$  は、回帰関数  $f_q(X)$  ( $q = 1, \dots, Q$ ) に従って混合状態の推定を行うことができる。学習により得られた推定器は、推定器記憶部 5200 に記憶される。

【0075】

ステップ S2300 において評価部 2400 は、ステップ S2200 で得られた推定器の識別精度を、撮影情報とともに評価し、領域設定器を生成する。例えば、評価部 2400 は、教師情報及び撮影情報が関連付けられている検証画像を用いて、それぞれの推定器の識別精度を評価することができる。そして、評価部 2400 は、所定の撮影情報が関連付けられている識別画像の判定を行う際に良好な識別精度が得られるように、特定の撮影情報に対応する推定器を示す情報を生成することができる。

10

【0076】

撮影情報の中には、学習画像の画素ごとに得られる情報がある。また、撮影情報の組み合わせにより新たな撮影情報を得ることもできる。例えば、画素位置  $p$  におけるレンズ面から被写体までの距離  $Z(p)$  と、光学系の焦点距離  $f$  と、が撮影情報として得られた場合、像倍率  $S(p)$  を算出することができる。

【数 2】

$$S(p) = \frac{f}{Z(p)}$$

20

【0077】

また、光学系の  $F$  値、焦点距離  $f$ 、撮影時の合焦距離  $Z_f$ 、及び画素位置  $p$  における被写体までの距離  $Z(p)$  が撮影情報として得られた場合、各画素位置におけるボケ量  $B(p)$  を得ることができる。

【数 3】

$$B(p) = \frac{f^2 |Z(p) - Z_f|}{FZ_f(Z(p) - f)}$$

【0078】

さらに、RAW 画像の各画素位置  $p$  における値  $r(p)$ 、 $g(p)$ 、及び  $b(p)$ 、露出時間  $T$ 、ゲイン  $G$ 、並びに絞り量  $F$  が撮影情報として得られた場合、画素位置  $p$  における入射光量  $BV(p)$  の絶対値を得ることができる。

30

【数 4】

$$BV(p) = \log_2 \frac{F^2(0.3r(p) + 0.6g(p) + 0.1b(p))}{0.3GT}$$

【0079】

以下、撮影情報として画素位置  $p$  におけるボケ量  $B(p)$  を用いて、領域設定器を生成する場合について説明する。もっとも、用いる撮影情報はこれには限定されず、像倍率  $S(p)$  又は入射光量  $BV(p)$  等の他の撮影情報を用いてもよい。また、複数の撮影情報を組み合わせてもよく、例えばボケ量  $B(p)$  と入射光量  $BV(p)$  を組み合わせて用いてもよい。

40

【0080】

まず、評価部 2400 は、ボケ量  $B$  を複数のピンに区切り、領域設定パターン  $q$  に関するテーブルを生成する。この例では、ボケ量  $B$  が 2 未満、2 以上 3 未満、3 以上 4 未満、4 以上、の 4 つのピンに区切られている。また、領域設定パターン  $q$  としては  $3 \times 3$ 、 $9 \times 9$ 、及び  $15 \times 15$  の 3 種類が用いられており、 $3 \times 4$  のテーブルが得られる。

【0081】

次に、評価部 2400 は、確認データを確認データ記憶部 5400 から読み込む。確認データは、学習データと同様、複数の確認画像、それぞれの確認画像についてのクラストラ

50

ベルデータ、及び撮影情報を含む。ここでは、確認画像の総枚数を  $N_v$  と表し、 $v$  番目の確認画像を  $I_v$  ( $v = 1, \dots, N_v$ ) と表す。

【0082】

評価部2400は、領域設定パターン $q$ のそれぞれに従って、確認画像中の識別単位となる領域 $i$ における特徴量を抽出し、対応する推定器に入力する。こうして、領域設定パターン $q$ を用いた場合の確認画像 $I_v$ 中の領域 $i$ の混合状態の推定値 $y_{q,i}$  ( $X_{i,v}$ )を得ることができる。このとき、混合状態教師情報 $c_q(v, i)$ に対する二乗誤差は下記のように表すことができる。

【数5】

$$Err_{v,q}(i) = (y_{q,i}(X_{i,v}) - c_q(v, i))^2$$

10

【0083】

また、ボケ量 $B$ と領域設定パターン $q$ との組み合わせに対するピン $(B, q)$ における二乗誤差平均 $MSE(B, q)$ は下記のように表される。

【数6】

$$MSE(B, q) = \frac{\sum_v \sum_i \delta_B(v, i) Err_{v,q}(i)}{\sum_v \sum_i \delta_B(v, i)}$$

ここで $\delta_B(v, i)$ は、確認画像 $I_v$ の領域 $i$ の中心位置におけるボケ量がピン $B$ の範囲内であるときに1、そうでないときに0を返すものとする。

20

【0084】

そして、ピン $(B, q)$ に関する信頼度 $T(B, q)$ は、1から二乗平均平方根誤差を減じた値として定義できる。

【数7】

$$T(B, q) = 1 - \sqrt{MSE(B, q)}$$

【0085】

このようにして、評価部2400は、各ピン $(B, q)$ に対する信頼度 $T(B, q)$ のテーブルを得ることができる。こうして得られたテーブルの例を下に示す。評価部2400は、このようにして得られたテーブルを、領域設定器として設定器記憶部5300に記憶する。

30

【表1】

[表1]

		ボケ量 $B(p)$ [pix]			
		2未満	2以上 ~3未満	3以上 ~4未満	4以上
領域 サイズ $q$	3×3	0.9	0.8	0.3	0.2
	9×9	0.6	0.5	0.8	0.5
	15×15	0.4	0.3	0.5	0.7

40

【0086】

本実施形態では、得られたテーブルが領域設定器として設定器記憶部5300に格納された。一方、評価部2400は、信頼度 $T(B, q)$ の値を教師情報として、ボケ量 $B$ に対する信頼度 $T$ を回帰値として出力する回帰関数 $g_q(B)$ を各領域設定パターン $q$ に対して生成し、これを領域設定器として用いてもよい。

【0087】

このように得られた混合状態推定器及び領域設定器を用いて、入力画像の混合状態を推定する処理について、図2(E)のフローチャートを用いて説明する。ステップS1100において画像取得部1100は、撮像装置により得られた画像データと撮影情報とを取

50

得する。

【 0 0 8 8 】

ステップ S 1 4 0 0 において領域設定部 1 4 0 0 は、設定器記憶部 5 3 0 0 から領域設定器を読み込み、撮影情報に従って使用する領域設定パターンを決定する。例えば、領域設定部 1 4 0 0 は、下式に従って、入力画像 I の各領域 i に対して、撮影情報として得られたボケ量  $B(i)$  から得られる信頼度 T が最も大きくなる領域設定パターン  $q_{win}$  を選ぶことができる。なお、ボケ量  $B(i)$  は、入力画像 I の領域 i の中心位置におけるボケ量を指す。具体的な処理は特に限定されないが、例えば、1つの領域設定パターンに従って入力画像 I を複数の領域に分割し、1つの領域を別の領域設定パターンに従って細分化した方が信頼度が高くなる場合に、この細分化を行うことができる。別の例としては、ボケ量が類似している領域を連結し、それぞれの連結領域についてボケ量に従う領域設定パターンを用いて領域分割を行うことができる。

10

【 数 8 】

$$q_{win}(i) = \arg \max_q T(B(i), q)$$

【 0 0 8 9 】

ステップ S 1 2 0 0 において、推定部 1 2 0 0 は、推定器記憶部 5 2 0 0 から推定器を読み込み、入力画像の各位置における混合状態を推定する。具体的には、推定部 1 2 0 0 は、各位置 p において設定された所定領域の画像の特徴量を抽出し、抽出された特徴量を推定器に入力することにより、この位置 p における混合状態を推定することができる。ここで、各位置 p についての所定領域は、ステップ S 1 4 0 0 で決定された領域設定パターン  $q_{win}$  に従って設定される。上述のように、本実施形態においては、複数の領域設定パターンのそれぞれに対応する推定器が生成されている。したがって、推定部 1 2 0 0 は、ステップ S 1 4 0 0 で決定された領域設定パターンに従って複数の推定器から選択された推定器を用いることができる。例えば、位置 p における推定器としては  $y_{q_{win}}$  が選択され、位置 p における所定領域の混合状態の推定値は、 $y_{q_{win}}(X_i)$  として得られる。図 3 ( D ) は、画像上の位置によって領域設定方法を変えた場合の例を示し、それぞれの矩形が推定器に入力される 1 つの領域を示す。

20

【 0 0 9 0 】

ステップ S 1 3 0 0 に係る処理は実施形態 1 と同様であるため、説明は省略する。本実施形態のように、撮影情報を利用して混合状態を推定する識別単位となる領域の設定方法を変えることにより、より誤差の少ない混合状態の推定を行うことができる。

30

【 0 0 9 1 】

[ 実施形態 4 ]

実施形態 1 ~ 3 では、識別単位となる所定領域における混合状態を推定した。実施形態 4 では、得られた混合状態の推定結果を用いて、領域を細分化することにより、詳細な領域分割結果を得る方法について説明する。学習装置及び画像処理装置の基本的な構成は実施形態 1 と同様であり、説明を省略する。

【 0 0 9 2 】

以下、学習時の処理について図 2 ( A ) のフローチャートに従って説明する。ステップ S 2 1 0 0 においてデータ取得部 2 1 0 0 は、学習データ記憶部 5 1 0 0 から、学習画像と混合状態の教師情報と学習データとして読み込む。

40

【 0 0 9 3 】

ステップ S 2 2 0 0 で学習部 2 2 0 0 は、実施形態 3 と同様の処理を行う。すなわち、識別単位としては様々な大きさの領域が用意される。例えば、複数の領域設定パターンのそれぞれに従って、 $1 \times 1$ 、 $3 \times 3$ 、 $9 \times 9$ 、及び  $15 \times 15$  等、異なるサイズの矩形領域を複数パターン用意することができる。そして、学習部 2 2 0 0 は、それぞれの領域サイズに対応する推定器の学習を、実施形態 3 と同様にそれぞれの領域サイズについて得られた混合状態の教師情報を用いて行うことができる。すなわち、領域サイズのインデックスを  $q$  とし、領域サイズの総数を  $Q$  とすると、学習によって  $Q$  種類の推定器  $y_q$  ( $q = 1$

50

, ... , Q) を得る事ができる。一例として、それぞれの推定器  $y_q$  は、回帰関数  $f_q(X)$  に従って混合状態の推定を行うことができる。学習により得られた推定器  $y_q$  は、推定器記憶部 5200 に書きこまれる。

#### 【0094】

次に、判定時の処理に関して、図 2 (F) のフローチャートに従って説明する。ステップ S1100 において、画像取得部 1100 は入力画像を取得する。ステップ S1200 において、推定部 1200 は、推定器を用いて入力画像中の所定領域における混合状態を推定する。ここで、推定部 1200 は、複数の領域設定パターンのうちの第 1 の領域設定パターンを用いて領域設定を行う。すなわち、推定部 1200 は、第 1 の領域設定パターンに従う大きさの第 1 の対象画像について混合状態を判定する。本実施形態において、識別単位としては、Q 種類の領域サイズのうち最も大きいサイズが用いられる。前述の例では、識別単位として  $15 \times 15 \text{ pixel}$  が選択され、また推定器としては  $15 \times 15 \text{ pixel}$  に対応する推定器が用いられる。

10

#### 【0095】

そして、推定部 1200 は、入力画像の第 1 の部分にある第 1 の対象画像について推定された混合状態を示す情報に従って、第 1 の部分の混合状態を再判定するか否かを判定する。例えば、推定部 1200 は、混合状態の推定を行った所定領域について混合状態の再判定を行うか否かを判定する。例えば、推定部 1200 は、クラス純度が閾値以上である領域については、このクラス推定結果を採用する。

#### 【0096】

一方、推定部 1200 は、クラス純度が閾値よりも低い領域については、この領域について混合状態の再判定を行う。再判定を行うとの判定に応じて、推定部 1200 は、第 1 の部分にある、第 2 の領域設定パターンに従う大きさの第 2 の対象画像の混合状態を示す情報を出力する。ここで、第 2 の対象画像は第 1 の対象画像よりも小さい。すなわち、推定部 1200 は、クラス純度が閾値よりも低い領域について、より小さい識別単位に従って領域を再分割し、再分割された領域のそれぞれについて、再び推定器を用いて混合状態の推定を行う。推定部 1200 は、例えば、一段階小さい領域サイズを用いて再分割を行うことができる。上述のように、本実施形態においては、複数の領域設定パターンのそれぞれに対応する推定器が生成されている。したがって、推定部 1200 は、再分割に用いた領域設定パターンに従って複数の推定器から選択された推定器を用いることができる。

20

30

#### 【0097】

ここで、クラス純度とは、領域内における同一クラスラベルが割り振られている画素の割合を示す。例えば、実施形態 1 に示した面積比  $r$  の値が  $0.8$  以上又は  $0.2$  以下である場合に、クラス純度が高いと定義することができる。図 7 に示すマップを用いる場合、 $p_1 = 0.9$  かつ  $p_2 = 0.8$  の場合にクラス純度が高いと定義することもできる。

#### 【0098】

このように、クラス純度が低い領域については細分化及び混合状態の再推定を行うことにより、詳細な領域分割を行うことができる。領域が細分化できなくなるか、すべての領域のクラス純度が閾値以上になると、処理はステップ S1300 へと進むことができる。ステップ S1300 における処理は実施形態 1 と同様であるため、説明は省略する。このようにして得られた詳細な領域分割結果は、領域別のトーンマッピング又はホワイトバランス調整等の高画質化処理に利用することができる。

40

#### 【0099】

#### [実施形態 5]

実施形態 4 では、識別単位を細分化していくことによって詳細な領域分割結果を算出したが、領域分割の方法はこの方法には限定されない。実施形態 5 では、各領域についての混合状態の推定結果を利用して、画素単位のクラス判定を行うことにより、詳細な領域分割結果を得る方法について説明する。

#### 【0100】

本実施形態に係る画像処理装置の基本構成を図 1 (F) に示す。画像取得部 1100 及

50

び推定部 1200 の機能は実施形態 1 と同様であるため説明を省略する。判定部 1500 は、対象画像の各画素の属性を判定する。判定部 1500 は、評価値に基づいて各画素の属性を決定し、この評価値が示す評価は、各画素の属性に基づいて計算される混合状態を示す情報と、推定部 1200 により得られた混合状態を示す情報と、の類似度が大きいほど高くなる。本実施形態においては、判定部 1500 は、混合状態推定結果及び画像情報に基づいて、入力画像の各画素のクラスラベルを推定する。出力部 1300 は、入力画像の各画素について推定されたクラスラベルを示す情報を出力する。

#### 【0101】

本実施形態に係る判定処理の詳細を図 2 (G) に従って説明する。ステップ S1100 及び S1200 における処理は、実施形態 1 と同様であるため説明を省略する。ステップ S1500 において判定部 1500 は、ステップ S1200 にて推定された各領域の混合状態を利用して、入力画像の各画素についてクラスを推定する。例えば、推定された各画素のクラスに従って求められる混合状態が、ステップ S1200 にて推定された混合状態に近くなるように、各画素のクラスの推定を行うことができる。各画素のクラスの推定には、さらに各画素の色情報を用いて、例えば同じクラスに属する画素の色が類似するように、又は異なるクラスに属する画素の色が類似しないように、行うことができる。

#### 【0102】

入力画像の各画素のクラスを推定する方法の一例として、CRF (Conditional Random Field、条件付き確率場) のような繰り返し処理を利用する場合について以下説明する。CRF は、複数のノードからなるグラフに対して、対となるノード間の類似度による pairwise potential と、各ノードの持つ unary potential を考慮して、各ノードの状態を安定した状態まで逐次推移させていく方法である。画像の画素判別に CRF を使う場合は、各ノードが画像の各画素に対応する CRF モデルを使うことができる。

#### 【0103】

入力画像 I 上の画素 i のクラスラベル  $c_i$  の条件付き確率は、下式で表すことができる。

#### 【数 9】

$$\log P(c_i | I, \theta_\phi, \theta_\psi) = \varphi_i(c_i, I; \theta_\phi) + \sum_{j \in \mathcal{E}_i} \phi(c_i, c_j, g_{ij}(I); \theta_\psi) - \log Z(\theta, I)$$

ここで、右辺第一項の  $\varphi_i$  は unary potential、右辺第二項の  $\phi$  は pairwise potential を表す。  $\theta_\phi$  及び  $\theta_\psi$  はそれぞれのパラメータであり、後述する学習処理において算出される。  $\mathcal{E}_i$  は、画素 i の近傍画素の集合である。  $g_{ij}$  は画素 i と画素 j との相互関係を表す関数であり、Z は正規化項である。判定部 1500 は、このモデル式に従って各画素のクラスラベルを更新していくことによって、画像全体のポテンシャルが高い状態へと判定結果を収束させていく。

#### 【0104】

pairwise potential は下式で表すことができる。

#### 【数 10】

$$g_{ij}(I) = \exp\left(-\beta \|x_i - x_j\|^2\right)$$

$$\phi(c_i, c_j, g_{ij}(I); \theta_\psi) = -\theta_\psi^T g_{ij}(I) \delta(c_i \neq c_j)$$

ここで、  $x_i$  及び  $x_j$  はそれぞれ画素 i 及び画素 j の色情報であり、RGB 値を持つ 3 次元ベクトルで表わされる。  $\beta$  はユーザの定義するハイパーパラメータであって、  $\beta = 1$  などと設定することができる。このように、pairwise potential は、時刻 t において違うクラスに属する画素の色が類似する場合に評価が低くなるように設定することができる。

10

20

30

40

50

## 【 0 1 0 5 】

unary potential は下記で表すことができる。

## 【 数 1 1 】

$$\varphi_i(c_i, I; \theta_\varphi) = \|y_c(\mathbf{X}_i) - L_c^i(t)\|$$

ここで  $y_c(\mathbf{X}_i)$  は、画素位置  $i$  におけるクラス  $c$  に関する混合状態推定値である。  
 $y_c(\mathbf{X}_i)$  は、画素位置  $i$  が含まれる所定領域について推定部 1200 が推定した混合状態に基づいて算出することができ、例えば所定領域内における、クラス  $c$  の面積比、エッジ画素率、又はクラスラベル配置パターン等でありうる。このように、unary potential は、時刻  $t$  において各画素の属性に基づいて計算される混合状態と、推定部 1200 により得られた混合状態と、の類似度が大きいほど評価が高くなる。

10

## 【 0 1 0 6 】

$L_c^i(t)$  は、CRF に従って画素単位のクラスラベルが遷移していったときの、時刻  $t$  における、画素  $i$  が含まれる所定領域のクラス  $c$  の混合状態である。 $L_c^i(t)$  は、推定部 1200 が推定する混合状態と同種の情報であり、推定部 1200 が時刻  $t$  において所定領域内の各画素について推定されているクラスを参照して算出できる。実施形態 1 で説明した混合状態の例に従って、以下に具体的な例を挙げる。例えば、遷移途中の時刻  $t$  において、画素  $i$  が含まれる所定領域内におけるクラスラベル  $c$  が割り当てられている画素をカウントすることにより、クラス  $c$  の面積比  $r(t)$  を求めることができる。また、所定領域内におけるクラスラベルの配置に従ってエッジ画素を抽出してカウントすることにより、エッジ画素率  $e(t)$  を求めることができる。さらに、所定領域内におけるクラスラベルの配置を、図 7 に示されるマップのどれに最も近いかを判定することにより、クラスラベル配置パターン  $p(t)$  を求めることができる。実施形態 1 で説明したように、 $L_c^i(t)$  は、これら時刻  $t$  における混合状態の組み合わせによって表すこともできる。

20

## 【 0 1 0 7 】

このように、推移中の時刻  $t$  における画素レベルでのクラスラベル配置に基づいて決定された時刻  $t$  における所定領域内の混合状態と、所定領域についての混合状態推定値と、の類似度を、unary potential として表現することができる。具体的には、時刻  $t$  における所定領域内の混合状態と、所定領域についての混合状態推定値と、の類似度が大きいほど評価が高くなるように、unary potential を表現することができる。

30

## 【 0 1 0 8 】

本実施形態における学習処理について、図 2 (A) に従って説明する。ステップ S2100 でデータ取得部 2100 は、学習画像と教師データとを取得する。ステップ S2200 で学習部 2200 は、実施形態 1 と同様に推定器の学習を行う。また、学習部 2200 は、入力画像の各画素のクラスを推定する際に用いるパラメータ（例えば上述の  $\theta_\varphi$  及び  $\theta_\phi$ ）の値を決定する。学習部 2200 は、複数の学習画像と、学習画像の各画素のクラスを示すクラスラベルのデータを用いて、この処理を行うことができる。クラスラベルのデータとしては、例えば、実施形態 2 に従って図 9 (C) のように作成されたものを用いることができる。本実施形態において学習部 2200 は、全学習画像に対するポテンシャルが最大となるように  $\theta_\varphi$  及び  $\theta_\phi$  の値を算出することができる。すなわち、下式をそれぞれ最大化する  $\theta_\varphi$  と  $\theta_\phi$  の値を、勾配法等によって求めることができる。

40

## 【 数 1 2 】

$$\sum_n \sum_i \varphi_i(c_i, I_n; \theta_\varphi)$$

$$\sum_n \sum_i \sum_{j \in \mathcal{E}_i} \phi(c_i, c_j, g_{ij}(I); \theta_\phi)$$

50

## 【0109】

学習部2200は、得られたパラメータを、推定器とともに推定器記憶部5200に格納する。本実施形態では 及び の値が推定器記憶部5200に記憶され、判定部1500によって上述のように用いられる。こうして得られた画素ごとのクラスラベルのデータは、実施形態4と同様にして、領域ごとに高画質化処理を行う場合等に利用することができる。

## 【0110】

混合状態の推定結果を利用して画素単位のクラス判定を行う方法は、上記の方法には限定されない。例えば、実施形態2と同様に、クラスが確定している領域を用いて求められた各クラスの混合ガウス分布と、上記の混合状態の類似度と、に基づいて、画素単位のクラス判定を行うこともできる。

10

## 【0111】

本実施形態に係る処理は、混合状態として、面積比、エッジ画素率、及びクラスラベル配置パターンのいずれを用いても可能であるし、利用可能な混合状態がこれらに限られるわけでもない。また、複数の表現を組み合わせて表現された混合状態を用いることにより、判定精度を向上させることができる。例えば、面積比に加えてエッジ画素率を用いることにより、建物と空との境界のように輪郭が単純な場合と、枝と空との境界のように輪郭が複雑な場合と、を区別することが可能となる。

## 【0112】

本実施形態において、混合状態を示す情報は推定部1200の処理により得られた。しかしながら、判定部1500は、異なる方法により得られた混合状態を示す情報を取得し、同様の方法で各画素の属性を判定することもできる。

20

## 【0113】

(その他の実施例)

本発明は、上述の実施形態の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、ASIC)によっても実現可能である。

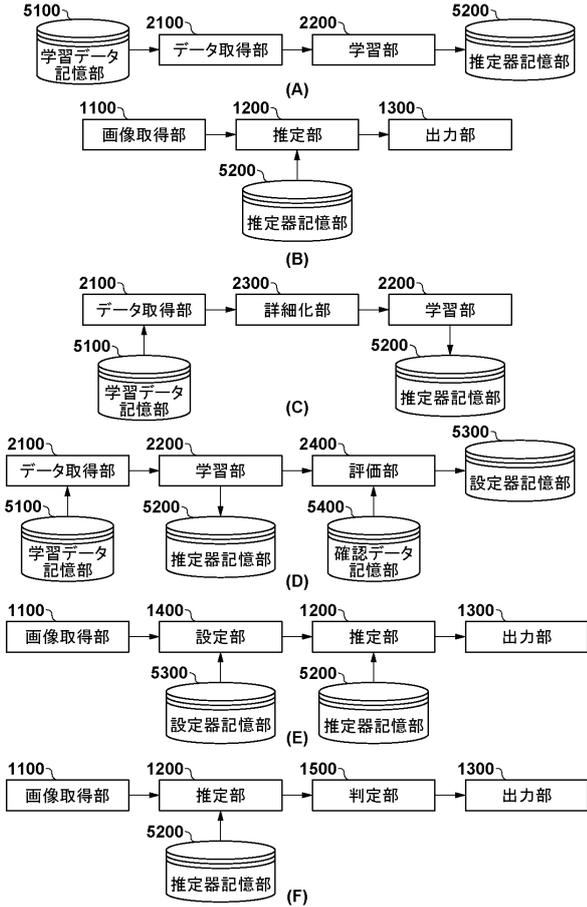
## 【符号の説明】

## 【0114】

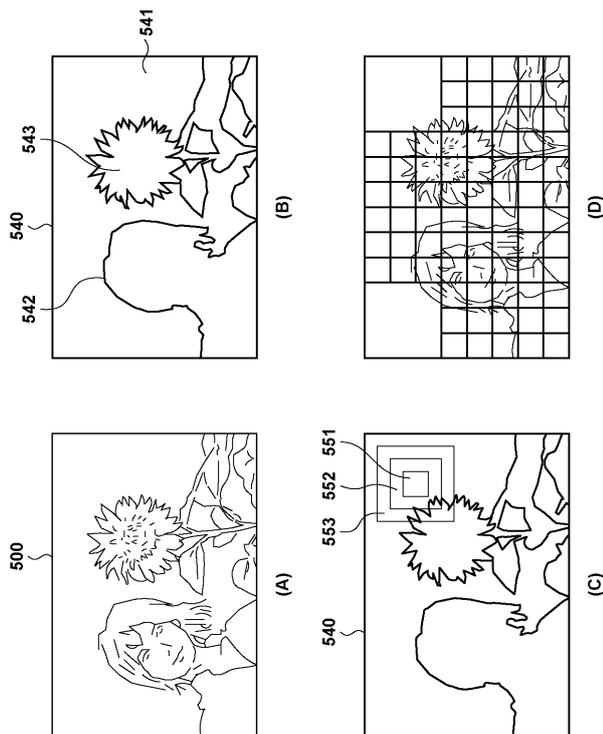
1100：画像取得部、1200：推定部、1300：出力部、1400：領域設定部、1500：判定部、2100：データ取得部、2200：学習部、2300：詳細化部

30

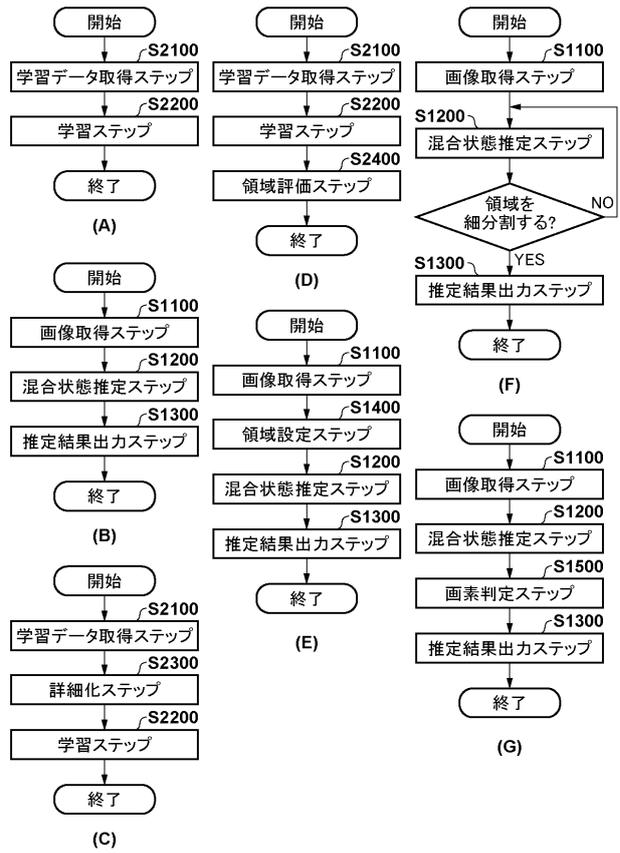
【図1】



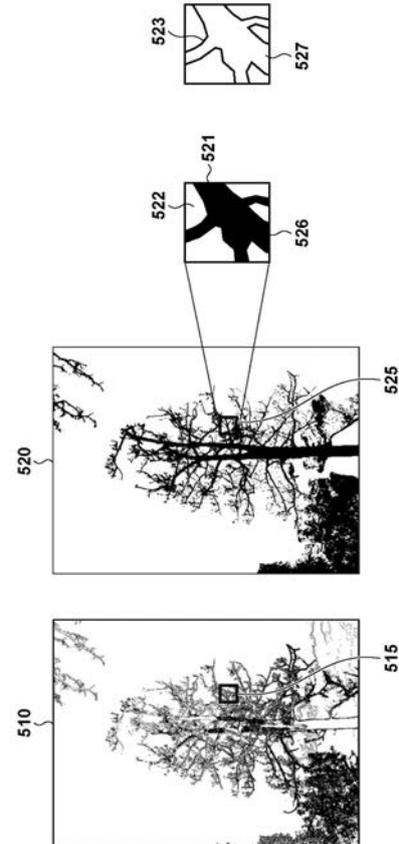
【図3】



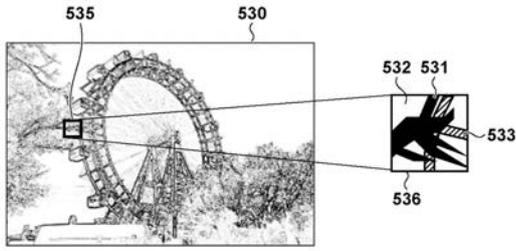
【図2】



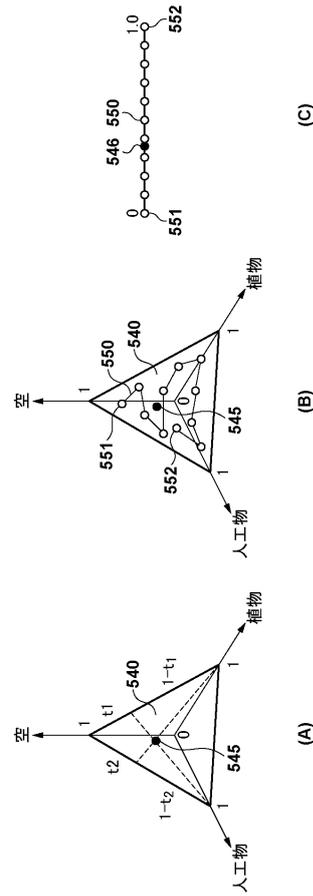
【図4】



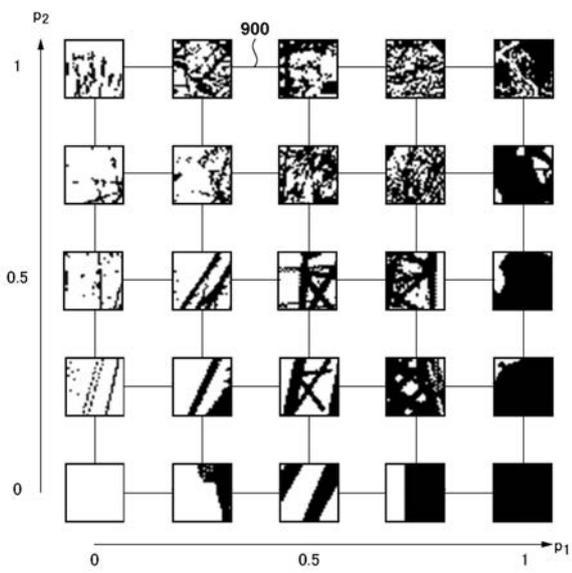
【 図 5 】



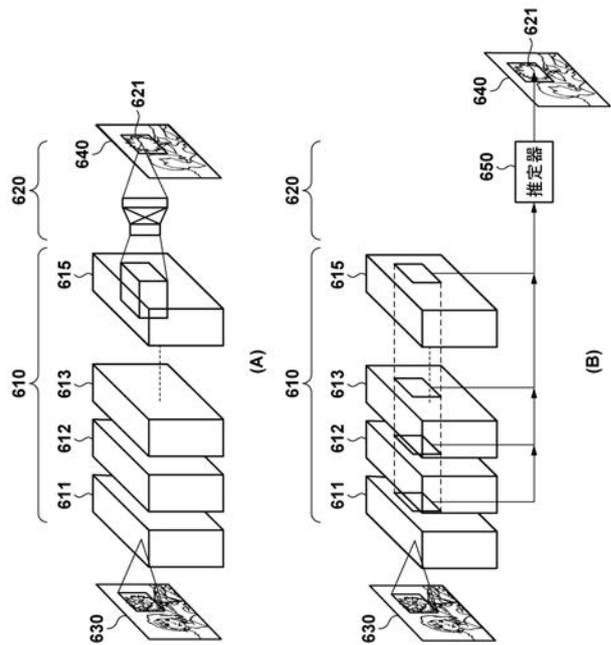
【 図 6 】



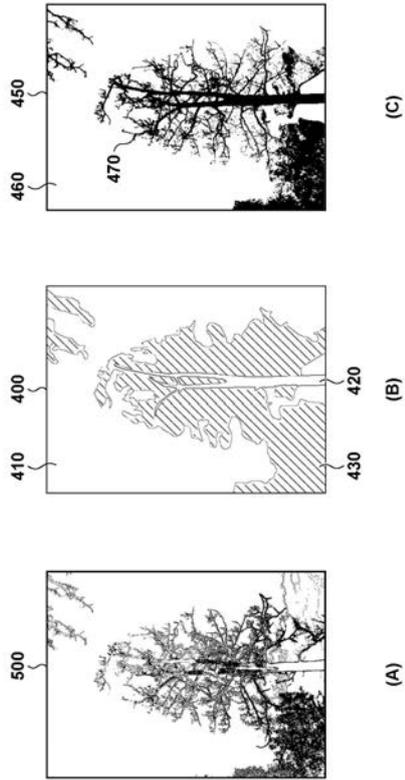
【 図 7 】



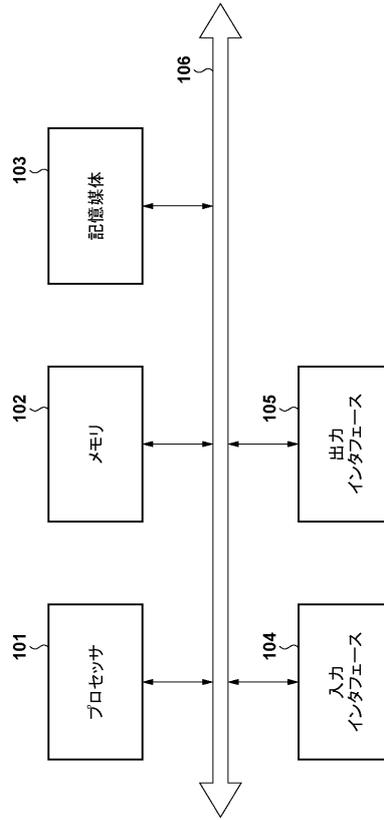
【 図 8 】



【図 9】



【図 10】



---

フロントページの続き

(72)発明者 青葉 雅人

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

(72)発明者 馬養 浩一

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

Fターム(参考) 5C122 DA01 EA06 FD01 FD13 FF01 FF18 FF23 FH09 FH10 FH11  
HA88 HB01 HB09 HB10  
5L096 EA39 FA59 FA81 JA03 JA11 KA04 MA07