

(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2009-258953
(P2009-258953A)

(43) 公開日 平成21年11月5日(2009.11.5)

(51) Int.Cl.	F 1	テーマコード (参考)
G06T 1/00 (2006.01)	G06T 1/00 200A	5B050
G06F 17/30 (2006.01)	G06F 17/30 170B	5B075
	G06F 17/30 210D	

審査請求 未請求 請求項の数 30 O L (全 23 頁)

(21) 出願番号 特願2008-106546 (P2008-106546)
(22) 出願日 平成20年4月16日 (2008. 4. 16)

特許法第30条第1項適用申請有り 平成20年2月7日 国立大学法人電気通信大学主催の「電気通信大学卒業論文発表会」に文書をもって発表

(71) 出願人 504133110
国立大学法人電気通信大学
東京都調布市調布ヶ丘一丁目5番地1
(74) 代理人 110000626
特許業務法人 英知国際特許事務所
(72) 発明者 柳井 啓司
東京都調布市調布ヶ丘1丁目5番地1 国立大学法人 電気通信大学内
Fターム(参考) 5B050 BA15 BA17 EA04 EA08 EA12 GA08
5B075 ND06 NK39 NR12 PQ05 PR06 UU40

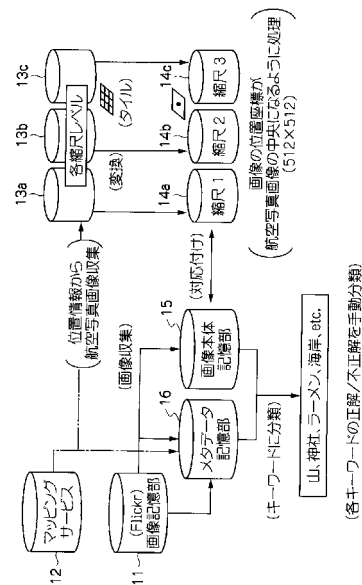
(54) 【発明の名称】 画像処理方法、その方法を実行するプログラム、記憶媒体、撮像機器、画像処理システム

(57) 【要約】

【課題】 GPSにより取得した撮影場所の位置情報（緯度経度）付きのデジタル画像に対する、一般的な物体やシーン等の被写体認識の精度を向上させ、デジタル写真の自動分類や検索を可能にする。

【解決手段】 認識対象（被写体）の画像に加え、撮影位置を中心とした小領域に対応する、様々な縮尺の航空写真画像および/または地図画像の画像特徴量を、認識対象のデジタル画像の画像特徴量に付加して利用する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

画像内の認識対象を分類するための分類器を用いて、認識対象画像における前記認識対象を分類する認識処理を含む画像処理方法であって、

認識対象画像を入力するステップ(S105)と、

前記認識対象画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成するステップ(S110)と、

前記分類器を用いて認識結果を得るステップ(S135)と、

前記認識対象画像における認識対象の有無を判断するステップ(S140、S145、S150)と、

を含むことを特徴とする画像処理方法。

10

【請求項 2】

請求項 1 記載の画像処理方法において、さらに、

前記小領域パッチ画像に基づき画像特徴量を抽出するステップ(S115)と、

前記抽出された画像特徴量からヒストグラムを作成するステップ(S120)と、

前記作成されたヒストグラムに最も近い特徴ベクトルをコードブックより選択するステップ(S125)と、

前記選択された特徴ベクトルを正規化するステップ(S130)と、

を備えたことを特徴とする画像処理方法。

20

【請求項 3】

前記小領域パッチ画像を生成するステップ(S110)が、前記航空写真画像および/または地図画像から生成された、一つの画像から小領域パッチ画像を生成するステップであることを特徴とする請求項 1 又は 2 記載の画像処理方法。

【請求項 4】

前記小領域パッチ画像を生成するステップ(S110)が、前記航空写真画像および/または地図画像から生成された、縮尺の異なる複数の画像から小領域パッチ画像を生成するステップであることを特徴とする請求項 1 又は 2 記載の画像処理方法。

【請求項 5】

前記小領域パッチ画像を生成するステップ(S110)が、前記認識対象画像、および、前記航空写真画像および/または地図画像から生成された、縮尺の異なる複数の画像から前記小領域パッチ画像を生成するステップであることを特徴とする請求項 1 又は 2 記載の画像処理方法。

30

【請求項 6】

前記ヒストグラムを作成するステップ(S120)が、前記抽出された複数の画像特徴量から生成されたそれぞれのヒストグラムを接続して、一つのヒストグラムを生成するステップであることを特徴とする請求項 2 乃至 5 のいずれか一項に記載の画像処理方法。

【請求項 7】

前記ヒストグラムを作成するステップ(S120)が、前記認識対象画像の画像特徴量から生成されたヒストグラムと、前記小領域パッチ画像の画像特徴量から生成されたヒストグラムと、を接続して、一つのヒストグラムを生成するステップであることを特徴とする請求項 2 乃至 5 のいずれか一項に記載の画像処理方法。

40

【請求項 8】

前記分類器が、

学習画像および該学習画像の分類を入力するステップ(S91)と、

前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成するステップ(S92)と、

前記分類器を作成するステップ(S96)と、

によって生成されることを特徴とする請求項 1 乃至 7 のいずれか一項に記載の画像処理方法。

【請求項 9】

50

請求項 8 記載の画像処理方法において、さらに、
前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成するステップ(S 9 2)と、
前記小領域パッチ画像の画像特徴量を抽出するステップ(S 9 3)と、
前記抽出された画像特徴量からコードブックを作成するステップ(S 9 4)と、
前記コードブックを用いて、前記抽出された画像特徴量からヒストグラムを作成するステップ(S 9 5)と、を有することを特徴とする画像処理方法。

【請求項 10】

前記小領域パッチ画像を生成するステップ(S 9 2)が、前記学習画像の撮影位置に対応する、縮尺の異なる複数の航空写真画像および/または地図画像から、縮尺の異なる複数の小領域パッチ画像を生成するステップ(S 9 2)であることを特徴とする請求項 9 記載の画像処理方法。

10

【請求項 11】

前記小領域パッチ画像を生成するステップ(S 9 2)が、前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から一つの小領域パッチ画像を生成するステップ(S 9 2)であることを特徴とする請求項 9 記載の画像処理方法。

【請求項 12】

前記小領域パッチ画像を生成するステップ(S 9 2)が、前記認識対象画像、および、前記航空写真画像および/または地図画像から生成された、縮尺の異なる複数の画像から前記小領域パッチ画像を生成するステップであることを特徴とする請求項 9 記載の画像処理方法。

20

【請求項 13】

前記ヒストグラムを作成するステップ(S 9 5)が、前記抽出された複数の画像特徴量から生成されたそれぞれのヒストグラムを接続して、一つのヒストグラムを生成するステップであることを特徴とする請求項 9 記載の画像処理方法。

【請求項 14】

前記ヒストグラムを作成するステップ(S 9 5)が、前記認識対象画像の画像特徴量から生成されたヒストグラムと、前記小領域パッチ画像の画像特徴量から生成されたヒストグラムと、を接続して、一つのヒストグラムを生成するステップであることを特徴とする請求項 9 記載の画像処理方法。

30

【請求項 15】

前記航空写真画像および/または地図画像が、ネットワークを介してアクセス可能なデータベースに格納されていることを特徴とする請求項 1 乃至 14 のいずれか一項に記載の画像処理方法。

【請求項 16】

前記認識対象画像が位置情報を保有しており、該位置情報に基づき、前記航空写真画像および/または地図画像を、前記認識対象画像と対応させることを特徴とする請求項 1 乃至 15 のいずれか一項に記載の画像処理方法。

【請求項 17】

請求項 1 乃至 16 のいずれか一項に記載の画像処理方法を、コンピュータ、または画像分類機能付き撮像機器、または画像処理システム、に実行させるための画像処理用プログラム。

40

【請求項 18】

請求項 17 記載の画像処理用プログラムを、コンピュータにより読み出され実行可能なプログラムとして記憶した記憶媒体。

【請求項 19】

請求項 1 乃至 16 のいずれか一項に記載の画像処理方法を実行可能に構成された、画像分類機能付き撮像機器。

【請求項 20】

画像内の認識対象を分類するための分類器と、認識対象画像における前記認識対象を分

50

類する画像認識手段と、を備えた画像処理システムであって、

前記画像認識手段は、

認識対象画像を入力する入力部と、

前記認識対象画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成する小領域パッチ画像生成部と、

前記分類器を用いて認識結果を得る認識結果取得部と、

前記認識対象画像における認識対象の有無を判断する判断部と、

を有することを特徴とする画像処理システム。

【請求項 2 1】

請求項 2 0 記載の画像処理システムにおいて、さらに、

画像特徴量を抽出する画像特徴量抽出部と、

前記画像特徴量抽出部から抽出された画像特徴量からヒストグラムを作成するヒストグラム作成部と、

前記作成されたヒストグラムに最も近い特徴ベクトルをコードブックより選択する特徴ベクトル選択部と、

前記特徴ベクトル選択部から選択された特徴ベクトルを正規化する正規化部と、

前記正規化部により正規化された特徴ベクトルに基づき、前記分類器を用いて認識結果を得る認識結果取得部と、

を有することを特徴とする画像処理システム。

【請求項 2 2】

前記画像特徴量抽出部は、前記小領域パッチ画像生成部により生成された一つの小領域パッチ画像の画像特徴量を抽出することを特徴とする請求項 2 0 又は 2 1 記載の画像処理システム。

【請求項 2 3】

前記画像特徴量抽出部は、前記小領域パッチ画像生成部により生成された縮尺の異なる複数の小領域パッチ画像の画像特徴量をそれぞれ抽出し、

前記ヒストグラム作成部は、前記画像特徴量抽出部により抽出された複数の画像特徴量から生成されたそれぞれのヒストグラムを接続して、一つのヒストグラムを生成することを特徴とする請求項 2 0 又は 2 1 記載の画像処理システム。

【請求項 2 4】

前記画像特徴量抽出部は、前記認識対象画像の画像特徴量を抽出すると共に、前記小領域パッチ画像の画像特徴量を抽出し、

前記ヒストグラム作成部は、前記認識対象画像の画像特徴量から生成されたヒストグラムと、前記小領域パッチ画像の画像特徴量から生成されたヒストグラムと、を接続して一つのヒストグラムを生成することを特徴とする請求項 2 0 又は 2 1 記載の画像処理システム。

【請求項 2 5】

前記分類器は、

学習画像および該学習画像の分類を入力する手段と、

前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成する手段と、

前記小領域パッチ画像を用いて分類器を作成する手段と、

によって生成されることを特徴とする請求項 2 0 乃至 2 4 のいずれか一項に記載の画像処理システム。

【請求項 2 6】

前記分類器は、

学習画像および該学習画像の分類を入力する手段と、

前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成する手段と、

前記小領域パッチ画像の画像特徴量を抽出する手段と、

10

20

30

40

50

前記抽出された画像特徴量からコードブックを作成する手段と、
前記コードブックを用いて、前記抽出された画像特徴量からヒストグラムを作成する手段と、

前記ヒストグラムを用いて分類器を作成する手段と、
によって生成されることを特徴とする請求項 20 乃至 24 のいずれか一項に記載の画像処理システム。

【請求項 27】

前記分類器は、
学習画像および該学習画像の分類を入力する手段と、
前記学習画像の撮影位置に対応する、縮尺の異なる複数の航空写真画像および / または地図画像から、縮尺の異なる複数の小領域パッチ画像を生成する手段と、
前記複数の小領域パッチ画像の画像特徴量をそれぞれ抽出する手段と、
前記抽出された複数の画像特徴量からコードブックを作成する手段と、
前記コードブックを用いて、前記抽出された複数の画像特徴量からそれぞれのヒストグラムを作成し、これらヒストグラムを接続して一つのヒストグラムを生成する手段と、
前記一つのヒストグラムを用いて分類器を作成する手段と、
によって生成されることを特徴とする請求項 20 乃至 24 のいずれか一項に記載の画像処理システム。

【請求項 28】

前記分類器は、
学習画像および該学習画像の分類を入力する手段と、
前記学習画像の撮影位置に対応する航空写真画像および / または地図画像から一つの小領域パッチ画像を生成する手段と、
前記学習画像および前記小領域パッチ画像の画像特徴量をそれぞれ抽出する手段と、
前記抽出された複数の画像特徴量からコードブックを作成する手段と、
前記コードブックを用いて、前記抽出された複数の画像特徴量からそれぞれのヒストグラムを作成し、これらヒストグラムを接続して一つのヒストグラムを生成する手段と、
前記一つのヒストグラムを用いて分類器を作成する手段と、
によって生成されることを特徴とする請求項 20 乃至 24 のいずれか一項に記載の画像処理システム。

【請求項 29】

前記航空写真画像および / または地図画像が、ネットワークを介してアクセス可能なデータベースに格納されていることを特徴とする請求項 20 乃至 28 のいずれか一項に記載の画像処理システム。

【請求項 30】

前記認識対象画像が位置情報を保持していることを特徴とする請求項 20 乃至 29 のいずれか一項に記載の画像処理システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像処理方法、およびその方法を実行するプログラム、記憶媒体、撮像機器、画像処理システムに関する。詳しくは、画像内の認識対象（被写体）を分類するための画像処理技術に関する。

【背景技術】

【0002】

近年、デジタルカメラ、カメラ付き携帯電話などの普及や、ハードディスク、その他記憶媒体の大容量化によって、一般の個人が大量にデジタル画像を保有、蓄積することが出来るようになった。

しかしながら、撮影されたデジタル画像の蓄積先、例えば、パーソナル・コンピュータ（PC）、デジタルカメラ、カメラ付き携帯電話などのデジタル機器は、撮影、蓄積され

10

20

30

40

50

た画像中の被写体（認識対象）を判別する機能をもっていない。

よって、画像の取り扱いに関するデジタル機器と人とのセマンティックギャップは狭まることはなく、現状では、大量の画像データの分類や検索には、人手の介入が不可欠である。

人手により、撮影画像の内容に関するメタデータを記述することも可能であるが、手間がかかるため、撮影した画像すべてに意味、内容等に関するメタデータを記載することは、現実的ではない。

【0003】

画像の意味、内容に基づく処理を人手の介入なしに実現するためには、被写体、例えば、「ライオン」「自動車」「花」「山」「夕焼け」などの一般的な対象の認識（一般画像認識）を行なう必要がある。

現実世界で撮影された画像に対して、コンピュータなどのデジタル機器が、その画像中に含まれる被写体を一般的な名称で認識することを「一般物体認識（generic object recognition）」と呼び、画像認識の研究において最も重要な課題の一つである（例えば、非特許文献1、2など参照）。

【0004】

一般に、現実世界で撮影された画像に対する物体認識には、大きく分けて、identification（同定）と、classification（分類）の、2種類の認識がある。

identificationは個々の物体（the object）を区別する認識であり、入力された画像とデータベース中のモデルの照合を行い、どのモデルに対応する物体が画像中に存在するかどうかを出力結果とする。

一方、classificationは物体の種類（an object）を区別する認識で、人間が決めた分類（class）と画像中の物体（被写体）とを対応付け、物体のクラス名（多くの場合は一般名称）を出力結果とする。

「物体認識」はidentificationを指すのが一般的であるが、「一般物体認識」はclassificationの認識を意味するものであり、本明細書においても、これらの用語の定義に基づいて説明する。

【0005】

現在、デジタル画像に対する「一般物体認識」の研究が急速に進歩しつつある。ここでいう「一般物体認識」での対象画像とは、例えば、デジタルカメラやカメラ付携帯電話などで撮影したデジタル写真などの画像であり、認識対象はそうした画像中の「ライオン」「自動車」「花」「山」「夕焼け」などの各種の物体やシーンなどの被写体である。

【0006】

「一般画像認識」では、画像のみの持つ情報（画像データ）から認識するのが最も基本的な方法であるが、近年においては、デジタル画像の撮影時に、デジタルカメラもしくはカメラ付き携帯電話によって自動的に埋め込まれた付加情報（メタデータ）を、認識に利用する研究が提案されている。

例えば、撮影された時間を用いれば、画像データだけでは難しい、「夕日」であるか「朝日」であるかの区別は、容易に行うことが可能となる。

【0007】

また、非特許文献3に示されたように、メタデータとして、撮影時間のほか、撮影時におけるフラッシュの利用の有無や、レンズの焦点距離等のデータを、画像認識に利用する提案がなされている。しかし、従来の各種文献等においては、位置情報の利用について開示されていない。

メタデータ中の重要な情報として、位置情報がある。位置情報は通常GPS（Global Positioning System）によって取得するが、最近のデジタルカメラや携帯電話にはGPSが内蔵されたものもあり、撮影した画像に位置情報を、メタデータとして埋め込むことが可能な撮像機器が多く登場している。

【0008】

また、独立したGPSデバイスをデジタルカメラと一緒に持ち歩くことによって、撮影

10

20

30

40

50

時の位置情報を記録し、PC（パーソナル・コンピュータ）により、デジタル画像のファイルに、付加情報として位置情報を埋め込むことも可能である。

また、画像ファイル中の位置情報を、画像認識に利用することも一部で試みられている（例えば、特許文献1参照）。

【0009】

しかし、位置情報は緯度経度の2つの数値のみからなる情報であり、それ自体では一般物体認識の手がかりに利用することは困難であり、はかばかしい成果を挙げていない。この困難さは、位置情報を、どのように認識の手がかりとして利用するかが簡単でないとの問題に起因するものである。

【0010】

【非特許文献1】柳井啓司：「一般物体認識の現状と今後」、情報処理学会コンピュータビジョン・イメージメディア研究会招待講演予稿、CVM2006、CVM155-17(2006年)

【非特許文献2】柳井啓司：「一般物体認識の現状と今後」、情報処理学会論文誌：コンピュータビジョン・イメージメディア、Vol.48, No.SIG16(CVIM19), pp.1-24, 2007.

【非特許文献3】M. Boutell and J. Luo: Bayesian Fusion of Camera Metadata Cues in Semantic Scene Classification, Proceeding of Computer Vision and Pattern Recognition, pp. 623-630, 2004.

【特許文献1】特開2007-41762号（第18頁、第4図）

【発明の開示】

【発明が解決しようとする課題】

【0011】

このような事情に鑑み、本発明は、認識対象画像の分類に際し、認識対象画像と併せて、撮影位置付近の航空写真画像および地図画像を画像認識の手がかりの一部として利用することで、画像の認識精度を向上させることを目的とする。

【0012】

また、本発明は、デジタルカメラやカメラ付き携帯電話で撮影した位置情報付きのデジタル写真の自動分類、検索を可能にすることを目的とする。

【0013】

さらに、本発明は、デジタルカメラやカメラ付き携帯電話で撮影した画像の自動タグ付け、自動説明文生成、自動アルバム作成といった、様々な応用を可能にすることを目的とする。

【課題を解決するための手段】

【0014】

本発明者等は、画像が撮影された場所の位置情報が、時間情報などと比較して、一般物体認識のための大きな手がかりとなることに着目した。例えば、「海」の画像は海の近くでなければ撮影できないし、「ライオン」の画像は、アフリカでの撮影である等の特別な状況を除いて、動物園以外では撮影することがほとんどできない。

【0015】

このような着目に基づき本発明者等は鋭意検討を続け、認識対象の画像特徴量と併せて、撮影場所の位置情報を示す航空写真または地図の画像情報の画像特徴量を用いることで、一般画像認識の精度を向上させることができるとの知見を得て、本発明を完成するに至った。

【0016】

すなわち、本発明は、画像内の認識対象を分類するための分類器を用いて、認識対象画像における前記認識対象を分類する認識処理を含む画像処理方法であって、基本的には、

認識対象画像を入力するステップ（S105）と、

前記認識対象画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成するステップ（S110）と、

前記分類器を用いて認識結果を得るステップ（S135）と、

前記認識対象画像における認識対象の有無を判断するステップ（S140、S145、

10

20

30

40

50

S 1 5 0) と、
を含むことを特徴とする。

【 0 0 1 7 】

また、さらに、
前記小領域パッチ画像に基づき画像特徴量を抽出するステップ (S 1 1 5) と、
前記抽出された画像特徴量からヒストグラムを作成するステップ (S 1 2 0) と、
前記作成されたヒストグラムに最も近い特徴ベクトルをコードブックより選択するステップ (S 1 2 5) と、
前記選択された特徴ベクトルを正規化するステップ (S 1 3 0) と、
を備えたことを特徴としてもよい。

10

【 0 0 1 8 】

前記小領域パッチ画像を生成するステップ (S 1 1 0) が、前記航空写真画像および /
または地図画像から生成された、一つの画像から小領域パッチ画像を生成するステップ
であるとしてもよい。

【 0 0 1 9 】

前記小領域パッチ画像を生成するステップ (S 1 1 0) が、前記航空写真画像および /
または地図画像から生成された、縮尺の異なる複数の画像から小領域パッチ画像を生成す
るステップであるとしてもよい。

【 0 0 2 0 】

前記小領域パッチ画像を生成するステップ (S 1 1 0) が、前記認識対象画像、および
、前記航空写真画像および / または地図画像から生成された、縮尺の異なる複数の画像か
ら前記小領域パッチ画像を生成するステップであるとしてもよい。

20

【 0 0 2 1 】

前記ヒストグラムを作成するステップ (S 1 2 0) が、前記抽出された複数の画像特徴
量から生成されたそれぞれのヒストグラムを接続して、一つのヒストグラムを生成するス
テップであるとしてもよい。

【 0 0 2 2 】

前記ヒストグラムを作成するステップ (S 1 2 0) が、前記認識対象画像の画像特徴量
から生成されたヒストグラムと、前記小領域パッチ画像の画像特徴量から生成されたヒス
トグラムと、を接続して、一つのヒストグラムを生成するステップであるとしてもよい。

30

【 0 0 2 3 】

また、前記分類器が、学習画像および該学習画像の分類を入力するステップ (S 9 1)
と、

前記学習画像の撮影位置に対応する航空写真画像および / または地図画像から小領域パ
ッチ画像を生成するステップ (S 9 2) と、

前記分類器を作成するステップ (S 9 6) と、
によって生成されてもよい。

【 0 0 2 4 】

また、さらに、
前記学習画像の撮影位置に対応する航空写真画像および / または地図画像から小領域パ
ッチ画像を生成するステップ (S 9 2) と、
前記小領域パッチ画像の画像特徴量を抽出するステップ (S 9 3) と、
前記抽出された画像特徴量からコードブックを作成するステップ (S 9 4) と、
前記コードブックを用いて、前記抽出された画像特徴量からヒストグラムを作成するス
テップ (S 9 5) と、を備えてもよい。

40

【 0 0 2 5 】

前記小領域パッチ画像を生成するステップ (S 9 2) が、前記学習画像の撮影位置に対
応する、縮尺の異なる複数の航空写真画像および / または地図画像から、縮尺の異なる複
数の小領域パッチ画像を生成するステップ (S 9 2) であるとしてもよい。

【 0 0 2 6 】

50

前記小領域パッチ画像を生成するステップ(S92)が、前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から一つの小領域パッチ画像を生成するステップ(S92)であるとしてもよい。

【0027】

前記小領域パッチ画像を生成するステップ(S92)が、前記認識対象画像、および、前記航空写真画像および/または地図画像から生成された、縮尺の異なる複数の画像から前記小領域パッチ画像を生成するステップであるとしてもよい。

【0028】

前記ヒストグラムを作成するステップ(S95)が、前記抽出された複数の画像特徴量から生成されたそれぞれのヒストグラムを接続して、一つのヒストグラムを生成するステップであるとしてもよい。

10

【0029】

前記ヒストグラムを作成するステップ(S95)が、前記認識対象画像の画像特徴量から生成されたヒストグラムと、前記小領域パッチ画像の画像特徴量から生成されたヒストグラムと、を接続して、一つのヒストグラムを生成するステップであるとしてもよい。

【0030】

前記航空写真画像および/または地図画像が、ネットワークを介してアクセス可能なデータベースに格納されていることとしてもよい。

【0031】

前記認識対象画像が位置情報を保有しており、該位置情報に基づき、前記航空写真画像および/または地図画像を、前記認識対象画像と対応させることとしてもよい。

20

【0032】

前記した画像処理方法を、コンピュータ、または画像分類機能付き撮像機器、または画像処理システム、に実行させるための画像処理用プログラムとして実現することもできる。

【0033】

前記した画像処理用プログラムを、コンピュータにより読み出され実行可能なプログラムとして記憶した記憶媒体として提供することもできる。

【0034】

前記の画像処理方法を実行可能に構成された、画像分類機能付き撮像機器として提供することもできる。

30

【0035】

また、本発明は、画像内の認識対象を分類するための分類器と、認識対象画像における前記認識対象を分類する画像認識手段と、を備えた画像処理システムであって、基本的には、

前記画像認識手段は、

認識対象画像を入力する入力部と、

前記認識対象画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成する小領域パッチ画像生成部と、

前記分類器を用いて認識結果を得る認識結果取得部と、

40

前記認識対象画像における認識対象の有無を判断する判断部と、
を有することを特徴とする。

【0036】

また、さらに、

画像特徴量を抽出する画像特徴量抽出部と、

前記画像特徴量抽出部から抽出された画像特徴量からヒストグラムを作成するヒストグラム作成部と、

前記作成されたヒストグラムに最も近い特徴ベクトルをコードブックより選択する特徴ベクトル選択部と、

前記特徴ベクトル選択部から選択された特徴ベクトルを正規化する正規化部と、

50

前記正規化部により正規化された特徴ベクトルに基づき、前記分類器を用いて認識結果を得る認識結果取得部と、を有することとしてもよい。

【0037】

前記画像特徴量抽出部は、前記小領域パッチ画像生成部により生成された一つの小領域パッチ画像の画像特徴量を抽出することとしてもよい。

【0038】

前記画像特徴量抽出部は、前記小領域パッチ画像生成部により生成された縮尺の異なる複数の小領域パッチ画像の画像特徴量をそれぞれ抽出し、

前記ヒストグラム作成部は、前記画像特徴量抽出部により抽出された複数の画像特徴量から生成されたそれぞれのヒストグラムを接続して、一つのヒストグラムを生成すること

10

【0039】

前記画像特徴量抽出部は、前記認識対象画像の画像特徴量を抽出すると共に、前記小領域パッチ画像の画像特徴量を抽出し、

前記ヒストグラム作成部は、前記認識対象画像の画像特徴量から生成されたヒストグラムと、前記小領域パッチ画像の画像特徴量から生成されたヒストグラムと、を接続して一つのヒストグラムを生成することとしてもよい。

【0040】

前記分類器は、

学習画像および該学習画像の分類を入力する手段と、

20

前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成する手段と、

前記小領域パッチ画像を用いて分類器を作成する手段と、
によって生成されることとしてもよい。

【0041】

前記分類器は、

学習画像および該学習画像の分類を入力する手段と、

前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から小領域パッチ画像を生成する手段と、

30

前記小領域パッチ画像の画像特徴量を抽出する手段と、

前記抽出された画像特徴量からコードブックを作成する手段と、

前記コードブックを用いて、前記抽出された画像特徴量からヒストグラムを作成する手段と、

前記ヒストグラムを用いて分類器を作成する手段と、
によって生成されることとしてもよい。

【0042】

前記分類器は、

学習画像および該学習画像の分類を入力する手段と、

前記学習画像の撮影位置に対応する、縮尺の異なる複数の航空写真画像および/または地図画像から、縮尺の異なる複数の小領域パッチ画像を生成する手段と、

40

前記複数の小領域パッチ画像の画像特徴量をそれぞれ抽出する手段と、

前記抽出された複数の画像特徴量からコードブックを作成する手段と、

前記コードブックを用いて、前記抽出された複数の画像特徴量からそれぞれのヒストグラムを作成し、これらヒストグラムを接続して一つのヒストグラムを生成する手段と、

前記一つのヒストグラムを用いて分類器を作成する手段と、
によって生成されることとしてもよい。

【0043】

前記分類器は、

学習画像および該学習画像の分類を入力する手段と、

前記学習画像の撮影位置に対応する航空写真画像および/または地図画像から一つの小

50

領域パッチ画像を生成する手段と、

前記学習画像および前記小領域パッチ画像の画像特徴量をそれぞれ抽出する手段と、

前記抽出された複数の画像特徴量からコードブックを作成する手段と、

前記コードブックを用いて、前記抽出された複数の画像特徴量からそれぞれのヒストグラムを作成し、これらヒストグラムを接続して一つのヒストグラムを生成する手段と、

前記一つのヒストグラムを用いて分類器を作成する手段と、

によって生成されることとしてもよい。

【0044】

前記航空写真画像および/または地図画像が、ネットワークを介してアクセス可能なデータベースに格納されていることが望ましい。また、前記認識対象画像が位置情報を保持しているように構成することもできる。

【発明の効果】

【0045】

本発明によれば、認識対象画像の分類に際して、認識対象画像と併せて、撮影位置付近の航空写真画像および地図画像を画像認識の手がかりの一部として利用することにより、一般画像認識の精度を向上させることが可能となる。

【0046】

また、本発明によれば、前記利点の応用により、デジタルカメラやカメラ付き携帯電話等の撮像機器で撮影した位置情報付きのデジタル写真の自動分類、検索することが可能となる。

【0047】

また、本発明によれば、デジタルカメラやカメラ付き携帯電話の撮像機器で撮影した画像の自動タグ付け、自動説明文生成、自動アルバム作成といった、様々な応用が可能になるなど、多くの効果を有する。

【発明を実施するための最良の形態】

【0048】

以下、実施形態例について説明する。

【0049】

図1に、本例の画像処理システムの概要を示す。この画像処理システムは、システム外部の位置情報付き画像記憶部11より画像を収集して画像本体記憶部15に蓄積すると共に、画像に付加されている位置情報をメタデータ記憶部16へ格納する。また、画像の撮影位置に対応してマッピングサービス記憶部12より、位置情報に対応する位置の航空写真画像または地図画像を取り出し、異なる縮尺の小領域パッチ画像を生成し、各縮尺の記憶部13a、13b、13cへ格納するようになっている。

【0050】

システム外部の位置情報付き画像記憶部11として、ネットワークを介してアクセス可能なデータベースを用いてもよい。例えば、インターネット上に公開されたソーシャルサイトである「Flicker(登録商標)」などをあげることができる。

「Flicker」は撮影画像を投稿(アップロード)、共有(ダウンロード)することが出来るソーシャルサイトであり、毎日100万以上の画像が投稿されるといわれている。本明細書において、以下、Flickerへ投稿(アップロード)された画像を「Flicker画像」と言う。Flickerでは、撮影画像を投稿する場合、その位置情報を付加させることを強く推奨している。従って、今後「Flicker」に投稿される位置情報付きの画像データ、すなわち、「Flicker画像」は増え続けると予想される。

【0051】

マッピングサービス記憶部12としても、ネットワークを介してアクセス可能なデータベースを用いてもよい。例えば、インターネット上に公開された国土交通省などの公的な検索サービス、あるいは、民間の検索サービス(例えば、「Google(登録商標)」、「Yahoo(登録商標)」等)が提供する航空写真画像や地図画像を用いるとよい。

航空写真画像や地図画像は、位置情報(緯度経度)と対応しているので、画像の特徴を持

10

20

30

40

50

った位置情報とみなせる。したがって、位置情報を記述、識別する客観的な手段に成りうる。

【0052】

以下、画像、航空写真画像、地図画像を、「Flicker」と「Google」から収集した場合について説明する。

【0053】

なお、上記データベースでは現在、航空写真画像と地図画像の独立した画像を提供している。これらの航空写真画像は、256×256ピクセルの「タイル」と呼ばれる地図情報の断片から形成されているものも存在する。

これらの航空写真画像と地図画像についてはその画像を拡大、縮小して閲覧することが出来る。現在公開されている範囲では、0～19の、20通りのズームレベルをもつものも存在する。

【0054】

対象となる画像を認識するにあたり、画像から特徴を抽出する必要がある。以下、本例における画像の特徴を抽出する方法について説明する。

画像の特徴を記述する手法としては、画素値の統計や固有値を記述するものから、局所的な特徴を記述するものまで多種にわたる。

本例では、特徴抽出のために局所特徴の一種であるSIFT特徴を用いる。また、この局所特徴を簡潔に記述するために後述するBag of Keypoints手法を用いてデータをベクトル量子化する(図3参照)。

なお、別な手法として、似ている画像同士では画像を構成する色が似ているという仮定に基づき、画像中に色がどのような割合で含まれているのかを比較する色ヒストグラム法(color histogram method)を用いて特徴抽出を行うことも可能である。量子化された色空間に形成された色ヒストグラムを使用すると、色情報が各画素に割り当てられた元の画像よりもマッチングに使用される情報量が減少して、演算量も減ることが期待できる。

【0055】

SIFT(Scale Invariant Feature Transform)とは、1999年にDavid Loweによって提案された特徴点とそれに付随する特徴ベクトルの抽出法であり、特徴点周りの局所画像パターンを128次元特徴ベクトルで表現する。

SIFT特徴は、画像の拡大縮小、回転や視点の変化のいずれに対してもロバスト(強靱)であるとの性質がある。SIFT特徴の抽出は、特徴点の抽出とその特徴点における特徴ベクトルの抽出の2つのステップに分けることができる。

【0056】

具体的には、図4、図5に示すように、キーワードにて分類したFlicker画像に対応する位置情報を含んでいる航空写真画像1枚と、その周囲の航空写真画像8枚(合計9枚)を、ズームレベル10、12、14のそれぞれについてマッピングサービス記憶部12より取り出す。

それぞれのズームレベルについて、まず9枚の航空写真画像のタイル(256×256ピクセル)を3×3で結合する。このとき、位置情報を含む航空写真画像を中心に配置する。この位置情報が正方形の中心になるように、この結合された航空写真画像から512×512ピクセルの正方形部分を切り抜き、これをFlicker画像に対応する航空写真画像(小領域パッチ画像)とする。

【0057】

なお航空写真画像についても、Flicker画像と同様に、SIFT特徴を抽出する。

本例では、SIFT特徴を抽出するために、SIFT++というツールを用いた。このツールにおけるアルゴリズムは、SIFTを提唱したLoweアルゴリズムとほぼ同一である。

【0058】

SIFT特徴における特徴点の抽出について、次に述べるGRID点抽出で行う。GRID点抽出では、格子状に点を配置し、SIFT特徴ベクトル計算のための特徴点として利用する。

10

20

30

40

50

【0059】

G R I D点抽出によるS I F T特徴の抽出手順は以下ようになる。

- 1．格子点の間隔を決定する。本例においては、画像に対して、10画素ごとにG R I D点抽出を行い、それらの点に基づきS I F T特徴量を計算することとした。
- 2．画像から格子点を抽出し、予め決められた複数のスケールで、それぞれの点について勾配方向を計算する。格子点の総数は画像の画素数と格子点の間隔に依存する。
- 3．抽出した特徴点に対して、S I F T特徴量を計算する。

【0060】

なお、本例では、SIFT++を用いてG R I D点抽出を行うために、G R I D点を抽出する処理を予め実装しておき、これらの点を明示的に指定するオプションを用いることによって実装することができる。

10

【0061】

次に、抽出された特徴点における特徴ベクトルの抽出について、Bag of Keypointsの手法で行う。

Bag of Keypointsモデルとは、画像を局所特徴の集合と捉えた手法である。局所特徴をベクトル量子化し、Visual Wordsと呼ばれる特徴ベクトルを生成する。それらをまとめたものをコードブックと呼び、それを記述子として画像全体の特徴ベクトルを生成する。これにより、画像をVisual Wordsの集合 (bag) として表現することができる。

【0062】

Bag of keypointsの画像認識の流れは以下の通りである(図3参照)。

20

- 1．全画像データから特徴点を抽出する。
- 2．それをベクトル量子化し、コードブックを作成する。
- 3．コードブックをもとに、学習画像の特徴ベクトルを生成する。
- 4．同様にテスト画像の特徴ベクトルも生成し、分類器により画像がどのカテゴリに属するか決定する。

【0063】

コードブックの生成手順について、図3、図6を参照しながら説明する。

まず、Visual Wordsを生成するために、G R I D点におけるS I F T特徴を用いて全ての画像から局所特徴を抽出する。次に、抽出したもののうち、学習画像についての局所特徴をベクトル量子化し、各々のクラスタの中心を求めることによりVisual Wordsを求め、コードブックとする。

30

【0064】

ベクトル量子化は、最も単純なクラスタリング手法であるk-Means法を用いる。

これは、クラスタ数kと、各クラスタの初期の重心(これはランダムでもよい)を予め決めておき、重心と各ベクトルとの距離の平均が最小になるように反復して重心を更新していく手法である。コードブックの大きさは、クラスタ数kに依存する。

本例では、k=300に固定してベクトル量子化を行なった。k-Means法におけるクラスタリング処理では、ベクトル間の距離を計測する必要があるが、その距離尺度として、本例ではユークリッド距離を用いた。

【0065】

40

学習データ作成のためのコードブックの作成手順は以下の通りである。一つの画像について一つのコードブックが作成される。

- 1．各キーワードのグループについて、正例画像(O K画像)と負例画像(N G画像)を明確にする。
- 2．各キーワードのグループの全ての画像から抽出された特徴量を用いて、コードブックを作成する。
- 3．各キーワードから抽出したコードブックを用いて、そのキーワードの画像について、コードブック についてのヒストグラムを作成する。

【0066】

各画像に対応する航空写真画像のコードブックの作成手順は以下の通りである。

50

1. 各画像が含んでいる位置情報から、対応する航空写真画像（本例では三種類のズームレベル）を探し、それぞれ $3 \times 3 = 9$ 枚ずつ用意する。

2. 各ズームレベル航空写真画像について、位置（緯度経度）が中央になるように 256×256 ピクセルの正方形に切り抜く。

3. 航空写真画像について、対応する画像のキーワードに対してズームごとに独立でグループ化し、画像の場合と同様に特徴量、コードブックの順に求める。

【0067】

この時点で、一つのキーワードに対して、キーワードの画像、レベル10、12、14の各航空写真画像のグループに関する、4種類のコードブックを求めることになる。

コードブックの作成は、クラスタリング処理を伴う。一つのキーワードについて、精度の高いコードブックを作成するには、著しく膨大な量のデータをクラスタリングする必要がある。これは処理時間に対してトレードオフの問題となる。

本例においては、クラスタリングに用いる特徴点を、10分の1の確率で特徴点を抽出するよう絞込みを行い処理の高速化を図った。

【0068】

本例では、このクラスタリングに関し、学習と分類の手段である分類器として、SVM (Support Vector Machine) を用いる。

SVMは、ニューロンのモデルとして最も単純な線形しきい素子を用いて、2クラスのパターン識別器を構成する手法である。

カーネル学習法と組み合わせると非線形の識別器になる。この拡張はカーネルトリックと呼ばれる手法で、このカーネルトリックにより、現在知られている多くの手法の中でも最も認識性能の優れた学習モデルの一つであると考えられている。

なお、別の手法として、当該技術分野に公知の方法から選択することができ、最近傍法を用いても良い。最近傍法とは、補間処理の手法の一つであり、ある画素の周辺で一番近い画素の値を設定する手法である。より具体的には、例えば、

「<http://www.microsoft.com/japan/msdn/academic/Articles/Algorithm/04>」にその処理内容が記載されている。最近傍法は、処理速度が高速であるというメリットを有している。

【0069】

本例では、このSVMを実行するツールとして、SVM lightを用いる。

学習と分類のために用いるSVMへの入力ベクトルは、位置情報（緯度経度）のベクトルと、コードブックに関するヒストグラム (bag) によって構成される。

【0070】

まず各画像に対して、それぞれのグループのコードブックに関するヒストグラムを作成する。コードブックはSIFT特徴の代表ベクトルを指定されたクラスタ数だけ記述したデータであるから、各画像に対応するSIFT特徴のそれぞれについて、コードブックから「距離が最も近い」ベクトルを探し、そのベクトルに対して投票することによってヒストグラムを作成することができる。ベクトルの距離を計測する尺度として、本例ではユークリッド距離を用いた。

【0071】

全ての画像に関してヒストグラムの完成が完了した時点で、それぞれのキーワードにおいて、各ヒストグラム (bag) と位置情報のそれぞれの組合せで、各ベクトルを結合したものを作成する。ただし、航空写真画像については、3つの各レベルにそれぞれ独立して行うので、一つのキーワードに対して9通りのベクトルパターンを作成することになる（図2、図7、図8参照）。

【0072】

本例では、一つのキーワードに対する各画像に、9通りのベクトルパターンを作成した。各ベクトルパターンのグループのそれぞれについて、対応する画像の手動分類結果に基づき、正例画像と負例画像の2つのグループに分ける。それぞれのグループから、一定の枚数をランダムに抽出して、Cross Validationの手法によって学習と分類を行う。

10

20

30

40

50

【0073】

次に、本例における学習処理（分類器）の動作について、図9、図10を参照して説明する。

学習処理を行なうプログラムの動作が開始されると（S90）、ある対象（例えば、「山」「海」「ライオン」など。本明細書においてはこれらを総称して「分類」とも言う。）についてその対象が学習画像に含まれていることが予めわかっている正例画像、および、その対象が画像に含まれていないことが予めわかっている負例画像が予め蓄積されている画像を、写真本体記憶部15より読み出す（S91）。

分類器の学習精度を高めるため、正例画像、負例画像の枚数は、それぞれ100枚以上が好ましい。

その際、読み出した画像の縦横のいずれか一方または縦横の両方が480画素以上の場合、画像の縦横比を維持しながら、縦横のいずれもが480画素未満となるように画像を縮小することが好ましい。

【0074】

次に、S91で読み出した画像の撮影位置情報をメタデータ記憶部16より読み出し、当該位置情報に対応する航空写真画像を用いて、小領域パッチ画像を作成する（S92）。

その際、画像の位置座標が航空写真画像の中央となるように対応づけを行なう（図1の14a~14c）。

小領域パッチ画像としては、縮尺の異なる航空写真画像または地図画像を用いる。

分類器の精度を向上させるため、3つ以上の異なる縮尺を用いることが好ましい。

【0075】

次に、画像認識を実行する分類器を生成する。分類器を生成する方法としては、この種分野で通常用いられる手法を用いることができる。例えば、図9に示すように、前記S91で得られた小領域パッチ画像の白黒ビットマップデータ（例えば、256×256）をそのまま65536（=256×256）次元の特徴ベクトルとして、分類器（例えば、SVMなど）へ入力して分類器を生成する（S96）こともできる。

【0076】

また、他の例として、図10に示すように、次に、前記S91およびS92で得られた全ての画像について、画像特徴量を抽出する（S93）手法とすることも、画像認識精度をさらに向上させる効果を奏するために好適である。

画像特徴量として、本発明の効果を奏するためには特に限定はなく、SIFT特徴量、Haar特徴量のどちらを用いてもよいが、SIFT特徴量を用いる場合について以下に説明する。

【0077】

S93において、格子点（GRID点）を特徴点として設定する。

処理データ量と精度向上のトレードオフの観点より、各画像について縦横10画素間隔で格子点を設定することが好ましい。さらに、前記特徴点の近傍領域において、輝度勾配の方向ヒストグラム（「SIFT特徴ベクトル」とも言う）を算出する。

その際、近傍領域の範囲を4通り設定し、一つの特徴点から4つのSIFT特徴ベクトルを算出することが、精度向上の観点から好ましい。

以上の処理により、1枚の学習画像から約数千個のSIFT特徴ベクトルが求められる。

【0078】

次に、コードブックを作成する（S94）。この処理によって、典型的な例では、数百万個程度のすべてのSIFT特徴ベクトルから代表SIFTベクトルを300個程度求めて、コードブック21a~21dを作成することになる。

より具体的には、数百万個程度のすべてのSIFT特徴ベクトルから、ランダムサンプリングにより、1万個程度を選択する。

次に、選択された1万個程度のSIFT特徴ベクトルから、300個の代表ベクトルを

10

20

30

40

50

、クラスタ分析により求める。クラスタリング法としては特に限定はなく、k-meansクラスタリング法を用いても良い。

k-meansクラスタリング法(k平均法)とは、分散最適手法の一つで、分割の良さの評価関数を求め、その評価関数を最小化するように、k個のクラスタを分割する代表的な手法である。

【0079】

次に、ヒストグラム22a~22dを作成する(S95)。画像ごとに、抽出された数千個の各SIFT特徴ベクトルに最も近いコードブックのベクトルを求める。典型的な例では、コードブックに関する300次元ヒストグラムを作成することになる。

さらに、ヒストグラムの要素の合計が1となるように正規化する。この正規化されたものが、画像を表すbag of keypointsベクトルとなる。

【0080】

次に、分類器を生成する(S96)。上記の処理により得られた正例画像のbag of keypointsベクトル、負例画像のbag of keypointsベクトルを学習データとして分類器へ入力することで分類器を生成する。分類器としては、SVMを用いても良い。

なお、分類器の生成(S96)は、色ヒストグラム法により得られた正例画像のヒストグラム、負例画像のヒストグラムを学習データとして、分類器へ入力することで分類器を生成する。分類器の実現方法は、当該技術分野に公知の方法から選択することができ、前述した最近傍法を用いても良い。

学習処理を実行するプログラム(S90)は、CPUにより、上記した処理を、全画像について実行する(S97)。

【0081】

次に、本例の認識処理の動作について、図11、図12を参照して説明する。

認識処理を行なうプログラムの動作が開始されると(S100)、認識対象画像が入力される(S105)。

その際、読み出した画像の縦横のいずれか一方または縦横の両方が480画素以上の場合は画像の縦横比を維持しながら、縦横のいずれもが480画素未満となるように画像を縮小する。

【0082】

次に、S105で読み出した画像の撮影位置情報をメタデータ記憶部16より読み出し、当該位置情報に対応する航空写真画像または地図画像を用いて、小領域パッチ画像を作成する(S110)。その際、画像の位置座標が航空写真画像または地図画像の中央となるように対応づけを行なう。小領域パッチ画像としては、縮尺の異なる航空写真画像または地図画像を用いる。分類器の精度を向上させるため、3つ以上の異なる縮尺を用いることが望ましい。

【0083】

次に、画像を分類器により判定する手法については、この分野で通常用いられる手法を用いることができる。例えば、図12に示すように、前記S110で得られた小領域パッチ画像の白黒ビットマップデータ(例えば、256×256)をそのまま65536(=256×256)次元の特徴ベクトルとして、分類器(例えば、SVMなど)へ入力して分類器による判定を行なう(S135)手法を用いることもできる。

【0084】

また、他の例として、図11に示すように、次に、前記S105およびS110で得られた全ての画像について、画像特徴量を抽出する(S115)こととすることも、画像認識精度をさらに向上させる効果を奏するために好適である。画像特徴量として、本発明の効果を奏するためには特に限定はなく、SIFT特徴量を用いても良く、また、Haar特徴量を用いても良い。以下、説明を容易にするために、SIFT特徴量を用いる場合を例に説明を行なう。

【0085】

S115において、格子点(GRID点)を特徴点として設定する。処理データ量と精

10

20

30

40

50

度向上のトレードオフの観点より、各画像について縦横10画素間隔で格子点を設定するのが望ましい。さらに、前記特徴点の近傍領域において輝度勾配の方向ヒストグラム(「SIFT特徴ベクトル」とも言う)を算出する。その際、近傍領域の範囲を4通り設定し、一つの特徴点から4つのSIFT特徴ベクトルを算出することが、精度向上の観点より望ましい。

以上の処理により、1枚の学習画像から約数千個のSIFT特徴ベクトルが求められることになる。

【0086】

次に、コードブックを検索する(S125)。当該処理によって、典型的な例では、前記1枚の学習画像から抽出された約数千個のSIFT特徴ベクトルにコードブック中で「距離が最も近い」ベクトルを探し、そのベクトルに対して投票することによってヒストグラムを作成する。ベクトルの距離を計測する尺度として、本例ではユークリッド距離を用いた。このようにして、コードブックに関する300次元のヒストグラムを得ることになる。

10

【0087】

次に、前記ヒストグラムを正規化する(S130)。前記典型的な例では、各300次元のヒストグラムは要素の合計が1となるように正規化されることによって、認識対象画像を表すbag-of-keypointsベクトルが得られる。

次に、当該得られたbag-of-keypointsベクトルを分類器に入力し、認識対象画像に対する認識結果値を得る(S135)。分類器として、前記学習処理によって学習済みであるサポートベクターマシン(SVM)を用いるのが好適である。

20

【0088】

次に、前記認識結果値を判定し、正であれば(S140:下方向)認識対象画像は予め指定された分類の対象物体を含むと判断され(S145)、また、負であれば(S140:右方向)認識対象画像は予め指定された分類の対象物体を含まないものと判断される(S150)。

認識処理を実行するプログラムS100は、CPUにより以上の処理を全画像について実行する(S155)。

【0089】

次に、試験例について説明する。

30

【0090】

本例をコンピュータソフトウェアとして実現し、インターネット(Web)にアクセス可能なパーソナル・コンピュータ(PC)にて実施した。Fllickrから収集した日本国内の位置情報を含む画像約5000枚を用いて、本例による一般画像認識の精度を確認した。

各画像について本試験のために、5種類のキーワード(景色、ラーメン、山、神社、海岸)を与えた。

本試験では、使用する航空写真画像のズームレベルは10、12、14の3種類とする。

Fllickrで収集した各画像について、収集した航空写真画像を用いて、その位置情報を表す航空写真画像を対応づけた。

40

【0091】

試験データセットから、特徴量を抽出し、コードブックとヒストグラムを作成することによって、SVMへの入力データを作成する。

これにより、画像と航空写真画像、位置情報のデータセットから、一つのキーワードについて各グループのデータが作成される。

本試験では、それぞれのグループから、200枚をランダムに抽出して、SVMへの入力データセットを作成した。

また、本試験では、テストデータ全体をより客観的に評価するため、学習と分類の方法としてCross Validationを用いた(図13参照)。

【0092】

50

具体的な手順を以下に説明する。まず認識対象物が画像中に写っている正例画像（「OKデータ」とも言う）と、認識対象物が画像中に写っていない負例画像（「NGデータ」とも言う）を等分割する。各枠に含まれるテストセットの数はすべて同じとした。本試験では、試験データを5つに等分割するので、枠の中には20枚が入ることになる。

すなわち、前述のSVMへの入力データの作成の際、学習データと分類データがそれぞれ等しく成るようにランダム抽出を行う。

【0093】

学習データを分割した後、学習データと分類データをそれぞれ組み替えて実験を行い、1つのグループについて、5通りの結果を得た。

SVMによる出力結果から、情報検索の評価と同様に、再現率（Recall）と、適合率（Precision）を求めることができる。また、再現率と適合率の両方を考慮した指標として、F値と、再現率 - 適合率グラフを求めることができる。

特に、再現率 - 適合率グラフでは、両方のトレードオフの関係を検証可能である。

【0094】

本試験では、学習と分類に、Cross Validationを採用している（図13参照）。5つのfoldで行ったので、上記のような評価方法において、5通りの結果が出力されることになる。

【0095】

本試験では、平均適合率（Average Precision）により実験結果を評価した。

1つのキーワードにおいて、9通りのグループのデータセットをそれぞれ評価する。さらに、それぞれのグループを5つのfoldに分割しているので、5つのfoldのそれぞれの平均適合率を求め、これらとこれらの平均を示す。1つのキーワードにおいて、合計で45個の平均適合率を計算する。

【0096】

試験結果は、それぞれの平均適合率を100倍して、有効数字4桁で示す（表1参照）。

ただし、各表において、画像をI、位置情報をG、10（表1での(1)）、12（表1での(2)）、14（表1での(3)）を航空写真画像のそれぞれのレベルとして、グループの組み合わせを示す。

【0097】

【表1】

結果(5回の平均)

Fold	I	I+G	G	I+(1)	I+(2)	I+(3)	(1)	(2)	(3)	diff
ラーメン	65.49	72.89	73.71	79.10	73.36	67.94	67.78	74.31	56.67	+13.61
海岸	74.34	86.15	81.73	73.99	72.73	64.60	74.18	73.16	65.61	+11.81
神社	56.22	59.04	53.86	56.86	58.15	68.31	62.17	50.41	56.52	+12.09
山	67.29	71.64	69.77	67.64	65.47	67.57	72.38	73.79	70.59	+6.50
景色	78.76	78.22	77.80	79.06	81.92	76.87	65.71	71.62	70.88	+3.16

【0098】

キーワード「景色」については、画像と航空写真画像を統合したもののうち、レベル10（表1でのI+(1)）と、レベル12（表1でのI+(2)）で精度が向上していることが確認された。

これらの位置情報には、都会や都市部全体的な割合が多く、都市部の局所特徴量のみでは、風景を識別するのは困難である。したがって、風景写真の画像と航空写真画像の統合

した結果が最も精度が高くなると考えられる。

【0099】

キーワード「ラーメン」については、タイトルや説明などのメタデータから、ラーメン店で撮影された画像が多い。したがって、「景色」の場合と同様に、位置情報が比較的都市部に集中する。航空写真画像レベル10(表1での(1))との組合せで最も精度が高くなるのは、このズームレベルに対して都市部の特徴が現れやすいためと考えられる。

【0100】

なお、本試験では、航空写真画像のコードブックを作成する際に、画像と区別した。これは、画像の特徴量との混乱を避けるためである。しかし、1種類のズームレベルの航空写真画像をあえて混合して、各ズームごとに独立したコードブックを作成するという方法も可能である。

10

【0101】

以上のように、本試験では、インターネットのWeb上から収集した、位置情報付きの画像と、位置情報に対応する航空写真画像を用いて、本発明の画像処理方法により、画像認識の精度が向上することが確認された。

【0102】

以上、本発明の実施形態例を図面等に基づき説明したが、本発明は前記した例に限定されるものではなく、特許請求範囲記載の技術的思想の範疇において種々の変更が可能であることは言うまでもない。

【産業上の利用可能性】

20

【0103】

本発明の画像処理技術、特に、分類対象の対象物(被写体)がデジタル画像に含まれるか否かを判断する技術は、インターネットなどのネットワークとそれに接続したパーソナル・コンピュータ(PC)上、または、一般家庭内やオフィス内に構築したLANに接続したPC上、または、個人ユーザなどが使うスタンドアローンのPC上などにおいて、一般画像認識の精度を向上させることができる。

また、デジタルカメラやカメラ付き携帯電話等の撮像機器に組み込むことによって、撮影した位置情報付きデジタル画像の自動分類、検索に寄与する。

また、デジタルカメラ等で撮像された画像の自動タグ付け、自動説明文生成、自動アルバム作成など、様々な応用技術、応用製品を提供可能となる。

30

そして、独立した製品として、または、他の製品に組み込むソフトウェアとして、または、インターネット上で利用可能なシステムとして、など、好適に利用され得る。

【図面の簡単な説明】

【0104】

【図1】本発明に係る画像処理システムの概要を示す概念図。

【図2】分類器への入力データを示す概念図。

【図3】Bag of Keypoints手法の説明図。

【図4】航空写真画像(または地図画像)の収集方法例の説明図。

【図5】航空写真画像(または地図画像)の処理方法例の説明図。

【図6】コードブックの生成の一例の説明図。

40

【図7】対象画像の分類を示すキーワードごとのSVMへの入力ベクトル例の説明図。

【図8】ベクトルパターン例の説明図。

【図9】学習処理(分類器)の動作を示すフローチャートの一例。

【図10】学習処理(分類器)の動作を示すフローチャートの一例。

【図11】認識処理の動作を示すフローチャートの一例。

【図12】認識処理の動作を示すフローチャートの一例。

【図13】クロスバリデーション手法の概念図。

【符号の説明】

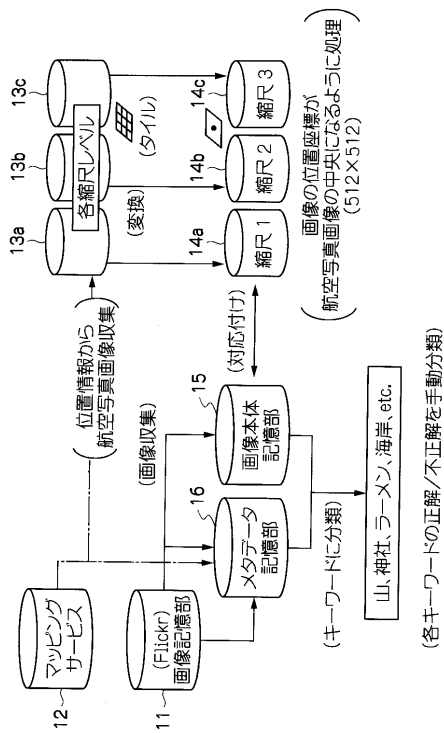
【0105】

11: 画像記憶部

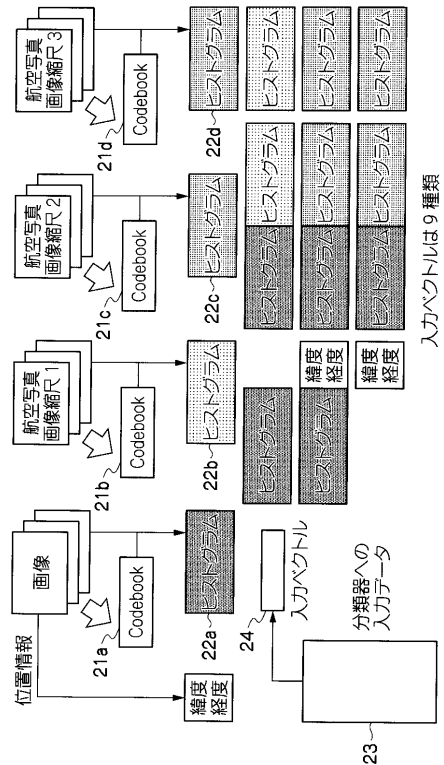
50

- 1 2 : マッピングサービス記憶部
- 1 3 a ~ 1 3 c : 位置情報に対応した各縮尺の航空写真画像または地図図形の記憶部
- 1 4 : 位置座標調整済みの各縮尺の航空写真画像または地図図形の記憶部
- 1 5 : 画像本体記憶部
- 1 6 : メタデータ記憶部
- 2 1 : コードブック記憶部
- 2 2 : ヒストグラム記憶部
- 2 3 : 分類器への入力データ記憶部
- 2 4 : 入力ベクトル記憶部

【 図 1 】

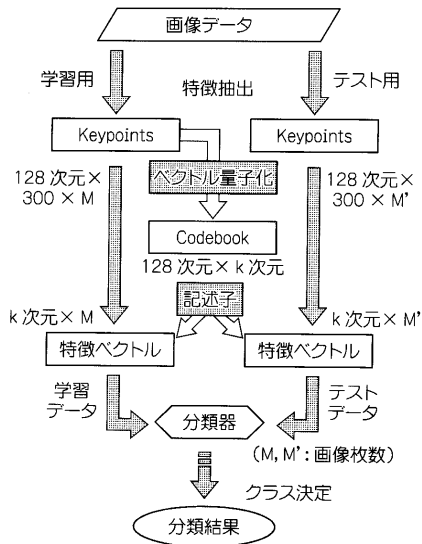


【 図 2 】



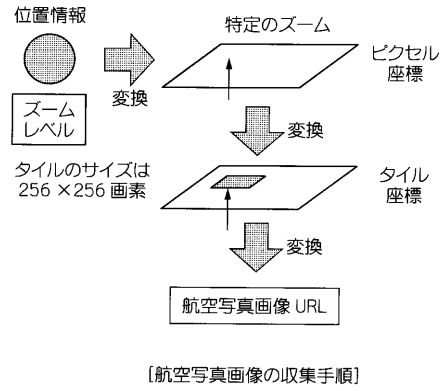
[SVM(分類器)への入力データ]

【 図 3 】

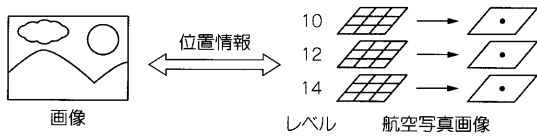


[Bag of Keypoints 手法]

【 図 4 】

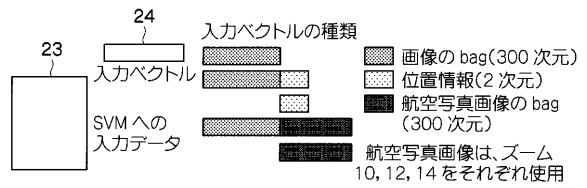


【 図 5 】



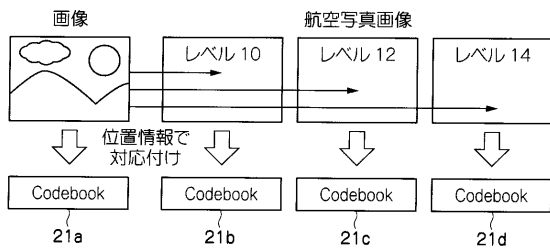
[航空写真画像の処理方法。1つの画像に対応する位置情報を表す航空写真画像を生成する]

【 図 7 】



[キーワードごとの SVM(分類器)への入力ベクトル]

【 図 6 】



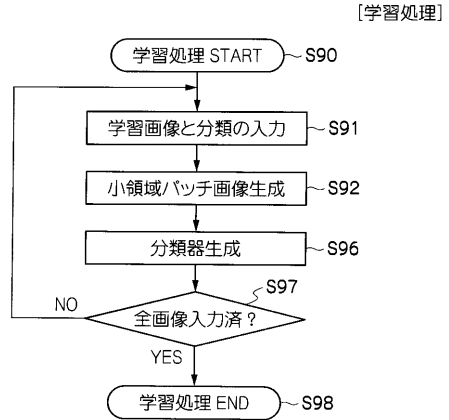
[Codebook の生成]

【 図 8 】

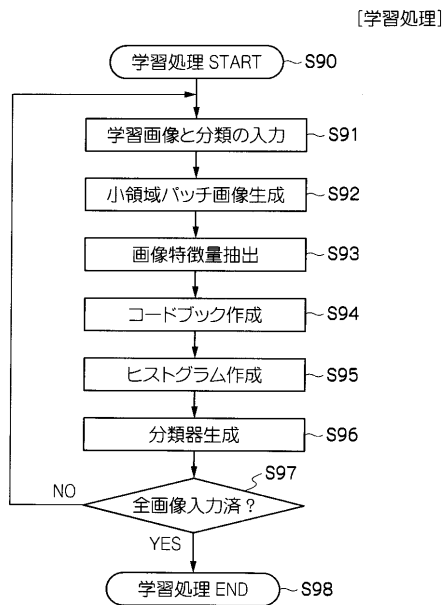
画像のみ	画像+位置情報	位置情報のみ
OK データ	OK データ	OK データ
NG データ	NG データ	NG データ
画像+航空写真画像 レベル 10	画像+航空写真画像 レベル 12	画像+航空写真画像 レベル 14
OK データ	OK データ	OK データ
NG データ	NG データ	NG データ
航空写真画像のみ レベル 10	航空写真画像のみ レベル 12	航空写真画像のみ レベル 14
OK データ	OK データ	OK データ
NG データ	NG データ	NG データ

[ベクトルパターンを2つのグループに分割]

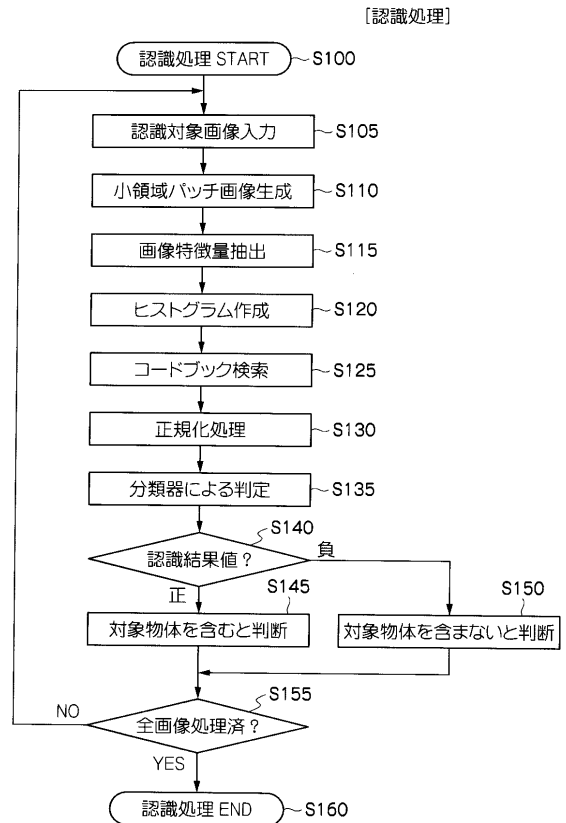
【 図 9 】



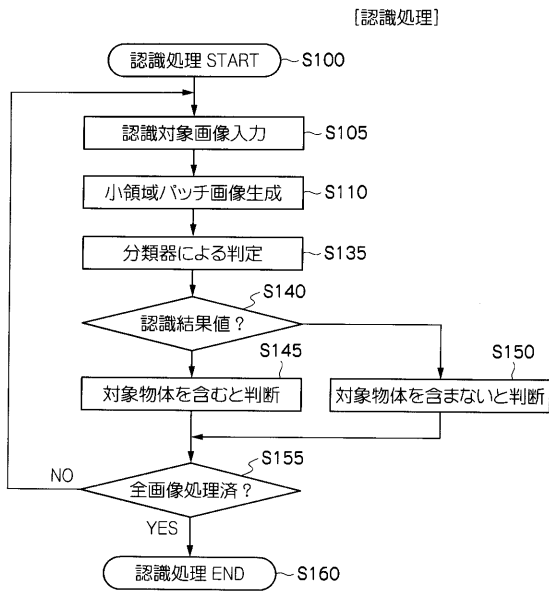
【 図 1 0 】



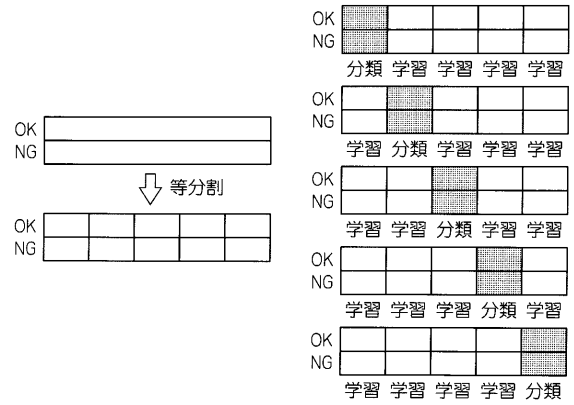
【 図 1 1 】



【 図 1 2 】



【 図 1 3 】



[Cross validation の手法]