

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号  
特許第5963609号  
(P5963609)

(45) 発行日 平成28年8月3日(2016.8.3)

(24) 登録日 平成28年7月8日(2016.7.8)

(51) Int.Cl.

F I

G O 6 F 17/30 (2006.01)

G O 6 T 1/00 (2006.01)

G O 6 F 17/30 3 5 0 C

G O 6 F 17/30 1 7 0 B

G O 6 T 1/00 2 0 0 E

請求項の数 11 (全 22 頁)

(21) 出願番号	特願2012-184572 (P2012-184572)	(73) 特許権者	000001007
(22) 出願日	平成24年8月23日 (2012.8.23)		キヤノン株式会社
(65) 公開番号	特開2014-41560 (P2014-41560A)		東京都大田区下丸子3丁目30番2号
(43) 公開日	平成26年3月6日 (2014.3.6)	(74) 代理人	100076428
審査請求日	平成27年7月3日 (2015.7.3)		弁理士 大塚 康德
		(74) 代理人	100112508
			弁理士 高柳 司郎
		(74) 代理人	100115071
			弁理士 大塚 康弘
		(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、画像処理方法

(57) 【特許請求の範囲】

【請求項 1】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性及び認識尤度を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置であって、

クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている画像を特定する特定手段と、

前記特定手段が特定した画像と関連づけて前記データベースに登録されている認識尤度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関する情報と、該類似度と、を画像検索結果として出力する出力手段と

を備えることを特徴とする画像処理装置。

【請求項 2】

前記出力手段は、前記特定した画像に関する情報と、前記類似度と、のセットを、類似度が高い順で出力することを特徴とする請求項 1 に記載の画像処理装置。

【請求項 3】

前記出力手段は、

前記クエリ画像に対する画像認識処理の結果、画像認識できたオブジェクトが 1 つも存在しなかった場合には、前記データベースに登録されている画像のうち、前記クエリ画像に対する画像認識処理により得られた前記クエリ画像の特徴量と類似する特徴量を有する

画像に関する情報を出力す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】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性と特徴量と認識尤度とを、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置であって、

前記データベースに属性と特徴量と認識尤度とが登録されていない新規オブジェクトを認識するために生成された検出器と、前記データベースに登録済みの特徴量と、を用いて、前記データベースに登録済みの画像から該新規オブジェクトを認識し、該認識した新規オブジェクトの属性と特徴量と認識尤度とを、該登録済みの画像と関連付けて前記データベースに登録する認識手段と、

10

クエリ画像に対して、前記新規オブジェクトの認識処理を含む画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている画像を特定する特定手段と、

前記特定手段が特定した画像と関連づけて前記データベースに登録されている認識尤度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関する情報と、該類似度と、を画像検索結果として出力する出力手段と

を備えることを特徴とする画像処理装置。

【請求項 5】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性及び領域を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置であって、

20

クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている領域を特定する特定手段と、

前記特定手段が特定した領域について画像認識処理で得られた特徴量と、前記クエリ画像において該領域に対応する領域について画像認識処理で得られた特徴量と、の類似度を求める手段と、

前記特定手段が特定した領域と関連付けて前記データベースに登録されている画像に関する情報と、前記類似度と、を画像検索結果として出力する出力手段と

を備えることを特徴とする画像処理装置。

30

【請求項 6】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性及び認識尤度を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置であって、

クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている画像を特定する特定手段と、

前記クエリ画像に対する画像認識処理で得られる特徴量と、前記特定手段が特定した画像に対する画像認識処理で得られる特徴量と、の類似度を求める手段と、

前記特定手段が特定した画像と関連づけて前記データベースに登録されている認識尤度と、前記類似度と、の重み付け加算により得られる総合類似度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関する情報と、該類似度と、を画像検索結果として出力する出力手段と

40

を備えることを特徴とする画像処理装置。

【請求項 7】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性及び認識尤度を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置が行う画像処理方法であって、

前記画像処理装置の特定手段が、クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベース

50

に登録されている画像を特定する特定工程と、

前記画像処理装置の出力手段が、前記特定工程で特定した画像と関連づけて前記データベースに登録されている認識尤度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関する情報と、該類似度と、を画像検索結果として出力する出力工程と

を備えることを特徴とする画像処理方法。

【請求項 8】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性と特徴量と認識尤度とを、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置が行う画像処理方法であって、

10

前記画像処理装置の認識手段が、前記データベースに属性と特徴量と認識尤度とが登録されていない新規オブジェクトを認識するために生成された検出器と、前記データベースに登録済みの特徴量と、を用いて、前記データベースに登録済みの画像から該新規オブジェクトを認識し、該認識した新規オブジェクトの属性と特徴量と認識尤度とを、該登録済みの画像と関連付けて前記データベースに登録する認識工程と、

前記画像処理装置の特定手段が、クエリ画像に対して、前記新規オブジェクトの認識処理を含む画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている画像を特定する特定工程と、

前記画像処理装置の出力手段が、前記特定工程で特定した画像と関連づけて前記データベースに登録されている認識尤度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関する情報と、該類似度と、を画像検索結果として出力する出力工程と

20

を備えることを特徴とする画像処理方法。

【請求項 9】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性及び領域を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置が行う画像処理方法であって、

前記画像処理装置の特定手段が、クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている領域を特定する特定工程と、

30

前記画像処理装置の類似度を求める手段が、前記特定工程で特定した領域について画像認識処理で得られた特徴量と、前記クエリ画像において該領域に対応する領域について画像認識処理で得られた特徴量と、の類似度を求める工程と、

前記画像処理装置の出力手段が、前記特定工程で特定した領域と関連付けて前記データベースに登録されている画像に関する情報と、前記類似度と、を画像検索結果として出力する出力工程と

を備えることを特徴とする画像処理方法。

【請求項 10】

それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の 1 以上のオブジェクトのそれぞれの属性及び認識尤度を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置が行う画像処理方法であって、

40

前記画像処理装置の特定手段が、クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている画像を特定する特定工程と、

前記画像処理装置の類似度を求める手段が、前記クエリ画像に対する画像認識処理で得られる特徴量と、前記特定工程で特定した画像に対する画像認識処理で得られる特徴量と、の類似度を求める工程と、

前記画像処理装置の出力手段が、前記特定工程で特定した画像と関連づけて前記データベースに登録されている認識尤度と、前記類似度と、の重み付け加算により得られる総合類似度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関す

50

る情報と、該類似度と、を画像検索結果として出力する出力工程とを備えることを特徴とする画像処理方法。

【請求項 11】

コンピュータを、請求項 1 乃至 6 の何れか 1 項に記載の画像処理装置の各手段として機能させるためのコンピュータプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、局所特徴を用いたオブジェクト単位での画像検索のための画像処理技術に関するものである。

10

【背景技術】

【0002】

画像の局所的な特徴量（局所特徴量）を用いて類似画像を検索する方法が提案されている。この方法では、まず、画像から特徴的な点（局所特徴点）を抽出する（非特許文献 1）。そして、該局所特徴点とその周辺の画素群とに基づいて、該局所特徴点に対応する特徴量（局所特徴量）を計算する（非特許文献 2）。画像の検索は、局所特徴量同士のマッチングによって行う。

【0003】

従来の検索結果の表示方法は、多くの場合、クエリ画像と検索先画像（サンプル画像）との間の類似度を局所特徴量のマッチングにより算出し、類似度順にサムネイル画像を表示する。

20

【0004】

他方、局所特徴量を用いて機械学習を行う画像認識技術も提案されている（非特許文献 3）。これは、複数の学習画像を用意して機械学習することにより、共通に存在するオブジェクトを検出する技術であり、デジタルスチルカメラ等における顔検出技術等で実用化されている。上記、局所特徴量を用いた画像検索に比べ、機械学習による効果により、誤って認識される割合が比較的少ないという特性を持つ。しかしながら、画像認識技術は前述の様に、予め機械学習を行う必要があり、万物に対応する事は非現実的である。

【0005】

画像認識技術と画像特徴量検索を併用して認識精度の向上を行う提案もなされている（特許文献 1）。これは、再生した画像データから顔画像を検出し、検出した顔画像から個人識別情報を抽出することを、通常の顔画像、サングラスをかけた顔画像、帽子を被った顔画像の夫々について行う。そして、それらが同一人物のものであるとしてそれらを相互に関連付けて記憶する。その後、撮像した画像から顔を検出し、顔画像の特徴点付近に動きがあった場合、動きのあった位置と、動きの方向とに基づいて個人識別情報を切り替える。そして、帽子やサングラスの部分を含めて個人の識別を行う。したがって、個人の識別精度や、顔の追従精度を高めることができる。

30

【先行技術文献】

【非特許文献】

【0006】

40

【非特許文献 1】C.Harris and M.J. Stephens, "A combined corner and edge detector," In Alvey Vision Conference, pages 147-152, 1988.

【非特許文献 2】David G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision, 60, 2 (2004), pp.91-110.

【非特許文献 3】柳井, "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とイメージメディア Nov.2007, Vol.48 No.SIG 16(CVIM 19)

【特許文献】

【0007】

【特許文献 1】特開2011-76439号公報

【発明の概要】

50

## 【発明が解決しようとする課題】

## 【0008】

特許文献1では、顔検出結果とその近傍の画像特徴を含めて同一のものであると判断しており、検索処理ではないものの、顔検出結果とその他の部分の画像特徴量を合わせてマッチングをしている。しかし、顔は器官検出ベースの特徴点特徴量、顔の近傍の帽子などは形状と色特徴でそれぞれ異なる特徴であり、顔検出に用いる特徴量と画像特徴量検索に用いる画像特徴量の共通化の概念は無く、画像認識と画像特徴量比較が独立した処理となっている。

## 【0009】

局所特徴量を用いた検索の対象となるオブジェクトを選ばない性質と、画像認識の対象画像オブジェクトが限られるが誤った検出が比較的少ないという性質を持ち合わせた、良好な検索結果をもたらす画像検索を実現できていない。画像から抽出した局所特徴量を画像認識に流用して効率化することも行っていない。

## 【0010】

画像認識技術は、ある認識ターゲットに対して多量の学習データを用いて機械学習を行う事で、画像特徴の検索に比べ精度の高い検索が期待できる。しかし、学習画像はアピランスの異なるものを与える事が通常であり、認識尤度は即、クエリ画像との類似性を示すものでは無いので、画像認識の結果を用いた検索だけでは不十分である。

## 【0011】

更に、画像認識の対象は時間が経つとともに増加するので、過去に登録した画像に対して新しい画像認識結果を適応する場合には、通常、再登録処理を行う必要があり、又、当然画像データが参照できる環境で無ければならない。反対に、局所特徴量を用いた検索では、機械的に見て特徴量が類似している検索結果を出してしまうので、その検索結果を人間が見ると全く違うオブジェクトであることがある。

## 【0012】

本発明は、上記の課題に鑑みてなされたものであり、局所特徴量を用いた検索の対象となるオブジェクトを選ばない性質と、画像認識の誤りが比較的少ない性質とを併せ持った画像検索を実現する技術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】

## 【0013】

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、それぞれの画像について、該画像に対する画像認識処理によって得られた該画像中の1以上のオブジェクトのそれぞれの属性及び認識尤度を、該画像と関連付けて登録しているデータベースにアクセス可能な画像処理装置であって、

クエリ画像に対して画像認識処理を行うことで該クエリ画像中のオブジェクトの属性を取得し、該取得した属性と関連付けて前記データベースに登録されている画像を特定する特定手段と、

前記特定手段が特定した画像と関連づけて前記データベースに登録されている認識尤度を、前記クエリ画像に対する該特定した画像の類似度とし、該特定した画像に関する情報と、該類似度と、を画像検索結果として出力する出力手段と

を備えることを特徴とする。

## 【発明の効果】

## 【0014】

本発明の構成によれば、局所特徴量を用いた検索の対象となるオブジェクトを選ばない性質と、画像認識の誤りが比較的少ない性質とを併せ持った画像検索を実現することができる。

## 【図面の簡単な説明】

## 【0015】

【図1】画像処理装置のハードウェア構成例を示すブロック図。

【図2】画像処理装置の機能構成例を示すブロック図。

10

20

30

40

50

【図3】画像特徴抽出・認識処理部202(202')の機能構成例を示すブロック図。

【図4】特徴点/局所特徴量算出部303が行う処理のフローチャート。

【図5】スケールスペースを説明する図。

【図6】画像登録処理のフローチャート。

【図7】DB204内のインデックススキーマを示す図。

【図8】登録画像及びクエリ画像を示す図。

【図9】DB204に登録される各インデックスを示す図。

【図10】画像検索処理のフローチャート。

【図11】画像検索処理のフローチャート。

【図12】画像検索処理のフローチャート。

【図13】DBインデックスを更新する処理のフローチャート。

【発明を実施するための形態】

【0016】

以下、添付図面を参照し、本発明の好適な実施形態について説明する。なお、以下説明する実施形態は、本発明を具体的に実施した場合の一例を示すもので、特許請求の範囲に記載の構成の具体的な実施例の1つである。

【0017】

<はじめに>

第1の実施形態の説明の前に、各実施形態で述べる概要を説明する。登録画像より特徴点を検出しその特徴点の局所特徴量を抽出し、更に画像認識処理し、画像中の認識対象のメタデータを抽出するとともに、これらを合わせて記憶する。予め、画像認識の準備として、認識対象を含んだ画像、含まない画像を学習用に準備し、機械学習を行っておく。

【0018】

そして、画像認識処理を行った結果の認識対象のメタデータには少なくとも認識対象の属性を含む。更にメタデータには、認識対象の領域情報と、認識尤度の少なくともいずれか1つを記憶しておく。尚、局所特徴には特徴点の位置情報を含む事とする。検索時は、クエリ画像を登録時と同様、画像認識処理し画像中の認識対象のメタデータを抽出する。

【0019】

画像認識処理で用いる局所特徴量を、画像検索処理で用いる局所特徴量と共通化することにより、局所特徴量の類似度の比較結果と画像認識の結果が大きく異ならないようにする。例えば、局所特徴量が輝度情報から得られるもので、画像認識が色情報を用いたものであると、それぞれでヒットする画像の見かけが大きく異なるものになってしまう。

【0020】

クエリ画像に対して画像認識処理を行い、認識対象のメタデータが得られた場合に、認識対象のメタデータ及び認識対象の領域の局所特徴量の少なくとも片方を用いてマッチングするが、認識対象が無かった場合に局所特徴のみを用いてマッチングする。これにより、画像認識の誤検出の少なさを生かしながらも、画像認識の対象が有限である影響を画像特徴量検索処理で緩和した画像検索処理が可能となる。

【0021】

また、クエリ画像に対して画像認識処理を行って認識対象のメタデータが得られた場合に、同じ認識対象の属性を持つ画像に絞り込み、認識対象の領域中の局所特徴比較を行い、認識対象が無かった場合に局所特徴を用いて画像特徴量比較処理する方法も有る。

【0022】

このメリットとしては、登録画像とクエリ画像の認識領域中の局所特徴量を比較することにより、認識尤度だけでは解決できなかったクエリとの類似性を反映しつつ、認識結果を絞り込みに使用した効率の良い検索処理を行う事が可能になる。

【0023】

更に次のような方法もある。即ち、クエリ画像に対して画像認識処理を行い、認識対象のメタデータが得られた場合、認識対象の属性を用いてマッチングし、認識尤度を得て更に局所特徴を用いて画像特徴量比較を行って類似度を得て認識尤度と類似度の重み付け和

10

20

30

40

50

で総合類似度を算出する。そして、認識対象が無かった場合に局所特徴を用いて画像特徴量検索処理する。これにより、処理コストは増えるものの、画像認識が失敗した時にも、検索漏れを減らす事が可能となる。

【 0 0 2 4 】

なお、クエリ画像で認識率の低い認識対象が検出された場合には、認識尤度への重みを低くし、認識率が高い認識対象の場合には高い値の重みとすると良好な結果が得られる。更に、類似度の重みが零の場合には局所特徴を用いた画像特徴比較処理を行わないと、無駄な処理を防ぐ事が出来る。

【 0 0 2 5 】

ところで、メタデータを用いてマッチングを行う際、認識メタデータの比較は同一属性であるかを判断し、尤度をその類似度へ換算する事が最も簡便な方法である。また、局所特徴を用いてマッチングを行う際、画像認識領域に含まれる局所特徴量のマッチングで求めた類似度を用いることで、真にクエリ画像の対応する領域との類似性で比較を行う事が可能となる。

10

【 0 0 2 6 】

もちろん、尤度と画像全体の特徴点に対する局所特徴量のマッチングで求めた類似度を統合した類似度を算出しても良い。このメリットは、画像認識で得た尤度を緩やかに検索結果に反映する事が可能となり、画像認識の失敗の影響を緩和し、よりロバストな検索が可能である事である。

【 0 0 2 7 】

20

ところで、画像認識処理と画像特徴検索処理において共通の局所特徴を用いる事により、画像認識の対象が増えた場合でも既に抽出した局所特徴を用いる事により、登録済みの画像を再解析を行う事無く画像認識を行う事が可能となる。その結果得たメタデータを追加記憶する事が可能となる。

【 0 0 2 8 】

これにより、画像認識対象が増えた場合に、過去に登録した画像を最初から解析する事無く、安価な処理コストで増えた画像認識対象を反映した画像検索用のDBを更新する事が可能となる。

【 0 0 2 9 】

[ 第 1 の実施形態 ]

30

先ず、本実施形態に係る画像処理装置のハードウェア構成例について、図1のブロック図を用いて説明する。なお、図1に示した構成はあくまでも一例であり、画像処理装置が行うものとして後述する各処理を実行可能な構成であれば、如何なる構成を採用しても良い。例えば、画像を撮像して本装置に入力するためのデジタルカメラを本装置に接続するようにしても良い。また、本装置で検索結果として得た情報を外部の機器に対して出力可能な構成を加えても良い。

【 0 0 3 0 】

CPU101は、RAM105やROM106に格納されているコンピュータプログラムやデータを用いて処理を実行することで、画像処理装置全体の動作制御を行うと共に、画像処理装置が行うものとして後述する各処理を実行する。

40

【 0 0 3 1 】

入力装置102は、アルファベットキー、ひらがなキー、カタカナキー、句点等の文字記号入力キー、カーソル移動を指示するカーソル移動キー等のような、各種の機能キーを有するキーボードや、マウス、スティックポイントなどにより構成されている。ユーザはこの入力装置102を操作することで各種の指示をCPU101に対して入力することができる。

【 0 0 3 2 】

出力装置103は、CRTや液晶画面等により構成されている表示装置であり、CPU101による処理結果を画像や文字などでもって表示することができる。

【 0 0 3 3 】

50

記憶装置 104 には、OS（オペレーティングシステム）や、後述するデータベース、画像処理装置が行うものとして後述する各処理を CPU 101 に実行させるためのコンピュータプログラムやデータが保存されている。また、既知の情報として後述する各情報も、この記憶装置 104 に保存されている。記憶装置 104 に保存されているコンピュータプログラムやデータは、CPU 101 による制御に従って適宜 RAM 105 にロードされ、CPU 101 による処理対象となる。なお、この記憶装置 104 には、ハードディスク、CD-ROM、DVD-ROM、フラッシュメモリ等を用いることができる。

#### 【0034】

RAM 105 は、記憶装置 104 からロードされたコンピュータプログラムやデータを一時的に記憶するためのエリア、CPU 101 が各種の処理を実行する際に用いるワークエリア、等を有する。即ち、RAM 105 は、各種のエリアを適宜提供することができる。ROM 106 には、画像処理装置の設定データやブートプログラムなどが格納されている。上記の各部はバス（BUS）107 に接続されている。

#### 【0035】

次に、本画像処理装置への画像登録処理と、登録した画像から目的の画像をクエリ画像に基づいて検索する画像検索処理と、について説明する。画像登録処理及び画像検索処理の為に使用する画像処理装置の機能の構成例について図 2 のブロック図を用いて説明する。

#### 【0036】

画像特徴抽出・認識処理部 202 は、記憶装置 104 からロードされた、若しくは外部機器から入力された登録画像 201 から特徴点（局所特徴点）とその特徴量（局所特徴量）を抽出し、抽出した特徴量を用いて登録画像 201 に対する画像認識処理を行う。そして画像特徴抽出・認識処理部 202 は、この画像認識処理の結果をメタデータとして出力する。このメタデータには、登録画像 201 中における特徴点の座標位置と該特徴点における特徴量、登録画像 201 中で認識できたオブジェクトの属性を示す属性情報、画像認識尤度、該オブジェクトの領域を示す領域情報、が含まれている。画像特徴抽出・認識処理部 202 の詳細については、図 3 を用いて後述する。

#### 【0037】

本実施形態では、局所特徴量を共通化する事により、局所特徴量の比較結果と画像認識の結果が大きく異ならないようにする。それぞれの結果が大きく異なってしまう例としては、局所特徴量が輝度情報から得られるもので、画像認識で色情報を用いると、これらをハイブリッドした場合にそれぞれでヒットする画像の見かけが大きく異なり不自然に結果となってしまう。

#### 【0038】

画像特徴・認識結果登録処理部 203 は、画像特徴抽出・認識処理部 202 が出力したメタデータを、このメタデータの生成元である登録画像 201 と関連付けてデータベース（DB）204 に登録する。このように、画像処理装置 200 は、登録画像 201 が入力されるたびに、該登録画像 201 から上記のメタデータを生成し、該生成したメタデータを該登録画像 201 と関連付けて DB 204 に登録する、いわゆる画像登録処理を行う。

#### 【0039】

画像特徴抽出・認識処理部 202' は、受け付ける画像が画像特徴抽出・認識処理部 202 と異なるだけで、画像特徴抽出・認識処理部 202 と同様の構成を有し且つ同様の動作を行うものである。画像特徴抽出・認識処理部 202' は、記憶装置 104 からロードされた、若しくは外部機器から入力されたクエリ画像 205 から特徴点とその特徴量を抽出し、抽出した特徴量を用いて該クエリ画像 205 に対する画像認識処理を行う。そして画像特徴抽出・認識処理部 202' は、この画像認識処理の結果をメタデータとして出力する。このメタデータには、クエリ画像 205 中における特徴点の座標位置と該特徴点における特徴量、クエリ画像 205 中で認識できたオブジェクトの属性を示す属性情報、画像認識尤度、該オブジェクトの領域を示す領域情報、が含まれている。画像特徴抽出・認識処理部 202' の詳細については、図 3 を用いて後述する。

10

20

30

40

50



## 【 0 0 4 0 】

画像比較処理部 2 0 6 は、画像特徴抽出・認識処理部 2 0 2 ' によってクエリ画像 2 0 5 から生成されたメタデータを用いて D B 2 0 4 から画像検索を行い、画像検索結果（検索処理結果）2 0 7 を出力する。ここでいう検索は、局所特徴量同士を比較する検索ではなく、画像認識結果のメタデータが一致または類似するかどうか比較する検索である。画像比較処理部 2 0 6 の詳細については後述する。

## 【 0 0 4 1 】

D B 2 0 4 は、画像処理装置内に備えることに限るものではなく、画像処理装置がアクセス可能に該画像処理装置に直接的若しくは間接的に接続されていれば、その接続形態は特定の接続形態に限るものではない。

10

## 【 0 0 4 2 】

なお、図 2 に示した各部はハードウェアで構成しても良いが、本実施形態では、D B 2 0 4 を除く各部（認識辞書 3 0 5 は除く）はコンピュータプログラムで実装し、D B 2 0 4 及び認識辞書 3 0 5 は記憶装置 1 0 4 で実装する。また、以下の説明では、図 2 の各部を動作の主体として説明するが、実際には、C P U 1 0 1 が対応する機能部のコンピュータプログラムを実行することで対応する動作が実行される。

## 【 0 0 4 3 】

次に、画像特徴抽出・認識処理部 2 0 2 （2 0 2 ' ）の機能構成例について、図 3 のブロック図を用いて説明する。

## 【 0 0 4 4 】

20

画像入力部 3 0 2 は、入力画像 3 0 1 （画像特徴抽出・認識処理部 2 0 2 の場合は登録画像 2 0 1、画像特徴抽出・認識処理部 2 0 2 ' の場合はクエリ画像 2 0 5 に相当する）を受けると、この入力画像 3 0 1 を特徴点 / 局所特徴量算出部 3 0 3 に転送する。

## 【 0 0 4 5 】

特徴点 / 局所特徴量算出部 3 0 3 は、画像入力部 3 0 2 から転送された入力画像 3 0 1 から、特徴点とその特徴量を算出する。特徴点 / 局所特徴量算出部 3 0 3 の動作の詳細については後述する。

## 【 0 0 4 6 】

画像認識処理部 3 0 4 は、認識辞書 3 0 5 を参照しながら、特徴点 / 局所特徴量算出部 3 0 3 が求めた特徴点とその特徴量を用いて、入力画像 3 0 1 に対する画像認識処理を行う。認識辞書 3 0 5 には複数の画像認識対象の情報が登録されている。そして画像認識処理部 3 0 4 は、画像認識処理の結果として上記のメタデータを局所特徴量及び画像認識結果 3 0 6 として出力する。

30

## 【 0 0 4 7 】

次に、特徴点 / 局所特徴量算出部 3 0 3 についてより詳細に説明する。従来技術として説明したように、画像の局所的な特徴量（局所特徴量）を用いたマッチングにより類似画像を検索する方法、画像特徴量検索が提案されている。

## 【 0 0 4 8 】

局所特徴量を利用する手法においては、局所特徴量を回転不変、拡大・縮小不変となる複数の要素で構成される情報として定義する。これにより、画像を回転させたり、拡大又は縮小させたりした場合であっても、検索を可能にする。

40

## 【 0 0 4 9 】

局所特徴量は一般的にベクトルとして表現される。ただし、局所特徴量が回転不変、拡大・縮小不変であることは理論上の話であり、実際のデジタル画像においては、線形補間をするので、画像の回転や拡大・縮小処理前の局所特徴量と処理後の対応する局所特徴量との間に若干の変動が生じる。

## 【 0 0 5 0 】

回転不変の局所特徴量を算出するために、たとえば非特許文献 2 では、局所特徴点周辺の局所領域の画素パターンから主方向を算出し、局所特徴量算出時に主方向を基準に局所領域を回転させて方向の正規化を行う。

50

## 【 0 0 5 1 】

また、拡大・縮小不変の局所特徴量を算出するために、異なるスケールの画像を内部で生成し、各スケールの画像からそれぞれ局所特徴点の抽出と局所特徴量の算出を行う。ここで、内部で生成した一連の異なるスケールの画像集合は一般的にスケールスペースと呼ばれ、図5にその概要を示す。各特徴点の特徴量は、方向の正規化の際の回転角度および特徴量を算出したスケールスペースも記憶する。

## 【 0 0 5 2 】

次に、特徴点 / 局所特徴量算出部 3 0 3 が行う処理について、同処理のフローチャートを示す図4を用いて説明する。ステップ S 4 0 1 では、画像入力部 3 0 2 から転送された入力画像 3 0 1 を受け取り、ステップ S 4 0 2 では、この入力画像 3 0 1 の輝度画像を生成する。

10

## 【 0 0 5 3 】

ステップ S 4 0 3 では、ステップ S 4 0 2 で生成した輝度画像の縮小画像を生成し、ステップ S 4 0 4 では、この縮小画像から特徴点を抽出し、ステップ S 4 0 5 では、この特徴点における特徴量を抽出する。

## 【 0 0 5 4 】

次に、画像比較処理部 2 0 6 の詳細について説明する。特徴点 / 局所特徴量の比較に基づく画像の照合方法にはいろいろあるが、ここでは以下の文献で説明がなされている R A N S A C を利用した方法を説明する。

## 【 0 0 5 5 】

20

M . A . F i s c h l e r a n d R . C . B o l l e s , “ R a n d o m s a m p l e c o n s e n s u s : A p a r a d i g m f o r m o d e l f i t t i n g w i t h a p p l i c a t i o n s t o i m a g e a n a l y s i s a n d a u t o m a t e d c a r t o g r a p h y , ” C o m m u n . A C M , n o . 2 4 , v o l . 6 , p p . 3 8 1 - 3 9 5 , J u n e 1 9 8 1 .

クエリ画像の各特徴点に対し、DB 2 0 4 が管理しているそれぞれの登録画像の特徴点で、特徴間距離が最小となるものをペアで記述する。次に、クエリ画像から3個の特徴点をランダムに選択し、それぞれの特徴間距離が最小となる登録画像の特徴点群との間で、その座標の対応からアフィン変換行列を求める。このアフィン変換行列を用い、クエリ画像の残りの特徴点の座標を登録画像の座標に変換し、その近傍に上記特徴間距離が最小となるペアの特徴点が存在するかを確認し、存在すれば1票投票し、存在しなければ投票しない。最終的に、この投票数が所定の値に達した場合には、クエリ画像と登録画像は部分一致する領域が存在すると判断し、その投票数が多いほど一致する領域が大きいと考える。

30

## 【 0 0 5 6 】

他方、投票数が所定の値に達しない場合には、新たにクエリ画像から3個の特徴点をランダムに選択し、アフィン変換行列を求める処理から再度処理を行うが、この再処理は定められた反復カウント数以内で繰り返す。もし、反復カウント数に達しても投票数が所定の値を超えなければ、部分一致する領域が存在しないと判断して比較の処理を終了する。そして、部分一致する領域が存在する場合、上記求めたアフィン変換行列と特徴間距離が最小となるもののペアを用い、クエリ画像中の着目した特徴点と対応する特徴点を求めることが出来る。

40

## 【 0 0 5 7 】

更に、その特徴点を求める際の方向の正規化の際の回転角度からクエリ画像と登録画像の回転角度関係を求めることが出来る。また更に、その特徴点を求める際のスケールスペースを用いて、クエリ画像と登録画像の拡張縮小関係を知る事が出来る。

## 【 0 0 5 8 】

また、上記の通り、非特許文献3には、一般物体認識技術として次の様な技術が記されている。ある認識対象カテゴリに属する画像群から特徴量を算出したものを学習データと

50

する。その分布モデルを学習データを用いて最尤推定で推定し、未知のデータに対して分布モデルを利用して事後確立を計算する。そして、事後確率最大化 (Maximum A Posteriori, MAP) 推定によってどのカテゴリに属するかを決定し、これを認識結果とする。その他、様々なモデルや機械学習方式を組み合わせた方法が数多く存在し、本実施形態で用いる画像認識処理は、局所特徴量を用いて行うものであればその機械学習方法を問わない。

【0059】

次に、画像処理装置200が行う画像登録処理について、同処理のフローチャートを示す図6を用いて説明する。ステップS601では、画像入力部302は、記憶装置104からロードされた、若しくは外部機器から入力された登録画像201を取得する。ステップS602では、特徴点/局所特徴量算出部303は、ステップS601で取得した登録画像201から特徴点とその特徴量を抽出する。

10

【0060】

ステップS603では、画像認識処理部304は、ステップS602で抽出した特徴量を用いて、登録画像201に対する画像認識処理を行う。画像認識処理には機械学習済みの検出器を用いる。画像認識対象が複数種ある場合には、それに合わせて検出器も複数種用意しておく必要がある。

【0061】

各検出器については、事前に正解画像として認識したいカテゴリの様々なアピアランスの画像を用意し、また非正解画像としてカテゴリ外の画像を用意し、機械学習を行う。SIFTの様な回転/拡大縮小不変の局所特徴量を用いる事により、回転、拡大や縮小した認識対象のオブジェクトの有る画像を認識する事が可能である。

20

【0062】

更に、その画像中のどこにオブジェクトが存在するかに関しては、オブジェクトが存在すると判断した画像に対し、マスク領域を設定して画像内をスキャンさせ、マスクサイズを小さくしていった検出器に入力し、その認識尤度を求める。そして、認識尤度が予め定めた閾値よりも高くなったマスク領域に、認識対象となるオブジェクトが存在すると推定する。

【0063】

そしてステップS604では、画像特徴・認識結果登録処理部203は、ステップS603で行った画像認識処理の結果をメタデータとして、登録画像201と関連付けてDB204に登録する。

30

【0064】

上記の画像登録処理によってDB204に登録された情報について、即ち、DB204内のインデックススキーマについて、図7を用いて説明する。図7(a)は、画像管理インデックスの構成例を示しており、これは登録画像ごとにDB204に登録されているものである。

【0065】

「画像ID」のフィールドには、登録画像201に固有のIDが登録され、このIDは例えば登録画像201に固有の整数値である。後述する検索の結果は、画像IDと類似度とのセットで表される。

40

【0066】

「画像データファイル名」のフィールドには、登録画像201が登録されている登録先(例えばDB204)における該登録画像201のパス名が登録される。もちろん、登録画像201の登録先はDB204に限るものではなく、他の装置であっても良い。

【0067】

「認識結果」のフィールドには、登録画像201に対する画像認識処理により認識された各オブジェクトのID(認識ID)が登録される。認識IDは、図7(b)に示す認識結果管理インデックスへのリンク情報でもある。「認識結果」のフィールドは、即ち可変長のフィールドになる。

【0068】

50

「画像特徴量」のフィールドには、登録画像 201 から抽出された特徴量の ID (画像特徴 ID) が登録される。画像特徴 ID は、図 7 (c) に示す画像特徴インデックスへのリンク情報でもある。「画像特徴量」のフィールドは、即ち可変長のフィールドに成る。

【0069】

次に、認識結果管理インデックスの構成例について、図 7 (b) を用いて説明する。認識結果管理インデックスは、認識されたオブジェクトごとに DB 204 に登録されるものである。

【0070】

「認識 ID」のフィールドには、登録画像 201 から認識されたオブジェクトに固有の ID が登録され、この ID は例えば、このオブジェクトに固有の整数値である。

10

【0071】

「認識属性」のフィールドには、登録画像 201 から認識されたオブジェクトの属性を示す属性値が登録される。図 7 では、オブジェクトの属性が「人体」であれば属性値は「1」、オブジェクトの属性が「車」であれば属性値は「2」となっている。属性の種類や数はこれに限るものではなく、属性が N 種類あれば、属性値は 1 ~ N の範囲を取る。

【0072】

「認識尤度」のフィールドには、登録画像 201 から認識されたオブジェクトの認識尤度が登録され、例えば 0 ~ 100 (%) の範囲の数値で表される。もちろん、認識尤度の表現方法はこれに限るものではない。

【0073】

20

「認識領域」のフィールドには、ステップ S 603 で行った画像認識処理で求めた認識尤度が予め定めた閾値よりも高くなったマスク領域 (= 認識領域) を示す領域情報が登録されている。ここではマスク領域の左上隅の頂点及び右下隅の頂点のそれぞれの座標位置が登録されている。もちろん、領域情報の表現方法はこれに限るものではない。

【0074】

「認識領域中の画像特徴量」のフィールドには、「認識領域」のフィールドで定義している領域情報が示す領域 (認識領域) 内のそれぞれの特徴点の特徴量の画像特徴 ID が登録されている。「認識領域中の画像特徴量」のフィールドに登録されている情報は第 1 の実施形態では用いず、第 2 の実施形態で用いるので、このフィールドに登録されている情報については第 2 の実施形態で詳しく説明する。

30

【0075】

次に、画像特徴インデックスの構成例について、図 7 (c) を用いて説明する。画像特徴インデックスは、特徴点の特徴量ごとに DB 204 に登録されるもので、1 つの特徴点に対する特徴量を 1 レコードで記憶する。

【0076】

「画像特徴 ID」のフィールドには、特徴量に固有の ID が登録される。「画像 ID」のフィールドには、該特徴量が抽出された登録画像の画像 ID が登録される。「特徴点座標」のフィールドには、登録画像上における該特徴点の座標位置が登録される。「局所特徴量」のフィールドには、ベクトル表現で表現される特徴量が登録される。

【0077】

40

ここで、図 8 (a) ~ (c) のそれぞれに示す登録画像 (登録画像 1 ~ 3) を対象にして画像特徴抽出・認識処理部 202 及び画像特徴・認識結果登録処理部 203 を動作させた場合に、DB 204 に登録される各インデックスを図 9 (a) ~ (c) に示す。

【0078】

図 8 (a) に示した登録画像 1 (画像 ID = 1 の登録画像) には、車のオブジェクト 801 ~ 803 と、バイクのオブジェクト 804 が含まれている。ここで、本実施形態では、画像認識において、様々な種類の画像を学習し、これら 3 つの車の画像が認識されるものとし、バイクは画像認識対象とせず、認識されないものとする。図 8 (b)、(c) のそれぞれに示した登録画像 2, 3 (それぞれ画像 ID = 2, 3 の登録画像) には、車のオブジェクト 805, 人のオブジェクト 806 が含まれている。

50

## 【 0 0 7 9 】

このとき、図 9 ( a ) に示す如く、DB 2 0 4 に登録される画像管理インデックスにおいて、「画像 ID」のフィールドには、それぞれの登録画像の画像 ID ( 1 ~ 3 ) が登録されている。

## 【 0 0 8 0 】

画像 ID = 1 が登録されている行 ( レコード ) における「画像データファイル名」のフィールドには、画像 ID = 1 の登録画像のパス名が登録されている。また、同行の「認識結果」のフィールドには、オブジェクト 8 0 1 ~ 8 0 3 のそれぞれの ID ( 認識 ID ) が登録されており、同行の「画像特徴量」のフィールドには、画像 ID = 1 の登録画像から抽出された特徴量の ID が登録されている。なお、上記の取り、オブジェクト 8 0 4 については画像認識は行われませんが、その特徴量の ID は「画像特徴量」のフィールドに登録されている。図 9 では、登録画像 1 から抽出した全特徴点に対する局所特徴量が 8 0 0 個検出されたとしているため、画像特徴 ID を 1 ~ 8 0 0 の間で付与する。

10

## 【 0 0 8 1 】

画像 ID = 2 が登録されている行における「画像データファイル名」のフィールドには、画像 ID = 2 の登録画像のパス名が登録されている。また、同行の「認識結果」のフィールドには、オブジェクト 8 0 5 の ID ( 認識 ID ) が登録されており、同行の「画像特徴量」のフィールドには、画像 ID = 2 の登録画像から抽出された特徴量の ID が登録されている。図 9 では、登録画像 2 から抽出した全特徴点に対する局所特徴量が 2 0 0 個検出されたとしているため、画像特徴 ID を 8 0 1 ~ 1 0 0 0 の間で付与する。

20

## 【 0 0 8 2 】

画像 ID = 3 が登録されている行における「画像データファイル名」のフィールドには、画像 ID = 3 の登録画像のパス名が登録されている。また、同行の「認識結果」のフィールドには、オブジェクト 8 0 6 の ID ( 認識 ID ) が登録されており、同行の「画像特徴量」のフィールドには、画像 ID = 3 の登録画像から抽出された特徴量の ID が登録されている。図 9 では、登録画像 3 から抽出した全特徴点に対する局所特徴量が 2 0 0 個検出されたとしているため、画像特徴 ID を 1 0 0 1 ~ 1 2 0 0 の間で付与する。

## 【 0 0 8 3 】

また、図 9 ( b ) に示す如く、DB 2 0 4 に登録される認識結果管理インデックスにおいて、「認識 ID」のフィールドには、オブジェクト 8 0 1 ~ 8 0 3、8 0 5 ~ 8 0 6 のそれぞれの ID、即ち、1 ~ 5 が登録されている。図 9 ( b ) に示す如く、登録画像 1 から認識した結果が認識 ID = 1 ~ 3 のレコードに登録されており、登録画像 2 から認識した結果が認識 ID = 4 のレコードに登録されており、登録画像 3 から認識した結果が認識 ID = 5 のレコードに登録されている。

30

## 【 0 0 8 4 】

認識 ID = 1 が登録されている行における「認識属性」のフィールドには、認識 ID = 1 のオブジェクト、即ち車のオブジェクト 8 0 1 の属性を示す属性値が登録されている。また、同行の「画像 ID」のフィールドには、このオブジェクト 8 0 1 の認識元である登録画像 1 の ID が登録されており、同行の「認識尤度」のフィールドには、登録画像 1 からオブジェクト 8 0 1 を画像認識した際の認識尤度が登録されている。また、同行の「認識領域」のフィールドには、オブジェクト 8 0 1 について画像認識処理で求めた認識尤度が予め定めた閾値よりも高くなったマスク領域を示す領域情報が登録されている。また、同行の「認識領域中の画像特徴量」のフィールドには、オブジェクト 8 0 1 について画像認識処理で求めた認識尤度が予め定めた閾値よりも高くなったマスク領域内のそれぞれの特徴点の特徴量の画像特徴 ID が登録されている。認識 ID = 2 ~ 5 のそれぞれのレコードも同様の構成を有している。

40

## 【 0 0 8 5 】

今回、車を学習するに当たり、ピックアップトラックの画像を学習に多めに用いたので、ピックアップトラックのオブジェクト 8 0 2 に対する認識尤度が最も高く 90 % となっている。また、スポーツカーのオブジェクト 8 0 1 に対する認識尤度は 85 %、バスのオブジ

50

エクト 803 に対する認識尤度は80%となっている。他方、スポーツカーのオブジェクト 805 に対する認識尤度は87%、人体のオブジェクト 806 に対する認識尤度は90%の尤度となっている。

【0086】

認識領域に関しては、各認識IDに対応するオブジェクト801～803、オブジェクト805～806のそれぞれについて、登録画像1～3のうち該当する画像中の位置を記憶する。

【0087】

また、図9(c)に示す如く、画像特徴ID=1～1200のそれぞれに対し、画像特徴インデックスが生成される。画像特徴ID=1が登録されている行における「画像ID」のフィールドには、画像特徴ID=1の特徴量の抽出元である登録画像のIDが登録されている。同行の「特徴点座標」のフィールドには、画像特徴ID=1の特徴点の登録画像1上の座標位置が登録されている。同行の「局所特徴量」のフィールドには、画像特徴ID=1に対応する特徴量のベクトルが登録されている。画像特徴ID=2～1200のそれぞれのレコードも同様の構成を有している。

【0088】

ところで、「認識領域中の画像特徴量」のフィールドに関しては上記の通り、第2の実施形態で使用するものであるが、上記認識領域中に含まれる特徴点を画像特徴ID群で記憶する。

【0089】

次に、図8(d)に示したクエリ画像に類似する画像をDB204から検索する画像検索処理について、同処理のフローチャートを示す図10を用いて説明する。なお、この処理を拡張すれば、入力されたクエリ画像中のオブジェクト毎に、該オブジェクトをクエリとして類似するオブジェクトを含む画像を検索することもできる。

【0090】

ステップS1001では、画像入力部302は、記憶装置104からロードされた、若しくは外部機器から入力された図8(d)のクエリ画像を取得し、特徴点/局所特徴量算出部303に転送する。なお、画像入力部302は、取得したクエリ画像を出力装置103の画面に表示させた後、ユーザが入力装置102を用いて指定したオブジェクト領域内の画像を、改めてクエリ画像として特徴点/局所特徴量算出部303に転送するようにしても良い。

【0091】

ステップS1002では、特徴点/局所特徴量算出部303は、ステップS1001で取得したクエリ画像から特徴点とその特徴量を抽出する。ステップS1003では、画像認識処理部304は、ステップS1002で抽出した特徴量を用いて、クエリ画像に対する画像認識処理を行う。

【0092】

ステップS1004では、画像認識処理部304は、クエリ画像中のオブジェクトに対する画像認識処理が成功したか否か(例えば属性が特定可能なオブジェクトが1以上あったか否か)を判断する。この判断の結果、画像認識が成功したオブジェクトがあった場合には、処理はステップS1005に進み、無かった場合には、処理はステップS1006に進む。

【0093】

ステップS1005では、画像比較処理部206は先ず、ステップS1003における画像認識処理で求めたオブジェクトの属性値と同じ属性値が登録されたレコードを、DB204内の認識結果管理インデックスから特定する。ここでは図8(d)に示す車が写っている画像をクエリ画像としている。そのため、この車が画像認識されてその結果、オブジェクトの属性値は「2」であるため、この場合、認識結果管理インデックスからは、属性値「2」が登録されているレコード、即ち、認識ID=1～4のレコードが特定される。ここで、認識ID=1～3は何れも登録画像1に対するものであるため、認識ID=1

10

20

30

40

50

～ 3 のレコードに登録されている認識尤度のうち、最も大きい認識尤度を、登録画像 1 に対する認識尤度とする。従って、ここでは、クエリ画像に対する類似度が 90 % の検索結果画像が登録画像 1、クエリ画像に対する類似度が 87 % である検索結果画像が登録画像 2、として、検索結果を得る。

【 0094 】

そして画像比較処理部 206 は、認識尤度の高い順、即ち、登録画像 1、登録画像 2、の順番で画像 ID と類似度と、を出力する。この場合（画像 ID、類似度）＝（1, 90）、（2, 87）の順に出力する。なお、認識尤度が 0 から 100 % で表されない場合には、0 から 100 % に成る様に換算して類似度とする。これにより、画像特徴の比較処理無く高速に検索可能である。

10

【 0095 】

一方、クエリ画像として図 8（e）の画像を用いた場合、ステップ S1004 では、画像認識処理部 304 は、クエリ画像中のオブジェクトに対する画像認識処理が成功していないと判断し、処理はステップ S1006 に進む。

【 0096 】

ステップ S1006 では、画像比較処理部 206 は、図 9（a）の画像管理インデックスに登録されている各画像 ID のレコードを順次参照する。そして画像比較処理部 206 は、該参照したレコード中の各画像特徴 ID の画像特徴インデックス中の「局所特徴量」のフィールドに登録されているベクトルと、クエリ画像から求めた特徴量と、を用いて上記の RANSAC を用いたマッチングを行う。そして、クエリ画像から求めた特徴量と類似する順に、画像 ID とその類似度と、を検索結果として出力する。

20

【 0097 】

このような処理により、特徴量を用いた画像検索のオブジェクトを選ばない性質と、画像認識の対象オブジェクトを選ぶが誤った検出が比較的少ないという性質を組み合わせる事ができた。その効果として、画像認識可能なオブジェクトがクエリ画像の場合には、極めて高速かつ、ノイズの少ない良好な検索結果をもたらすことができた。

【 0098 】

〔 第 2 の実施形態 〕

第 1 の実施形態では、クエリ画像が認識可能な場合には、上記のような処理を行うことで、高速に類似画像検索が可能である。しかし、第 1 の実施形態によれば、最も類似度が高いオブジェクトは画像認識尤度が最も高い、登録画像 1 中のオブジェクト 802 となる。これは、クエリ画像にアピアランスが似ているは登録画像 2 中のオブジェクト 805 と思う多くのユーザの期待とは異なる結果である。本実施形態では、このような課題を解決するものである。

30

【 0099 】

本実施形態は、画像登録処理は第 1 の実施形態と同様であるが、画像検索処理において以下の点が第 1 の実施形態と異なる。以下に、第 1 の実施形態との差分のみについて説明し、以下で特に触れない限りは、第 1 の実施形態と同様であるものとする。本実施形態に係る画像検索処理について、同処理のフローチャートを示す図 11 を用いて説明する。

【 0100 】

ステップ S1101 では、画像入力部 302 は、記憶装置 104 からロードされた、若しくは外部機器から入力された図 8（d）のクエリ画像を取得し、特徴点 / 局所特徴量算出部 303 に転送する。なお、画像入力部 302 は、取得したクエリ画像を出力装置 103 の画面に表示させた後、ユーザが入力装置 102 を用いて指定したオブジェクト領域内の画像を、改めてクエリ画像として特徴点 / 局所特徴量算出部 303 に転送するようにしても良い。

40

【 0101 】

ステップ S1102 では、特徴点 / 局所特徴量算出部 303 は、ステップ S1101 で取得したクエリ画像から特徴点とその特徴量を抽出する。ステップ S1103 では、画像認識処理部 304 は、ステップ S1102 で抽出した特徴量を用いて、クエリ画像に対す

50

る画像認識処理を行う。

【 0 1 0 2 】

ステップ S 1 1 0 4 では、画像認識処理部 3 0 4 は、クエリ画像中のオブジェクトに対する画像認識処理が成功したか否か（例えば属性が特定可能なオブジェクトがあったか否か）を判断する。この判断の結果、画像認識が成功したオブジェクトがあった場合には、処理はステップ S 1 1 0 5 に進み、無かった場合には、処理はステップ S 1 1 0 7 に進む。

【 0 1 0 3 】

ステップ S 1 1 0 5 では画像比較処理部 2 0 6 は上記のステップ S 1 0 0 5 と同様にして、ステップ S 1 1 0 3 における画像認識処理で求めたオブジェクトの属性値と同じ属性値が登録されたレコードを、DB 2 0 4 内の認識結果管理インデックスから特定する。ここでは図 8 ( d ) に示す車が写っている画像をクエリ画像としている。そのため、この車が画像認識されてその結果、オブジェクトの属性値は「 2 」であるため、この場合、認識結果管理インデックスからは、属性値「 2 」が登録されているレコード、即ち、認識 ID = 1 ~ 4 のレコードが特定される。なお、認識 ID = 1 ~ 4 のレコードのうち、認識尤度が閾値以上のレコードのみを以降の対象としても良い。

【 0 1 0 4 】

ステップ S 1 1 0 6 で画像比較処理部 2 0 6 は、ステップ S 1 1 0 5 で特定した各レコードについて、次のような処理を行う。即ち、該レコード中の画像特徴 ID に対応する画像特徴インデックス内の特徴量と、該レコード内の「認識領域」のフィールドで定義されている領域に対応するクエリ画像内の領域中の特徴量と、のマッチングを行って類似度を求める。これにより、局所特徴量を比較する画像およびその領域を絞り込む事が出来、比較処理のコストを大幅に低減可能である。マッチングに関しては、上記の R A N S A C 法を用いる事が可能である。

【 0 1 0 5 】

或いは、マッチングに関しては多次元の、S I F T と言えば 128 次元の局所特徴量をクラスタリングし、各クラスに属す特徴点個数のヒストグラムを作成し、ヒストグラムの正規化相関か或いはインターセクションを用いる事で、類似度の算出が可能である。

【 0 1 0 6 】

そして画像比較処理部 2 0 6 は、第 1 の実施形態と同様にして、より高い類似度を求めたレコード順に、該レコード内の画像 ID と、この類似度と、を検索結果として出力する。

【 0 1 0 7 】

このように、登録画像とクエリ画像の認識領域中の局所特徴量を比較する事により、認識尤度だけでは解決できなかったクエリとの類似性を反映しつつ、認識結果を絞り込みに使用した効率の良い検索処理を行う事が可能になる。

【 0 1 0 8 】

一方、クエリ画像として図 8 ( e ) の画像を用いた場合、ステップ S 1 0 0 4 では、画像認識処理部 3 0 4 は、クエリ画像中のオブジェクトに対する画像認識処理が成功していないと判断し、処理はステップ S 1 1 0 7 に進む。

【 0 1 0 9 】

ステップ S 1 1 0 7 では、画像比較処理部 2 0 6 は、図 9 ( a ) の画像管理インデックスに登録されている各画像 ID のレコードを順次参照する。そして画像比較処理部 2 0 6 は、該参照したレコード中の各画像特徴 ID の画像特徴インデックス中の「局所特徴量」のフィールドに登録されているベクトルと、クエリ画像から求めた特徴量と、を用いて上記の RANSAC を用いたマッチングを行う。そして、クエリ画像から求めた特徴量と類似する順に、画像 ID とその類似度と、を検索結果として出力する。

【 0 1 1 0 】

[ 第 3 の実施形態 ]

本実施形態では、第 1 の実施形態及び第 2 の実施形態において、クエリ画像、登録画像

10

20

30

40

50



のどちらか一方において画像認識がうまく機能せず、結果として検索漏れを起こすという課題を解決する。

【 0 1 1 1 】

本実施形態は、画像登録処理は第 1 の実施形態と同様であるが、画像検索処理において以下の点が第 1 の実施形態と異なる。以下に、第 1 の実施形態との差分のみについて説明し、以下で特に触れない限りは、第 1 の実施形態と同様であるものとする。

【 0 1 1 2 】

本実施形態に係る画像検索処理について、同処理のフローチャートを示す図 1 2 を用いて説明する。なお、ステップ S 1 2 0 1 ~ S 1 2 0 4 , S 1 2 0 8 はそれぞれ、図 1 0 のステップ S 1 0 0 1 ~ S 1 0 0 4 , S 1 0 0 6 と同じ処理ステップであり、ステップ S 1 0 0 5 の代わりにステップ S 1 2 0 5 ~ S 1 2 0 7 を行う点が第 1 の実施形態と異なる。

10

【 0 1 1 3 】

ステップ S 1 2 0 5 では、画像比較処理部 2 0 6 は先ず、ステップ S 1 2 0 3 における画像認識処理で求めたオブジェクトの属性値と同じ属性値が登録されたレコードを、DB 2 0 4 内の認識結果管理インデックスから特定する。そして画像比較処理部 2 0 6 は、特定したレコードから認識尤度を読み出す。

【 0 1 1 4 】

ステップ S 1 2 0 6 では、画像比較処理部 2 0 6 は、ステップ S 1 2 0 5 で特定した各レコードについて次のような処理を行う。即ち、該レコード中の画像 ID に対応する画像管理インデックスから「画像特徴量」のフィールド中の画像特徴 ID を読み出し、該画像特徴 ID に対応する特徴量を画像特徴インデックスから読み出し、クエリ画像中の特徴量とのマッチングを行う。このマッチングには、上記の RANSAC を用い、これにより特徴量間の類似度を求める。

20

【 0 1 1 5 】

ステップ S 1 2 0 7 では、画像比較処理部 2 0 6 は、ステップ S 1 2 0 5 で特定した各レコードについて次のような処理を行う。即ち、該レコードについてステップ S 1 2 0 5 読み出した認識尤度  $LH$  と、該レコードについてステップ S 1 2 0 6 で求めた類似度  $Sim$  と、をそれぞれに対する重み  $w_1$  ,  $w_2$  を用いて重み付け加算を行うことで総合類似度を求める。この重み付け加算の式を以下に示す。

【 0 1 1 6 】

総合類似度  $Sim_{total} = w_1 \times LH + w_2 \times Sim$ , 但し  $w_1 + w_2 = 1$

30

そして、類似度の代わりに総合類似度を適用してステップ S 1 0 0 5 の処理を行うことで、画像認識で失敗した画像についても、局所特徴量比較で高いスコアを出した画像であれば検索結果に反映する事が可能となる。

【 0 1 1 7 】

また、 $w_1$  は認識対象毎に変えても良く、例えば、クエリで認識率の低い認識対象が検出された場合には低い値の  $w_1$  を、認識率が高い認識対象の場合には高い値の  $w_1$  を用いると良い。画像認識が可能なオブジェクトを持つ画像に対して、認識尤度が緩やかに効く様に画像認識尤度と類似度の重み付け和で総合評価を得る。

【 0 1 1 8 】

このように、本実施形態によれば、局所特徴量を用いた画像検索のオブジェクトを選ばない性質と、画像認識の対象画像オブジェクトを選ぶが誤った検出が比較的少ないという性質を見合わせた、良好な検索結果をもたらす画像検索を実現する。

40

【 0 1 1 9 】

また、上記処理により、第 1 の実施形態よりは処理コストが高くなるものの、より検索漏れの少なく且つ画像認識結果を緩やかに反映したロバストな検索が可能となる。

【 0 1 2 0 】

[ 第 4 の実施形態 ]

画像認識の対象は時間が経過すると共に増加するので、登録済みの画像に対しても新しい画像認識結果を適応する事が好ましい。もちろん、新たな画像認識対象で学習した検出

50

器を用い、最初から画像を登録する方法も有るが、登録済みの画像数が膨大で有る場合、それは現実的で無い。また、DBと画像を別個に管理している場合には、登録画像の参照が困難な場合も有る。

#### 【0121】

そこで、本実施形態では、画像認識対象の増加を反映したDBインデックスの更新を登録済みの画像を再解析する事無く、DB内で閉じた処理で効率的に画像認識対象の増加を反映する方法について説明する。各登録画像について、DBインデックスを更新するために行う処理について、同処理のフローチャートを示す図13を用いて説明する。

#### 【0122】

ステップS1301で画像特徴・認識結果登録処理部203は、登録済みの登録画像のうちの1つに着目し、該着目した登録画像の画像管理インデックス中の画像特徴IDを用いて対応する画像特徴インデックスから、特徴点の座標位置やその特徴量を読み出す。

10

#### 【0123】

ステップS1302では、画像特徴・認識結果登録処理部203は、ステップS1301で読み出した情報を用いて、上記着目した登録画像に対する画像認識処理を行う。このとき画像認識に用いる検出器は、以前よりも新たなオブジェクトの認識が可能になっている。

#### 【0124】

ステップS1303では、画像特徴・認識結果登録処理部203は、上記着目した登録画像からこれまでに認識した何れのオブジェクトとも異なる位置から新たなオブジェクトを認識したか否かを判断する。この判断の結果、新たなオブジェクトを認識していない場合には、本処理を終了するが、新たなオブジェクトを認識した場合には、処理はステップS1304に進む。

20

#### 【0125】

ステップS1304では、画像特徴・認識結果登録処理部203は、上記着目した登録画像の画像管理インデックス中の「認識結果」のフィールドに、現在の最大認識IDに1を加えた認識IDを追加登録する。

#### 【0126】

また、画像特徴・認識結果登録処理部203は、認識結果管理インデックスに1つレコードを追加し、この追加したレコードに以下に列挙する情報を登録する。即ち、ステップS1304で追加登録した認識ID、新たなオブジェクトの属性値、上記着目した登録画像の画像ID、新たなオブジェクトの認識尤度、新たなオブジェクトの認識領域の座標、新たなオブジェクトの認識用特徴量の画像特徴ID、を登録する。このように、本実施形態によれば、DB内で閉じた処理で、安価な処理コストで新たに追加した画像認識対象を反映したDBに更新する事が可能となる。

30

#### 【0127】

##### [第5の実施形態]

上記の各実施形態では、特徴量としてSIFTを用いたが、画像認識と画像特徴量検索とで共有可能な局所特徴量であれば、その種類は問わない。また、画像認識においては、MAP推定を用いた機械学習により説明したが、当然、生成モデルに基づく方法としてEMアルゴリズムを利用し反復学習する方法や、判別モデルに基づく方法でSVMを用いても良い。また、登録画像の局所特徴量に対して複数の画像認識処理を行い更新する処理も可能である。

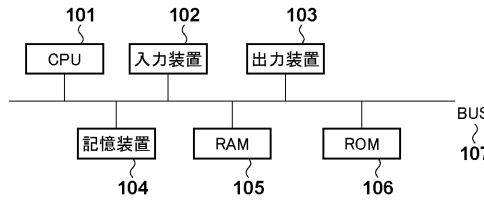
40

#### 【0128】

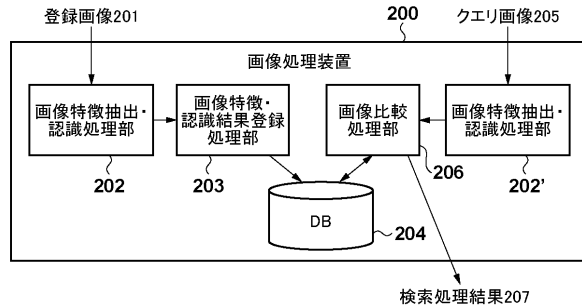
##### (その他の実施例)

また、本発明は、以下の処理を実行することによっても実現される。即ち、上述した実施形態の機能を実現するソフトウェア(プログラム)を、ネットワーク又は各種記憶媒体を介してシステム或いは装置に供給し、そのシステム或いは装置のコンピュータ(またはCPUやMPU等)がプログラムを読み出して実行する処理である。

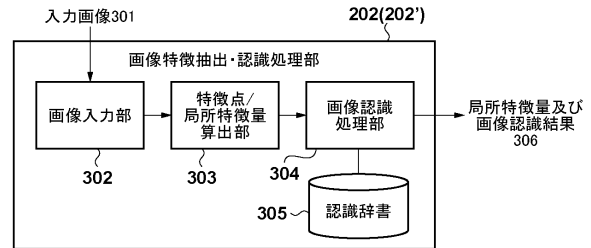
【図 1】



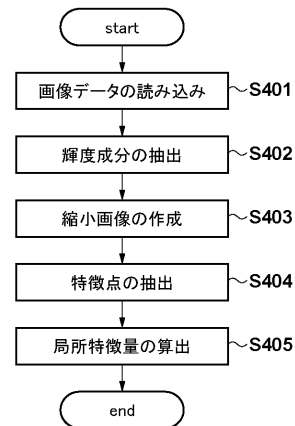
【図 2】



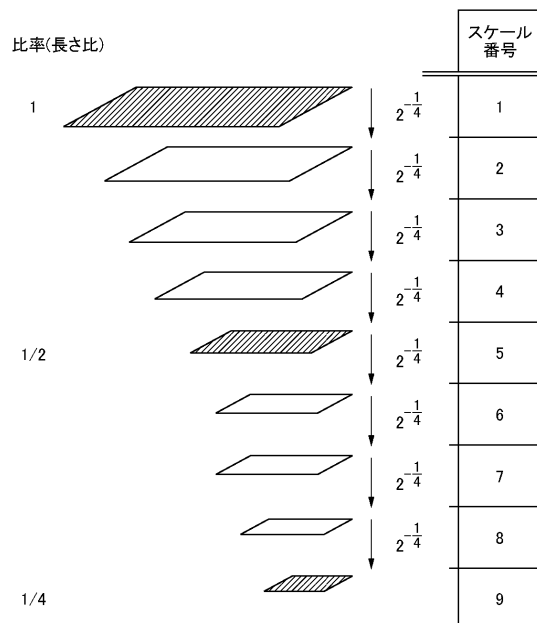
【図 3】



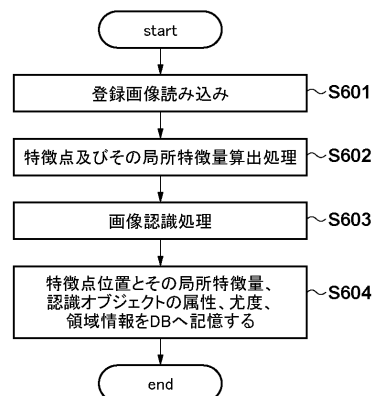
【図 4】



【図 5】



【図 6】



【図 7】

画像ID	画像データファイル名	認識結果	画像特徴量
ユニークな 整数値	画像データへのパス	認識ID群	画像特徴ID群

(a)

画像ID	認識属性	画像ID	認識尤度	認識領域	認識領域中の画像特徴量
ユニークな 整数値	属性の 種別のID 人体: 1 車: 2	ユニークな 整数値	0~100% の数値	(x1, y1)・ (x2, y2)	画像特徴ID群

(b)

画像特徴ID	画像ID	特徴点座標	局所特徴量
ユニークな 整数値	ユニークな 整数値	(x, y)	局所特徴ベクトル

(c)

【図 9】

画像ID	画像データファイル名	認識結果	画像特徴量
1	C:\data\image1.jpg	1,2,3	1,2,3,....,800
2	C:\data\image2.jpg	4	801,802,803,....,1000
3	C:\data\image3.jpg	5	1001,1002,1003,....,1200

(a)

認識ID	認識属性	画像ID	認識尤度	認識領域	認識領域中の画像特徴量
1	2	1	85	100,100,200,200	1,2,3,....,200
2	2	1	90	300,100,400,200	201,202,203,....,400
3	2	1	80	100,400,250,500	401,402,403,....,600
4	2	2	87	100,100,200,200	801,802,803,....,1000
5	1	3	90	150,50,200,150	1001,1002,1003,....,1200

(b)

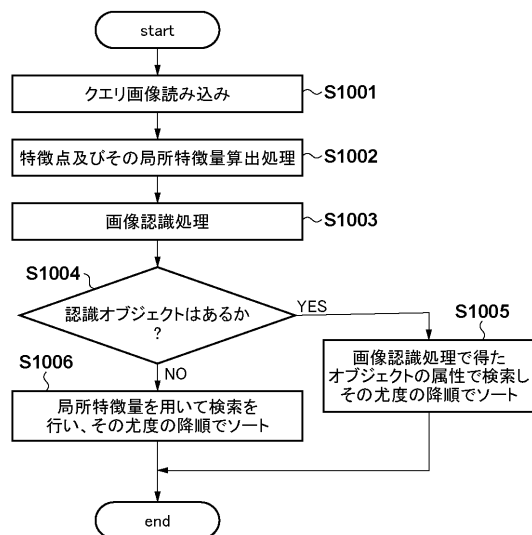
画像特徴ID	画像ID	特徴点座標	局所特徴量
1	1	150,100	v1,v2,....,v128

~

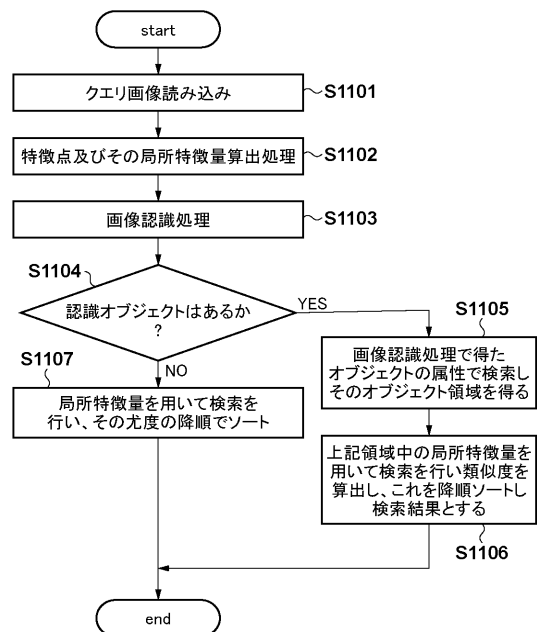
画像特徴ID	画像ID	特徴点座標	局所特徴量
1200	2	(x, y)	v1,v2,....,v128

(c)

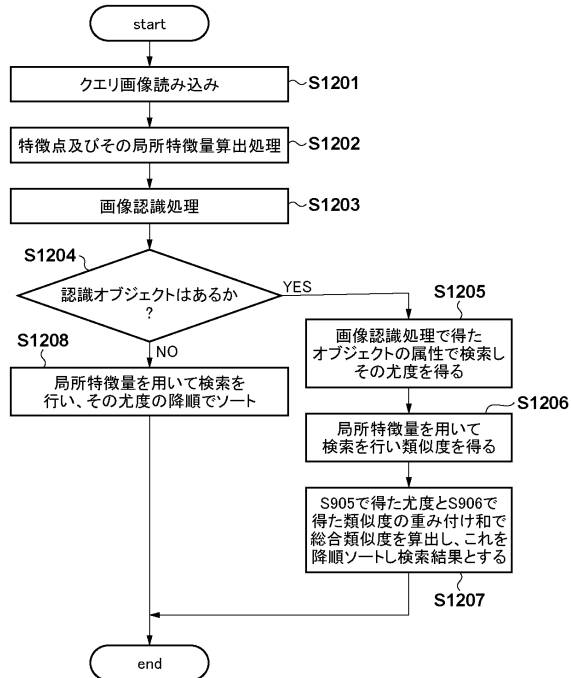
【図 10】



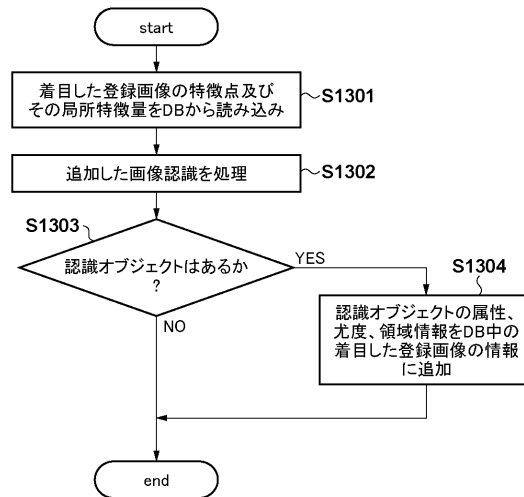
【図 11】



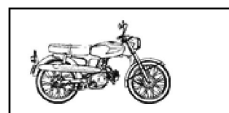
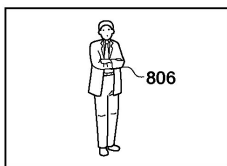
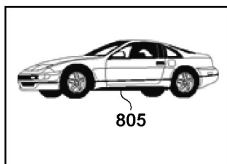
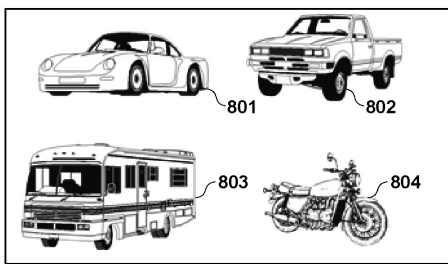
【図 12】



【図 13】



【図 8】



---

フロントページの続き

(72)発明者 椎山 弘隆  
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 小太刀 慶明

(56)参考文献 特開2010-26603(JP,A)  
特開2011-2966(JP,A)  
特表2010-518507(JP,A)  
特開平11-306322(JP,A)  
国際公開第2008/016102(WO,A1)  
特開平11-039317(JP,A)  
特開2003-187229(JP,A)

(58)調査した分野(Int.Cl., DB名)  
G06F 17/30  
G06T 1/00