

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6202938号  
(P6202938)

(45) 発行日 平成29年9月27日(2017.9.27)

(24) 登録日 平成29年9月8日(2017.9.8)

(51) Int.Cl. F 1  
G 0 6 T 7 / 0 0 (2017.01) G 0 6 T 7 / 0 0 3 0 0 F

請求項の数 10 (全 18 頁)

(21) 出願番号	特願2013-172620 (P2013-172620)	(73) 特許権者	000001007 キヤノン株式会社 東京都大田区下丸子3丁目30番2号
(22) 出願日	平成25年8月22日(2013.8.22)	(74) 代理人	100090273 弁理士 園分 孝悦
(65) 公開番号	特開2015-41293 (P2015-41293A)	(72) 発明者	野本 敦夫 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
(43) 公開日	平成27年3月2日(2015.3.2)	(72) 発明者	山本 貴久 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
審査請求日	平成28年8月12日(2016.8.12)	(72) 発明者	佐藤 博 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

最終頁に続く

(54) 【発明の名称】 画像認識装置および画像認識方法

(57) 【特許請求の範囲】

【請求項1】

登録画像および入力画像のそれぞれから物体画像を取得する取得手段と、  
前記取得手段によって取得された物体画像から、画素値の大小関係に基づく第一特徴量を抽出する第一特徴量抽出手段と、  
前記第一特徴量抽出手段によって抽出された前記登録画像に係る第一特徴量を画素値の変化の勾配に応じて変化させた第二特徴量を抽出する第二特徴抽出手段と、  
前記登録画像に係る第一および第二特徴量を辞書に登録する登録手段と、  
前記辞書に登録された特徴量と前記入力画像の特徴量とを比較することにより前記入力画像の物体と前記登録画像の物体との類似度を算出する算出手段とを備えたことを特徴とする画像認識装置。

10

【請求項2】

前記算出手段は、前記第一特徴量抽出手段によって抽出された前記入力画像に係る第一特徴量と、前記辞書に登録された第一特徴量および第二特徴量のそれぞれとを比較して類似度を算出することを特徴とする請求項1に記載の画像認識装置。

【請求項3】

前記第二特徴抽出手段は、前記入力画像に係る第一特徴量を画素値の変化の勾配に応じて変化させた第二特徴量をも抽出し、

前記算出手段は、前記第一および第二特徴量抽出手段によって抽出された前記入力画像に係る第一特徴量および第二特徴量と、前記辞書に登録された第一特徴量および第二特徴

20

量のそれぞれと比較して類似度を算出することを特徴とする請求項 1 に記載の画像認識装置。

【請求項 4】

前記第二特徴抽出手段は、画素値の変化の勾配が小さいほど前記第一特徴量のパターンが大きく変化するように前記第二特徴量を抽出することを特徴とする請求項 1 ~ 3 の何れか 1 項に記載の画像認識装置。

【請求項 5】

前記第一特徴量抽出手段は、前記登録画像および前記入力画像の局所領域から特徴量を抽出し、

前記算出手段は、前記局所領域毎に局所類似度を算出し、前記算出した局所類似度を統合した類似度を算出することを特徴とする請求項 1 ~ 4 の何れか 1 項に記載の画像認識装置。

10

【請求項 6】

前記第一および第二特徴量抽出手段は、LBP特徴量を抽出することを特徴とする請求項 1 ~ 5 の何れか 1 項に記載の画像認識装置。

【請求項 7】

前記第一および第二特徴量抽出手段は、HOG特徴量を抽出することを特徴とする請求項 1 ~ 5 の何れか 1 項に記載の画像認識装置。

【請求項 8】

登録画像および入力画像のそれぞれから物体画像を取得する取得工程と、  
前記取得工程において取得された物体画像から、画素値の大小関係に基づく第一特徴量を抽出する第一特徴量抽出工程と、  
前記第一特徴量抽出工程において抽出された前記登録画像に係る第一特徴量を画素値の変化の勾配に応じて変化させた第二特徴量を抽出する第二特徴抽出工程と、  
前記登録画像に係る第一および第二特徴量を辞書に登録する登録工程と、  
前記辞書に登録された特徴量と前記入力画像の特徴量とを比較することにより前記入力画像の物体と前記登録画像の物体との類似度を算出する算出工程とを備えたことを特徴とする画像認識方法。

20

【請求項 9】

登録画像および入力画像のそれぞれから物体画像を取得する取得工程と、  
前記取得工程において取得された物体画像から、画素値の大小関係に基づく第一特徴量を抽出する第一特徴量抽出工程と、  
前記第一特徴量抽出工程において抽出された前記登録画像に係る第一特徴量を画素値の変化の勾配に応じて変化させた第二特徴量を抽出する第二特徴抽出工程と、  
前記登録画像に係る第一および第二特徴量を辞書に登録する登録工程と、  
前記辞書に登録された特徴量と前記入力画像の特徴量とを比較することにより前記入力画像の物体と前記登録画像の物体との類似度を算出する算出工程とをコンピュータに実行させることを特徴とするプログラム。

30

【請求項 10】

請求項 9 に記載のプログラムを記憶したことを特徴とするコンピュータ読み取り可能な記憶媒体。

40

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、特に、異なる照明条件によって物体認識を行うために用いて好適な画像認識装置、画像認識方法、プログラムおよび記憶媒体に関する。

【背景技術】

【0002】

従来、画像に含まれる物体を、予め辞書に登録された物体と比較することにより物体を識別する物体認識技術が数多く提案されている。この物体認識技術の代表例としては、顔

50

認識技術が挙げられ、顔認識技術の応用例としては、画像データベースの中から特定の人物を探し出す検索用途や、デジタルカメラ等の撮影パラメタを人物毎に自動で切り換える機器制御用途などがある。これらの用途に用いられる画像は実環境下で撮影されるため、様々な撮影環境に頑健な顔認識技術が求められる。特に、辞書と入力画像とで照明条件が異なる場合に顔認証の精度を向上させることが大きな課題となっている。

【0003】

一般には、様々な照明条件で撮影された多数の顔画像を辞書に登録しておくことが望ましい。しかし、入力画像の照明条件を網羅するような顔画像を集めて辞書に登録するのは難しい。また、様々な照明条件で撮影すると、撮影対象となるユーザの負担が大きく、先に挙げた顔認証技術の応用例において、利便性を大きく損ねることになる。

10

【0004】

このような課題に対して、任意の照明条件の顔認識用情報（顔画像または特徴量）を生成し、顔認証に利用する方法が提案されている。例えば特許文献1には、入力画像と顔立体形状モデルとを用いて顔の表面反射率を推定し、その表面反射率から任意の照明条件の顔画像を生成する手法が提案されている。この方法によれば、入力画像と似た照明条件の顔画像を生成することにより照明条件の変化に頑健な顔認識を実現すると記載されている。

【0005】

また、特許文献2には、抽出した特徴量に対して座標変換を施すことによって、任意の照明条件の特徴量を生成する手法が提案されている。この手法では、特徴量に対する変換を行うことによりメモリ容量および計算コストを低減しつつ、照明条件の変化に頑健な顔認識を実現すると記載されている。

20

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特開2002-24830号公報

【特許文献2】特許第4951995号公報

【特許文献3】特開2009-211177号公報

【非特許文献】

【0007】

【非特許文献1】P. Viola and M. Jones, "Robust real-time face detection", In page II: 747, 2001.

【非特許文献2】N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. Conference on Computer Vision and Pattern Recognition (CVPR), 2005.

30

【発明の概要】

【発明が解決しようとする課題】

【0008】

しかしながら、上記の方法では、入力画像と辞書に登録されている登録画像とで照明の方向が異なる場合に依然として類似度の判定精度が低下する問題がある。特許文献1に記載の方法では、顔画像と顔立体形状モデルとを使用するため、照明条件の変化に伴う変換の質が表面反射率の推定精度によって左右されてしまう。また、顔立体形状モデルを用いる点や画像ベースで変換・生成処理を行うため、計算コストが膨大となる。また、特許文献2に記載の方法では、特徴量に対する処理により少ない計算コストを実現しているものの、座標変換によって特徴量を一律に変換してしまい、このため、顔の器官領域など、個体差が生じやすい部分のエッジまで変換してしまい、認識精度の低下を招く可能性がある。

40

【0009】

50

本発明は前述の問題点に鑑み、計算コストを抑えて照明条件の変化に頑健な物体認識を実現できるようにすることを目的としている。

【課題を解決するための手段】

【0010】

本発明の画像認識装置は、登録画像および入力画像のそれぞれから物体画像を取得する取得手段と、前記取得手段によって取得された物体画像から、画素値の大小関係に基づく第一特徴量を抽出する第一特徴量抽出手段と、前記第一特徴量抽出手段によって抽出された前記登録画像に係る第一特徴量を画素値の変化の勾配に応じて変化させた第二特徴量を抽出する第二特徴抽出手段と、前記登録画像に係る第一および第二特徴量を辞書に登録する登録手段と、前記辞書に登録された特徴量と前記入力画像の特徴量とを比較することにより前記入力画像の物体と前記登録画像の物体との類似度を算出する算出手段とを備えたことを特徴とする。

10

【発明の効果】

【0011】

本発明によれば、計算コストを抑えて照明条件の変化に頑健な物体認識を実現することができる。

【図面の簡単な説明】

【0012】

【図1】本発明の実施形態に係る画像認識装置のハードウェア構成例を示すブロック図である。

20

【図2】実施形態における画像認識装置の機能構成例を示すブロック図である。

【図3】実施形態における辞書データを登録する処理手順の一例を示すフローチャートである。

【図4】第一特徴抽出処理の詳細な手順の一例を示すフローチャートである。

【図5】注目画素とその周辺画素との関係を表わした模式図である。

【図6】第二特徴抽出処理の詳細な手順の一例を示すフローチャートである。

【図7】第1の実施形態における入力画像を識別する処理手順の一例を示すフローチャートである。

【図8】局所類似度を算出する詳細な処理手順の一例を示すフローチャートである。

【図9】統合類似度を算出する詳細な処理手順の一例を示すフローチャートである。

30

【図10】第1の実施形態において、局所類似度および統合類似度を算出する処理を模式的に説明するための図である。

【図11】第2の実施形態における入力画像を識別する処理手順の一例を示すフローチャートである。

【図12】第2の実施形態において、局所類似度および統合類似度を算出する処理を模式的に説明するための図である。

【発明を実施するための形態】

【0013】

(第1の実施形態)

以下、図面を参照しながら本発明の第1の実施形態について詳細に説明する。以下の説明では、物体画像から物体認識を行う例として顔認識について説明する。

40

【0014】

<ハードウェア構成>

図1は、本実施形態に係る画像認識装置100のハードウェア構成例を示すブロック図である。

図1に示すように、本実施形態に係る画像認識装置100は、CPU101、ROM102、RAM103、2次記憶装置104、撮像素子105、信号処理装置106、外部出力装置107、および接続バス108を備えている。

【0015】

CPU101は、ROM102やRAM103に格納された制御プログラムを実行する

50

ことにより、画像認識装置100全体の制御を行う。ROM102は不揮発性メモリであり、制御プログラムや各種パラメータデータを記憶する。制御プログラムは、CPU101で実行され、後述する各処理を実行するための手段として、当該装置を機能させる。RAM103は揮発性メモリであり、画像データや制御プログラムおよびその実行結果を一時的に記憶する。2次記憶装置104は、ハードディスクやフラッシュメモリーなどの書き換え可能な2次記憶装置であり、画像情報や制御プログラム、各種設定内容などを記憶する。これらの情報はRAM103に出力され、CPU101がプログラムの実行に利用する。

#### 【0016】

撮像素子105は、CCDセンサやCMOSセンサで構成され、被写体像の光を電気信号に変換する。信号処理装置106は撮像素子105から取得した電気信号を処理し、デジタル信号に変換する信号処理回路である。このデジタル信号は、画像データとして、RAM103または2次記憶装置104へ出力される。外部出力装置107は、CRTやFT液晶などのモニタであり、RAM103または2次記憶装置104から取得した画像データや制御プログラムの実行結果等を表示する。接続バス108は、これらの各構成を接続して相互にデータの入出力を行う。

#### 【0017】

なお、本実施形態では、後述する処理を、CPU101を用いてソフトウェアで実現することとするが、その処理の一部または全部をハードウェアで実現するようにしても構わない。ハードウェアとして専用回路(ASIC)やプロセッサ(リコンフィギュラブルプロセッサ、DSP、CPU)などを用いることができる。また、本実施形態の画像認識装置100は、撮像素子105や信号処理装置106を備えていないパーソナルコンピュータ(PC)を用いて実現してもよいし、画像認識専用装置として実現するようにしても構わない。また、ネットワークまたは各種記憶媒体を介して取得したソフトウェア(プログラム)をパーソナルコンピュータ等の処理装置(CPU、プロセッサ)にて実行してもよい。

#### 【0018】

##### <機能構成>

図2は、本実施形態における画像認識装置100の制御プログラムによる機能構成例を示すブロック図である。

図2に示すように、本実施形態に係る画像認識装置100は、顔画像取得部201、顔特徴点抽出部202、局所画像取得部203、第一特徴抽出部204、および第一特徴抽出パラメータ保持部205を備えている。さらに、第二特徴抽出部206、第二特徴抽出パラメータ保持部207、特徴量射影部208、辞書データ登録部209、局所類似度算出部210、統合類似度算出部211、および認識結果出力部212を備えている。

#### 【0019】

顔画像取得部201は、RAM103または2次記憶装置104から画像データを取得し、その画像中から切り出した顔画像を、顔特徴点抽出部202および局所画像取得部203へ出力する。顔特徴点抽出部202は、顔画像取得部201から取得した顔画像から、顔の器官位置を抽出し、局所画像取得部203へ出力する。局所画像取得部203は、顔画像取得部201から取得した顔画像と顔特徴点抽出部から取得した顔の器官位置の情報とを用いて、所定のサイズの局所領域を設定し、第一特徴抽出部204および第二特徴抽出部206へ、局所画像を出力する。

#### 【0020】

第一特徴抽出部204は、局所画像取得部203から取得した局所画像と、第一特徴抽出パラメータ保持部205が保持する所定の特徴抽出パラメータとを用いて第一特徴量を抽出する。第二特徴抽出部206は、局所画像取得部203から取得した局所画像と、第二特徴抽出パラメータ保持部207が保持する所定の特徴抽出パラメータとを用いて、第二特徴量を取得する。

#### 【0021】

特徴量射影部 208 は、第一特徴抽出部 204 または第二特徴抽出部 206 からそれぞれ第一特徴量または第二特徴量を取得し、特徴量の射影処理を行い、辞書データ登録部 209 または局所類似度算出部 210 に出力する。以下、第一特徴量を射影したベクトルを第一射影特徴量と呼び、第二特徴量を射影したベクトルを第二射影特徴量と呼ぶ。辞書データ登録部 209 は、特徴量射影部 208 から取得した第一および第二射影特徴量を辞書として登録する。辞書データ登録部 209 は、登録した辞書（以下、登録辞書）を局所類似度算出部 210 に出力する。

#### 【0022】

局所類似度算出部 210 は、取得した第一射影特徴量と、辞書データ登録部 209 から取得した登録辞書とを比較し、局所画像毎の局所類似度を算出し、統合類似度算出部 211 へ出力する。統合類似度算出部 211 は、局所類似度算出部 210 から取得した局所類似度を人物 ID 毎の類似度に統合した統合類似度を算出し、認識結果出力部 212 へ出力する。認識結果出力部 212 は、取得した人物 ID 毎の統合類似度から、最終的な識別結果を算出し、RAM 103 または外部出力装置 107 に出力する。

10

#### 【0023】

<全体の処理>

本実施形態に係る画像認識装置 100 は、辞書データ登録処理と識別処理とを行う。辞書データ登録処理では、識別したい人物の顔画像を登録する。一方、識別処理では、入力画像と辞書データ登録処理で登録された辞書とを用いて識別処理を行う。以下、それぞれについて説明する。

20

#### 【0024】

<辞書データ登録処理>

図 3 は、本実施形態における辞書データを登録する処理手順の一例を示すフローチャートである。

始めに、顔画像取得部 201 は、RAM 103 または 2 次記憶装置 104 から画像データを取得する (S300)。そして、取得した画像データから人物の顔を検出する処理を行う (S301)。画像中から人物の顔を検出する方法については、公知の技術を用いればよい。例えば、非特許文献 1 に記載されているような技術を用いることができる。

#### 【0025】

続いて、顔画像取得部 201 は、取得した画像データから顔が検出されたか否かを判定する (S302)。この判定の結果、顔が 1 つも検出されなかった場合は、処理を終了する。一方、顔が検出された場合は、顔画像取得部 201 は、検出された顔の位置情報を元に顔領域を切り出した顔画像データを取得する。なお、顔領域を切り出す際に、画像データの水平方向の軸に対して顔の両目をつないだ直線が平行になるよう画像を回転させる処理を施し、画像の面内回転を補正しておく。こうすることにより、以後の処理で、顔の面内回転による余計な影響を除去することができる。さらに、両目をつなぐ線分が顔画像の横幅に占める割合が一定になるよう拡大・縮小を行う顔サイズ正規化処理を施すことにより、どのような顔でも一定の目幅で以後の処理を実行することができる。そして、顔特徴点抽出部 202 は、顔画像データから器官の位置を示す顔特徴点を抽出する (S303)。抽出する顔特徴点としては、両目や口といった顔の個人差が出やすい器官を選択する。顔特徴点を抽出する方法については、公知の技術を用いればよい。例えば、特許文献 3 に記載されているような技術を用いることができる。

30

40

#### 【0026】

続いて、抽出した顔特徴点を基に第一特徴抽出処理を行う (S304)。この処理では、顔特徴点を基準に定めた局所領域毎に特徴量を抽出するため、局所領域の数だけ特徴量が抽出される。この処理の詳細については後述する。

#### 【0027】

続いて、特徴量射影部 208 は、第一特徴抽出部 204 において局所領域毎に抽出されたすべての第一特徴量に対して射影処理を行う (S305)。射影時に使用する射影行列は、学習データ（大量の顔画像データ）から事前に計算しておく。計算には主成分分析（

50

Principal Component Analysis、PCA)を用いる。以下、射影にPCAを用いた場合の具体的な処理方法について説明する。まず、入力の特徴量ベクトル $X$ を以下の式(1)により定義する。

【0028】

【数1】

$$X = [x_1 \ x_2 \ x_3 \ \cdots \ x_m]^T \quad \cdots(1)$$

【0029】

ここで $T$ は転置を表し、 $x$ はLBP特徴量の構成要素成分を表し、 $m$ は特徴量の次元数を表す。次に、あらかじめ学習しておいた固有空間への射影行列 $V$ を以下の式(2)により表す。

【0030】

【数2】

$$V = \begin{bmatrix} V_{11} & \cdots & V_{1m} \\ \vdots & \ddots & \vdots \\ V_{m1} & \cdots & V_{mm} \end{bmatrix} \quad \cdots(2)$$

10

20

【0031】

さらに、次元削減を行うために行列 $V$ の $n$ 行以降を削除した行列 $V'$ を以下の式(3)により表す。

【0032】

【数3】

$$V' = \begin{bmatrix} V_{11} & \cdots & V_{1m} \\ \vdots & \ddots & \vdots \\ V_{n1} & \cdots & V_{nm} \end{bmatrix} \quad \cdots(3)$$

30

【0033】

以上のように式(2)および式(3)に示した行列を射影処理に用いる。ただし $n < m$ とする。また、射影後の特徴量 $P$ は、以下の式(4)により算出される。

【0034】

【数4】

$$P = V' \cdot (X - A) = [p_1, p_2, p_3, \cdots, p_n]^T \quad \cdots(4)$$

【0035】

ここで、 $A$ は学習データの平均ベクトルである。

【0036】

主成分分析は、特徴量群の分布を解析し、分散最大基準で基底を計算する手法であり、特徴量を元の特徴空間よりも情報的価値(分散最大基準での)が高い空間で表現することができる。主成分分析により得られた固有空間に特徴量を射影することにより、学習に用いた画像セットをよりよく表現できるだけでなく、情報的価値の低い次元を切り落とすことにより特徴量の次元数を削減する効果も期待できる。なお、削減する次元の量を調整するには式(3)の変数 $n$ を調整すればよい。ここでは削減する次元数を実験的に定めるが、累積寄与率を基に定めてもよい。固有空間の計算方法としては、この他にもLocality Preserving Projection(LPP)を用いてもよい。また

40

50

、それ以外に、教師付きの固有空間学習方法である Local Fisher Discriminant Analysis (LFDA)、などの公知の例を用いてもよい。

【0037】

続いて、顔特徴点を基に第二特徴量抽出処理を行う (S306)。この処理では、S304における第一特徴抽出処理と同様に、局所領域毎に特徴量が抽出されるため、局所領域の数だけ特徴量が抽出される。この処理の詳細については後述する。続いて、特徴量射影部208は、第二特徴量に対して射影処理を行う (S307)。具体的な処理の内容は、S305の第一特徴量に対する射影処理と同様であるため、ここでの説明は省略する。

【0038】

続いて、辞書データ登録部209は、第一射影特徴量および第二射影特徴量を辞書に登録する (S308)。このとき、第一射影特徴量および第二射影特徴量は同一人物 (同一の画像) から抽出した特徴量であるため、同一の人物IDを割り振り、辞書に登録する。

【0039】

次に、まだ処理していない顔が存在するか否かを判定する (S309)。この判定の結果、まだ処理していない顔が存在する場合は、次の顔について、S303に戻って処理を繰り返す。一方、全ての顔について処理が完了した場合は、処理を終了する。

【0040】

<第一特徴抽出処理>

次に、S304の第一特徴抽出処理について説明する。図4は、図3のS304における第一特徴抽出処理の詳細な手順の一例を示すフローチャートである。

始めに、局所画像取得部203は、顔画像取得部201から顔画像データを取得する (S400)。続いて局所画像取得部203は、顔特徴点抽出部202から顔特徴点を取得する (S401)。続いて、第一特徴抽出部204は、第一特徴抽出パラメタ保持部205から第一特徴抽出パラメタを取得する (S402)。

【0041】

続いて、局所画像取得部203は、取得した顔画像および顔特徴点を用いて局所領域を設定し、局所画像を取得する (S403)。局所領域を設定する際には、顔特徴点を基準点として、そこから所定画素数の距離にある点を中心点とする所定サイズの矩形領域を局所領域として設定する。このとき、一つの基準点に対して複数の局所領域を設定してよい。また、どの局所領域を局所画像として取得するかについては、事前に学習で定めるようにしてもよい。

【0042】

続いて、第一特徴抽出部204は、局所画像と第一特徴抽出パラメタとを用いて第一特徴量を抽出する (S404)。このとき、第一特徴抽出部204は、抽出した特徴量を局所画像毎に固有のラベルと対応付けて保持する。本実施形態では、特徴量としてLBP (Local Binary Pattern) 特徴量を抽出するが、その他公知の例を用いてもよい。LBP特徴量は、隣り合う画素値の大小関係をコード化した特徴量である。画素間の大小関係のみに着目するため、画像の全体的な照明変化に頑健な特徴量として知られている。そのため、様々な照明条件が想定される顔画像の識別等に適している。LBP特徴量の抽出方法は以下の通りである。

【0043】

図5は、LBP特徴量を抽出するための注目画素とその周辺画素との関係を表わした模式図である。ここで、注目画素  $g_c$  の画素値を  $I_c$  とし、周辺の  $p$  番目の画素  $g_p$  の画素値を  $I_p$  とする。図5に示すように、注目画素  $g_c$  の真上が  $g_1$ 、右上が  $g_2$  となり、時計周りに  $g_8$  まで続いている。このとき、注目画素  $g_c$  におけるLBPコード値  $LBP_c$  は、第一特徴抽出パラメタである以下の式 (5) により算出される。

【0044】

10

20

30

40

【数 5】

$$LBP_c = \sum_{p=0}^{P-1} s(I_p - I_c)2^p, \quad s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \dots (5)$$

【0045】

ここで、PはLBPコードのビット数を示しており、図5に示す例ではP = 8である。式(5)では、注目画素とその周辺8画素との差分が0以上である場合は1(ビットを立てる)を当てはめ、0より小さい場合は0(ビットを立てない)を当てはめる。このように数値を連結したコードを8ビットの数値として扱う。このLBPコード値LBP<sub>c</sub>をLBP特徴量と呼び、本実施形態における第一特徴量とする。

10

【0046】

次に、全ての局所画像の特徴抽出を完了したか否かを判定する(S405)。この判定の結果、特徴抽出が完了した場合は、局所画像取得部203は、使用した局所画像を保持し(S406)、第一特徴抽出処理を終了する。なお、保持した局所画像は第二特徴抽出処理で使用するため、この処理では保持したままにしておく。一方、まだ特徴抽出していない局所画像がある場合は、S403に戻り、残りの局所画像に対して処理を繰り返す。

【0047】

&lt;第二特徴抽出処理&gt;

20

次に、第二特徴抽出処理について説明する。図6は、図3のS306における第二特徴抽出処理の詳細な手順の一例を示すフローチャートである。

始めに、第二特徴抽出部206は、第二特徴抽出パラメタ保持部207から第二特徴抽出パラメタを取得する(S601)。本実施形態では、この第二特徴抽出パラメタを事前に3種類用意しておく。第二特徴抽出パラメタの数は、多いほど精度向上に寄与する傾向があるが、多くすると特徴抽出時の計算量が多くなってしまうため、目的に応じて第二特徴抽出パラメタの数を設定する。ここで取得する第二特徴抽出パラメタは、第一特徴抽出パラメタとは異なるパラメタである。具体的には、式(5)に示したLBP特徴量の算出式と比べて、以下の式(6)に示すように注目画素値をI<sub>c</sub>と周辺画素値I<sub>p</sub>との比較にオフセット値を導入する。

30

【0048】

【数 6】

$$LBP_c(\theta) = \sum_{p=0}^{P-1} s(I_p - I_c - \theta)2^p, \quad s(x) = \begin{cases} 1, & \text{if } x \geq 0 \\ 0, & \text{otherwise} \end{cases} \dots (6)$$

【0049】

このようにオフセット値を導入することによって、顔の個人差はそのままにし、照明方向が異なる顔画像の特徴量を疑似的に生成することができる。以下、そのオフセット値の詳細について説明する。

40

【0050】

顔画像において、顔にあたる照明の方向が変わると、肌領域では画素値の変化の勾配が小さく起伏が乏しいため、画素値の大小関係が照明方向に応じて変化する。そのため、LBP特徴量のように画素値の大小関係を利用する特徴量のパターンも変化してしまう。一方、顔の器官領域における器官のエッジは、画素値の変化の勾配が大きく画素値の大小関係の差が大きいため、照明方向が変わっても画素値の大小関係が逆転することはなく、特徴量のパターンは大きく変化しない。

【0051】

つまり、画素値の大小関係を利用する特徴量で照明方向の異なる顔を表現するには、元の特徴量と比較して、肌領域では照明方向による変化があり、器官領域ではその変化が少

50

ない特徴量を生成しなければならない。そこで、式(6)において、オフセット値を、肌領域では特徴量のパターンが変化し、器官領域ではあまり変化しないような値に設定することによって、あたかも照明方向を変えたかのような特徴量を抽出することができる。オフセット値をどのような値に設定するかについては、表現したい照明条件によるため実験的に定めるのがよい。

#### 【0052】

一つの例として、入力画像と登録辞書とで照明方向が大きく異なる場合(順光の顔と逆光の顔との比較など)では、オフセット値を負の値にすることによって、大小関係が逆転したかのような特徴量を抽出することができる。また、本実施形態ではオフセット値を含む式を第二特徴抽出パラメタとし、第一特徴量を変更した特徴量を第二特徴量としたが、LBP特徴量のその他のパラメタを変更して第二特徴量としてもよい。例えば、式(5)のpの順番、つまりコード化の順番などを変えてもよい。

10

#### 【0053】

続いて、第二特徴抽出部206は、局所画像取得部203から局所画像を取得する(S602)。前述の第一特徴抽出処理のS406において、局所画像取得部203は局所画像を保持しているので、保持している局所画像をそのまま取得することになる。

#### 【0054】

続いて、第二特徴抽出部206は、前述した第二特徴抽出パラメタと局所画像とを用いて第二特徴量を抽出する(S603)。このとき、第一特徴抽出処理と同様に、抽出した特徴量を局所画像毎に固有のラベルと対応付けて保持するようにする。

20

#### 【0055】

次に、全ての局所画像について処理を完了したか否かを判定する(S604)。この判定の結果、処理していない局所画像が存在する場合はS602に戻り、全ての局所画像について処理を完了した場合は、次のS605に進む。最後に、全ての第二特徴抽出パラメタで第二特徴量の抽出が完了したか否かを判定する(S605)。この判定の結果、全ての第二特徴抽出パラメタで終了した場合は、第二特徴抽出処理を終了する。一方、まだ処理が完了していない第二特徴抽出パラメタがある場合は、S601に戻り、残りの特徴抽出パラメタに対して処理を継続する。

#### 【0056】

< 識別処理フロー >

30

次に、識別処理について説明する。図7は、本実施形態における入力画像を識別する処理手順の一例を示すフローチャートである。なお、図7のS700~S705は図3の辞書データ登録処理におけるS300~S305と同様であるため、説明は省略する。以下、S706の局所類似度算出処理から説明する。

#### 【0057】

S705の処理が終了すると、局所類似度算出部210は、S705で算出した第一射影特徴量と辞書(第一射影特徴量および第二射影特徴量)とを比較することにより局所類似度を算出する(S706)。この処理の詳細については後述する。続いて、統合類似度算出部211は、取得した局所類似度から辞書に登録されている人物ID毎の類似度を算出する(S707)。この処理の詳細についても後述する。

40

#### 【0058】

次に、まだ処理していない顔画像データが存在するか否かを判定する(S708)。この判定の結果、まだ処理していない顔画像データが存在する場合は、S703に戻り、次の顔画像データについて処理を継続する。一方、全ての顔画像データの処理を完了した場合は、認識結果出力部212は、顔認識結果をRAM103または外部出力装置107に出力する(S709)。この処理では、人物ID毎の類似度のうち最も高いものを選択し、その類似度が所定値よりも大きい場合は、その類似度に対応する人物IDを顔認識結果として出力する。なお、類似度が所定値以下の場合は、辞書中の誰でもないという顔認識結果を出力する。また、S702の判定の結果、入力画像から顔が1つも検出されなかった場合も、その旨の結果を出力する。

50

## 【 0 0 5 9 】

< 局所類似度算出処理 >

次に、局所類似度算出処理について説明する。図 8 は、図 7 の S 7 0 6 における局所類似度を算出する詳細な処理手順の一例を示すフローチャートである。なお、図 8 に示す各処理は局所類似度算出部 2 1 0 により行われる。

## 【 0 0 6 0 】

始めに、局所類似度算出部 2 1 0 は、特徴量射影部 2 0 8 から、図 7 の S 7 0 5 により得られた特徴量（第一射影特徴量）を取得する（S 8 0 0）。続いて、辞書データ登録部 2 0 9 から登録辞書を取得する（S 8 0 1）。

## 【 0 0 6 1 】

続いて、取得した特徴量と登録辞書とから局所類似度を算出する（S 8 0 2）。前述の通り、射影特徴量は局所画像の数だけ存在するので、局所画像毎に局所類似度を算出することとなる。また、特徴量には局所画像毎に固有のラベルが対応付けられているので、同一のラベルを持つ特徴量同士で局所類似度を算出する。局所類似度を算出する方法としては、特徴ベクトル間の内積に基づくコサイン類似度を用いる。すなわち、入力画像の特徴量を  $V_I$ 、登録辞書の特徴量を  $V_R$  とし、 $V_I$  と  $V_R$  とがなす角を  $\alpha$  とすると、コサイン類似度  $S$  は、以下の式（7）により算出される。

## 【 0 0 6 2 】

【数 7】

$$S = \cos\alpha = \frac{V_I \cdot V_R}{|V_I||V_R|} \quad \dots(7)$$

## 【 0 0 6 3 】

ここで、 $V_I \cdot V_R$  は  $V_I$  と  $V_R$  との内積、 $|V_I|$  および  $|V_R|$  はそれぞれのベクトルのノルムを表す。本実施形態ではコサイン類似度を用いているため、局所類似度の範囲が -1 から +1 の間となり、異なる局所領域で算出された類似度を統合する際に都合がよい。

## 【 0 0 6 4 】

次に、全ての局所画像について局所類似度を算出したか否かを判定する（S 8 0 3）。この判定の結果、全ての局所画像について局所類似度を算出した場合は次の処理に進み、そうでない場合は S 8 0 2 に戻り、残りの局所画像について局所類似度を算出する。次に、全登録辞書について局所類似度の算出が完了したか否かを判定する（S 8 0 4）。この判定の結果、全登録辞書について局所類似度を算出した場合は、算出した局所類似度を出し（S 8 0 5）、この類似度算出処理を終了する。この時点で、登録されている画像毎に複数の局所類似度が算出されている状態となっている。一方、局所類似度の算出が完了していない登録辞書がある場合は、S 8 0 1 に戻り、残りの登録辞書に対して処理を継続する。

## 【 0 0 6 5 】

< 統合類似度算出処理 >

次に、統合類似度算出処理について説明する。図 9 は、図 7 の S 7 0 7 における統合類似度を算出する詳細な処理手順の一例を示すフローチャートである。なお、図 9 に示す各処理は、統合類似度算出部 2 1 1 により行われる。

## 【 0 0 6 6 】

始めに、統合類似度算出部 2 1 1 は、局所類似度算出部 2 1 0 から局所類似度を取得する（S 9 0 0）。続いて、登録辞書の顔画像毎に存在する局所類似度を、人物 ID 毎の局所類似度に統合する（S 9 0 1）。図 10 は、人物 ID 毎の局所類似度を算出する際のある局所領域についての処理例を示した模式図である。図 10 に示す例では、人物 ID = 1 の人物の局所領域 A について、人物 ID 毎の局所類似度を算出する場合について示している。

## 【 0 0 6 7 】

図 1 0 に示すように、人物 I D = 1 の局所領域 A における局所類似度の中から、最も高い類似度を選択し、その類似度をその人物 I D の局所領域 A の局所類似度とする。なお、図 1 0 に示す登録画像 3 は人物 I D が異なるので、この計算には使用しない。このように全局所領域について同様の計算を施す。

## 【 0 0 6 8 】

一般的に、顔の個人差を示すような複雑なパターンの場合、類似度が偶然にも高くなってしまう可能性は低い。このため、局所類似度の中で最も高い値を選択することにより、他人に比べて統合後の類似度も高くなることが期待できる。また、本実施形態では、第二特徴量を 3 種類抽出して、第一特徴量とは異なる照明方向の特徴量を疑似的に辞書に登録している。これらの情報を用いて人物 I D 毎の局所類似度を算出することにより、照明方向により頑健な類似度を算出することができる。

10

## 【 0 0 6 9 】

続いて、人物 I D 毎に複数ある局所類似度を、人物 I D 毎に一つの統合類似度に統合する ( S 9 0 2 )。統合の方法としては、局所類似度の平均値を用いる。なお、極端に高いまたは低い局所類似度はノイズである可能性があるため、それらを除去した上で平均値をとってもよい。最後に、算出した統合類似度を出力し ( S 9 0 3 )、この統合類似度算出処理を終了する。

## 【 0 0 7 0 】

以上のように本実施形態によれば、辞書に第一特徴量と 3 種類の第二特徴量とを登録して、疑似的に異なる照明方向の辞書を生成するようにしている。これにより、入力画像の照明方向が、登録辞書の第一特徴量とは異なる場合においても、第二特徴量で表現した照明方向の情報を用いることができ、照明方向の変化に頑健な顔認識を実現できる。また、特徴量のパラメタを変更することによって疑似的に辞書を生成するため、画像そのものを疑似的に生成する方法などと比較して、メモリ容量および計算コストを低減することができる。また、本実施形態では、局所画像から特徴量を抽出する方法について述べたが、顔画像全体から特徴量を抽出してもよい。

20

## 【 0 0 7 1 】

( 第 2 の実施形態 )

第 1 の実施形態では、辞書に登録する顔画像のみから第二特徴量を抽出したが、本実施形態では、類似度を算出する入力画像からも第二特徴量を抽出し、入力画像、登録辞書ともに疑似的に特徴量を増す例について説明する。また、第 1 の実施形態では、特徴量として L B P 特徴量を用いる場合について説明したが、他の画素値の大小関係に基づいた特徴量の一例として、Histogram of Oriented Gradient ( H O G ) 特徴量を用いる例について説明する。以下、第 1 の実施形態と同じ処理を行う場合は説明を省略し、処理の異なる部分のみを説明する。また、第 1 の実施形態で説明した図 1 に示したハードウェア構成、および図 2 に示した機能構成については、本実施形態においても同様であるため、説明を省略する。

30

## 【 0 0 7 2 】

< 全体の処理 >

本実施形態における辞書データ登録処理については第 1 の実施形態と同様であるため、説明は省略する。以下、本実施形態に係る識別処理について図 1 1 を参照しながら説明する。

40

## 【 0 0 7 3 】

< 識別処理 >

図 1 1 は、本実施形態における入力画像を識別する処理手順の一例を示すフローチャートである。なお、S 1 1 0 0 ~ S 1 1 0 3 は、それぞれ図 7 の S 7 0 0 ~ S 7 0 3 と同様であるため、説明は省略する。ここで、S 1 1 0 4 の第一特徴量を抽出する際に、H O G 特徴量を抽出する処理について説明する。まず、画像の座標 ( x , y ) の画素 ( 輝度 ) 値を  $I ( x , y )$  とした場合に、輝度勾配は x 方向、y 方向で、以下の式 ( 8 ) により表さ

50

れる。

【 0 0 7 4 】

【 数 8 】

$$\begin{aligned} f_x(x, y) &= I(x + 1, y) - I(x - 1, y) \\ f_y(x, y) &= I(x, y + 1) - I(x, y - 1) \end{aligned} \quad \dots(8)$$

【 0 0 7 5 】

次に、これらの輝度勾配から、勾配強度  $m(x, y)$ 、および勾配方向  $\vartheta(x, y)$  は、それぞれ以下の式 (9) および式 (10) により表される。 10

【 0 0 7 6 】

【 数 9 】

$$m(x, y) = \sqrt{f_x(x, y)^2 + f_y(x, y)^2} \quad \dots(9)$$

$$\vartheta(x, y) = \tan^{-1} \frac{f_y(x, y)}{f_x(x, y)} \quad \dots(10)$$

20

【 0 0 7 7 】

次に、勾配方向の範囲 180 度を 20 度毎に 9 分割したそれぞれの区間をビンとして、勾配方向  $(x, y)$  の頻度ヒストグラムを画像の所定の矩形領域 (セル) 毎に作成する。そして、 $3 \times 3$  の 9 個のセルを 1 ブロックとして、ブロック毎に正規化を行う。正規化は 1 セルずつずらしながら全領域に対して行い、正規化の度にヒストグラムを連結していく。この処理を画像全体にわたって行って得た連結ヒストグラムが HOG 特徴量である。このように、通常は、ブロック毎の正規化を 1 セルずつずらしながら行うが、ずらす量を増やすことにより計算量および特徴量の次元数を軽減してもよい。この HOG 特徴量は、画像中の物体の回転、スケール変化に対しては不変ではないものの、局所的な幾何学的変化や明るさの変化には不変な特徴量として知られている。勾配方向ヒストグラム特徴量についての詳細は非特許文献 2 に開示されている。本実施形態では、局所画像から取得したこの HOG 特徴量を第一特徴量とする。次の S 1 1 0 5 については図 7 の S 7 0 5 と同様であるため、説明は省略する。 30

【 0 0 7 8 】

次に、S 1 1 0 6 の第二特徴量を抽出する際に、HOG 特徴量に関する第二特徴量を抽出する処理について説明する。第二特徴量を算出する際には、前述の式 (8) の勾配計算においてオフセット値  $\rho_1$ 、 $\rho_2$  を導入し、以下の式 (11) を用いる。

【 0 0 7 9 】

【 数 10 】

$$\begin{aligned} f_x(x, y, \rho_1) &= I(x + 1, y) - I(x - 1, y) + \rho_1 \\ f_y(x, y, \rho_2) &= I(x, y + 1) - I(x, y - 1) + \rho_2 \end{aligned} \quad \dots(11)$$

40

【 0 0 8 0 】

これ以外の点については第一特徴量を取得した手順と同様であるため、説明は省略する。輝度勾配の計算をこのように変更することにより画素の大小関係を変化させ、あたかも照明の方向が変化したかのような HOG 特徴量を抽出することができる。次の S 1 1 0 7 50

については図3のS307と同様であるため、説明は省略する。次のS1108およびS1109については図7のS706およびS707と異なる点についてのみ説明する。

【0081】

図12は、図11のS1108およびS1109の処理を模式的に説明するための図である。第1の実施形態で示した図10と比較して、入力画像においても第二特徴量を抽出し、S1108ではその局所類似度も算出する。このため、画像毎の局所類似度の数が飛躍的に増えている。S1109においても、最も高い局所類似度を選択し、最終的な統合類似度を算出する。S1110およびS1111については、それぞれ図7のS708およびS709と同様である。

【0082】

以上のように本実施形態によれば、識別処理においても第二特徴量を抽出するようにしたので、最終的な統合類似度を算出する際に使用できる情報量を増やすことができ、照明方向に対する頑健性のさらなる向上が期待できる。常に第二特徴量を抽出し、飛躍的に増えた局所類似度を算出しなければならないため、第1の実施形態と比較して、計算量は増大する。そこで、この場合にはユーザが許容できる計算量となるよう第二特徴量を抽出するためのパラメタの数を調整し、その分の計算量を抑えることも可能である。

【0083】

(その他の実施形態)

前述した各実施形態では、顔認識について説明したが、一般の物体に対しても、領域によって画素値の変化の勾配が異なる場合には第二特徴量を同様に抽出することができる。このため、顔認識のみならず、一般の物体認識でも適用することができる。

【0084】

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。この場合、そのプログラム、及び該プログラムを記憶したコンピュータ読み取り可能な記憶媒体は本発明を構成することになる。

【符号の説明】

【0085】

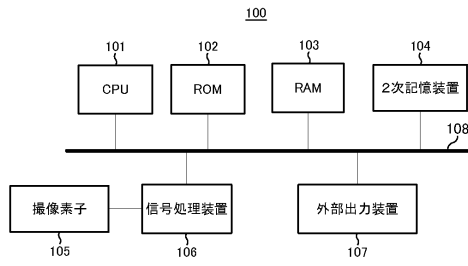
- 201 顔画像取得部
- 204 第一特徴抽出部
- 206 第二特徴抽出部
- 209 辞書データ登録部
- 210 局所類似度算出部
- 211 統合類似度算出部

10

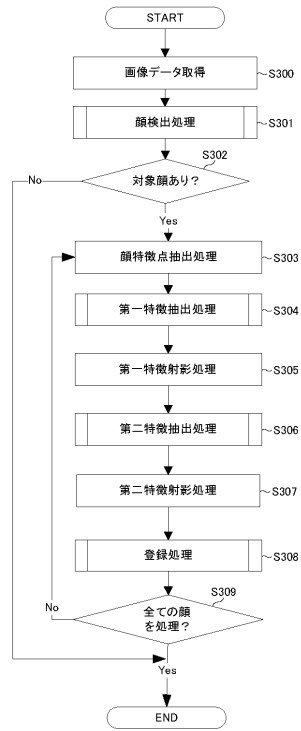
20

30

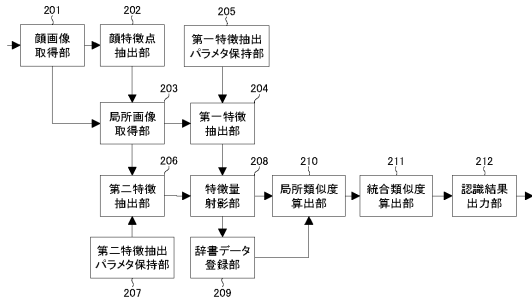
【図1】



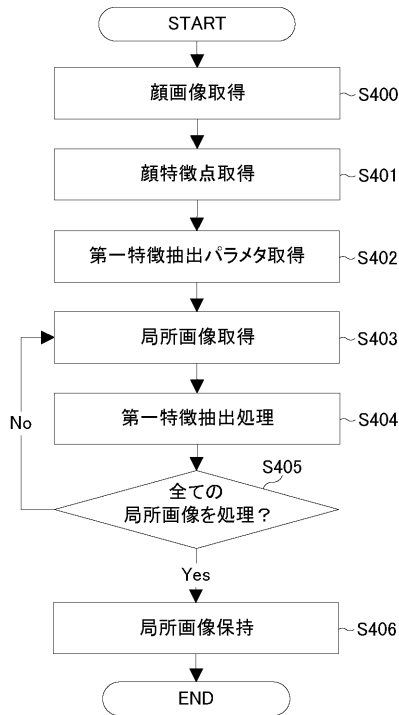
【図3】



【図2】



【図4】

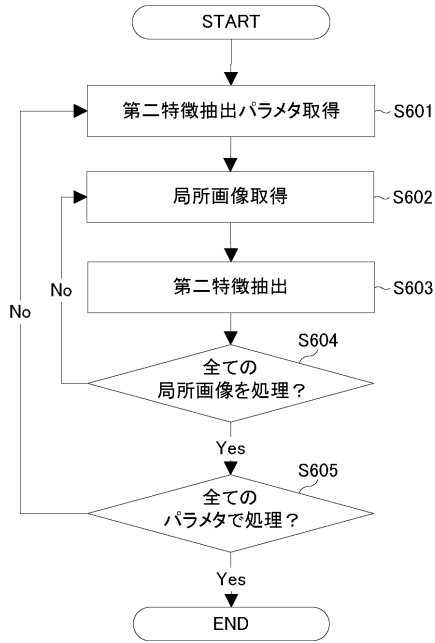


【図5】

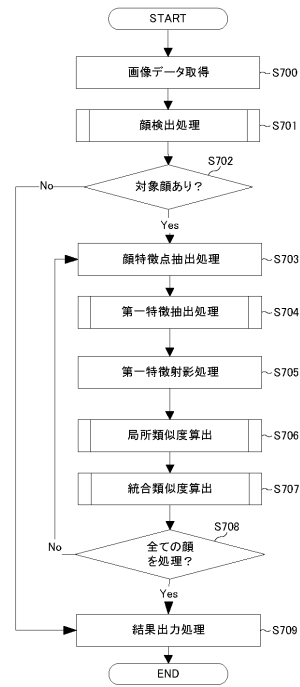
注目画素:  $g_c$

		$g_8$	$g_1$	$g_2$			
		$g_7$	$g_c$	$g_3$			
		$g_6$	$g_5$	$g_4$			

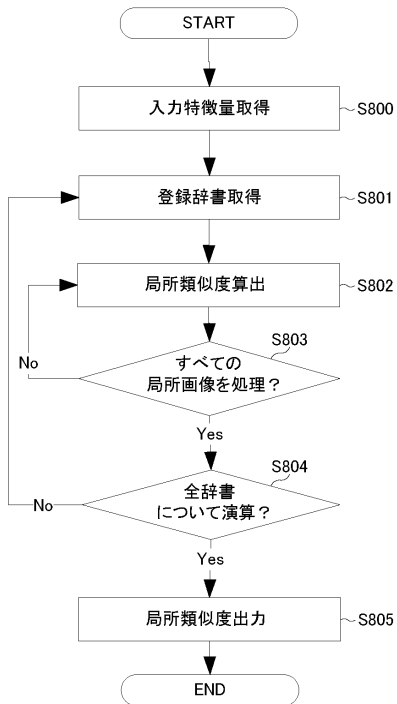
【図6】



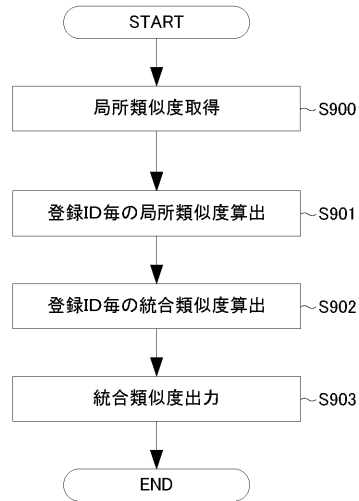
【図7】



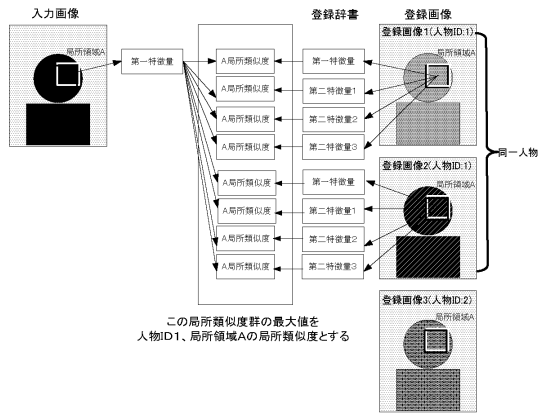
【図8】



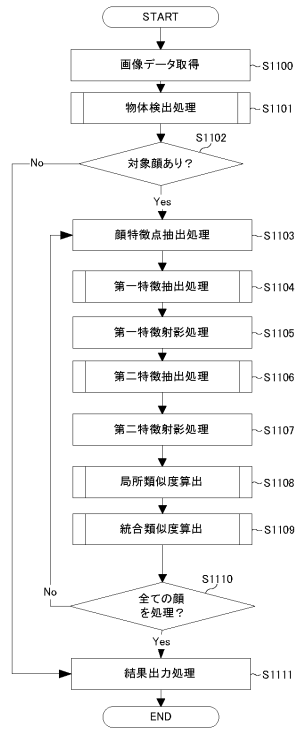
【図9】



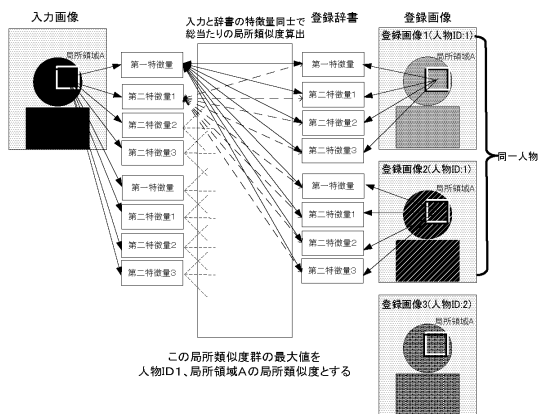
【図10】



【図11】



【図12】



---

フロントページの続き

- (72)発明者 金田 雄司  
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 中野 俊亮  
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 鈴木 崇士  
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 真木 健彦

- (56)参考文献 特開2011-008631(JP,A)  
特開2008-217706(JP,A)  
特開2013-030104(JP,A)  
特開2002-024830(JP,A)  
松島 千佳, 物体検出のためのRelational HOG特徴量とワイルドカードを用いた  
バイナリーのマスクング, 電子情報通信学会論文誌 (J94-D) 第8号, 日本, 社団法人  
電子情報通信学会, 2011年 8月 1日, Vol.J94-D No.8, P.1172-1182, ISSN 1880-4535

- (58)調査した分野(Int.Cl., DB名)  
G06T 7/00 - 7/90