

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第5737909号
(P5737909)

(45) 発行日 平成27年6月17日 (2015. 6. 17)

(24) 登録日 平成27年5月1日 (2015. 5. 1)

(51) Int. Cl.

F I

HO 4 N 5/232 (2006. 01)
 HO 4 N 5/225 (2006. 01)
 HO 4 N 7/18 (2006. 01)
 GO 6 T 1/00 (2006. 01)

HO 4 N 5/232 C
 HO 4 N 5/225 C
 HO 4 N 7/18 K
 GO 6 T 1/00 3 4 O A

請求項の数 3 (全 15 頁)

(21) 出願番号 特願2010-250207 (P2010-250207)
 (22) 出願日 平成22年11月8日 (2010. 11. 8)
 (65) 公開番号 特開2012-104964 (P2012-104964A)
 (43) 公開日 平成24年5月31日 (2012. 5. 31)
 審査請求日 平成25年11月8日 (2013. 11. 8)

(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
 (74) 代理人 100090273
 弁理士 國分 孝悦
 (72) 発明者 東條 洋
 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 吉川 康男

最終頁に続く

(54) 【発明の名称】 画像処理装置、画像処理方法及びプログラム

(57) 【特許請求の範囲】

【請求項 1】

同一の物体の異なる方向に対応した複数の辞書のデータを格納する辞書記憶手段と、
 前記複数の辞書のそれぞれに対してフレーム画像における適用領域を、少なくとも2つの辞書の適用領域に重複する領域を含んで設定する設定手段と、

前記複数の辞書のそれぞれを用いて当該辞書に対して設定された適用領域において前記物体を認識する認識手段と

を備え、

前記認識手段は、少なくとも2つの辞書で重複する適用領域に対して、連続する複数のフレーム画像においてフレーム画像ごとに使用する辞書を切り替え、そのフレーム画像ごとの認識結果を統合することを特徴とする画像処理装置。

【請求項 2】

同一の物体の異なる方向に対応した複数の辞書のそれぞれに対してフレーム画像における適用領域を、少なくとも2つの辞書の適用領域に重複する領域を含んで設定する設定ステップと、

前記複数の辞書のそれぞれを用いて当該辞書に対して設定された適用領域において前記物体を認識する認識ステップと

を備え、

前記認識ステップでは、少なくとも2つの辞書で重複する適用領域に対して、連続する複数のフレーム画像においてフレーム画像ごとに使用する辞書を切り替え、そのフレーム

10

20

画像ごとの認識結果を統合することを特徴とする画像処理方法。

【請求項 3】

請求項 2 に記載の画像処理方法の各ステップをコンピュータに実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、カメラから取得されるフレーム画像等から物体を認識する技術に関するものである。

【背景技術】

10

【0002】

従来より、店舗等の入り口や通路を通行する人をカメラで撮影し、撮影した映像より人物の顔の位置を検出して、通過した人数を計測したり、予め登録されている人物の顔であるかを認識したりする技術が開示されている。

【0003】

このような所定領域における通行人をカメラ映像から自動的にカウントする技術としては、例えば、下記の特許文献 1 がある。この特許文献 1 では、通路の上方から真下に向けてカメラを設置する。カメラ上方から見た人物の頭の形状が円であることから、カメラ映像から円形の物体を抽出することで人物を検知、カウントするようにしている。

【0004】

20

一方、近年、画像から顔を検出する技術の実用化が進んでいる。このような技術を利用して、後述する図 1 に示すように通路の前方にカメラを設置して、カメラ映像から顔を検出することで人物をカウントすることも可能である。

【0005】

ここで、カメラが広い範囲を撮影するほど、カメラと人物の位置関係で、顔の向きが異なって撮影されることになる。そして、顔の向きが変化すると、顔の特徴が異なってくる。従って、認識が困難になる。

【0006】

この課題に対応するために、下記の特許文献 2 では、顔の向きに応じて認識辞書を用意し、フレーム画像を複数の領域に分け、それぞれの領域で適用する認識辞書を変更していた。ここで、例として図 11 の場合を挙げる。

30

【0007】

図 11 は、カメラからの距離に応じて撮影される顔の向きが異なることを説明する模式図である。

図 11 において、1101 は通路の天井であり、1102 は床である。1103 がカメラであり、天井 1101 に設置され、通路を斜め上より撮影している。1104 のようにカメラ 1103 から遠い位置に人物がいた場合、撮影された顔の垂直方向の向きは小さい角度になるが、1105 のようにカメラ 1103 から近い位置に人物がいた場合、顔の垂直方向の向きは大きな角度になる。

【0008】

40

図 12 は、顔の特徴が見え方によって変化することを説明する模式図である。

1201 が、人物 1104 のようにカメラ 1103 から遠い位置に人物がいる時であり、1202 が、人物 1105 のようにカメラ 1103 に近い位置に人物がいるときである。この図 12 からわかるように、顔の特徴が見え方によって変化する。

【0009】

図 13 は、従来技術による課題を説明するための模式図である。

そこで、図 13 のように、フレーム画像 1301 を、1302 と 1303 の 2 つの領域に分け、1302 の領域に対しては角度が小さい顔を認識する辞書を用い、1303 の領域に対しては角度が大きな顔を認識する辞書を用いるようにしている。即ち、前記例において、1304 が、人物 1104 の位置に人物がいるときに認識される顔のフレーム画像

50

内の位置であり、認識辞書は角度が小さい顔用の辞書が用いられる。また、1305が、人物1105の位置に人物がいるときに認識される顔のフレーム画像内の位置であり、認識辞書は角度が大きい顔用の辞書が用いられる。

【先行技術文献】

【特許文献】

【0010】

【特許文献1】特開平4-199487号公報

【特許文献2】特開2007-25767号公報

【発明の概要】

【発明が解決しようとする課題】

10

【0011】

しかしながら、特許文献2では、各認識辞書を適用する領域の境界にあたる位置に顔があると認識が難しくなる。前述の例では、図13の1306の位置に顔があるような場合である。このような位置での顔の角度は、中間的な角度となる。つまり、角度が小さい顔用の辞書からも、角度が大きい顔用の辞書からも、その特徴の変動が最も大きくなる。従って、どちらの認識辞書を用いても認識精度が低くなってしまう傾向にあるため、認識が難しくなるのである。

【0012】

本発明は、このような問題点に鑑みてなされたものであり、認識する物体の向きが画面内で変化する場合であっても、精度良く物体を認識することができる仕組みを提供することを目的とする。

20

【課題を解決するための手段】

【0013】

前述した目的を達成するために、本発明は、同一の物体の異なる方向に対応した複数の辞書のデータを格納する辞書記憶手段と、前記複数の辞書のそれぞれに対してフレーム画像における適用領域を、少なくとも2つの辞書の適用領域に重複する領域を含んで設定する設定手段と、前記複数の辞書のそれぞれを用いて当該辞書に対して設定された適用領域において前記物体を認識する認識手段とを備え、前記認識手段は、少なくとも2つの辞書で重複する適用領域に対して、連続する複数のフレーム画像においてフレーム画像ごとに使用する辞書を切り替え、そのフレーム画像ごとの認識結果を統合することを特徴とする画像処理装置等、を提供する。

30

【発明の効果】

【0014】

本発明によれば、認識する物体の向きが画面内で変化する場合であっても、精度良く物体を認識することができる。

【図面の簡単な説明】

【0015】

【図1】本発明の実施形態に係る画像処理装置の設置例を示す模式図である。

【図2】本発明の実施形態に係る画像処理装置のハードウェア構成の一例を示すブロック図である。

40

【図3】本発明の実施形態に係る画像処理装置の機能構成の一例を示すブロック図である。

【図4】本発明の実施形態に係る画像処理装置による画像処理方法の処理手順の一例を示すフローチャートである。

【図5】本発明の実施形態を示し、顔の角度を説明する模式図である。

【図6】本発明の実施形態を示し、認識辞書A及びBの適用できる顔の角度と認識精度との関係を示した模式図である。

【図7】本発明の実施形態を示し、認識辞書A及びBの適用できる領域を説明する模式図である。

【図8】画像から顔パターンの探索を行う方法を説明する図である。

50

【図 9】本発明の実施形態を示し、辞書と照合領域を切り替える動作を説明するフローチャートである。

【図 10】本発明の実施形態を示し、軌跡の生成とカウントの一例を示した模式図である。

【図 11】カメラからの距離に応じて撮影される顔の向きが異なることを説明する模式図である。

【図 12】顔の特徴が見え方によって変化することを説明する模式図である。

【図 13】従来技術による課題を説明するための模式図である。

【発明を実施するための形態】

【0016】

10

以下に、図面を参照しながら、本発明を実施するための形態（実施形態）について説明する。

【0017】

なお、以下に挙げる実施形態は、通路を通過する人数を計測する例で説明する。

図 1 は、本発明の実施形態に係る画像処理装置の設置例を示す模式図である。

101 は、通路の天井であり、102 は通路の床である。103 は通路を通行している人物である。104 は撮像部（カメラ）であり、人物 103 を斜め上から撮影できるように、天井 101 に設置してある。105 は LAN ケーブルであり、撮像部 104 で撮像される映像を送信する。106 は、映像を解析し、計数する画像処理装置となる PC である。

20

【0018】

図 2 は、本発明の実施形態に係る画像処理装置 106 のハードウェア構成の一例を示すブロック図である。

図 2 において、201 は CPU であり、本実施形態の画像処理装置 106 における各種制御を実行する。202 は ROM であり、本画像処理装置 106 の立ち上げ時に実行されるブートプログラムや各種データを格納する。203 は RAM であり、CPU 201 が処理するための制御プログラムを格納するとともに、CPU 201 が各種制御を実行する際の作業領域を提供する。204 はキーボード、205 はマウスであり、ユーザによる各種入力操作環境を提供する。

【0019】

30

206 は外部記憶装置であり、ハードディスクやフレキシブルディスク、光ディスク、磁気ディスク、光磁気ディスク、磁気テープ等で構成される。ただし、外部記憶装置 206 は、制御プログラムや各種データを全て ROM 202 に持つようにすれば、必ずしも必要な構成要素ではない。本実施形態においては、本発明の処理に係る制御プログラムは、ROM 202（或いは外部記憶装置 206）に格納されているものとする。

【0020】

207 は表示装置であり、ディスプレイなどで構成され、結果等をユーザに対して表示する。208 はネットワークインターフェース（NIC）であり、ネットワーク上の撮像部 104 と LAN ケーブル 105 を介した通信を可能とする。209 はビデオインターフェース（ビデオ I/F）であり、撮像部 104 と同軸ケーブルを介したフレーム画像の取り込みを可能とする。また、210 は上記の各構成を接続するバスである。

40

【0021】

図 3 は、本発明の実施形態に係る画像処理装置 106 の機能構成の一例を示すブロック図である。

10 は、撮像レンズ、及び、CCD、CMOS などの撮像センサからなる撮像手段である。この撮像手段 10 は、図 1 の撮像部 104 に相当するものである。

【0022】

30 は、画像取得手段であり、撮像手段 10 で撮像した画像データを所定時間間隔で取得し、時間的に連続した複数フレーム単位で出力する。フレーム画像は、撮像部 104 から、LAN ケーブル 105 を介して http プロトコルのパケットデータとして送られ、

50

画像処理装置 106 上のネットワークインターフェース 208 を介して取得する。或いは、105 を同軸ケーブルで構成し、画像処理装置 106 上のビデオインターフェース 209 で取得するようにしてもよい。

【0023】

40 は、物体認識手段であり、画像取得手段 30 で取得した画像データに所望の物体が映っているかどうかを認識処理する。具体的に、物体認識手段 40 は、物体に係る辞書を用いて所定方向の物体を認識する処理を行う。

50 は、認識結果分析・出力手段であり、物体認識手段 40 で認識した結果を分析し、分析した結果を、例えば、表示装置 207 に表示するように出力する。

【0024】

60 は、物体辞書記憶手段であり、物体認識手段 40 で用いる所望の認識対象に対応する物体辞書を記憶したメモリである。具体的に、物体辞書記憶手段 60 は、同一の物体の異なる方向に対応した複数の辞書のデータを格納する。物体辞書は、数多くの所定方向の物体パターンから機械学習により予め求められたものである。外部記憶装置 206 に記憶され、プログラムの起動時などに RAM 203 に読み込まれる。なお、本実施形態では、顔の垂直方向の角度に応じた、複数の認識辞書を用意しているものとする。もちろん、角度が異なる方向（水平方向等）の角度に応じて認識辞書を分けてもよいが、説明を簡単にするために、以降、垂直方向の角度に応じた認識辞書の例で説明する。

70 は、辞書・照合領域設定手段であり、物体辞書記憶手段 60 に記憶されている複数の認識辞書から使用する認識辞書の選択と、選択された認識辞書を切替えて使用して照合を行ってフレーム画像内の領域ごとに物体認識手段 40 に設定する。即ち、辞書・照合領域設定手段 70 は、フレーム画像の領域ごとに複数の辞書を切り替えて物体認識手段 40 に適用する切替え手段を構成する。

【0025】

80 は、辞書適用領域決定手段であり、物体辞書記憶手段 60 に記憶された認識辞書に対するフレーム画像内の照合領域を決定する。

90 は、辞書適用領域記憶手段あり、辞書適用領域決定手段 80 で決定された認識辞書に対するフレーム画像内の照合領域を記憶する。

【0026】

図 3 の画像取得手段 30、物体認識手段 40、辞書・照合領域設定手段 70、辞書適用領域決定手段 80 は、例えば、図 2 の CPU 201 及び ROM 202（或いは外部記憶装置 206）に格納されている制御プログラム、並びに RAM 203 から構成されている。また、認識結果分析・出力手段 50 は、例えば、図 2 の CPU 201 及び ROM 202（或いは外部記憶装置 206）に格納されている制御プログラム、RAM 203、並びに、表示装置 207 から構成されている。また、物体辞書記憶手段 60、辞書適用領域記憶手段 90 は、例えば、図 2 の外部記憶装置に構成される。

【0027】

図 4 は、本発明の実施形態に係る画像処理装置 106 による画像処理方法の処理手順の一例を示すフローチャートである。

【0028】

まず、ステップ S400 において、辞書適用領域決定手段 80 は、フレーム画像内で適用できる、顔の角度の範囲が異なる複数の認識辞書のそれぞれが適用できる領域を決定する。これは、フレーム画像内の顔の位置と、その位置に顔が存在するときの顔の角度との関係より決定できる。なお、本実施形態では、垂直方向の角度であるものとして説明する。

【0029】

顔の角度は、人物が直立しているときの水平方向を 0 度として現される、図 5 の として示した角度である。

図 5 は、本発明の実施形態を示し、顔の角度を説明する模式図である。

ここで、図 5 の 501 は通路の天井、502 は通路の床、503 は撮像部（カメラ）、

10

20

30

40

50

504は人物である。

【0030】

図5にも示している通り、は天井に沿った直線と、カメラ503から人物504の顔まで引いた直線(図5の実線)とのなす角に等しい。従って、カメラ503から人物504までの距離(図5中のX)と(カメラ503の地面からの高さ-顔の地面からの高さ)(図5中のY)を求めることにより、は、以下の数式(1)で求めることができる。

$$= \tan^{-1}((\text{カメラ503の地面からの高さ} - \text{顔の地面からの高さ}) / \text{カメラ503から人物までの距離}) \quad \dots (1)$$

【0031】

の範囲は、どの範囲を撮影できるかによるので、カメラ503の画角と設置時のカメラ503の角度で決まる。図5の例では、点線が撮影できる範囲である。人物504の顔の高さは、例えば平均的な身長的人物を想定すればよいので、カメラ503の設置条件(カメラの高さ、カメラの角度、カメラの画角)が決まれば、計算により求めることができる。そこで、カメラ503の設置条件をユーザに入力させ、数式(1)によってフレーム画像内の顔の位置と顔の角度の関係を求めることができる。

【0032】

また、フレーム画像内の任意の位置(y)と顔の角度の関係式($\theta = f(y)$)を定義し、設置時にいくつかの値を入力して、 $f(y)$ を求めるようにしてもよい。例えば、フレーム画像内の任意の位置(y)と顔の角度が以下の数式(2)の一次式で表せるとする。

$$\theta = ay + b \quad \dots (2)$$

そうすると、2つ以上のフレーム画面内の位置と顔の角度を予め計測しておけば、係数a、定数bを求めることができる。よって、実際にカメラ503の前に人物504を2地点以上立たせ、それぞれのフレーム画面内の位置とカメラ503からの距離を入力するようにして、係数a、定数bを求めるようにしてもよい。

【0033】

次に、認識辞書について説明する。

認識辞書は、その適用できる顔の角度の範囲が重なるように作成されたものを用いる。

図6は、本発明の実施形態を示し、認識辞書A及びBの適用できる顔の角度と認識精度との関係を示した模式図である。

【0034】

図6では、横軸に認識辞書が適用可能な顔の角度、縦軸に認識精度を示しており、601は認識辞書A、602は認識辞書Bを示している。このとき、認識辞書A(601)の適用可能な顔の角度の範囲は603(1)から604(3)、認識辞書B(602)の適用可能な顔の角度の範囲は605(2)から606(4)である。なお、 $1 < 2 < 3 < 4$ である。

【0035】

既に、[発明が解決しようとしている課題]で述べたように、適用できる顔の角度の範囲の端に近くなると認識精度が落ちてくる。そこで、本実施形態では、2つの認識辞書によって、認識精度の低下を補完しあえるように、適用可能な顔の角度の範囲(画像の領域)が重複するように辞書を作成する。図6の例では、605(2)から604(3)の範囲が相当する。重なりのおおきさは、認識精度の許容範囲をどの程度にするかによって決定できる。このようにして作成された2つの認識辞書A及びBを用いることによって、常に認識精度の高い顔検出を可能とする。また、認識辞書A及びBを適用するフレーム画像内の領域は、前述のフレーム画像内の顔の位置と顔の角度との関係を用いて、次のように重複するように決められる。

【0036】

図7は、本発明の実施形態を示し、認識辞書A及びBの適用できる領域を説明する模式図である。

図7の例では、図7(a)の701(斜線領域)は、顔の位置が上端から下端に向かう

10

20

30

40

50

に従って、顔の角度が 1 から 3 まで変化する。そこで、この領域については、認識辞書 A (6 0 1) が適用可能である。

【 0 0 3 7 】

図 7 (b) の 7 0 2 (斜線領域) は顔の位置が上端から下端に向かうに従って、顔の角度が角度 2 から 4 まで変化する。そこで、この領域については、認識辞書 B (6 0 2) が適用可能である。図 7 (a) 及び (b) を重ねて描くと、図 7 (c) のようになり、7 0 4 (横線領域) の部分は 2 < 3 の領域であり、認識辞書 A 及び認識辞書 B の両方が適用可能である。即ち、領域 7 0 4 においては、画像の領域に用いられる認識辞書の適用領域は、重複している。

【 0 0 3 8 】

以上のようにして、顔の角度の範囲が異なる複数の認識辞書のそれぞれが適用できる領域を決定することができる。これらの領域は、辞書適用領域記憶手段 9 0 に保存される。

【 0 0 3 9 】

ここで、再び、図 4 の説明に戻る。

ステップ S 4 0 0 の処理が終了すると、ステップ S 4 0 1 に進む。

ステップ S 4 0 1 に進むと、画像処理装置 1 0 6 は、処理を終了するか否かを判断する。

【 0 0 4 0 】

ステップ S 4 0 1 の判断の結果、電源 OFF やキーボード 2 0 4 やマウス 2 0 5 を介してユーザから処理の終了の指示があると、本フローチャートの処理を終了する。

【 0 0 4 1 】

一方、ステップ S 4 0 1 の判断の結果、ユーザから処理の終了の指示がなかった場合、ステップ S 4 0 2 に進む。即ち、ユーザから処理の終了の指示があるまで、ステップ S 4 0 2 ~ ステップ S 4 0 6 の処理を繰り返し行う。

【 0 0 4 2 】

ステップ S 4 0 2 に進むと、画像取得手段 3 0 は、撮像手段 1 0 へ入力された映像から、前述した方法によりフレーム画像として取得する。

ここで読み込まれた画像データは、例えば、8 ビットの画素により構成される 2 次元配列のデータであり、R、G、B の 3 つの面により構成される。このとき、画像データが J P E G 等の方式により圧縮されている場合には、画像データを所定の解凍方式にしたがって解凍し、R G B 各画素により構成される画像データとする。さらに、本実施形態では、R G B データを輝度データに変換し、輝度画像データを以後の処理に適用するものとし、画像メモリ (例えば、図 2 の外部記憶装置 2 0 6) に格納する。画像データとして Y C r C b のデータを入力する場合には、Y 成分をそのまま輝度データとしてもよい。

【 0 0 4 3 】

続いて、ステップ S 4 0 3 において、物体認識手段 4 0 は、内部の画像メモリに転送された画像データから、辞書・照合領域設定手段 7 0 で設定された辞書データと照合を行い、所望の物体を認識する。

【 0 0 4 4 】

ここで、まず、一般的な物体認識方法について説明する。

公知技術 1 や公知技術 2 で提案されている方法が知られている。

例えば、公知技術 1 では、ニューラル・ネットワークにより画像中の顔パターンを検出する技術である。以下、その方法について簡単に説明する。

【 0 0 4 5 】

まず、顔の検出を対象とする画像データをメモリに読み込み、顔と照合する所定の領域を読み込んだ画像中から切り出す。そして、切り出した領域の画素値の分布を入力としてニューラル・ネットワークによる演算で 1 つの出力を得る。このとき、ニューラル・ネットワークの重み、閾値が膨大な顔画像パターンと非顔画像パターンにより予め学習されており、例えば、ニューラル・ネットワークの出力が 0 以上なら顔、それ以外は非顔であると判別する。ここで、重みや閾値が辞書データとなる。そして、ニューラル・ネットワー

10

20

30

40

50

クの入力である顔と照合する画像パターンの切り出し位置を、例えば、図 8 に示すように、画像全域から縦横順次に走査していくことにより、画像中から顔を検出する。

図 8 は、画像から顔パターンの探索を行う方法を説明する図である。

具体的には、画像全域 801 を縦横順次に走査して、照合するパターン 802 を抽出し。この照合するパターン 802 に対して、顔判別処理 803 を行う。

【0046】

また、処理の高速化に着目した例としては、公知技術 2 がある。この技術の中では、AdaBoost を使って多くの弱判別器を有効に組合せて顔判別の精度を向上させる一方、夫々の弱判別器を Haar タイプの矩形特徴量で構成し、しかも矩形特徴量の算出を、積分画像を利用して高速に行っている。また、AdaBoost 学習によって得た判別器を直列に繋ぎ、カスケード型の顔検出器を構成するようにしている。このカスケード型の顔検出器は、まず前段の単純な判別器を使って明らかに顔でないパターンの候補をその場で除去する。そして、それ以外の候補に対してのみ、より高い識別性能を持つ後段の複雑な判別器を使って顔かどうかの判定を行っている。これにより、すべての候補に対して複雑な判定を行う必要がないので高速である。なお、公知技術 1 と同様に判別器で用いる重みや閾値が辞書データとなる。

【0047】

次に、本実施形態において、特徴的な辞書・照合領域設定手段 70 等による動作について、図 9 のフローチャートを用いて説明する。

図 9 は、本発明の実施形態を示し、辞書と照合領域を切り替える動作を説明するフローチャートである。

【0048】

まず、ステップ S900 において、辞書・照合領域設定手段 70 は、全てのフレームについて処理が行われたか否かを判断する。この判断の結果、全てのフレームについて処理が行われた場合には、本フローチャートの処理を終了する。

【0049】

一方、全てのフレームについては未だ処理が行われていない場合には、ステップ S901 に進む。即ち、全てのフレームについては未だ処理が行われていない場合には、ステップ S901 ~ ステップ S909 又は S910 までの処理を繰り返す。

【0050】

続いて、ステップ S901 において、辞書・照合領域設定手段 70 は、物体辞書記憶手段 60 から読み込まれた複数の認識辞書の中から、認識辞書 A を選択し、物体認識手段 40 へ設定する。

【0051】

続いて、ステップ S902 において、辞書・照合領域設定手段 70 は、認識辞書 A の照合領域を、辞書適用領域記憶手段 90 から読み出して、物体認識手段 40 へ設定する。前述の通り認識辞書 A の照合領域は図 7 (a) の 701 になる。

【0052】

続いて、ステップ S903 において、物体認識手段 40 は、ステップ S901 で設定された認識辞書 A を用いて、ステップ S902 で設定された照合領域の範囲で、辞書との照合を行う。

【0053】

続いて、ステップ S904 において、辞書・照合領域設定手段 70 は、物体辞書記憶手段 60 から読み込まれた複数の認識辞書の中から、認識辞書 B を選択し、物体認識手段 40 へ設定する。

【0054】

続いて、ステップ S905 において、辞書・照合領域設定手段 70 は、認識辞書 B の照合領域を、辞書適用領域記憶手段 90 より読み出して、物体認識手段 40 へ設定する。前述の通り認識辞書 B の照合領域は図 7 (b) の 702 になる。

【0055】

10

20

30

40

50

続いて、ステップS906において、物体認識手段40は、ステップS904で設定された認識辞書Bを用いて、ステップS905で設定された照合領域の範囲で、辞書との照合を行う。

【0056】

次に、様々な大きさの顔の認識に対応するために、以降の処理で、フレーム画像を縮小して照合を繰り返す。

まず、ステップS907において、例えば、物体認識手段40（或いは辞書・照合領域設定手段70）は、縮小が十分で行われたか否かを判断する。ここでは、照合に用いる画像パターンと同じサイズまで縮小したとき、フレーム画像内で最大の顔を検出することになる。

10

【0057】

ステップS907の判断の結果、縮小が十分でない、即ち、照合に用いる画像パターンよりも小さくならない範囲で縮小可能であるときには、ステップS908へ進む。

【0058】

ステップS908に進むと、物体認識手段40は、所定の縮小率でフレーム画像を縮小する。

【0059】

続いて、ステップS909において、辞書・照合領域設定手段70は、認識辞書Aと認識辞書Bの照合領域を、ステップS908と同じ縮小率で縮小する。そして、その後、ステップS901へ戻る。

20

【0060】

以降、ステップS908において縮小されたフレーム画像に対して、ステップS901～ステップS906の処理を行う。ここで、ステップS902とステップS905で設定される照合領域には、ステップS909で縮小された領域を用いられる。

【0061】

以上のように、1枚のフレーム画像に対して、ステップS901～ステップS909の処理を繰り返す。

【0062】

一方、ステップS907の判断の結果、照合に用いる画像パターンと同じサイズまで縮小したときには、ステップS910へ進み、画像取得手段30は、次のフレーム画像を取得し、ステップS900へ戻る。

30

【0063】

以上の処理によって、顔の角度の適用範囲が重なっている2つの認識辞書A及びBを用いることで、顔がフレーム画像内のどの位置にあっても、高精度な認識ができるようになる。しかしながら、毎フレーム、認識辞書Aと認識辞書Bの両方を用いると、重なり部分（図7（c）の704）は2重に認識辞書との照合（ステップS903とステップS906）が行われることになる。これでは、演算コストが増大してしまう。そこで、以下の方法によって、この課題を回避することが可能である。

【0064】

Nフレーム目とN+1フレーム目で、使用する認識辞書と照合領域を変更する（ただし、Nには自然数が入る）。即ち、Nフレームでは、辞書・照合領域設定手段70は、認識辞書Aを設定し、照合は画像全体ではなく図7（a）の701の領域に対してのみ行う。そして、N+1フレームでは辞書・照合領域設定手段70は、認識辞書Bを設定し、図7（b）の702の領域に対してのみ行う。このように連続したフレームごとに使用する辞書と、照合の領域を切り替えながら、認識を行う。これにより、顔がフレーム画像内のどの位置にあってもNフレームかN+1フレームのどちらかで認識されることになる。また、フレームごとの照合の領域が制限されるので、フレーム画像の全領域について照合を行う場合に比べて、演算コストが少なくてすむ。

40

【0065】

なお、認識結果としては、NフレームとN+1フレームを論理和したものを使用すれば

50

よい。重なる部分において、Nフレーム目とN+1フレーム目の両方で認識されても、前後のフレーム間の時間差が十分に小さければ、位置はほとんど変わらないので、同じものであると判定することは容易である。

【0066】

ここで、再び、図4の説明に戻る。

ステップS403の処理が終了すると、続いて、ステップS404に進む。

ステップS404に進むと、認識結果分析・出力手段50は、現在から所定時間前までの間に検出された被写体領域をRAM203より読み出して、軌跡を生成する。これは、所定時間内に検出された複数ある顔のうち、どれが同一の人物の動きに対応するかを求める処理である。

10

【0067】

この処理の詳細について、図10を用いて説明する。

図10は、本発明の実施形態を示し、軌跡の生成とカウントの一例を示した模式図である。

【0068】

図10において、1001は撮像しているフレーム全体である。ここに、所定の時間に検出された顔の領域を、矩形で表現して重ね描きしている(1003~1005)。図10の例では、3フレーム分を重ね描きしており、最も古いフレームでは1003が、次のフレームでは1004が、その次の現在のフレームでは、1005が検出されているものとする。これらの軌跡を求める方法としては、各領域の中心を求め、各領域の中心間の距離が最小となるもの同士を同一の被写体とみなし、線分で接続するようにすればよい。このようにして求めた軌跡が、図10の例では1009となる。

20

【0069】

続いて、ステップS405において、認識結果分析・出力手段50は、ステップS404で作成された軌跡が、所定の条件を満たすかどうかをチェックし、条件を満たしていればカウントする。ここで所定の条件とは、例えば、図10に示した1002のような計測ラインを横切っているかどうか、である。計測ライン1002は、ユーザによってフレーム画面内に設定される。図10の例では、軌跡1009が計測ライン1002を横切っているので、1とカウントされる。もし、まだ、計測ライン1002を横切っていない軌跡が存在すれば、この時点では、カウントされない。

30

【0070】

続いて、ステップS406において、認識結果分析・出力手段50は、カウントした結果をユーザに対して表示する。

【0071】

その後、再び、ステップS401に戻る。

【0072】

以上のように、適用範囲が重なる辞書を複数用意し、奇数フレームと偶数フレームで使用する辞書と照合範囲を切り替えるようにした。これにより、認識する物体の向きが画面内で変化する場合であっても精度よく認識することができる。

【0073】

本実施形態では、辞書・照合領域設定手段70が使用する認識辞書と照合する範囲をフレームごとに交互に変更したが、次のような方法であっても良い。即ち、図7(c)の703の範囲については、認識辞書Aを用いて照合を行う。重なる領域704に対しても、まずは認識辞書Aを用いて照合を行うが、このとき、認識辞書との照合の結果得られる確からしさ(尤度)を元に、尤度マップを作成する。尤度は、辞書との照合の際の、閾値処理を施す前の演算結果より得られる。

40

【0074】

次に、尤度マップを参照し、尤度が所定値以下の部分に対してのみ認識辞書Bを用いて照合を行う。図7(c)の705に対しては、認識辞書Bを用いて照合を行う。このようにすると、図7(c)の704については部分的に2重に照合を行うことになるが、公知

50

技術２のカスケード型の判別器を用いる場合には、尤度が低いものは、前段の判別器で判定できるため、照合にかかる時間は極めて少なくなる。従って、演算コストは大きく増えることはない。

【００７５】

本実施形態では、顔の位置を検出する例で説明したが、人体全体や上半身、頭部など人物の様々な部位や、自動車や自転車など、様々な物体においても適用可能である。また、特定の人物であるかどうかを個人の顔の特徴から弁別するような場合でも適用可能である。

【００７６】

本実施形態では、顔の垂直方向の角度で説明したが、もちろん、水平方向の角度に関しても同様である。

10

【００７７】

また、本実施形態では、認識結果分析・出力手段５０は、通路を通行する人数をカウントする例を説明した。しかしながら、所定のエリアの混雑率を計測したり、動線を分析したり、特定の人物に対してアラームを発生させるなど、様々な用途に適用可能である。

【００７８】

また、本実施形態では、ＰＣである画像処理装置１０６において認識、計数、表示まで行うように構成したが、これに限ったものではない。例えば、物体認識手段４０から辞書適用領域記憶手段９０までの全てをチップに納め、撮像部１０４と一体化させることにより、計数結果のみＬＡＮケーブル１０５を介して画像処理装置１０６にて受信し、計数結果を閲覧するようにしてもよい。或いは、物体認識手段４０、辞書・照合領域設定手段７０、物体辞書記憶手段６０、辞書適用領域決定手段８０、辞書適用領域記憶手段９０を撮像部１０４と一体化する。そして、認識結果のみＬＡＮケーブル１０５を介して画像処理装置１０６にて受信して画像処理装置１０６において計数するようにしてもよい。

20

なお、本実施形態は、コンピュータ内でプログラムを実行することによっても実現することができることは当然である。

【００７９】

また、本実施形態では、フレーム画像の領域ごとに複数の辞書を切り替える例を説明したが、例えば、画像が時間的に連続した画像であって、画像ごとに使用する辞書を切り替える形態も適用可能である。

30

【００８０】

（その他の実施形態）

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。

このプログラム及び当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録媒体は、本発明に含まれる。

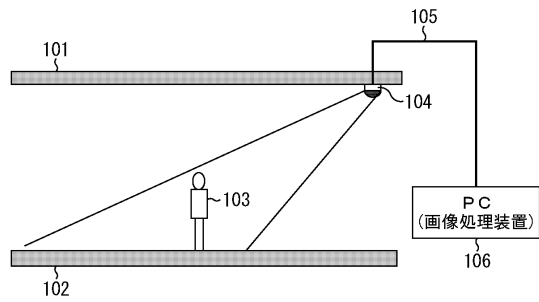
【符号の説明】

【００８１】

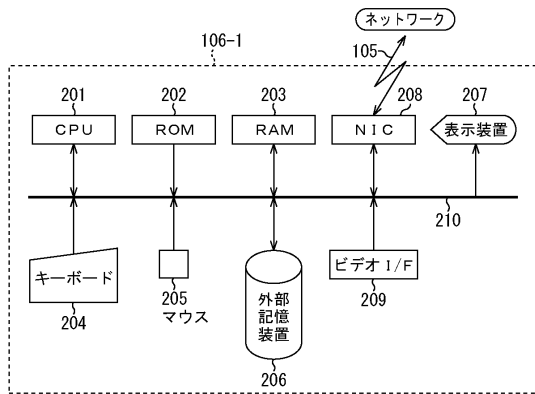
40

１０ 撮像手段、３０ 画像取得手段、４０ 物体認識手段、５０ 認識結果分析・出力手段、６０ 物体辞書記憶手段、７０ 辞書・照合領域設定手段、８０ 辞書適用領域決定手段、９０ 辞書適用領域記憶手段

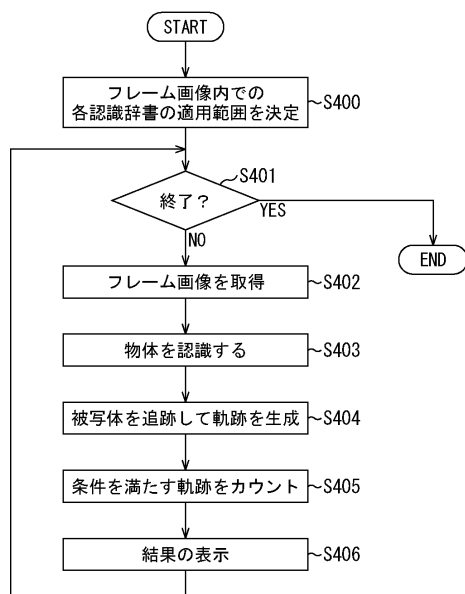
【図 1】



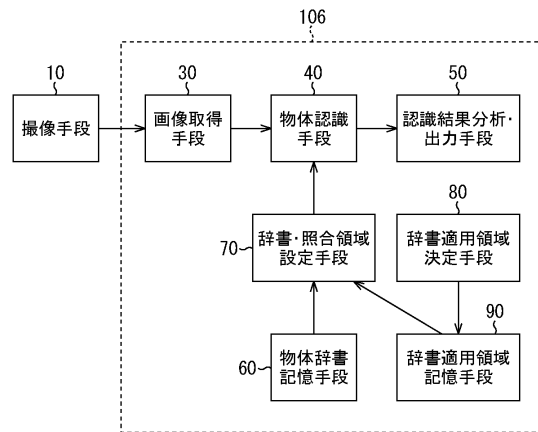
【図 2】



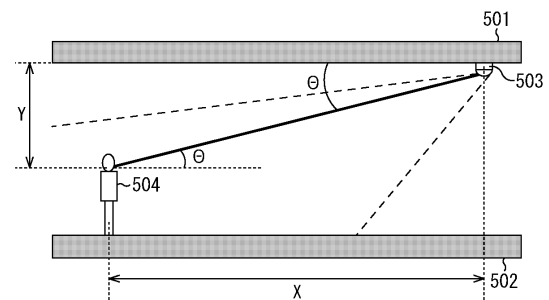
【図 4】



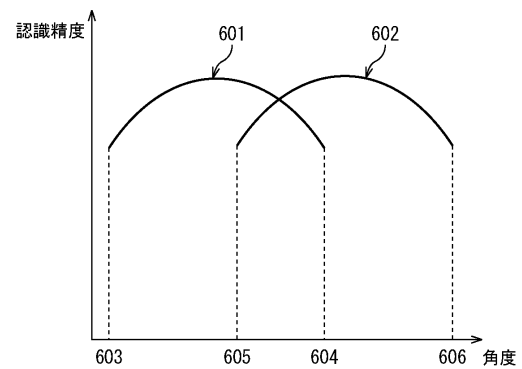
【図 3】



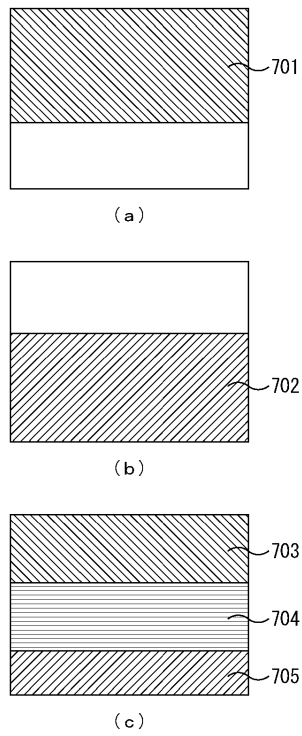
【図 5】



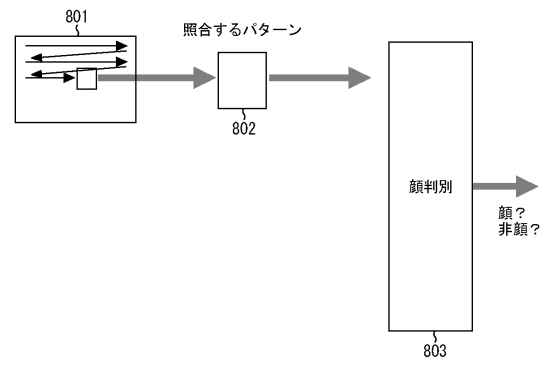
【図 6】



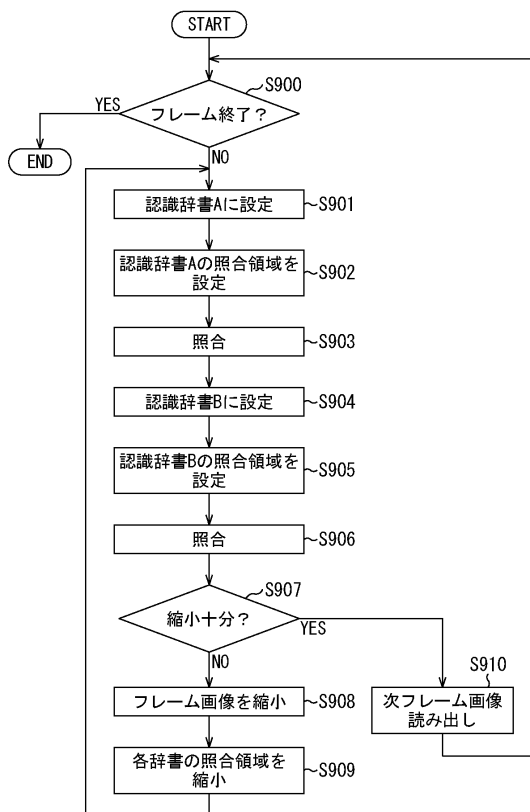
【図 7】



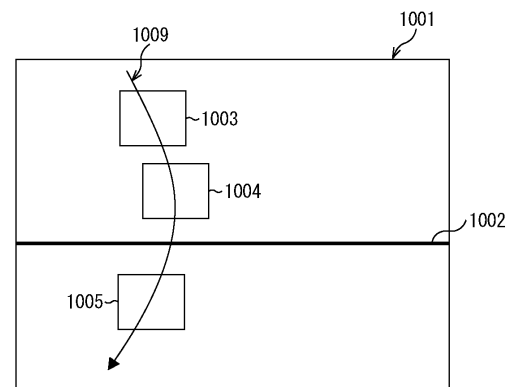
【図 8】



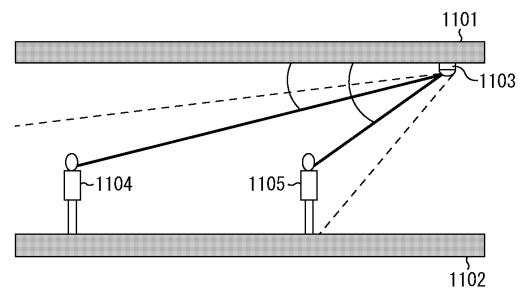
【図 9】



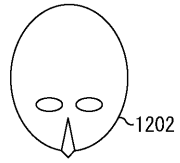
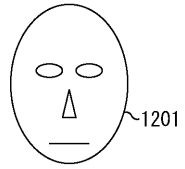
【図 10】



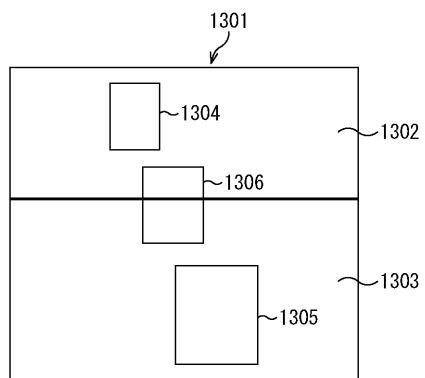
【図 11】



【図 12】



【図 13】



フロントページの続き

(56)参考文献 特開2007-025767(JP,A)
特開2009-230284(JP,A)
特開2009-239871(JP,A)
特開2000-276596(JP,A)
特開昭61-292003(JP,A)
特開2008-040736(JP,A)
特開平04-367987(JP,A)

(58)調査した分野(Int.Cl., DB名)

| | |
|------|-------|
| H04N | 5/232 |
| G06T | 1/00 |
| H04N | 5/225 |
| H04N | 7/18 |