

(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2006-293644

(P2006-293644A)

(43) 公開日 平成18年10月26日(2006.10.26)

(51) Int. Cl.		F I			テーマコード (参考)
G06F 21/20	(2006.01)	G06F 15/00	330F	4C038	
G06T 7/00	(2006.01)	G06T 7/00	510B	5B043	
A61B 5/117	(2006.01)	A61B 5/10	320A	5B285	
		A61B 5/10	320Z		
		A61B 5/10	322		
審査請求 未請求 請求項の数 16 O L (全 17 頁)					

(21) 出願番号 特願2005-112663 (P2005-112663)
 (22) 出願日 平成17年4月8日 (2005.4.8)

(71) 出願人 000001007
 キヤノン株式会社
 東京都大田区下丸子3丁目30番2号
 (74) 代理人 100076428
 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
 (72) 発明者 真継 優和
 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

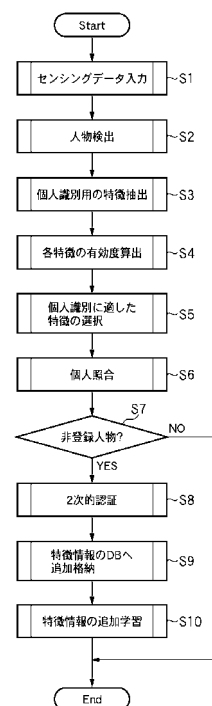
(54) 【発明の名称】 情報処理装置、情報処理方法

(57) 【要約】

【課題】 認証の対象人物から得られる特徴検出信号の有効性などを評価し、照明条件や環境の変動にロバストであってかつ照合精度を高く維持するために、適切な特徴の選択を行う人物認識技術を提供すること。

【解決手段】 画像情報、音声情報から複数種の特徴量を収集し (S3)、収集したそれぞれの特徴量のうち個人認識処理のために用いる特徴量を特定し (S4, S5)、特定した特徴量を用いて個人認識処理を行う (S6)。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

個人認識処理を行う情報処理装置であって、
 人が存在する空間の環境情報を取得する取得手段と、
 前記取得手段が取得した環境情報から複数種の特徴量を収集する収集手段と、
 前記収集手段が収集したそれぞれの特徴量について、有効値を求める計算手段と、
 前記計算手段が計算したそれぞれの特徴量に基づいて、個人認識処理のために用いる特徴量を特定する特定手段と、
 前記特定手段が特定した特徴量を用いて個人認識処理を行う認識手段と
 を備えることを特徴とする情報処理装置。

10

【請求項 2】

前記取得手段は、前記空間の画像情報、前記空間における音声情報、前記空間内に存在する人の生体情報のうち 1 以上を取得することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 3】

前記取得手段は、前記空間の画像情報を取得し、前記収集手段は、人物の容姿もしくは人物の動作に関する画像特徴量を収集することを特徴とする請求項 2 に記載の情報処理装置。

【請求項 4】

前記取得手段は、前記空間の音声情報を取得し、前記収集手段は、非言語に関する音声特徴量を収集することを特徴とする請求項 2 に記載の情報処理装置。

20

【請求項 5】

前記取得手段は、前記空間内に存在する人の生体情報として、虹彩パターン、指紋パターン、静脈パターン、遺伝子配列情報、臭気の少なくとも 1 つを取得することを特徴とする請求項 2 に記載の情報処理装置。

【請求項 6】

前記計算手段は、
 階層型ニューラルネットワークの最終層において、それぞれの特徴量を検出するためのモジュールの正規化した出力レベル V_i を取得する処理を全てのクラス i について行う手段と、

30

全てのクラスの出力レベルの総和 $S = V_1 + V_2 + \dots + V_N$ を求める手段と、
 クラス i の特徴量に対する有効値 Y_i を、クラス i の出現頻度を k_i とすると、 $Y_i = k_i \times V_i / S$ を計算することにより求める手段と
 を備えることを特徴とする請求項 1 に記載の情報処理装置。

【請求項 7】

前記特定手段は、所定の閾値以上の有効度を有する特徴量を、個人認識処理のために用いる特徴量として特定することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 8】

前記有効度は、特徴量の収集頻度またはレベルの少なくとも 1 つに基づいて算出されることを特徴とする請求項 7 に記載の情報処理装置。

40

【請求項 9】

前記特定手段は、前記収集手段により収集された複数の特徴量から所定数の特徴量を所定の確率的重みで特定することを特徴とする請求項 1 に記載の情報処理装置。

【請求項 10】

前記確率的重みは、前記収集手段により収集された特徴量の分布に基づいて決定されることを特徴とする請求項 9 に記載の情報処理装置。

【請求項 11】

更に、個人毎の特徴量群を保持する保持手段を備え、
 前記認識手段は、前記特定手段が特定した特徴量が誰のものであるのかを特定すべく、前記特定手段が特定した特徴量と、前記保持手段が保持するそれぞれの個人の特徴量群と

50

を照合する照合手段と

を備えることを特徴とする請求項 1 に記載の情報処理装置。

【請求項 1 2】

更に、前記特定手段が特定した特徴量が、前記保持手段が保持するそれぞれの個人の特徴量群の何れとも合致しなかった場合には、前記特定手段が特定した特徴量を前記保持手段に登録する登録手段を備えることを特徴とする請求項 1 1 に記載の情報処理装置。

【請求項 1 3】

前記特定手段が特定した特徴量が、前記保持手段が保持するそれぞれの個人の特徴量群の何れとも合致しなかった場合に、前記認識手段が、当該特定した特徴量とは異なる特徴量に基づく認識を行なうことを特徴とする請求項 1 1 に記載の情報処理装置。

10

【請求項 1 4】

個人認識処理を行う情報処理方法であって、
人が存在する空間の環境情報を取得する取得工程と、
前記取得工程で取得した環境情報から複数種の特徴量を収集する収集工程と、
前記収集工程で収集したそれぞれの特徴量について、有効値を求める計算工程と、
前記計算工程で計算したそれぞれの特徴量に基づいて、個人認識処理のために用いる特徴量を特定する特定工程と、
前記特定工程で特定した特徴量を用いて個人認識処理を行う認識工程と
を備えることを特徴とする情報処理方法。

【請求項 1 5】

コンピュータに請求項 1 4 に記載の情報処理方法を実行させることを特徴とするプログラム。

20

【請求項 1 6】

請求項 1 5 に記載のプログラムを格納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。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、個人認識技術に関するものである。

【背景技術】

30

【0002】

個人の照合を行う認証技術としてバイオメトリクスがあり、指紋認証（特許文献 1 を参照）、虹彩認証（特許文献 2 を参照）、声紋認証（特許文献 3 を参照）、顔認証（特許文献 4 を参照）などが用いられている。

【0003】

その高精度化のために複数の照合用特徴を用いて認証を行う方式（特許文献 5 , 6 を参照）がある。

【0004】

例えば、特許文献 7 に記載の個体識別装置では、顔画像認識とアイリス認識を併用している。

40

【0005】

特許文献 8 に開示される生態的特徴の認証方法及び装置、記録媒体では、複数の生体特徴の照合スコアを統合して得られる特徴空間での識別規則に基づいて認証を行っている。

【0006】

特許文献 9 の本人認証システムでは指紋、顔、音声、筆跡のうち、2種類以上の認証方式を組み合わせる確率的に認証を行う構成を開示している。

【0007】

特許文献 10 の個人認証装置では、1 次認証で異なる複数方式による認証処理を行い、その結果を融合して最終的に判定を行う方式を開示している。

【0008】

50

特許文献 11 の個人認証システムでは、生体情報として音声、顔画像の一部若しくは全体、及び体臭を用い、第 1 の生体情報の類似度（または特徴量）と第 2 の生体情報の類似度とを用いて認証の判断を行っている。

【0009】

特許文献 12 に示される生態情報による本人認証システムでは、顔紋情報の適合率が閾値以上の登録対象者が複数存在するときに声紋情報を用いて適合率の評価を行うか、顔紋情報の適合率と声紋情報の適合率とを掛け合わせて閾値処理するなどして認証評価を行っている。

【特許文献 1】特開平 11 - 96363 号公報

【特許文献 2】特開平 11 - 13164 号公報

【特許文献 3】特開平 11 - 73196 号公報

【特許文献 4】特開平 11 - 161791 号公報

【特許文献 5】特開 2001 - 331801 号公報

【特許文献 6】特開 2003 - 186847 号公報

【特許文献 7】特開 2000 - 102524 号公報

【特許文献 8】特開平 11 - 253426 号公報

【特許文献 9】特開 2000 - 242786 号公報

【特許文献 10】特開 2001 - 109855 号公報

【特許文献 11】特開 2000 - 148985 号公報

【特許文献 12】特開 2001 - 338295 号公報

【発明の開示】

【発明が解決しようとする課題】

【0010】

しかしながら上述の複数の照合用特徴を用いて認証を行う技術では、照合に用いる特徴の検出信号の品質を考慮していないため、例えば特徴の計測条件が良好でなく SN 比が悪い場合などでは、品質のよい特徴検出信号だけを用いて照合する場合と比べて誤り発生率が高くなるという問題を少なからず残している。

【0011】

例えば、同一人物であっても見る角度や距離、その人物の心身の状態、照明条件によって見え方が大きく異なる。また、撮影条件が良好でない場合にも人物の認識をする必要がある場合、顔やその他の生体情報を一元的に組み合わせて、或いは決まった方法で認識対象を絞り込むようにして認識を行っても精度が高くないことが多い。

【0012】

また従来では、認証のために対象者をカメラの前に正対させたり、音声指示に従って声を出したりするなど認証されている事を意識させる状況を作り出すことが必要となっており、様々な条件下で被認証者に負担を掛けずに自然に人物認識（認証）を高い精度で行うことが出来なかった。

【0013】

本発明は以上の問題に鑑みてなされたものであり、認証の対象人物から得られる特徴検出信号の有効性などを評価し、照明条件や環境の変動にロバストであってかつ照合精度を高く維持するために、適切な特徴の選択を行う人物認識技術を提供することを目的とする。

【0014】

また、この有効性に基づいて認識対象の絞込みを効率的に行う認識技術を提供する。

【0015】

また、本発明は人物の歩き方、しぐさ、身振り、咳払い、くしゃみなど動作に関する特徴も人物認識の補助情報として絞込みに用いることにより認証に要する時間を短縮し、かつ認識精度を向上させる技術を提供することも目的とする。

【0016】

更に、自然にかつ認証されていることを意識させずに高精度な個人識別を行う人物認識

10

20

30

40

50

技術を提供する。

【課題を解決するための手段】

【0017】

本発明の目的を達成するために、例えば本発明の情報処理装置は以下の構成を備える。

【0018】

即ち、個人認識処理を行う情報処理装置であって、
人が存在する空間の環境情報を取得する取得手段と、
前記取得手段が取得した環境情報から複数種の特徴量を収集する収集手段と、
前記収集手段が収集したそれぞれの特徴量について、有効値を求める計算手段と、
前記計算手段が計算したそれぞれの特徴量に基づいて、個人認識処理のために用いる特
徴量を特定する特定手段と、
前記特定手段が特定した特徴量を用いて個人認識処理を行う認識手段と
を備えることを特徴とする。

【0019】

本発明の目的を達成するために、例えば本発明の情報処理方法は以下の構成を備える。

【0020】

即ち、個人認識処理を行う情報処理方法であって、
人が存在する空間の環境情報を取得する取得工程と、
前記取得工程で取得した環境情報から複数種の特徴量を収集する収集工程と、
前記収集工程で収集したそれぞれの特徴量について、有効値を求める計算工程と、
前記計算工程で計算したそれぞれの特徴量に基づいて、個人認識処理のために用いる特
徴量を特定する特定工程と、
前記特定工程で特定した特徴量を用いて個人認識処理を行う認識工程と
を備えることを特徴とする。

【発明の効果】

【0021】

本発明の構成により、認証の対象人物から得られる特徴検出信号の有効性などを評価し、
照明条件や環境の変動にロバストであってかつ照合精度を高く維持するために、適切な
特徴の選択を行うことができる。

【発明を実施するための最良の形態】

【0022】

以下添付図面を参照して、本発明を好適な実施形態に従って詳細に説明する。

【0023】

[第1の実施形態]

図2は、人物像の画像や音声を取得し、取得した画像、音声に基づいて人物を認識する
ためのシステムの機能構成を示すブロック図である。同図に示す如く、本実施形態に係る
システムは、センサ部111、認識処理装置1000とで構成されており、且つこの認識
処理装置1000は人物検出部150、特徴情報検出部112、学習分類部113、デー
タベース部115、照合部118、2次的認証部119により構成されている。

【0024】

センサ部111は現実空間の画像情報、音声情報を収集するためのものであり、より具
体的には、現実空間の画像（動画像／静止画像）を撮像するためのビデオカメラや、現実
空間中の音声を収集する為のマイクなどにより構成されている。本実施形態では現実空間
における環境情報として画像情報と音声情報とを収集すべく、センサ部111は図9に示
す如く、現実空間の画像情報を収集する画像センシング部1111と、現実空間中の音声
情報を収集する音声センシング部1112とで構成されている。図9は、本実施形態に係
るセンサ部111の機能構成を示すブロック図である。

【0025】

なお、現実空間の画像情報、音声情報を収集する手段についてはこれに限定するもので
はなく、例えば外部装置に保持されている現実空間の画像情報、音声情報を収集するよう

10

20

30

40

50

にしても良い。

【 0 0 2 6 】

図 2 に戻って、人物検出部 1 5 0 は、センサ部 1 1 1 によって取得した画像から、人物像の領域を特定する処理を行う。図 7 は、人物検出部 1 5 0 の機能構成を示す図である。同図に示す如く、人物検出部 1 5 0 は階層型のニューラルネットワーク（階層的神経回路網モデルの一つである Convolutional ネットワークの構造 (LeCun, Y. and Bengio, Y., 1995, "Convolutional Networks for Images Speech, and Time Series" in Handbook of Brain Theory and Neural Networks (M. Arbib, Ed.), MIT Press, pp. 255-258)）により構成されている。

【 0 0 2 7 】

データ入力層 7 0 1 にはセンサ部 1 1 1 によって取得した画像が入力される。j 番目の特徴検出層 (1, j) は、前段の層からの入力パターン中の局所的な特徴（一種類以上）を所定の解像度で検出するニューロン素子（検出予定の特徴量の種類によって決まる共通の局所受容野構造を有する）から構成される。

【 0 0 2 8 】

k 番目の特徴統合（特徴プーリング）層 (2, k) は、所定の受容野構造（以下、受容野とは直前の層の出力素子との結合範囲を、受容野構造とはその結合荷重の分布を意味する）を有するニューロン素子からなり、特徴検出層からの複数ニューロン出力（特徴統合層ニューロンの局所受容野内）の統合処理として局所平均化、最大出力抽出等によるサブサンプリングなどの演算を行う。特徴統合層各ニューロンの各受容野も検出される特徴クラスに対応する各モジュール内で共通の構造を有している。詳細は、特開 2002 - 8032 号公報、特開 2002 - 358500 号公報などを参照されたい。そして最終出力層からは、画像中の人物の検出位置情報が出力される。

【 0 0 2 9 】

特徴情報検出部 1 1 2 は、人物検出部 1 5 0 が特定した位置における人物像を個人認識するために適したクラスの特徴量、上記音声センシング部 1 1 1 2 から取得した音声情報から個人認識する為に適したクラスの特徴量を求める処理を行う。特徴情報検出部 1 1 2 は図 3 に示す如く、特徴検出部 1 1 2 0、特徴有効度評価部 1 1 2 1、特徴選択部 1 1 2 2 により構成されている。図 3 は特徴情報検出部 1 1 2 の機能構成を示すブロック図である。

【 0 0 3 0 】

特徴検出部 1 1 2 0 は、人物検出部 1 5 0 が特定した位置における人物像に係る様々な特徴量、音声センシング部 1 1 1 2 から取得した音声情報から個人認識を行うための必要な様々な特徴量を求める。

【 0 0 3 1 】

特徴有効度評価部 1 1 2 1 は、特徴検出部 1 1 2 0 が求めたそれぞれの特徴量に対して有効値を求める。

【 0 0 3 2 】

特徴選択部 1 1 2 2 は、特徴有効度評価部 1 1 2 1 が求めたそれぞれの特徴量の有効値に基づいて、人物検出部 1 5 0 が特定した位置における人物（音声センシング部 1 1 1 2 から取得した音声情報に含まれている人物）を個人認識する為に用いる特徴量を決定する。

【 0 0 3 3 】

照合部 1 1 8 は、特徴情報検出部 1 1 2 によって得られた特徴量（特徴選択部 1 1 2 2 が決定した特徴量）と、データベース部 1 1 5 に個人毎に登録されている特徴量とを照合して、個人認識する。当然、特徴情報検出部 1 1 2 によって得られた特徴量（特徴選択部 1 1 2 2 が決定した特徴量）が、データベース部 1 1 5 に個人毎に登録されている特徴量の何れにも該当しない場合には、「人物検出部 1 5 0 が特定した位置における人物像はデータベース部 1 1 5 には登録されていない」と判断するのであるが、このような場合には、学習分類部 1 1 3 は、特徴情報検出部 1 1 2 によって得られた特徴量をデータベース部

10

20

30

40

50

115に登録すると共に、特徴情報検出部112にこの特徴量を用いて人物検出部150が特定した位置における人物像を学習させ、その学習結果をデータベース部115に登録する。これにより、次回、この特徴量を照合部118に提示するとこの人物が認識できるようにする。

【0034】

データベース部115は、それぞれの個人について、クラス毎に特徴量が登録されている。クラスとしては、複数の角度から見た人物の顔画像、もしくは、顔の3次元形状モデルと離散的な角度から見た顔画像、人物の服装、所持品、髪型に関する画像特徴、身振り、歩き方に関する画像特徴、及び、発話、笑い、咳、くしゃみなどでの音声特徴などを含み、それぞれのクラスにおける特徴量が個人毎にデータベース部115に登録されている。

10

【0035】

本実施形態で用いられるクラスとしては、顔、髪型、後姿、体型、姿勢、服装、持ち物などについて目や口など人体構成パーツの局所的形状、輪郭形状、色成分特徴、テクスチャ特徴量、目や口など人体構成パーツの配置情報等の各クラスに関する静止画ベース画像特徴、および身振り・歩き方などに関する人体構成パーツの時空間ドメインでの軌跡情報などの動画ベースの画像特徴、更には、咳やくしゃみなどの非言語音声、発話時の音声に関する声紋等の音声特徴などが用いられる。声紋を用いた個人識別は、本出願人による特開平7-302351号公報などに記載の技術を用いればよい。

【0036】

また、照合部118についてさらに説明すると、照合部118は画像情報についてデータベース部115と照合する処理と、音声情報についてデータベース部115と照合する処理とを行う。

20

【0037】

画像情報の照合については、顔、髪型、後ろ姿、体型、服装、持ち物などの静止画ベース特徴、身振り・歩き方のくせなどの動画ベース特徴について照合を行い、これらの特徴について個人認識を行う。例えば、服装特徴の認識においては、服装のテクスチャ特徴、色成分特徴、服装の型（ワンピース、ズボン、スカート、シャツ、セーター、コートなどの各クラス）の輪郭若しくは特定視点位置からの画像について局所特徴の空間配置情報（本出願人の特許第3078166号公報参照）などが照合の際に用いられる特徴クラスである。持ち物やその他の静止画ベース照合も同様にして行う。身振りや歩き方のくせに関する特徴検出方法は、公知の技術（例えば、特開2002-8042号公報、特開2002-344946号公報など）を用いればよく、ここでの主眼とするところではないのでその説明は省略する。

30

【0038】

2次的認証部119は、照合部118が個人を認識すると、更にその個人についてより詳細に認識処理を行うべく、その他の特徴量を用いて認識処理を行う。

【0039】

図6は、上記認識処理装置1000の基本構成を示すブロック図である。認識処理装置1000としては一般のPC（パーソナルコンピュータ）やWS（ワークステーション）などが適用可能であるが、これに限ったものではない。

40

【0040】

同図において801はCPUで、RAM802やROM803に格納されているプログラムやデータを用いて本装置全体の制御を行うと共に、認識処理装置1000が行う後述の各処理を実行する。なお図2に示した人物検出部150、特徴情報検出部112、学習分類部113、照合部118、2次的認証部119は何れもCPU801が有する機能一部として動作する。

【0041】

802はRAMで、外部記憶装置806からロードされたプログラムやデータ、インターフェース807を介して受信したプログラムやデータを一時的に記憶するためのエリアを備えると共に、CPU801が各種の処理を実行する際に使用するワークエリアも備え

50

る。

【 0 0 4 2 】

8 0 3 は R O M で、本装置の設定データやブートプログラムなどを格納する。

【 0 0 4 3 】

8 0 4 は操作部で、キーボードやマウスなどにより構成されており、本装置の操作者が操作することで各種の指示を C P U 8 0 1 に対して入力することができる。

【 0 0 4 4 】

8 0 5 は表示部で、C R T や液晶画面などにより構成されており、C P U 8 0 1 による処理結果を画像や文字などもでって表示することができる。

【 0 0 4 5 】

8 0 6 は外部記憶装置で、ハードディスクドライブ装置などの大容量情報記憶装置により構成されており、ここに O S (オペレーティングシステム) や C P U 8 0 1 に、認識処理装置 1 0 0 0 が行う後述の各処理を実行させるためのプログラムやデータが保存されている。また、ここには図 2 に示したデータベース部 1 1 5 が保持する情報群、即ち、それぞれの個人について、クラス毎の特徴量が登録されている。そしてこれらの一部若しくは全部は C P U 8 0 1 の制御に従って R A M 8 0 2 にロードされ、C P U 8 0 1 の処理対象となる。

【 0 0 4 6 】

8 0 7 は I / F で、本装置が外部機器とのデータ通信を行うために使用するものであり、例えばここに上記ビデオカメラやマイクなどを接続することができるし、その他にも、現実空間の画像情報や音声情報を保持する装置にネットワークを介して接続することもできる。

【 0 0 4 7 】

8 0 8 は上述の各部を繋ぐバスである。

【 0 0 4 8 】

次に、上記構成を備える認識処理装置 1 0 0 0 が行う本実施形態に係る個人認識処理について、同処理のフローチャートを示す図 1 を用いて以下、説明する。なお、同図のフローチャートに従った処理を C P U 8 0 1 に実行させるためのプログラムやデータは外部記憶装置 8 0 6 に保存されており、これを適宜 C P U 8 0 1 の制御に従って R A M 8 0 2 にロードし、C P U 8 0 1 がこれを用いて処理を行うことで、本装置は以下説明する各処理

【 0 0 4 9 】

先ず、センサ部 1 1 1 から認識処理装置 1 0 0 0 に対して画像情報、音声情報(まとめてセンシングデータと呼称する場合がある)が入力されるので、C P U 8 0 1 はこれを検知してこれらの情報を R A M 8 0 2 に一時的に記憶する(ステップ S 1)。

【 0 0 5 0 】

次に C P U 8 0 1 は人物検出部 1 5 0 として機能し、R A M 8 0 2 に記憶した画像から人物の頭部、若しくは顔面の位置を主として検出し、検出した位置に基づいて画像中における人物領域を特定する(ステップ S 2)。

【 0 0 5 1 】

次に、C P U 8 0 1 は特徴情報検出部 1 1 2 として機能し、ステップ S 2 で特定した人物領域から予め決まったクラス毎の特徴量を検出すると共に、センサ部 1 1 1 から取得した音声情報から予め決まったクラス毎の特徴量を検出し(ステップ S 3)、検出したそれぞれの特徴量毎に有効値を求め(ステップ S 4)、求めたそれぞれの有効値に基づいて以下の処理で使用する特徴量を選別する(ステップ S 5)。

【 0 0 5 2 】

ステップ S 3 では、顔面、髪型、腕や手、脚部、服装、持ち物などの静止画ベースの特徴量、身振り・歩き方などの動画ベースの特徴量、音声情報を検出する。なお、これらの抽出は全て完全に行われる必要は無く、他の物体による遮蔽や照明条件などによって抽出されなくてもよい。

10

20

30

40

50

【 0 0 5 3 】

次に、上記ステップ S 4 における処理について説明する。本ステップでは上述の通り、特徴量毎に有効値なるものを求めるのであるが、その求める方法の一例について以下に説明する。

【 0 0 5 4 】

先ず、図 7 に示した階層型ニューラルネットワークの最終層において、クラス i ($1 \leq i \leq N$) の特徴量を検出するためのモジュールの出力レベル V_i (0 (最小) から 1 (最大) に正規化したもの) を取得する処理を全てのクラス (全ての i) について行う。本実施形態では音声情報についてはこのような階層型ニューラルネットワークを用いていないが、このような処理を行う場合には、音声情報について予め決められたクラスを認識する
10
ような階層型ニューラルネットワークを用意しておき、これに取得した音声情報を入力する。これにより、音声情報のそれぞれの暮らすにおける出力レベルを求めることができる。これはその他の情報 (例えば生体情報) についても同様である。

【 0 0 5 5 】

次に、全てのクラスの出力レベルの総和 S ($S = V_1 + V_2 + \dots + V_N$) を求める。また、それぞれのクラスについて、予め出現頻度を求めておき、そしてクラス i の特徴量に対する有効値 Y_i を、クラス i の出現頻度を k_i とすると、

$$Y_i = k_i \times V_i / S$$

を計算することにより求めることができる。これにより、それぞれのクラスについて相対的な有効度 (= 有効値) を求めることができる。なお、それぞれのクラスについて有効値
20
を求める方法についてはこれに限定するものではない。

【 0 0 5 6 】

例えば、上記ニューラルネットワークにおいて、自己組織化マップで得られた隣接、若しくは近接する各クラスの代表的局所特徴は、人間が視覚的に認識しても互いに類似度が高い場合がある。そこで (所定の特徴空間で互いに直交するような) 互いに類似度の低い代表的クラスを選択する。具体的には、特徴の類似度を評価するための距離尺度 (特徴ベクトル空間のユークリッド距離など) を導入して互いに所定閾値以上に離れた特徴を選択するか、或いは人間の視覚認識により代表的な特徴を選択する。この結果得られるそれぞれの特徴クラスにはある程度異なる変動 (サイズ変動、向き変動、照明変動など) 下での複数の局所特徴が含まれている。
30

【 0 0 5 7 】

よってこのようにして以下の処理で用いる特徴量を選択するようにしても良い。

【 0 0 5 8 】

次に、上記ステップ S 5 における処理について説明する。本ステップでは上述の通り、ステップ S 4 で求めた各クラスの有効値に基づいて、以下の処理で用いる特徴量を選択するのであるが、例えば、所定の閾値以上の有効値を有する特徴量を以下の処理で用いる特徴量として選択する。このとき用いる閾値は、それぞれの有効値の分布の統計量 (平均値、分散値など) を調べて適応的に設定しても良い。本実施形態では、有効度の分布の平均値を閾値として用いる。
40

【 0 0 5 9 】

有効値を用いた特徴量の選択の例としては、遠方から人物が歩いて近づいて来るような場合には、顔画像は顔のサイズが小さすぎて人物の識別には適さないことが多い。そのような場合には、顔特徴クラスは有効度が低く、歩き方や容姿全体、髪型、服装などの特徴は有効度が高くなり、歩き方をはじめとする特徴を用いて対象人物を絞り込むことが可能である。

【 0 0 6 0 】

また、人物の後ろ姿しか観測されない場合には正面顔などから得られる特徴の有効度は低いことは言うまでもない。このように予めどのような順番で特徴の選択と絞り込みを行うかは決まっておらず、環境の条件やセンシング情報取得の条件、認識対象人物の置かれた状況などによって決まる上述した有効度を評価することにより決まる。
50

【0061】

図1に戻って次にCPU801は照合部118として機能し、ステップS5で選択した特徴量が誰のものであるのかを特定するために、ステップS5で選択した特徴量とデータベース部115に登録されている個人毎の特徴量との照合処理を行う(ステップS6)。ステップS5で選択した特徴量には画像特徴と音声特徴とがあるが、画像特徴の照合については、特徴クラスが静止画ベースか動画ベースかによらずに画像データの入力には動画で行う。時系列データとして複数の画像データを入力して照合を行い、例えば特定人物と判定された画像データ数の割合が所定値以上であれば残りは特徴抽出が良好に行われなかったための誤判定とみなす。これは、撮影条件の変動に伴う誤認識発生の低減を図るためである。

10

【0062】

そしてこの照合処理の結果、データベース部115に登録されている個人毎の特徴量の何れにもステップS5で選択した特徴量が合致していない場合には処理をステップS7を介してステップS8に進め、CPU801は2次的認証部119として機能し、2次的認証処理を開始する(ステップS8)。即ち、データベース部115に登録されている個人毎の特徴量の何れにもステップS5で選択した特徴量が合致していないかどうかは判定誤差があるために、厳密には判定できない。そこで更に、認証処理(2次的認証処理)を行うことで、より厳密に判定する。

【0063】

2次的認証では虹彩認証、遺伝子配列に基づく認証、指紋認証などの生体情報ベースで行う認証、パスワードや生年月日その他本人確認のための属性情報を示すテキスト情報を用いた認証を行う。このテキスト情報は、非登録人物と判定された人物に対して入力指示などを行って入力する。

20

【0064】

そしてこの2次的認証処理の結果、データベース部115に登録されている個人毎の特徴量の何れかにステップS5で選択した特徴量が合致している場合には、ステップS5で選択した特徴量をデータベース部115に登録し(ステップS9)、更に、ステップS5で選択した特徴量が、この特徴量に合致する個人(特定個人)の特徴量として認識するように、認識モデルに対する学習処理を行う(ステップS10)。この学習処理については、公知の手法(特開平5-89076号公報、特開平5-108596号公報など参照)を用いればよい。このとき、照明条件、湿度、温度などを計測するセンサ出力を環境情報としてデータベースに格納してもよい。これは特徴選択部1122が選択した特徴が環境情報(例えば、照度が低いため顔画像データが人物認識に有用でない場合など)を考慮すると不適切と判断できる場合に対応して、後で同様な環境条件で人物認識する場合に特徴選択を適切に行えるようにする為である。即ち、特徴選択部1122は、特徴クラスの有効度と環境情報の両者を参照して認識に用いる特徴クラスを選択してもよい。どの環境情報を格納するかは、照合に用いた特徴クラスに依存する。例えば、顔画像から抽出される特徴クラスの場合には、照度及び露光量などの撮影条件が環境情報として用いられる。

30

【0065】

以上の説明により、本実施形態によって、照明条件や環境の変動にロバストであってかつ照合精度を高く維持することができる。また認識対象の絞込みを効率的に行うことができる。

40

また、本発明は人物の歩き方、しぐさ、身振り、咳払い、くしゃみなど動作に関する特徴も人物認識の補助情報として絞込みに用いることにより認証に要する時間を短縮し、かつ認識精度を向上させることができる。

【0066】

なお、本実施形態では現実空間における環境情報として画像情報、音声情報を収集し、それぞれの情報の特徴量のうち使用する特徴量を決定したが、環境情報として用いる情報はこれに限定するものではなく、画像情報のみ、音声情報のみであっても良いし、その他の情報と適宜組み合わせ用いるようにしても良い。

50

【 0 0 6 7 】

また、本実施形態では図 2 に示した一部は CPU 8 0 1 が有する機能として説明したが、各部を専用のハードウェアでもって構成するようにしても良い。

【 0 0 6 8 】

〔 第 2 の実施形態 〕

本実施形態では、家庭内などの生活環境やオフィスなどの特定環境において通常若しくは一定時間帯にその場所に居る人物の認識、或いはその人物の状況（正常／異常）認識を行う場合について説明する。

【 0 0 6 9 】

より詳しくは、顔画像などの画像ベース特徴の検出手段（撮像装置など）や声紋などの音声ベース特徴の検出手段（マイクロフォンなど）、及び以下に例示する生体情報の検出手段（センサ手段）を環境内に分散して設置する。ここで生体情報としては、虹彩パターン、指紋パターン、静脈パターン、遺伝子配列情報、臭気の少なくとも一つを用い、個々に対応するセンサからのセンシングデータを処理して抽出する。これらはいずれも個人を特定する生体センシングデータとして従来から扱われてきた（匂いに基づく個体識別技術は、特開平 9 - 131335 号公報等を参照）。人物の状況認識は主に撮像装置から得られる画像データを処理して行う。

【 0 0 7 0 】

人物の認識や人物の状況認識は、それらセンサから得られるセンシングデータと第 1 の実施形態で説明した画像情報、音声情報を併用して行う。

【 0 0 7 1 】

図 4 は、本実施形態に係る個人認識システムの機能構成を示すブロック図である。同図に示す如く、本実施形態に係るシステムは、画像入力部 2 2 1 0、音声入力部 2 2 1 1、認識処理装置 4 0 0、そして生体情報入力部 2 3 0 とで構成されており、且つこの認識処理装置 4 0 0 は、人物検出部 2 2 5 0、特徴情報検出部 2 5 1 2、学習分類部 2 5 1 3、データベース部 2 5 1 5、照合部 2 5 1 8 により構成されている。

【 0 0 7 2 】

画像入力部 2 2 1 0、音声入力部 2 2 1 1 はそれぞれ画像センシング部 1 1 1 1、音声センシング部 1 1 1 2 と同様のものであり、それぞれ画像情報、音声情報を収集するためのものである。生体情報入力部 2 3 0 は同図に示す如く、虹彩パターン入力部 2 2 1 2、指紋パターン入力部 2 2 1 3、静脈パターン入力部 2 2 1 4、臭気入力部 2 2 1 5、遺伝子配列情報入力部 2 2 1 6 を有する。なお、生体情報入力部 2 3 0 はこれら全てによって構成されることに限定するものではなく、何れか 1 つのみを用いるようにしても良いし、複数個のものを適宜組み合わせる構成するようにしても良い。また、これ以外の生体情報をセンシングするセンサを生体情報入力部 2 3 0 に加えるようにしても良い。

【 0 0 7 3 】

認識処理装置 4 0 0 の構成は認識処理装置 1 0 0 0 の基本構成から 2 次的認証部 1 1 9 を省いた構成となっており、且つ認識処理装置 4 0 0 を構成する各部は基本的には認識処理装置 1 0 0 0 を構成する各部と同様の処理を行う。よって特徴情報検出部 2 5 1 2 は第 1 の実施形態と同様に、図 3 に示す構成を有する。また、認識処理装置 4 0 0 の基本構成については第 1 の実施形態と同様に、図 6 に示した構成を有する。

【 0 0 7 4 】

図 5 は、認識処理装置 4 0 0 が行う個人認識処理のフローチャートである。まず、上記ステップ S 1、S 2 における処理を行って、画像入力部 2 2 1 0 から得た画像中における人物領域を特定する（ステップ S 2 1）。また第 1 の実施形態と同様に音声情報を音声入力部 2 2 1 1 によって取得している。

【 0 0 7 5 】

次に、生体情報入力部 2 3 0 によって空間中に存在する人物からの生体情報を収集する（ステップ S 2 2）。

【 0 0 7 6 】

10

20

30

40

50

そしてステップ S 2 3 ~ ステップ S 2 6 の処理はそれぞれ上記ステップ S 3 ~ ステップ S 6 の処理と同様にして行う。なお、ステップ S 2 3 ~ ステップ S 2 6 の処理では画像情報、音声情報に加えて生体情報が加わっている点が第 1 の実施形態とは異なる。よってステップ S 2 3 では、以下に説明するように特定種類のパターンの検出を行ってもよい。

【 0 0 7 7 】

本実施形態では、対象人物に本システムによって認識がなされていることを意識させないために特別な動作をこの人物に要求せず、また各種センサの動作状態は通常の被認識対象である人物には分からないようにしておく。人物の居る環境の中で取得可能なセンシングデータのみを用いて人物の認識を行うものとする。

【 0 0 7 8 】

そのために有効度評価部 2 1 2 1 は各センサからのセンシングデータから有用な信号が取り込まれたか否かをモニタリングする機能も有する。例えば有効度評価部 2 1 2 1 は、虹彩パターン入力部 2 2 1 2 から虹彩に相当するパターンが入力されたかを判定する虹彩判定部、画像入力部 2 2 1 0 からの出力に対しては顔検出部、後ろ姿検出部、髪検出部、服装検出部等を備え、それぞれの検出レベルが閾値以上の時だけ個人識別の為の特徴検出レベルの有効度評価を行う。なお、各検出部は、対応するセンサ部に内蔵されるように設定されていても良い。

【 0 0 7 9 】

なお、画像を撮像するセンサ、指紋センサ、虹彩画像入力を行うセンサは、小型であって壁面や家具、事務機、机、マウス、キーボード、リモコンなどに埋め込まれている。静脈パターン、遺伝子配列のセンシング部等はその機能に応じて専用センシング部として設置されていてもよい。

【 0 0 8 0 】

例えば、遺伝子配列のセンシング部は、対象人物の体毛、爪、体液などを自動的に採取する装置により採取されたサンプルを用いて遺伝子配列の解析を行う。例えばこの人物から自然に抜け落ちた髪の毛を採取して遺伝子配列の解析を行う。採取されたサンプルがこの人物のものであることの確認は、特徴検出部 2 1 2 0 に内蔵される画像認識部により行い、例えばその人物から髪の毛が抜け落ちた瞬間を識別して行う。なお、遺伝子配列情報の解析部は遠隔地にあり、採取部のみがこの人物の存在する環境に備わっているものでもよい。採取部の例としては、小型の掃除ロボットやハンド・アイ協調動作ロボットなどの形態がある。

【 0 0 8 1 】

静脈パターンのセンシング部の構成としては（特開 2001 - 184507 号公報など参照）例えば、マウスの表面が所定の受光部をなし、所定の位置に設定された光源からの光を通じて静脈パターンに対応する透過光の分布を受光するなどの構成でよい。虹彩パターンのセンシング部としては、自律移動ロボットの目その他の画像入力部が用いられる。

【 0 0 8 2 】

特徴選択部 2 1 2 2 は、有効度評価値に基づいて第 1 の実施形態と同様に個人識別用の特徴の選択を行う。

【 0 0 8 3 】

図 8 は、図 1 若しくは図 5 に示したシステムを搭載する信号処理装置（例えば、カメラ、ビデオカメラ、スキャナ、複写機、プリンタなど）の基本構成を示す図である。

【 0 0 8 4 】

同図において 9 は装置本体で、撮影レンズおよびズーム撮影用駆動制御機構を含む結像光学系 9 2、CCD 又は CMOS イメージセンサ 9 3、撮像パラメータの計測部 9 4、映像信号処理回路 9 5、記憶部 9 6、画像入力動作の制御、画像入力条件の制御などの制御用信号を発生する制御信号発生部 9 7、表示ディスプレイ 9 8、センサ情報やユーザの操作情報を入力する入力部 9 9、記録媒体 9 10、そして図 2 若しくは図 4 に示したシステム（人物認識装置）9 1 を備える。

【 0 0 8 5 】

この信号処理装置は、入力部 99 からのセンサ情報に基づいて、人物認識装置 91 により特定のユーザを認識し、そのユーザだけに使用許可を与える機能、ユーザごとに使える機能を制限できる機能を備える。予め登録された人物との照合を行う際に用いられるのは、顔など頭部の認識、指紋認証、虹彩認証、声紋認証などである。制御信号発生部 97 は、その人物の認識結果に基づき、本装置に対するアクセス制御（ユーザごとの扱える機能の制限など）を行う制御信号を発生する。なお、顔などを認識対象とする場合は、認識対象の画像を結像光学系 92 から入力するようにしてもよい。

【0086】

そして人物認識装置 91 の動作は第 1 の実施形態、第 2 の実施形態の何れの動作であっても良い。例えば、図 1 に示す処理フローにおいて各種特徴の有効度が評価され、顔画像ベース特徴が選択され、個人識別が行われた結果、非登録人物と判定された場合、制御信号発生部 97 は、2 次認証処理としてディスプレイ 98 に、この人物のパスワードを入力部 99 のテンキーボタンなどから入力するように指示する表示制御信号や、指紋認証を行うように指示する表示制御信号を発生してもよい。このとき、個人識別用の特徴クラスの追加学習を行う。

10

【0087】

また、非登録人物と判定された場合に本装置の機能を自動停止するような制御信号を発生しても良い。

【0088】

また、本装置の不正使用防止ほか、入退室管理など様々な用途に用いることができる。特に、以上説明したような制御を行うことにより、ユーザとしての人物を確実に認識し、不正使用や盗難を防止することができる。

20

【0089】

[その他の実施形態]

また、本発明の目的は、前述した実施形態の機能を実現するソフトウェアのプログラムコードを記録した記録媒体（または記憶媒体）を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（または CPU や MPU）が記録媒体に格納されたプログラムコードを読み出し実行することによっても、達成されることは言うまでもない。この場合、記録媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記録した記録媒体は本発明を構成することになる。

30

【0090】

また、コンピュータが読み出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているオペレーティングシステム（OS）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【0091】

さらに、記録媒体から読み出されたプログラムコードが、コンピュータに挿入された機能拡張カードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張カードや機能拡張ユニットに備わる CPU など実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

40

【0092】

本発明を上記記録媒体に適用する場合、その記録媒体には、先に説明したフローチャートに対応するプログラムコードが格納されることになる。

【図面の簡単な説明】

【0093】

【図 1】本発明の第 1 の実施形態に係る個人認識処理のフローチャートである。

【図 2】人物像の画像や音声を取得し、取得した画像、音声に基づいて人物を認識するた

50

めのシステムの機能構成を示すブロック図である。

【図 3】特徴情報検出部 112 の機能構成を示すブロック図である。

【図 4】本発明の第 2 の実施形態に係る個人認識システムの機能構成を示すブロック図である。

【図 5】本発明の第 2 の実施形態に係る個人認識処理のフローチャートである。

【図 6】認識処理装置 1000 の基本構成を示すブロック図である。

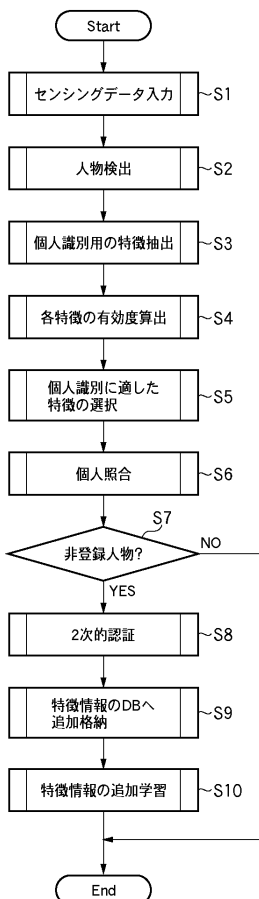
【図 7】人物検出部 150 の機能構成を示す図である。

【図 8】図 1 若しくは図 5 に示したシステムを搭載する信号処理装置（例えば、カメラ、ビデオカメラ、スキャナ、複写機、プリンタなど）の基本構成を示す図である。

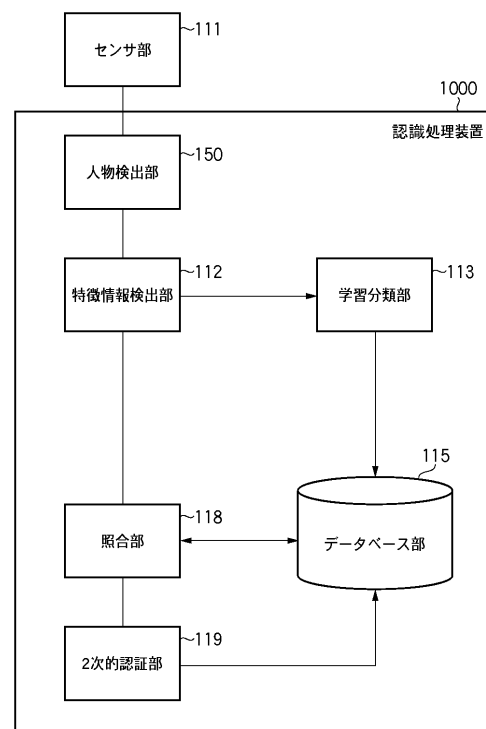
【図 9】本発明の第 1 の実施形態に係るセンサ部 111 の機能構成を示すブロック図である。

10

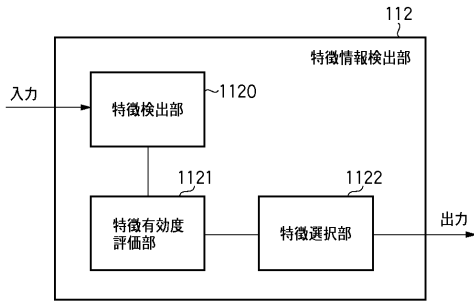
【図 1】



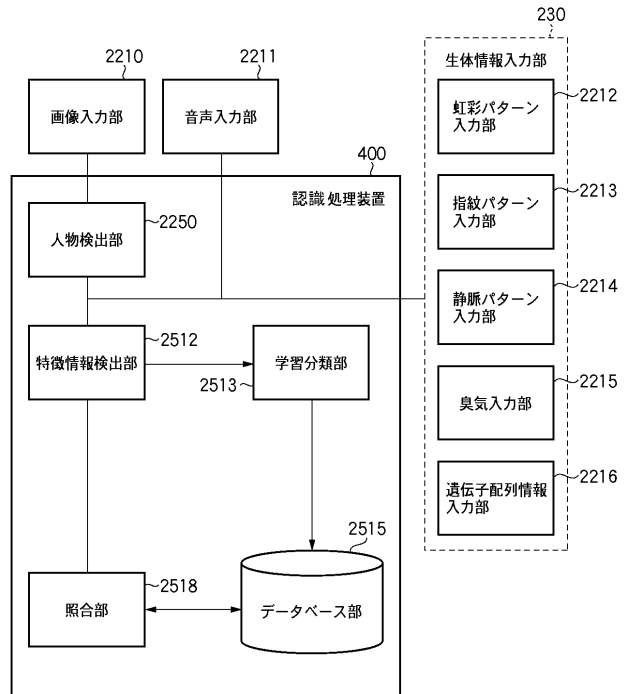
【図 2】



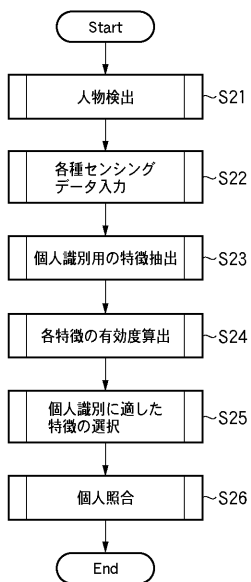
【図 3】



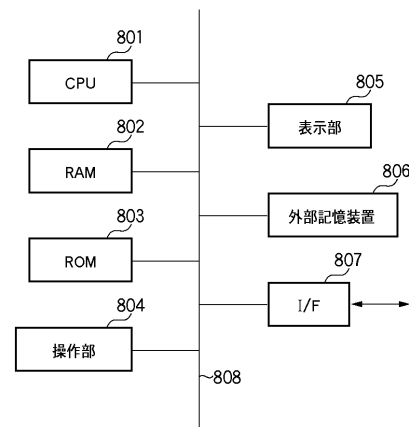
【図 4】



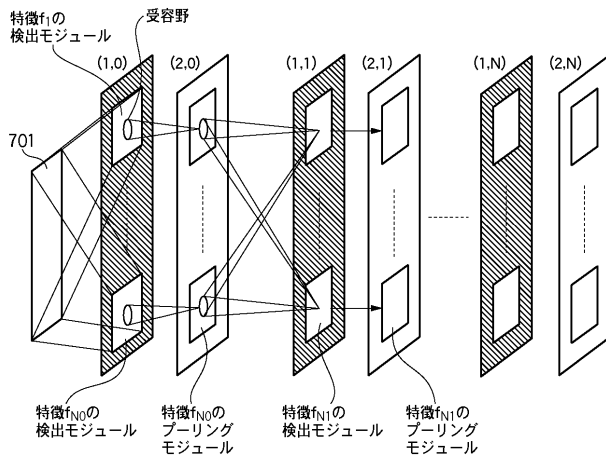
【図 5】



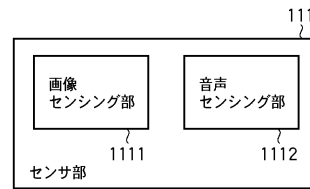
【図 6】



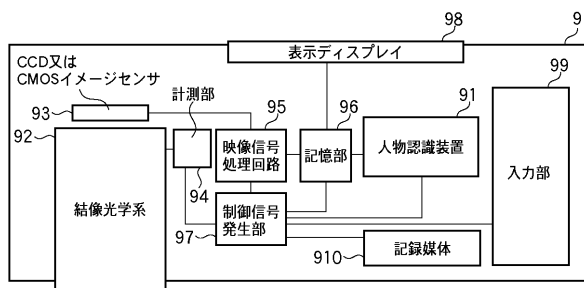
【図 7】



【図 9】



【図 8】



フロントページの続き

- (72)発明者 森 克彦
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 加藤 政美
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 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
- (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 御手洗 裕輔
東京都大田区下丸子3丁目30番2号 キヤノン株式会社内

Fターム(参考) 4C038 FF01 FF05 FG01 VA07 VB02 VB03 VB04 VB11 VB12 VB21
VB31 VC01 VC05
5B043 AA09 BA02 BA04 DA04 GA01 GA17
5B285 CB12 CB23