

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4590717号  
(P4590717)

(45) 発行日 平成22年12月1日(2010.12.1)

(24) 登録日 平成22年9月24日(2010.9.24)

(51) Int.Cl.	F I		
<b>G06T 7/00</b> (2006.01)	G06T 7/00	350B	
<b>A63H 11/00</b> (2006.01)	G06T 7/00	100C	
<b>B25J 5/00</b> (2006.01)	A63H 11/00	Z	
<b>B25J 13/00</b> (2006.01)	B25J 5/00	C	
<b>G06T 1/00</b> (2006.01)	B25J 13/00	Z	

請求項の数 9 (全 22 頁) 最終頁に続く

(21) 出願番号	特願2000-351897 (P2000-351897)	(73) 特許権者	000002185
(22) 出願日	平成12年11月17日(2000.11.17)		ソニー株式会社
(65) 公開番号	特開2002-157596 (P2002-157596A)		東京都港区港南1丁目7番1号
(43) 公開日	平成14年5月31日(2002.5.31)	(74) 代理人	100067736
審査請求日	平成19年3月9日(2007.3.9)		弁理士 小池 晃
		(74) 代理人	100086335
			弁理士 田村 榮一
		(74) 代理人	100096677
			弁理士 伊賀 誠司
		(72) 発明者	大久保 厚志
			東京都品川区北品川6丁目7番35号 ソニー株式会社内
		(72) 発明者	佐部 浩太郎
			東京都品川区北品川6丁目7番35号 ソニー株式会社内

最終頁に続く

(54) 【発明の名称】 顔識別装置及び顔識別方法

(57) 【特許請求の範囲】

【請求項1】

撮像手段と、

上記撮像手段により撮像された画像内の輝度パターンにより顔の顔データを検出する顔データ検出手段と、

上記顔データ検出手段が検出した顔データに含まれる色分布を元に上記撮像された画像内で変化する顔を追跡する顔追跡手段と、

上記顔データ検出手段が検出した上記顔データに基づいて、特定顔を識別する顔識別手段とを備え、

上記顔データ検出手段は、上記顔追跡手段による顔の追跡によって得られる顔の推定位置について顔データの検出処理を行い、

上記顔追跡手段は、色空間及び画像上位置の両方向からトラッキング制約をかけて追跡を行い、

上記顔識別手段は、上記特定顔との一致確率を用いた顔の識別結果を時系列的に蓄積し、蓄積された識別結果を統計的に処理することにより顔識別を行う

顔識別装置。

【請求項2】

上記顔データ検出手段は、顔と非顔とを識別するサポートベクターマシンである請求項1記載の顔識別装置。

【請求項3】

10

20

上記顔追跡手段は、画像内の顔の変化として、画像内の顔の移動或いは形の変化のうちの少なくとも1つを追跡する請求項1又は2記載の顔識別装置。

【請求項4】

上記顔追跡手段は、上記色分布を正規化して得た正規化分布により当該顔を追跡する請求項1乃至3のいずれか1項に記載の顔識別装置。

【請求項5】

予め得ている登録顔データと上記顔データ検出手段により得た顔データとを差分し、学習差分値を予め得て、上記学習差分値により特徴空間を形成する学習をする学習手段を備え、

上記顔識別手段は、予め得ている全ての登録顔データと上記顔データ検出手段が検出した上記顔データとを差分して得た検出差分値に基づいて特定顔を識別し、上記特徴空間における上記検出差分値に基づいて特定人物を識別する請求項1乃至4のいずれか1項に記載の顔識別装置。

10

【請求項6】

上記登録顔データがデータベースとして記憶される記憶手段と、画像撮像モードにより撮像した顔画像を上記データベースに登録顔データとして記憶する撮像モード撮像手段とを備える請求項1乃至5のいずれか1項に記載の顔識別装置。

【請求項7】

動物に模した外観形状を有し自律的に行動するロボット装置に内蔵され、

上記撮像手段は、上記動物の目の機能を果たす請求項1乃至6のいずれか1項に記載の顔識別装置

20

【請求項8】

コンピュータが、

撮像手段により撮像する撮像工程と、

上記撮像工程にて撮像された画像内の輝度パターンにより顔の顔データを検出する顔データ検出工程と、

上記顔データ検出工程にて検出した顔データに含まれる色分布を元に上記撮像された画像内で変化する顔を追跡する顔追跡工程と、

上記顔データ検出工程にて検出した上記顔データに基づいて、特定人物を識別する人物識別工程とを実行し、

30

上記顔データ検出工程では、上記顔追跡工程での顔の追跡によって得られる顔の推定位置について顔データの検出処理を行い、

上記顔追跡工程では、色空間及び画像上位置の両方向からトラッキング制約をかけて追跡を行い、

上記顔識別工程では、上記特定顔との一致確率を用いた顔の識別結果を時系列的に蓄積し、蓄積された識別結果を統計的に処理することにより顔識別を行う

顔識別方法。

【請求項9】

自律的に行動するロボット装置が上記撮像手段を備えている請求項8記載の顔識別方法。

40

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、顔識別装置及び顔識別方法に関し、詳しくは、自律的に行動し、且つ顔を識別することができるロボット装置に適用可能な顔識別装置及び顔識別方法に関する。

【0002】

【従来の技術】

顔認識に関する研究は、パターン認識理論を検証する上でベンチマーク的な位置付けを持っており、古くから多くの手法が開発されてきた。そこで想定されていたアプリケーションは、セキュリティシステムにおける認証や大規模データベースからの人物検索等、静的

50

な環境下で正確性を要するものがほとんどである。

【0003】

【発明が解決しようとする課題】

ところで、近年、外観形状が犬等の動物に模して形成され、エンターテインメントとしてのロボット装置が提供されている。このロボット装置は、外部からの情報（例えば、周囲環境の情報等）や内部の状態（例えば、感情状態等）等に応じて目や脚等を自律的に動作させることで、動物のような仕草を表出させている。

【0004】

このようなロボット装置の出現により、多少不正確でも動的に変化する環境下で一定時間内に応答できるようなヒューマンインターフェース技術が要求されてきており、その一つとしてロボット装置による顔識別が求められている。例えば、顔識別を利用すれば、ロボット装置は、多くの中からユーザ（飼い主）を識別することができ、よりエンターテインメント性のあるものとなる。

【0005】

こうしたロボット装置に搭載される顔識別のアプリケーションでは、ある与えられた1枚のシーンの中の人物を識別する問題に加え、次のような問題を解決する必要がある。

(1) ロボット装置自身が移動するため、環境の変化やその多様性を許容しなくてはならない。

【0006】

(2) 人間とロボット装置の位置関係も変化するため、インタラクション中には人間を視野に入れ続ける必要がある。

【0007】

(3) 数多くのシーン画像から人物の識別に使える画像を選び出し、総合的に判断しなくてはならない。

【0008】

(4) ある時間内に応答しなくてはならない。

【0009】

そこで、本発明は、上述の種々の問題を解決することができる顔識別装置及び顔識別方法を提供することを目的としている。

【0010】

【課題を解決するための手段】

本発明に係る顔識別装置は、上述の課題を解決するために、撮像手段と、上記撮像手段により撮像された画像内の輝度パターンにより顔の顔データを検出する顔データ検出手段と、上記顔データ検出手段が検出した顔データに含まれる色分布を元に上記撮像された画像内で変化する顔を追跡する顔追跡手段と、上記顔データ検出手段が検出した上記顔データに基づいて、特定顔を識別する顔識別手段とを備え、上記顔データ検出手段は、上記顔追跡手段による顔の追跡によって得られる顔の推定位置について顔データの検出処理を行い、上記顔追跡手段は、色空間及び画像上位置の両方向からトラッキング制約をかけて追跡を行い、上記顔識別手段は、上記特定顔との一致確率を用いた顔の識別結果を時系列的に蓄積し、蓄積された識別結果を統計的に処理することにより顔識別を行う。

【0011】

このロボット装置は、画像内で変化する顔を追跡しながら当該顔の顔データを検出して、検出した顔データに基づいて、特定顔を特定する。

【0012】

また、本発明に係る顔識別方法は、上述の課題を解決するために、コンピュータが、撮像手段により撮像する撮像工程と、上記撮像工程にて撮像された画像内の輝度パターンにより顔の顔データを検出する顔データ検出工程と、上記顔データ検出工程にて検出した顔データに含まれる色分布を元に上記撮像された画像内で変化する顔を追跡する顔追跡工程と、上記顔データ検出工程にて検出した上記顔データに基づいて、特定人物を識別する人物識別工程とを実行し、上記顔データ検出工程では、上記顔追跡工程での顔の追跡によ

10

20

30

40

50

て得られる顔の推定位置について顔データの検出処理を行い、上記顔追跡工程では、色空間及び画像上位置の両方向からトラッキング制約をかけて追跡を行い、上記顔識別工程では、上記特定顔との一致確率を用いた顔の識別結果を時系列的に蓄積し、蓄積された識別結果を統計的に処理することにより顔識別を行う。

【0013】

このような顔識別方法は、画像内で変化する顔を追跡しながら当該顔の顔データを検出して、検出した顔データに基づいて、特定顔を特定する。

【0014】

【発明の実施の形態】

以下、本発明の実施の形態について図面を用いて詳細に説明する。この実施の形態は、本発明を、外観形状が犬等の動物に模して形成されたロボット装置に適用したものである。ロボット装置は、外部からの情報（例えば、周囲環境の情報等）や内部の状態（例えば、感情状態等）等に応じて目や脚等を自律的に動作させることで、動物のような仕草を表出させている。そして、このロボット装置は、ユーザ（飼い主）等の顔を識別することができるものとして構成されている。

【0015】

実施の形態では、まず、ロボット装置の構成について説明して、その後、ロボット装置における本発明の適用部分とされる顔識別について詳細に説明する。

【0016】

（1）本実施の形態によるロボット装置の構成

図1に示すように、例えば「犬」を模した形状のいわゆるペットロボットとされ、胴体部ユニット2の前後左右にそれぞれ脚部ユニット3A、3B、3C、3Dが連結されると共に、胴体部ユニット2の前端部及び後端部にそれぞれ頭部ユニット4及び尻尾部ユニット5が連結されて構成されている。

【0017】

胴体部ユニット2には、図2に示すように、CPU（Central Processing Unit）10、DRAM（Dynamic Random Access Memory）11、フラッシュROM（Read Only Memory）12、PC（Personal Computer）カードインターフェース回路13及び信号処理回路14が内部バス15を介して相互に接続されることにより形成されたコントロール部16と、このロボット装置1の動力源としてのバッテリー17とが収納されている。また、胴体部ユニット2には、ロボット装置1の向きや動きの加速度を検出するための角速度センサ18及び加速度センサ19なども収納されている。

【0018】

また、頭部ユニット4には、外部の状況を撮像するためのCCD（Charge Coupled Device）カメラ20と、使用者からの「撫でる」や「叩く」といった物理的な働きかけにより受けた圧力を検出するためのタッチセンサ21と、前方に位置する物体までの距離を測定するための距離センサ22と、外部音を集音するためのマイクロホン23と、鳴き声等の音声を出力するためのスピーカ24と、ロボット装置1の「目」に相当するLED（Light Emitting Diode）（図示せず）となどがそれぞれ所定位置に配置されている。ここで、CCDカメラ20は、顔識別において顔画像を撮像する撮像手段を構成する。

【0019】

さらに、各脚部ユニット3A～3Dの関節部分や各脚部ユニット3A～3D及び胴体部ユニット2の各連結部分、頭部ユニット4及び胴体部ユニット2の連結部分、並びに尻尾部ユニット5の尻尾5Aの連結部分などにはそれぞれ自由度数分のアクチュエータ $25_1 \sim 25_n$ 及びポテンシオメータ $26_1 \sim 26_n$ が配設されている。例えば、アクチュエータ $25_1 \sim 25_n$ はサーボモータを構成として有している。サーボモータの駆動により、脚部ユニット3A～3Dが制御されて、目標の姿勢或いは動作に遷移する。

【0020】

そして、これら角速度センサ18、加速度センサ19、タッチセンサ21、距離センサ22、マイクロホン23、スピーカ24及び各ポテンシオメータ $26_1 \sim 26_n$ などの各種

10

20

30

40

50

センサ並びにLED及び各アクチュエータ $25_1 \sim 25_n$ は、それぞれ対応するハブ $27_1 \sim 27_n$ を介してコントロール部16の信号処理回路14と接続され、CCDカメラ20及びバッテリー17は、それぞれ信号処理回路14と直接接続されている。

【0021】

信号処理回路14は、上述の各センサから供給されるセンサデータや画像データ及び音声データを順次取り込み、これらをそれぞれ内部バス15を介してDRAM11内の所定位置に順次格納する。また信号処理回路14は、これと共にバッテリー17から供給されるバッテリー残量を表すバッテリー残量データを順次取り込み、これをDRAM11内の所定位置に格納する。

【0022】

このようにしてDRAM11に格納された各センサデータ、画像データ、音声データ及びバッテリー残量データは、この後CPU10がこのロボット装置1の動作制御を行う際に利用される。

10

【0023】

實際上CPU10は、ロボット装置1の電源が投入された初期時、胴体部ユニット2の図示しないPCカードスロットに装填されたメモリカード28又はフラッシュROM12に格納された制御プログラムをPCカードインターフェース回路13を介して又は直接読み出し、これをDRAM11に格納する。

【0024】

また、CPU10は、この後上述のように信号処理回路14よりDRAM11に順次格納される各センサデータ、画像データ、音声データ及びバッテリー残量データに基づいて自己及び周囲の状況や、使用者からの指示及び働きかけの有無などを判断する。

20

【0025】

さらに、CPU10は、この判断結果及びDRAM11に格納した制御プログラムに基づいて続く行動を決定すると共に、当該決定結果に基づいて必要なアクチュエータ $25_1 \sim 25_n$ を駆動させることにより、頭部ユニット4を上下左右に振らせたり、尻尾部ユニット5の尻尾5Aを動かせたり、各脚部ユニット3A~3Dを駆動させて歩行させるなどの行動を行わせる。

【0026】

また、この際CPU10は、必要に応じて音声データを生成し、これを信号処理回路14を介して音声信号としてスピーカ24に与えることにより当該音声信号に基づく音声を外部に出力させたり、上述のLEDを点灯、消灯又は点滅させる。

30

【0027】

このようにしてこのロボット装置1においては、自己及び周囲の状況や、使用者からの指示及び働きかけに応じて自律的に行動し得るようになされている。

【0028】

(2) 制御プログラムのソフトウェア構成

ここで、ロボット装置1における上述の制御プログラムのソフトウェア構成は、図3に示すようになる。この図3において、デバイス・ドライバ・レイヤ30は、この制御プログラムの最下位層に位置し、複数のデバイス・ドライバからなるデバイス・ドライバ・セット31から構成されている。この場合、各デバイス・ドライバは、CCDカメラ20(図2)やタイマ等の通常のコンピュータで用いられるハードウェアに直接アクセスすることを許されたオブジェクトであり、対応するハードウェアからの割り込みを受けて処理を行う。

40

【0029】

また、ロボティック・サーバ・オブジェクト32は、デバイス・ドライバ・レイヤ30の最下位層に位置し、例えば上述の各種センサやアクチュエータ $25_1 \sim 25_n$ 等のハードウェアにアクセスするためのインターフェースを提供するソフトウェア群でなるバーチャル・ロボット33と、電源の切換えなどを管理するソフトウェア群でなるパワーマネージャ34と、他の種々のデバイス・ドライバを管理するソフトウェア群でなるデバイス・ド

50

ライバ・マネージャ 35 と、ロボット装置 1 の機構を管理するソフトウェア群でなるデザイン・ロボット 36 とから構成されている。

【0030】

マネージャ・オブジェクト 37 は、オブジェクト・マネージャ 38 及びサービス・マネージャ 39 から構成されている。オブジェクト・マネージャ 38 は、ロボティック・サーバ・オブジェクト 32、ミドル・ウェア・レイヤ 40、及びアプリケーション・レイヤ 41 に含まれる各ソフトウェア群の起動や終了を管理するソフトウェア群であり、サービス・マネージャ 39 は、メモリカード 28 (図 2) に格納された接続ファイルに記述されている各オブジェクト間の接続情報に基づいて各オブジェクトの接続を管理するソフトウェア群である。

10

【0031】

ミドル・ウェア・レイヤ 40 は、ロボティック・サーバ・オブジェクト 32 の上位層に位置し、画像処理や音声処理などのこのロボット装置 1 の基本的な機能を提供するソフトウェア群から構成されている。また、アプリケーション・レイヤ 41 は、ミドル・ウェア・レイヤ 40 の上位層に位置し、当該ミドル・ウェア・レイヤ 40 を構成する各ソフトウェア群によって処理された処理結果に基づいてロボット装置 1 の行動を決定するためのソフトウェア群から構成されている。

【0032】

なお、ミドル・ウェア・レイヤ 40 及びアプリケーション・レイヤ 41 の具体的なソフトウェア構成をそれぞれ図 4 に示す。

20

【0033】

ミドル・ウェア・レイヤ 40 は、図 4 に示すように、騒音検出用、温度検出用、明るさ検出用、音階認識用、距離検出用、姿勢検出用、タッチセンサ用、動き検出用及び色認識用の各信号処理モジュール 50 ~ 58 並びに入力セマンティクスコンバータモジュール 59 などを有する認識系 60 と、出力セマンティクスコンバータモジュール 68 並びに姿勢管理用、トラッキング用、モーション再生用、歩行用、転倒復帰用、LED 点灯用及び音再生用の各信号処理モジュール 61 ~ 67 などを有する出力系 69 とから構成されている。

【0034】

認識系 60 の各信号処理モジュール 50 ~ 58 は、ロボティック・サーバ・オブジェクト 32 のバーチャル・ロボット 33 により DRAM 11 (図 2) から読み出される各センサデータや画像データ及び音声データのうちの対応するデータを取り込み、当該データに基づいて所定の処理を施して、処理結果を入力セマンティクスコンバータモジュール 59 に与える。ここで、例えば、バーチャル・ロボット 33 は、所定の通信規約によって、信号の授受或いは変換をする部分として構成されている。

30

【0035】

入力セマンティクスコンバータモジュール 59 は、これら各信号処理モジュール 50 ~ 58 から与えられる処理結果に基づいて、「うるさい」、「暑い」、「明るい」、「ボールを検出した」、「転倒を検出した」、「撫でられた」、「叩かれた」、「ドミソの音階が聞こえた」、「動く物体を検出した」又は「障害物を検出した」などの自己及び周囲の状況や、使用者からの指令及び働きかけを認識し、認識結果をアプリケーション・レイヤ 41 (図 2) に出力する。

40

【0036】

アプリケーション・レイヤ 41 は、図 5 に示すように、行動モデルライブラリ 70、行動切換えモジュール 71、学習モジュール 72、感情モデル 73 及び本能モデル 74 の 5 つのモジュールから構成されている。

【0037】

行動モデルライブラリ 70 には、図 6 に示すように、「バッテリー残量が少なくなった場合」、「転倒復帰する」、「障害物を回避する場合」、「感情を表現する場合」、「ボールを検出した場合」などの予め選択されたいくつかの条件項目にそれぞれ対応させて、それぞれ独立した行動モデル  $70_1 \sim 70_n$  が設けられている。

50

## 【0038】

そして、これら行動モデル70<sub>1</sub>～70<sub>n</sub>は、それぞれ入力セマンティクスコンバータモジュール59から認識結果が与えられたときや、最後の認識結果が与えられてから一定時間が経過したときなどに、必要に応じて後述のように感情モデル73に保持されている対応する情動のパラメータ値や、本能モデル74に保持されている対応する欲求のパラメータ値を参照しながら続く行動をそれぞれ決定し、決定結果を行動切換えモジュール71に出力する。

## 【0039】

なお、この実施の形態の場合、各行動モデル70<sub>1</sub>～70<sub>n</sub>は、次の行動を決定する手法として、図7に示すような1つのノード(状態)NODE<sub>0</sub>～NODE<sub>n</sub>から他のどのノードNODE<sub>0</sub>～NODE<sub>n</sub>に遷移するかを各ノードNODE<sub>0</sub>～NODE<sub>n</sub>に間を接続するアークARC<sub>1</sub>～ARC<sub>n</sub>に対してそれぞれ設定された遷移確率P<sub>1</sub>～P<sub>n</sub>に基づいて確率的に決定する有限確率オートマトンと呼ばれるアルゴリズムを用いる。

10

## 【0040】

具体的に、各行動モデル70<sub>1</sub>～70<sub>n</sub>は、それぞれ自己の行動モデル70<sub>1</sub>～70<sub>n</sub>を形成するノードNODE<sub>0</sub>～NODE<sub>n</sub>にそれぞれ対応させて、これらノードNODE<sub>0</sub>～NODE<sub>n</sub>ごとに図8に示すような状態遷移表80を有している。

## 【0041】

この状態遷移表80では、そのノードNODE<sub>0</sub>～NODE<sub>n</sub>において遷移条件とする入力イベント(認識結果)が「入力イベント名」の行に優先順に列記され、その遷移条件についてのさらなる条件が「データ名」及び「データ範囲」の行における対応する列に記述されている。

20

## 【0042】

したがって、図8の状態遷移表80で表されるノードNODE<sub>100</sub>では、「ボールを検出(BALL)」という認識結果が与えられた場合に、当該認識結果と共に与えられるそのボールの「大きさ(SIZE)」が「0から1000」の範囲であることや、「障害物を検出(OBSTACLE)」という認識結果が与えられた場合に、当該認識結果と共に与えられるその障害物までの「距離(DISTANCE)」が「0から100」の範囲であることが他のノードに遷移するための条件となっている。

## 【0043】

また、このノードNODE<sub>100</sub>では、認識結果の入力がない場合においても、行動モデル70<sub>1</sub>～70<sub>n</sub>が周期的に参照する感情モデル73及び本能モデル74にそれぞれ保持された各情動及び各欲求のパラメータ値のうち、感情モデル73に保持された「喜び(JOY)」、「驚き(SURPRISE)」若しくは「悲しみ(SADNESS)」のいずれかのパラメータ値が「50から100」の範囲であるときには他のノードに遷移することができるようになっている。

30

## 【0044】

また、状態遷移表80では、「他のノードへの遷移確率」の欄における「遷移先ノード」の列にそのノードNODE<sub>0</sub>～NODE<sub>n</sub>から遷移できるノード名が列記されていると共に、「入力イベント名」、「データ値」及び「データの範囲」の行に記述された全ての条件が揃ったときに遷移できる他の各ノードNODE<sub>0</sub>～NODE<sub>n</sub>への遷移確率が「他のノードへの遷移確率」の欄内の対応する箇所それぞれに記述され、そのノードNODE<sub>0</sub>～NODE<sub>n</sub>に遷移する際に出すべき行動が「他のノードへの遷移確率」の欄における「出力行動」の行に記述されている。なお、「他のノードへの遷移確率」の欄における各行の確率の和は100[%]となっている。

40

## 【0045】

したがって、図8の状態遷移表80で表されるノードNODE<sub>100</sub>では、例えば「ボールを検出(BALL)」し、そのボールの「SIZE(大きさ)」が「0から1000」の範囲であるという認識結果が与えられた場合には、「30[%]」の確率で「ノードNODE<sub>120</sub>(node 120)」に遷移でき、そのとき「ACTION1」の行動が出力されること

50

となる。

【0046】

各行動モデル $70_1 \sim 70_n$ は、それぞれこのような状態遷移表80として記述されたノード $NODE_0 \sim NODE_n$ がいくつも繋がるようにして構成されており、入力セマンティクスコンバータモジュール59から認識結果が与えられたときなどに、対応するノード $NODE_0 \sim NODE_n$ の状態遷移表を利用して確率的に次の行動を決定し、決定結果を行動切換えモジュール71に出力するようになされている。

【0047】

図5に示す行動切換えモジュール71は、行動モデルライブラリ70の各行動モデル $70_1 \sim 70_n$ からそれぞれ出力される行動のうち、予め定められた優先順位の高い行動モデル $70_1 \sim 70_n$ から出力された行動を選択し、当該行動を実行すべき旨のコマンド(以下、これを行動コマンドという。)をミドル・ウェア・レイヤ40の出力セマンティクスコンバータモジュール68に送付する。なお、この実施の形態においては、図6において下側に表記された行動モデル $70_1 \sim 70_n$ ほど優先順位が高く設定されている。

10

【0048】

また、行動切換えモジュール71は、行動完了後に出力セマンティクスコンバータモジュール68から与えられる行動完了情報に基づいて、その行動が完了したことを学習モジュール72、感情モデル73及び本能モデル74に通知する。

【0049】

一方、学習モジュール72は、入力セマンティクスコンバータモジュール59から与えられる認識結果のうち、「叩かれた」や「撫でられた」など、使用者からの働きかけとして受けた教示の認識結果を入力する。

20

【0050】

そして、学習モジュール72は、この認識結果及び行動切換えモジュール71からの通知に基づいて、「叩かれた(叱られた)」ときにはその行動の発現確率を低下させ、「撫でられた(誉められた)」ときにはその行動の発現確率を上昇させるように、行動モデルライブラリ70における対応する行動モデル $70_1 \sim 70_n$ の対応する遷移確率を変更する。

【0051】

他方、感情モデル73は、「喜び(joy)」、「悲しみ(sadness)」、「怒り(anger)」、「驚き(surprise)」、「嫌悪(disgust)」及び「恐れ(fear)」の合計6つの情動について、各情動ごとにその情動の強さを表すパラメータを保持している。そして、感情モデル73は、これら各情動のパラメータ値を、それぞれ入力セマンティクスコンバータモジュール59から与えられる「叩かれた」及び「撫でられた」などの特定の認識結果と、経過時間及び行動切換えモジュール71からの通知となどに基づいて周期的に更新する。

30

【0052】

具体的には、感情モデル73は、入力セマンティクスコンバータモジュール59から与えられる認識結果と、そのときのロボット装置1の行動と、前回更新してからの経過時間となどに基づいて所定の演算式により算出されるそのときのその情動の変動量を $E[t]$ 、現在のその情動のパラメータ値を $E[t]$ 、その情動の感度を表す係数を $k_0$ として、(1)式によって次の周期におけるその情動のパラメータ値 $E[t+1]$ を算出し、これを現在のその情動のパラメータ値 $E[t]$ と置き換えるようにしてその情動のパラメータ値を更新する。また、感情モデル73は、これと同様にして全ての情動のパラメータ値を更新する。

40

【0053】

【数1】

$$E[t+1] = E[t] + ke \times \Delta E[t] \quad \dots (1)$$

10

## 【0054】

なお、各認識結果や出力セマンティクスコンバータモジュール68からの通知が各情動のパラメータ値の変動量  $E[t]$  にどの程度の影響を与えるかは予め決められており、例えば「叩かれた」といった認識結果は「怒り」の情動のパラメータ値の変動量  $E[t]$  に大きな影響を与え、「撫でられた」といった認識結果は「喜び」の情動のパラメータ値の変動量  $E[t]$  に大きな影響を与えるようになっている。

## 【0055】

ここで、出力セマンティクスコンバータモジュール68からの通知とは、いわゆる行動のフィードバック情報（行動完了情報）であり、行動の出現結果の情報であり、感情モデル73は、このような情報によっても感情を変化させる。これは、例えば、「吠える」といった行動により怒りの感情レベルが下がるといったようなことである。なお、出力セマンティクスコンバータモジュール68からの通知は、上述した学習モジュール72にも入力されており、学習モジュール72は、その通知に基づいて行動モデル70<sub>1</sub> ~ 70<sub>n</sub>の対応する遷移確率を変更する。

20

## 【0056】

なお、行動結果のフィードバックは、行動切換えモジュレータ71の出力（感情が付加された行動）によりなされるものであってもよい。

30

## 【0057】

一方、本能モデル74は、「運動欲(exercise)」、「愛情欲(affection)」、「食欲(appetite)」及び「好奇心(curiosity)」の互いに独立した4つの欲求について、これら欲求ごとにその欲求の強さを表すパラメータを保持している。そして、本能モデル74は、これらの欲求のパラメータ値を、それぞれ入力セマンティクスコンバータモジュール59から与えられる認識結果や、経過時間及び行動切換えモジュール71からの通知などに基づいて周期的に更新する。

## 【0058】

具体的には、本能モデル74は、「運動欲」、「愛情欲」及び「好奇心」については、認識結果、経過時間及び出力セマンティクスコンバータモジュール68からの通知などに基づいて所定の演算式により算出されるそのときのその欲求の変動量を  $I[k]$ 、現在のその欲求のパラメータ値を  $I[k]$ 、その欲求の感度を表す係数  $k_i$  として、所定周期で(2)式を用いて次の周期におけるその欲求のパラメータ値  $I[k+1]$  を算出し、この演算結果を現在のその欲求のパラメータ値  $I[k]$  と置き換えるようにしてその欲求のパラメータ値を更新する。また、本能モデル74は、これと同様にして「食欲」を除く各欲求のパラメータ値を更新する。

40

## 【0059】

## 【数2】

$$I[k+1] = I[k] + k_i \times \Delta I[k] \quad \dots (2)$$

10

## 【0060】

なお、認識結果及び出力セマンティクスコンバータモジュール68からの通知などが各欲求のパラメータ値の変動量  $I[k]$  にどの程度の影響を与えるかは予め決められており、例えば出力セマンティクスコンバータモジュール68からの通知は、「疲れ」のパラメータ値の変動量  $I[k]$  に大きな影響を与えるようになっている。

## 【0061】

なお、本実施の形態においては、各情動及び各欲求（本能）のパラメータ値がそれぞれ0から100までの範囲で変動するように規制されており、また係数  $k_e$ 、 $k_i$  の値も各情動及び各欲求ごとに個別に設定されている。

20

## 【0062】

一方、ミドル・ウェア・レイヤ40の出力セマンティクスコンバータモジュール68は、図4に示すように、上述のようにしてアプリケーション・レイヤ41の行動切換えモジュール71から与えられる「前進」、「喜ぶ」、「鳴く」又は「トラッキング（ボールを追いかける）」といった抽象的な行動コマンドを出力系69の対応する信号処理モジュール61～67に与える。

## 【0063】

そしてこれら信号処理モジュール61～67は、行動コマンドが与えられると当該行動コマンドに基づいて、その行動を行うために対応するアクチュエータ25<sub>1</sub>～25<sub>n</sub>（図2）に与えるべきサーボ指令値や、スピーカ24（図2）から出力する音の音声データ及び又は「目」のLEDに与える駆動データを生成し、これらのデータをロボティック・サーバ・オブジェクト32のバーチャル・ロボット33及び信号処理回路14（図2）を順次介して対応するアクチュエータ25<sub>1</sub>～25<sub>n</sub>又はスピーカ24又はLEDに順次送出する。

30

## 【0064】

このようにしてロボット装置1においては、制御プログラムに基づいて、自己（内部）及び周囲（外部）の状況や、使用者からの指示及び働きかけに応じた自律的な行動を行うことができるようになっている。

40

## 【0065】

（3）ロボット装置による顔識別

（3-1）ロボット装置に適用される顔識別の原理

上述したように、ロボット装置1は、外部からの情報や内部の状態に応じて自律的に行動する。このようなロボット装置1において、従来における顔識別の問題点の解決と、顔識別のタスクを遂行とを、次の3つの技術により実現している。

## 【0066】

（i）複雑なシーンからの顔の検出

（ii）顔の実時間トラッキング

（iii）顔の識別

50

顔の検出手法は、対象物の識別に色、動き及びパターンを使うものに大別できるが、複雑なシーンの中から精度良く顔を切り出すためには顔のパターンを使うのが最も高性能である。しかしながら、シーン全体にわたり全てのスケールの顔を探索するのは、非常に処理が重いため、従来より、この手法は静止画に対してしか用いられていない。

【0067】

その一方で、実時間で顔を検出するシステムのほとんどは肌色の検出を行っている。しかし、色は照明条件により変化してしまうし、肌色にも人種や個人差があるため、単純な肌色認識のみでは有効な手段となり得ない。

【0068】

そこで、本発明の適用により、検出した顔パターンに含まれる色分布を元に顔の実時間トラッキングを行い、その動的変化に顔検出を適応していく手法を取る。また、推定した色分布から求めた顔領域に対してのみ、顔パターンの探索を行う。これにより、顔検出における演算時間の短縮を図る。

【0069】

さらに、顔の識別には、パターン探索により切り出した顔画像を用いて行う。そして、トラッキングに成功している間は同じ顔の識別結果として扱うことで複数の識別結果から総合的な判断を下すことを可能としている。

【0070】

例えば、図9に示すように、ロボット装置1において行う顔識別のための処理は、(i) 複雑なシーンからの顔の検出を、輝度パターンによる顔の検出(顔認識)により行い、(ii) 顔の実時間トラッキングを、色による顔の追跡(顔のトラッキング)により行い、そして、(iii) 顔の識別を、差分顔を利用した人物識別を行っている。

【0071】

例えば、ロボット装置1におけるこのような各処理は、モジュール或いはオブジェクトとして実現される。すなわち、ロボット装置1は、図9に示すように、顔の追跡モジュール $M_2$ 、顔の検出モジュール $M_1$ 及び顔の識別モジュール $M_3$ を備える。ここで、顔の追跡モジュール $M_2$ は、CCDカメラ20に撮像された画像内で変化する顔を追跡する顔追跡手段として機能し、顔の検出モジュール $M_1$ は、顔の追跡モジュール $M_2$ による顔の追跡情報に基づいて、撮像手段により撮像された画像内の顔の顔データを検出する顔データ検出手段として機能し、顔の識別モジュール $M_3$ は、顔の検出モジュール $M_1$ が検出した顔データに基づいて、特定顔を識別する顔識別手段として機能する。

【0072】

ここで、輝度パターンによる顔の検出では、入力画像中から顔を検出(認識)する処理を行う。具体的には、この顔の検出では、サポートベクターマシン(SVM)による顔、非顔の識別を行う。この処理は、通常、環境変化に強く、計算量が多く、姿勢変化に弱いといった特徴がある。ここで、環境変化としては、例えば、周囲の照明の変化が挙げられる。

【0073】

また、色による顔の追跡では、入力画像中の顔を追跡する処理を行う。具体的には、この顔の追跡では、顔の色分布の推定、顔領域の推定を行う。この処理は、通常、環境変化に弱く、計算量が少なく、姿勢変化に強いといった特徴がある。

【0074】

また、人物識別では、上述の顔の検出により認識された顔を特定の顔として識別する処理を行う。具体的には、この人物識別では、目、鼻の位置同定から位置合わせ(モーフィング)を行い、差分顔PCAから同一人物判定を行う。

【0075】

顔識別のシステムでは、以上のような処理を顔識別における各工程として適宜分担させ、相互に補完しあう関係とすることで、精度を高くした顔検出を可能としている。例えば、次のように各処理が補完しあう関係とされている。

【0076】

10

20

30

40

50

例えば、図9に示すように、色による顔の追跡では環境変化に弱い、輝度パターンによる顔の検出では環境に強いことを利用することで補完している。逆に、輝度パターンによる顔の検出では計算量が多く、姿勢変化に弱い、色による顔の追跡が計算量が少なく、姿勢変化に強いことを利用することが補完している。

【0077】

すなわち、概略すれば次のように言える。もともと計算量が多い処理とされる顔の検出を実時間で行うことは困難である。しかし、所定のタイミングにより一定期間行うとすれば、計算量の負担は軽減される。一方で、各タイミングにおいて入力画像内から顔の位置まで毎回検出したのでは、負担が大きい。

【0078】

そこで、計算量が少なく、姿勢変化に強い処理を利用して、入力画像内の顔の変化を実時間で追跡して、この追跡結果から得られる入力画像内の顔の推定位置についてだけ顔を検出する処理をすることとすれば、顔の位置を特定した状態での顔の検出が可能になる。すなわち、大雑把ではあるが早い処理と、信頼性は高いが遅い処理とを組み合わせ、役割を分担することで、システム全体においては各処理相互間で補完させ、これにより、協調して実時間による顔検出を可能としている。

【0079】

これにより、多くの顔検出結果を短時間で獲得でき、この獲得した顔検出結果に基づいて顔の識別を行い、そのような処理を統計的に処理することで、高精度の顔識別を可能にしている。

【0080】

ロボット装置1は、このような顔識別のシステムにより、シーン中から人間を見つけ出し(顔検出処理)、それを注視し(顔のトラッキング処理)、そこから得られた情報を用いた顔の識別による人物の特定(顔識別処理)に至るまで、全ての処理を自動的に行い、信頼性の高い顔識別を実現している。これによりロボット装置と人間との間で自然なインタラクションが可能になる。以下、顔の検出、顔の追跡及び顔の識別の各処理について具体的に説明する。

【0081】

(3-2) 顔の検出

顔の検出については、パターン認識の分野で現在最も学習汎化能力が高いと期待され、注目を集めているサポートベクターマシン(SVM)を用いて顔と非顔の識別を行っている。

【0082】

このような技術としては、B.Shoikopfらの報告(B.Shoikopf, C.Burges, A.Smola, "Advance in Kernel Methods Support Vector Learning", The MIT Press, 1999.)やV.Vapnikの報告(V.Vapnik, "The Nature of Statistical Learning Theory Second Edition", Springer, 1999.)に開示されている技術が挙げられる。本願発明者が行った予備実験の結果からは、この手法が主成分分析やニューラルネットワークを用いる手法と比べ、良好な結果を示すことが分かっている。

【0083】

上述のSVMは、識別関数に線形識別器(パーセプトロン)を用いた学習機械であり、カーネル関数を使うことで非線形空間に拡張することができる。また、識別関数の学習では、クラス間分離のマージンを最大にとるように行われ、その解は二次数理計画法を解くことで得られるため、グローバル解に到達できることを理論的に保証できる。この手法の詳細は次の通りである。入力画像は以下の手順で作成する。

【0084】

まず、シーンから20×20画素の画像を切り出し、さらに、切り出した画像から顔以外の背景部分を除くために画像の4隅をマスクする。次に、照明による撮像対象物の輝度が変化することを前提として、照明による輝度の勾配を補正して、ヒストグラム平滑化或いはコントラストノーマライゼーションを施す。

10

20

30

40

50

## 【 0 0 8 5 】

次に、顔検出に使用する識別関数の学習を行う。学習用のデータとして最初に顔データ、非顔データそれぞれ所定の枚数（例えば、30枚）を用いて、先ず、暫定的な識別関数を得る。それから、このようにして暫定的に得られた識別関数を様々なデータベース上の画像に試して顔の検出を行い、その結果、検出に成功したものを顔データ、失敗したものを非顔データとして学習データに追加し、さらに学習をし直す。

## 【 0 0 8 6 】

例えば、実施例としては、数百枚の画像に対してこの処理を繰り返し、最終的に顔データ389枚、非顔データ576枚を蓄積することができた。また、これらの顔データは全員ほぼ正面を向いていること以外は、人種、表情、髪型、眼鏡、ひげ、照明条件等の違いを含んでおり、これにより、これらの条件に左右されない識別器（例えば、識別モジュール）の学習が期待できる。この結果、得られた所定のサポートベクター、例えば255本のサポートベクターを用いる。

10

## 【 0 0 8 7 】

以上のような顔の検出は、後で詳述する顔の追跡処理により追従されている画像内の顔部分と推定される領域についてその処理を実施する。すなわち、顔の検出では、顔の追跡処理により限定された検索領域に対してのみ処理を実施する。これにより、信頼性の高い顔の検出処理を短時間で実行可能としている。

## 【 0 0 8 8 】

また、この顔の検出処理は、上述したように、ロボット装置1においてモジュール或いはオブジェクトとして構成される。

20

## 【 0 0 8 9 】

## ( 3 - 3 ) 顔の追跡

顔の追跡は、画像中で変化（特に移動）する顔を連続的に追従して、上述の顔の検出において顔の検出のための処理がなされる検索領域を限定するためのものである。特に、自律的に行動するロボット装置1のような場合には、画像内における顔の位置等が変化することから、このような画像内の顔の追従処理は顔識別において効果的に作用する。

## 【 0 0 9 0 】

この顔の追跡では、演算量の少なさ、顔の姿勢変化やオクルージョンに対する強さを考慮し、色領域の追跡をベースに行う。

30

## 【 0 0 9 1 】

例えば、Yangらの報告（J.Yang, W.Lu, A.Waibel, "Skin-Color Modeling and Adaptation", Technical Report of CMU, CMU-CS-97-146, 1997.）によれば、肌色の分布は輝度で正規化された（3）式及び（4）式により得られる（ $r, g$ ）平面上で、人物や人種に依らず正規分布に近い分布となることが実験的に知られている。

## 【 0 0 9 2 】

## 【 数 3 】

40

$$r = \frac{R}{R + G + B} \quad \dots (3)$$

## 【 0 0 9 3 】

## 【 数 4 】

$$g = \frac{G}{R + G + B} \quad \dots (4)$$

10

【 0 0 9 4 】

また、上述のYangらはこの色空間において肌色を表す正規分布のパラメータ推定値を、過去に測定されたデータの線形和の(5)式及び(6)式として表現し、重み係数  $\alpha_j$  ,  $\beta_j$  の更新規則を最尤推定により導出している。

【 0 0 9 5 】

【 数 5 】

$$\hat{\mu} = \sum_{j=1}^r \alpha_j m_j \quad \dots (5)$$

20

【 0 0 9 6 】

【 数 6 】

30

$$\hat{\sigma} = \sum_{j=1}^r \beta_j S_j \quad \dots (6)$$

【 0 0 9 7 】

ここで、(5)式は平均であり、(6)式は分散である。また、 $m_j$  ,  $s_j$  はそれぞれ、過去の  $j$  ステップ前の平均値と分散値である。また、 $r$  は最大履歴数である。

【 0 0 9 8 】

本手法では、この定式化に基づいて肌色領域の適応的トラッキングを行う。そして、本手法では、色モデルを用いて抽出した画素に対して画像領域上でセグメントを行う。具体的には、色モデルから得られた画素に対して、分布データを取得するための領域を指定する楕円の大きさ、楕円の位置、楕円の傾き等を変化させ、前フレーム中の顔の位置近傍で最も良く適合する楕円を求める。そして、求めた楕円内の画素を現在の肌色の分布データとして取得する。ここで、分布データを取得するための楕円について、その大きさ、位置及び傾きを変化させるのは、例えば、顔までの距離が変化したこと、顔が移動したこと、顔が傾いたこと等に対応するためである。

40

【 0 0 9 9 】

図 10 には、色空間と画像上位置の両方向からトラッキング制約をかける場合の処理例を

50

示している。顔領域が推定される画像平面における顔領域の色分布を、色空間上に写像する。そして、この色空間において、色の分布推定によるトラッキングを行い、色分布の時間変化を推定して、推定色を使って、画像平面上において顔領域を推定する。

#### 【0100】

このように色空間と画像上位置の両方向からトラッキング制約をかけることで、シーン中に似た色がある場合や照明条件が変動する場合でも比較的安定に顔の肌色を追跡できるようになる。

#### 【0101】

以上のような顔の追跡によって、画像内で移動等する顔を追従することができる。顔の検出では、この顔の追跡によって特定される領域についてその処理を実施しデータを得ている。このように、追従している領域に限定して顔の検出のための処理を実施することができるので、信頼性のある処理内容であっても計算量を少なくして実施することができ、短時間で処理結果を得ることができるようになる。

10

#### 【0102】

なお、この顔の追跡処理は、上述したように、ロボット装置1においてモジュール或いはオブジェクトとして構成される。

#### 【0103】

##### (3-4) 顔の識別

顔の識別では、上述の顔の検出によって得られた顔データを使用して、特定の顔を識別している。具体的には、この顔の識別では、与えられた2枚の顔画像に対してその画像中の人物が同一人物か、他人であるかの判定をすることで実現している。その判定については、2枚の顔画像の差分画像に基づいて行っている。なお、この顔の識別処理は、上述したように、ロボット装置1においてモジュール或いはオブジェクトとして構成される。

20

#### 【0104】

具体的には、顔の識別では、先ず、同一人物の2枚の顔画像の差分を学習段階として予め求めており、これを参照して、2枚の顔画像の差分結果から同一人物か別の人物であるかを判定している。

#### 【0105】

その学習段階では、同一人物の2枚の顔画像の差分を求め、同一人物の差分画像群 (intra-personal class) を生成し、これを固有値展開することで特徴空間とされる intra-personal class の固有空間を獲得している。例えば、このような学習手法の技術としては、B. Moghaddamらの報告 (B. Moghaddam, W. Wahid, A. Pentland, "Beyond Eigenfaces: Probabilistic Matching for Face Recognition", MIT Media Lab. Perceptual Computing Section Technical Report No.443, 1998.) に開示されている技術が挙げられる。

30

#### 【0106】

そして、このような学習を経て実現される顔識別の処理は次のようになる。先ず、2枚の顔画像の差分を正しく得るためには顔の位置を精度良く求める必要がある。また、顔の中の眼、鼻の位置は人によって異なるため、これらの位置を正規化する必要もある。例えば、ここでは、得られた顔画像に対し固有眼と固有鼻のマッチングを行うことで眼と鼻の位置同定を行い、検出された眼と鼻が定められた位置になるように、アフィン変換で回転・伸縮 (モーフィング処理) を行っている。

40

#### 【0107】

このようなモーフィング操作を経て、入力顔画像とデータベース中の登録顔画像データとされる顔画像との差分を生成する。ここで、データベースとして記憶されている顔画像は、予め登録されているものでもよい。さらに、データベースとして記憶される顔画像については更新或いは新たに蓄積していくこともできる。

#### 【0108】

例えば、ロボット装置1は、画像撮像を撮像するような画像撮像モードにおいて撮像して得た顔画像を、データベースに新たに登録することもできる。例えば、ロボット装置1は、ユーザの命令によって画像を撮像する画像撮像モードを備えており、このようなユーザ

50

の命令によって撮像した顔画像をデータベースに新たに登録することもできる。ここで、ユーザによる撮像命令としては、例えば、「顔撮って」といった音声による命令等が挙げられる。このようにすることで、エンターテインメント性を維持しながら、同時に顔画像のデータベースの構築が可能になる。また、例えば、データベース中に予め記憶しておく顔画像としては、各個人について正面を向いた顔画像とする。

#### 【0109】

次に、上述したような差分により得られた差分画像と上述の固有空間とに基づいて、人物を特定する。具体的には、差分画像とintra-personal class空間の距離 (distance from feature space, DFFS) 及び差分画像をintra-personal空間に投影した点と原点からの距離 (distance in feature space, DIFS) を計算する。例えば、同一人物の差分画像がintra-personal class空間内で正規分布をとると仮定すれば、DFFSとDIFSの加算値 (DFFS + DIFS) が分布からの距離を表すことになる。そして、この距離が小さいほど同一人物である確率が高いことになる。このような関係を利用して、入力顔画像と最も同一人物確率が高いデータベース画像をシーン中の人物と判断する。例えば、差分値が小さい場合には、差分画像はメリハリなく、いわゆるのっぺりとした画像として得られる。

10

#### 【0110】

さらに、顔の追跡によるトラッキング情報を利用して上述の確率を時系列的に蓄積している。すなわち、トラッキングに成功している間は各連続シーンごとについてのこのような顔の識別結果を時系列に取得して、このように得られる複数の識別結果が同一人物であるとする制約の下で、それら識別結果を統計的に処理する。この結果、一時的な誤認識があった場合でも、システム全体としては、その影響を受けにくいものとして、信頼の高い顔識別を実現する。

20

#### 【0111】

例えば、ロボット装置1は、上述したように自律的に行動しており、撮像画像中の顔の向きや大きさは一定であるとは言えない。さらに、データベースとして予め記憶されている顔画像については、略正面を向いた顔の数枚の顔画像を複数人物分用意しているだけである。

#### 【0112】

このような条件の下で、顔の識別をしたのでは、誤った識別結果を得る場合も多い。しかし、トラッキングに成功している間は各連続シーンごとに顔の識別処理を行い、多くの識別結果を取得して、さらにこの識別結果が同一人物のものであるとの制約の下で統計的に処理することにより、一時的に誤った識別結果を得た場合でも、システム全体としては信頼の高い顔識別が可能になる。

30

#### 【0113】

なお、あらゆる1人の人物についてあらゆるパターンの顔画像を持っていれば良いのだが、そのような横顔のデータを持っていたのでは保持すべきデータ量が膨大なものとなり、非現実である。

#### 【0114】

以上、ロボット装置1に適用される顔の検出、顔の追跡及び顔(人物)の識別として大別される顔識別について説明した。このロボット装置1においては、SVMによる濃淡顔パターンの検出、肌色のモデルと顔領域の適応更新を使ったトラッキング手法、差分顔固有値展開を使った人物識別の3つを組み合わせたものにより、顔識別を行っている。このような顔識別は、実時間、実環境でロバストに動作するトータルな顔識別システムとして実現される。

40

#### 【0115】

以上のように、計算量が少なく、姿勢変化に強い処理を利用して、入力画像内の顔の変化を実時間で追跡して、この追跡結果から得られる入力画像内の顔の推定位置についてだけ顔を検出する処理をするにより、顔の位置を特定した状態での顔の検出処理が可能になる。一方で、多くの顔検出結果を短時間で獲得できることから、これら顔検出結果に基づい

50

て顔識別を行い統計的に処理することで、高精度での顔識別を可能としている。

【0116】

ロボット装置1は、このような顔識別のシステムにより、シーン中から人間を見つけ出し（顔検出処理）、それを注視し（顔のトラッキング処理）、そこから得られた情報を用いた顔の識別による人物の特定（人物識別処理）に至るまで、全ての処理を自動的にを行い、信頼性の高い顔識別を実現する。これによりロボット装置と人間との間で自然なインタラクションが可能になる。以下、実施例について説明する。

【0117】

（3-5）実施例

図11及び図12に例として、顔の検出から顔の識別までの一連の流れを示す。条件としては、例えば、176×120の全シーンから数段階のスケージングを考慮した顔の検出に5sec程度かけ、色のトラッキングでは50msec周期で実行している場合を例としている。

【0118】

図11に示すように、顔の検出処理では、入力画像 $G_0$ から顔画像 $G_1$ を切り出す。そして、顔画像 $G_2$ のように画像の4隅をマスクして、背景部分を取り除く。一方、人物の特定する顔の識別処理では、モーフィング操作をした顔画像 $G_3$ を取得する。そして、このような処理は、顔の追跡処理に応じて時系列的に行われるようになる。

【0119】

図12に示すように、顔の識別処理では、上述したようにモーフィング操作により得た顔画像 $G_3$ について、データベースに予め学習段階で得ている顔画像 $DG_1, DG_2, DG_3, \dots$ との差分を得る。これにより、差分の顔画像 $SD_1, SD_2, SD_3, \dots$ がそれぞれ得られる。本例では、入力画像内に顔が含まれていた人物は、差分値が $P = 9.73e^{-7.2}$ となったデータベースの第3の顔画像 $DG_3$ の人物と同一人物である確率が最も高い結果となっている。例えば、本実施例では、同一人物の画像対169枚中123枚（72.8%）が同一人物であると識別され、別人物の画像対5609枚中4093枚（73.0%）が別人物であると識別する結果を得ている。

【0120】

【発明の効果】

本発明に係る顔認識装置は、撮像手段により撮像された画像内で変化する顔を追跡する顔追跡手段と、顔追跡手段による顔の追跡情報に基づいて、撮像手段により撮像された画像内の顔の顔データを検出する顔データ検出手段と、顔データ検出手段が検出した顔データに基づいて、特定顔を識別する顔識別手段とを備えることにより、画像内で変化する顔を追跡しながら当該顔の顔データを検出して、検出した顔データに基づいて、特定顔を特定することができる。

【0121】

また、本発明に係る顔識別方法は、撮像工程にて撮像された画像内で移動する顔を追跡する顔追跡工程と、顔追跡工程による顔の追跡情報に基づいて、撮像工程にて撮像された画像内の顔の顔データを検出する顔データ検出手段と、顔データ検出工程にて検出した顔データに基づいて、特定人物を識別する人物識別工程とを有することにより、画像内で変化する顔を追跡しながら当該顔の顔データを検出して、検出した顔データに基づいて、特定顔を特定することができる。

【図面の簡単な説明】

【図1】本発明の実施の形態であるロボット装置の外観構成を示す斜視図である。

【図2】上述のロボット装置の回路構成を示すブロック図である。

【図3】上述のロボット装置のソフトウェア構成を示すブロック図である。

【図4】上述のロボット装置のソフトウェア構成におけるミドル・ウェア・レイヤの構成を示すブロック図である。

【図5】上述のロボット装置のソフトウェア構成におけるアプリケーション・レイヤの構成を示すブロック図である。

【図6】上述のアプリケーション・レイヤの行動モデルライブラリの構成を示すブロック

10

20

30

40

50

図である。

【図7】ロボット装置の行動決定のための情報となる有限確率オートマトンを説明するために使用した図である。

【図8】有限確率オートマトンの各ノードに用意された状態遷移表を示す図である。

【図9】顔識別における各処理の説明に使用した図である。

【図10】顔のトラッキングの説明に使用した図である。

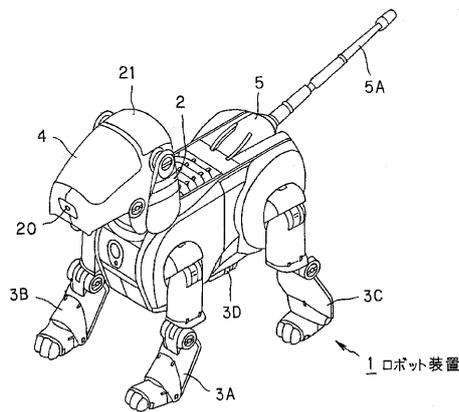
【図11】顔検出処理及び顔識別処理において行われる入力画像内から顔が切り出されモーフィングされるまでの処理の説明に使用した図である。

【図12】顔識別処理において行われる顔の差分画像を得る処理の説明に使用した図である。

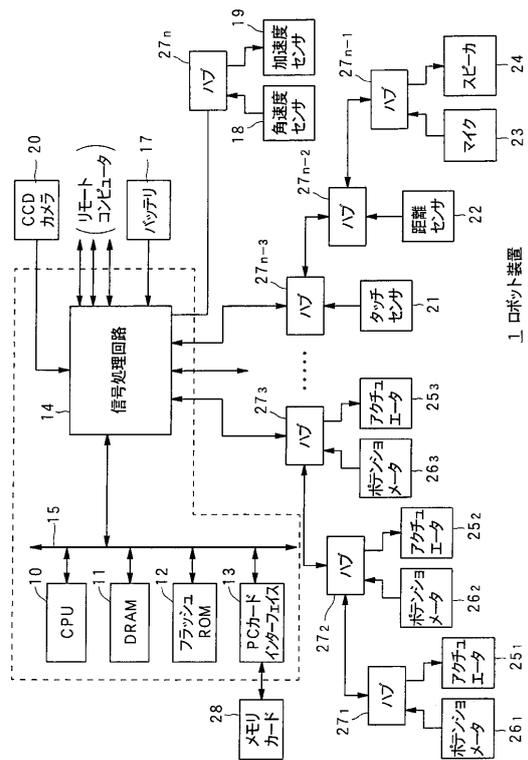
【符号の説明】

1 ロボット装置、M<sub>1</sub> 顔の検出モジュール、M<sub>2</sub> 顔の追跡モジュール、M<sub>3</sub> 顔の識別モジュール

【図1】



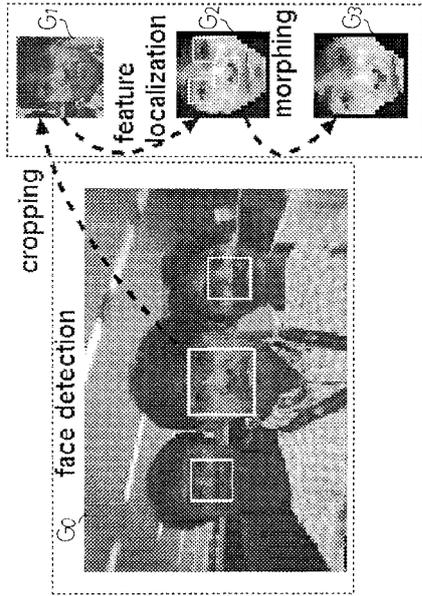
【図2】



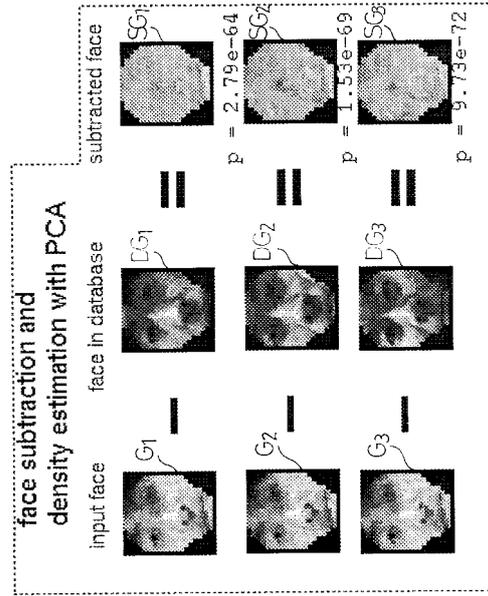




【 図 1 1 】



【 図 1 2 】



## フロントページの続き

(51)Int.Cl. F I  
G 0 6 T 7/20 (2006.01) G 0 6 T 1/00 3 4 0 A  
G 0 6 T 7/20 B

(72)発明者 河本 献太  
東京都品川区北品川6丁目7番35号 ソニー株式会社内  
(72)発明者 福地 正樹  
東京都品川区北品川6丁目7番35号 ソニー株式会社内

審査官 松尾 俊介

(56)参考文献 特開2000-210886(JP,A)  
特開平11-015979(JP,A)  
特開平10-293850(JP,A)  
特開平08-339445(JP,A)  
特開平10-269358(JP,A)  
特開平05-020442(JP,A)  
特開2000-030065(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 7/00~7/60  
A63H 11/00  
B25J 5/00  
B25J 13/00  
G06T 1/00