

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4329398号
(P4329398)

(45) 発行日 平成21年9月9日 (2009.9.9)

(24) 登録日 平成21年6月26日 (2009.6.26)

(51) Int. Cl.

F I

G 0 6 T 7/00 (2006.01)
 B 2 5 J 5/00 (2006.01)
 B 2 5 J 19/04 (2006.01)
 G 0 6 T 1/00 (2006.01)

G O 6 T 7/00 3 O O D
 B 2 5 J 5/00 F
 B 2 5 J 19/04
 G O 6 T 1/00 3 4 O A

請求項の数 14 (全 39 頁)

(21) 出願番号 特願2003-133601 (P2003-133601)
 (22) 出願日 平成15年5月12日 (2003.5.12)
 (65) 公開番号 特開2004-30629 (P2004-30629A)
 (43) 公開日 平成16年1月29日 (2004.1.29)
 審査請求日 平成18年5月10日 (2006.5.10)
 (31) 優先権主張番号 特願2002-136322 (P2002-136322)
 (32) 優先日 平成14年5月10日 (2002.5.10)
 (33) 優先権主張国 日本国 (JP)

(73) 特許権者 000002185
 ソニー株式会社
 東京都港区港南1丁目7番1号
 (74) 代理人 100067736
 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 森貞 英彦
 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

最終頁に続く

(54) 【発明の名称】 顔検出装置及び方法、並びにプログラム及び記録媒体

(57) 【特許請求の範囲】

【請求項 1】

入力画像から対象物の顔領域を抽出する顔検出装置において、
 撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と
 平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマ
 ッチング結果を生成するマッチング結果生成手段と、

上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽
 出する顔候補抽出手段と、

上記マッチング結果における上記顔候補に相当する入力画像領域から上記顔領域を識別
 する識別手段とを有し、

相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつ
 ずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上
 記テンプレートの移動に伴い配列された2次元配列であり、

上記顔候補抽出手段は、上記マッチング結果を複数の領域に分割し、この各分割領域毎
 に少なくとも相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点
 を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出する

顔検出装置。

【請求項 2】

上記分割領域のサイズは、上記テンプレートのサイズ以下である

請求項1記載の顔検出装置。

【請求項 3】

異なるサイズのテンプレートから上記マッチング結果生成手段に入力する上記テンプレートのサイズを決定するテンプレートサイズ決定手段を有する

請求項 1 記載の顔検出装置。

【請求項 4】

上記テンプレートサイズ決定手段は、予め検出された顔領域と同一サイズのテンプレートを選択する

請求項 3 記載の顔検出装置。

【請求項 5】

上記テンプレートサイズ決定手段は、上記入力画像における上記対象物との距離情報が入力され、この距離情報に基づき上記テンプレートを選択する

請求項 3 記載の顔検出装置。

【請求項 6】

上記マッチング結果生成手段は、上記異なるサイズのテンプレートが順次入力され各テンプレートに対応する上記マッチング結果を生成する

請求項 3 記載の顔検出装置。

【請求項 7】

上記顔候補抽出手段は、上記各分割領域の相関値の最大値のうち、所定の閾値以上のもののみを顔候補として抽出する

請求項 1 記載の顔検出装置。

【請求項 8】

上記顔候補抽出手段は、上記第 1 の顔候補近傍の点のうち、該第 1 の顔候補近傍の点に対応する上記テンプレートサイズの入力画像領域における肌色領域の占有率が所定の閾値以上である点を第 2 の顔候補として抽出する

請求項 1 記載の顔検出装置。

【請求項 9】

上記顔候補抽出手段は、予め学習した顔色情報を有し、上記第 1 の顔候補近傍の点のうち、該第 1 の顔候補近傍の点に対応する上記テンプレートサイズの入力画像領域における顔色領域の占有率が所定の閾値以上である点を第 2 の顔候補として抽出する

請求項 1 記載の顔検出装置。

【請求項 10】

入力画像から対象物の顔領域を抽出する顔検出方法において、

コンピュータが、

撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成工程と、

上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出工程と、

上記マッチング結果における上記顔候補に相当する入力画像領域から上記顔領域を識別する識別工程と

を実行し、

相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された 2 次元配列であり、

上記顔候補抽出工程では、上記マッチング結果を複数の領域に分割し、この各分割領域毎に少なくとも相関値の最大値を有する点を第 1 の顔候補とし、この第 1 の顔候補近傍の点を第 2 の顔候補として上記第 1 及び第 2 の顔候補を顔候補として抽出する顔検出方法。

【請求項 11】

上記顔候補抽出工程では、上記第 1 の顔候補近傍の点のうち、該第 1 の顔候補近傍の点

10

20

30

40

50

に対応する上記テンプレートサイズの入力画像領域における肌色領域の占有率が所定の閾値以上である点を第2の顔候補として抽出する

請求項10記載の顔検出方法。

【請求項12】

上記顔候補抽出工程では、予め学習した顔色情報を有し、上記第1の顔候補近傍の点のうち、該第1の顔候補近傍の点に対応する上記テンプレートサイズの入力画像領域における顔色領域の占有率が所定の閾値以上である点を第2の顔候補として抽出する

請求項10記載の顔検出方法。

【請求項13】

入力画像から人物の顔領域を抽出する動作をコンピュータに実行させるためのプログラムにおいて、

コンピュータに、

撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成手順と、

上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出手順と、

上記顔候補として抽出された点に相当する入力画像領域から上記顔領域を識別する識別手順と

を実行させるためのプログラムであって、

相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された2次元配列であり、

上記顔候補抽出手順では、上記マッチング結果を複数の領域に分割し、この各分割領域毎に少なくとも相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出する

プログラム。

【請求項14】

入力画像から人物の顔領域を抽出する動作をコンピュータに実行させるためのプログラムを記録したコンピュータ読み取り可能な記録媒体において、

コンピュータに、

撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成手順と、

上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出手順と、

上記顔候補として抽出された点に相当する入力画像領域から上記顔領域を識別する識別手順と

を実行させるためのプログラムを記録した記録媒体であって、

相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された2次元配列であり、

上記顔候補抽出手順では、上記マッチング結果を複数の領域に分割し、この各分割領域毎に少なくとも相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出する

プログラムを記録した記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力画像から対象物の顔を検出する顔検出装置及びその方法、並びに顔検出

10

20

30

40

50

する動作をコンピュータに実行させるためのプログラム及びこのプログラムを記録した記録媒に関する。

【 0 0 0 2 】

【従来の技術】

電氣的又は磁氣的な作用を用いて人間（生物）の動作に似た運動を行う機械装置を「ロボット」という。我が国においてロボットが普及し始めたのは、1960年代末からであるが、その多くは、工場における生産作業の自動化・無人化等を目的としたマニピュレータ及び搬送ロボット等の産業用ロボット（Industrial Robot）であった。

【 0 0 0 3 】

最近では、人間のパートナーとして生活を支援する、即ち住環境その他の日常生活上の様々な場面における人的活動を支援する実用ロボットの開発が進められている。このような実用ロボットは、産業用ロボットとは異なり、人間の生活環境の様々な局面において、個々に個性の相違した人間、又は様々な環境への適応方法を自ら学習する能力を備えている。例えば、犬又は猫のように4足歩行の動物の身体メカニズム及びその動作を模した「ペット型」ロボット、或いは、2足直立歩行を行う人間等の身体メカニズム及びその動作をモデルにしてデザインされた「人間型」又は「人間形」ロボット（Humanoid Robot）等のロボット装置は、既に実用化されつつある。

10

【 0 0 0 4 】

これらのロボット装置は、産業用ロボットと比較して、例えばエンターテインメント性を重視した様々な動作等を行うことができるため、エンターテインメントロボットと呼称される場合もある。また、そのようなロボット装置には、CCD（Charge Coupled Device）カメラ及びマイクロホン等の各種外部センサが搭載され、これら外部センサの出力に基づいて外部状況を認識して、外部からの情報及び内部の状態に応じて自律的に動作するものがある。

20

【 0 0 0 5 】

【発明が解決しようとする課題】

ところで、かかるエンターテインメント型のロボット装置において、対話中にその相手となる人間の顔や、移動中に視界内に入る人間の顔を検出して、その人間の顔を見ながら対話や動作を行うことができれば、人間が普段行う場合と同様に、その自然性から考えて最も望ましく、エンターテインメントロボット装置としてのエンターテインメント性をより一層向上させ得るものと考えられる。

30

【 0 0 0 6 】

従来、動画のような複雑な画像シーンの中から色や動きを使うことなく画像信号に基づく濃淡パターンのみを使って、人間の顔を検出する方法が数多く提案されている。

【 0 0 0 7 】

これらの顔検出方法としては、固有顔、ニューラル・ネットワーク及びサポートベクタマシン（SVM：Support Vector Machine）等のパターン認識の手法を利用して、予め顔パターンを学習させて識別器を生成する方法が挙げられる。

【 0 0 0 8 】

しかし、このパターン識別器を生成する方法によると、ロボット装置は、膨大なデータ量である顔画像データに対して学習によるパターン識別を行うにあたって、環境の変化や自己の姿勢及び表情の変化に対してロバスト性を示すが、その分当該パターン識別に要する演算量が増加するため、演算処理に要する時間が膨大になるという問題点があった。

40

【 0 0 0 9 】

実際に、撮像画像の中から人間の顔画像を検出するプロセス（以下、これを顔検出タスクという。）においては、当該撮像画像の中から顔画像を切り出しながら識別を行うため、撮像画像全体を様々なスケールでスキャンすることとなる。このため一回一回のパターン識別に要する演算処理を極力少なくすることが極めて重要となる。

【 0 0 1 0 】

例えば、サポートベクタマシンによるパターン認識を用いた顔検出タスクの場合、撮像画

50

像の中から400(=20×20)画素程度の切り出し画像から得られる400次元のベクトルに対して、数百通りのサポートベクタ(400次元)との内積演算が必要となる。これを大きさ(W、H)からなる全画面中で行うとなると、当該内積演算を(W-20+1)×(H-20+1)回繰り返さなくてはならないため、膨大な量の演算処理となる。

【0011】

また、顔検出タスクをロボット装置に利用する場合には、動画の中から十分に早く顔画像を検出しなければ、リアルタイム性が要求されるロボットの行動としてフィードバックすることが極めて困難となる。更に、ロボット装置の内部のCPUは、顔検出タスク以外にも常時実行しているタスクが多くあるため、これらのタスクに対して演算能力を費やしている分、当該顔検出タスクに対してフルに演算能力を費やすことは極めて困難である。

10

【0012】

本発明は、このような従来の実情に鑑みて提案されたものであり、演算量を低減してリアルタイム性を格段と向上し得る顔検出装置及びその方法、並びにプログラム及び記録媒体を提供することを目的とする。

【0013】

【課題を解決するための手段】

上述した目的を達成するために、本発明に係る顔検出装置は、入力画像から対象物の顔領域を抽出する顔検出装置において、撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成手段と、上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出手段と、上記マッチング結果における上記顔候補に相当する入力画像領域から上記顔領域を識別する識別手段とを有し、相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された2次元配列であり、上記顔候補抽出手段は、上記マッチング結果を複数の領域に分割し、この各分割領域毎に少なくとも相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出する。

20

【0014】

本発明においては、平均顔テンプレートと同一のサイズの顔が入力画像に存在する場合にテンプレートマッチングの相関値が最も大きくなることを利用し、マッチング結果の相関値の局所最大値に基づいて顔候補を抽出することにより、マッチング結果から単に閾値を設けて顔候補を検出する場合に比して、顔検出性能を維持しつつ顔候補の数を飛躍的に低減することができ、これにより、後段の識別手段における顔識別の演算量を削減し、入力画像から極めて高速に顔領域を検出することができる。

30

【0015】

また、上記顔候補抽出手段は、上記マッチング結果を、例えばテンプレートサイズ以下の大きさの複数の領域に分割しこの各分割領域毎に少なくとも相関値の最大値を顔候補として抽出することができる。

【0016】

40

更に、異なるサイズのテンプレートから上記マッチング結果生成手段に入力する上記テンプレートのサイズを決定するテンプレートサイズ決定手段を有し、上記テンプレートサイズ決定手段は、予め検出された顔領域と同一サイズのテンプレートを選択するか、又は入力画像における上記対象物との距離情報が入力され、この距離情報に基づきテンプレートを選択することにより、効率よくテンプレートサイズを決定することができる。

【0017】

更にまた、上記マッチング結果生成手段は、異なるサイズのテンプレートが順次入力され各テンプレートに対応するマッチング結果を生成することができ、これにより、入力画像に含まれる任意の大きさの顔を検出することができる。

【0018】

50

また、上記顔候補抽出手段は、上記各分割領域の相関値の最大値のうち、所定の閾値以上のみを顔候補として抽出することにより、更に顔候補を絞り込むことができる。

【0019】

更に、相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された2次元配列とすることができる。

【0020】

更にまた、上記顔候補抽出手段は、上記各分割領域の相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出することにより、顔検出性能を更に向上することができる。

10

【0021】

また、上記顔候補抽出手段は、上記第1の顔候補近傍の点のうち、該第1の顔候補近傍の点に対応する上記テンプレートサイズの入力画像領域における肌色領域の占有率が所定の閾値以上である点か、又は、予め学習した顔色情報を有し、上記第1の顔候補近傍の点のうち、該第1の顔候補近傍の点に対応する上記テンプレートサイズの入力画像領域における顔色領域の占有率が所定の閾値以上である点を第2の顔候補として抽出することができ、これにより、顔検出性能を高めつつ顔候補を絞り込むことができる。

【0022】

本発明に係る顔検出方法は、入力画像から対象物の顔領域を抽出する顔検出方法において、コンピュータが、撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成工程と、上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出工程と、上記マッチング結果における上記顔候補に相当する入力画像領域から上記顔領域を識別する識別工程とを実行し、相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された2次元配列であり、上記顔候補抽出工程では、上記マッチング結果を複数の領域に分割し、この各分割領域毎に少なくとも相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出する。

20

30

【0026】

本発明に係るプログラム及びこれを記録した記録媒体は、入力画像から人物の顔領域を抽出する動作をコンピュータに実行させるためのプログラム及びこれを記録した記録媒体において、コンピュータに、撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成手順と、上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出手順と、上記顔候補として抽出された点に相当する入力画像領域から上記顔領域を識別する識別手順とを実行させるためのプログラムであって、相関値の集合である上記マッチング結果は、上記入力画像をスキャンして所定画素ずつずらしながら移動させた上記テンプレートと上記入力画像との相関値の集合であって、上記テンプレートの移動に伴い配列された2次元配列であり、上記顔候補抽出手順では、上記マッチング結果を複数の領域に分割し、この各分割領域毎に少なくとも相関値の最大値を有する点を第1の顔候補とし、この第1の顔候補近傍の点を第2の顔候補として上記第1及び第2の顔候補を顔候補として抽出するプログラム及びこれを記録した記録媒体である。

40

【0027】

【発明の実施の形態】

(1) 本願発明の概要

本発明に係る顔検出装置及び顔検出方法は、入力画像（撮像画像）から人物の顔（顔領域）を検出するものである。この顔検出タスクは、大きく分けて2つの工程を有する。即ち

50

、入力画像に対してテンプレートマッチングを行って顔の候補を抽出する第1の工程と、この顔候補からSVM等により顔であるか否かの判定する第2の工程とからなる。

【0028】

ロボット装置等のCPU及びメモリ等のリソースが限られたシステムにおいて、リアルタイムに顔検出を行うためには、顔検出タスクにおける演算量を削減すること必要である。しかしながら、上述した如く、第2の工程で行われるSVM等を使用する顔判定は膨大な演算量を要する。従って、顔検出タスクにおいて、顔検出性能を保持しつつ顔検出に要する演算量を削減するためには、第1の工程において顔候補を絞り込むことが有効となる。

【0029】

ここで、第1の工程において、テンプレートマッチングを行って顔候補を抽出する際、例えばマッチングの相関値に対して閾値を設けて顔候補を抽出する方法がある。この方法では、顔候補の見逃しを軽減しようとした場合、相関値の閾値を上げる方法又はマッチングの際の間引きを減らす方法等をとることができるものの、例えば相関値の閾値を下げると多くの顔候補が抽出されるため、計算量が増大してしまい、ロボット装置等のリソースの限られた環境においては好ましくない場合がある。一方、閾値を上げると、顔候補の数を減らすことができるものの、本来顔である画像も候補から取り除いてしまい、顔画像を見逃してしまう場合がある。

【0030】

そこで、本願発明は、本来顔である画像を顔候補から見逃すことなく、演算量の増大を防止する、即ち、顔候補の数を有効に絞り込むことができる方法を提案するものである。具体的には、第1の工程では、入力画像と、所定サイズの平均的な顔画像を示すテンプレートとの相関をとった相関値の集合であるマッチング結果を生成し、このマッチング結果における相関値の局所最大値に基づき顔領域を抽出する。顔領域は、マッチング結果を例えばテンプレートと同一サイズの複数の領域に分割して各分割領域毎に少なくとも相関値の最大値を顔候補として抽出する。そして、第2の工程では、上記顔候補として抽出された点に相当する入力画像領域からSVM等により顔領域（顔画像）を識別（判定）するものである。

【0031】

（2）ロボットの構成

以下、本発明を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。この実施の形態は、本発明を、顔検出装置を搭載した2足歩行ロボット装置に適用したものである。先ず、顔検出手段を有するロボット装置の一例として、2足歩行の人間型ロボット装置について説明する。

【0032】

図1及び図2には、人間型ロボット装置200を前方及び後方の各々から眺望した様子を示している。更に、図3には、この人間型ロボット装置200が具備する関節自由度構成を模式的に示している。

【0033】

図1及び図2に示すように、人間型ロボット装置200は、脚式移動を行う左右2足の脚部ユニット201R、201Lと、胴体部ユニット202と、左右の腕部ユニット203R、203Lと、頭部ユニット204とで構成される。

【0034】

左右各々の脚部ユニット201R、201Lは、大腿部205R、205Lと、膝関節206R、206Lと、脛部207R、207Lと、足首208R、208Lと、足平209R、209Lとで構成され、股関節210R、210Lによって胴体部ユニット202の略最下端にて連結されている。また、左右各々の腕部ユニット203R、203Lは、上腕211R、211Lと、肘関節212R、212Lと、前腕213R、213Lとで構成され、肩関節214R、214Lによって胴体部ユニット202上方の左右各側縁にて連結されている。また、頭部ユニット204は、首関節255によって胴体部ユニット202の略最上端中央に連結されている。

【 0 0 3 5 】

頭部ユニット 2 0 4 と体幹部ユニット 2 0 2 との間の首間接は、図 3 に示すように、頭部ユニット 2 0 4 を支持する首関節ヨー軸 3 0 2 と、首関節ピッチ軸 3 0 3 と、首関節ロール軸 3 0 4 という 3 自由度を有している。

【 0 0 3 6 】

また、腕の関節は、肩関節ピッチ軸 3 0 8 と、肩関節ロール軸 3 0 9 と、上腕ヨー軸 3 1 0 と、肘関節ピッチ軸 3 1 1 と、前腕ヨー軸 3 1 2 と、手首関節ピッチ軸 3 1 3 と、手首関節ロール軸 3 1 4 と、手部 3 1 5 とで構成される。手部 3 1 5 は、実際には、複数本の指を含む多関節・多自由度構造体である。ただし、手部 3 1 5 の動作は人間型ロボット装置 2 0 0 の姿勢制御や歩行制御に対する寄与や影響が少ないので、本明細書ではゼロ自由度と仮定する。従って、各腕部は 7 自由度を有するとする。

10

【 0 0 3 7 】

また、胴体部ユニットは、胴体ピッチ軸 3 0 5 と、胴体ロール軸 3 0 6 と、胴体ヨー軸 3 0 7 という 3 自由度を有する。

【 0 0 3 8 】

また、下肢を構成する各々の脚部ユニットは、股関節ヨー軸 3 1 6 と、股関節ピッチ軸 3 1 7 と、股関節ロール軸 3 1 8 と、膝関節ピッチ軸 3 1 9 と、足首関節ピッチ軸 3 2 0 と、足首関節ロール軸 3 2 1 と、足部 3 2 2 とで構成される。本明細書中では、股関節ピッチ軸 3 1 7 と股関節ロール軸 3 1 8 との交点は、人間型ロボット装置 2 0 0 の股関節位置を定義する。人体の足部 3 2 2 は、実際には多関節・多自由度の足底を含んだ構造体であるが、人間型ロボット装置 2 0 0 の足底は、ゼロ自由度とする。従って、各脚部は、6 自由度で構成される。

20

【 0 0 3 9 】

以上を総括すれば、人間型ロボット装置 2 0 0 全体としては、合計で $3 + 7 \times 2 + 3 + 6 \times 2 = 32$ 自由度を有することになる。ただし、エンターテインメント向けの人間型ロボット装置 2 0 0 が必ずしも 32 自由度に限定される訳ではない。設計・制作上の制約条件や要求仕様等に応じて、自由度すなわち関節数を適宜増減することができることはいうまでもない。

【 0 0 4 0 】

上述したような人間型ロボット装置 2 0 0 がもつ各自由度は、実際にはアクチュエータを用いて実装される。外観上で余分な膨らみを排してヒトの自然体形状に近似させること、2 足歩行という不安定な構造体に対して姿勢制御を行うこと等の要請から、アクチュエータは小型且つ軽量であることが好ましい。

30

【 0 0 4 1 】

図 4 には、人間型ロボット装置 2 0 0 の制御システム構成を模式的に示している。同図に示すように、人間型ロボット装置 2 0 0 は、ヒトの四肢を表現した各機構ユニット 3 3 0 , 3 4 0 , 3 5 0 R / L , 3 6 0 R / L と、各機構ユニット間の協調動作を実現するための適応制御を行う制御ユニット 3 8 0 とで構成される（ただし、R 及び L の各々は、右及び左の各々を示す接尾辞である。以下同様）。

【 0 0 4 2 】

人間型ロボット装置 2 0 0 全体の動作は、制御ユニット 3 8 0 によって統括的に制御される。制御ユニット 3 8 0 は、CPU (Central Processing Unit) やメモリ等の主要回路コンポーネント（図示しない）で構成される主制御部 3 8 1 と、電源回路や人間型ロボット装置 2 0 0 の各構成要素とのデータやコマンドの授受を行うインターフェイス（何れも図示しない）等を含んだ周辺回路 3 8 2 とで構成される。この制御ユニット 3 8 0 の設置場所は、特に限定されない。図 4 では胴体部ユニット 3 4 0 に搭載されているが、頭部ユニット 3 3 0 に搭載してもよい。或いは、人間型ロボット装置 2 0 0 外に制御ユニット 3 8 0 を配備して、人間型ロボット装置 2 0 0 の機体とは有線若しくは無線で通信するようにしてもよい。

40

【 0 0 4 3 】

50

図 4 に示した人間型ロボット装置 200 内の各関節自由度は、それぞれに対応するアクチュエータによって実現される。すなわち、頭部ユニット 330 には、首関節ヨー軸 302、首関節ピッチ 303、首関節ロール軸 304 の各々を表現する首関節ヨー軸アクチュエータ A_2 、首関節ピッチ軸アクチュエータ A_3 、首関節ロール軸アクチュエータ A_4 が配設されている。

【0044】

また、頭部ユニット 330 には、外部の状況を撮像するための CCD (Charge Coupled Device) カメラが設けられているほか、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、外部音を集音するためのマイク、音声を出力するためのスピーカ、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等が配設されている。

10

【0045】

また、胴体部ユニット 340 には、胴体ピッチ軸 305、胴体ロール軸 306、胴体ヨー軸 307 の各々を表現する胴体ピッチ軸アクチュエータ A_5 、胴体ロール軸アクチュエータ A_6 、胴体ヨー軸アクチュエータ A_7 が配設されている。また、胴体部ユニット 340 には、この人間型ロボット装置 200 の起動電源となるバッテリーを備えている。このバッテリーは、充放電可能な電池によって構成されている。

【0046】

また、腕部ユニット 350 R/L は、上腕ユニット 351 R/L と、肘関節ユニット 352 R/L と、前腕ユニット 353 R/L に細分化されるが、肩関節ピッチ軸 308、肩関節ロール軸 309、上腕ヨー軸 310、肘関節ピッチ軸 311、前腕ヨー軸 312、手首関節ピッチ軸 313、手首関節ロール軸 314 の各々を表現する肩関節ピッチ軸アクチュエータ A_8 、肩関節ロール軸アクチュエータ A_9 、上腕ヨー軸アクチュエータ A_{10} 、肘関節ピッチ軸アクチュエータ A_{11} 、肘関節ロール軸アクチュエータ A_{12} 、手首関節ピッチ軸アクチュエータ A_{13} 、手首関節ロール軸アクチュエータ A_{14} が配備されている。

20

【0047】

また、脚部ユニット 360 R/L は、大腿部ユニット 361 R/L と、膝ユニット 362 R/L と、脛部ユニット 363 R/L に細分化されるが、股関節ヨー軸 316、股関節ピッチ軸 317、股関節ロール軸 318、膝関節ピッチ軸 319、足首関節ピッチ軸 320、足首関節ロール軸 321 の各々を表現する股関節ヨー軸アクチュエータ A_{16} 、股関節ピッチ軸アクチュエータ A_{17} 、股関節ロール軸アクチュエータ A_{18} 、膝関節ピッチ軸アクチュエータ A_{19} 、足首関節ピッチ軸アクチュエータ A_{20} 、足首関節ロール軸アクチュエータ A_{21} が配備されている。各関節に用いられるアクチュエータ A_2, A_3, \dots は、より好ましくは、ギア直結型で且つサーボ制御系をワンチップ化してモータ・ユニット内に搭載したタイプの小型 AC サーボ・アクチュエータで構成することができる。

30

【0048】

頭部ユニット 330、胴体部ユニット 340、腕部ユニット 350、各脚部ユニット 360 等の各機構ユニット毎に、アクチュエータ駆動制御部の副制御部 335, 345, 355 R/L, 365 R/L が配備されている。更に、各脚部 360 R, L の足底が着床したか否かを検出する接地確認センサ 391 及び 392 を装着するとともに、胴体部ユニット 340 内には、姿勢を計測する姿勢センサ 393 を装備している。

40

【0049】

接地確認センサ 391 及び 392 は、例えば足底に設置された近接センサ又はマイクロ・スイッチ等で構成される。また、姿勢センサ 393 は、例えば、加速度センサとジャイロ・センサの組み合わせによって構成される。

【0050】

接地確認センサ 391 及び 392 の出力によって、歩行・走行等の動作期間中において、左右の各脚部が現在立脚又は遊脚何れの状態であるかを判別することができる。また、姿勢センサ 393 の出力により、胴体部分の傾きや姿勢を検出することができる。

【0051】

50

主制御部 381 は、各センサ 391 ~ 393 の出力にตอบสนองして制御目標をダイナミックに補正することができる。より具体的には、副制御部 335, 345, 355 R/L, 365 R/L の各々に対して適応的な制御を行い、人間型ロボット装置 200 の腕部、胴体、及び脚部が協調して駆動する全身運動パターンを実現できる。

【0052】

人間型ロボット装置 200 の機体上での全身運動は、足部運動、ZMP (Zero Moment Point) 軌道、胴体運動、上肢運動、腰部高さ等を設定するとともに、これらの設定内容にしたがった動作を指示するコマンドを各副制御部 335, 345, 355 R/L, 365 R/L に転送する。そして、各々の副制御部 335, 345, ... 等では、主制御部 381 からの受信コマンドを解釈して、各アクチュエータ A_2 , A_3 ... 等に対して駆動制御信号を出力する。ここでいう「ZMP」とは、歩行中の床反力によるモーメントがゼロとなる床面上の点のことであり、また、「ZMP 軌道」とは、例えば人間型ロボット装置 200 の歩行動作期間中に ZMP が動く軌跡を意味する。

10

【0053】

歩行時には、重力と歩行運動に伴って生じる加速度によって、歩行系から路面には重力と慣性力、並びにこれらのモーメントが作用する。いわゆる「ダランベールの原理」によると、それらは路面から歩行系への反作用としての床反力、床反力モーメントとバランスする。力学的推論の帰結として、足底接地点と路面の形成する支持多角形の辺上或いはその内側にピッチ及びロール軸モーメントがゼロとなる点、すなわち「ZMP (Zero Moment Point)」が存在する。

20

【0054】

脚式移動ロボットの姿勢安定制御や歩行時の転倒防止に関する提案の多くは、この ZMP を歩行の安定度判別の規範として用いたものである。ZMP 規範に基づく 2 足歩行パターン生成は、足底着地点を予め設定することができ、路面形状に応じた足先の運動学的拘束条件を考慮しやすい等の利点がある。また、ZMP を安定度判別規範とすることは、力ではなく軌道を運動制御上の目標値として扱うことを意味するので、技術的に実現可能性が高まる。なお、ZMP の概念並びに ZMP を歩行ロボットの安定度判別規範に適用する点については、Miomir Vukobratovic 著「LEGGED LOCOMOTION ROBOTS」(加藤一郎外著『歩行ロボットと人工の足』(日刊工業新聞社))に記載されている。

【0055】

一般には、4 足歩行よりもヒューマノイドのような 2 足歩行のロボットの方が、重心位置が高く、且つ、歩行時の ZMP 安定領域が狭い。従って、このような路面状態の変化に伴う姿勢変動の問題は、2 足歩行ロボットにおいてとりわけ重要となる。

30

【0056】

以上のように、人間型ロボット装置 200 は、各々の副制御部 335, 345, ... 等が、主制御部 381 からの受信コマンドを解釈して、各アクチュエータ A_2 , A_3 ... に対して駆動制御信号を出力し、各ユニットの駆動を制御している。これにより、人間型ロボット装置 200 は、目標の姿勢に安定して遷移し、安定した姿勢で歩行できる。

【0057】

また、人間型ロボット装置 200 における制御ユニット 380 では、上述したような姿勢制御のほかに、加速度センサ、タッチセンサ、接地確認センサ等の各種センサ、及び CCD カメラからの画像情報、マイクからの音声情報等を統括して処理している。制御ユニット 380 では、図示しないが加速度センサ、ジャイロ・センサ、タッチセンサ、距離センサ、マイク、スピーカ等の各種センサ、各アクチュエータ、CCD カメラ及びバッテリーが各々対応するハブを介して主制御部 381 と接続されている。

40

【0058】

主制御部 381 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インターフェイスを介して DRAM 内の所定位置に順次格納する。また、主制御部 381 は、バッテリー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 DRAM 内の所定位置に格納する。DR

50

ＡＭ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、主制御部３８１がこの人間型ロボット装置２００の動作制御を行う際に利用される。

【００５９】

主制御部３８１は、人間型ロボット装置２００の電源が投入された初期時、制御プログラムを読み出し、これをＤＲＡＭに格納する。また、主制御部３８１は、上述のように主制御部３８１よりＤＲＡＭ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等を判断する。更に、主制御部３８１は、この判断結果及びＤＲＡＭに格納した制御プログラムに基づいて自己の状況に応じて行動を決定するとともに、当該決定結果に基づいて必要なアクチュエータを駆動させることにより人間型ロボット装置２００に、いわゆる「身振り」、「手振り」といった行動をとらせる。

10

【００６０】

従って、人間型ロボット装置２００は、制御プログラムに基づいて自己及び周囲の状況を判断し、使用者からの指示及び働きかけに応じて自律的に行動できる。また、人間型ロボット装置２００は、ＣＣＤカメラにおいて撮像された画像から抽出した文字の発音のしかた（読み方）を、抽出された文字から推定される読み方と集音マイクにおいて集音された音声とをマッチングして決定する。従って、人間型ロボット装置２００の音声認識の精度が向上し、新規単語が音声認識用辞書に登録できる。

【００６１】

（２）適用例

20

以下、本発明の実施の形態の説明に先立ち、理解を容易とするために、上述した２足歩行ロボット装置において、入力画像に対してテンプレートマッチングを行って顔候補を抽出し、この顔候補から顔を判定する顔検出装置を搭載した他の例について説明する（特願２００２－０７４９０７号参照、以下、適用例という。）。

【００６２】

以下に説明する第１及び第２の適用例においても、本願発明と同様にテンプレートマッチング等によりおおまかに顔候補を抽出した後、ＳＶＭ等により顔候補の中から顔であるかを判定することにより、この後段における演算量を削減してロボット装置の顔検出のリアルタイム性を向上させるものである。

【００６３】

30

（２－１）第１の適用例

（２－１－１）ロボットの内部構成

第１の適用例においては、胴体部ユニット２０４の胴体下部を形成する腰ベースの背面側には、図４及び図５に示すように、当該ロボット装置２００全体の動作制御を司る主制御部３８１、電源回路及び通信回路等の周辺回路３８１、バッテリー４５（図５）等がボックスに収納されてなる制御ユニット４２が配設されている。

【００６４】

そしてこの制御ユニット４２は、各構成ユニット（胴体部ユニット２０２、頭部ユニット２０４、各腕部ユニット２０３Ｒ、２０３Ｌ及び各脚部ユニット２０９Ｒ、２０９Ｌ）内にそれぞれ配設された各副制御部３３５、３５０Ｒ、３５０Ｌ、３６５Ｒ、３６５Ｌと接続されており、これら副制御部３３５、３５０Ｒ、３５０Ｌ、３６５Ｒ、３６５Ｌに対して必要な電源電圧を供給したり、これら副制御部３３５、３５０Ｒ、３５０Ｌ、３６５Ｒ、３６５Ｌと通信を行ったりすることができるようになされている。

40

【００６５】

また、各副制御部３３５、３５０Ｒ、３５０Ｌ、３６５Ｒ、３６５Ｌは、それぞれ対応する構成ユニット内の各アクチュエータ $A_2 \sim A_{21}$ と接続されており、当該構成ユニット内の各アクチュエータ $A_2 \sim A_{21}$ を主制御部３８１から与えられる各種制御コマンドに基づいて指定された状態に駆動し得るようになされている。

【００６６】

更に、頭部ユニット２０４には、図５に示すように、このロボット装置２００の「目」と

50

して機能するＣＣＤ（Charge Coupled Device）カメラ５０及び「耳」として機能するマイクロホン５１及びタッチセンサ５２等からなる外部センサ部５３と、「口」として機能するスピーカ５４と等がそれぞれ所定位置に配設され、制御ユニット３８０内には、バッテリーセンサ５５及び加速度センサ５６等からなる内部センサ部５７が配設されている。

【００６７】

そして外部センサ部５３のＣＣＤカメラ５０は、周囲の状況を撮像し、得られた画像信号Ｓ１Ａを主制御部３８１に送出してフレーム単位で内部メモリ３８１Ａに順次記憶する一方、マイクロホン５１は、ユーザから音声入力として与えられる「歩け」、「座れ」又は「ボールを追いかける」等の各種命令音声を集音し、かくして得られた音声信号Ｓ１Ｂを主制御部３８１に送出するようになされている。

10

【００６８】

更に、タッチセンサは、頭部ユニット３の上部に設けられており、ユーザからの「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し、検出結果を圧力検出信号Ｓ１Ｃとして主制御部３８１に送出する。

【００６９】

更にまた、内部センサ部５７のバッテリーセンサ５５は、バッテリー４５のエネルギー残量を所定周期で検出し、検出結果をバッテリー残量検出信号Ｓ２Ａとして主制御部３８１に送出する。一方、加速度センサ５６は、３軸方向（ｘ軸、ｙ軸及びｚ軸）の加速度を所定周期で検出し、検出結果を加速度検出信号Ｓ２Ｂとして主制御部３８１に送出する。

【００７０】

20

主制御部３８１は、外部センサ部５３のＣＣＤカメラ５０、マイクロホン５１及びタッチセンサ５２等からそれぞれ供給される画像信号Ｓ１Ａ、音声信号Ｓ１Ｂ及び圧力検出信号Ｓ１Ｃ等（以下、これらをまとめて外部センサ信号Ｓ１と呼ぶ）と、内部センサ部５７のバッテリーセンサ５５及び加速度センサ等からそれぞれ供給されるバッテリー残量検出信号Ｓ２Ａ及び加速度検出信号Ｓ２Ｂ等（以下、これらをまとめて内部センサ信号Ｓ２という。）に基づいて、ロボット装置２００の周囲及び内部の状況や、ユーザからの指令、ユーザからの働きかけの有無等を判断する。

【００７１】

そして主制御部３８１は、この判断結果と、予め内部メモリ３８１Ａに格納されている制御プログラムと、そのとき装填されている外部メモリ５８に格納されている各種制御パラメータとに基づいて続く行動を決定し、決定結果に基づく制御コマンドを対応する副制御部３３５、３５０Ｒ、３５０Ｌ、３６５Ｒ、３６５Ｌに送出する。この結果、この制御コマンドに基づき、その副制御部３３５、３５０Ｒ、３５０Ｌ、３６５Ｒ、３６５Ｌの制御のもとに、対応するアクチュエータ $A_2 \sim A_{21}$ が駆動され、かくして頭部ユニット２０４を上下左右に揺動させたり、腕部ユニット２０３Ｒ、２０３Ｌを上にあげたり、歩行する等の行動がロボット装置２００により発現されることとなる。

30

【００７２】

この際、主制御部３８１は、必要に応じて所定の音声信号Ｓ３をスピーカ５４に与えることにより当該音声信号Ｓ３に基づく音声を外部に出力させたり、外見上の「目」として機能する頭部ユニット２０４の所定位置に設けられたＬＥＤに駆動信号を出力することによりこれを点滅させる。

40

【００７３】

このようにしてこのロボット装置２００においては、周囲及び内部の状況や、ユーザからの指令及び働きかけの有無等に基づいて自律的に行動することができるようになされている。

【００７４】

（２－１－２）顔検出タスク機能に関する主制御部３８１の処理

次に、このロボット装置２００に搭載された顔検出タスク機能について説明する。このロボット装置２００には、ＣＣＤカメラ５０を介して内部メモリ３８１Ａに記憶されるフレーム画像について、当該フレーム画像の中から人間の顔画像を検出する顔検出タスク機能

50

が搭載されている。そしてこの顔検出タスク機能は、主制御部 3 8 1 における各種処理により実現されている。

【 0 0 7 5 】

ここで、かかる顔検出タスク機能に関する主制御部 3 8 1 の処理内容を機能的に分類すると、図 6 に示すように、入力画像スケール変換部 6 0、ウィンドウ切出部 6 1、テンプレートマッチング部 6 2、前処理部 6 3、パターン識別部 6 4 及び重なり判定部 6 5 に分けることができる。

【 0 0 7 6 】

入力画像スケール変換部 6 0 は、CCD カメラ 5 0 (図 5) からの画像信号 S 1 A に基づくフレーム画像を内部メモリ 3 8 1 A から読み出して、当該フレーム画像を縮小率が相異なる複数のスケール画像に変換する。この適用例の場合、2 5 3 4 4 (= 176 × 144) 画素からなるフレーム画像に対して、これを 0 . 8 倍ずつ順次縮小して 5 段階 (1 . 0 倍、0 . 8 倍、0 . 6 4 倍、0 . 5 1 倍、0 . 4 1 倍) のスケール画像 (以下、これを第 1 ~ 第 5 のスケール画像と呼ぶ) に変換する。

【 0 0 7 7 】

続くウィンドウ切出部 6 1 は、第 1 ~ 第 5 のスケール画像のうち、まず第 1 のスケール画像に対して、画像左上を起点として順に画像右下まで、適当な画素 (例えば 2 画素) 分を右側又は下側にずらしながらスキャンするようにして、4 0 0 (= 20 × 20) 画素の矩形領域 (以下、この領域をウィンドウ画像と呼ぶ) を順次切り出す。

【 0 0 7 8 】

その際、ウィンドウ切出部 6 1 は、第 1 のスケール画像から切り出した複数のウィンドウ画像のうち先頭のウィンドウ画像を後段のテンプレートマッチング部 6 2 に送出する。

【 0 0 7 9 】

テンプレートマッチング部 6 2 は、ウィンドウ切出部 6 1 から得られた先頭のウィンドウ画像について、正規化相関法や誤差二乗法等の演算処理を実行してピーク値をもつ関数曲線に変換した後、当該関数曲線に対して認識性能が落ちない程度に十分に低い閾値を設定して、当該閾値を基準として当該ウィンドウ画像が顔画像か否かを判断する。

【 0 0 8 0 】

本適用例の場合、テンプレートマッチング部 6 2 では、例えば 1 0 0 人程度の人物の平均からなる平均的な顔画像をテンプレートとして、かかる顔画像か否かの判断基準となる閾値を設定するようになされている。これにより当該ウィンドウ画像について、テンプレートとなる平均的な顔画像との大まかなマッチングをとり得るようになされている。

【 0 0 8 1 】

このようにしてテンプレートマッチング部 6 2 は、ウィンドウ切出部 6 1 から得られたウィンドウ画像について、テンプレートによるマッチングをとり、顔画像であると判断された場合には、当該ウィンドウ画像をスコア画像として後段の前処理部 6 3 に送出する一方、顔画像でないと判断された場合には、当該ウィンドウ画像をそのまま後段の重なり判定部 6 5 に送出する。

【 0 0 8 2 】

この時点で顔画像であると判断されたウィンドウ画像 (スコア画像) には、実際には顔画像以外の判断誤りの画像が大量に含まれるが、日常のシーンの中では顔に類似した背景画像が多く存在することはあまりないため、ほとんどのウィンドウ画像は顔画像ではないと判断されることとなり極めて有効である。

【 0 0 8 3 】

実際に上述した正規化相関法や誤差二乗法等の演算処理は、後段の前処理部及びパターン識別部における演算処理と比較すると、演算量が 1 0 分の 1 から 1 0 0 分の 1 程度で済むと共に、実験上この段階で顔画像以外の画像を 8 0 [%] 以上はふるい落とすことができることが確認されたため、主制御部 3 8 1 全体としては大幅な演算量の削減につながる事がわかる。

【 0 0 8 4 】

前処理部 6 3 は、テンプレートマッチング部 6 2 から得られたスコア画像について、矩形領域でなる当該スコア画像から人間の顔画像とは無関係である背景部分に相当する 4 隅の領域を除去するために、当該 4 隅の領域を切り取ったマスクを用いて、4 0 0 (= 20 × 20) 画素あるスコア画像から 3 6 0 画素分を抽出する。

【 0 0 8 5 】

そして前処理部 6 3 は、撮像時の照明により濃淡で表される被写体の傾き条件を解消すべく、当該抽出した 3 6 0 画素分のスコア画像のうち顔画像として最適な部位を基準とする平面を形成するように、例えば平均二乗誤差 (R S M : Root Mean Square) 等による算出方法を用いて当該 3 6 0 画素の濃淡値に補正をかける。

【 0 0 8 6 】

続いて前処理部 6 3 は、当該 3 6 0 画素分のスコア画像のコントラストを強調した結果をヒストグラム平滑化処理を行うことにより、C C D カメラ 5 0 のゲインや照明の強弱によらずに検出できるようにする。

【 0 0 8 7 】

次いで、前処理部 6 3 は、後述するガボア・フィルタリング (Gabor Filtering) 処理を行うことにより、当該 3 6 0 画素分のスコア画像をベクトル変換し、得られたベクトル群を更に 1 本のパターンベクトルに変換する。

【 0 0 8 8 】

以下において、ガボア・フィルタリング処理について説明する。まず人間の視覚細胞には、ある特定の方位に対して選択性を持つ細胞が存在することが既に判っている。これは、垂直の線に対して反応する細胞と、水平の線に反応する細胞で構成される。ガボア・フィルタリングは、これと同様に、方位選択性を持つ複数のフィルタで構成される空間フィルタである。

【 0 0 8 9 】

ガボア・フィルタは、ガボア関数によって空間表現される。ガボア関数 $g (x , y)$ は、下記式 (1) に示すように、コサイン成分からなるキャリア $s (x , y)$ と、2 次元ガウス分析状のエンベロープ $W_r (x , y)$ とで構成される。

【 0 0 9 0 】

【 数 1 】

$$g(x, y) = s(x, y)W_r(x, y) \quad \dots(1)$$

【 0 0 9 1 】

キャリア $s (x , y)$ は、複素関数を用いて、下記式 (2) のように表現される。ここで、座標値 (u_0 , v_0) は空間周波数を表し、また P はコサイン成分の位相を表す。

【 0 0 9 2 】

【 数 2 】

$$s(x, y) = \exp(j(2\pi(u_0x + v_0y) + P)) \quad \dots(2)$$

【 0 0 9 3 】

ここで、上記式 (2) に示すキャリアは、下記式 (3) に示すように、実数成分 $Re (s (x , y)$ と虚数成分 $Im (s (x , y))$ に分離することができる。

【 0 0 9 4 】

【 数 3 】

$$\begin{cases} \operatorname{Re}(s(x, y)) = \cos(2\pi(u_0x + v_0y) + P) \\ \operatorname{Im}(s(x, y)) = \sin(2\pi(u_0x + v_0y) + P) \end{cases} \quad \dots(3)$$

【 0 0 9 5 】

一方、2次元ガウス分布からなるエンベロープは、下記式(4)のように表現される。

【 0 0 9 6 】

【数4】

10

$$W_r(x, y) = K \exp\left(-\pi\left(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2\right)\right) \quad \dots(4)$$

【 0 0 9 7 】

ここで、座標軸 (x_0, y_0) はこの関数のピークであり、定数 a 及び b はガウス分布のスケール・パラメータである。また、添え字 r は、下記式(5)に示すような回転操作を意味する。

【 0 0 9 8 】

20

【数5】

$$\begin{cases} (x-x_0)_r = (x-x_0)\cos\theta + (y-y_0)\sin\theta \\ (y-y_0)_r = -(x-x_0)\sin\theta + (y-y_0)\cos\theta \end{cases} \quad \dots(5)$$

【 0 0 9 9 】

従って、上記式(2)及び(4)より、ガボア・フィルタは、下記式(6)に示すような空間関数として表現される。

30

【 0 1 0 0 】

【数6】

$$g(x, y) = K \exp\left(-\pi\left(a^2(x-x_0)_r^2 + b^2(y-y_0)_r^2\right)\right) \exp\left(j(2\pi(u_0x + v_0y) + P)\right) \quad \dots(6)$$

【 0 1 0 1 】

40

本適用例に係る前処理部は、8種類の方向と3通りの周波数を採用して、合計24個のガボア・フィルタを用いて顔抽出処理を行う。

【 0 1 0 2 】

ガボア・フィルタのレスポンスは、 G_i を i 番目のガボア・フィルタとし、 i 番目のガボアの結果(Gabor Jet)を J_i とし、入力イメージを I とすると、下記式(7)で表される。この式(7)の演算は、実際には高速フーリエ変換を用いて高速化することができる。

【 0 1 0 3 】

【数7】

$$J_i(x, y) = G_i(x, y) \oplus I(x, y) \quad \dots(7)$$

【 0 1 0 4 】

作成したガボア・フィルタの性能を調べるためには、フィルタリングして得られた画素を再構築することによって行う。再構築されたイメージHは、下記式(8)のように表される。

【 0 1 0 5 】

10

【数8】

$$H(x, y) = \sum_{i=1}^n a_i J_i(x, y) \quad \dots(8)$$

【 0 1 0 6 】

そして、入力画像Iと再構築された画像HとのエラーEは、下記式(9)のように表される。

20

【 0 1 0 7 】

【数9】

$$E = \frac{1}{2} \|I(x, y) - H(x, y)\|^2 = \frac{1}{2} \sum_{x, y} (I(x, y) - H(x, y))^2 \quad \dots(9)$$

【 0 1 0 8 】

このエラーEを最小にするような最適なaを求めることにより再構築することができる。なお、ガボア・フィルタリングに関しては、認識タスクに応じてフィルタの種類を変更するようにしてもよい。

30

【 0 1 0 9 】

低周波でのフィルタリングでは、フィルタリング後のイメージすべてをベクトルとして持っているのは冗長である。そこで、ダウンサンプリングして、ベクトルの次元を落とすようにしてもよい。ダウンサンプリングされた24種類のベクトルを一行に並べ、長いベクトル(上述したパターンベクトル)にする。

【 0 1 1 0 】

パターン識別部64は、外部から供給される学習用のデータすなわち教師データを用いて、暫定的な識別関数を得た後、当該識別関数を前処理部63からパターンベクトルとして得られた360画素分のスコア画像に試して顔の検出を行う。そして、検出に成功したものを顔データとして出力する。また検出に失敗したものを非顔データとして学習データに追加して、更に学習をし直す。

40

【 0 1 1 1 】

本適用例では、パターン識別部64における顔認識に関して、パターン認識の分野で最も学習汎化能力が高いとされるサポートベクタマシン(Support Vector Machine: SVM)を用いて該当する顔か否かの識別を行う。

【 0 1 1 2 】

サポートベクタマシン自体に関しては、例えばB.sholkopf外著の報告(B.Sholkopf, C.Burges, A.Smola, "Advance in Kernel Support Vector learning", The MIT Press, 1999.)を挙げることができる。本願出願人が行った予備実験の結果からは、サポートベクタ

50

マシンによる顔認識方法は、主成分分析（PCA）やニューラル・ネットワークを用いる手法に比べ、良好な結果を示すことが判っている。

【0113】

サポートベクタマシンは、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2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することができる。

【0114】

通常、パターン認識の問題は、テスト・サンプル $x = (x_1, x_2, \dots, x_n)$ に対して、下記式（10）で与えられる識別関数 $f(x)$ を求めることである。

10

【0115】

【数10】

$$f(x) = \sum_{j=1}^n W_j X_j + b \quad \dots(10)$$

【0116】

ここで、サポートベクタマシンの学習用の教師ラベルを下記式（11）のようにおく。

20

【0117】

【数11】

$$y = (y_1, y_2, \dots, y_n) \quad \dots(11)$$

【0118】

すると、サポートベクタマシンにおける顔パターンの認識を下記式（12）に示す制約条件の下での重み因子 w の2乗の最小化する問題としてとらえることができる。

【0119】

30

【数12】

$$y_i (w^T x_i + b) \geq 1 \quad \dots(12)$$

【0120】

このような制約のついた問題は、ラグランジュの未定定数法を用いて解くことができる。すなわち、下記式（13）に示すラグランジュをまず導入し、次いで、下記式（14）に示すように、 b 、 w の各々について偏微分する。

40

【0121】

【数13】

$$L(w, b, a) = \frac{1}{2} \|W\|^2 - \sum_{j=1}^n a_j (y_j ((x_j^T w + b) - 1)) \quad \dots(13)$$

【0122】

【数14】

50

$$\frac{\partial L}{\partial b} = \frac{\partial L}{\partial w} = 0 \quad \cdots(14)$$

【 0 1 2 3 】

この結果、サポートベクタマシンにおける顔パターンの識別を下記式 (1 5) に示す 2 次計画問題としてとらえることができる。

【 0 1 2 4 】

【 数 1 5 】

10

$$\max \sum a_i - \frac{1}{2} \sum a_i a_i y_i y_i' x_j \quad \cdots(15)$$

$$\text{制約条件: } a_i \geq 0, \sum a_i y_i = 0$$

【 0 1 2 5 】

特徴空間の次元数が、訓練サンプルの数よりも少ない場合は、スクラッチ変数 0 を導入して、制約条件を下記式 (1 6) のように変更する。最適化については、下記式 (1 7) の目的関数を最小化する。

20

【 0 1 2 6 】

【 数 1 6 】

$$y_1 (w' x_1 + b) \geq 1 - \xi_1 \quad \cdots(16)$$

【 0 1 2 7 】

【 数 1 7 】

30

$$\frac{1}{2} \|w\|^2 + C \sum \xi_i \quad \cdots(17)$$

【 0 1 2 8 】

上記式 (1 7) において、C は、制約条件をどこまで緩めるかを指定する係数であり、実験的に値を決定する必要がある。ラグランジュ定数 a に関する問題は下記式 (1 8) のように変更される。

40

【 0 1 2 9 】

【 数 1 8 】

$$\max \sum a_i - \frac{1}{2} \sum a_i a_i y_i y_i' x_j \quad \cdots(18)$$

$$\text{制約条件: } 0 \leq a_i \leq C, \sum a_i y_i = 0$$

【 0 1 3 0 】

しかし、上記式 (1 8) のままでは、非線型の問題を解くことはできない。そこで、本適

50

用例では、カーネル関数 $K(x, x')$ を導入して、一旦、高次元の空間に写像して（カーネル・トリック）、その空間で線形分離することになっている。従って、元の空間では非線形分離していることと同等となる。

【 0 1 3 1 】

カーネル関数は、ある写像 を用いて下記式（ 1 9 ）のように表される。

【 0 1 3 2 】

【 数 1 9 】

$$K(x, y) = \Phi(x)^T \Phi(x') \quad \dots(19)$$

10

【 0 1 3 3 】

また、上記（ 1 0 ）式に示した識別関数も、下記式（ 2 0 ）のように表すことができる。

【 0 1 3 4 】

【 数 2 0 】

$$\begin{aligned} f(\Phi(x)) &= w^T \Phi(x) + b \\ &= \sum a_i y_i K(x, x_i) + b \end{aligned} \quad \dots(20)$$

20

【 0 1 3 5 】

また学習に関しても、下記式（ 2 1 ）に示す 2 次計画問題としてとらえることができる。

【 0 1 3 6 】

【 数 2 1 】

$$\max \sum a_i - \frac{1}{2} \sum a_i a_j y_i y_j K(x_i, x_j) \quad \dots(21)$$

$$\text{制約条件: } 0 \leq a_i \leq C, \sum a_i y_i = 0$$

30

【 0 1 3 7 】

カーネルとしては、下記式（ 2 2 ）に示すガウシアン・カーネル（ R B F : Radius Basic Function ）等を用いることができる。

【 0 1 3 8 】

【 数 2 2 】

$$K(x, x') = \exp\left(-\frac{\|x - x'\|^2}{\sigma^2}\right) \quad \dots(22)$$

40

【 0 1 3 9 】

このようにパターン識別部 6 4 は、前処理部 6 3 から与えられたスコア画像に基づくパターンベクトルについて、当該スコア画像内に顔データが存在するか否かを判断し、存在する場合のみ当該スコア画像の画像領域における左上位置（座標）及びその大きさ（縦横の画素数）と、当該スコア画像の切出し元となるスケール画像のフレーム画像に対する縮小率（すなわち上述の 5 段階のうちの該当する段階）とをリスト化し、これをリストデータ

50

として内部メモリ 381A に格納する。

【0140】

この後、パターン識別部 64 は、ウィンドウ切出部 61 に対して、第 1 のスケール画像のうち先頭のウィンドウ画像の顔検出が終了した旨を通知することにより、当該ウィンドウ切出部 61 から第 1 のスケール画像のうち次にスキャンされたウィンドウ画像をテンプレートマッチング部 62 に送出させる。

【0141】

そしてテンプレートマッチング部 62 は、当該ウィンドウ画像がテンプレートにマッチングした場合のみスコア画像として前処理部 63 に送出する。前処理部 63 は、当該スコア画像をパターンベクトルに変換してパターン識別部 64 に送出する。パターン識別部 64 は、パターンベクトルから識別結果として得られた顔データに基づいてリストデータを生成して内部メモリ 381A に格納する。

10

【0142】

このようにウィンドウ切出部 61 おいて第 1 のスケール画像から切り出した全てのウィンドウ画像について、スキャン順にテンプレートマッチング部 62、前処理部 63 及びパターン識別部 64 の各処理を行うことにより、当該第 1 のスケール画像から撮像結果に存在する顔画像を含むスコア画像を複数検出することができる。

【0143】

この後、パターン識別部 64 は、入力画像スケール変換部 60 に対して、第 1 のスケール画像の顔検出が終了した旨を通知することにより、当該入力画像スケール変換部 60 から第 2 のスケール画像をウィンドウ切出部 61 に送出させる。

20

【0144】

そして第 2 のスケール画像についても、上述した第 1 のスケール画像と同様の処理を行って、当該第 2 のスケール画像から撮像結果に存在する顔画像を含むスコア画像を複数検出した後、第 3 ~ 第 5 のスケール画像についても同様の処理を順次行う。

【0145】

かくしてパターン識別部 64 は、撮像画像であるフレーム画像を 5 段階に縮小した第 1 ~ 第 5 のスケール画像について、当該撮像画像内に存在する顔画像を含むスコア画像をそれぞれ複数検出した後、その結果得られたリストデータをそれぞれ内部メモリ 381A に格納する。この場合、元のフレーム画像内での顔画像のサイズによっては、全くスコア画像が得られない場合もあるが、少なくとも 1 以上 (2 又は 3 以上でもよい) のスケール画像でスコア画像が得られれば、顔検出処理を続行することとする。

30

【0146】

ここで、各スケール画像において顔画像を含む複数のスコア画像は、ウィンドウ切出部 61 におけるスキャンが 2 画素ずつずらして行われたため、実際に顔がある領域とその近傍領域とで高い相関性があり、隣接するスコア画像同士で相互に重なり合う画像領域を含むこととなる。

【0147】

そこで続く重なり判定部 65 は、内部メモリ 381A に格納されている第 1 ~ 第 5 のスケール画像ごとに複数のリストデータをそれぞれ読み出して、当該各リストデータに含まれるスコア画像同士を比較して、相互に重なり合う領域を含むか否かを判定する。

40

【0148】

その際、重なり判定部 65 は、図 7 に示すように、2 つのスコア画像 P1、P2 の左上角部の位置 (座標) がそれぞれ (X_A, Y_A) 、 (X_B, Y_B) であり、その大きさ (縦横の画素数) がそれぞれ $H_A \times L_A$ 、 $H_B \times L_B$ であるとき、 $dX (= X_B - X_A)$ 、 $dY (= Y_B - Y_A)$ として、下記式 (23) を満たすか否かによって、当該スコア画像 P1、P2 同士が重なり合うか否かを判定することができる。

【0149】

【数 23】

$$\begin{aligned} (L_A - dX) \times (L_B + dX) > 0 \\ \text{かつ} (H_A - dY) \times (H_B + dY) > 0 \end{aligned} \quad \dots(23)$$

【0150】

重なり判定部65は、当該判定結果に基づいて、スコア画像同士で重なり合う領域を除去することにより、各スケール画像において、最終的に複数のスコア画像を互いに重なることなく寄せ集めた単一の画像領域として得ることができ、当該画像領域を顔決定データとして新たに内部メモリ381Aに格納する。

10

【0151】

また重なり判定部65は、テンプレートマッチング部62において顔画像でないと判断された場合には、そのまま何もすることなく、内部メモリ381Aの格納も行わない。

【0152】

(2-1-3)第1の適用例における動作及び効果

以上の構成において、このロボット装置200では、CCDカメラ50により撮像したフレーム画像を縮小率が異なる複数のスケール画像に変換した後、当該各スケール画像の中からそれぞれ所定サイズのウィンドウ画像を所定画素ずつずらすようにスキャンしながら1枚ずつ切り出す。

20

【0153】

このウィンドウ画像について、平均的な顔画像を表すテンプレートを用いてマッチングをとって大まかに顔画像であるか否かを判断するようにして、明らかに顔画像でないウィンドウ画像を除去することにより、後段の顔検出処理に要する演算量及び時間をその分減少させることができる。

【0154】

続いてテンプレートマッチングで顔画像であると判断されたウィンドウ画像(すなわちスコア画像)について、当該スコア画像の矩形領域の4隅部分を除去した後、濃淡補正及び続くコントラスト強調の平滑化を行い、更に1本のパターンベクトルに変換する。

【0155】

そして当該パターンベクトルについて、元のスコア画像内での顔検出を行って顔データ又は非顔データを判断し、顔データが存在するスコア画像の画像領域の位置(座標)及びその大きさ(画素数)と、当該スコア画像の切出し元となるスケール画像のフレーム画像に対する縮小率とをリスト化したリストデータを生成する。

30

【0156】

このように各スケール画像毎にそれぞれ全てのスコア画像についてリストデータを生成した後、当該各リストデータに含まれるスコア画像同士を比較して、相互に重なり合う領域を除去した顔決定データを求めることにより、元のフレーム画像から顔画像を検出することができる。

【0157】

このような顔検出タスク処理のうち特にテンプレートマッチング処理は、比較的構成が簡易な演算器にもたやすく実装できる上に、画像圧縮等で利用されるブロックマッチングの手法と類似する処理であることからCPUを用いた高速処理を行うハードウェアが数多く存在する。従ってテンプレートマッチング処理に関してはさらなる高速化が可能である。

40

【0158】

以上の構成によれば、このロボット装置200において、CCDカメラ50により撮像したフレーム画像について顔画像を検出する顔検出タスク処理の際、当該フレーム画像を異なる縮小率で縮小した各スケール画像の中からそれぞれ所定サイズのウィンドウ画像を所定画素ずつずらすようにスキャンしながら1枚ずつ切り出した後、平均的な顔画像を表すテンプレートを用いてマッチングをとって大まかに顔画像であるか否かを判断するよう

50

にして、明らかに顔画像でないウィンドウ画像を除去するようにしたことにより、当該テンプレートマッチングで顔画像であると判断されたスコア画像に対する種々の顔検出処理に要する演算量及び時間をその分減少させることができ、ロボット装置200全体の制御を司る主制御部381の処理負担を軽減させることができ、かくしてリアルタイム性を格段と向上し得るロボット装置200を実現できる。

【0159】

(2-2) 第2の適用例

(2-2-1) ロボットの内部構成

第2の適用例において、胴体部ユニット202(図1)の胴体下部を形成する腰ベースの背面側には、図5との対応部分に同一符号を付した図8に示すように、上述した主制御部381に加えて、この主制御部381と接続されたバスを介して内部メモリ381A、DMA(Direct Memory Access)コントローラ70及び演算処理部(DSP(Digital Signal Processor)及び画像処理用エンジン等)71が相互に接続されており、それ以外は第1の適用例と同様に構成されている。

10

【0160】

この場合、主制御部381のみが内部メモリ381Aの管理を行うのではなく、DMAコントローラ70が主制御部381又は演算処理部71にデータを転送する。一方、当該主制御部381又は演算処理部71から得られた演算結果をDMAコントローラ70を介して内部メモリ381Aに格納するようになされている。

20

【0161】

このようにデータ転送をDMAコントローラ70が行うと共に、データ演算を主制御部381又は演算処理部71に振り分けて互いに独立に行わせることにより、主制御部381にかかる負担を低減させ得るようになされている。

【0162】

(2-2-2) 顔検出タスク機能に関する主制御部381の処理

第2の適用例の場合、顔検出タスク機能は、DMAコントローラ70を介した主制御部381及び演算処理部71における各種処理により実現されている。

【0163】

ここで、かかる顔検出タスク機能に関する主制御部381及び演算処理部71の各処理内容を機能的に分類すると、図9に示すように、入力画像スケール変換部80、ウィンドウ切出部81、テンプレートマッチング部82、重なり判定部83、スケール変換及び切出部84、前処理部85及びパターン識別部86に分けることができる。

30

【0164】

このうち入力画像スケール変換部80、ウィンドウ切出部81及びテンプレートマッチング部82は、演算処理部71が処理すると共に、重なり判定部83、スケール変換及び切出部84、前処理部85及びパターン識別部86は、主制御部381が処理するようになされている。演算処理部71と主制御部381との処理の切り換えはDMAコントローラ70が行う。

【0165】

まず、入力画像スケール変換部80は、CCDカメラ50(図8)からの画像信号S1Aに基づくフレーム画像を内部メモリ381Aから読み出して、当該フレーム画像を縮小率が相異なる複数のスケール画像(上述した第1~第5のスケール画像)に変換する。

40

【0166】

続くウィンドウ切出部81は、第1~第5のスケール画像のうち、まず第1のスケール画像に対して、画像左上を起点として順に画像右下まで、適当な画素(例えば2画素)分を飛ばしながらスキャンするようにして、400(=20×20)画素の矩形領域でなるウィンドウ画像を順次切り出す。

【0167】

その際、ウィンドウ切出部81は、第1のスケール画像から切り出した複数のウィンドウ画像のうち先頭のウィンドウ画像を後段のテンプレートマッチング部82に送出する。

50

【0168】

テンプレートマッチング部82は、ウィンドウ切出部81から得られた先頭のウィンドウ画像について、平均的な顔画像をテンプレートとし、大まかなマッチングをとって、当該テンプレートとの相関度（テンプレートのうちマッチングする画素数の割合）を相関度データとして当該ウィンドウ画像と共にDMAコントローラ70を介して内部メモリ381Aに格納する。

【0169】

具体的には、図10に示すように、任意のスケール画像F1から切り出したウィンドウ画像W1について、平均的な顔画像であるテンプレートT1をとり、当該テンプレートT1との相関度をとった結果として、当該相関度を表す画像R1が検出され、これを相関度データとして内部メモリ381Aに格納するようになされている。

10

【0170】

この第2の適用例の場合には、ウィンドウ画像について顔画像であると判断されたスコア画像以外をふるい落とす第1の適用例とは異なり、高速処理を実現すべく、当該ウィンドウ画像をスコア画像か否かにかかわらず後段の重なり判定部83に送出する。

【0171】

この後、テンプレートマッチング部82は、ウィンドウ切出部81に対して、第1のスケール画像のうち先頭のウィンドウ画像のテンプレートマッチングが終了した旨を通知することにより、当該ウィンドウ切出部81から第1のスケール画像のうち次にスキャンされたウィンドウ画像をテンプレートマッチング部82に送出させる。

20

【0172】

このようにテンプレートマッチング部82は、ウィンドウ切出部81において第1のスケール画像から切り出した全てのウィンドウ画像について、スキャン順にそれぞれ相関度データを生成する。

【0173】

この後、テンプレートマッチング部82は、入力画像スケール変換部80に対して、第1のスケール画像の顔検出が終了した旨を通知することにより、当該入力画像スケール変換部80から第2のスケール画像をウィンドウ切出部81に送出させる。

【0174】

そして第2のスケール画像についても、上述した第1のスケール画像と同様の処理を行って、当該第2のスケール画像から全てのウィンドウ画像に対応する相関度データを生成した後、第3～第5のスケール画像についても同様の処理を順次行う。

30

【0175】

かくしてテンプレートマッチング部82は、撮像画像であるフレーム画像を5段階に縮小した第1～第5のスケール画像について、それぞれ切り出した複数のウィンドウ画像に対応する相関度データを生成して、当該ウィンドウ画像と共にDMAコントローラ70を介して内部メモリ381Aに格納する。

【0176】

以上が演算処理部71における処理内容であり、DMAコントローラ70の制御に応じて、主制御部381に切り換えられ、以下は当該主制御部381における処理内容となる。

40

【0177】

まず、重なり判定部83は、図11の重なり判定処理手順R11をステップSP0から実行し、第1～第5のスケール画像毎に、当該各スケール画像内のウィンドウ画像を切り出す際のスキャン順と同じ順番で、ウィンドウ画像をこれに付随する相関度データと共にDMAコントローラ70を介して内部メモリ381Aから読み出す（ステップSP1）。

【0178】

続いて重なり判定部83は、該当するスケール画像における複数のウィンドウ画像について、スキャン順にウィンドウ画像同士を上述した式(23)を用いて比較して、相互に重なり合う領域を含むか否かを判定する（ステップSP2）。

【0179】

50

そして重なり判定部 8 3 は、相互に重なり合う領域を含む場合のみ、当該ウィンドウ画像同士をそれぞれ対応する相関度の大小で比較して、相関度の大きいほうをスコア画像として、内部メモリ 3 8 1 A 内に設定する候補リストに追加して格納する（ステップ S P 3）

。

【 0 1 8 0 】

そして重なり判定部 8 3 は、該当するスケール画像内の全てのウィンドウ画像について、相関度を比較した場合の大きい方をスコア画像として候補リストに追加する（ステップ S P 4）。一方、小さい方は候補リストに加えることなく又は既に候補リストに存在する場合には削除する（ステップ S P 5）ようにして順次候補リストの書き換えを行う。

【 0 1 8 1 】

やがて重なり判定部 8 3 は、該当するスケール画像内の全てのウィンドウ画像について、比較的相関度が高い複数のスコア画像を候補リストとして決定した後、他のスケール画像についても比較的相関度が高い複数のスコア画像を候補リストとして決定する。なお、候補リストとして追加される比較的相関度が高いスコア画像の数は、各スケール画像毎にそれぞれ 1 0 以下に絞られることが実験結果として得ることができた。

【 0 1 8 2 】

このようにして重なり判定部 8 3 では、第 1 ～ 第 5 のスケール画像について、スコア画像を相関度の最も高いピーク値のみならず、そのピーク値の近傍領域をも対象として候補リストに加えるようにしたことにより、テンプレートマッチング部 8 2 でのマッチング結果と後段のパターン識別部 8 6 におけるサポートベクタマシンの処理結果とのずれをその分低下させることができる。

【 0 1 8 3 】

實際上、かかる重なり判定部 8 3 における候補リストの作成処理は、数回の加算及び乗算によって実行可能であるため、後段の前処理部 8 5 及びパターン識別部 8 6 による演算処理と比較すると、極めて高速に短時間に処理することができる。そして候補リストとして残ったスコア画像について前処理部 8 5 及びパターン識別部 8 6 による処理を行えば、当該前処理部 8 5 及びパターン識別部 8 6 においても処理負担を軽減させることができる。

【 0 1 8 4 】

この後、スケール変換及び切出部 8 4 は、重なり判定部 8 3 において各スケール画像毎に候補リストとして内部メモリ 3 8 1 A に格納された複数のスコア画像について、当該内部メモリ 3 8 1 A から D M A コントローラ 7 0 を介して第 1 のスケール画像内でかつ候補リストの決定順に 1 つずつスコア画像を読み出して、前処理部 8 5 に送出する。

【 0 1 8 5 】

前処理部 8 5 は、スケール変換及び切出部 8 4 から得られたスコア画像について、矩形領域でなる当該スコア画像（4 0 0 画素）から人間の顔画像とは無関係である背景部分に相当する 4 隅の領域を除去するために、当該 4 隅の領域を切り取ったマスクを用いて 3 6 0 画素分を抽出する。

【 0 1 8 6 】

そして前処理部 8 5 は、撮像時の照明により濃淡で表される被写体の傾き条件を解消すべく、当該抽出した 3 6 0 画素分のスコア画像のうち顔画像として最適な部位を基準とする平面を形成するように画素単位で濃淡値に補正をかける。

【 0 1 8 7 】

続いて前処理部 8 5 は、当該 3 6 0 画素分のスコア画像のコントラストを強調した結果をヒストグラム平滑化処理を行うことにより、C C D カメラ 5 0 のゲインや照明の強弱によらずに検出できるようにする。

【 0 1 8 8 】

次いで、前処理部 8 5 は、上述したガボア・フィルタリング処理を行うことにより、当該 3 6 0 画素分のスコア画像をベクトル変換し、得られたベクトル群を更に 1 本のパターンベクトルに変換する。

【 0 1 8 9 】

パターン識別部 8 6 は、上述したサポートベクタマシンを用いて、前処理部 8 5 からパターンベクトルとして得られた 3 6 0 画素分のスコア画像に試して顔か否かの識別を学習しながら行う。

【 0 1 9 0 】

このようにパターン識別部 8 6 は、前処理部 8 5 から与えられたスコア画像に基づくパターンベクトルについて、当該スコア画像内での顔データに相当する画像領域の位置（座標）及びその大きさ（画素数）と、当該スコア画像の切出し元となるスケール画像のフレーム画像に対する縮小率（すなわち上述の 5 段階のうちの該当する段階）とをリスト化し、これをリストデータとして内部メモリ 3 8 1 A に格納する。

【 0 1 9 1 】

この後、パターン識別部 8 6 は、スケール変換及び切出部 8 4 に対して、第 1 のスケール画像のうち最初のスコア画像の顔検出が終了した旨を通知することにより、当該スケール変換及び切出部 8 4 から第 1 のスケール画像のうち次のスコア画像を前処理部 8 5 に送出する。前処理部 8 5 は、当該スコア画像をパターンベクトルに変換してパターン識別部 8 6 に送出する。パターン識別部 8 6 は、パターンベクトルから得られた顔データに基づいてリストデータを生成して内部メモリ 3 8 1 A に格納する。

【 0 1 9 2 】

このようにスケール変換及び切出部 8 4 において第 1 のスケール画像内で候補リストにある全てのスコア画像について、順番に前処理部 8 5 及びパターン識別部 8 6 の各処理を行うことにより、当該第 1 のスケール画像から撮像結果に存在する顔画像を検出することができる。

【 0 1 9 3 】

この後、パターン識別部 8 6 は、スケール変換及び切出部 8 4 に対して、第 1 のスケール画像の顔検出が終了した旨を通知することにより、当該第 2 のスケール画像についても、上述した第 1 のスケール画像と同様の処理を行って、当該第 2 のスケール画像から撮像結果に存在する顔画像を検出した後、第 3 ～ 第 5 のスケール画像についても同様の処理を順次行う。

【 0 1 9 4 】

かくしてパターン識別部 8 6 は、撮像画像であるフレーム画像を 5 段階に縮小した第 1 ～ 第 5 のスケール画像について、当該撮像画像内に存在する顔画像をそれぞれ検出した後、その結果得られた顔データに基づいてそれぞれリストデータを生成して内部メモリ 3 8 1 A に格納する。

【 0 1 9 5 】

（ 2 - 2 - 3 ）第 2 の適用例における動作及び効果

以上の構成において、このロボット装置 2 0 0 では、CCD カメラ 5 0 により撮像したフレーム画像を縮小率が異なる複数のスケール画像に変換した後、当該各スケール画像の中からそれぞれ所定サイズのウィンドウ画像を所定画素ずつずらすようにスキャンしながら 1 枚ずつ切り出す。

【 0 1 9 6 】

このウィンドウ画像について、平均的な顔画像を表すテンプレートを用いてマッチングをとって当該テンプレートとの相関度を表す相関度データを生成する。このように各スケール画像毎にそれぞれ全てのウィンドウ画像についてスキャン順にそれぞれ相関度データを生成する。

【 0 1 9 7 】

ここまです演算処理部 7 1 において実行すると共にこれ以降の処理を主制御部 3 8 1 において実効することにより、その分当該主制御部 3 8 1 の処理負担を軽減することができる。

【 0 1 9 8 】

続いて各スケール画像における複数のウィンドウ画像について、スキャン順にウィンドウ画像同士を比較して、相互に重なり合う領域を含む場合には、更に相関度を比較して大き

10

20

30

40

50

い方のみをスコア画像として残るように候補リストに追加する。この結果、各スケール画像毎に、比較的相関度が高い複数のスコア画像を候補リストとして決定することができ、後段の顔識別処理の際に識別精度を一層向上させることができる。

【0199】

続いて各スケール画像毎に候補リストとして内部メモリ381Aに格納された複数のスコア画像について、当該スコア画像の矩形領域の4隅部分を除去した後、濃淡補正及び続くコントラスト強調の平滑化を行い、更に1本のパターンベクトルに変換する。

【0200】

そして当該パターンベクトルについて、元のスコア画像内での顔検出を行って顔データ又は非顔データを判断し、顔データが存在するスコア画像の画像領域の位置(座標)及びその大きさ(画素数)と、当該スコア画像の切出し元となるスケール画像のフレーム画像に対する縮小率とをリスト化したリストデータを生成する。

10

【0201】

このように各スケール画像毎にそれぞれ全てのスコア画像についてリストデータを生成することにより、元のフレーム画像から顔画像を検出することができる。

【0202】

かくして顔検出タスク処理に関して、ロボット装置200の全体の制御を司る主制御部381に加えて、演算処理部71を設け、内部メモリ381Aに対するデータの読み書きをDMAコントローラ70の制御を介して主制御部381及び演算処理部71の双方が行い得るようにしたことにより、当該主制御部381及び演算処理部71がそれぞれ処理を分配することができる分、主制御部381にかかる演算量及び演算時間の負担を格段と低減させることができる。

20

【0203】

以上の構成によれば、このロボット装置200において、CCDカメラ50により撮像したフレーム画像について顔画像を検出する顔検出タスク処理の際、当該フレーム画像を相異なる縮小率で縮小した各スケール画像の中からそれぞれ所定サイズのウィンドウ画像を所定画素ずつずらすようにスキャンしながら1枚ずつ切り出した後、平均的な顔画像を表すテンプレートを用いてマッチングをとって当該テンプレートとの相関度を表す相関度データをそれぞれ生成し、当該ウィンドウ画像同士で相互に重なり合う領域を含む場合には相関度を比較して大きい方のみを候補リストに残るようにしたことにより、当該候補リストにある比較的相関度が高い複数のスコア画像に対する種々の顔検出処理に要する演算量及び時間をその分減少させることができるのみならず、顔識別処理の際の識別精度を一層向上させることができる。

30

【0204】

これに加えて顔検出タスク処理に関して、ロボット装置200の全体の制御を司る主制御部381と演算処理部71とで処理を分担して行うようにしたことにより、当該主制御部381にかかる演算量及び演算時間の負担を格段と低減させることができ、かくしてリアルタイム性を格段と向上し得るロボット装置200を実現できる。

【0205】

(3)実施の形態

40

次に、本発明の実施の形態について説明する。上述したように、テンプレートマッチングを行って顔候補を抽出し(第1の工程)、この顔候補の中からSVM等により顔領域を判定して(第2の工程)顔領域を検出する方法において、第1の工程においては、単純に正規化相関値の代償により顔候補を決定しているため、顔候補の見逃しを軽減しようとした場合、閾値を上げる方法又は間引きを減らす方法をとることができるものの、閾値を下げると演算量が増大してしまい、ロボット装置等のリソースの限られた環境においては好ましくない場合がある。一方、閾値を上げると、第2の工程において顔判定するための候補画像が減るため、演算量を減らすことができるものの、本来顔である画像も候補画像から取り除いてしまい、顔画像を見逃してしまう場合がある。

【0206】

50

そこで、本願発明者等は、テンプレートと同一サイズの顔領域（顔画像）が入力画像内に存在する場合、この顔画像とテンプレートとの相関をとれば、テンプレートサイズ近傍では最も相関値が大きくなることに着目し、顔領域の候補を絞り込む際に、局所的な絞り込みを行うアルゴリズムを使用することにより、本来顔である画像を見逃すことなく顔候補画像を低減して後段の第２の工程にて顔判定する計算量を低減する方法を見出した。具体的には、入力画像と所定サイズの平均顔のテンプレートとの正規化相関をとった相関値の集合であるマッチング結果における相関値の局所最大値に基づき候補となる顔領域を抽出するものである。以下、この顔検出方法について詳細に説明する。

【０２０７】

本実施の形態における顔検出装置は、上述の第１及び第２の適用例と同様の２足歩行のロボット装置等のＣＰＵ及びメモリ等のリソースが限られたシステムにおいて、性能を保持しつつ顔検出に要する計算量を削減する場合に好適である。本実施の形態においても、上述の第１及び第２の適用例と同様、図１乃至図４に示すロボット装置２００に搭載されるものとし、その内部構成は、上述の適用例と同様とし、その詳細な説明は省略する。

【０２０８】

（３－１）顔検出タスク機能に関する主制御部３８１の処理

本実施の形態におけるロボット装置２００に搭載された顔検出タスク機能について説明する。上述した如く、ロボット装置２００には、ＣＣＤカメラ５０を介して内部メモリ３８１Ａに記憶されるフレーム画像について、当該フレーム画像の中から人間の顔画像を検出する顔検出タスク機能が搭載されている。そしてこの顔検出タスク機能は、主制御部３８１における各種処理により実現されている。

【０２０９】

ここで、かかる顔検出タスク機能に関する主制御部３８１の処理内容を機能的に分類すると、図１２に示すように、テンプレートサイズ及び入力画像スケール変換決定部９０、ウィンドウ切り出し部９１、テンプレートマッチング部９２、スケール変換及び切り出し部９３、前処理部９４、及びパターン識別部９５に分けることができる。

【０２１０】

入力画像スケール変換決定部９０は、第１の適用例と同様に、ＣＣＤカメラ５０（図５）からの画像信号Ｓ１Ａに基づくフレーム画像を内部メモリ３８１Ａから読み出して、当該フレーム画像を縮小率が相異なる複数のスケール画像（第１～第５のスケール画像）に変換すると共に、後段のテンプレートマッチング部９２で使用するテンプレートサイズ、即ち、顔画像のサイズを選択する（以下、これを第１のテンプレートサイズという）。

【０２１１】

続くウィンドウ切出部９１は、第１～第５のスケール画像のうち、まず第１のスケール画像に対して、画像左上を起点として順に画像右下まで、適当な画素（例えば２画素）分を飛ばしながらスキャンするようにして、４００（＝２０×２０）画素の矩形領域でなるウィンドウ画像を順次切り出す。

【０２１２】

その際、ウィンドウ切出部９１は、第１のスケール画像から切り出した複数のウィンドウ画像のうち先頭のウィンドウ画像を、スケール変換決定部９１で選択した第１のテンプレートサイズのテンプレートと共に後段のテンプレートマッチング部９２に送出する。

【０２１３】

テンプレートマッチング部９２は、ウィンドウ切出部９１から得られた先頭のウィンドウ画像について、第１のテンプレートサイズを有する平均的な顔画像をテンプレートとして大まかなマッチングをとり、当該テンプレートとの相関値（テンプレートのうちマッチングする画素数の割合）の集合をマッチング結果として当該ウィンドウ画像と共にＤＭＡコントローラ７０を介して内部メモリ３８１Ａに格納する。

【０２１４】

即ち、図１３（ａ）に示すように、任意のスケール画像から切り出した、例えば高さ（ｙ軸方向の辺の長さ）hei_s×幅（ｘ軸方向の辺の長さ）weid_sのウィンドウ画像（スケー

10

20

30

40

50

ル変換後の入力画像) W_2 について、図 13 (b) に示すように、例えば高さ $hei_t \times$ 幅 wid_s である第 1 のテンプレートサイズを有する平均的な顔画像であるテンプレート T_2 を使用し、ウィンドウ画像 W_2 をスキャンし、所定画素 (例えば 1 画素) ずつずらしながら移動させたテンプレート T_2 と上記入力画像との相関値の集合であるマッチング結果を求める。このマッチング結果は、テンプレート T_2 の移動に伴い相関値が 2 次元に配列されたものであり、図 14 に示すように、当該相関値を表す高さ $hei_r \times$ 幅 wid_r のテンプレートマッチング結果画像 R_2 が得られる。ここで、テンプレートマッチング結果画像 R_2 の高さ hei_r は、 $hei_s - (hei_t + 1)$ であり、画像 R_2 の幅 wid_s は、 $wid_s - (wid_t + 1)$ となる。

【0215】

10

次に、このテンプレートマッチング結果画像 R_2 を所定のサイズ、例えば第 1 のテンプレートサイズと同一の大きさに分割し、各第 1 のテンプレートサイズに仕切られた分割領域毎に相関値の最大値を有する点 (位置) を求め、これら各分割領域から得られた最大値を示す点のうち、所定の閾値以上のものを顔候補として抽出する。

【0216】

即ち、平均顔のテンプレートを使用して正規化相関をしようした場合、必ずしも任意のパターンより、顔画像の方が相関値が高くなるという保証はないものの、テンプレートと同一のサイズの顔画像が存在する場合は、テンプレートサイズ近傍の大きさでは相関値が最大値をとることから、相関値が分割領域内で最大値となり、且つ所定の閾値以上の点を顔候補として抽出することにより、単にテンプレートマッチングの結果、相関値が所定の閾値以上であるものを顔候補として抽出する場合に比して、顔候補をより有効に絞り込むことができる。

20

【0217】

次に、この顔候補として抽出された点を顔候補 (第 1 の顔候補) とすると共に、この顔候補近傍の点も顔候補 (第 2 の顔候補) として抽出する。後段のパターン認識部 95 において、上述した SVM を使用して顔判定する場合、SVM は 1 画素ずれるのみで顔検出ができなくなるほど敏感であるため、本実施の形態においては、顔候補の近傍も検索範囲 (第 2 の顔候補) とすることにより、顔候補の近傍を重点的に検索して顔検出性能を向上することができる。ここで第 1 の顔候補として抽出された点の上下左右に隣接する 8 点全てを検索範囲とすると、後段の計算量が増加するため、所定の条件を満たす場合にのみ、検索範囲に指定することにより、後段の計算量の増加を抑えつつ検出性能を向上させることができる。

30

【0218】

即ち、第 1 の顔候補として抽出された点に隣接する点において、これらの点に対応するテンプレートサイズの入力画像領域において、肌色領域の占有率 (肌色占有率) が所定の閾値以上である場合、又は、予め検出された (又は予め学習された) 顔色情報を有し、この顔色領域の占有率 (顔色占有率) が所定の閾値以上である場合にのみ、その点を検索範囲に設定することができる。ここで、肌色占有率は、例えば、肌色カラーテーブルを使用し、この肌色と比較することにより求めることができる。

【0219】

40

この後、テンプレートマッチング部 92 は、ウィンドウ切出部 91 に対して、第 1 のスケール画像のうち先頭のウィンドウ画像のテンプレートマッチングが終了した旨を通知することにより、当該ウィンドウ切出部 91 から第 1 のスケール画像のうち次にスキャンされたウィンドウ画像をテンプレートマッチング部 92 に送出させる。

【0220】

このようにテンプレートマッチング部 92 は、ウィンドウ切出部 91 において第 1 のスケール画像から切り出した全てのウィンドウ画像について、スキャン順にそれぞれ顔候補を検出する。

【0221】

以上が演算処理部 71 における処理内容であり、DMA コントローラ 70 の制御に応じて

50

、主制御部 3 8 1 に切り換えられ、以下は当該主制御部 3 8 1 における処理内容となる。

【 0 2 2 2 】

スケール変換及び切出部 9 3 は、各スケール画像毎に候補リストとして内部メモリ 3 8 1 A に格納された複数のスコア画像について、当該内部メモリ 3 8 1 A から D M A コントローラ 7 0 を介して第 1 のスケール画像内でかつ候補リストの決定順に 1 つずつスコア画像を読み出して、前処理部 9 4 に送出する。

【 0 2 2 3 】

前処理部 9 4 は、スケール変換及び切出部 9 3 から得られたスコア画像について、矩形領域でなる当該スコア画像 (4 0 0 画素) から人間の顔画像とは無関係である背景部分に相当する 4 隅の領域を除去するために、当該 4 隅の領域を切り取ったマスクを用いて 3 6 0 画素分を抽出する。

10

【 0 2 2 4 】

そして前処理部 9 4 は、撮像時の照明により濃淡で表される被写体の傾き条件を解消すべく、当該抽出した 3 6 0 画素分のスコア画像のうち顔画像として最適な部位を基準とする平面を形成するように画素単位で濃淡値に補正をかける。

【 0 2 2 5 】

続いて前処理部 9 4 は、当該 3 6 0 画素分のスコア画像のコントラストを強調した結果をヒストグラム平滑化処理を行うことにより、C C D カメラ 5 0 のゲインや照明の強弱によらずに検出できるようにする。

【 0 2 2 6 】

20

次いで、前処理部 9 4 は、上述したガボア・フィルタリング処理を行うことにより、当該 3 6 0 画素分のスコア画像をベクトル変換し、得られたベクトル群を更に 1 本のパターンベクトルに変換する。

【 0 2 2 7 】

パターン識別部 (識別手段) 9 5 は、上述したサポートベクタマシンを用いて、前処理部 9 4 からパターンベクトルとして得られた 3 6 0 画素分のスコア画像に試して顔か否かの識別を学習しながら行う。

【 0 2 2 8 】

このようにパターン識別部 9 5 は、前処理部 9 4 から与えられたスコア画像に基づくパターンベクトルについて、当該スコア画像内での顔データに相当する画像領域の位置 (座標) 及びその大きさ (画素数) と、当該スコア画像の切出し元となるスケール画像のフレーム画像に対する縮小率 (すなわち上述の 5 段階のうちの該当する段階) と、テンプレートのサイズとをリスト化し、これをリストデータとして内部メモリ 3 8 1 A に格納する。

30

【 0 2 2 9 】

この後、パターン識別部 9 5 は、スケール変換及び切出部 9 3 に対して、第 1 のスケール画像のうち最初のスコア画像の顔検出が終了した旨を通知することにより、当該スケール変換及び切出部 9 3 から第 1 のスケール画像のうち次のスコア画像を前処理部 9 4 に送出する。前処理部 9 5 は、当該スコア画像をパターンベクトルに変換してパターン識別部 9 5 に送出する。パターン識別部 9 5 は、パターンベクトルから得られた顔データに基づいてリストデータを生成して内部メモリ 3 8 1 A に格納する。

40

【 0 2 3 0 】

このようにスケール変換及び切出部 9 3 において、第 1 のスケール画像内で候補リストにある全てのスコア画像について、順番に前処理部 9 4 及びパターン識別部 9 5 の各処理を行うことにより、当該第 1 のスケール画像から撮像結果に存在する顔画像を検出することができる。

【 0 2 3 1 】

即ち、テンプレートマッチング部 9 2 は、入力画像スケール変換部 9 0 に対して、第 1 のスケール画像及び第 1 のテンプレートサイズのテンプレートを使用した顔検出が終了した旨を通知することにより、当該入力画像スケール変換部 9 0 から、第 1 のスケール画像及び第 2 のテンプレートサイズのテンプレートをウィンドウ切り出し部 9 1 に送出させる。

50

第2のテンプレートサイズを使用した場合も、上述した第1のテンプレートサイズのテンプレートを使用した場合と同様の処理を行って全てのテンプレートサイズに対応する顔候補を検出した後、テンプレートマッチング部92は、入力画像スケール変換部90の第1のスケール画像に対して全てのテンプレートサイズを使用した顔検出が終了した旨を通知することにより、当該入力画像スケール変換決定部90から第2のスケール画像をウィンドウ切出部91に送出させる。

【0232】

そして第2のスケール画像についても、上述した第1のスケール画像と同様の処理を行って、当該第2のスケール画像から全てのウィンドウ画像に対応する顔候補を検出した後、第3～第5のスケール画像についても同様の処理を順次行う。

10

【0233】

かくしてテンプレートマッチング部92は、撮像画像であるフレーム画像を5段階に縮小した第1～第5のスケール画像及び複数のテンプレートサイズのテンプレートについて、それぞれ切出した複数のウィンドウ画像に対応する顔候補を抽出して、当該ウィンドウ画像と共にDMAコントローラ70を介して内部メモリ381Aに格納する。

【0234】

ここで、本実施の形態においては、任意の大きさのテンプレートを使用することができるが、使用するテンプレートサイズを切り替えて、テンプレートサイズを選択することにより、入力画像に対して準備できる全てのテンプレートサイズに対して演算をする場合に比べて、演算量を減らして高効率化することができる。例えば、一度顔が検出された場合に、次に顔検出する際はそのテンプレートサイズを使用することができる。また、例えば、ロボット装置に設けられた距離センサを使用し、この距離センサからの距離情報に基づき入力画像に含まれる対象物との間の距離を認識することにより、対象物の顔領域の大きさを予測してテンプレートサイズを選択する対象距離切り替え手段を設ける等することができ、目的に応じてテンプレートサイズを切り替えることができる。

20

【0235】

(3-2) 実施の形態における動作

以上の構成において、このロボット装置200では、CCDカメラ50により撮像したフレーム画像を縮小率が異なる複数のスケール画像に変換した後、当該各スケール画像の中からそれぞれ所定サイズのウィンドウ画像を所定画素ずつずらすようにスキャンしながら1枚ずつ切り出す。

30

【0236】

このウィンドウ画像について、平均的な顔画像を表すテンプレートを用いてマッチングをとって当該テンプレートとの相関値の集合であるマッチング結果画像を生成する。このように各スケール画像毎にそれぞれ全てのウィンドウ画像についてスキャン順にそれぞれマッチング結果画像を生成する。以下、マッチング結果画像から顔候補を検出する工程について詳細に説明する。

【0237】

図15は、テンプレートマッチング部92において、テンプレートマッチング結果画像R2から顔候補となる画素を検出する各処理工程を示すフローチャートである。図15に示すように、まず、テンプレートマッチング結果画像R2が入力されると、マッチング結果画像R2をテンプレートサイズに分割し、その分割領域の1つ、例えば0 x wid_t - 1、0 y hei_t - 1において、最も相関値が高い点(座標)を抽出する(ステップS P11)。以下、マッチング結果画像R2をテンプレートサイズに分割した領域を分割領域r n、分割領域r nにおいて、相関値が最も大きい点(座標)をlocal_max(x, y)という。ここでは、この各分割領域内において最も相関値が高い画素を抽出するが、本実施の形態においては、マッチング結果画像において分割された分割領域を左から右へ一行ずつ順に処理を行う場合について説明する。

40

【0238】

次に、local_max(x, y)が所定の閾値(th1)より大きいかな否かを判定し(ステッ

50

ステップS P 1 2)、大きい場合は、顔候補として追加する(ステップS P 1 3)。上述した如く、入力画像スケール変換決定部90においては、スケールと共に入力画像に含まれると想定される顔の大きさのテンプレートサイズを選択するが、テンプレートサイズは異なる大きさの複数種類あり、複数種類ある各テンプレートサイズ毎にマッチング結果画像R2を算出して顔候補を抽出すると、同一の点が抽出される場合がある。従って、ステップS P 1 3において、顔候補として同一の点がある場合、即ち、異なるテンプレートサイズで顔候補を抽出した際に既に抽出されている場合はこの点は追加しない。

【0239】

次に、顔候補として抽出された点に対応するテンプレートサイズの入力画像領域において、この画像領域内に含まれる肌色画素の占有率を求める。本実施の形態においては、肌色画素の占有率を求める際に、肌色カラーテーブル100を参照する。そして、この肌色画素占有率が所定の閾値(th2)より大きいかなかを判定する(ステップS P 1 4)。大きい場合は、このlocal_max(x, y)の周辺、例えば上下左右の8近傍点を顔候補として追加する(ステップS P 1 5)。ここで、ステップS P 1 3と同様に、既にこれらの8近傍点が既に顔候補として抽出されている場合は、候補に追加しない。

【0240】

ステップS P 1 2でlocal_max(x, y)が閾値th1未満だった場合、ステップS P 1 4でlocal_max(x, y)に相当する入力画像における肌色画素占有率が閾値th2未満であった場合、及びステップS P 1 5で顔候補の追加が終了した後は、いずれもステップS P 1 6に進み、次の顔候補を抽出するために次の分割領域に移り、処理を進める。

【0241】

先ず、マッチング結果画像R2において、x方向にテンプレートサイズ分、即ち、wid_tだけずれた隣の分割領域に移る(ステップS P 1 6)。次に、wid_tだけずれたx座標(x + wid_t)の分割領域において、そのx座標がマッチング結果画像の幅(x方向の辺)wid_rより大きい場合は、分割領域がマッチング結果画像に含まれないことを示し、次の行に移り、0 ≤ x ≤ wid_t - 1であって、y方向にテンプレートサイズ分、即ち、hei_tだけずれた隣の分割領域に移る(ステップS P 1 8)。次に、分割領域のy座標がマッチング結果画像の高さ(y方向の辺)hei_rより大きいかなかを判定し(ステップS P 1 9)、大きい場合は、マッチング結果画像における全ての分割領域の相関値の最大値を求めたことを示し、処理を終了する。

【0242】

一方、ステップS P 1 7及びステップS P 1 8において、分割領域がマッチング結果画像に含まれると判定された場合は、再びステップS P 1 1に戻り、その分割領域内で最も高い相関値を有する点を抽出する。

【0243】

本実施の形態においては、マッチング結果画像R2をテンプレートサイズに区切った分割領域における相関値の最大値を求めているため、ステップS P 1 6において、隣接する分割領域に移る場合は、x方向にwid_tだけずれるものとしたが、マッチング結果画像R2は、テンプレートサイズ以下のサイズであれば、任意の大きさに分割することができる。その際、分割する画像の大きさの幅(x方向の辺)wid_step、高さ(y方向)hei_stepとすると、ステップS P 1 6及びステップS P 1 8において、夫々x方向にwid_step、又はy方向にhei_step移動することにより、次の分割領域に進むことができる。

【0244】

こうして、テンプレートマッチング部92において、顔候補が抽出される。ここまです演算処理部71において実行すると共にこれ以降の処理を主制御部381において実効することにより、その分当該主制御部381の処理負担を軽減することができる。以降のスケール変換及び切り出し部93、前処理部94及びパターン認識部95を経て顔データを抽出する処理は、上述の第2の適用例と同様である。

【0245】

(3-3) 本実施の形態の効果

10

20

30

40

50

図 16 は、テンプレートマッチング部 92 において、ウィンドウ画像 W2 から顔候補として検出された点を示す図である。図 16 において、白で示す点が、図 14 に示すマッチング結果画像 R2 から顔候補として抽出された点である。比較として、図 17 は、マッチング結果画像 R2 において、閾値以上である点を全て顔候補として抽出した例を示す図である。図 17 に示す図と比較すると、本実施の形態において、テンプレートマッチング部 92 にて顔候補として抽出される点が飛躍的に少なくなっているのがわかる。これにより、後段の処理における計算量を飛躍的に削減することができる。

【0246】

本実施の形態においては、ウィンドウ画像について、平均的な顔画像を表すテンプレートを用いてマッチングをとって大まかに顔画像であるか否かを判断する際に、テンプレートマッチング結果画像を所定のサイズに仕切り、相関値の最大値を顔候補として抽出して明らかに顔画像でないウィンドウ画像を除去することにより、本来顔である領域を見逃すことなく、後段の顔検出処理に要する演算量及び時間を減少させることができ、かくしてリアルタイム性を格段と向上した顔検出装置及びこれを搭載したロボット装置を実現することができる。

10

【0247】

また、相関値が最大となる点と共にその周囲においても顔検索範囲とすることにより、顔検出精度を向上することができる。更に、所定の閾値以上の肌色占有率又は顔の色占有率を有する場合のみ、顔検索範囲として設定することにより、顔検出精度を保ちつつ顔候補を減らして後段の演算量を減らすことができる。更にまた、テンプレートのサイズを適宜切り替えることにより、更に演算量を減らすことができる。

20

【0248】

(4) 他の実施の形態

なお上述の実施の形態においては、本発明を図 1 のように構成された 2 足歩行型のロボット装置 200 に適用するようにした場合について述べたが、本発明はこれに限らず、この他種々のロボット装置及びロボット装置以外のこの他種々の装置に広く適用することができる。例えば、ロボット装置は、4 足歩行であってもよく、更に、移動手段は、脚式移動方式に限定されない。

【0249】

また、CCD カメラ (撮像手段) 50 による撮像結果として得られるフレーム画像を複数の異なる縮小率でなるスケール画像に縮小変換するスケール変換手段として、図 12 に示す主制御部 381 の処理機能のうちの入力画像スケール変換部 90 を適用するようにした場合について述べたが、本発明はこれに限らず、要は、フレーム画像を複数の異なる縮小率で縮小変換した複数のスケール画像を得ることができれば、この他種々の構成のものを適用するようにしてもよい。

30

【0250】

更に、1 枚のフレーム画像を 0.8 倍ずつ縮小変換した 5 種類のスケール画像を適用するようにしたが、主制御部 381 の処理能力の範囲内であれば、2 ~ 4 種類でも 6 種類以上でも良く、また縮小率も自由に設定するようにしてもよい。この場合、複数のスケール画像を基にして顔検出処理を行えば、撮像結果であるフレーム画像に実際に写っている顔のサイズにかかわらず (すなわちロボット装置 200 と人間との距離)、当該フレーム画像から直接顔画像を検出する場合よりも極めて高い確率で顔画像を検出することができる。

40

【0251】

また、CCD カメラ (撮像手段) 50 による撮像結果として得られるフレーム画像内を、400 画素のウィンドウ画像を 1 画素ずつずらすようにスキャンしながら順次抽出する抽出手段を、図 12 に示す主制御部 381 の処理機能のうちのウィンドウ切出部 91 により構成するようにした場合について述べたが、本発明はこれに限らず、要は、フレーム画像又はスケール画像から所定サイズのウィンドウ画像を所定画素ずつずらすようにスキャンしながら順次抽出する得ることができれば、この他種々の構成のものを適用するようにし

50

てもよい。

【 0 2 5 2 】

この場合、スキャンの順番は矩形画像の左上位置から右下位置まで以外にも当該矩形画像全体を走査することができればどのように設定してもよい。またスキャン時にずらす画素数は1画素でも3画素以上でも良く、ウィンドウ画像のサイズも400画素以外に種々の縦横比の所望の画素数に設定するようにしてもよい。

【 0 2 5 3 】

更に、ウィンドウ画像から当該顔画像を識別する識別手段として、図12に示す主制御部381の処理機能のうちの前処理部95及びパターン識別部96を適用するようにした場合について述べたが、本発明はこれに限らず、顔画像を検出することができればこの他種々の方法を用いた構成のものに広く適用することができる。

10

【 0 2 5 4 】

更にまた、抽出されたウィンドウ画像を、平均的な顔画像を表すテンプレートを用いてマッチングをとり、当該テンプレートとの相関度を検出する相関度検出手段として、図12に示す主制御部381の処理機能のうちのテンプレートマッチング部92を適用するようにした場合について述べたが、本発明はこれに限らず、要は、当該テンプレートを基準として相関度を検出することができれば、この他種々の方法を用いた構成のものに広く適用することができる。

【 0 2 5 5 】

更に、ロボット装置200の制御全体を司る制御手段として主制御部381を適用すると共に、当該主制御部381と接続され、その制御に応じて情報を読み書きするメモリ手段として内部メモリ381Aを適用し、更に図12に示す主制御部381の処理機能のうち、入力画像スケール変換部90、ウィンドウ切出部91、テンプレートマッチング部92による各処理を実行する演算処理手段として演算処理部71(DMAコントローラ70)を適用するようにした場合について述べたが、本発明はこれに限らず、主制御部381と演算処理部71とで顔検出処理の処理機能(90~95)を選択的に分担して内部メモリ381Aを介して切り換えて実行することができれば、その際の分担内容は、主制御部381にかかる演算量及び演算時間の負担を低減させるように自由に設定することができる。また、主制御部381のみで処理を行ってもよい。

20

【 0 2 5 6 】

【発明の効果】

以上詳細に説明したように本発明に係る顔検出装置は、入力画像から対象物の顔領域を抽出する顔検出装置において、撮像手段による撮像結果として得られるフレーム画像を入力画像とし、この入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成手段と、上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出手段と、上記顔候補として抽出された点に相当する入力画像領域から顔画像を識別する識別手段とを有するので、顔候補抽出手段により極めて効率よく且つ正確に顔候補を絞り込むことができるため、後段の顔識別手段における演算量を低減して、リアルタイムの顔検出を可能とする。

30

40

【 0 2 5 7 】

本発明に係るロボット装置は、供給された入力情報に基づいて動作を行う自律型のロボット装置において、撮像手段と、上記撮像手段による撮像結果として得られるフレーム画像を入力画像とし該入力画像から人物の顔領域を抽出する顔検出手段を有し、上記顔検出手段は、上記入力画像と平均的な顔画像を示す所定サイズのテンプレートとの相関をとった相関値の集合であるマッチング結果を生成するマッチング結果生成手段と、上記マッチング結果から相関値の局所最大値を求めこの局所最大値に基づき顔候補を抽出する顔候補抽出手段と、上記顔候補として抽出された点に相当する入力画像領域から顔画像を識別する識別手段とを具備するので、顔候補抽出手段により極めて効率よく且つ正確に顔候補を絞り込み、顔検出の演算量を低減して、リアルタイムの顔検出を可能とすることにより、口

50

ボット装置のエンタテインメント性を向上することができる。

【図面の簡単な説明】

【図 1】本発明の実施の形態における人間型ロボット装置を前方からみた外観を示す斜視図である。

【図 2】本発明の実施の形態における人間型ロボット装置を後方からみた外観を示す斜視図である。

【図 3】本発明の実施の形態における人間型ロボット装置の自由度構成モデルを示す模式図である。

【図 4】本発明の実施の形態における人間型ロボット装置の制御システム構成を示す模式図である。

【図 5】第 1 の適用例におけるロボット装置の内部構成の説明に供するブロック図である。

【図 6】第 1 の適用例におけるロボット装置の顔検出手段を示すブロック図である。

【図 7】重なり判定処理の説明に供する略線的な平面図である。

【図 8】第 2 の適用例におけるロボット装置の内部構成の説明に供するブロック図である。

【図 9】第 2 の適用例におけるロボット装置の顔検出手段を示すブロック図である。

【図 10】テンプレートマッチングによる相関性の検出の説明に供する模式図である。

【図 11】重なり判定処理手順をその工程順に示すフローチャートである。

【図 12】本発明の実施の形態におけるロボット装置の顔検出タスク機能を示すブロック図である。

【図 13】(a) 及び (b) は、夫々入力画像 (ウィンドウ画像) 及びテンプレートを示す模式図である。

【図 14】入力画像 (ウィンドウ画像) とテンプレートとから求めた相関値の集合であるマッチング結果画像を示す図である。

【図 15】本発明の実施の形態における顔検出手段のテンプレートマッチング部における処理工程を示すフローチャートである。

【図 16】本発明の実施の形態における顔検出タスク機能のテンプレートマッチング部がマッチング結果画像から顔候補を抽出した結果を示す図である。

【図 17】マッチング結果画像において、所定の閾値以上のものを顔候補として抽出した結果を示す図である。

【符号の説明】

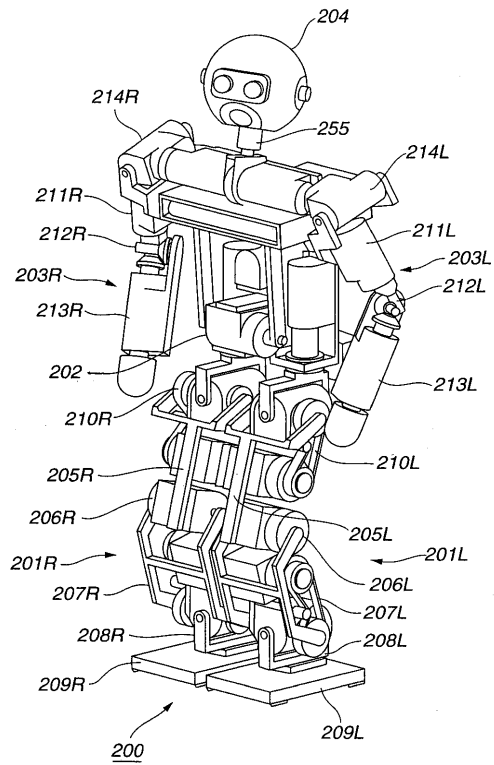
1 ロボット装置、 381 メイン制御部、 381A 内部メモリ、 50 CCD カメラ、 60、80、90 入力画像スケール変換部、 61、81、91 ウィンドウ切出部、 62、82、92 テンプレートマッチング部、 63、85、94 前処理部、 64、86、95 パターン識別部、 65、83 重なり判定部、 70 DMA コントローラ、 71 演算処理部、 84、93 スケール変換及び切出部、 RT 1 重なり判定処理手順、 S1A 画像信号

10

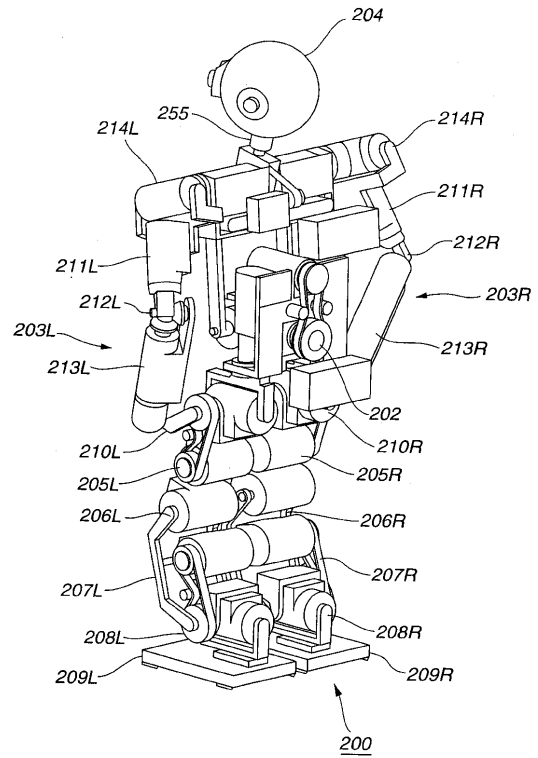
20

30

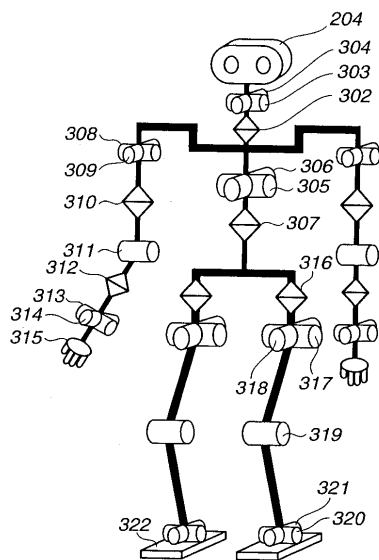
【図 1】



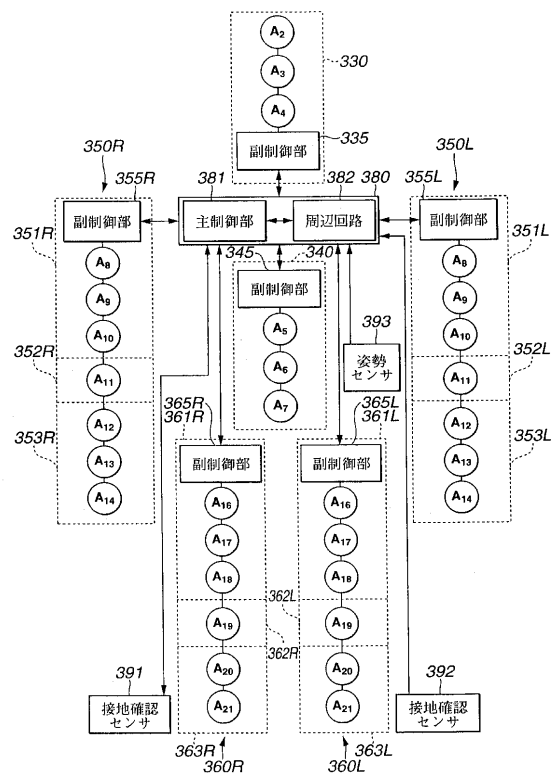
【図 2】



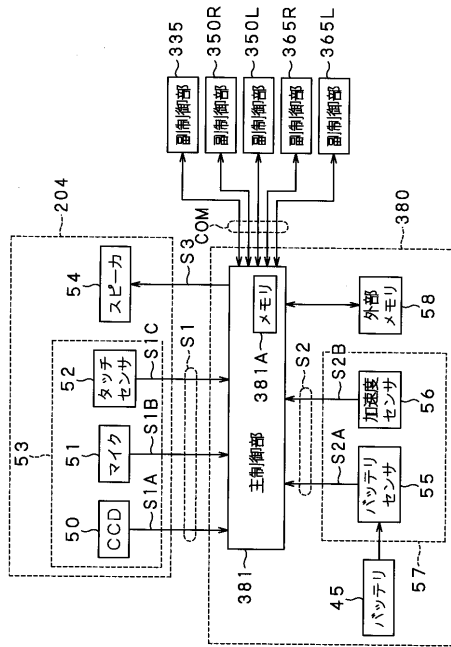
【図 3】



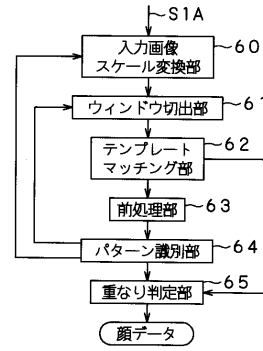
【図 4】



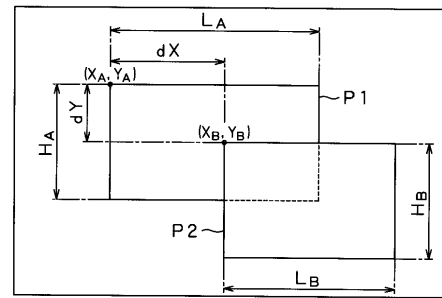
【図 5】



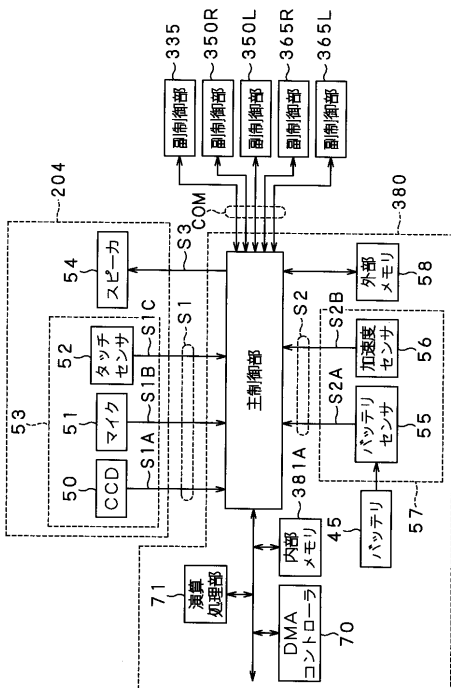
【図 6】



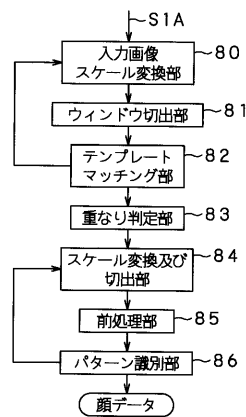
【図 7】



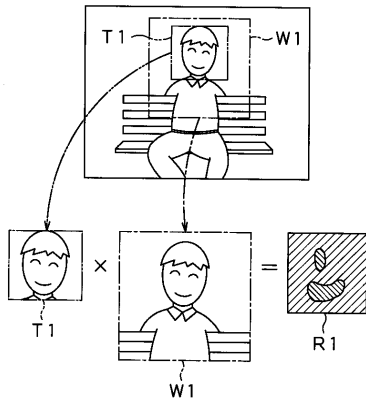
【図 8】



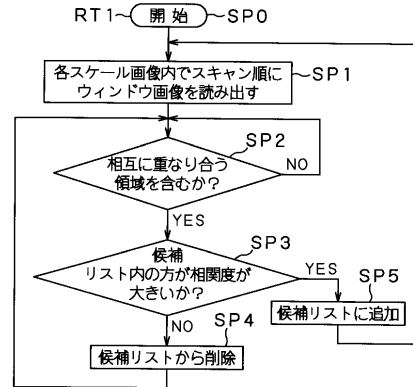
【図 9】



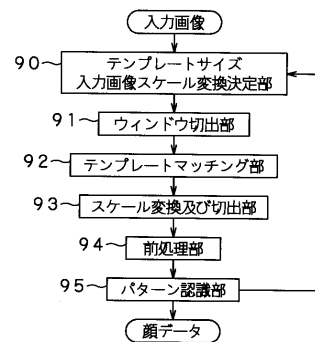
【図 10】



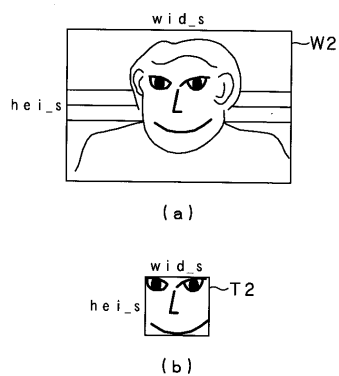
【図 11】



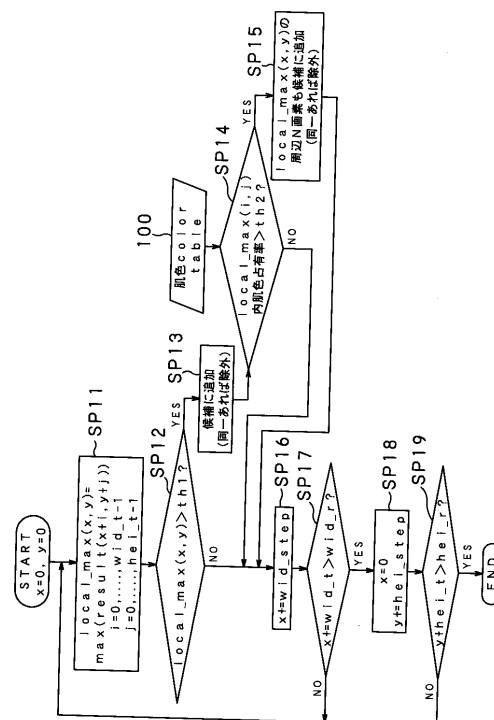
【図 12】



【図 13】



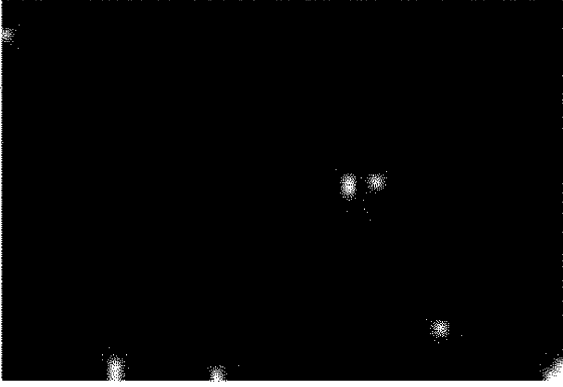
【図 15】



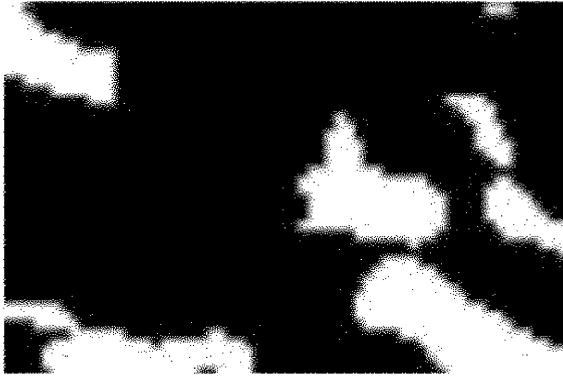
【図 14】



【図 16】



【図 17】



フロントページの続き

(56)参考文献 特開 2 0 0 2 - 0 5 1 3 1 6 (J P , A)
特開 2 0 0 1 - 0 6 7 4 5 9 (J P , A)
特開昭 5 7 - 1 3 4 7 6 3 (J P , A)
特開 2 0 0 0 - 1 2 3 1 8 5 (J P , A)
特開 2 0 0 0 - 2 7 6 5 9 6 (J P , A)
特開平 1 1 - 3 5 3 4 8 4 (J P , A)
特開平 1 0 - 2 2 2 6 7 8 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G06T 7/00 ~ 7/60

B25J 5/00

B25J 19/04

G06T 1/00