

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4093273号
(P4093273)

(45) 発行日 平成20年6月4日(2008.6.4)

(24) 登録日 平成20年3月14日(2008.3.14)

(51) Int. Cl. F I
G06T 7/60 (2006.01) G06T 7/60 150D
G06T 7/00 (2006.01) G06T 7/00 350B
 G06T 7/00 300E

請求項の数 16 (全 30 頁)

| | |
|--|---|
| <p>(21) 出願番号 特願2006-68047 (P2006-68047) (22) 出願日 平成18年3月13日 (2006.3.13) (65) 公開番号 特開2007-249280 (P2007-249280A) (43) 公開日 平成19年9月27日 (2007.9.27) 審査請求日 平成19年8月10日 (2007.8.10)</p> <p>早期審査対象出願</p> | <p>(73) 特許権者 000002945 オムロン株式会社 京都市下京区塩小路通堀川東入南不動堂町 801番地 (74) 代理人 100085006 弁理士 世良 和信 (74) 代理人 100106622 弁理士 和久田 純一 (74) 代理人 100125357 弁理士 中村 剛 (72) 発明者 木下 航一 京都府京都市下京区塩小路通堀川東入南不 動堂町801番地 オムロン株式会社 内</p> <p>審査官 新井 則和</p> <p style="text-align: right;">最終頁に続く</p> |
|--|---|

(54) 【発明の名称】 特徴点検出装置、特徴点検出方法および特徴点検出プログラム

(57) 【特許請求の範囲】

【請求項1】

入力された画像から、所定の対象物の画像における複数の特徴点の3次元位置を推定する特徴点検出装置であって、

前記複数の特徴点にそれぞれ対応する複数のノードの3次元位置を定めた3次元モデルを作成するモデル作成手段と、

前記3次元モデルによって定められた各ノードを前記画像上に射影する射影手段と、

前記射影手段によって射影された各ノードの射影点の付近の一つ又は複数のサンプリング点から特徴量を取得し、一つの射影点に基づいて取得された一つ又は複数の特徴量を一つのノード特徴量として取得する特徴量取得手段と、

学習用の画像において、複数のノードがそれぞれ対応する特徴点の正しい位置に3次元配置された正解モデルと、複数のノードのうち少なくともいずれかが対応する特徴点の正しい位置からずれた位置に3次元配置された誤差モデルとの差、及び、前記誤差モデルを学習用の画像に射影した射影点から取得されるノード特徴量、に基づいて算出される誤差推定行列を予め記憶している記憶手段と、

前記記憶手段に記憶される前記誤差推定行列と、前記特徴量取得手段によって取得された各ノードのノード特徴量とに基づいて、現在の3次元モデルとそれぞれに対応する特徴点の位置のずれを示す誤差推定量を取得する誤差推定量取得手段と、

前記誤差推定量取得手段によって得られた誤差推定量と、現在の3次元モデルとに基づいて、入力された画像における各特徴点の3次元位置を推定する推定手段とを備える特徴

点検出装置。

【請求項 2】

繰り返し処理の終了を判定する終了判定手段をさらに備え、

前記モデル作成手段は、前記誤差推定量が得られた場合は、この誤差推定量に基づいて現在の各ノードの位置を移動させて新たな 3 次元モデルを作成し、

前記モデル作成手段と前記誤差推定量取得手段とは、繰り返して処理を実行し、

前記推定手段は、前記終了判定手段によって終了が判定された時点での誤差推定量と、この時点での 3 次元モデルとに基づいて、入力された画像における各特徴点の 3 次元位置を推定する請求項 1 に記載の特徴点検出装置。

【請求項 3】

前記終了判定手段は、前記誤差推定量取得手段によって得られた誤差推定量が閾値以下になった場合に、繰り返し処理の終了と判定する請求項 2 に記載の特徴点検出装置。

【請求項 4】

前記モデル作成手段は、

繰り返し処理の初期段階においては、前記誤差推定量に基づいて、現在の各ノードの相対位置関係を保ったまま移動及び/又は回転させて姿勢を変化させた新たな 3 次元モデルを作成し、

3 次元モデルの姿勢に関する誤差が所定の閾値以下になった後は、前記誤差推定量に基づいて、現在の各ノードの相対位置も変化させて新たな 3 次元モデルを作成する請求項 2 又は 3 に記載の特徴点検出装置。

【請求項 5】

前記対象物は人の顔である請求項 1 ~ 4 のいずれかに記載の特徴点検出装置。

【請求項 6】

前記入力された画像から少なくとも人の顔又は人の顔の要素の位置を検出する顔検出手段をさらに備え、

前記配置手段は、前記検出された人の顔又は人の顔の要素の位置に基づいて、前記 3 次元モデルにおける各ノードの 3 次元位置を決定する請求項 5 に記載の特徴点検出装置。

【請求項 7】

前記顔検出手段は、人の顔の向きをさらに検出し、

前記モデル作成手段は、3 次元モデルにおける各ノードの 3 次元位置を定めた 3 次元ノードモデルを有しており、前記顔検出手段によって検出された顔の向きに基づいて、前記 3 次元ノードモデルを変形させて使用する請求項 6 に記載の特徴点検出装置。

【請求項 8】

前記顔検出手段は、人の顔の大きさをさらに検出し、

前記モデル作成手段は、3 次元モデルにおける各ノードの 3 次元位置を定めた 3 次元ノードモデルを有しており、前記顔検出手段によって検出された顔の大きさに基づいて、前記 3 次元ノードモデルを変形させて使用する請求項 6 に記載の特徴点検出装置。

【請求項 9】

前記顔検出手段は、人の顔の向きをさらに検出し、

前記モデル作成手段は、3 次元モデルにおける各ノードの 3 次元位置を定めた 3 次元ノードモデルを、顔の向きに応じて複数有しており、

前記記憶手段は、それぞれに対応する向きの顔が撮像された前記学習用画像を用いて予め得られた各 3 次元ノードモデルに対応する前記誤差推定行列を記憶し、

前記モデル作成手段は、前記顔検出手段によって検出された顔の向きに基づいて、使用する 3 次元ノードモデルを選択し、

前記誤差推定量取得手段は、前記モデル作成手段によって選択された 3 次元ノードモデルに対応する前記誤差推定行列を前記記憶手段から読み出し使用する請求項 6 に記載の特徴点検出装置。

【請求項 10】

前記顔検出手段は、人の顔の大きさをさらに検出し、

前記顔検出手段は、人の顔の大きさをさらに検出し、

10

20

30

40

50

前記モデル作成手段は、3次元モデルにおける各ノードの3次元位置を定めた3次元ノードモデルを、顔の大きさに応じて複数有しており、

前記記憶手段は、それぞれに対応する大きさの顔が撮像された前記学習用画像を用いて予め得られた各3次元ノードモデルに対応する前記誤差推定行列を記憶し、

前記モデル作成手段は、前記顔検出手段によって検出された顔の大きさに基づいて、使用する3次元ノードモデルを選択し、

前記誤差推定量取得手段は、前記モデル作成手段によって選択された3次元ノードモデルに対応する前記誤差推定行列を前記記憶手段から読み出し使用する請求項6に記載の特徴点検出装置。

【請求項11】

ノードの射影点付近の前記サンプリング点は、ノードの射影点の周囲に離散的に複数配置されている請求項1～10のいずれかに記載の特徴点検出装置。

【請求項12】

人の顔における複数の特徴点の3次元配置を表す3次元顔形状モデルを用いて、入力画像における顔の特徴点の3次元位置を推定する特徴点検出装置であって、

前記3次元顔形状モデルを作成するモデル作成手段と、

前記モデル作成手段によって作成された3次元顔形状モデルの各ノードを入力された画像平面上に射影する射影手段と、

前記射影手段によって射影された各ノードの射影点の位置の付近で離散的に複数の特徴量を取得し、一つの射影点に基づいて取得された複数の特徴量を一つのノード特徴量として取得する特徴量取得手段と、

各ノードが正しい位置に配置された正解3次元顔形状モデルと、少なくともいずれかのノードが正しい位置からずれて配置された誤差3次元顔形状モデルとにおける各ノードの位置の差を示すベクトルの集合、及び、前記誤差3次元顔形状モデルを学習用の画像に射影した射影点から取得されるノード特徴量の集合、を入力とした正準相関分析により予め取得された変換ベクトル行列を誤差推定行列として記憶する記憶手段と、

前記記憶手段に記憶される誤差推定行列と、前記特徴量取得手段によって取得されたノード特徴量の集合とに基づいて、3次元顔形状モデルの各ノードの現在位置と顔の特徴点の位置とのずれを示す誤差推定量を取得する誤差推定量取得手段と、

前記誤差推定量取得手段によって得られた前記誤差推定量に基づいて、現在の3次元顔形状モデルの各ノード位置を変更する変更手段と、

前記変更手段によって位置が変更された後の各ノード位置を特徴点の3次元位置として推定する推定手段と

を備える特徴点検出装置。

【請求項13】

入力された画像から、所定の対象物の画像における複数の特徴点の3次元位置を推定する特徴点検出方法であって、

情報処理装置が、

前記複数の特徴点にそれぞれ対応する複数のノードの3次元位置を定めた3次元モデルを作成するステップと、

前記3次元モデルによって定められた各ノードを前記画像上に射影するステップと、

前記射影手段によって射影された各ノードの射影点の付近の一つ又は複数のサンプリング点から特徴量を取得し、一つの射影点に基づいて取得された一つ又は複数の特徴量を一つのノード特徴量として取得するステップと、

学習用の画像に基づいて予め取得されている、複数のノードがそれぞれ対応する特徴点の正しい位置に3次元配置された正解モデルと、複数のノードのうち少なくともいずれかが対応する特徴点の正しい位置からずれた位置に3次元配置された誤差モデルとの差、及び、前記誤差モデルを学習用の画像に射影した射影点から取得されるノード特徴量、に基づいて算出される誤差推定行列と、前記特徴量を取得するステップにおいて取得された各ノードのノード特徴量とに基づいて、現在の3次元モデルとそれぞれに対応する特徴点

10

20

30

40

50

の位置のずれを示す誤差推定量を取得するステップと、

前記誤差推定量と、現在の3次元モデルとに基づいて、入力された画像における各特徴点の3次元位置を推定するステップと
 を実行する特徴点検出方法。

【請求項14】

情報処理装置に、入力された画像から、所定の対象物の画像における複数の特徴点の3次元位置を推定させるための特徴点検出プログラムであって、

情報処理装置に対して、

前記複数の特徴点にそれぞれ対応する複数のノードの3次元位置を定めた3次元モデルを作成するステップと、

前記3次元モデルによって定められた各ノードを前記画像上に射影するステップと、

前記射影手段によって射影された各ノードの射影点の付近の一つ又は複数のサンプリング点から特徴量を取得し、一つの射影点に基づいて取得された一つ又は複数の特徴量を一つのノード特徴量として取得するステップと、

学習用の画像に基づいて予め取得されている、複数のノードがそれぞれ対応する特徴点の正しい位置に3次元配置された正解モデルと、複数のノードのうち少なくともいづれかが対応する特徴点の正しい位置からずれた位置に3次元配置された誤差モデルとの差、及び、前記誤差モデルを学習用の画像に射影した射影点から取得されるノード特徴量、に基づいて算出される誤差推定行列と、前記特徴量を取得するステップにおいて取得された各ノードのノード特徴量とに基づいて、現在の3次元モデルとそれぞれに対応する特徴点
 の位置のずれを示す誤差推定量を取得するステップと、

前記誤差推定量と、現在の3次元モデルとに基づいて、入力された画像における各特徴点の3次元位置を推定するステップと
 を情報処理装置に実行させる特徴点検出プログラム。

【請求項15】

人の顔における複数の特徴点の3次元配置を表す3次元顔形状モデルを用いて、入力画像における顔の特徴点の3次元位置を推定する特徴点検出方法であって、

情報処理装置が、

前記3次元顔形状モデルを作成するステップと、

作成された3次元顔形状モデルの各ノードを入力された画像平面上に射影するステップと、

射影された各ノードの射影点の付近で離散的に複数の特徴量を取得し、一つの射影点に基づいて取得された複数の特徴量を一つのノード特徴量として取得するステップと、

学習用の画像に基づいて予め取得されている、各ノードが正しい位置に配置された正解3次元顔形状モデルと、少なくともいづれかのノードが正しい位置からずれて配置された誤差3次元顔形状モデルとにおける各ノードの位置の差を示すベクトルの集合、及び、前記誤差3次元顔形状モデルを学習用の画像に射影した射影点から取得されるノード特徴量の集合、を入力とした正準相関分析により予め取得される変換ベクトル行列と、前記特徴量を取得するステップにおいて取得されたノード特徴量の集合とに基づいて、3次元顔形状モデルの各ノードの現在位置と顔の特徴点の位置とのずれを示す誤差推定量を取得する
 ステップと、

取得された前記誤差推定量に基づいて、現在の3次元顔形状モデルの各ノード位置を変更するステップと、

位置が変更された後の各ノード位置を特徴点の3次元位置として推定するステップと
 を実行する特徴点検出方法。

【請求項16】

情報処理装置に、人の顔における複数の特徴点の3次元配置を表す3次元顔形状モデルを用いて、入力画像における顔の特徴点の3次元位置を推定させる特徴点検出プログラムであって、

情報処理装置に対して、

10

20

30

40

50

前記 3 次元顔形状モデルを作成するステップと、
作成された 3 次元顔形状モデルの各ノードを入力された画像平面上に射影するステップと、

射影された各ノードの射影点の付近で離散的に複数の特徴量を取得し、一つの射影点に基づいて取得された複数の特徴量を一つのノード特徴量として取得するステップと、

学習用の画像に基づいて予め取得されている、各ノードが正しい位置に配置された正解 3 次元顔形状モデルと、少なくともいずれかのノードが正しい位置からずれて配置された誤差 3 次元顔形状モデルとにおける各ノードの位置の差を示すベクトルの集合、及び、前記誤差 3 次元顔形状モデルを学習用の画像に射影した射影点から取得されるノード特徴量の集合、を入力とした正準相関分析により予め取得される変換ベクトル行列と、前記特徴量を取得するステップにおいて取得されたノード特徴量の集合とに基づいて、3 次元顔形状モデルの各ノードの現在位置と顔の特徴点の位置とのずれを示す誤差推定量を取得するステップと、

取得された前記誤差推定量に基づいて、現在の 3 次元顔形状モデルの各ノード位置を変更するステップと、

位置が変更された後の各ノード位置を特徴点の 3 次元位置として推定するステップとを情報処理装置に実行させる特徴点検出プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像の中から対象物の特徴点を検出するための装置や方法やプログラム等に適用されて有効な技術に関する。

【背景技術】

【0002】

画像の中から対象物の特徴点を検出する技術として、顔の特徴点を検出する技術がある。以下、顔を対象物とした場合の背景技術について述べる。

【0003】

人物が撮像された画像から顔の特徴点位置を得ることにより、例えば、その人物の目の開閉情報、顔姿勢、視線方向、顔の表情などを推定することができる。また、これらの推定結果から、その人物の状態をより詳細に知ることが可能となる。そして、人物の状態を詳細に把握することにより、マン・マシン・インタフェースの性能向上や新たなサービスの提供等を図ることができる。このように、顔の特徴点位置を正確に知ることが、マン・マシン・インタフェースの開発の上で重要な課題となっている。

【0004】

このようなマン・マシン・インタフェースの例として、運転者モニタリングシステムがある。このシステムでは、運転者の目の開閉、顔姿勢、視線方向などが逐次観察される。そして、この観察結果に基づいて運転者の疲労度合い等が判断され、状況に応じた的確なアドバイスが可能となる。

【0005】

また、他のマン・マシン・インタフェースの例として、動画ビデオカメラや静止画カメラなどへの応用がある。これらの装置において、被撮影者の顔姿勢を知ることにより、撮影条件を変化させるなど、多彩な処理が可能となる。この他、目や口などの特徴点の画像を詳細に解析することで、個人の識別を高い精度で行うことも可能となる。

【0006】

画像から特徴点位置を得る方法（以下、「特徴点検出方法」と呼ぶ）としては、テンプレートマッチングによるもの、及びその応用によるものが一般的である。これらの方法では、各特徴点の一般的な情報（以下、「特徴量」と呼ぶ）が予め保存される。そして、画像内の一部の領域から得られる特徴量と、保存されている特徴量とが比較されることにより、取得すべき特徴点であるか否かが判断される。このとき使用される特徴量の具体例としては、画像の輝度値ベクトルがある。また、両特徴量の比較手段としては、正規化相関

10

20

30

40

50

やユークリッド距離を使用するものが一般的である。

【0007】

このような特徴点検出技術の例として、分離度フィルタによって予め探索点数を減らした後、パターンマッチングにより特徴点候補を検出する技術が提案されている（特許文献1参照）。この技術では、特徴点候補が検出された後、幾何的な拘束条件を適用し、最も人間の顔らしいと判断される候補点の組み合わせが出力される。

【0008】

しかしこの技術では、ほぼ正面を向いた、撮影条件のよい人物画像の特徴点しか検出することができない。従って、特徴点の一部が遮蔽物によって隠れてしまった画像や、予め保存されている特徴点の取得時と撮影条件（例えば照明条件）が大きく異なる画像からは、特徴点を正確に検出することは困難であった。

10

【0009】

また、顔の特徴点位置を得ることにより顔姿勢を推定する技術もある（特許文献2, 3参照）。このような技術では、顔の特徴点位置が取得された後、その全体配置や特徴量を用いて推定する手法が一般的である。例えば、目や口や眉や鼻を示す特徴点の座標と顔姿勢との関係が予めルックアップテーブルとして保存されている。そして、画像から取得された特徴点の座標に対応する顔姿勢が、このルックアップテーブルから判断され、推定結果として出力される。この他にも、顔全体又は顔の特徴量のテンプレートを複数の顔の向きに対応して用意し、これらのテンプレートとマッチングをとることによって顔姿勢を求める手法もある。しかしながら、これらの手法においても、顔姿勢の推定が正確に実行できるか否かは、顔の特徴点位置の精度に依存する。従って、顔の特徴点位置の取得を正確に実行できなければ、顔姿勢の推定も正確に行うことができなかった。

20

【0010】

このような問題を解決する技術として、ASM (Active Shape Model) と呼ばれる技術がある（非特許文献1参照）。ASMでは、予め、学習用の多数の顔画像について特徴点の位置が取得され、顔形状モデルが作成され保持される。顔形状モデルは、各特徴点に対応するノードによって構成される。なお、顔形状モデルの詳細については、発明を実施するための最良の形態の欄で説明する。

【0011】

次に、ASMにより画像から特徴点の位置を検出する処理について説明する。まず、処理対象となる画像の適当な初期位置に、この顔形状モデルが配置される。次に、顔形状モデルの各ノードについて、ノード周辺の複数の特徴量の取得が行われる。取得された複数の特徴量と予めそのノードに対応づけて保持されている特徴量とが比較される。これら複数の特徴量取得位置のうち、各ノードに対応する特徴量に最も近い特徴量が取得された位置（即ち、各ノードに対応する特徴点である可能性が最も高いと判断される位置）に各ノードが移動される。この時点で、顔形状モデルの各ノードの位置は、初期の位置からずれている。そこで、変形されたノード集合が、顔形状モデルに射影されることにより整形される。各ノード周りにおける特徴量の取得以降の処理が、所定の回数、又は一定の条件（収束条件）が満たされるまで繰り返し実行される。そして、各ノードの最終的な位置が、各特徴点の位置と判断される。

30

40

【0012】

ASMでは、上記のように、各ノードの位置が移動した後に、顔形状モデルへの射影が行われる。この処理により、各ノードの位置関係が顔らしい形状を保ったまま、特徴点の正確な位置を検出することができる。即ち、一般の人の顔の形状からは到底ありえない位置に偶然的に特徴点と似た特徴量を有する部分が存在した場合にも、このような点を特徴点として誤検出することが防止される。

【0013】

しかしながら、従来のASMでは、各ノードにおいて周囲の探索処理（特徴量の取得と、その比較）を行う必要があった。このため、多くの計算時間を必要とするという欠点があった。また、従来のASMには、ロバスト性が低いという問題があった。即ち、処理対

50

象となる画像中の顔の向きが、初期位置に配置された顔形状モデルにおいて想定されている顔の向きと大きく異なる場合、特徴点の検出に失敗するケースが多いという問題があった。

【0014】

A S Mのこのような問題を解決するための技術として、A A M (Active Appearance Model)がある(非特許文献2参照)。A A Mでは、以下のようにして特徴点の位置が求められる。まず、A S Mの場合と同様に、学習用の多数の顔画像について特徴点の位置が取得され、顔形状モデルが作成される。次に、各ノード位置の平均値が取得され、平均位置のノードの集合により構成される平均形状モデルが作成される。次に、学習用の顔画像において各特徴点により構成される複数のパッチが作成され、各パッチが平均形状モデルに射影され、シェイプフリーイメージが作成される(この処理を「形状修正処理」と呼ぶ)。そして、シェイプフリーイメージの集合について主成分分析が実施されることにより、顔輝度値モデル(シェイプフリー顔輝度値モデル)が作成される。次に、顔形状モデルを正解位置から一定ずつ各方向に微小にずらした際のシェイプフリー顔輝度値ベクトルが求められる。これらの集合について線形回帰計算が行われる。これにより、微小にずれた顔輝度値ベクトルから顔形状モデルの正解点への移動・変形方向、量を推定することが可能になる。A A Mでは、以上の処理が学習処理として、予め実行される。

10

【0015】

次に、A A Mにより画像から特徴点の位置を検出する処理について説明する。まず、処理対象となる画像の適当な初期位置に、顔形状モデルが配置される。次に、配置された顔形状モデルのノード位置に基づいてパッチが作成され、各パッチにおける輝度分布がサンプリングされる。そして、サンプリングされた輝度分布が射影されることにより、シェイプフリー顔輝度値モデルが作成される。

20

【0016】

次に、このシェイプフリー顔輝度値モデルから、予め求められている回帰式により、顔形状モデルの移動・変形量が推定される。顔形状モデルは、この推定結果に従って移動・変形される。以上の処理が、所定の回数、又は一定の条件(収束条件)が満たされるまで繰り返し実行される。そして、各ノードの最終的な位置が、各特徴点の位置と判断される。

30

【0017】

A A Mによれば、上記のように、各ノード周りの探索処理を行うことなく各特徴点の位置を検出することができる。このため、A S Mの場合と異なり、各ノード周りの探索処理に時間を要することが無く、計算時間を削減できる。また、形状モデルを変形させることにより探索を行うため、A S Mの場合と同様に、各ノードの位置関係が顔らしい形状を保ったまま、特徴点の正確な位置を検出することが可能となる。

【特許文献1】特開平9-251534号公報

【特許文献2】特開2000-97676号公報

【特許文献3】特開2003-141551号公報

【非特許文献1】A. Lanitis, C. J. Taylor, T. F. Cootes, "Automatic Interpretation and Coding of Face Images Using Flexible Models. IEEE PAMI Vol. 19 No. 7 pp. 743-756, July 1997.

40

【非特許文献2】T. F. Cootes, G. J. Edwards and C. J. Taylor. "Active Appearance Models", IEEE PAMI, Vol. 23, No. 6, pp. 681-685, 2001

【発明の開示】

【発明が解決しようとする課題】

【0018】

50

しかしながら、A A Mには以下のような問題点がある。

【 0 0 1 9 】

A A Mによって各特徴点の位置検出の精度を維持するためには、各パッチの輝度分布を取得する際に均質で密度の高いサンプリングを行う必要がある。このため、そのサンプリングや、輝度分布を平均形状モデルへ射影する際の計算量が多大となり、多くの計算時間を要してしまう。さらに、A A Mでは、正解点周りの微小なずれによる線形回帰に基づいて、顔形状モデルの移動量や変形量が推定される。このため、大きなずれに対しては正しい推定を行うことができず、正しい結果を得ることができない。

【 0 0 2 0 】

このように、従来のA S MやA A Mを用いた処理では、探索処理や、均質且つ高密度なサンプリングによる輝度分布の形状修正処理などにおいて多くの計算量を必要とするため、高速に特徴点の検出を行うことができなかった。しかしながら、上記したように、特徴点の検出結果はマン・マシン・インタフェース等において入力データとして利用されるため、応答性が求められる場合が多い。このため、高速に特徴点の検出を行う事への要求が高まっていることも事実である。

【 0 0 2 1 】

そこで本発明はこれらの問題を解決し、画像から特徴点の位置を高速に検出することを可能とする装置やプログラム等を提供することを目的とする。

【課題を解決するための手段】

【 0 0 2 2 】

従来は、A A Mのように、形状修正処理を行うことによって、誤差推定（形状モデルにおける各ノードの正解点への移動方向や移動量を推定すること）を行うことが一般であった。従って、これまで、誤差推定を行うにあたり形状修正処理を省くことは検討されてこなかった。しかし、従来のA A Mにおいて、計算量の増大を招いていた要因を検討すると、必要とされている形状修正処理がその要因として挙げられる。そこで、本発明者は、このような形状修正処理を行うことなく特徴点の検出を可能とする装置について検討した。

【 0 0 2 3 】

また、従来は、誤差推定の精度を高く維持するため、各ノードの周囲で高密度なサンプリングが実施されることが一般であった。例えば、従来のA A Mでは、各ノードによって構成されるパッチ内において均質且つ高密度なサンプリング（具体的には、通常1万点以上のサンプリング）が必要とされていた。従って、これまで、誤差推定を行うにあたり低密度なサンプリングを実施することは検討されてこなかった。しかし、従来のA A Mにおいて、計算量の増大を招いていた要因の一つとして、このような均質で高密度なサンプリングも挙げることができる。そこで、本発明者は、このような均質且つ高密度なサンプリングを行うことなく特徴点の検出を可能とする装置について検討した。

【 0 0 2 4 】

また、従来は特徴点の位置を画像平面（2次元）上で捉えていた。したがって、対象物（例えば、人の顔）の姿勢を得るためには、画像平面上での特徴点の位置が決定された後に、この特徴点の位置に基づいて対象物の姿勢を推定する必要があった。形状モデルを3次元で表すことは検討されてこなかった。しかし、形状モデルとして3次元モデルを用いた場合は、検出された特徴点の位置から直接対象物の姿勢を求めることができる。そこで、本発明者は、3次元形状モデルを用いて特徴点の検出を行う装置について検討した。

【 0 0 2 5 】

これらの検討の結果、本発明者は以下のような考えに至った。まず、従来のA A Mにおける誤差推定では、各ノードにより構成されるパッチにおいてサンプリングが実施されていた。このため、パッチのように面としてサンプリングの結果を捉えると、誤差推定の結果に基づくノードの移動に伴う形状変形を形状修正処理によって修正しなければ処理の精度を維持することが難しかった。しかし、各ノードの周囲の複数の点について非均質且つ低密度のサンプリングを行う場合、形状が変形しても点は変形しないため、またサンプリング点が低密度であるため、ノードの移動に伴う形状変形の影響を受けにくい。このため

10

20

30

40

50

、形状修正処理を特に実施することなく処理の精度を維持することが可能となる。このような検討結果に基づき、本発明者は以下のような特徴点検出装置を提案する。

【0026】

本発明の第一の態様は、入力された画像から、所定の対象物の画像における複数の特徴点の3次元位置を推定する特徴点検出装置である。ここでいう対象物は、形状モデルを設定できるものであればどのような対象物であっても良い。例えば、対象物は、人の顔、人の全体（全身）、レントゲン画像やCT（Computed Tomography）画像における臓器であっても良い。言い換えれば、大きさの個人差がある対象物や基本的な形が変わらずに変形する対象物について本技術は適用可能である。また、他の例として、自動車などの工業製品のように変形しない剛体の対象物であっても、形状モデルを設定す

10

【0027】

本発明の第一の態様に係る特徴点検出装置は、モデル作成手段、射影手段、特徴量取得手段、記憶手段、誤差推定量取得手段、及び推定手段を備える。モデル作成手段は、複数の特徴点にそれぞれ対応する複数のノードの3次元位置を定めた3次元モデルを作成する。各ノードは、検出の対象となる特徴点各々に対応づけられている。また、作成される3次元モデルは、3次元モデルの各ノードを入力画像上に射影した射影点が、各ノードの対応する特徴点に近いと想定される位置であることが望ましい。従って、このような基準に基づいて作成される3次元モデルの初期位置及び姿勢が予め定められると良い。例えば、対象物は画面中央に正面向きに撮像されることが多いという経験に基づいて、画面中央を初期位置とし、正面向きとすることができる。

20

【0028】

射影手段は、3次元モデルによって定められた各ノードを入力画像に射影し、その射影点の画像平面における位置を算出する。

【0029】

特徴量取得手段は、射影手段によって射影された各ノードの射影点の付近の一つ又は複数のサンプリング点から取得された一つ又は複数の特徴量を、3次元モデルの各ノードのノード特徴量として取得する。特徴量の具体例として、画素の輝度値、Sobelフィルタ特徴量、HarrWavelet特徴量、GaborWavelet特徴量やこれらを複合した値がある。また、サンプリング点の配置は、非均質で低密度であっても良い。

30

【0030】

記憶手段は、予め実行されている学習結果を記憶する。この学習結果は、複数の学習用の画像について学習処理を行うことにより取得される。具体的には、以下に示す二つの値に基づいて、誤差推定行列が予め学習結果として取得される。一方の値は、複数のノードがそれぞれに対応する特徴点の正しい3次元位置に配置された3次元モデルである正解モデルと、複数のノードのうち少なくともいずれかが対応する特徴点の正しい位置からずれた位置に配置された3次元モデルである誤差モデルとの差である。他方の値は、誤差モデルを学習用の画像に射影した射影点に基づいて取得された上述のノード特徴量である。このような二つの値に基づく誤差推定行列は、例えば正準相関分析や、線形回帰、線形重回帰、非線形重回帰などの分析手法によって取得される。

40

【0031】

誤差推定量取得手段は、記憶手段に記憶される誤差推定行列と、特徴量取得手段によって取得された各ノードのノード特徴量とに基づいて、現在の3次元モデルの各ノードの位置とそれぞれに対応する特徴点の位置とのずれを示す誤差推定量を取得する。

【0032】

推定手段は、誤差推定量取得手段によって得られた誤差推定量と、現在の3次元モデルの各ノードの位置とに基づいて、入力された画像における各特徴点の3次元位置を推定す

50

る。即ち、特徴点検出装置は、推定手段によって推定された各3次元位置において各特徴点を検出したこととなる。

【0033】

本発明の第一の態様によれば、第一に、形状修正処理が必要とされない。言い換えれば、本発明の第一の態様によれば、形状修正処理が実施されない。このため、これまで形状修正処理に費やされていた多くの処理を省き、処理の高速化を図ることが可能となる。

【0034】

本発明の第一の態様によれば、第二に、各ノードの周囲において探索処理をする必要はなく、各ノードによって構成されるパッチにおいて面としての高密度なサンプリングを行う必要もない。このため、従来技術によるこのようなサンプリングによって必要とされていた多くの計算量を削減することが可能となる。

【0035】

特徴点の検出処理が高速化されることにより、上記したマン・マシン・インタフェースを高速に動作させることが可能となる。例えば、運転手モニタリングシステムはその性質から応答性が要求されるが、従来実現されていなかったこの要求を解決することが可能となる。

【0036】

本発明の第一の態様は、繰り返し処理の終了を判定する終了判定手段をさらに備えても良い。終了判定手段は、例えば繰り返し処理が実行された回数をカウントし、所定の回数に達した場合に終了を判定するように構成されても良い。終了判定手段がさらに備えられる場合、モデル作成手段は、誤差推定量が得られた場合は、この誤差推定量に基づいて現在の各ノードの位置を移動させて新たな3次元モデルを作成する。そして、モデル作成手段と誤差推定量取得手段とは、繰り返して処理を実行し、推定手段は、終了判定手段によって終了が判定された時点での誤差推定量と、この時点での3次元モデルとに基づいて、入力された画像における各特徴点の3次元位置を推定する。

【0037】

このように構成された本発明の第一の態様によれば、誤差推定量に基づいて各ノードの3次元位置が移動された新たな3次元モデルに基づいて、再び誤差推定量が求められ、この処理が繰り返される。そして、終了判定手段によって、この繰り返し処理の終了が判定された時点での誤差推定量と、この時点での3次元モデルとに基づいて、各特徴点の3次元位置が推定される。従って、3次元モデルに基づいた一度の誤差推定量の取得により各特徴点の3次元位置が推定される場合に比べて、その精度が向上する。

【0038】

本発明の第一の態様における終了判定手段は、誤差推定量取得手段によって得られた誤差推定量が閾値以下になった場合に、繰り返し処理の終了と判定するように構成されても良い。このように構成されることにより、各特徴点の位置をより正確に検出することが可能となる。

【0039】

また、本発明の第一の態様におけるモデル作成手段は、上記の繰り返し処理を行う場合には、繰り返し回数によって、誤差推定量に基づいて新たな3次元モデルを作成する処理を変更しても良い。例えば、モデル作成手段は、繰り返し処理の初期段階においては、誤差推定量に基づいて、現在の各ノードの相対位置関係を保ったまま各ノードを移動及び/又は回転させて、3次元モデルの姿勢を変化させて新たな3次元モデルを作成する。そして、3次元モデルの姿勢に関する誤差が所定の閾値以下になった後は、誤差推定量に基づいて、現在の各ノードの相対位置も変化させて新たな3次元モデルを作成するようにしても良い。このような構成をとることによって、入力画像中に現れる対象物の向きの変化に対してロバストな特徴点検出を行うことができる。

【0040】

本発明の第一の態様は、所定の対象物として人の顔を対象とすることが好適である。そして、本発明の第一の態様は、入力された画像から少なくとも人の顔又は人の顔の要素（

10

20

30

40

50

顔に含まれる器官や部分を指し、例えば目、鼻、口、額などである)の位置を検出する顔検出手段をさらに備えても良い。この場合、モデル作成手段は、検出された人の顔又は人の顔の要素の位置に基づいて、3次元モデルにおける各ノードの3次元位置を決定する。このように構成されることにより、各ノードの初期位置が適切な位置、即ちそれぞれが対応する特徴点に近い位置となる。このため、各特徴点の位置をより正確に検出することが可能となる。

【0041】

本発明の第一の態様における顔検出手段は、人の顔の向きをさらに検出するように構成されても良い。この場合、モデル作成手段は、各ノードの3次元位置を定めた3次元ノードモデルを有しており、顔検出手段によって検出された顔の向きに基づいて、3次元ノードモデルを変形させて使用するよう構成されても良い。例えば、モデル作成手段は、顔検出手段によって顔の向きが左向きとして検出された場合には、3次元ノードモデルの各ノードを、左向きの顔にマッチするように回転させればよい。モデル作成手段は、他の方向の顔が検出された場合にも、それぞれの方向について上記のような回転処理を行うように構成されても良い。このように構成されることにより、予め用意された3次元ノードモデルを変形せずに使用する場合に比べて、各特徴点の位置をより正確に検出することが可能となる。

【0042】

本発明の第一の態様における顔検出手段は、人の顔の大きさをさらに検出するように構成されても良い。この場合、モデル作成手段は、3次元モデルにおける各ノードの3次元位置を定めた3次元ノードモデルを有しており、顔検出手段によって検出された顔の大きさに基づいて、3次元ノードモデルを変形させて使用するよう構成されても良い。例えば、モデル作成手段は、顔検出手段によって検出された顔の大きさに基づいて、3次元ノードモデルを拡大又は縮小して使用するよう構成されても良い。このように構成されることにより、予め用意された3次元ノードモデルを変形せずに使用する場合に比べて、各特徴点の位置をより正確に検出することが可能となる。

【0043】

本発明の第一の態様における顔検出手段は、人の顔の向きをさらに検出するように構成されても良い。この場合、モデル作成手段は、3次元モデルにおける各ノードの3次元位置を定めた3次元ノードモデルを、顔の向きに応じて複数有するよう構成されても良い。例えば、モデル作成手段は、右向きの顔、正面向きの顔、左向きの顔、上向きの顔、下向きの顔それぞれに対応する3次元ノードモデルを有するよう構成されても良い。また、記憶手段は、それぞれに対応する向きの顔が撮像された学習用画像を用いて予め得られた各3次元ノードモデルに対応する誤差推定行列を記憶するよう構成されても良い。また、モデル作成手段は、顔検出手段によって検出された顔の向きに基づいて、使用する3次元ノードモデルを選択するよう構成されても良い。そして、誤差推定量取得手段は、モデル作成手段によって選択された3次元ノードモデルに対応する誤差推定行列を記憶手段から読み出し使用するよう構成されても良い。

【0044】

このように構成された本発明の第一の態様では、顔検出手段によって検出された顔の向きに特化した3次元ノードモデル及びこの3次元ノードモデルに対応する誤差推定行列を用いて特徴点検出が行われる。従って、常に一つの3次元ノードモデルが使用される場合に比べて、各特徴点の位置をより正確に検出することが可能となる。

【0045】

本発明の第一の態様における顔検出手段は、人の顔の大きさをさらに検出するように構成されても良い。この場合、モデル作成手段は、3次元モデルにおける各ノードの3次元位置を定めた3次元ノードモデルを、顔の大きさに応じて複数有するよう構成されても良い。また、記憶手段は、それぞれに対応する大きさの顔が撮像された学習用画像を用い

10

20

30

40

50

て予め得られた各3次元ノードモデルに対応する誤差推定行列を記憶するように構成されても良い。また、モデル作成手段は、顔検出手段によって検出された顔の大きさに基づいて、使用する3次元ノードモデルを選択するように構成されても良い。また、誤差推定量取得手段は、モデル作成手段によって選択された3次元ノードモデルに対応する誤差推定行列を記憶手段から読み出し使用するように構成されても良い。

【0046】

このように構成された本発明の第一の態様では、顔検出手段によって検出された顔の大きさに特化した3次元ノードモデル及びこの3次元ノードモデルに対応する誤差推定行列を用いて特徴点検出が行われる。従って、常に一つの3次元ノードモデルが使用される場合に比べて、各特徴点の位置をより正確に検出することが可能となる。

10

【0047】

本発明の第一の態様における特徴量取得手段がノード特徴量を取得するサンプリング点は、ノードの射影点の周囲に離散的に配置されても良い。本発明の第一の態様における特徴量取得手段がノード特徴量を取得するサンプリング点は、ノードの射影点に近いほど密であり、ノードの射影点から遠いほど疎であるように構成されても良い。また、本発明の第一の態様における特徴量取得手段がノード特徴量を取得するサンプリング点は、レティナ構造を有するように構成されても良い。また、本発明の第一の態様における特徴量取得手段は、複数の異なるレティナ構造を記憶し、各ノードの射影点に応じたレティナ構造に従って複数のサンプリング点を決定するように構成されても良い。

20

【0048】

本発明の第二の態様は、人の顔における複数の特徴点の3次元配置を表す3次元顔形状モデルを用いて、入力画像における顔の特徴点の3次元配置を推定する特徴点検出装置である。本発明の第二の態様に係る特徴点検出装置は、モデル作成手段、射影手段、特徴量取得手段、記憶手段、誤差推定量取得手段、変更手段、及び推定手段を備える。

【0049】

モデル作成手段は、人の顔における複数の特徴点の3次元配置を表す3次元顔形状モデルを作成する。

【0050】

射影手段は、モデル作成手段によって作成された3次元顔形状モデルの各ノードを入力された画像平面上に射影し、射影点の画像平面における位置を決定する。

30

【0051】

特徴量取得手段は、射影手段によって射影された3次元顔形状モデルの各ノードの射影点の位置の付近で一つ又は複数の特徴量を取得し、一つの射影点に基づいて取得された複数の特徴量を一つのノード特徴量として取得する。

【0052】

記憶手段は、各ノードが正しい位置に配置された3次元顔形状モデルである正解3次元顔形状モデルと、少なくともいずれかのノードが正しい位置からずれて配置された3次元顔形状モデルである誤差3次元顔形状モデルとにおける各ノードの位置の差を示すベクトルの集合、及び、誤差3次元顔形状モデルを学習用の画像に射影して得られるノード特徴量の集合、を入力とした正準相関分析により予め取得された変換ベクトル行列を誤差推定行列として記憶する。

40

【0053】

誤差推定量取得手段は、記憶手段に記憶される誤差推定行列と、特徴量取得手段によって取得されたノード特徴量の集合とに基づいて、3次元顔形状モデルの各ノードの現在位置と顔の特徴点の位置とのずれを示す誤差推定量を取得する。

【0054】

変更手段は、誤差推定量取得手段によって得られた誤差推定量に基づいて、現在の3次元顔形状モデルの各ノード位置を変更する。

50

【 0 0 5 5 】

そして、推定手段は、変更手段によって位置が変更された後の各ノード位置を特徴点の位置として推定する。

【 0 0 5 6 】

このように構成された本発明の第二の態様によっても、本発明の第一の態様と同様の作用や効果を奏することができる。

【 0 0 5 7 】

第一～第二の態様は、プログラムが情報処理装置によって実行されることによって実現されても良い。即ち、本発明は、上記した第一～第二の態様における各手段が実行する処理を、情報処理装置に対して実行させるためのプログラム、或いは当該プログラムを記録した記録媒体として特定することができる。また、本発明は、上記した各手段が実行する処理を情報処理装置が実行する方法をもって特定されても良い。また、ここに示す情報処理装置とは、パーソナルコンピュータ、デジタルカメラ、デジタルビデオカメラ、防犯カメラ、デジタルカメラを搭載した携帯電話機、デジタルビデオレコーダ、画像処理を行うサーバ（ネットワークに接続されたサーバであっても良い）、デジタルフォトプリンタ、ミニラボ機などである。

10

【 発明の効果 】

【 0 0 5 8 】

本発明によれば、画像から対象物の特徴点の位置を検出する際に、形状修正処理や高密度なサンプリングを必要としない。このため、特徴点の検出処理をより高速に実現することが可能となる。

20

【 発明を実施するための最良の形態 】

【 0 0 5 9 】

以下では、対象物を人の顔とし、人の顔の特徴点の3次元位置を推定する特徴点検出装置について説明する。以下の説明において、顔画像とは、少なくとも人物の顔の一部または全部の画像を含む画像である。従って、顔画像とは、人物全体の画像を含んでも良いし、人物の顔だけや上半身だけの画像を含んでも良い。また、顔画像は、複数の人物についての画像を含んでも良い。さらに、顔画像は、その背景に人物以外の風景（背景：被写体として注目された物も含む）や模様などのいかなるパターンを含んでも良い。

30

【 0 0 6 0 】

本発明者は、ASMにおいて形状モデルとして3次元形状モデルを用い、さらに可変レティナ構造による特徴量サンプリングと、正準相関分析による誤差推定行列学習とを組み合わせることにより、高速な特徴点検出を実現した。また、この組み合わせにより、ロバスト性の高い特徴点検出も実現することができる。また、この組み合わせにより、特徴点の検出と同時に顔姿勢の推定まで行うことができる。以下、本発明者の提案する技術について説明する。

【 0 0 6 1 】

まず、特徴点の検出に必要な学習処理について説明する。次に、この学習処理の結果に基づいて、本発明者が提案する特徴点検出を実施する特徴点検出装置1について、構成例や動作例などを説明する。次に、特徴点検出装置1の構成に当たり採用可能な変形例について説明する。

40

【 0 0 6 2 】

〔 学習処理 〕

まず、特徴点検出装置1を動作させるために必要となる学習処理について説明する。この学習処理は、特徴点検出装置1によって顔画像から特徴点の位置を検出するために予め実施しておく必要のある処理である。まず、本学習処理のフローを説明するにあたって必要となる処理、3次元顔形状モデルの取得、3次元顔形状モデルの画像平面への射影、レティナサンプリング、及び誤差推定行列の取得について説明する。その後、学習処理の具体的なフローの例について説明する。

50

【 0 0 6 3 】

3次元顔形状モデルの取得

学習処理では、まず、複数の学習用顔画像（以下、学習処理の説明において「顔画像」と呼ぶ）と、各顔画像における特徴点の3次元座標を用意する。特徴点は、例えば、レーザスキャナやステレオカメラなどの技術によって取得することができるが、その他どのような技術を用いても良い。この特徴点抽出処理は、学習処理の精度を高めるためにも、人間によって実施されることが望ましい。図1(a)は、顔画像から抽出された特徴点の例を示す図である。図1(a)において、斜線を含む円の中心位置が特徴点となる。具体的には、両目の両端（目頭と目尻）及び中心部分、左右のほお骨部分（眼窩底部分）、鼻の頂点、左右の口角、口の中心、鼻の左右端点と左右の口角の中間点などを特徴点とすることができる。なお、図1(a)は特徴点の位置を説明するために平面的な図として示しているが、実際には図1(b)に示すように3次元座標として特徴点が抽出される。各特徴点 p_i の座標を、 $p_i(x_i, y_i, z_i)$ とする。このとき、 i は、1から n （ n は特徴点の数を示す）の値を示す。

10

【0064】

次に、各顔画像についての特徴点配置ベクトル X を数1のように定義する。ある顔画像 j についての特徴点配置ベクトルは、 X_j と記す。なお、 X の次元数は $3n$ である。

【0065】

【数1】

$$X = [x_1, y_1, z_1, x_2, y_2, z_2, \dots, x_n, y_n, z_n]^T$$

20

【0066】

次に、取得された全ての特徴点配置ベクトル X を、適当な基準に基づき正規化する。このときの正規化の基準は、設計者によって適宜決定されて良い。以下、正規化の具体例について説明する。例えば、ある顔画像 j についての特徴点配置ベクトル X_j について、点 $p_1 \sim p_n$ の重心座標を p_G とすると、重心 p_G を原点とする座標系に各点を移動させた後、数2によって定義される L_m を用いて、その大きさを正規化することができる。具体的には、 L_m によって移動後の座標値を割ることにより、大きさを正規化することができる。ここで、 L_m は、重心から各点までの直線距離の平均値である。

【0067】

30

【数2】

$$L_m = \frac{1}{n} \sum_{i=1}^n \sqrt{(x_i - x_G)^2 + (y_i - y_G)^2 + (z_i - z_G)^2}$$

【0068】

また、回転に対しては、例えば両目の中心を結ぶ直線が一定方向を向くように特徴点座標に対して回転変換を行うことにより、正規化することができる。以上の処理は、回転、拡大・縮小の組み合わせで表現できるため、正規化後の特徴点配置ベクトル x は数3のように表すことができる（相似変換）。

40

【0069】

【数 3】

$$\mathbf{x} = s\mathbf{R}_x\mathbf{R}_y\mathbf{R}_z\mathbf{X} + \mathbf{t}$$

$$\left(\begin{array}{l} \mathbf{R}_x = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \cos\theta & -\sin\theta \\ 0 & \sin\theta & \cos\theta \end{bmatrix}, \mathbf{R}_y = \begin{bmatrix} \cos\phi & 0 & \sin\phi \\ 0 & 1 & 0 \\ -\sin\phi & 0 & \cos\phi \end{bmatrix}, \mathbf{R}_z = \begin{bmatrix} \cos\psi & -\sin\psi & 0 \\ \sin\psi & \cos\psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \\ \mathbf{t} = \begin{bmatrix} t_x \\ t_y \\ t_z \end{bmatrix} \end{array} \right) \quad 10$$

【0070】

次に、この正規化特徴点配置ベクトルの集合に対し、主成分分析を行う。主成分分析は例えば以下のように行うことができる。まず、数4に示される式に従って平均ベクトル（平均ベクトルはxの上部に水平線を記すことにより示される）を取得する。なお、数4において、Nは、顔画像の数、即ち特徴点配置ベクトルの数を示す。

【0071】

【数4】

$$\bar{\mathbf{x}} = \frac{1}{N} \sum_{j=1}^N \mathbf{x}_j$$

20

【0072】

そして、数5に示されるように、全ての正規化特徴点配置ベクトルから平均ベクトルを差し引くことにより、差分ベクトル \mathbf{x}'_j を取得する。画像jについての差分ベクトルは、 \mathbf{x}'_j と示される

【0073】

【数5】

$$\mathbf{x}'_j = \mathbf{x}_j - \bar{\mathbf{x}}$$

30

【0074】

上記した主成分分析の結果、固有ベクトルと固有値との組が3n個得られる。任意の正規化特徴点配置ベクトルは、数6に示される式によって表すことができる。

【0075】

【数6】

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}\mathbf{b}$$

40

【0076】

ここで、Pは固有ベクトル行列を示し、bは形状パラメタベクトルを示す。それぞれの値は数7に示される通りである。なお、 \mathbf{e}_i は、固有ベクトルを示す。

【0077】

【数7】

$$\mathbf{P} = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_{3n}]^T$$

$$\mathbf{b} = [b_1, b_2, \dots, b_{3n}]$$

【0078】

50

実際には、固有値の大きい上位 k 次元までの値を用いることにより、任意の正規化特徴点配置ベクトル \mathbf{x} は数 8 のように近似して表すことができる。以下、固有値の大きい順に、 \mathbf{e}_i を第 i 主成分と呼ぶ。

【 0 0 7 9 】

【数 8】

$$\mathbf{x} = \bar{\mathbf{x}} + \mathbf{P}'\mathbf{b}'$$

$$\mathbf{P}' = [\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_k]'$$

$$\mathbf{b}' = [b_1, b_2, \dots, b_k]$$

10

【 0 0 8 0 】

なお、実際の顔画像に顔形状モデルをあてはめる（フィッティングさせる）際には、正規化特徴点配置ベクトル \mathbf{x} に対して相似変換（平行移動，回転）を行う。相似変換のパラメタを $s_x, s_y, s_z, s_\theta, s_\phi, s_\psi$ とすると、形状パラメタとあわせて、モデルパラメタ \mathbf{k} を数 9 のように表すことができる。

【 0 0 8 1 】

【数 9】

$$\mathbf{k} = [s_x, s_y, s_z, s_\theta, s_\phi, s_\psi, b_1, b_2, \dots, b_k]$$

【 0 0 8 2 】

20

このモデルパラメタ \mathbf{k} によって表される 3 次元顔形状モデルが、ある顔画像上の特徴点位置にほぼ正確に一致する場合に、そのパラメタをその顔画像における 3 次元正解モデルパラメタと呼ぶ。正確に一致しているか否かは、設計者により設定される閾値や基準に基づいて判断される。

【 0 0 8 3 】

射影処理

3 次元形状モデルは、2 次元平面に射影することにより 2 次元画像上での処理が可能になる。3 次元形状を 2 次元平面に射影する方法としては、平行投影法、透視投影法などの各種の手法が存在する。ここでは、透視投影法のうち単点透視投影を例に説明する。もっとも、他のどのような手法を使用しても同様の効果を得ることができる。 $z = 0$ 平面への単点透視投影行列は、数 10 に示すとおりである。

30

【 0 0 8 4 】

【数 10】

$$\mathbf{T} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & r \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

【 0 0 8 5 】

40

ここで、 $r = -1/z_c$ であり、 z_c は z 軸上の投影中心を表す。これにより、3 次元座標 $[x, y, z]$ は数 11 に示すように変換され、 $z = 0$ 平面上の座標系で数 12 のように表される。

【 0 0 8 6 】

【数 1 1】

$$\begin{bmatrix} x & y & z & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & r \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} x & y & 0 & rz+1 \end{bmatrix}$$

【0 0 8 7】

【数 1 2】

$$\begin{bmatrix} x^* & y^* \end{bmatrix} = \begin{bmatrix} \frac{x}{rz+1} & \frac{y}{rz+1} \end{bmatrix}$$

10

【0 0 8 8】

以上の処理により、3次元形状モデルは2次元平面に射影される。

【0 0 8 9】

レティナサンプリング

次に、特徴量のサンプリングについて説明する。特徴量のサンプリングは、画像上に射影された顔形状モデルに対し可変レティナ構造を組み合わせて行う。レティナ構造とは、ある着目したい特徴点（ノード）の周囲に放射状に離散的に配置されたサンプリング点の構造のことである。図2は、レティナ構造の例を示す図である。図2では、注目したい特徴点の座標は原点であり、その他の座標に位置する各点がレティナ構造におけるサンプリング点を示す。レティナ構造によるサンプリングを実施することにより、特徴点周りの情報を、低次元で効率的にサンプリングすることが可能となる。本学習処理では、3次元顔形状モデルから2次元平面に射影された顔形状モデル（以下、2次元顔形状モデルという）の各ノードの射影点（各点p）においてレティナ構造によるサンプリングを実施する。なお、レティナ構造によるサンプリングとは、レティナ構造に従って定められたサンプリング点においてサンプリングを実施することを言う。

20

【0 0 9 0】

図3は、2次元顔形状モデルのある一つのノードに対してレティナ構造を適用した場合のサンプリング点の分布例を示す図である。図3において、斜線を含む円は顔形状モデルのノードを示し、黒丸はレティナ構造におけるサンプリング点を示す。

30

【0 0 9 1】

レティナ構造は、i番目のサンプリング点の座標を $q_i(x_i, y_i)$ とすると、数13のように表すことができる。

【0 0 9 2】

【数 1 3】

$$\mathbf{r} = [\mathbf{q}_1^T, \mathbf{q}_2^T, \dots, \mathbf{q}_m^T]^T$$

【0 0 9 3】

従って、例えばある点p(x_p, y_p)について、レティナ構造によるサンプリングを行うことにより得られるレティナ特徴量 f_p は、数14のように表すことができる。

40

【0 0 9 4】

【数 1 4】

$$\mathbf{f}_p = [f(\mathbf{p} + \mathbf{q}_1), \dots, f(\mathbf{p} + \mathbf{q}_m)]^T$$

【0 0 9 5】

ただし、 $f(p)$ は、点p（サンプリング点p）での特徴量を示す。また、レティナ構造における各サンプリング点では、画像の輝度値，Sobelフィルタ特徴量，Harr Wavelet特徴量，Gabor Wavelet特徴量やこれらを複合した値が、そのサンプリング点における特徴量として取得される。特徴量が多次元の場合、レティナ特徴

50

量は数 15 のように表すことができる。

【 0 0 9 6 】

【 数 1 5 】

$$\mathbf{f}_p = [f_1(\mathbf{p} + \mathbf{q}_1^{(1)}), \dots, f_D(\mathbf{p} + \mathbf{q}_1^{(D)}), \dots, f_1(\mathbf{p} + \mathbf{q}_m^{(1)}), \dots, f_D(\mathbf{p} + \mathbf{q}_m^{(D)})]^T$$

【 0 0 9 7 】

ここで、 D は特徴量の次元数、 $f_d(\mathbf{p})$ は、点 \mathbf{p} での第 d 次元の特徴量を表す。また、 $\mathbf{q}_i^{(d)}$ は第 d 次元に対するレティナ構造の、 i 番目のサンプリング座標を示す。

【 0 0 9 8 】

なお、レティナ構造は、顔形状モデルのスケールに応じてその大きさを変化させることができる。例えば、平行移動パラメタ s_z に反比例させて、レティナ構造の大きさを変化させることができる。このとき、レティナ構造 r は数 16 のように表すことができる。なお、 r は適当な固定値である。また、レティナ構造は、顔形状モデルにおける他のパラメタに応じて回転や形状変化させても良い。また、レティナ構造は、顔形状モデルの各ノードによってその形状（構造）が異なるように設定されても良い。また、レティナ構造は中心点一点のみの構造であっても良い。すなわち、特徴点（ノード）のみをサンプリング点とする構造もレティナ構造に含まれる。

【 0 0 9 9 】

【 数 1 6 】

$$\mathbf{r} = \alpha s_z^{-1} [\mathbf{q}_1^T, \mathbf{q}_2^T, \dots, \mathbf{q}_m^T]^T$$

【 0 1 0 0 】

あるモデルパラメタによって定まる 3 次元顔形状モデルにおいて、射影平面上に射影された各ノードの射影点ごとに上記のサンプリングを行って得られたレティナ特徴量を一列に並べたベクトルを、その 3 次元顔形状モデルにおけるサンプリング特徴量 \mathbf{f} と呼ぶ。サンプリング特徴量 \mathbf{f} は数 17 のように表すことができる。数 17 において、 n は顔形状モデルにおけるノードの数を示す。

【 0 1 0 1 】

【 数 1 7 】

$$\mathbf{f} = [\mathbf{f}_{p1}^T, \mathbf{f}_{p2}^T, \dots, \mathbf{f}_{pn}^T]^T$$

【 0 1 0 2 】

なお、サンプリング時に、各ノードにおいて正規化を行う。例えば、特徴量が 0 から 1 の範囲に収まるようにスケール変換を行うことにより正規化を行っても良い。また、一定の平均や分散をとるように変換を行うことにより正規化を行っても良い。また、特徴量によっては正規化を行わなくても良い場合がある。

【 0 1 0 3 】

誤差推定行列の取得

次に、本学習処理において実行する誤差推定行列の取得について説明する。誤差推定行列の取得は、正準相関分析 (Canonical Correlation Analysis) を用いることにより実施する。正準相関分析は、二つの次元の異なる変量間の相関関係を求める手法の一つである。正準相関分析により、顔形状モデルの各ノードが誤った位置（検出すべき特徴点と異なる位置）に配置されてしまった際にどの方向に修正すべきかの相関関係について学習結果を得ることができる。

【 0 1 0 4 】

まず、学習用の顔画像の特徴点の 3 次元位置情報から 3 次元顔形状モデルを作成する。または、学習用の顔画像の 2 次元正解座標点から 3 次元顔形状モデルを作成する。そして、3 次元顔形状モデルから正解モデルパラメタを作成する。この正解モデルパラメタを、乱数などにより一定範囲内ですらすることにより、少なくともいずれかのノードが特徴点の 3 次元位置からずれているずれ配置モデルを作成する。図 4 は、ずれ配置モデルを学習用

10

20

30

40

50

顔画像に射影したときの各ノードの射影点の例を示す図である。図4において、斜線を含む円が射影点の位置を示す。図4(a)は、ある顔画像における正解モデルパラメタを学習用顔画像に射影することによって得られる各ノードの射影点の位置を示す図である。図4(b)は、図4(a)に示された正解モデル(正解モデルパラメタによって表される3次元顔形状モデル)をずらすことによって作成されたずれ配置モデルを学習用顔画像に射影することによって得られる各ノードの射影点の位置を示す図である。そして、ずれ配置モデルに基づいて取得したサンプリング特徴量と、ずれ配置モデルと正解モデルとの差と、を組として、相関関係についての学習結果を取得する。以下、具体的な処理を説明する。

【0105】

10

まず、二組の変量ベクトル x と y を数18のように定義する。 x は、ずれ配置モデルに対するサンプリング特徴量を示す。 y は、正解モデルパラメタ (k_{opt}) とずれ配置モデルパラメタ (ずれ配置モデルを示すパラメタ: k_{dif}) との差を示す。

【0106】

【数18】

$$x = [x_1, x_2, \dots, x_p]^T$$

$$y = [y_1, y_2, \dots, y_q]^T = k_{opt} - k_{dif}$$

【0107】

20

二組の変量ベクトルは、予め各次元ごとに平均0、分散1に正規化する。正規化するために用いたパラメタ(各次元の平均、分散)は、後述する特徴点の検出処理において必要となる。以下、それぞれを x_{ave} , x_{var} , y_{ave} , y_{var} とし、正規化パラメタと呼ぶ。

【0108】

次に、二つの変量に対する線形変換を数19のように定義した場合、 u , v 間の相関を最大にするような a , b を求める。

【0109】

【数19】

$$u = a_1 x_1 + \dots + a_p x_p = \mathbf{a}^T \mathbf{x}$$

$$v = b_1 y_1 + \dots + b_q y_q = \mathbf{b}^T \mathbf{y}$$

30

【0110】

上記 a と b とは、 x , y の同時分布を考え、その分散共分散行列 を数20のように定義した場合に、数21に示す一般固有値問題を解いたときの最大固有値に対する固有ベクトルとして得られる。

【0111】

【数20】

$$\Sigma = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

40

【0112】

【数21】

$$\left(\Sigma_{XY} \Sigma_{YY}^{-1} \Sigma_{YX} - \lambda^2 \Sigma_{XX} \right) \mathbf{A} = 0$$

$$\left(\Sigma_{YX} \Sigma_{XX}^{-1} \Sigma_{XY} - \lambda^2 \Sigma_{YY} \right) \mathbf{B} = 0$$

【0113】

これらのうち、次元の低い方の固有値問題を先に解く。例えば、1番目の式を解いて得られる最大固有値が λ_1 、対応する固有ベクトルが \mathbf{a}_1 であった場合、ベクトル \mathbf{b}_1 は、

50

数 2 2 に表される式によって得られる。

【 0 1 1 4 】

【数 2 2】

$$\mathbf{b}_1 = \frac{1}{\lambda_1} \Sigma_{YY}^{-1} \Sigma_{YX} \mathbf{a}_1$$

【 0 1 1 5 】

このようにして求められた λ_1 を第一正準相関係数と呼ぶ。また、数 2 3 によって表される u_1, v_1 を第一正準変量と呼ぶ。

【 0 1 1 6 】

【数 2 3】

$$u_1 = \mathbf{a}_1^T \mathbf{x}$$

$$v_1 = \mathbf{b}_1^T \mathbf{y}$$

【 0 1 1 7 】

以下、2番目に大きい固有値に対応する第二正準変量、3番目に大きい固有値に対応する第三正準変量というように、固有値の大きさに基づいて順に求めていく。なお、後述する特徴点の検出処理に用いるベクトルは、固有値がある一定以上の値(閾値)を有する第 M 正準変量までのベクトルとする。このときの閾値は、設計者によって適宜決定されて良い。以下、第 M 正準変量までの変換ベクトル行列を、 \mathbf{A}' 、 \mathbf{B}' とし、誤差推定行列と呼ぶ。 \mathbf{A}' 、 \mathbf{B}' は、数 2 4 のように表すことができる。

【 0 1 1 8 】

【数 2 4】

$$\mathbf{A}' = [\mathbf{a}_1, \dots, \mathbf{a}_M]$$

$$\mathbf{B}' = [\mathbf{b}_1, \dots, \mathbf{b}_M]$$

【 0 1 1 9 】

\mathbf{B}' は、一般に正方行列とはならない。しかし、特徴点の検出処理において逆行列が必要となるため、 \mathbf{B}' に対し擬似的に 0 ベクトルを追加し、正方行列 \mathbf{B}'' とする。正方行列 \mathbf{B}'' は数 2 5 のように表すことができる。

【 0 1 2 0 】

【数 2 5】

$$\mathbf{B}'' = [\mathbf{b}_1, \dots, \mathbf{b}_M, \mathbf{0}, \dots, \mathbf{0}]$$

【 0 1 2 1 】

なお、誤差推定行列を求めることは、線形回帰、線形重回帰、非線形重回帰などの分析手法によっても可能である。しかし、正準相関分析を用いることにより、小さな固有値に対応する変量の影響を無視することが可能となる。従って、誤差推定に影響しない要素の影響を排除することが可能となり、より安定した誤差推定が可能となる。よって、係数効果を必要としないのであれば、正準相関分析ではなく上記した他の分析手法を用いて誤差推定行列の取得を実施することも可能である。また、誤差推定行列は、SVM、RVMなどの手法によって取得することも可能である。

【 0 1 2 2 】

学習処理の流れ

以上、本学習処理に必要な三つの処理、即ち顔形状モデルの取得、レティナサンプリング、及び誤差推定行列の取得について説明した。次に、これらの処理を用いて学習処理を行う際のフローについて説明する。図 5 は、学習処理のフローの例を示すフローチャートである。以下、図 5 を用いて学習処理の流れの例について説明する。なお、以下に説明する処理は、学習処理プログラムがインストールされた情報処理装置によって実行される。この情報処理装置は、接続された記憶装置から学習用の顔画像及び各顔画像において

10

20

30

40

50

予め抽出されている各特徴点の3次元座標位置(特徴点の正解座標)を読み出すことが可能に構成される。この情報処理装置は、上記した記憶装置からのみならず、ネットワークを介することにより、又は記録媒体にアクセスすること等によって、上記データを読み出すことが可能に構成されても良い。

【0123】

まず、変数 i を定義し、1を代入する(S01)。次に、予め特徴点の3次元位置が取得されている学習用の顔画像のうち、 i 番目の顔画像(Img_i)を読み込む(S02)。ここでは、 i に1が代入されているため1番目の顔画像(Img_1)が読み込まれる。次に、顔画像 Img_i の特徴点の正解座標の集合を読み出し、正解モデルパラメタ k_{opt} を取得し3次元顔形状モデル(正解モデル)を作成する(S03)。次に、正解モデルパラメタ k_{opt} に基づいてずれ配置モデルパラメタ k_{dif} を作成し、ずれ配置モデルを作成する(S04)。このずれ配置モデルの作成は乱数を発生させて所定の範囲内で正解モデルからずらすことが好ましい。次に、ずれ配置モデルを学習画像上に射影する(S05)。そして、ずれ配置モデルが射影された2次元形状モデルに基づいて、レティナ構造を用いたサンプリングを実行し、サンプリング特徴量 f_i を取得する(S06)。また、正解モデルパラメタ k_{opt} とずれ配置モデルパラメタ k_{dif} とに基づいて形状モデルの誤差(ずれ) dp_i を取得する(S07)。ここで、全ての学習用の顔画像について処理が完了したか否か判断する(S08)。例えば、 i の値と、学習用の顔画像の数を示す n とを比較することにより判断することができる。処理していない顔画像がある場合には(S08-*No*)、 i の値をインクリメントし(S09)、新たな i の値に基づいてS02以降の処理を実行する。一方、全ての顔画像について処理が完了したと判断した場合(S08-*Yes*)、各顔画像について得られたサンプリング特徴量 f_i と形状モデルの誤差 dp_i との集合について正準相関分析を実行する(S10)。そして、予め定められた閾値よりも小さい固定値に対応する不要な相関行列を削除し(S11)、最終的に誤差推定行列を取得する(S12)。

【0124】

上記の処理では、各学習用画像に対してずれ配置モデルを1つしか作成していないが、複数個のずれ配置モデルを作成しても良い。すなわち、学習用の画像に対してS03からS07の処理を複数回(例えば10~100回)実行しても良い。

【0125】

〔特徴点検出装置〕

次に、特徴点検出装置1について説明する。特徴点検出装置1は、上記した学習処理によって得られる誤差推定行列や正規化パラメタ等に基づいて、入力された顔画像から特徴点を検出する。まず、特徴点検出装置1の装置構成について説明する。

【0126】

装置構成

例えば、特徴点検出装置1は、ハードウェア的には、バスを介して接続されたCPU(中央演算処理装置)、主記憶装置(RAM)、補助記憶装置などを備えるように構成されても良い。この場合は、プログラムがCPUによって実行されることにより、特徴点検出装置1が実現される。ここで言う補助記憶装置は、不揮発性記憶装置を用いて構成される。不揮発性記憶装置とは、いわゆるROM(Read-Only Memory: EPROM(Erasable Programmable Read-Only Memory), EEPROM(Electrically Erasable Programmable Read-Only Memory)、マスクROM等を含む)、FeRAM(Ferroelectric RAM)、ハードディスク等を指す。

【0127】

図6は、特徴点検出装置1の機能ブロック例を示す図である。特徴点検出装置1は、補助記憶装置に記憶された各種のプログラム(OS, アプリケーション等)が主記憶装置にロードされCPUにより実行されることによって、画像入力部2, 顔検出部3, 特徴量取得部4, 学習結果記憶部5, 及び誤差推定部6等を含む装置として機能する。また、特徴

点検出装置 1 の全部又は一部は、専用のチップとして構成されても良い。次に、特徴点検出装置 1 が含む各機能部について説明する。

【 0 1 2 8 】

画像入力部 2 は、顔画像のデータを特徴点検出装置 1 へ入力するためのインタフェースとして機能する。画像入力部 2 によって、顔画像のデータが特徴点検出装置 1 へ入力される。画像入力部 2 は、特徴点検出装置 1 へ顔画像のデータを入力するためのどのような既存技術を用いて構成されても良い。

【 0 1 2 9 】

例えば、ネットワーク（例えばローカル・エリア・ネットワークやインターネット）を介して顔画像のデータが特徴点検出装置 1 へ入力されても良い。この場合、画像入力部 2 はネットワークインタフェースを用いて構成される。また、デジタルカメラやスキャナやパーソナルコンピュータや記録装置（例えばハードディスクドライブ）等から顔画像のデータが特徴点検出装置 1 へ入力されても良い。この場合、画像入力部 2 は、デジタルカメラやパーソナルコンピュータや記録装置などと特徴点検出装置 1 とをデータ通信可能に接続する規格（例えば USB（Universal Serial Bus）や SCSI（Small Computer System Interface）等の有線接続や bluetooth（登録商標）等の無線接続の規格）に応じて構成される。また、記録媒体（例えば各種フラッシュメモリやフロッピー（登録商標）ディスクや CD（Compact Disk）や DVD（Digital Versatile Disc、Digital Video Disc））に記録された顔画像のデータが特徴点検出装置 1 へ入力されても良い。この場合、画像入力部 2 は、記録媒体からデータを読み出す装置（例えばフラッシュメモリリーダーやフロッピー（登録商標）ディスクドライブや CD ドライブや DVD ドライブ）を用いて構成される。

【 0 1 3 0 】

また、特徴点検出装置 1 がデジタルカメラ等の撮像装置又はデジタルカメラ等の撮像装置を備える各種装置（例えば PDA（Personal Digital Assistant）や携帯電話機）の内部に含まれ、撮像された画像が顔画像のデータとして特徴点検出装置 1 へ入力されても良い。この場合、画像入力部 2 は、CCD（Charge-Coupled Devices）や CMOS（Complementary Metal-Oxide Semiconductor）センサ等を用いて構成されても良いし、CCD や CMOS センサなどによって撮像された顔画像のデータを特徴点検出装置 1 に入力させるためのインタフェースとして構成されても良い。また、特徴点検出装置 1 が、プリンタやディスプレイ等の画像出力装置の内部に含まれ、出力データとしてこれらの画像出力装置に入力された画像が顔画像のデータとして特徴点検出装置 1 へ入力されても良い。この場合、画像入力部 2 は、これらの画像出力装置に入力された顔画像のデータを特徴点検出装置 1 において取り扱い可能なデータに変換する装置などを用いて構成される。

【 0 1 3 1 】

また、画像入力部 2 は上記された複数の場合に応じることが可能となるように構成されても良い。

【 0 1 3 2 】

顔検出部 3 は、画像入力部 2 を介して入力された顔画像から、人の顔を検出する。顔検出部 3 は、例えば、顔全体の輪郭に対応した基準テンプレートを用了テンプレートマッチングによって顔を検出するように構成されても良い。また、顔検出部 3 は、顔の構成要素（目，鼻，耳など）に基づくテンプレートマッチングによって顔を検出するように構成されても良い。また、顔検出部 3 は、クロマキー処理によって頭部などの頂点を検出し、この頂点に基づいて顔を検出するように構成されても良い。また、顔検出部 3 は、肌の色に近い領域を検出し、その領域を顔として検出するように構成されても良い。また、顔検出部 3 は、ニューラルネットワークを使って教師信号による学習を行い、顔らしい領域を顔として検出するように構成されても良い。また、顔検出部 3 による顔検出処理は、その他、既存のどのような技術が適用されることによって実現されても良い。

【0133】

顔検出部3は、検出された顔の位置を特定し、その位置を特徴量取得部4へ渡す。このとき、顔検出部3は、顔の傾きや顔の大きさを取得して、その値を特徴量取得部4にさらに渡すように構成されても良い。

【0134】

特徴量取得部4は、顔検出部3によって検出された人の顔の画像からサンプリング特徴量を取得する。以下、特徴量取得部4の具体的な処理について説明する。

【0135】

特徴量取得部4は、顔検出部3によって検出された顔について初めてサンプリング特徴量を取得する場合には、まず、3次元顔形状モデルにおける各ノードの3次元位置を決定し、その形状モデルのパラメタ(初期パラメタ) k_{init} を取得する。この3次元形状モデルは、例えば検出された顔の中心と3次元顔形状モデルの中心とが一致するような形状であっても良い。また、3次元形状モデルは、例えば顔検出部3によって顔矩形(検出された顔を囲む矩形)が取得される場合には、この顔矩形のある頂点(例えば左上のかど)から所定の位置に顔形状モデルの所定のノードが配置されるような形状であっても良い。特徴量取得部4が作成する3次元顔形状モデルは、検出された顔に対し適当にノードが配置されるのであれば、その他どのように設計者によって定義されても良い。

【0136】

また、初期パラメタ k_{init} とは、数9によって表されるモデルパラメタ k のうち、初期値によって表されるモデルパラメタをいう。初期パラメタ k_{init} には、適当な値が設定されて良い。ただし、一般的な顔画像から得られる平均的な値を初期パラメタ k_{init} に設定することにより、様々な顔の向きや表情変化などに対応することが可能となる。従って、例えば、相似変換のパラメタ $s_x, s_y, s_z, s_r, s_t, s_s$ については、学習処理の際に用いた顔画像の正解モデルパラメタの平均値を用いても良い。また、例えば、形状パラメタ b については、ゼロとしても良い。また、顔検出部3によって顔の向きの情報が得られる場合には、この情報を用いて初期パラメタを設定しても良い。その他、設計者が経験的に得た他の値をもって初期パラメタとしても良い。

【0137】

次に、特徴量取得部4は、設定された初期パラメタによって表される顔形状モデルを入力された顔画像に射影する。そして、特徴量取得部4は、射影された形状モデルを用いて、レティナ構造に基づいたサンプリングを実施し、サンプリング特徴量 f を取得する。そして、特徴量取得部4は、得られたサンプリング特徴量 f を、誤差推定部6へ渡す。

【0138】

一方、特徴量取得部4は、顔検出部3によって検出された顔についてサンプリング特徴量を取得するのが二度目以降の場合には、誤差推定部6によって得られた新たなモデルパラメタ k (即ち、正解モデルパラメタの推定値 k_{i+1})によって表される顔形状モデルについて、サンプリング特徴量 f を取得する。そして、この場合も、特徴量取得部4は、得られたサンプリング特徴量 f を、誤差推定部6へ渡す。

【0139】

学習結果記憶部5は、不揮発性記憶装置を用いて構成され、上記した学習処理によって得られる学習結果を記憶している。具体的には、学習結果記憶部5は、誤差推定行列、正規化パラメタなどを記憶している。また、必要に応じて、学習結果記憶部5は、初期パラメタ k_{init} 又は初期パラメタ k_{init} を構成する値の一部(例えば、学習処理の際に用いた顔画像の正解モデルパラメタにおける相似変換パラメタの平均値など)をさらに記憶しても良い。

【0140】

誤差推定部6は、特徴量取得部4によって取得されたサンプリング特徴量 f 、及び学習結果記憶部5に記憶される誤差推定行列や正規化パラメタなどに基づいて、誤差量を推定し新たなモデルパラメタ k を取得する。以下、誤差推定部6の具体的な処理例について説明する。

10

20

30

40

50

【 0 1 4 1 】

まず、誤差推定部 6 は、正規化パラメタ (x_{ave} , x_{var}) を用いて、特徴量取得部 4 により取得されたサンプリング特徴量 f を正規化し、正準相関分析を行うためのベクトル x を取得する。そして、誤差推定部 6 は、数 2 6 により表される式に基づいて第一～第 M 正準変量を算出することにより、変量 u を取得する。

【 0 1 4 2 】

【 数 2 6 】

$$\mathbf{u} = [u_1, \dots, u_M]^T = \mathbf{A}'^T \mathbf{x}$$

【 0 1 4 3 】

次に、誤差推定部 6 は、数 2 7 に表される式を用いて、正規化誤差推定量 y を取得する。なお、数 2 7 において、 B' が正方行列でない場合には、 B'^{-1} は B' の擬似逆行列である。

【 0 1 4 4 】

【 数 2 7 】

$$\mathbf{y} = \mathbf{B}'^{-1} \mathbf{u}'$$

【 0 1 4 5 】

次に、誤差推定部 6 は、正規化誤差推定量 y に対し、正規化パラメタ (y_{ave} , y_{var}) を用いて復元処理を行い、誤差推定量 k_{err} を取得する。 k_{err} は、現在のモデルパラメタ k_i から正解モデルパラメタ k_{opt} までの誤差推定量である。従って、正解モデルパラメタの推定値 k_{i+1} は、現在のモデルパラメタ k_i に誤差推定量 k_{err} を加算することにより取得できる。ただし、 k_{err} は誤差を含んでいる可能性がある。このため、より安定した検出を行うために、数 2 8 に表される式によって正解モデルパラメタの推定値 k_{i+1} を取得することができる。数 2 8 において、 σ は適当な固定値であり、設計者によって適宜決定されて良い。また、 σ は、例えば i の変化に従って変化しても良い。

【 0 1 4 6 】

【 数 2 8 】

$$\mathbf{k}_{i+1} = \mathbf{k}_i + \frac{\mathbf{k}_{err}}{\sigma}$$

【 0 1 4 7 】

誤差推定部 6 は、上記のサンプリング、誤差推定を繰り返し正解モデルパラメタの推定値 k_i を正解パラメタに近づけていくことが好ましい。このような繰り返し処理を行う場合には、誤差推定部 6 は、推定値 k_i を得る度に、終了判定を行う。例えば、誤差推定部 6 は、 i の値が閾値を超えた場合に、処理を終了すると判定しても良い。また、例えば誤差推定部 6 は、数 2 9 によって表される k の値が閾値以下になった場合に、処理を終了すると判定しても良い。また、誤差推定部 6 は、取得された k_{i+1} の値が正常範囲内であるか否かに基づいて終了判定を行っても良い。例えば、誤差推定部 6 は、取得された k_{i+1} の値が、明らかに人の顔の画像における正解位置を示すものでない場合には、エラーを出力することにより終了することを決定しても良い。また、誤差推定部 6 は、取得された k_{i+1} によって表されるノードの一部が、処理対象の画像からはみでてしまった場合にも、エラーを出力することにより終了することを決定しても良い。

【 0 1 4 8 】

【 数 2 9 】

$$\Delta k = k_{i+1} - k_i$$

【 0 1 4 9 】

誤差推定部 6 は、処理を続行すると判定した場合、取得された正解モデルパラメタの推

10

20

30

40

50

定値 k_{i+1} を特徴量取得部 4 へ渡す。一方、誤差推定部 6 は、処理を終了すると判定した場合、その時点で得られている正解モデルパラメタの推定値 k_i (又は k_{i+1} であっても良い) を、最終推定パラメタ k_{est} として出力する。

【0150】

動作例

図 7 は、特徴点検出装置 1 の動作例を示すフローチャートである。以下、図 7 を用いて特徴点検出装置 1 の動作例について説明する。まず、特徴点検出装置 1 は、画像入力部 2 を介して、処理対象となる画像を読み込む (S20)。次に、顔検出部 3 は、処理対象となる画像から人の顔を検出する。次に、特徴量取得部 4 は、初期パラメタ k_{init} に基づいた 3 次元顔形状モデルを、顔検出部 3 による顔検出処理の結果に従って得られる初期位置に配置する (S21)。次に、特徴量取得部 4 は、変数 i を定義し 1 を代入する。また、特徴量取得部 4 は、 k_i を定義し、初期パラメタ k_{init} を代入する (S22)。そして、特徴量取得部 4 は k_i で表される 3 次元顔形状モデルを処理対象の画像上に射影する (S23)。特徴量取得部 4 は、射影された 2 次元形状モデルに従って、レティナ構造に基づき、サンプリング特徴量 f を取得する (S24)。

10

【0151】

次に、誤差推定部 6 は、取得されたサンプリング特徴量 f や学習結果記憶部 5 に記憶される誤差推定式などにより、3 次元形状モデル k_i と正解モデルパラメタとの推定誤差 k_{err} を取得する (S25)。誤差推定部 6 は、推定誤差に基づいて正解モデルパラメタの推定値 k_{i+1} を数 28 にしたがって算出する (S26)。次に、誤差推定部 6 は、 k を k_{i+1} と k_i との差として算出し (S27)、 E を k の二乗として算出する (S28)。そして、誤差推定部 6 は終了判定を行う。

20

【0152】

終了判定において、誤差推定部 6 は、まず取得された k_{i+1} の値が正常範囲内であるか否か判断する (S29)。 k_{i+1} の値が正常範囲内でない場合 (S29 - No)、誤差推定部 6 は、不図示の出力装置などにエラーを出力し (S30)、特徴点検出装置 1 としての処理を終了する。一方、 k_{i+1} の値が正常範囲内である場合 (S29 - Yes)、S28 の処理で取得された E の値が閾値 () を超えているか否か判断する (S31)。 E が閾値を超えていない場合 (S31 - Yes)、誤差推定部 6 は、処理が収束したとして、 k_{est} を出力し (S32)、特徴点検出装置 1 としての処理を終了する。

30

【0153】

一方、 E が閾値を超えている場合 (S31 - No)、誤差推定部 6 は、 k_{i+1} の値を特徴量取得部 4 に渡す。そして、特徴量取得部 4 は、この k_{i+1} の値に基づいて新たな 3 次元顔形状モデルを作成する (S33)。この後、誤差推定部 6 は、 i の値をインクリメントする (S34)。特徴量取得部 4 は、新たな 3 次元顔形状モデルを処理対象画像に射影してサンプリング特徴量 f の取得などを行う。即ち、この場合、新たな 3 次元顔形状モデルに基づいて S23 以降の処理が繰り返される。

【0154】

作用 / 効果

従来の ASM では、顔形状モデルの各ノード周囲において探索処理を繰り返す必要があった。また、従来の AAM では、各ノードによって構成されるパッチ内で均質且つ高密度なサンプリングを行い、さらにこの結果に対し形状修正処理を行う必要があった。具体的には、従来の AAM では通常 1 万点以上のサンプリングが必要とされていた。即ち、特徴点の検出処理には処理速度や応答性が要求されているにも関わらず、従来のこれらの技術では、計算量を削減することが十分には実現されていなかった。

40

【0155】

一方、特徴点検出装置 1 では、第一に、形状修正処理を必要としない。さらに、特徴点検出装置 1 では、レティナ構造に基づく非均質且つ低密度なサンプリングによって、特徴点の検出処理が行われる。具体的には、通常数百点程度のサンプリングで検出処理が可能である。このように、特徴点検出装置 1 によれば、特徴点の検出に要する計算量を従来に

50

比べて大幅に削減することが可能となる。また、サンプリング点が少なくなるため、学習結果として保存しておくデータ量を抑制することも可能となる。

【0156】

また、特徴点検出装置1では、誤差推定行列を正準相関分析によって求めるため、予測誤差の原因となる余分な要素の影響を予め排除することができる。また、線形演算のみによって誤差推定を行うことができるため、高速に誤差推定処理を行うことができる。

【0157】

また、特徴点検出装置1は顔形状モデルを使用して特徴点の探索を行うため、処理対象の顔画像において、一部の特徴点が隠れていたり、学習画像の顔と大きく異なっていた場合でも安定して特徴点検出を行うことができる。さらに、顔形状モデルとして3次元顔形状モデルを用いているため、より精度良く安定した結果を得ることができる。

【0158】

また、3次元顔形状モデルを用いているため、検出過程もしくは検出結果から顔姿勢（顔が向いている方向やその角度）を直接求めることができる。すなわち、特徴点検出装置1が出力した顔形状モデルパラメタに含まれるパラメタ s_x , s_y , s_z から顔の姿勢を求めることが可能である。

【0159】

変形例

学習結果記憶部5は、予め複数の顔の状態を示す初期パラメタ k_{init} を記憶しているても良い。例えば、学習結果記憶部5は、横を向いた顔や上を向いた顔などに対応する初期パラメタ k_{init} の値を記憶しているても良い。そして、特徴量取得部4は、取得される k_{i+1} の値によっては、使用する顔形状モデルを変更しても良い。具体的には、まず、各初期パラメタが担当する顔形状モデルの範囲を予め定めておく。特徴量取得部4は、誤差推定部6によって新たに取得された k_{i+1} の値が採用した初期パラメタの担当範囲からずれた場合には、この k_{i+1} の値を破棄して、採用された初期パラメタ k とは異なる別の初期パラメタを学習結果記憶部5から読み出し、その値に基づいてS21以降の処理を再度実行するように構成されても良い。

【0160】

また、特徴量取得部4は、例えば、特徴点の検出処理が繰り返し実行される場合には、ある時点で得られた k_{i+1} の値に応じて次に入力される顔画像に対する検出処理で用いる顔形状モデルを変更しても良い。処理対象となっている画像に含まれる顔が横を向いている場合などには、 k_{i+1} で表されるモデルを用いて処理を続行するよりも、横を向いた顔の顔形状モデルを示す初期パラメタを用いて処理を行った方がより精度の高い結果を得られる場合がある。従って、このような場合には、上記の変形例に従って構成することが有効である。

【0161】

このような構成により、より大きな顔姿勢の変化に対応することができるようになる。

【0162】

また、特徴量取得部4は、初期パラメタ k_{init} ではなく、レティナ構造や、誤差推定行列などを切り替えるように構成されても良い。例えば、初期の正解モデルパラメタ推定は形状パラメタを固定して移動・回転のパラメタだけを変化させて行い、移動・回転パラメタがほぼ決定した後に形状パラメタの推定を行う処理を行うことが考えられる。すなわち、初期のモデルパラメタ推定においては、誤差推定量 k_{err} のうちから移動と回転に関するパラメタ s_x , s_y , s_z , s_{θ_x} , s_{θ_y} , s_{θ_z} を利用してモデルパラメタを正解パラメタに近づけていく。そして、移動と回転に関するパラメタが収束したら誤差推定量 k_{err} から形状パラメタ b_1 , b_2 , \dots , b_k を利用してモデルパラメタを正解パラメタに近づけていく。このような構成により、より顔向きの変化にロバストな特徴点検出・顔姿勢推定が可能となる。また、一度に使用するパラメタの次元数を減らすことができるので処理を高速化することができる。なお、上記のように移動と回転に関するパラメタが修正してから初めて形状パラメタを近似させる構成以外も、繰り返し処理の回数に応

じて形状パラメタの変化を大きくしていく構成をとることができる。

【0163】

また、誤差推定部6は、終了判定を行うことなく、その時点で得られた k_{i+1} の値を k_{est} として出力しても良い。

【0164】

また、特徴点検出装置1は、顔検出部3を備えることなく構成されても良い。この場合、特徴量取得部4は、例えば画面中央を初期位置とすることにより処理を行うように構成されても良い。また、顔検出部3が特徴点検出装置1の外部に設けられ、顔画像とともに検出された顔の位置情報が特徴点検出装置1に入力されても良い。

【0165】

また、顔検出部3は、人の顔そのものではなく、人の目や口や他の特徴点を検出するように構成されても良い。この場合、特徴量取得部4は、これらの特徴点の位置に基づいて初期位置を決定するように構成されても良い。

【0166】

また、特徴量取得部4は、レティナ構造以外の構造に基づいてサンプリング点の配置を行っても良い。例えば、特徴量取得部4は、ノードの射影点の周囲に離散的に配置された複数の点から特徴量を取得しても良い。

【0167】

また、特徴点検出装置1は、人の顔を対象物として人の顔の特徴点の位置を推定するものとして説明したが、対象物は人の顔以外のものであっても良い。例えば、特徴点検出装置は1、人の全体(全身)やレントゲン画像やCT画像における臓器など、大きさに個人差がある対象物や基本的な形が変わらずに変形する対象物を対象として、特徴点の位置を推定しても良い。また、特徴点検出装置1は、自動車などの工業製品のように変形しない剛体に対しても形状モデルを設定することができるため、このような剛体を対象物としてその特徴点の位置を推定しても良い。

【図面の簡単な説明】

【0168】

【図1】顔画像から抽出された特徴点の例を示す図である。

【図2】レティナ構造の例を示す図である。

【図3】顔形状モデルのある一つのノードに対してレティナ構造を適用した場合のサンプリング点の分布例を示す図である。

【図4】ずれ配置モデルの例を示す図である。

【図5】学習処理の処理例を示すフローチャートである。

【図6】特徴点検出装置の機能ブロックの例を示す図である。

【図7】特徴点検出装置の動作例を示すフローチャートである。

【符号の説明】

【0169】

- | | |
|---|---------|
| 1 | 特徴点検出装置 |
| 2 | 画像入力部 |
| 3 | 顔検出部 |
| 4 | 特徴量取得部 |
| 5 | 学習結果記憶部 |
| 6 | 誤差推定部 |

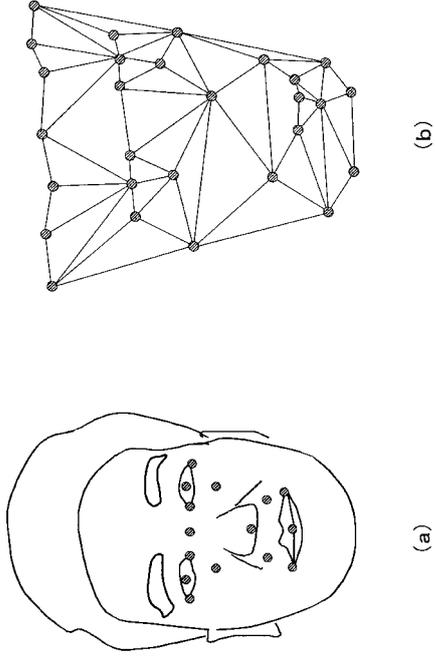
10

20

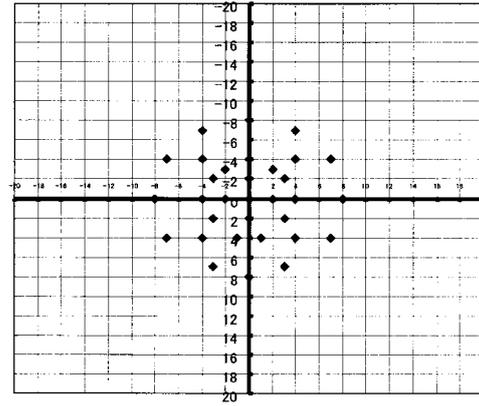
30

40

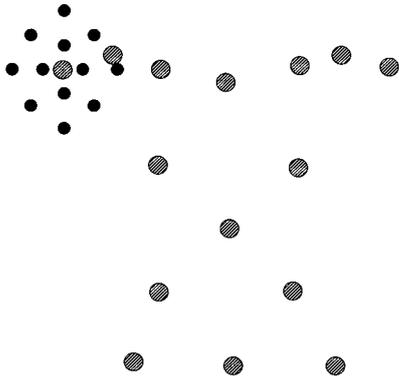
【図 1】



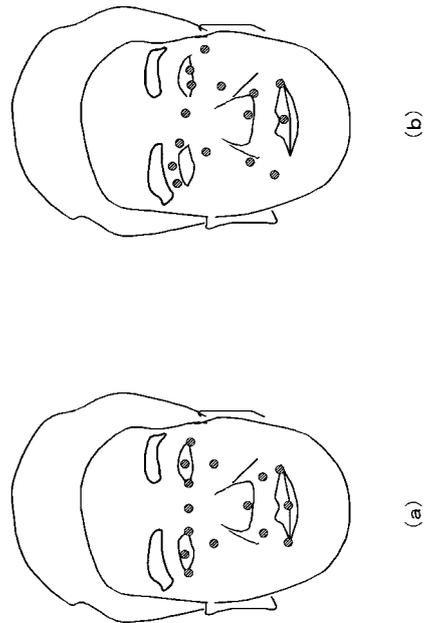
【図 2】



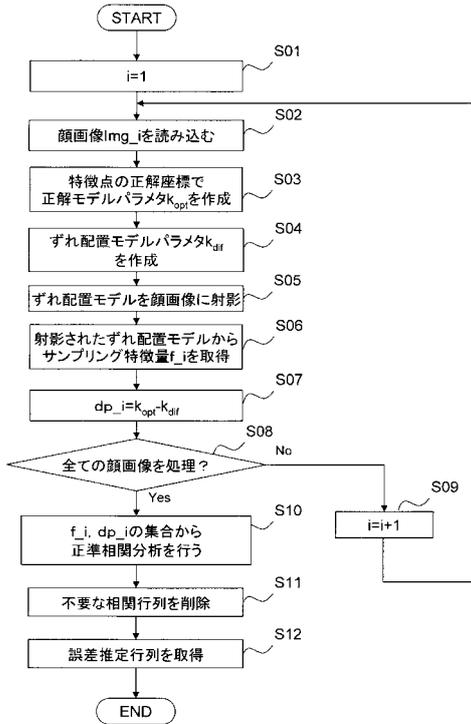
【図 3】



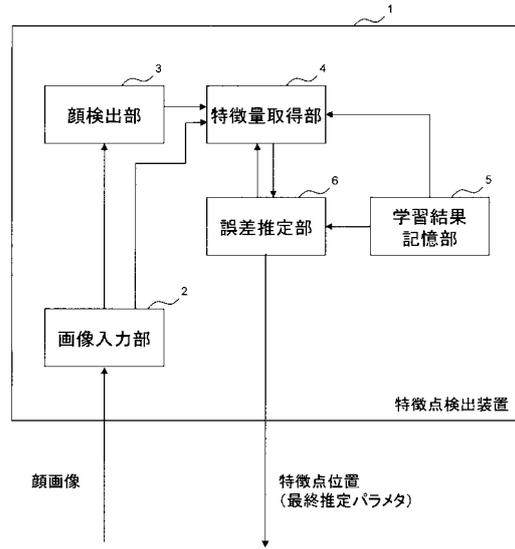
【図 4】



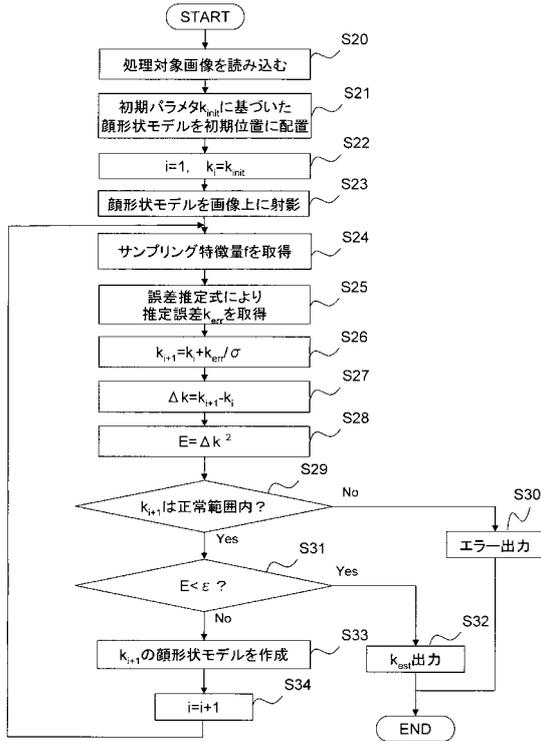
【図5】



【図6】



【図7】



フロントページの続き

- (56)参考文献 Andreas Lanitis, Chris J. Taylor, and Timothy F.Cootes , Automatic Interpretation and Coding of Face Images Using Flexible Models , IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE , 1997年 7月 , VOL.19, NO.7 , p.743-756
Timothy F. Cootes, Gareth J. Edwards, and Christopher J. Taylor , Active Appearance Models , IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE , 2001年 6月 , VOL.23, NO.6 , p.681-685

(58)調査した分野(Int.Cl. , DB名)

G 0 6 T 7 / 6 0

G 0 6 T 7 / 0 0