

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4755202号
(P4755202)

(45) 発行日 平成23年8月24日(2011.8.24)

(24) 登録日 平成23年6月3日(2011.6.3)

(51) Int. Cl.		F I			
G06T	1/00	(2006.01)	G06T	1/00	340A
G06T	7/00	(2006.01)	G06T	7/00	300D

請求項の数 23 (全 14 頁)

(21) 出願番号	特願2007-555707 (P2007-555707)	(73) 特許権者	501253316
(86) (22) 出願日	平成18年2月20日 (2006.2.20)		ミツビシ・エレクトリック・アールアンド ディー・センター・ヨーロッパ・ビーヴィ MITSUBISHI ELECTRIC R&D CENTRE EUROPE B. V. イギリス国、サリー・ジーユー2・5ワイ ディ、ギルドフォード、ザ・サリー・リサ ーチ・パーク、フレデリック・サンガー・ ロード 20 20 Frederick Sanger Road, The Surrey R esearch Park, Guild ford, Surrey GU2 5Y D, Great Britain 最終頁に続く
(65) 公表番号	特表2008-530701 (P2008-530701A)		
(43) 公表日	平成20年8月7日 (2008.8.7)		
(86) 国際出願番号	PCT/GB2006/000591		
(87) 国際公開番号	W02006/087581		
(87) 国際公開日	平成18年8月24日 (2006.8.24)		
審査請求日	平成21年2月17日 (2009.2.17)		
(31) 優先権主張番号	05250970.0		
(32) 優先日	平成17年2月21日 (2005.2.21)		
(33) 優先権主張国	欧州特許庁 (EP)		

(54) 【発明の名称】 顔特徴の検出方法

(57) 【特許請求の範囲】

【請求項 1】

画像中の物体の検出方法であって、前記物体は顔特徴群を備え、前記顔特徴群は両目を含み、

前記方法は、

前記物体の1つの目の特徴である第1の特徴と、前記物体の両目間の特徴である第2の特徴について、それぞれの単純化された外観を定義する第1および第2のテンプレートをを用いることと、

前記第1および第2のテンプレートにマッチする各特徴を識別するために、前記第1および第2のテンプレートを用いて同時にテンプレートマッチングを行うことと、

3つ一組の特徴を有する領域を少なくとも1つを検出するために、前記第1および第2のテンプレートのそれぞれについて前記テンプレートマッチングの結果を処理することであって、前記3つ一組の特徴のうち2つの特徴は前記第1のテンプレートにマッチし1つの特徴は前記第2のテンプレートにマッチする、前記テンプレートマッチングの結果を処理することと

を含む、画像中の物体の検出方法。

【請求項 2】

前記第1のテンプレートは、左目の特徴および右目の特徴のためのものであり、前記第2のテンプレートは、両目間の特徴のためのものである、請求項1に記載の方法。

【請求項 3】

10

20

左目、右目および両目間の3つ一組の特徴を有する、前記検出された少なくとも1つの領域を、口領域の存在によって検証することを含む、請求項2に記載の方法。

【請求項4】

閉じた目用のテンプレートおよび開いた目用のテンプレートの双方を用いることを含む、請求項1に記載の方法。

【請求項5】

1つの物体に関する前記少なくとも2つの異なるテンプレートの各々の前記テンプレートマッチングの結果を結合することを含む、請求項1に記載の方法。

【請求項6】

同一の物体に対する前記テンプレートマッチングの結果にそれぞれ関連する、複数の信頼度マップを導出することを含む、請求項1に記載の方法。

10

【請求項7】

最初に元画像をダウンサンプリングすることと、
関心物体を含む可能性がある領域を少なくとも1つ検出する前記ステップの後に、前記検出された領域に関して前記元画像の前記テンプレートマッチングを行うことと
を含む、請求項1～6のいずれか一項に記載の方法。

【請求項8】

異なる解像度についての前記テンプレートマッチングの結果を結合することを含む、請求項7に記載の方法。

【請求項9】

20

複数の解像度でテンプレートマッチングを行うことと、
異なる解像度に関して複数の信頼度マップを導出することと
を含む、請求項1～8のいずれか一項に記載の方法。

【請求項10】

前記第1および第2のテンプレートについて前記テンプレートマッチングの結果を処理する前記ステップは、前記信頼度マップを結合することを含む、請求項6又は9に記載の方法。

【請求項11】

前記結合することは、ピクセル毎の乗算、加算、又は最大値選択を含む、請求項10に記載の方法。

30

【請求項12】

前記結合することの後に閾値処理及びモルフォロジー後処理が続く、請求項11に記載の方法。

【請求項13】

連結性解析を、請求項11に記載の方法から得られる2値画像に対して、行うことを含む、請求項1～12に記載の方法。

【請求項14】

領域配置解析を行うことを含む、請求項1～13のいずれか一項に記載の方法。

【請求項15】

前記領域配置解析は対称性の検査を含む、請求項14に記載の方法。

40

【請求項16】

テンプレートマッチングの結果を、
テクスチャ情報、及び、前記画像中の他の物体に関する情報のうちの、1つ又は複数と結合することを含む、請求項1～15のいずれか一項に記載の方法。

【請求項17】

請求項1に記載の方法を用いる、画像中の顔領域候補の選択方法であって、
左目、右目及び両目間の3つ一組の領域を考慮することと、
口領域を用いて前記3つ一組の領域を検証することと
を含む、画像中の顔領域候補の選択方法。

【請求項18】

50

請求項 1 ~ 17 のいずれか一項に記載の方法を実行する装置。

【請求項 19】

請求項 1 ~ 17 のいずれか一項に記載の方法を実行するようにプログラムされる制御装置。

【請求項 20】

請求項 19 に記載の制御装置と、画像を記憶する記憶手段とを備える装置。

【請求項 21】

請求項 1 ~ 17 のいずれか一項に記載の方法を実行させるためのコンピュータプログラム。

【請求項 22】

請求項 1 ~ 17 のいずれか一項に記載の方法を実行するコンピュータシステム。

【請求項 23】

請求項 1 ~ 17 のいずれか一項に記載の方法を実行させるためのコンピュータプログラムを記憶したコンピュータ読み取り可能な記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、画像中の物体、特に顔特徴を検出又は位置特定する方法及び装置に関する。

【背景技術】

【0002】

顔特徴の検出は、顔認識システム、コンテンツベースの画像検索、ビデオ符号化、テレビ会議、群衆監視、及び知的人間コンピュータインタフェース等の分野における多くのコンピュータビジョンアプリケーションに必要な最初のステップである。

【0003】

雑然としたシーン中の顔を位置特定する典型的な顔検出問題が与えられると、低レベルの解析がまず、グレースケール及び色等のピクセル特性を用いる視覚特徴の分割を扱う。これらの特徴は低レベルの性質を持つため、この解析から生成される特徴は通常曖昧である。特徴解析では、顔形状に関する情報を用いて、視覚特徴が顔及び顔特徴のより包括的な概念に整理される。特徴解析により、特徴の曖昧さが低減され、顔及び顔特徴の位置が決定される。

【0004】

画像中のグレーレベル情報はしばしば顔特徴を識別するために用いられる。眉、瞳孔、及び唇等の特徴は通常、周囲の顔領域よりも暗く見える。この特性を利用して、様々な顔部品を区別することができる。

【0005】

いくつかの最近の顔特徴抽出アルゴリズム (P.J.L. Van Beek、M.J.T. Reinders、B. Sankur、J.C.A. Van Der Lubbe 著「Semantic Segmentation of Videophone Image Sequences」(Proc. of SPIE Int. Conf. on Visual Communications and Image Processing, 1992, pp. 1182-1193)、H.P. Graf、E. Cosatto、D. Gibson、E. Petajan、及び M. Kocheisen 著「Multi-modal System for Locating Heads and Faces」(IEEE Proc. of 2nd Int. Conf. on Automatic Face and Gesture Recognition, Vermont, Oct. 1996, pp. 277-282)、並びに K.M. Lam 及び H. Yan 著「Facial Feature Location and Extraction for Computerised Human Face Recognition」(Int. Symposium on Information Theory and Its Applications, Sydney, Australia, Nov. 1994)) は、分割された顔領域内のグレー最小値を探索する。これらのアルゴリズムにおいて、入力画像はまず、コントラスト伸長及びグレースケールモルフォロジー (濃度階調形態素) ルーチンにより強調されて局所的な暗いパッチの品質が改善され、それにより検出が容易になる。暗いパッチの抽出は、低レベルグレースケールの閾値処理により達成される。

【0006】

C. Wong、D. Kortenkamp、及び M. Speich 著「A Mobile Robot that Recognises People

10

20

30

40

50

」(IEEE Int. Conf. on Tools with Artificial Intelligence, 1995)の論文は、色解析により間接的に得られる顔候補中の暗い顔領域も探すロボットの実装を記載している。このアルゴリズムは、重み付けした人間の目のテンプレートを利用して、一对の目の考えられる位置を判定する。

【0007】

R. Hoogenboom及びM. Lew著「Face Detection Using Local Maxima」(IEEE Proc. of 2nd Int. Conf. on Automatic Face and Gesture Recognition, Vermont, Oct. 1996, pp. 334-339)では、8つの暗い近傍ピクセルで囲まれた1つの明るいピクセルで定義される極大値が代わりに用いられ、鼻先のような明るい顔のスポットが指示される。検出点は次に特徴テンプレートと位置合わせされ、相関測定が行われる。

10

【0008】

一方、G. Yang及びT.S. Huang著「Human Face Detection in a Complex Background」(Pattern Recognition, 27, 1994, 53-63)の論文は、モザイク(ピラミッド)画像における顔のグレースケール挙動を研究している。顔画像の解像度をサブサンプリング又は平均化により徐々に低くしていくと、顔の巨視的特徴が消える。低解像度では、顔領域は均一になる。この観察に基づき、階層的な顔検出の枠組みが提案されている。低解像度画像から始めて、均一な領域を探索する一組の規則により顔候補が確立される。顔候補は次に、高解像度における極小値を用いて、顕著な顔特徴の存在により検証される。この技法は、回転不変顔検出システム(X.-G. Lv, J. Zhou, 及びC.-S. Zhang著「A Novel Algorithm for Rotated Human Face Detection」(IEEE Conference on Computer Vision and Pattern Recognition, 2000))に採用され、このアルゴリズムの拡張が、C. Kotropoulos及びI. Pitas著「Rule-based Face Detection in Frontal Views」(Proc. Int. Conf. on Acoustic, Speech and Signal Processing, 1997)に提示されている。

20

【0009】

S. Kawato及びN. Tetsutani著「Real-time Detection of Between-the-Eyes with a Circle Frequency Filter」(ACCV2002: The 5th Asian Conference on Computer Vision, Melbourne, 2002)の論文は、両目間の点はほとんどの人に共通しており、広範囲の顔の向きで見えると主張している。さらに、この点の周囲のパターンはどんな表情でも極めて安定している。著者らは、独自のリング周波数フィルタによるこの特徴の検出アルゴリズムを提案している。両目間の点を検出した後、近傍の顔特徴の検出タスクは大幅に簡略化されるため、本発明の提案でもこのアイデアを用いる。

30

【0010】

米国特許第6690814号明細書は、画像シーケンスからの顔及び顔特徴の検出に関する。顔検出タスクは、新たな顔画像を事前に記録されている背景画像と比較する際に、カメラ位置の特別な配置により簡略化される。画像減算及び閾値処理を用いて、顔領域が位置特定され、目のような顔特徴が顔領域の中の暗い領域として抽出される。様々なサイズの特別な照合フィルタが暗い領域に適用される。出力値が最大となる特定のフィルタサイズを瞳孔のサイズと見なす。フィルタリング後、フィルタ出力がガウスフィルタにより平滑化され、結果として得られる極大値が瞳孔候補となる。カメラ位置の特別な配置が、この方法の考えられる用途を制限する。本発明の提案は任意の事前に定義されたカメラ位置に依存しない。

40

【0011】

米国特許第6681032号明細書は、顔検出モジュール及び目検出モジュールを含む顔認識システムを記載している。この方法の本発明の提案する方法との主な違いは、顔検出方法が、肌の色の分布の抽出及びその事前に保存されている情報との照合に完全に依拠していることである。目検出モジュールは、関心領域の、参照画像セットから構築された固有目テンプレートとの照合に基づく。

【0012】

米国特許第6611613号明細書は、顔画像の目領域は、強いグレー特徴(色成分の最大値と最小値の差が小さい)を共通して含むという観察に基づく目検出方法を含む。抽

50

出された領域は次に、幾何学的特徴（小ささ、アスペクト比）及びテクスチャ特徴（強い水平方向エッジの存在）により検証される。抽出された一対の目は、顔領域のサイズ、向き及び位置を決め、さらに他の顔特徴（眉、鼻腔、口）により検証される。

【0013】

米国特許第6381345号明細書は、テレビ会議用途に適した目検出方法を記載している。まず、ガウシアンフィルタを用いて画像がぼかされ、次に、両目が大きな暗い領域として抽出され、次に、幾何学的特徴を用いて両眉が除去される。その後、明度閾値処理を用いて両目がセグメンテーションされ、両目のパラメータが求められる。この方法の、本発明による提案との主な違いは、顔が常に画像の中心に位置しているものと仮定し、一対の目のみを抽出することである。またこの方法は、所定の一組の明度閾値を用いるが、これは、様々な照明条件下では使用が難しいことが分かった。

10

【0014】

米国特許第6151403号明細書の方法は、本発明の方法とは異なる以下のステップ、すなわち、(a)色を基に画像中の考えられる肌領域を判定すること、(b)モルフォロジー演算を用いて肌領域内で谷領域を決定すること、(c)谷領域において適用される相互相関を用いてテンプレートマッチングを行うことから成る。

【0015】

米国特許第6130617号明細書に記載される目検出方法は、顔画像を2値化するステップと、対で存在する候補領域を抽出するステップと、全ての対の中から1つの候補対を鼻腔と判定するステップと、鼻腔に対して正三角形を形成する残りの候補対を目の候補対として設定するステップと、より小さな正三角形を形成する候補対を目と判定するステップとを含む。この手法は、主要な顔特徴としての鼻腔の検出に基づく。この特徴は、カメラの向きの特別な配置（上向きの光軸）下でのみ安定していることが分かった。本発明の提案では、両目間領域を主要な顔特徴として用いる。この特徴は、様々な顔の向き及びカメラ向きで安定している。

20

【0016】

米国特許第5870138号明細書は、顔及び顔特徴の検出にHSV色空間を用いる方法を記載している。顔領域の検出にはH成分及びS成分のみが用いられる。口は、顔領域内でバンドパスフィルタを用いてS成分及びV成分から検出される。顔領域内のV成分は正規化され、目テンプレートとの相関を用いて目が位置特定される。領域追跡を用いて探索エリアが縮小される。

30

【0017】

米国特許第5715325号明細書は、顔検出の最終段階で目特徴を用いる人物特定システムを記載している。まず、画像の解像度が低減及び正規化され、照明の変化が補償される。次に、画像は事前に記憶されている背景画像と比較され、バイナリ関心マスク(binary interest mask)が作成される。顔領域はテンプレートマッチングにより決定され、マッチングスコアが閾値を上回ると、ニューラルネットワークに基づくさらなる目位置特定手順が行われる。

【0018】

米国特許第6278491号明細書及び米国特許第5990973号明細書は、赤目検出方法及び赤目軽減方法を記載している。これらの方法の主な目的は、デジタル写真撮影におけるフラッシュの使用により生じる赤目を自動的に見つけることである。これらの方法は顔検出ステップ及び目検出ステップを含むが、これらの方法の主な欠点は、カラーで高解像度のデジタル画像でしか上手く機能しないことである。独特な特徴すなわち赤い瞳孔が目の検出に用いられる。また、これらの方法は、1枚のデジタル写真の後処理用に設計されており、計算量の多いアルゴリズム（例えば、マルチスケール及び多回転テンプレートマッチング）を用いるためにリアルタイムでのビデオ処理には向かない場合がある。

40

【0019】

本発明が取り組む課題は、複雑な環境（例えば低品質画像及び雑然とした背景）でのロバストな顔特徴検出である。

50

【 0 0 2 0 】

既存の目検出方法の1つの課題は、画像中の一貫した領域対（左目及び右目）の探索のために複雑で信頼度が低いことである。単純な画像でも多くの候補領域対が見つかる可能性があり、この曖昧さをより高レベルの判定規則により解決しなければならない。本明細書では、目検出は画像中に頻繁に現れない特徴に依拠する。この特徴は、左目領域、両目間領域及び右目領域を含む3つ一組の領域である。この3つ一組の領域はさらに、口のような他の顔特徴領域の存在により検証されるため、目の検出は曖昧さが遥かに少なくなり、時間もかからなくなる。

【 0 0 2 1 】

本発明の態様は、添付の特許請求の範囲に記載されている。本発明は、適切な装置を用いて画像に対応する信号を処理することを含む。

10

【 0 0 2 2 】

本発明の一実施の形態により提案される解決策は以下のステップを含む。

1．顔特徴テンプレートの設計及び簡略化。このテンプレートは、顔特徴の全体的な外観のみを、暗い領域及び明るい領域の集合として表す。各顔特徴は一組の異なるテンプレートを有し得る。

2．低レベル画像処理段階：

次のテンプレートマッチングに必要な時間がテンプレートサイズに関係なくなるように積分画像への画像変換。

ピクセル毎のテンプレートマッチングにより、顔特徴毎に複数の信頼度マップを得る

20

信頼度マップの結合、及び信頼度レベルの高い顔特徴領域の抽出。

3．顔特徴の解析及び選択段階：

複数組の「左目」、「両目間」、「右目」の候補の領域配置の解析及び抽出。

高信頼度レベルと、テクスチャ特徴と、他の顔特徴領域の存在とに基づき最良の顔特徴候補の選択。

4．顔特徴検出の階層的戦略

画像及びテンプレートをダウンサンプリングしたバージョンへのステップ1、2の適用

検出された顔特徴を含む関心領域（ROI）の抽出

30

テンプレートの元のバージョン及びROI内の画像へのステップ1、2の適用により顔特徴を正確に位置特定する。

様々な解像度からの結果を結合する。本方法がそれらの解像度のうちの1つで失敗したら他の解像度での結果を用いる。

【 0 0 2 3 】

本発明のさらなる態様は以下を含む。

1．統一した枠組みの中で各顔特徴に任意数のテンプレートを用いる。

2．図2に示す4つの特定のテンプレート（「両目間領域」、「水平方向の顔特徴」及び「開いた目」）を用いる。

3．統計データ解析及び信号対雑音比の計算に基づくテンプレートマッチング方法。

40

4．安定した3つ一組の領域 左目、両目間、右目 を顔領域の主な特徴として用いる。

5．高信頼度レベルと、テクスチャ特徴と、他の顔特徴領域の存在とに基づき最良の顔特徴候補の選択方法。

【 0 0 2 4 】

提案する方法は、いくつかの重要で有用な特性を有する。第1に、本方法は、リアルタイム処理を可能にする顔特徴テンプレートの設計及び簡略化の手法を説明する。第2に、既知の積分画像技法を用いることによって、テンプレートマッチング方法の低コストでリアルタイムの拡張が達成される。

【 0 0 2 5 】

50

本発明の一実施形態が添付図面に関して説明される。

【0026】

関心の顔特徴の全体的な外観が単純なテンプレートにより符号化される(図1及び図2)。テンプレートは、ピクセル値の分布がより暗い場所(黒領域)やより明るい場所(白領域)を示す領域から成る。領域形状の定性的な推定のために、2値化法を用いることができる。図1a及び図1bは、テンプレート形状の定性的な推定のために関心の顔特徴(両目間領域)を有する画像及び対応する2値化を示す。図1の例において、関心の特徴(両目間領域)は、2つの暗い楕円領域のように見える(図1bの2値化画像から導出した図1cのテンプレートを参照)。リアルタイム処理の要件により、全ての領域は矩形であることが好ましい。これにより、図1(d)に示すようにテンプレートがさらに単純化される。

10

【0027】

両目間領域の2つの異なるテンプレートを図2(a)及び図2(b)に示す。

【0028】

以下の文章及びブロック図では、これらのテンプレートを「両目間マスク1」及び「両目間マスク2」と呼ぶ。図2(c)に示す水平方向の顔特徴テンプレートは、閉じた目、口、鼻腔、及び眉を検出するのに役立つ。図2(d)のテンプレートは、開いた目の検出用に特別に設計されている(明るい近傍の中の暗い瞳孔)。以下の文章及びブロック図では、図2(c)及び図2(d)のテンプレートもそれぞれ「目マスク1」及び「目マスク2」と呼ぶ。

20

【0029】

図2に示すテンプレートは簡略化されたテンプレートである。これらのテンプレートは、各顔特徴の全体的な外観のみを暗い領域及び明るい領域(矩形)の集合として表す。

【0030】

顔特徴の階層的な検出方法のブロック図及びデータフローを図3に示す。

【0031】

まず、元画像及び全てのテンプレートをダウンサンプリングする(S11)。速度の理由から、4つの近傍ピクセルの平均化及び2分の1の画像縮小を用いるが、任意の他の画像サイズ変更方法を用いることができる。テンプレートをダウンサンプリングするには、それらの矩形の座標を2で割り、最も近い整数値に切り上げるか又は切り捨てる。

30

【0032】

次のステップS12において、画像及びテンプレートをダウンサンプリングしたバージョンに、顔特徴検出アルゴリズムを適用する。これにより、計算時間が4分の1に短縮されるが、顔特徴位置の精度も低下する場合がある。多くの場合、眼鏡等のまぎらわしい細部が、下げた解像度では見えなくなり得るため、目はダウンサンプリングした画像でより容易に検出できる。逆の状況も起こり得る。すなわち、低解像度では眉が閉じた目のように見え、閉じた目がほとんど見えなくなる可能性がある。この場合、眉が通常は最終的な検出結果となる。口は、元の解像度で検出することができれば、通常は低解像度でも検出することができる。これは、目の代わりに眉が検出されても、目及び口を含む顔領域は正確に検出されることを意味する。

40

【0033】

低解像度での検出結果に関する考えられる問題を考慮して、検出される特徴は、関心領域の抽出(S13)のみに用いられる。

【0034】

最後に、同じ検出アルゴリズムを元の解像度のテンプレート及びROI内の画像に適用して顔特徴を正確に位置特定する(S12)。大まかな口位置のような低解像度での検出の結果のいくつかをこのステップで用いることができる。このステップの計算時間はROIサイズに比例し、このROIサイズは通常、元画像のサイズより小さい。

【0035】

顔特徴検出アルゴリズム(図3にS12で示す)のブロック図及びデータフローを図4

50

に示す。このアルゴリズムには2つの主なステップがあり、第1のステップは、以下を含む低レベル画像処理である。

1) 積分画像の計算 (S 2 1)。テンプレートは異なるサイズの矩形の集合として表されるため、これらの矩形内の統計特徴 (平均及び分散) の高速計算のために特別な画像前処理が必要である。画像の積分表現への変換はそのような特徴の高速計算を4回のピクセル参照 (すなわち矩形の角) のみにより提供する。積分画像の計算は既知の従来技術の技法である。簡単に言えば、この計算は、以下のように定義される積分画像 $Sum(x, y)$ 及び $SumQ(x, y)$ を含む。

【0036】

【数1】

$$Sum(x, y) = \sum_{a \leq x} \sum_{b \leq y} I(a, b) = I(x, y) + S(x-1, y) + S(x, y-1) - S(x-1, y-1)$$

$$SumQ(x, y) = \sum_{a \leq x} \sum_{b \leq y} I^2(a, b) = I^2(x, y) + SQ(x-1, y) + SQ(x, y-1) - SQ(x-1, y-1)$$

【0037】

ここで、 $I(x, y)$ は $x, y \geq 0$ の場合に元画像であり、 $x, y < 0$ の場合に $I(x, y) = 0$ である。

2) テンプレート毎に行われるテンプレートマッチング (S 2 2)。この手順は、各ピクセル近傍の統計的仮説試験に基づく。このステップの結果は一組の信頼度マップである。2つのマップが「両目間」領域が存在する尤度を示し、もう2つのマップが考えられる目領域を示す。信頼度マップの各ピクセルは仮説試験の結果を含み、画像領域とテンプレートとの間の類似度として考えることができる。

3) 信頼度マップの結合 (S 2 3)。この実施態様では信頼度マップのピクセル毎の乗算を用いるが、加算又は最大値選択等の他の結合方法も適用可能である。ピクセル毎の乗算結果の範囲は信頼度マップの数値フォーマットに依存する。固定小数点表現の場合、さらなる除算又はビットシフトが必要となる。

4) 信頼度マップの分割 (S 2 4)。各信頼度マップは、高信頼度の領域を低信頼度の背景と分離するために分割される。類似度は信号対雑音比 (SNR) としても解釈することができ、信頼度マップを閾値処理する可能性を開く。本実施態様では、信号レベルが雑音レベルのすぐ上であることを示す $SNR \sim 1.1$ の閾値を用いる。モルフォロジー演算 (morphological operations) 又は小/大領域の除去等により洗練された分割方法もこの段階で適用可能である。

【0038】

アルゴリズムの第2のステップは、信頼度の高い画像領域を検出することを目的とする解析から成る。まず、閾値処理された信頼度マップに適用される連結成分ラベリングアルゴリズム (S 2 5) によりそのような領域を全て抽出する。次に、全ての考えられる3つ一組の領域 (左目領域、両目間領域、右目領域) を繰り返し、対称性を大まかに調べる (S 2 6)。そして最後に、テクスチャ特徴と、口及び鼻腔等の他の顔特徴領域の存在とにより検証される合計信頼度レベルが高い3つ一組の領域を選択する (S 2 7)。信頼度レベルの極大値を正確な目位置と見なす。

【0039】

好適な実施形態におけるテンプレートマッチングは以下のように行われる。

【0040】

テンプレート (図2) により定義される画像からの2つのピクセル群への統計的 t 検定の適用により以下の類似度が得られる (いくつかの等価変換は省く)。

【0041】

10

20

30

40

【数 2】

$$\left(\frac{\text{信号}}{\text{雑音}} \right)^2 = \left(\frac{\text{群間の差}}{\text{群のばらつき}} \right)^2 = \frac{|R|\sigma^2(R)}{|B|\sigma^2(B) + |W|\sigma^2(W)} - 1$$

【0042】

上記の式から定数を除去すると、本発明の類似度が得られる。

【0043】

【数 3】

10

$$SNR = \frac{|R|\sigma^2(R)}{|B|\sigma^2(B) + |W|\sigma^2(W)}$$

【0044】

ここで、テンプレート R は 2 つの領域の集合として表される： $R = B \cup W$ 。ここで、B は「黒」領域であり、W は「白」領域であり、 $\sigma^2(Q)$ は領域 Q における画像値の分散であり、 $|Q|$ は領域 Q 内のピクセル数を指示する。積分画像を用いた場合、任意の矩形領域 Q の $\sigma^2(Q)$ の計算は $2 * |Q|$ 回ではなく $2 * 4$ 回のピクセル参照を必要とする。ここでピクセル参照とは、ピクセル値を得るためのメモリ内の 2D 画像アレイへの 1 回の

20

【0045】

図 5 は、「両目間」テンプレート及び「目」テンプレートのテンプレートマッチング結果を示す。

【0046】

図 6 は、信頼度マップから抽出された一組の領域、並びに「両目間」テンプレート及び「目」テンプレートの両方の 2 つの信頼度マップの結合結果に適用された連結領域ラベリングの結果を示す。各領域はそれを囲むボックスで示される。図 7 は、対称性に基づく領域配置解析の結果を示す。この結果は、左目特徴、両目間特徴、右目特徴の候補を含む。本発明では、左目と右目との間の距離は、口と両目の中間との間の距離にほぼ等しいと仮定する。2 つの目候補を得た状態で、 $d \times d$ の寸法の矩形探索エリアを決める。ここで、 d は目位置間の距離である。この領域と両目との間の垂直距離は $d / 2$ として選択する。最も高い信頼度マップ値を含む探索エリア内の領域が口領域の候補として選択される（図 8）。

30

【0047】

テンプレートマッチング結果及び連結領域ラベリング結果を用いて、本アルゴリズムは、高い信頼度マップ値、領域対称性及び高い勾配密度に基づき最良の目候補を選択する。左目、両目間、右目及び口に対応する正しい領域組を選択するために、図 9 に示すアルゴリズムを用いる。

【0048】

40

図 9 では以下の記号を用いる：

$b(x, y)$ は「両目間」信頼度マップであり、

$B = \{ B_1, \dots, B_n \}$ は、 $b(x, y)$ から抽出された一組の連結領域である。

$e(x, y)$ は「目」信頼度マップであり、

$E = \{ E_1, \dots, E_m \}$ は、 $e(x, y)$ から抽出された一組の連結領域である。なお、E 組は両目及び口の候補領域を含む。

インデックス i, j, k はそれぞれ、現在の左目領域、右目領域及び両目間領域を指定する。

B 組の領域毎に最大値の位置及びその値 $b_{max}(x_{max}, y_{max})$ が事前に計

50

算されており、 b_{max} は領域組の合計スコアを計算するために用いられ (図 9)、 (x_{max}, y_{max}) は考えられる各両目間領域の中心を示す。

E 組の各領域は以下の事前に計算された特徴を有する。

$e_{max}(x_{max}, y_{max})$ - 最大値の位置及びその値。 e_{max} は、領域組の合計スコアを計算するために用いられ (図 9)、 (x_{max}, y_{max}) は目の中心を示す。

G_E 、 G_M - 領域 P 内で以下のように計算されるテクスチャ測定値。

【0049】

【数4】

$$G_E = \frac{1}{|P|} \sum_{(x,y) \in P} |I(x+1, y) - I(x, y)|, \quad G_M = \frac{1}{|P|} \sum_{(x,y) \in P} |I(x, y+1) - I(x, y)|$$

10

【0050】

両テクスチャ測定値は領域組の合計スコアに寄与する (図 9)。

【0051】

いくつかの代替形態及び発展形態を以下に述べる。

【0052】

主な顔特徴の検出を助けるために鼻及び鼻腔等の他の顔特徴の検出も考慮された。この実施態様は、本方法が安定して機能する状況の範囲を制限する可能性がある。これは、これらの顔特徴の外観が照明条件及びカメラ位置により影響されるためである。しかし、テレビ会議等のいくつかの特定用途では、これらの特徴を含めることによりシステムがよりロバストになる。

20

【0053】

色に基づく肌分割は画像中の探索エリアをかなり限定し、顔特徴の候補数を低減することができる。この実施態様もまた、本方法が機能し得る状況の範囲 (グレースケール画像、暗い照明条件) を制限する可能性がある。

【0054】

顔及び顔特徴の追跡は一つのオプションとして実施されてきた (画像の代わりに任意の画像 ROI を用いることができる図 3、図 4 を参照)。このモードは、計算時間を大幅に短縮するが、1枚の画像の処理には適さない。

30

【0055】

本明細書において、「画像」という用語は、画像単位を記述するために用いられる。この画像単位は、解像度の変更、アップサンプリング、ダウンサンプリング等の処理の後のものを含み、又は積分画像と関連するものを含む。また、この用語は、フレーム、フィールド、ピクチャ等の他の同様の用語、又は画像、フレーム等のサブユニット若しくは領域にも適用される。ピクセル及びブロック又はピクセル群という用語は、適宜互換可能に用いられ得る。本明細書において、画像という用語は、文脈から明らかな場合を除いて、画像全体又は画像の一領域を意味する。同様に、画像の一領域は画像全体を意味し得る。画像は、フレーム又はフィールドを含み、静止画、フィルム若しくはビデオ等の画像シーケンス中の1つの画像、若しくは関連画像群中の1つの画像に関連する。

40

【0056】

画像はグレースケール画像又はカラー画像であってもよく、又は別のタイプのマルチスペクトル画像、例えば、IR、UV若しくは他の電磁画像、若しくは音響画像等であってもよい。

【0057】

本発明は、例えばコンピュータシステムにおいて適切なソフトウェア及び/又はハードウェアの変更を行って実施することができる。例えば、本発明は、プロセッサ又は制御装置等の制御手段又は処理手段、メモリ、磁気記憶装置、CD、DVD等といった画像記憶手段を含むデータ記憶手段、ディスプレイ若しくはモニタ又はプリンタ等のデータ出力手

50

段、キーボード等のデータ入力手段、及びスキャナ等の画像入力手段、又はこのような構成要素の任意の組み合わせを追加構成要素と共に有する、コンピュータ等を用いて実施することができる。本発明の態様は、ソフトウェア形態及び/又はハードウェア形態で、又は特定用途向け装置において提供するか、又はチップ等の特定用途向けモジュールを提供することができる。本発明の一実施形態による装置内のシステムの構成要素は、例えばインターネットを通じて他の構成要素と離れた場所に設けてもよい。

【0058】

本件出願は、同一日に出願された「統計的テンプレートマッチングによる高速な物体検出方法 (Fast Method Of Object Detection By Statistical Template Matching)」と題する同一出願人による同時係属中の出願に関連し、この内容は参照により本明細書中に援用される。本明細書中に記載される実施形態の、統計学的テンプレートマッチングや、積分画像の導出および使用等の特定の態様のさらなる詳細は、上記の同時係属中の出願に見出すことができる。

10

【図面の簡単な説明】

【0059】

【図1】図1 aは顔特徴の画像を示す。図1 bは図1 aの画像を2値化したものを示す。図1 cおよび図1 dは図1 aの顔特徴用のテンプレートを示す。

【図2】図2 a ~ 図2 dは他の顔特徴用のテンプレートである。

【図3】顔特徴の検出方法を示すブロック図である。

【図4】顔特徴検出アルゴリズムを示すブロック図である。

20

【図5】元画像及びテンプレートマッチング結果を示す対応する画像を含む。

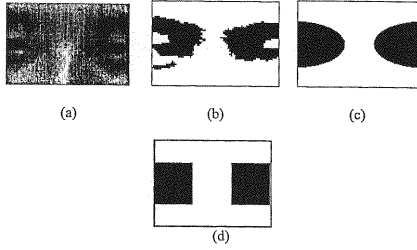
【図6】複数のテンプレートマッチング結果から成る閾値処理した信頼度マップの連結領域をラベル付けした結果を示す画像である。

【図7】3つ一組の特徴の検出結果を示す画像である。

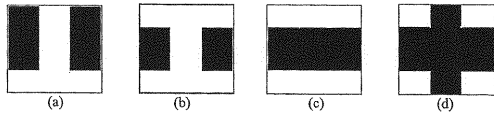
【図8】特徴検出を示す画像である。

【図9】信頼度マップ、領域対称性及びテクスチャ測定値に基づく領域解析及び顔特徴選択アルゴリズムを示すブロック図である。

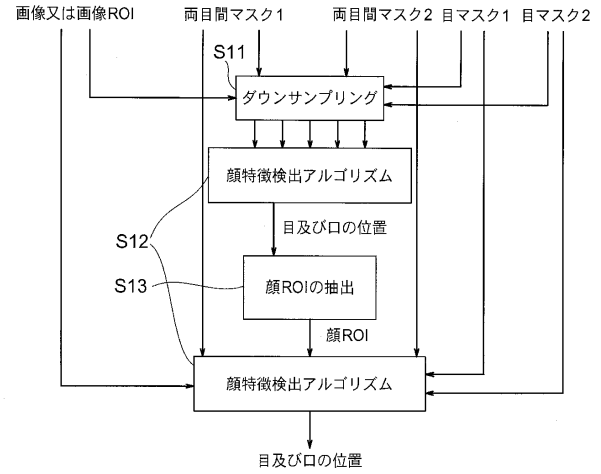
【図1】



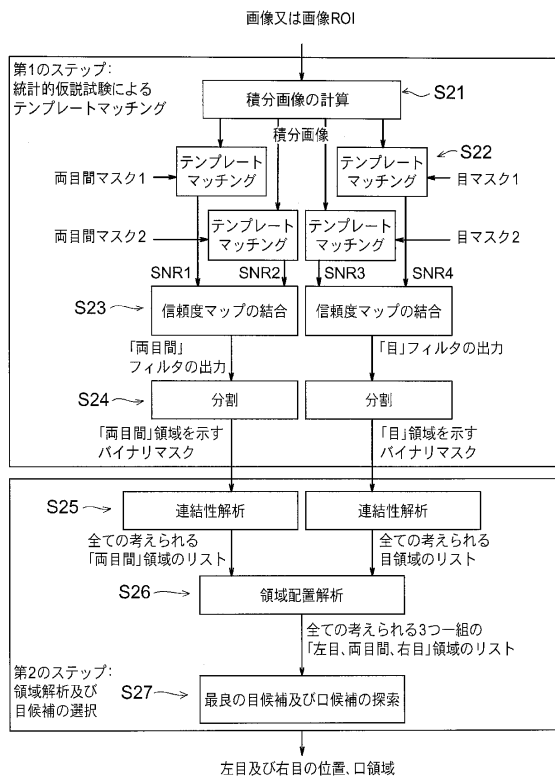
【図2】



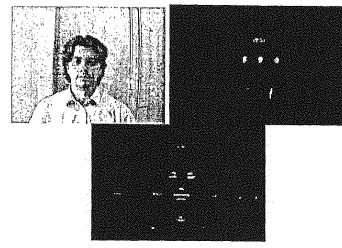
【図3】



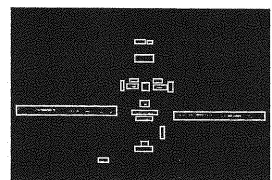
【図4】



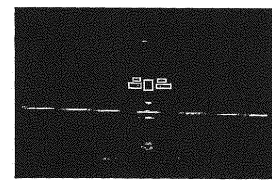
【図5】



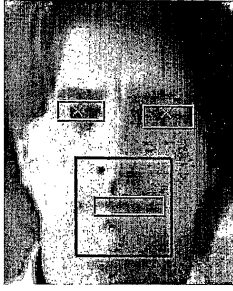
【図6】



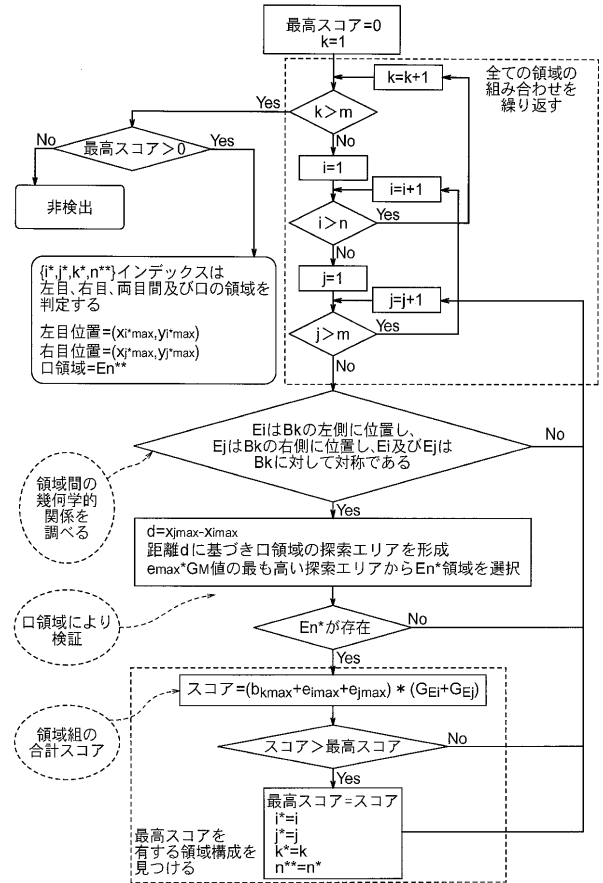
【図7】



【図8】



【図9】



フロントページの続き

(73)特許権者 000006013

三菱電機株式会社
東京都千代田区丸の内二丁目7番3号

(74)代理人 100110423

弁理士 曾我 道治

(74)代理人 100084010

弁理士 古川 秀利

(74)代理人 100094695

弁理士 鈴木 憲七

(74)代理人 100111648

弁理士 梶並 順

(74)代理人 100147500

弁理士 田口 雅啓

(74)代理人 100166235

弁理士 大井 一郎

(72)発明者 シビルヤコフ、アレクサンダー

イギリス国、サリー・ジュー-2・9ティージェイ、ギルドフォード、リンカーン・ロード 24

(72)発明者 ボバー、ミロスロー

イギリス国、サリー・ジュー-1・2エスイー、ギルドフォード、ウィッカム・ロード 21

審査官 岡本 俊威

(56)参考文献 特開2005-044330(JP,A)

特開2003-271933(JP,A)

特開2004-185611(JP,A)

特開平09-251534(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00

G06T 7/00-7/60