

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4166143号
(P4166143)

(45) 発行日 平成20年10月15日(2008.10.15)

(24) 登録日 平成20年8月8日(2008.8.8)

(51) Int. Cl. F I
G06T 1/00 (2006.01) G O 6 T 1/00 3 4 O A
G06T 7/60 (2006.01) G O 6 T 7/60 1 5 O B

請求項の数 18 (全 30 頁)

(21) 出願番号	特願2003-391148 (P2003-391148)	(73) 特許権者	393031586
(22) 出願日	平成15年11月20日(2003.11.20)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2004-185611 (P2004-185611A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成16年7月2日(2004.7.2)	(74) 代理人	100064746
審査請求日	平成17年5月17日(2005.5.17)		弁理士 深見 久郎
(31) 優先権主張番号	特願2002-338175 (P2002-338175)	(74) 代理人	100085132
(32) 優先日	平成14年11月21日(2002.11.21)		弁理士 森田 俊雄
(33) 優先権主張国	日本国(JP)	(74) 代理人	100083703
			弁理士 仲村 義平
		(74) 代理人	100096781
			弁理士 堀井 豊
		(74) 代理人	100098316
			弁理士 野田 久登
		(74) 代理人	100109162
			弁理士 酒井 将行

最終頁に続く

(54) 【発明の名称】 顔位置の抽出方法、およびコンピュータに当該顔位置の抽出方法を実行させるためのプログラムならびに顔位置抽出装置

(57) 【特許請求の範囲】

【請求項1】

人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備するステップと、

前記対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによるフィルタリング処理により眉間候補点の位置を抽出するステップと、

抽出された前記眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択するステップとを備える、顔位置の抽出方法。

【請求項2】

人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備するステップと、

前記対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによる、各前記矩形形状の相対的な明暗関係が所定の条件を満たす場合に前記眉間検出フィルタの中心を抽出するフィルタリング処理により眉間候補点の位置を抽出するステップと、

抽出された前記眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択するステップとを備える、顔位置の抽出方法。

【請求項3】

前記眉間検出フィルタは、1つの矩形形状を6分割したものである、請求項1または2に

記載の顔位置の抽出方法。

【請求項 4】

前記 6 つの矩形形状は、

鉛直方向に隣接する 2 つの第 1 の矩形形状と、

前記第 1 の矩形形状とは前記鉛直方向に所定量だけずれ、かつ前記鉛直方向に隣接する 2 つの第 2 の矩形形状と、

前記第 2 の矩形形状とは前記鉛直方向に所定量だけずれ、かつ前記鉛直方向に隣接する 2 つの第 3 の矩形形状とを含む、請求項 1 または 2 に記載の顔位置の抽出方法。

【請求項 5】

前記真の候補点を選択するステップは、

前記眉間候補点に対応する前記眉間検出フィルタを構成する矩形形状のうち、所定の 2 つの矩形形状に対応する前記対象画像に対して、目のパターン判別処理により、目の位置を検出するステップと、

前記検出された目の位置に基づいて、前記眉間候補点の位置を、2 つの目の midpoint の位置に修正するステップと、

前記修正された眉間候補点位置を中心に 2 つの目が水平となるように入力画像を回転するステップと、

前記回転された入力画像について、前記修正された眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択するステップとを含む、請求項 1 または 2 に記載の顔位置の抽出方法。

【請求項 6】

デジタルデータを準備するステップは、前記対象画像をステレオ画像として準備するステップを含み、

前記真の候補点を選択するステップは、

前記ステレオ画像に基づいて検出される前記眉間候補点の観測点からの距離に応じて、前記眉間候補点のうちから真の候補点を選択するステップを含む、請求項 1 または 2 に記載の顔位置の抽出方法。

【請求項 7】

コンピュータに、対象となる画像領域内の顔位置を抽出する方法を実行させるためのプログラムであって、前記プログラムは、

人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備するステップと、

前記対象となる画像領域内において、6 つの矩形形状の結合した眉間検出フィルタによるフィルタリング処理により眉間候補点の位置を抽出するステップと、

抽出された前記眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択するステップとを備える、プログラム。

【請求項 8】

コンピュータに、対象となる画像領域内の顔位置を抽出する方法を実行させるためのプログラムであって、前記プログラムは、

人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備するステップと、

前記対象となる画像領域内において、6 つの矩形形状の結合した眉間検出フィルタによる、各前記矩形形状の相対的な明暗関係が所定の条件を満たす場合に前記眉間検出フィルタの中心を抽出するフィルタリング処理により眉間候補点の位置を抽出するステップと、

抽出された前記眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択するステップとを備える、プログラム。

【請求項 9】

10

20

30

40

50

前記眉間検出フィルタは、1つの矩形形状を6分割したものである、請求項7または8に記載のプログラム。

【請求項10】

前記6つの矩形形状は、

鉛直方向に隣接する2つの第1の矩形形状と、

前記第1の矩形形状とは前記鉛直方向に所定量だけずれ、かつ前記鉛直方向に隣接する2つの第2の矩形形状と、

前記第2の矩形形状とは前記鉛直方向に所定量だけずれ、かつ前記鉛直方向に隣接する2つの第3の矩形形状とを含む、請求項7または8に記載のプログラム。

【請求項11】

前記真の候補点を選択するステップは、

前記眉間候補点に対応する前記眉間検出フィルタを構成する矩形形状のうち、所定の2つの矩形形状に対応する前記対象画像に対して、目のパターン判別処理により、目の位置を検出するステップと、

前記検出された目の位置に基づいて、前記眉間候補点の位置を、2つの目の中心の位置に修正するステップと、

前記修正された眉間候補点位置を中心に2つの目が水平となるように入力画像を回転するステップと、

前記回転された入力画像について、前記修正された眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択するステップとを含む、請求項7または8に記載のプログラム。

【請求項12】

デジタルデータを準備するステップは、前記対象画像をステレオ画像として準備するステップを含み、

前記真の候補点を選択するステップは、

前記ステレオ画像に基づいて検出される前記眉間候補点の観測点からの距離に応じて、前記眉間候補点のうちから真の候補点を選択するステップを含む、請求項7または8に記載のプログラム。

【請求項13】

人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備する撮影手段と、

前記対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによるフィルタリング処理により眉間候補点の位置を抽出する手段と、

抽出された前記眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択する選択手段とを備える、顔位置抽出装置。

【請求項14】

人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備する撮影手段と、

前記対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによる、各前記矩形形状の相対的な明暗関係が所定の条件を満たす場合に前記眉間検出フィルタの中心を抽出するフィルタリング処理により眉間候補点の位置を抽出する手段と、

抽出された前記眉間候補点の位置を中心として、所定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のうちから真の候補点を選択する選択手段とを備える、顔位置抽出装置。

【請求項15】

前記眉間検出フィルタは、1つの矩形形状を6分割したものである、請求項13または14に記載の顔位置抽出装置。

【請求項16】

前記6つの矩形形状は、

10
20
30
40
50

鉛直方向に隣接する 2 つの第 1 の矩形形状と、
前記第 1 の矩形形状とは前記鉛直方向に所定量だけずれ、かつ前記鉛直方向に隣接する
2 つの第 2 の矩形形状と、
前記第 2 の矩形形状とは前記鉛直方向に所定量だけずれ、かつ前記鉛直方向に隣接する
2 つの第 3 の矩形形状とを含む、請求項 1 3 または 1 4 に記載の顔位置抽出装置。

【請求項 1 7】

前記選択手段は、
前記眉間候補点に対応する前記眉間検出フィルタを構成する矩形形状のうち、所定の 2
つの矩形形状に対応する前記対象画像に対して、目のパターン判別処理により、目の位置
を検出する手段と、
前記検出された目の位置に基づいて、前記眉間候補点の位置を、2 つの目の midpoint の位置
に修正する手段と、
前記修正された眉間候補点位置を中心に 2 つの目が水平となるように入力画像を回転す
る手段と、
前記回転された入力画像について、前記修正された眉間候補点の位置を中心として、所
定の大きさで前記対象画像を切り出し、パターン判別処理に応じて、前記眉間候補点のう
ちから真の候補点を選択する手段とを含む、請求項 1 3 または 1 4 に記載の顔位置抽出装
置。

【請求項 1 8】

前記撮影手段は、前記対象画像をステレオ画像として準備する手段を含み、
前記選択手段は、前記ステレオ画像に基づいて検出される前記眉間候補点の観測点から
の距離に応じて、前記眉間候補点のうちから真の候補点を選択する手段を含む、請求項 1
3 または 1 4 に記載の顔位置抽出装置。

【発明の詳細な説明】

【技術分野】

【0 0 0 1】

この発明はカメラ等からの画像を処理する画像処理に関し、特に、画像中の人物の顔を
抽出するための画像認識の分野に関する。

【背景技術】

【0 0 0 2】

通信により、遠隔地にいる複数の人間で会議を行なう TV 会議システムが実用化されて
いる。しかしこれらシステムにおいて、映像そのものを送ると通信データ量が增大する
という問題点がある。そのために、たとえば対象となる人物の視線、顔の向き、表情等に関
する特徴データを各地で抽出し、抽出したデータのみを互いに送信する技術が研究されて
いる。受信側では、このデータに基づいて仮想的な人物の顔面の画像を生成して表示す
る。これにより、通信データ量を削減しながら、TV 会議を効率良く行なえる。

【0 0 0 3】

さらに、このような画像中から人物を検出する技術は、ヒューマンコンピュータインタ
ラクション、ジェスチャー認識、セキュリティーなどの分野の発展に欠かせない技術とし
ても、盛んに研究されている。

【0 0 0 4】

これらの人物検出技術の応用においては、1) 検出率が高い、2) 照明環境の変化に強
い、3) 実時間で動作する、という条件を満たす安定したシステムを構築する必要がある
。さらに今後、高品質画像（一画面を構成する画素数の多い画像）を対象にした実時間人
物検出の必要性が高まってくると考えられ、今後はさらに、より高速な人物検出アルゴ
リズムの開発が必要になる。

【0 0 0 5】

人物を検出するには、まず顔を検出する手法が有効である。顔は表情など重要な情報を
有しており、顔を検出できれば四肢の位置を推測して探索することが容易になる。

【0 0 0 6】

10

20

30

40

50

これまでに、肌色情報を用いた顔検出システムは報告が数多くなされている（たとえば、特許文献1、非特許文献1～非特許文献4を参照）。

【0007】

これらの手法は、画像から肌色領域を抽出し顔候補領域を求める。顔候補領域を限定できることから、処理の範囲が限定され、計算量を大幅に減らすことができるので、高速なシステムを構築することが可能である。しかし、色情報を利用する手法は照明環境の変動に弱く、一般的な環境下で動作させる場合、安定した性能を期待できない。

【0008】

一方、色情報を用いない（濃淡情報を用いる）顔検出手法では、テンプレートマッチングやニューラルネットワーク等の学習的手法を利用した手法が数多く報告されている（たとえば、非特許文献5～非特許文献6を参照）。これらの手法は高い検出率と照明環境に対するロバスト性が特徴である。たとえば、非特許文献5に開示された技術においては、ニューラルネットワークを応用し、非常に高い検出率を実現している。

【0009】

しかし、これらの手法は、サイズを変えながら画像全体にわたってテンプレート（モデル）とのマッチングをとる必要があり、計算量が多いという問題がある。そのため、画素サイズが大きくなった場合には、計算量が飛躍的に増加するため、実時間システムを構築することは非常に困難である。

【0010】

一方、非特許文献7に開示された技術では、分割領域の平均明るさの明暗関係から顔を検出するが、その領域が額から顎まで分布していて16分割領域があり、まともにヘアスタイルや髭の影響を受けてしまうという問題がある。

【非特許文献1】川戸慎二郎、鉄谷信二、"リング周波数フィルタを利用した眉間の実時間検出" 信学論(D-I I), vol. J84-D-I I, no.12, pp.2577-2584, Dec.2001.

【非特許文献2】川戸慎二郎、鉄谷信二、"目のリアルタイム検出と追跡", 信学技報, PRMU2000-63, pp.15-22, Sept.2000.

【非特許文献3】Chai D, Ngan K.N. "テレビ電話アプリケーションにおける肌色マップを用いた顔の分割" IEEE トランザクション オン サーキット アンド システムズ フォー ビデオ テクノロジー, 第9巻, No.4, pp.551-564, 1990 (Chai, D. and Ngan, K.N.: "Face Segmentation Using Skin-Color Map in Videophone Application," IEEE Trans. on Circuits and Systems For Video Technology, Vol.9, No., pp.551-564, 1990)

【非特許文献4】J. ヤン, A. ワイベル, "実時間の顔追跡器", プロシーディングス 3rd IEEE ワークショップ オン アプリケーション オブ コンピュータ ビジョン, pp.142-147, 1996年 (J. Yang and A. Waibel, "A real-time face tracker," Proc. 3rd IEEE Workshop on Application of Computer Vision, pp.142-147, 1996)

【非特許文献5】H. ローリー, S. バルージャ, T. カナダ, "ニューラルネットワークによる顔検知" IEEE トランザクション パターン アナリシス アンド マシン インテリジェンス, 第20巻, no.1, pp.23-38, 1月 1998年 (H. Rowly, S. Baluja, and T. Kanada, "Neural-Network-Based Face Detection," IEEE Trans. Pattern Analysis and Machine Intelligence, vol.20, no.1, pp.23-38, Jan.1998)

【非特許文献6】E. ジェルマス, B. K. ロウ "顔検知: サーベイ" コンピュータ ビジョン アンド イメージ アンダスタンディング, 83(3), pp.236-274, 2001年 (E. Hjelmás and B. K. Low, "Face Detection: A survey," Computer Vision and Image Understanding, 83(3), pp.236-274, 2001)

【非特許文献7】ブライアン スカセツラティ "フォビーティッド アクティブ ビジョン システムに対する顔検出による目の検出" プロシーディングス AAAI 1

10

20

30

40

50

998年 pp. 969 - 976 (Brian Scassellati, "Eye Finding via Face Detection for a Foveated, Active Vision System", Proc. AAAI '98, pp.969-976)

【特許文献1】特開2001-52176号公報明細書 上述した特許文献1に開示された技術では、安定した顔の特徴点として両目の間の点(以下では眉間(Between-the-Eyes)と呼ぶ)に着目している。つまり、眉間の周囲は、額部と鼻筋は相対的に明るく、両サイドの目と眉の部分は暗いパターンになっており、それを検出するリング周波数フィルタを用いている。

【0011】

しかしリング周波数フィルタでは、肌色領域を抽出して、領域を限定する前処理が必要な点と、前髪が眉までかかっているような顔は上述のようなパターンが現れないために、検出できない場合があるという問題があった。

10

【発明の開示】

【発明が解決しようとする課題】

【0012】

それゆえに本発明の目的は、照明の状況や人物の髪型の影響等を抑えて、画像情報から顔画像を抽出することが可能な顔位置抽出装置、そのための方法および当該方法をコンピュータを用いて実現するためのプログラムを提供することである。

【0013】

さらに、本発明の他の目的は、照明の状況や人物の髪型の影響等を抑えて、顔の眉間の位置を特定して実時間でその位置を追跡することが可能な顔位置抽出装置、そのための方法および当該方法をコンピュータを用いて実現するためのプログラムを提供することである。

20

【課題を解決するための手段】

【0014】

この発明のある局面に従うと、顔位置の抽出方法であって、人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備するステップと、対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによるフィルタリング処理により眉間候補点の位置を抽出するステップと、抽出された眉間候補点の位置を中心として、所定の大きさで対象画像を切り出し、パターン判別処理に応じて、眉間候補点のうちから真の候補点を選択するステップとを備える。

30

【0015】

好ましくは、眉間検出フィルタは、1つの矩形形状を6分割したものである、請求項1記載の顔位置の抽出方法。

【0016】

好ましくは、6つの矩形形状は、鉛直方向に隣接する2つの第1の矩形形状と、第1の矩形形状とは鉛直方向に所定量だけずれ、かつ鉛直方向に隣接する2つの第2の矩形形状と、第2の矩形形状とは鉛直方向に所定量だけずれ、かつ鉛直方向に隣接する2つの第3の矩形形状とを含む。

【0017】

好ましくは、真の候補点を選択するステップは、眉間候補点に対応する眉間検出フィルタを構成する矩形形状のうち、所定の2つの矩形形状に対応する対象画像に対して、目のパターン判別処理により、目の位置を検出するステップと、検出された目の位置に基づいて、眉間候補点の位置を、2つの目の midpoint の位置に修正するステップと、修正された眉間候補点位置を中心に2つの目が水平となるように入力画像を回転するステップと、回転された入力画像について、修正された眉間候補点の位置を中心として、所定の大きさで対象画像を切り出し、パターン判別処理に応じて、眉間候補点のうちから真の候補点を選択するステップとを含む。

40

【0018】

好ましくは、デジタルデータを準備するステップは、対象画像をステレオ画像として準備するステップを含み、真の候補点を選択するステップは、ステレオ画像に基づいて検出

50

される眉間候補点の観測点からの距離に応じて、眉間候補点のうちから真の候補点を選択するステップとを含む。

【0019】

この発明の他の局面に従うと、コンピュータに、対象となる画像領域内の顔位置を抽出する方法を実行させるためのプログラムであって、プログラムは、人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備するステップと、対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによるフィルタリング処理により眉間候補点の位置を抽出するステップと、抽出された眉間候補点の位置を中心として、所定の大きさに対象画像を切り出し、パターン判別処理に応じて、眉間候補点のうちから真の候補点を選択するステップとを備える。

10

【0020】

好ましくは、眉間検出フィルタは、1つの矩形形状を6分割したものである。

【0021】

好ましくは、6つの矩形形状は、鉛直方向に隣接する2つの第1の矩形形状と、第1の矩形形状とは鉛直方向に所定量だけずれ、かつ鉛直方向に隣接する2つの第2の矩形形状と、第2の矩形形状とは鉛直方向に所定量だけずれ、かつ鉛直方向に隣接する2つの第3の矩形形状とを含む。

【0022】

好ましくは、真の候補点を選択するステップは、眉間候補点に対応する眉間検出フィルタを構成する矩形形状のうち、所定の2つの矩形形状に対応する対象画像に対して、目のパターン判別処理により、目の位置を検出するステップと、検出された目の位置に基づいて、眉間候補点の位置を、2つの目の中点の位置に修正するステップと、修正された眉間候補点位置を中心に2つの目が水平となるように入力画像を回転するステップと、回転された入力画像について、修正された眉間候補点の位置を中心として、所定の大きさに対象画像を切り出し、パターン判別処理に応じて、眉間候補点のうちから真の候補点を選択するステップとを含む。

20

【0023】

好ましくは、デジタルデータを準備するステップは、対象画像をステレオ画像として準備するステップを含み、真の候補点を選択するステップは、ステレオ画像に基づいて検出される眉間候補点の観測点からの距離に応じて、眉間候補点のうちから真の候補点を選択するステップとを含む。

30

【0024】

この発明のさらに他の局面に従うと、顔位置抽出装置であって、人間の顔領域を含む対象画像領域内の各画素の値のデジタルデータを準備する撮影手段と、対象となる画像領域内において、6つの矩形形状の結合した眉間検出フィルタによるフィルタリング処理により眉間候補点の位置を抽出する手段と、抽出された眉間候補点の位置を中心として、所定の大きさに対象画像を切り出し、パターン判別処理に応じて、眉間候補点のうちから真の候補点を選択する選択手段とを備える。

【0025】

好ましくは、眉間検出フィルタは、1つの矩形形状を6分割したものである。

40

【0026】

好ましくは、6つの矩形形状は、鉛直方向に隣接する2つの第1の矩形形状と、第1の矩形形状とは鉛直方向に所定量だけずれ、かつ鉛直方向に隣接する2つの第2の矩形形状と、第2の矩形形状とは鉛直方向に所定量だけずれ、かつ鉛直方向に隣接する2つの第3の矩形形状とを含む。

【0027】

好ましくは、選択手段は、眉間候補点に対応する眉間検出フィルタを構成する矩形形状のうち、所定の2つの矩形形状に対応する対象画像に対して、目のパターン判別処理により、目の位置を検出する手段と、検出された目の位置に基づいて、眉間候補点の位置を、2つの目の中点の位置に修正する手段と、修正された眉間候補点位置を中心に2つの目

50

が水平となるように入力画像を回転する手段と、回転された入力画像について、修正された眉間候補点の位置を中心として、所定の大きさに対象画像を切り出し、パターン判別処理に応じて、眉間候補点のうちから真の候補点を選択する手段とを含む。

【 0 0 2 8 】

好ましくは、撮影手段は、対象画像をステレオ画像として準備する手段を含み、選択手段は、ステレオ画像に基づいて検出される眉間候補点の観測点からの距離に応じて、眉間候補点のうちから真の候補点を選択する手段を含む。

【 発明の効果 】

【 0 0 2 9 】

以上説明したとおり、本発明によれば、連続する画面情報から実時間で、人物の顔の位置、特に、眉間または目の位置を検出することができる。

【 発明を実施するための最良の形態 】

【 0 0 3 0 】

[実施の形態 1]

[ハードウェア構成]

以下、本発明の実施の形態 1 にかかる顔位置抽出装置について説明する。この顔位置抽出装置は、パーソナルコンピュータまたはワークステーション等、コンピュータ上で実行されるソフトウェアにより実現されるものであって、対象画像から人物の顔を抽出し、さらに人物の顔の映像から、眉間の位置および目の位置を検出するためのものである。図 1 に、この顔位置抽出装置の外観を示す。

【 0 0 3 1 】

図 1 を参照してこのシステム 2 0 は、C D - R O M (Compact Disc Read-Only Memory) ドライブ 5 0 および F D (Flexible Disk) ドライブ 5 2 を備えたコンピュータ本体 4 0 と、コンピュータ本体 4 0 に接続された表示装置としてのディスプレイ 4 2 と、同じくコンピュータ本体 4 0 に接続された入力装置としてのキーボード 4 6 およびマウス 4 8 と、コンピュータ本体 4 0 に接続された、画像を取込むためのカメラ 3 0 とを含む。この実施の形態の装置では、カメラ 3 0 としては C C D (固体撮像素子) を含むビデオカメラを用い、カメラ 3 0 の前にいてこのシステム 2 0 を操作する人物の眉間または目の位置を検出する処理を行なうものとする。

【 0 0 3 2 】

すなわち、カメラ 3 0 により、人間の顔領域を含む画像であって対象となる画像領域内の各画素の値のデジタルデータが準備される。

【 0 0 3 3 】

図 2 に、このシステム 2 0 の構成をブロック図形式で示す。図 2 に示されるようにこのシステム 2 0 を構成するコンピュータ本体 4 0 は、C D - R O M ドライブ 5 0 および F D ドライブ 5 2 に加えて、それぞれバス 6 6 に接続された C P U (Central Processing Unit) 5 6 と、R O M (Read Only Memory) 5 8 と、R A M (Random Access Memory) 6 0 と、ハードディスク 5 4 と、カメラ 3 0 からの画像を取込むための画像取込装置 6 8 とを含んでいる。C D - R O M ドライブ 5 0 には C D - R O M 6 2 が装着される。F D ドライブ 5 2 には F D 6 4 が装着される。

【 0 0 3 4 】

既に述べたようにこの顔位置抽出装置の主要部は、コンピュータハードウェアと、C P U 5 6 により実行されるソフトウェアとにより実現される。一般的にこうしたソフトウェアは C D - R O M 6 2 、 F D 6 4 等の記憶媒体に格納されて流通し、C D - R O M ドライブ 5 0 または F D ドライブ 5 2 等により記憶媒体から読取られてハードディスク 5 4 に一旦格納される。または、当該装置がネットワークに接続されている場合には、ネットワーク上のサーバから一旦ハードディスク 5 4 にコピーされる。そうしてさらにハードディスク 5 4 から R A M 6 0 に読出されて C P U 5 6 により実行される。なお、ネットワーク接続されている場合には、ハードディスク 5 4 に格納することなく R A M 6 0 に直接ロードして実行するようにしてもよい。

【 0 0 3 5 】

図 1 および図 2 に示したコンピュータのハードウェア自体およびその動作原理は一般的なものである。したがって、本発明の最も本質的な部分は、F D 6 4、ハードディスク 5 4 等の記憶媒体に記憶されたソフトウェアである。

【 0 0 3 6 】

なお、最近の一般的傾向として、コンピュータのオペレーティングシステムの一部として様々なプログラムモジュールを用意しておき、アプリケーションプログラムはこれらモジュールを所定の配列で必要な時に呼び出して処理を進める方式が一般的である。そうした場合、当該顔位置抽出装置を実現するためのソフトウェア自体にはそうしたモジュールは含まれず、当該コンピュータでオペレーティングシステムと協働してはじめて顔位置抽出装置が実現することになる。しかし、一般的なプラットフォームを使用する限り、そうしたモジュールを含ませたソフトウェアを流通させる必要はなく、それらモジュールを含まないソフトウェア自体およびそれらソフトウェアを記録した記録媒体（およびそれらソフトウェアがネットワーク上を流通する場合のデータ信号）が実施の形態を構成すると考えることができる。

10

【 0 0 3 7 】

〔 顔画像の抽出の基本的原理 〕

まず、本発明の手続きの概略をまとめると、顔を連続撮影したビデオ画像を処理するにあたり、横が顔幅、縦がその半分程度の大きさの矩形フィルタで画面を走査する。矩形は、たとえば、 3×2 に 6 分割されていて、各分割領域の平均明るさが計算され、それらの相対的な明暗関係がある条件を満たすとき、その矩形の中心を眉間候補とする。

20

【 0 0 3 8 】

連続した画素が眉間候補となるときは、それを取囲む枠の中心候補のみを眉間候補として残す。残った眉間候補を標準パターンと比較してテンプレートマッチング等を行なうことで、上述した手続きで得られた眉間候補のうちから、偽の眉間候補を捨て、真の眉間を抽出する。

【 0 0 3 9 】

以下、本発明の顔検出の手続きについて、さらに詳しく説明する。

【 0 0 4 0 】

(6 分割矩形フィルタ)

図 3 は、上述した 3×2 に 6 分割された矩形フィルタ（以下、「6 分割矩形フィルタ」と呼ぶ）を示す図である。

30

【 0 0 4 1 】

6 分割矩形フィルタは、1) 鼻筋は両目領域よりも明るい、2) 目領域は頬部よりも暗い、という顔の特徴を抽出し、顔の眉間位置を求めるフィルタである。点 (x , y) を中心として、横 i 画素、縦 j 画素 (i , j : 自然数) の矩形の枠を設ける。

【 0 0 4 2 】

図 3 のように、この矩形の枠を、横に 3 等分、縦に 2 等分して、6 個のブロック S 1 ~ S 6 に分割する。

【 0 0 4 3 】

図 4 は、このような 6 分割矩形フィルタを顔画像に当てはめた場合を示す概念図である。図 4 (a) は 6 分割矩形フィルタの形状を示し、図 4 (b) は 6 分割矩形フィルタを顔画像の両目領域および頬部に当てはめた状態を示す。

40

【 0 0 4 4 】

なお、鼻筋の部分が目の領域よりも通常は狭いことを考慮すると、ブロック S 2 および S 5 の横幅 w 2 は、ブロック S 1 , S 3 , S 4 および S 6 の横幅 w 1 よりも狭い方がより望ましい。好ましくは、幅 w 2 は幅 w 1 の半分とすることができる。図 5 は、このような場合の 6 分割矩形フィルタの構成を示す概念図である。

【 0 0 4 5 】

実施の形態 1 では、図 5 に示すような 6 分割矩形フィルタを用いるものとする。

50

【 0 0 4 6 】

また、ブロック S 1、S 2 および S 3 の縦幅 h 1 と、ブロック S 4、S 5 および S 6 の縦幅 h 2 とは、必ずしも同一である必要もない。ただし、以下の説明では、縦幅 h 1 と縦幅 h 2 とは等しいものとして説明する。

【 0 0 4 7 】

図 5 に示す 6 分割矩形フィルタにおいて、それぞれのブロック S i (1 ≤ i ≤ 6) について、画素の輝度の平均値「バー S i」(S i に上付きの “ - ” をつける) を求める。

【 0 0 4 8 】

ブロック S 1 に 1 つの目と眉が存在し、ブロック S 3 に他の目と眉が存在するものとする、以下の関係式 (1) が成り立つ。

【 0 0 4 9 】

【数 1】

$$\begin{aligned} \overline{S_1} < \overline{S_2} \quad \text{and} \quad \overline{S_1} < \overline{S_4} \\ \overline{S_3} < \overline{S_2} \quad \text{and} \quad \overline{S_3} < \overline{S_6} \quad \dots (1) \end{aligned}$$

【 0 0 5 0 】

図 6 は、このような 6 分割矩形フィルタを走査する対象となる画像を示す概念図である。

【 0 0 5 1 】

図 6 に示すとおり、顔画像を検知する対象画像は、横方向に M 画素、縦方向に N 画素の M × N 画素から構成される。原理的には、左上隅の画素 (0 , 0) から横方向および縦方向について順次 1 画素ずつずらせながら、上記 6 分割矩形フィルタを当てはめて、上記関係式 (1) の妥当性をチェックする作業を行えばよいことになる。しかしながら、このように 6 分割矩形フィルタをずらせるたびに、各ブロック内の輝度の平均値を求めるのでは、効率が悪い。

【 0 0 5 2 】

そこで、本発明では、矩形枠内の画素の総和を求める処理について、公知の文献 (P. Viola and M. Jones, “ Rapid Object Detection using a Boosted Cascade of Simple Features, ” Proc. Of IEEE Conf. CVPR, 1, pp.511-518, 2001) がにおいて開示されている、インテグラルイメージ (Integral Image) を利用した計算の高速化手法を取り入れる。

【 0 0 5 3 】

画像 i (x , y) から、「インテグラルイメージ」は、次式 (2) で定義される。

【 0 0 5 4 】

【数 2】

$$ii(x, y) = \sum_{x' \leq x, y' \leq y} i(x', y') \quad \dots (2)$$

【 0 0 5 5 】

インテグラルイメージは、以下の繰り返しで求めることができる。

【 0 0 5 6 】

【数 3】

$$\begin{aligned} s(x, y) &= s(x, y - 1) + i(x, y), \\ ii(x, y) &= ii(x - 1, y) + s(x, y) \quad \dots (3) \end{aligned}$$

【 0 0 5 7 】

s (x , y) は行の画素の総和を表わす。ただし s (x , - 1) = 0、ii (- 1 , y) = 0 とする。重要な点は、画像全体を一回走査するだけで、インテグラルイメージを求

10

20

30

40

50

めることができることである。

【0058】

インテグラルイメージを用いると、長方形領域内の画素の輝度値の総和を簡単に求めることができる。図7は、このようなインテグラルイメージを用いて、総和を求める長方形領域を示す図である。

【0059】

インテグラルイメージを用いて、図7に示す長方形Dの枠内の画素の輝度の総和 S_r は、以下のように4点の値の計算で求めることができる。

【0060】

【数4】

$$S_r = (ii(x, y) + ii(x - W, y - L)) - (ii(x - W, y) + ii(x, y - L)) \quad \dots (4)$$

【0061】

このように、インテグラルイメージを用いることによって、長方形領域内の画素の輝度値の総和、ひいては、画素の輝度値の平均を高速に求めることができるので、高速に6分割矩形フィルタに対する処理を行なうことが可能である。

【0062】

(眉間候補点の抽出処理)

以下では、上述した6分割矩形フィルタを用いて、眉間の候補点を抽出する処理を説明する。

【0063】

図8は、眉間の候補点を抽出する処理を説明するためのフローチャートである。

【0064】

図8を参照して、まず、初期化処理として、変数 m 、 n の値を、 $m = 0$ 、 $n = 0$ とする(ステップS100)。

【0065】

続いて、6分割フィルタの左上コーナーを画像の (m, n) 画素に合わせる(ステップS102)。さらに、ブロック S_i の領域の画素の平均濃度 \bar{S}_i を計算する(ステップS104)。

【0066】

次に、平均濃度 \bar{S}_i の値の大小が、式(1)による眉間候補条件を満たすかどうかテストする(ステップS106)。

【0067】

テスト条件を満たす場合は(ステップS108)、フィルタの中心点に相当する $(m + i/2, n + j/2)$ の位置の画素に眉間候補マークをつける(ステップS110)。一方、テスト条件を満たさない場合は(ステップS108)、処理はステップS112に移行する。

【0068】

ステップS112では、変数 m の値が1だけインクリメントされる。次に、変数 m の値が対象画像の中で横方向にフィルタが動ける範囲内であるかが判定される(ステップS114)。フィルタが動ける範囲内であるときは、処理はステップS102に復帰する。一方、フィルタが横方向に動ける限界になっているときは、変数 m の値を0にリセットし、変数 n の値を1だけインクリメントする(ステップS116)。

【0069】

次に、変数 n の値が対象画像の中で縦方向にフィルタが動ける範囲内であるかが判定される(ステップS118)。フィルタが動ける範囲内であるときは、処理はステップS102に復帰する。一方、フィルタが縦方向に動ける限界になっているときは、眉間候補マークのついて、画素の連結性を調べ、各連結要素ごとに連結要素の外形枠の中央の画素を眉間候補点とする(ステップS120)。ここで、「中央の画素」とは、特に限定されな

10

20

30

40

50

いが、たとえば、各連結要素の重心位置とすることができる。

【0070】

図9は、以上のような処理による眉間候補点の抽出結果を示す図である。

【0071】

図9(a)は、適用した6分割矩形フィルタの形状および大きさを示し、図9(b)は、眉間候補マークのついた連結要素をハッチングした領域として示す。

【0072】

なお、与えられた対象画像に対して、どのような大きさの6分割矩形フィルタを適用するかについては、たとえば、予め対象画像中の顔画像の大きさが分かっている場合は、その大きさに合わせて設定しておくことも可能である。あるいは、撮影対象となる範囲内(カメラ30からの距離)に人物が存在する場合の顔の大きさに対応して、予め幾種類かの大きさの6分割矩形フィルタを準備しておき、一番最初に顔を検出する際には、この複数種類の6分割矩形フィルタのうちから、順次違う大きさのものを選択して適用して、以下に説明するような顔検出の適合度が最も高いものを選ぶこととしてもよい。

【0073】

(目の候補点の抽出および真の眉間候補点の抽出)

以上のようにして抽出された眉間候補点には、真の眉間候補点以外に偽の眉間候補点も含まれる。そこで、以下に説明する手順で、真の眉間候補点を抽出する。

【0074】

まず、眉間候補点の情報に基づいて、目の位置の候補点を抽出する。

【0075】

そのために、複数の目の画像を顔画像データベースから抽出し、その平均画像を得る。

【0076】

図10は、このようにして得られた右目のテンプレートを示す図である。左目のテンプレートは、この右目テンプレートを水平方向に反転させればよい。

【0077】

この右目テンプレートおよび左目のテンプレートを用いて、図3に示した眉間候補点を中心とする6分割矩形フィルタのブロックS1およびS3の領域において、テンプレートマッチング処理を行えば、右目および左目の各々の候補点を抽出できる。

【0078】

図11は、このような目の候補点の抽出を行なった上で、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

【0079】

図11を参照して、まず、眉間候補抽出フィルタのブロックS1とS3のそれぞれの領域において、目のテンプレートともっとも良くマッチする点を探索し、左右の目の候補点とする(ステップS200)。

【0080】

次に、眉間候補点位置を左右の目の候補点の中心に修正する(ステップS202)。続いて、修正眉間候補点位置を中心に左右の目の候補点が水平に並ぶように入力画像を回転する(ステップS204)。

【0081】

回転後の修正眉間候補点を中心とするパターンと、後に説明するような手続きで予め形成されている眉間テンプレートとの類似度を計算する(ステップS206)。

【0082】

類似度が予め定めたしきい値以上かを判断し(ステップS208)、しきい値以上であれば、それを真の眉間候補点とする(ステップS210)。一方、しきい値未満であれば、それを偽の眉間候補点とする(ステップS212)。

【0083】

このような処理をすべての眉間候補点について行なう。

【0084】

10

20

30

40

50

図12は、図11のステップS200における目の候補点の抽出処理を説明するための図である。

【0085】

図12において、白丸は、修正前の眉間の候補点であり、白十字は、目の候補点を示す。

【0086】

(眉間テンプレート)

次に、図11のステップS206において使用する眉間テンプレートの形成方法について説明する。

【0087】

図13は、眉間テンプレートの形成手順を説明するためのフローチャートである。

【0088】

図13を参照して、まず、複数の顔画像データを準備する(ステップS300)。続いて、各顔画像について両目の位置をマウス等により、操作者が入力する(ステップS302)。

【0089】

さらに、以下は、計算機内部での処理として、両目の位置が水平となるように、両目の中点を中心に画像を回転して、向きの正規化を行なう(ステップS304)。両目の間隔が所定の距離となるように画像を拡大あるいは縮小して、サイズの正規化を行なう(ステップS306)。次に、両目の中点を中心とする眉間パターン $i \times j$ 画素を抽出する(ステップS308)。

【0090】

さらに、抽出した眉間パターンの平均濃度が所定の値、たとえばゼロとなるように、分散が他の所定の値、たとえば1.0になるように濃度を変換して、濃度の正規化を行なう(ステップS310)。

【0091】

正規化した多数の眉間パターンの平均パターンを計算して(ステップS312)、得られた平均パターンを眉間のテンプレートとする(ステップS314)。

【0092】

ただし、本発明では、ステップS314で得られた眉間テンプレートをさらに以下のよう

【0093】

すなわち、髪の毛が眉までかかっている人の場合は、額部が低い輝度値になるが、平均テンプレートは高い濃度値になっており、このままマッチング評価を行なうと、マッチング度が低くなってしまふ。そこで、髪型の影響を受けないように上から所定の画素数、たとえば、3画素の額にあたる部分は評価しない。たとえば、ステップS314で得られた眉間テンプレートが 32×16 画素のパターンであるならば、結局、 32×13 画素のパターンを用いてテンプレートマッチングを行なう。

【0094】

図14は、眉間テンプレートを説明するための図である。

【0095】

図14(a)は、図13のステップS314で得られた眉間テンプレートを示し、図14(b)は、額の影響を排除するための最終的な眉間テンプレートを示す。

【0096】

なお、テンプレートマッチングは、顔の向きによって照明のあたり方が異なる場合を考慮して、左右独立に評価を行なうことも可能である。このときは、上述した眉間テンプレートを左右に2分割して、それぞれテンプレートマッチングを行なえばよい。たとえば、上記例のような大きさの眉間テンプレートであれば、左右それぞれ、片側 16×13 画素のパターンを用いてテンプレートマッチングを行なってもよい。

【0097】

10

20

30

40

50

次に、図11のステップS206のテンプレートマッチングの処理をさらに詳しく説明する。

【0098】

図15は、ステップS206のテンプレートマッチングの手続きを説明するためのフローチャートである。

【0099】

図15を参照して、まず、眉間候補点を抽出して(ステップS400)、必要に応じて、眉間候補点を中心に回転を行ない、スケール補正を行なう(ステップS402)。

【0100】

次に、眉間候補点を中心として、テンプレートと同じサイズの画像を切り出す(ステップS404)。切り出した眉間候補パターンと眉間テンプレートとの相関値を計算して類似度とする(ステップS406)。

10

【0101】

なお、類似度の計算としては、切り出した眉間候補パターンの濃度を正規化(平均ゼロ、分散1.0)して、画素ごとにテンプレートの対応画素との差の2乗を計算し、その総和を求めることとしてもよい。すなわち、この場合、総和の値は、不類似度とみなせるので、この逆数により類似度を評価してもよい。

【0102】

図16は、このようにして対象画像から眉間および目の位置を抽出した例を示す図である。

20

【0103】

帽子をかぶり、かつ手で口を覆うという状態であるにも関わらず、眉間の位置(図中長方形の枠の中心)と目の位置(十字)が良好に検出されている。

【0104】

実施の形態1の本発明においては、濃淡情報を用いて6分割矩形フィルタにより、まず、眉間の候補点を抽出してから、最終的に目の位置を特定しているため、照明条件の変化に強く、かつ、高速な顔位置の抽出を行なうことができる。

【0105】

さらに、以上のような処理を、撮影されたビデオ画像の各フレームについて行なえば、動画像において、顔画像を追跡することも可能となる。

30

【0106】

このときは、既に顔画像が検出されている前フレームの情報に基づいて、原フレームにおいてフィルタ処理をする領域を絞り込むことも可能である。

【0107】

なお、以上の説明では、眉間の候補点を探索する際に用いるフィルタは、矩形形状を3×2に6分割した6分割矩形フィルタを用いることとした。

【0108】

ただし、顔画像が水平から傾いている場合にも対応可能とするためには、フィルタの形状は、図3や図5に示したものに限定されない。

【0109】

図17および図18は、このようなフィルタの他の形状を説明するための図である。

40

【0110】

すなわち、図17や図18に示すように、図1におけるブロックS2およびS5に対して、ブロックS1およびS4と、ブロックS3およびS5とを、互いに反対方向に上下に所定の量だけずらせることも可能である。

【0111】

この場合、ずれた量に対応する角度だけ、顔画像が傾いている場合にも良好に眉間の候補点を抽出できる。

【0112】

本明細書中では、図3および図5に示した形状のフィルタ(6分割矩形フィルタ)と、

50

図17や図18に示したようなフィルタとを総称して、「眉間検出フィルタ」と呼ぶことにする。

【0113】

[実施の形態2]

実施の形態1の図11において説明したとおり、眉間候補点から真の候補点を抽出する際には、一般には、眉間候補点の位置の修正および入力画像の回転等を行なう必要がある。ただし、テレビ会議のように画像中の人物の動きが比較的小さい場合には、真の候補点の抽出処理を簡略化することも可能である。

【0114】

図19は、このような実施の形態2の顔位置抽出装置において、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

10

【0115】

図19を参照して、まず、入力画像における眉間候補点を中心とするパターンと、予め形成されている眉間テンプレートとの類似度を計算する(ステップS500)。

【0116】

類似度が予め定めたしきい値以上かを判断し(ステップS502)、しきい値以上であれば、それを真の眉間候補点とする(ステップS504)。一方、しきい値未満であれば、それを偽の眉間候補点とする(ステップS506)。

【0117】

このような処理をすべての眉間候補点について行なう。

20

【0118】

その他の処理および構成は、実施の形態1の顔位置抽出装置と同様であるので、その説明は繰り返さない。

【0119】

このような構成でも、人物のカメラ30からの位置や顔の向きの変化が小さい場合は、実施の形態1と同様の効果を奏することができる。

【0120】

[実施の形態3]

実施の形態1および2では、カメラ30は1台で撮影を行なうこととしていた。

【0121】

これに対して、カメラ30をたとえば2台として、2眼ステレオ構成とすると、人物についての距離の情報も得ることができる。

30

【0122】

すなわち、6分割矩形フィルタで抽出した候補点の中から、真の顔候補点を検出する手法は、実施の形態3でも、原理的には、実施の形態1または2と同様の方法を用いることができる。

【0123】

ただし、実施の形態3の顔位置抽出装置では、検出できる顔の大きさの範囲をより広げるため、カメラ30を2眼ステレオ構成とし、距離情報に応じて顔候補領域を切り出すサイズを切り替える。

40

【0124】

顔候補領域の切り出すサイズを切り替えることで、平均顔テンプレートと同じ顔の大きさにスケールリングしてマッチングをとることができ、顔の大きさの検出範囲を広げることが可能である。

【0125】

実施の形態3では、上述のとおり、2眼ステレオ構成とし、候補点の視差情報を求める。顔の大きさは視差に反比例すると考えられるので、視差情報からその候補点を切り出すサイズを決定する。そのため、最適なサイズで顔候補領域を切り出してテンプレートとのマッチングを採ることができる。

【0126】

50

ここで、以下では、40人、各人10枚、計400枚の表情や向き照明条件などを少しずつ変化させた画像が納められている顔画像データベースに対する評価を交えて説明する。

【0127】

このデータベース内の顔画像は、画像サイズが 92×112 のモノクロ画像である。矩形サイズは、横が左右こめかみの間の画素数、縦が眉毛から鼻先までの画素数の大きさを基準とした。手動で計測した結果、顔画像(92×112)に対して、基準の矩形サイズは 60×30 とした。

【0128】

まず、図20は、異なるサイズの6分割矩形フィルタにより、同一の顔画像について、どの範囲で眉間候補点が検出可能であることを示す図である。

10

【0129】

図20では、矩形サイズを基準サイズから20%ずつ変化させながら、眉間抽出処理を行なっている。実験では、真の候補点抽出率と候補点の個数を調べた。候補点に真の候補点が含まれているかは、眉間付近に候補点が存在するかを目視で判断した。

【0130】

図20より、基準の矩形サイズ(60×30)での抽出率は、92.0%であり、有効に機能していると考えられる。一方、矩形サイズが 84×42 の場合には抽出率が非常に悪くなっており、矩形が大きすぎて顔の特徴を抽出できていないと考えられる。

【0131】

図20を参照すると、基準の矩形サイズから、0.6~1.2倍のサイズの矩形で眉間候補点の抽出ができることを確認できる。顔の大きさと矩形サイズとは単純な比例関係にあると考えられる。したがって、矩形フィルタは、基準の大きさの顔から、0.83~1.67倍の範囲の大きさの顔の眉間候補点が抽出できると考えられる。

20

【0132】

次に、人物の距離と切り出す顔候補領域の大きさの関係を求めるために、顔位置抽出装置で使用しているカメラ構成で人物の顔を撮影し、カメラ30との距離を変えながら、眉間位置の視差と、その顔に最適な顔を切り出すサイズを計測しておく。

【0133】

たとえば、視差は左右のカメラ30に写る人物の眉間の位置の横方向の画素数の差を手動で計測することにより得る。顔を切り出すサイズは、左右のこめかみの間の画素数を手動で測定する。特に限定されないが、6分割矩形フィルタの縦方向のサイズは横方向の半分と定めることができる。

30

【0134】

図21は、視差と最適な顔を切り出すサイズの関係を示す図である。

【0135】

この図21を基に、6分割矩形フィルタのサイズ、顔候補点を切り出すサイズ、視差と顔候補点を切り出すサイズの関係を決定する。

【0136】

図22は、図21より設定した6分割矩形フィルタサイズ、視差、候補点を切り出すサイズの関係を示す図である。あるサイズの6分割矩形フィルタが抽出できる顔候補領域を切り出すサイズが0.83~1.67倍の範囲を持つことを利用し、たとえば、 40×20 、 24×12 の2段階のフィルタサイズで全体をカバーできるように設定した。顔候補領域を切り出すサイズは視差5画素ごとに切り替えるように設定した。切り出すサイズは細かく設定する方が精度が高くなると考えられるが、平均顔テンプレートのマッチング処理はある程度の大きさに対する柔軟性があるため、この範囲での切り替えで十分である。図22では、例えば、矩形フィルタサイズが 40×20 のとき、ステレオマッチングの結果、視差が20であれば、 48×24 の大きさで候補点を切り出すという意味である。

40

【0137】

もしも、この表に当てはまらない視差が出てきた場合、または、どこにもマッチングし

50

なかった場合、その候補点は偽の候補点であるとして切り捨てる。

【0138】

以上の処理により、実施の形態3の顔位置抽出装置において、対象となる画像から眉間の候補点を抽出することができる。

【0139】

図23は、実施の形態3の顔位置抽出装置において、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

【0140】

図23を参照して、まず、候補点のカメラ30からの距離を2眼ステレオ方式により推定する(ステップS600)。

【0141】

次に、距離が予め定めた範囲内にあるかを判断する(ステップS602)。もしも、予め定めた範囲内に眉間の候補点がないならば、それは偽の候補点であると判断する(ステップS612)。

【0142】

一方、距離が予め定めた範囲内にある場合は、次に、距離に応じて、予め用意しておいたサイズの異なる眉間テンプレートを選択する(ステップS604)。

【0143】

入力画像における眉間候補点を中心とするパターンと、選択された眉間テンプレートとの類似度を計算する(ステップS606)。

【0144】

類似度が予め定めたしきい値以上かを判断し(ステップS608)、しきい値以上であれば、それを真の眉間候補点とする(ステップS610)。一方、しきい値未満であれば、それを偽の眉間候補点とする(ステップS612)。

【0145】

このような処理をすべての眉間候補点について行なう。

【0146】

その他の処理および構成は、実施の形態1の顔位置抽出装置と同様であるので、その説明は繰り返さない。

【0147】

このような構成では、人物のカメラ30からの距離も考慮して、真の候補点の抽出を行なうので、より高速に顔画像の位置検出を行なうことが可能である。したがって、この実施の形態3の処理を動画の各フレームに対して行なうことで、顔画像の追跡を行なうことも可能である。

【0148】

なお、実施の形態3においても、実施の形態1の図11において説明したとおり、眉間候補点から真の候補点を抽出する際に、目の位置を検出した上で眉間候補点の位置の修正および入力画像の回転等を行なうことも可能である。

【0149】

[実施の形態4]

実施の形態3では、予め準備しておいたサイズの異なる眉間テンプレートから眉間候補点のカメラ30からの距離に応じて、眉間テンプレートを選択した。

【0150】

しかしながら、眉間候補点のカメラ30からの距離に応じて、基準となる眉間テンプレートのサイズに合うように入力画像を縮小(または拡大)して、テンプレートマッチングを行なうことも可能である。

【0151】

図24は、このような実施の形態4の顔位置抽出装置において、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

【0152】

10

20

30

40

50

図24を参照して、まず、候補点のカメラ30からの距離を2眼ステレオ方式により推定する(ステップS700)。

【0153】

次に、距離が予め定めた範囲内にあるかを判断する(ステップS702)。もしも、予め定めた範囲内に眉間の候補点がないならば、それは偽の候補点であると判断する(ステップS712)。

【0154】

一方、距離が予め定めた範囲内にある場合は、次に、距離に応じて、眉間像がテンプレートサイズに合うように入力画像を縮小する(ステップS704)。

【0155】

入力画像の候補点を中心とする縮小パターンと眉間テンプレートとの類似度を計算する(ステップS706)。

【0156】

類似度が予め定めたしきい値以上かを判断し(ステップS708)、しきい値以上であれば、それを真の眉間候補点とする(ステップS710)。一方、しきい値未満であれば、それを偽の眉間候補点とする(ステップS712)。

【0157】

このような処理をすべての眉間候補点について行なう。

【0158】

その他の処理および構成は、実施の形態3の顔位置抽出装置と同様であるので、その説明は繰り返さない。

【0159】

このような構成では、人物のカメラ30からの距離も考慮して、真の候補点の抽出を行なうので、より高速に顔画像の位置検出を行なうことが可能である。したがって、この実施の形態4の処理を動画の各フレームに対して行なうことで、顔画像の追跡を行なうことも可能である。

【0160】

なお、実施の形態4においても、実施の形態1の図11において説明したとおり、眉間候補点から真の候補点を抽出する際に、目の位置を検出した上で眉間候補点の位置の修正および入力画像の回転等を行なうことも可能である。

【0161】

以上説明したような各実施の形態の処理で、時間軸について所定間隔で連続する画面情報、たとえば、連続するフレーム画像から、実時間で、眉間または目の位置を検出することができる。さらに、このような連続する画面情報の各々において、眉間または目の位置の検出を連続して行なっていくことで、眉間または目の位置のトラッキングを行なうことができる。

【0162】

[眉間の候補点の中から真の眉間を選択する処理の変形例]

以上説明した実施の形態では、顔位置の抽出処理において、眉間検出フィルタで画像中から眉間の候補点を抽出し、候補点の中から真の眉間を選択する、という処理を行なっている。

【0163】

この「真の眉間を検出する処理」は、言い換えると、複数の眉間候補点から、真の眉間に相当する候補点を選択するためのパターン判別処理を行なっていることに相当する。上述した実施の形態では、「眉間テンプレートとの類似度」に基づいて、パターン判別処理を行なうものとしたが、パターン判別の方法としては、必ずしもこのような方法に限定されるものではない。

【0164】

以下では、「眉間テンプレートとの類似度」によるパターン判別処理も含めて、このようなパターン判別処理として可能な変形例について説明する。

10

20

30

40

50

【 0 1 6 5 】

(1) パターンテンプレートとの類似度によるパターン判別処理

テンプレートを $f = \{ t_{ij} \}$ 、評価されるパターンを $f = \{ f_{ij} \}$ とすると、単純な類似度評価値 (q) としては、以下の式 (5) のような各対応画素値の差の絶対値の総和がある。

【 0 1 6 6 】

【数 5】

$$q = \sum |t_{ij} - f_{ij}| \quad \dots (5)$$

10

【 0 1 6 7 】

あるいは、以下の式 (6) のような差の絶対値の 2 乗和を用いることもできる。

【 0 1 6 8 】

【数 6】

$$q = \sum (t_{ij} - f_{ij})^2 \quad \dots (6)$$

【 0 1 6 9 】

式 (5) または (6) を用いる場合は値が小さいほど類似度が高いと判断することになる。

20

【 0 1 7 0 】

一方、他の評価値としては次式 (7) で表わされる正規化相関値を用いることもできる。

【 0 1 7 1 】

【数 7】

$$q = \frac{\sum (t_{ij} - \bar{t})(f_{ij} - \bar{f})}{\sqrt{\sum (t_{ij} - \bar{t})^2} \sqrt{\sum (f_{ij} - \bar{f})^2}} \quad \dots (7)$$

30

ここで、 \bar{t} 、 \bar{f} は、それぞれ $\{t_{ij}\}$ 、 $\{f_{ij}\}$ の平均値である。

【 0 1 7 2 】

式 (7) において、 $\{t_{ij}\}$ と、 $\{f_{ij}\}$ とが完全に一致していれば、 q の値は 1 であり、完全に反転パターン (明暗が逆) ならば q の値は - 1 になる。それ以外の場合は、 q の値は 1 と - 1 の間の値となる。式 (7) を用いるときは、 q の値が大きいほど類似度は高いという評価になる。

【 0 1 7 3 】

正規化相関値は、平均値からの差で評価しているので、全体的に明るさレベルがシフトしてもその評価に影響がない。また、例えば照明が暗くなると明るさの平均値が下がるだけでなく、明暗の差も小さくなる。この場合でも、分母の正規化項のおかげで q の値に影響がない。

40

【 0 1 7 4 】

また、以下の式 (8) で示されるように、テンプレートととして多くのサンプルパターン ($s^n = \{ s^n_{ij} \}$) の平均パターンを使うこともできる。

【 0 1 7 5 】

【数 8】

$$t_{ij} = \frac{\sum s_{ij}^n}{N} \quad \dots (8)$$

【0176】

この場合は、重み付きの類似度評価を行なうことができる。例えば、右目の右上部分や左目の左上部分は、人によっては前髪が下がっていたりいなかったりする。このため、その部分は、テンプレートと差があっても、あまり重要でないと考えられる。

【0177】

そこで、多くのサンプルパターンがある場合は、以下の式(9)に示すように、各画素位置で明るさがどれくらいサンプル間でばらついているかを示す分散をまず計算する。

【0178】

【数 9】

$$v_{ij} = \frac{\sum (s_{ij}^n - \bar{s}_{ij})^2}{N} \quad \dots (9)$$

【0179】

次に、その分散の逆数を重みづけに使用して、以下の式(10)に示すような評価値 q を用いた重み付き類似度評価を行なうこともできる。

【0180】

【数 10】

$$q = \sum \frac{(t_{ij} - f_{ij})^2}{v_{ij}} \quad \dots (10)$$

【0181】

あるいは、「右目位置に対象な位置には左目があって同じように黒いはず」であり、「その中央は鼻筋で明るいはず」というように、互いの画素間にも関係があって、その関係がどれくらいばらついているかを表わす指標である共分散を考慮して重み付けを行なうことができる。なお、これに対して、式(9)は、自己分散の場合である。

【0182】

このような共分散を考慮して重みづけをした類似度は、「マハラノビス距離」と呼ばれる。

【0183】

すなわち、 t_{ij} を 1 列にならべてベクトルのように表わすとすると、マハラノビス距離 q は、以下の式(11)のように表わされる。

【0184】

【数 11】

$$q = (t - f)^T \Sigma^{-1} (t - f) \quad \dots (11)$$

【0185】

ここで s^n の共分散行列である。このマハラノビス距離 q を用いても、パターンテンプレートとの類似度によるパターン判別処理を実施することができる。

【0186】

(2) 統計的パターン判別処理

10

20

30

40

50

眉間検出フィルタで画像中から眉間の候補点を抽出し、候補点の中から真の眉間を選択する、という処理は、言い換えれば、眉間の候補点の中から、顔のパターンに対応するか、あるいは、顔パターンではないのかを判別することで、真の眉間を抽出する、との手続きとみることにもできる。

【0187】

この場合、「顔」と「非顔」の判別処理には、統計的パターン判別処理を適用することができる。

【0188】

すなわち、統計的パターン判別処理は、多数の「顔」と「非顔」のサンプルが与えられたときに、それらのデータを元に「不明」のパターンを「顔」か「非顔」に判別するものである。これに対して、上述した類似度計算では「非顔」という概念は、必要ない。

10

【0189】

(2-1) 線形判別法

パターン $f = \{f_{ij}\}$ を、その画素値を一列にならべた $I \times J$ 次元のベクトルと考えると、1パターンは $I \times J$ 次元空間の1点と考えられる。

【0190】

以下の説明では、3次元以上は平面上に図示しにくいので、2次元の場合を例にとって説明する。

【0191】

図25は、「顔」のサンプルと「非顔」のサンプルの分布の一例を示す概念図である。

20

【0192】

図25に示すように、「顔」のサンプル(○)と「非顔」のサンプル(×)が分布していたとすると、「顔」(○)と「非顔」(×)を分離する直線 L_1 を予め求めておき、「不明」の入力パターンが直線 L_1 のどちらにあるかで、「顔」(○)と「非顔」(×)かを判定することができる。

【0193】

2次元では直線 $ax + by$ になるが、3次元では $ax + by + cz$ で表現される平面になる。より一般に、さらに高次元では各次元要素の線形結合で表わされる超平面となる。このような超平面による判別を、「線形判別法」と呼ぶ。

【0194】

一般には、一つの超平面で完全に「顔」(○)と「非顔」(×)を分離できるとはかぎらないものの、「顔」(○)の側に「非顔」(×)がくる誤りと、「非顔」(×)の側に「顔」(○)がくる誤りの合計が最小になるように超平面を決定しておく。

30

【0195】

(2-2) サポートベクターマシン

上述した線形判別法で誤りが最小になるように超平面を決定しても、実用上は、誤りが大きすぎる場合もあり得る。

【0196】

そのようなときであっても、例えば (x, y, z) の3次元の空間の点を $(x^2, y^2, z^2, xy, yz, zx)$ のようなより高次元(この場合6次元)の空間に写像してやると、その空間の超平面でうまく、上述したような「顔」(○)と「非顔」(×)とが分離できるようになる場合があることが知られている。しかも、サポートベクターマシンでは、実際には高次元の空間に写像することなく、もとの空間で写像先の高次元空間の超平面を計算することができる。

40

【0197】

サポートベクターマシンで顔の検出を行なう具体的な構成については、たとえば、文献: E.Osuna, R.Freund, and F.Girosi: "Training Support Vector Machines: an Application to Face Detection", Proc. of International Conference on Computer Vision and Pattern Recognition, pp.130-136(1997)に開示されている。

【0198】

50

以下では、サポートベクターマシンの概要について説明する。

【0199】

図26は、サポートベクターマシンが適用される写像先の高次元空間を示す図である。

【0200】

図26でも、高次元空間を2次元空間として説明している。

【0201】

サポートベクターマシンでは平行な超平面が2つ想定される。この2つの超平面は、1つは「非顔」(図では×)のサンプルに接する超平面P1であり、もう1つは「顔」(図では○)のサンプルに接する超平面P2のようなペアである。

【0202】

他のペアの超平面P3および超平面P4も考えられる。しかし、サポートベクターマシンでは、可能な超平面のペアの中で間隔が最大となるペアが採用される。この間隔が、判別の際の余裕と考えられ、余裕が最大となるようなペアが採用されることになる。

【0203】

図26に示すような超平面による、「顔」パターンと「非顔」パターンの判別は、超平面P1と超平面P2から等距離にある中間の超平面を、上述した線形判別におけるの判定のための超平面のようにみなして行なう。

【0204】

(2-3) ベイズ推定による判別

排反事象 H_1 (顔である)と H_2 (非顔である)があって、 A を任意の事象(切り出した濃淡パターン)としたとき、ベイズの定理は、以下の式で表わされる。

【0205】

【数12】

$$P(H_1 | A) = \frac{P(A | H_1)P(H_1)}{P(A | H_1)P(H_1) + P(A | H_2)P(H_2)} \quad \dots (12)$$

【0206】

ここで、 $P(H_1 | A)$ は A が生じた時にそれが H_1 である事後確率で、 $P(A | H_1)$ は、 H_1 の時に A が生じる事前確率である。 A が生じたとわかったあとで、それが H_1 である事後確率または H_2 である事後確率の両者を比較して、ベイズ判定では確率の高い方のパターンであると判定を行なう。ふたつの事後確率の比は、以下の式で表わされる。

【0207】

【数13】

$$\frac{P(H_1 | A)}{P(H_2 | A)} = \frac{P(A | H_1)P(H_1)}{P(A | H_2)P(H_2)} \quad \dots (13)$$

【0208】

式(9)が1より大きければ方1と判断することになる。式(9)は書き直せば、以下の式(10)となる。

【0209】

【数14】

$$\frac{P(A | H_1)}{P(A | H_2)} > \frac{P(H_2)}{P(H_1)} = \lambda \quad \dots (14)$$

【0210】

10

20

30

40

50

そこで、事象 H_1 と H_2 のサンプルをたくさん収集して、 $P(A|H_1)$ と $P(A|H_2)$ を推定しておき、 θ をしきい値パラメータとして、式(10)により判定すれば、事象 A を事象 H_1 と判断するか事象 H_2 と判断するかを決めることができる。

【0211】

ベイズ判定方法で顔を検出する方法については、たとえば、文献：H.Shneiderman and T.Kanade:"Probabilistic Modeling Of Local Appearance and Spatial Relationships for Object Recognition", Proc. of International Conference on Computer Vision and Pattern Recognition, pp.45-51(1998)に開示されている。

【0212】

この他、ニューラルネットワークによる判別により、「顔」と「非顔」の判別処理を行なうことも可能である。

10

【0213】

今回開示された実施の形態はすべての点で例示であって制限的なものではないと考えられるべきである。本発明の範囲は上記した説明ではなくて特許請求の範囲によって示され、特許請求の範囲と均等の意味および範囲内でのすべての変更が含まれることが意図される。

【図面の簡単な説明】

【0214】

【図1】本発明の1実施の形態にかかるシステムの外觀図である。

【図2】本発明の1実施の形態にかかるシステムのハードウェア的構成を示すブロック図である。

20

【図3】6分割矩形フィルタを示す図である。

【図4】6分割矩形フィルタを顔画像に当てはめた場合を示す概念図である。

【図5】6分割矩形フィルタの他の構成を示す概念図である。

【図6】分割矩形フィルタを走査する対象となる画像を示す概念図である。

【図7】インテグラルイメージを用いて、総和を求める長方形領域を示す図である。

【図8】眉間の候補点を抽出する処理を説明するためのフローチャートである。

【図9】眉間候補点の抽出結果を示す図である。

【図10】右目のテンプレートを示す図である。

【図11】目の候補点の抽出を行なった上で、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

30

【図12】図11のステップS200における目の候補点の抽出処理を説明するための図である。

【図13】眉間テンプレートの形成手順を説明するためのフローチャートである。

【図14】眉間テンプレートを説明するための図である。

【図15】ステップS206のテンプレートマッチングの手続きを説明するためのフローチャートである。

【図16】対象画像から眉間および目の位置を抽出した例を示す図である。

【図17】眉間検出フィルタの他の形状を説明するための第1の図である。

【図18】眉間検出フィルタの他の形状を説明するための第2の図である。

40

【図19】実施の形態2の顔位置抽出装置において、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

【図20】異なるサイズの6分割矩形フィルタにより、同一の顔画像について、どの範囲で眉間候補点を検出可能であることを示す図である。

【図21】視差と最適な顔を切り出すサイズの関係を示す図である。

【図22】図21より設定した6分割矩形フィルタサイズ、視差、候補点を切り出すサイズの関係を示す図である。

【図23】実施の形態3の顔位置抽出装置において、真の眉間候補点の抽出を行なう処理を説明するためのフローチャートである。

【図24】実施の形態4の顔位置抽出装置において、真の眉間候補点の抽出を行なう処理

50

を説明するためのフローチャートである。

【図25】「顔」のサンプルと「非顔」のサンプルの分布の一例を示す概念図である。

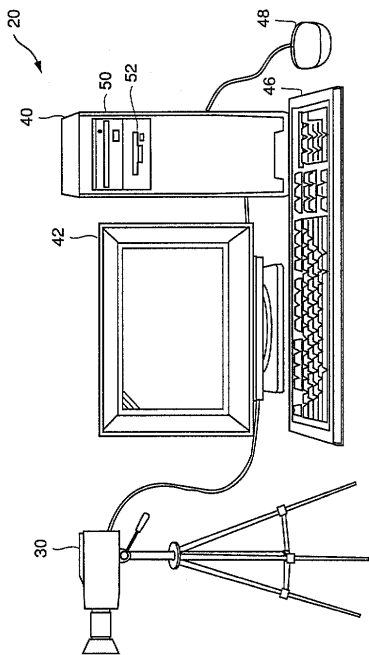
【図26】サポートベクターマシンが適用される写像先の高次元空間を示す図である。

【符号の説明】

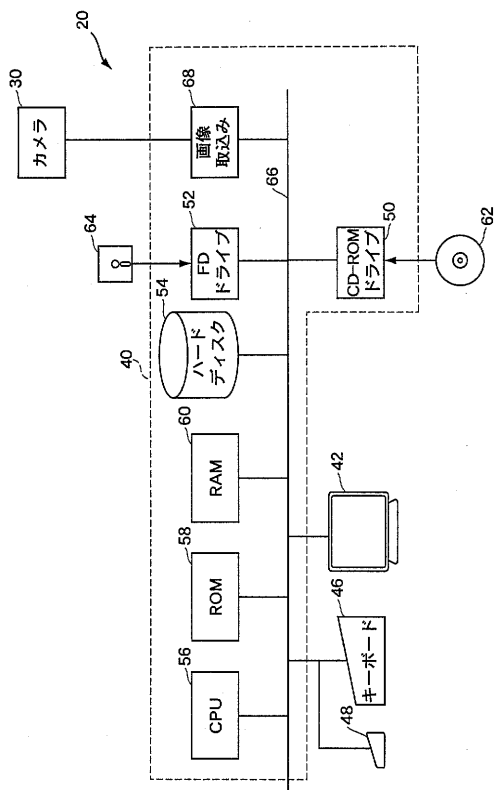
【0215】

20 顔位置抽出装置、30 カメラ、40 コンピュータ本体、42 モニタ。

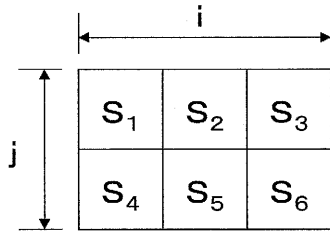
【図1】



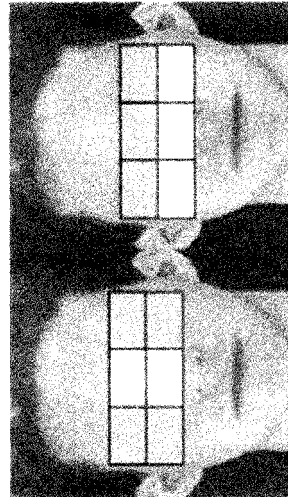
【図2】



【 図 3 】

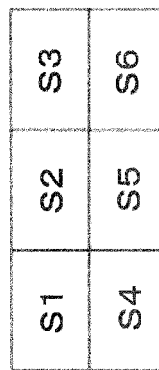


【 図 4 】

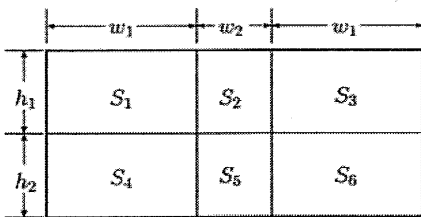


(b)

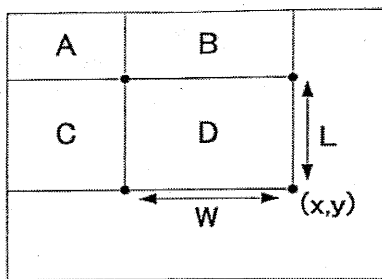
(a)



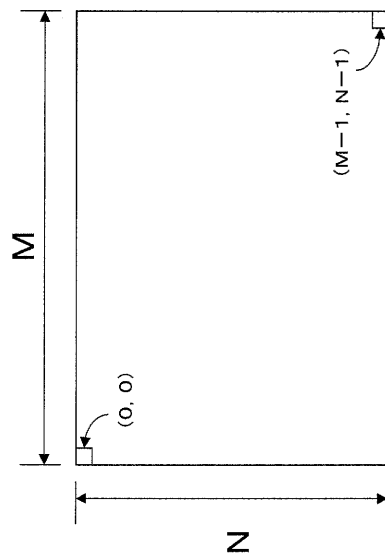
【 図 5 】



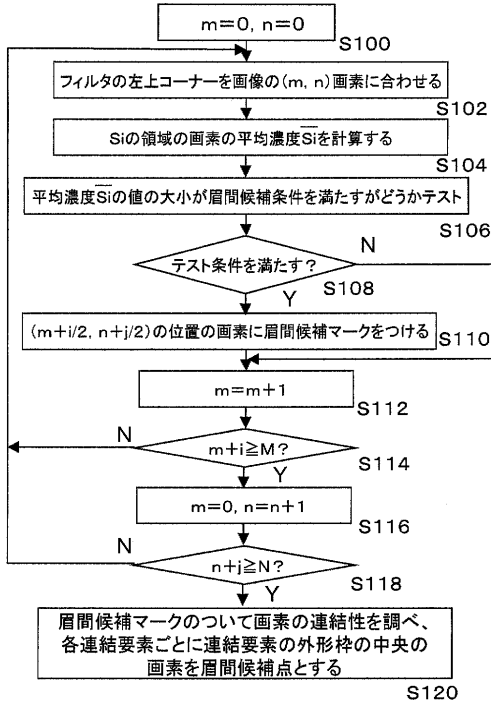
【 図 7 】



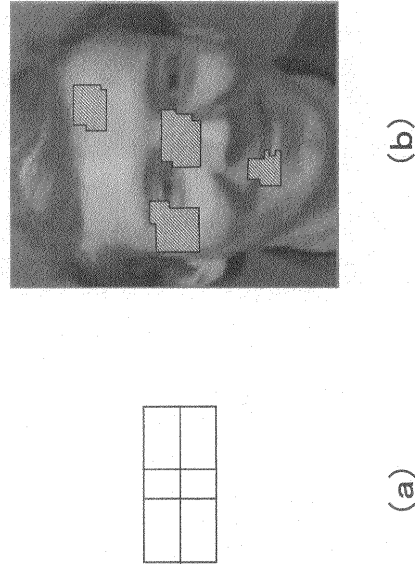
【 図 6 】



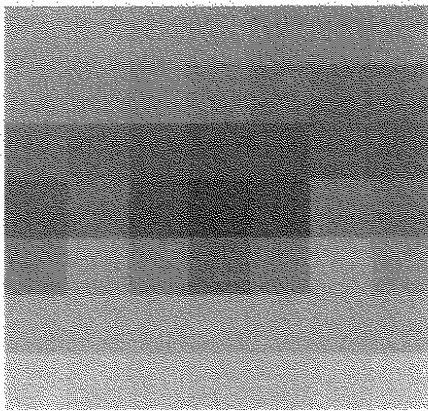
【図8】



【図9】

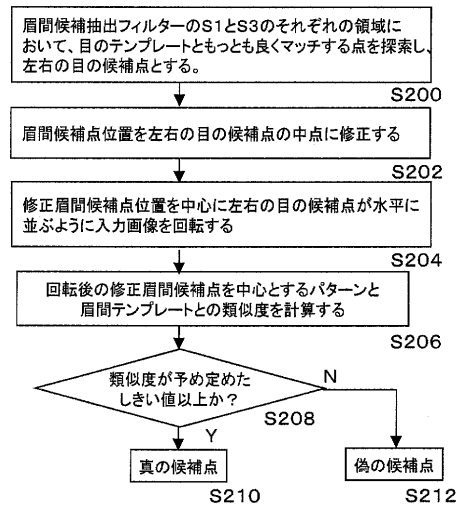


【図10】

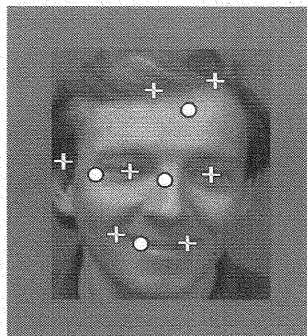


右目テンプレート

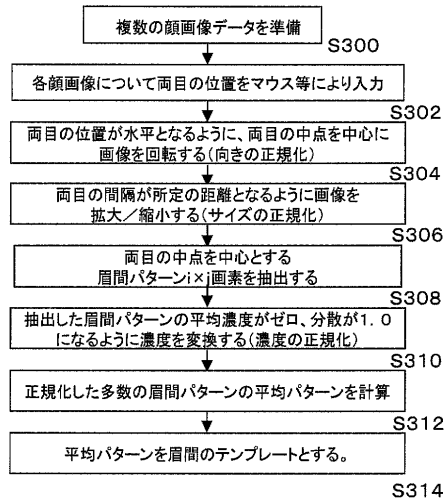
【図11】



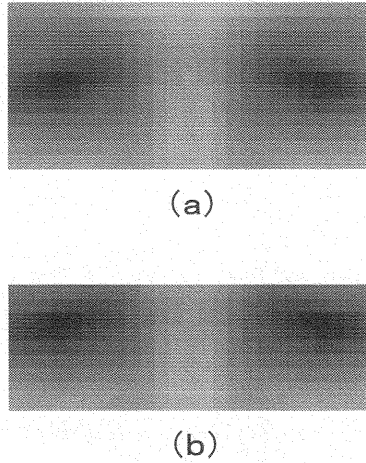
【図12】



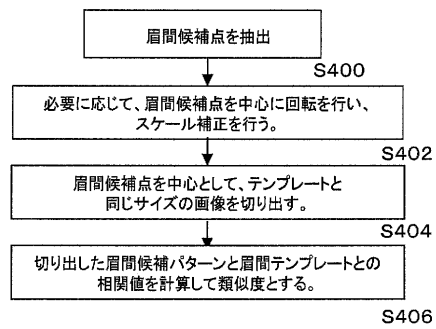
【図13】



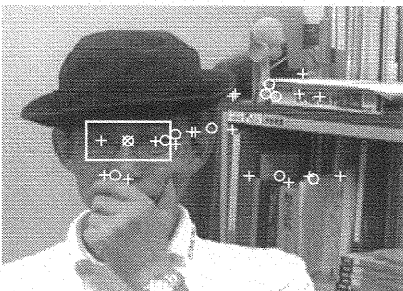
【図14】



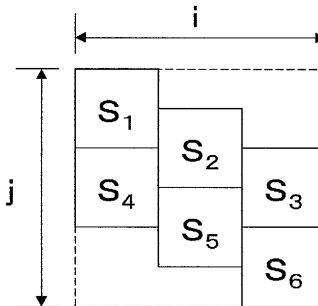
【図15】



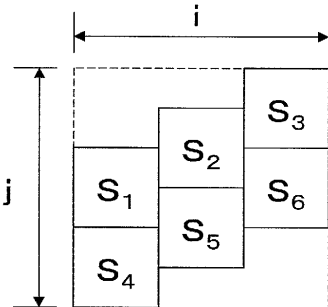
【図16】



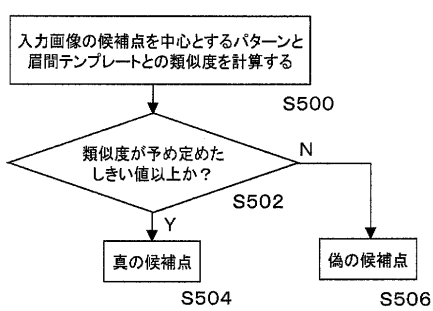
【図18】



【図17】



【図19】

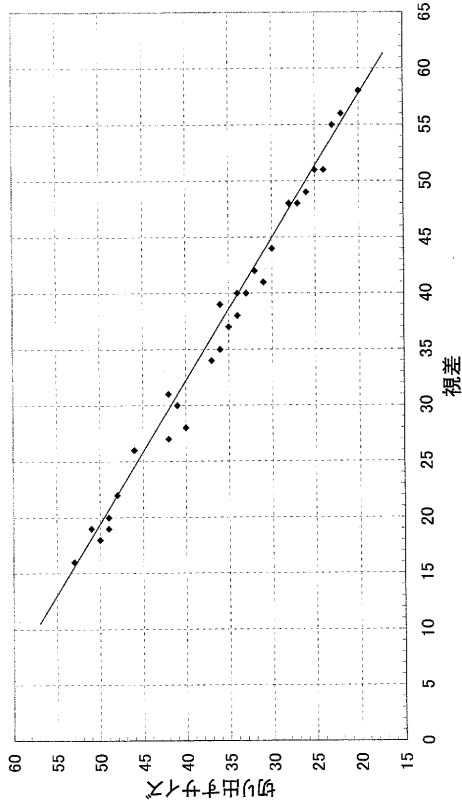


【図 20】

6分割矩形フィルタの性能実験結果 (総数 400)

矩形フィルタサイズ	サイズの倍率	倍率の逆数	失敗数 (個)	候補点数 (個 (平均))	抽出率 (%)
24 x 12	0.3	3.3	13	17	96.8
36 x 18	0.6	1.67	11	8	97.25
48 x 24	0.8	1.25	15	5	96.25
60 x 30 (基準)	1.0	1.0	33	4	92.0
72 x 36	1.2	0.83	42	3	89.5
84 x 42	1.4	0.71	192	3	52.0

【図 21】

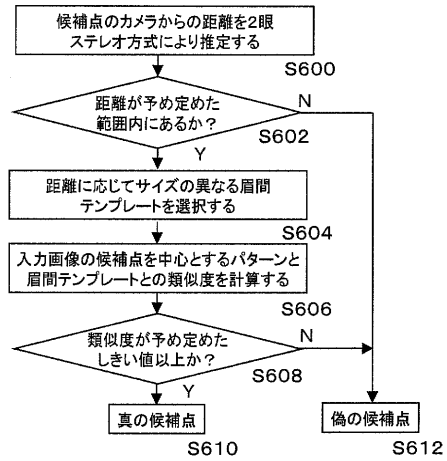


【図 22】

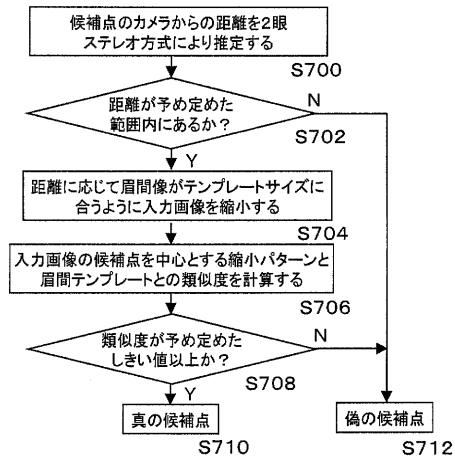
フィルタサイズと視差と切り出しサイズの関係

フィルタサイズ	フィルタサイズ	視差	切り出しサイズ
40 x 20	—	10 ~ 15	56 x 28
—	—	15 ~ 20	52 x 26
—	—	20 ~ 25	48 x 24
—	—	25 ~ 30	44 x 22
—	24 x 12	30 ~ 35	40 x 20
—	—	35 ~ 40	36 x 18
—	—	40 ~ 45	32 x 16
—	—	45 ~ 50	28 x 14
—	—	50 ~ 55	24 x 12
—	—	55 ~ 60	20 x 10

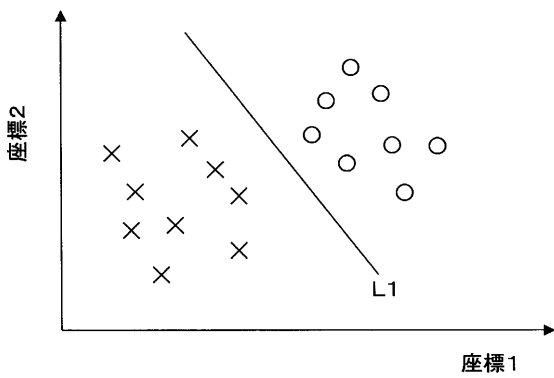
【図 23】



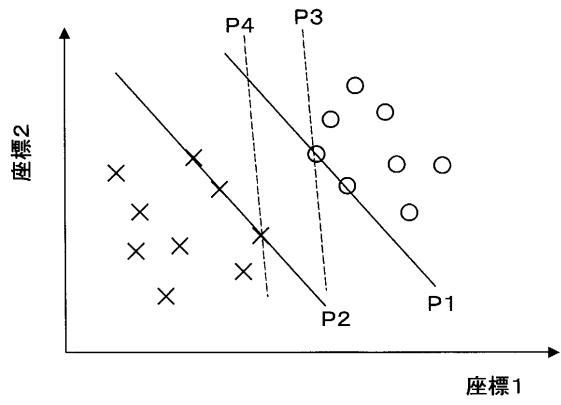
【図 2 4】



【図 2 5】



【図 2 6】



フロントページの続き

(72)発明者 川戸 慎二郎

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

(72)発明者 千田 康隆

京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内

審査官 岡本 俊威

(56)参考文献 特開2000-242768(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06T 1/00

G06T 7/00-7/60