

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4206053号
(P4206053)

(45) 発行日 平成21年1月7日(2009.1.7)

(24) 登録日 平成20年10月24日(2008.10.24)

(51) Int.Cl. F I
G06F 3/033 (2006.01) G O 6 F 3/033 3 1 O A
G06T 1/00 (2006.01) G O 6 T 1/00 3 4 O A

請求項の数 5 (全 21 頁)

<p>(21) 出願番号 特願2004-105250 (P2004-105250) (22) 出願日 平成16年3月31日 (2004. 3. 31) (65) 公開番号 特開2005-293061 (P2005-293061A) (43) 公開日 平成17年10月20日 (2005.10.20) 審査請求日 平成17年11月15日 (2005.11.15)</p> <p>(出願人による申告) 平成15年度通信・放送機構、研究テーマ「軽度脳障害者のための情報セラピーインタフェースの研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願</p>	<p>(73) 特許権者 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2 (74) 代理人 100064746 弁理士 深見 久郎 (74) 代理人 100085132 弁理士 森田 俊雄 (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</p>
---	--

最終頁に続く

(54) 【発明の名称】 ユーザインタフェース装置およびユーザインタフェースプログラム

(57) 【特許請求の範囲】

【請求項1】

ユーザ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獲得する撮影手段と、

前記対象となる画像領域内において、目の位置を抽出する目検出手段と、

前記検出された目の位置に基づき、予め記憶装置に登録された前記目を含む画像についての情報との比較に基づいて、目の開閉および視線方向を含む目の状態を判定する目状態判定手段と、

前記ユーザの口の位置を検出して、前記ユーザの口に対して、予め前記記憶装置に登録された前記口を含む画像についての情報との比較に基づいて、開状態および閉状態を含む口の状態を判定する口状態判定手段と、

検出された前記目の状態および前記口の状態の組み合わせに応じて、複数の所定の処理のうち、予め対応づけられた処理を選択して実行する処理手段とを備え、

前記処理手段は、前記口が前記開状態にあるときのみ、前記目の状態に応じた処理を実行し、

実行された前記処理に対応する結果を、出力する出力手段をさらに備える、ユーザインタフェース装置。

【請求項2】

前記口状態判定手段が判定する前記口の状態は、舌が出ている状態をさらに含み、

前記処理手段は、前記舌が出ている場合にのみ、前記目の状態に応じた処理を実行する

、請求項 1 に記載のユーザインタフェース装置。

【請求項 3】

前記予め記憶装置に登録された前記目を含む画像についての情報は、前記目の間隔で正規化された前記目を含む所定の範囲の画素の明度の平均値を含む、請求項 1 または 2 に記載のユーザインタフェース装置。

【請求項 4】

前記予め記憶装置に登録された前記口を含む画像についての情報は、前記目の間隔で正規化された前記口を含む所定の範囲の画素の明度の平均値を含む、請求項 1 または 2 に記載のユーザインタフェース装置。

【請求項 5】

コンピュータにユーザインタフェース処理を実行させるためのユーザインタフェースプログラムであって、

ユーザ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獲得するステップと、

前記対象となる画像領域内において、目の位置を抽出するステップと、

前記検出された目の位置に基づき、予め記憶装置に登録された前記目を含む画像についての情報との比較に基づいて、目の開閉および視線方向を含む目の状態を判定するステップと、

前記ユーザの口の位置を抽出するステップと、

前記ユーザの口に対して、予め前記記憶装置に登録された前記口を含む画像についての情報との比較に基づいて、開状態および閉状態を含む口の状態を判定するステップと、

検出された前記目の状態および前記口の状態の組み合わせに応じて、複数の所定の処理のうち、前記口が前記開状態にあるときのみ、前記目の状態に応じた処理を実行するように、予め対応づけられた処理を選択して実行するステップと、

実行された前記処理に対応する結果を、出力装置に出力するステップとをコンピュータに実行させる、ユーザインタフェースプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

この発明はカメラ等により撮影されたユーザの画像をコンピュータのユーザインタフェースとして利用するユーザインタフェース装置およびユーザインタフェースプログラムに関する。

【背景技術】

【0002】

コンピュータのユーザインタフェースにおいて、入力デバイスは、従来、たとえば、キーボードやマウスといった人間が手で操作するデバイスが用いられることが一般的である。

【0003】

一方、いわゆる音声認識技術を用いて、人間の発話により、直接コンピュータに文字等を入力するためのソフトウェアも開発され、汎用のコンピュータの音声入力機能を用いることで、実用的なレベルのものが販売されている。

【0004】

しかしながら、例えば、事故による脊椎損傷などで首から下の筋肉が動かせなくなり、発話も困難になった人などに対しては、以上のような従来の入力装置では、十分な入力動作を提供できない。

【0005】

そこで、コンピュータへの入力インタフェースとして、さらには、このようなコンピュータを用いた他人とのコミュニケーションのために、ユーザの視線を検出して文字を綴ることができる装置が提案されている（たとえば、特許文献 1 を参照）。

【0006】

10

20

30

40

50

しかし、視線のみによる操作は、たとえば、非特許文献1にも記載されているように、ポインティング操作に適していないだけでなく、ポインティングか単なる視線移動かの区別も難しいという問題がある。

【0007】

また、視線を解析するには目を拡大撮像する必要があるだけでなく、赤外光を照射するなど、カメラ以外の装置が必要になる。頭部に視線を検出する装置を装着するタイプの装置もあるが、障害者が自分で着脱できないので使いづらい。

【0008】

一方で、撮影された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。

10

【0009】

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。

【0010】

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2、非特許文献2～非特許文献3を参照)。

【0011】

ただし、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。また、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、といったような問題があった。

20

【非特許文献1】大野健彦、"視線を用いたインタフェース"、情報処理、2003年7月号、44巻7号、pp.726-732

【非特許文献2】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-II), vol. J84-D-II, no12, pp.2577-2584, Dec.2001.

30

【非特許文献3】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.

【特許文献1】特開2000-020196号公報明細書

【特許文献2】特開2001-52176号公報明細書

【発明の開示】

【発明が解決しようとする課題】

【0012】

本発明の目的は、顔のジェスチャーをスイッチとして、ユーザの手の動作を不要とし、コンピュータの安定なインタフェースとして利用することが可能なユーザインタフェース装置およびユーザインタフェースプログラムを提供することである。

40

【課題を解決するための手段】

【0013】

この発明のある局面に従うと、ユーザインタフェース装置であって、ユーザ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獲得する撮影手段と、対象となる画像領域内において、目の位置を抽出する目検出手段と、検出された目の位置に基づき、予め記憶装置に登録された目を含む画像についての情報との比較に基づいて、目の開閉および視線方向を含む目の状態を判定する目状態判定手段と、少なくとも検出された目の状態に応じて、複数の所定の処理のうち、予め対応づけられた処理を選択して実行する処理手段と、実行された処理に対応する結果を、出力する出力手段とを備える。

【0014】

50

好ましくは、ユーザの口の位置を検出して、ユーザの口に対して、予め記憶装置に登録された口を含む画像についての情報との比較に基づいて、開状態および閉状態を含む口の状態を判定する口状態判定手段をさらに備え、処理手段は、目の状態および口の状態の組み合わせに応じて、複数の所定の処理のうち、予め対応づけられた処理を選択して実行する。

【 0 0 1 5 】

好ましくは、予め記憶装置に登録された目を含む画像についての情報は、目の間隔で正規化された目を含む所定の範囲の画素の明度の平均値を含む。

【 0 0 1 6 】

好ましくは、予め記憶装置に登録された口を含む画像についての情報は、目の間隔で正規化された口を含む所定の範囲の画素の明度の平均値を含む。

【 0 0 1 7 】

この発明の他の局面に従うと、コンピュータにユーザインタフェース処理を実行させるためのユーザインタフェースプログラムであって、ユーザ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獲得するステップと、対象となる画像領域内において、目の位置を抽出するステップと、検出された目の位置に基づき、予め記憶装置に登録された目を含む画像についての情報との比較に基づいて、目の開閉および視線方向を含む目の状態を判定するステップと、少なくとも検出された目の状態に応じて、複数の所定の処理のうち、予め対応づけられた処理を選択して実行するステップと、実行された処理に対応する結果を、出力装置に出力するステップとをコンピュータに実行させる。

【 0 0 1 8 】

好ましくは、ユーザの口の位置を検出するステップと、ユーザの口に対して、予め記憶装置に登録された口を含む画像についての情報との比較に基づいて、開状態および閉状態を含む口の状態を判定するステップをさらに備え、実行するステップは、目の状態および口の状態の組み合わせに応じて、複数の所定の処理のうち、予め対応づけられた処理を選択して実行する。

【発明を実施するための最良の形態】

【 0 0 1 9 】

[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にかかるユーザインタフェース装置について説明する。このユーザインタフェース装置は、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、目および口の位置を検出し、さらに目や口の状態を判定して、それらの状態の組み合わせにより、コンピュータの入力装置としての機能を実現するためのものである。

【 0 0 2 0 】

図 1 は、本発明のユーザインタフェース装置が動作するシステム 1 0 0 の構成を示すブロック図である。

【 0 0 2 1 】

システム 1 0 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) 等の固体撮像素子を含むカメラを用い、カメラ 3 0 の前にいてこのシステム 1 0 0 を操作する人物の目および口の位置ならびに目および口の状態を検出する処理を行うものとする。

【 0 0 2 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 2 3 】

図 1 に示されるように、このシステム 1 0 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 2 4 】

既に述べたようにこのユーザインタフェース装置の主要部は、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 2 5 】

図 1 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。なお、記録媒体としては、この他に、メモ리카ードや、D V D (Digital Versatile Disc) - R O M のようなものであってもよい。この場合には、このような媒体に対応した読取り用のドライブ装置が、本体 4 0 に設けられる。

【 0 0 2 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、当該ユーザインタフェース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ユーザインタフェース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。

【 0 0 2 7 】

[ユーザインタフェース処理]

次に、システム 1 0 0 のユーザインタフェース装置としての動作について説明する。

【 0 0 2 8 】

なお、画像中から人間の顔を検出し、顔内から目の位置を検出する方法については、上述した非特許文献 1 ~ 2 等に開示されている。また、本願発明者による特許出願の特願 2 0 0 2 - 3 3 8 1 7 5 号や特願 2 0 0 3 - 3 9 1 1 4 8 号にも開示がある。

【 0 0 2 9 】

そこで、以下の説明では、画像中から目の位置の検出については行われているものとの前提で説明する。なお、その後、特願 2 0 0 2 - 3 3 8 1 7 5 号に開示されている目の位置の検出方法について、説明しておくことにする。

【 0 0 3 0 】

(システム起動時の処理)

システム 1 0 0 は、顔のジェスチャーをスイッチとして、コンピュータのインタフェースとして利用するものである。

10

20

30

40

50

【 0 0 3 1 】

そこで、検出した目や口の状態、例えば、目の状態を、「正面を見ている」、「上方を見ている」、「下方を見ている」、「右方向を見ている」、「左方向をみている」、「目をつむっている」、の6状態のデジタルスイッチとし、口の状態を、「閉じている」、「開けている」、「舌を出している」、の3状態のデジタルスイッチとする。組み合わせると18状態を表現することができる。もちろん、上述のような全ての場合を用いずに、より少ない組み合わせとしてもよい。たとえば、目の状態のみでユーザインタフェースを構成することも可能である。逆に、たとえば、口の状態に、たとえば、「口を大きく開ける」等の状態を付け加えて、さらに多くの組み合わせを表現することとしてもよい。

【 0 0 3 2 】

このために、システム100では、このような口の状態や目の状態をシステム起動時に予め各ユーザについて、登録しておく処理を行う。なお、一度登録すれば、基本的には、同一のユーザについては、2回目以降は、たとえば、ハードディスク54に登録された情報に基づいて、口の状態や目の状態を判断することができる。

【 0 0 3 3 】

図2は、このようなシステム起動時の登録処理について説明するためのフローチャートである。

【 0 0 3 4 】

まず、システム100が起動すると、ディスプレイ42上に、「口を閉じて、モニター画面中央を見て下さい」と表示する(ステップS102)。

【 0 0 3 5 】

後に説明する手続きにしたがって、目と口角を検出して、ディスプレイ42上の画像に検出位置を表示する(ステップS104)。

【 0 0 3 6 】

続いて、「システムは目と口を検出しています。よろしければ舌をだして合図してください。」とディスプレイ42に表示する(ステップS106)。

【 0 0 3 7 】

システム100は、口領域の平均濃度が所定のしきい値以上に明るく変化したことを検知して、変化前と変化後の平均濃度を、口閉じ、舌だしの状態として、たとえば、ハードディスク54に記憶する。同時に、当該ユーザの目間の基準パターン(正面視の状態での両目と目の間を含む矩形領域の基準パターン)と、基準目位置(目間の基準パターンにおける目の位置(黒目の位置)の相対データ:正面視に相当)と、開眼時目領域平均濃度を計算し、ハードディスク54に記憶する(S108)。

【 0 0 3 8 】

続いて、システム100は、「舌を検出しました。次は、口を少し開けて目を閉じ、1から3までゆっくり数えてください。」と表示する(S110)。

【 0 0 3 9 】

口領域の平均濃度が、口閉じ状態よりも暗く変化したことを検知して変化後の平均濃度を口開きの状態として記憶する。同時に、すでに獲得している目間基準パターンと基準目位置相対データから基準目位置を計算し、目領域の平均濃度を閉眼時目領域平均濃度としてハードディスク54に記憶する(ステップS112)。

順次、「右を見てください」、「左を見てください」、「上をみてください」、「下をみてください」と指示して、見ている方向に対応する、目位置の基準目位置からのずれ量を検出し、ハードディスク54に記憶する(ステップS114)。

【 0 0 4 0 】

そして、システム100は、登録処理が終了すると、「初期化処理を完了しました」とディスプレイ42に表示する(ステップS116)。

【 0 0 4 1 】

以後は、システム100は、新しい画像を入力する度に、目の状態を3ビットのデータ、口の状態を2ビットのデータとして判断し、これらの組み合わせに応じて予め定められ

10

20

30

40

50

た処理に対応した処理結果を出力する（ステップS 1 1 8）。

【 0 0 4 2 】

例えば、システム 1 0 0 を文字編集システムとして使用した場合、舌がでている状態での目の状態の変化でカーソルを移動し、口が開いた変化でクリック動作とみなして、文字を編集するような処理を行うことができる。口が閉じた状態での目の状態変化では、なにも動作させない。こうすることにより、入力を意図した目の状態変化と入力を意図しない、単なる視線移動による目の状態変化を区別することができる。

【 0 0 4 3 】

（口の位置検出フロー）

次に、口の位置の検出フローについて説明する。

10

【 0 0 4 4 】

上述のとおり、両目位置は、予め検出されているものとして、口検出のフローについて説明する。

【 0 0 4 5 】

カメラ 3 0 により撮影された画像は、システム 1 0 0 により、検出された目の位置に基づいて、両目が水平に並ぶように画像は回転補正されているものとする。

【 0 0 4 6 】

このとき、両目間の距離を L_e とすると、口は目の下方、例えば $0.7 L_e \sim 1.4 L_e$ の範囲にあり、口の横幅は、口を閉じた普通の状態でほぼ L_e に等しく、口を開けたり、口をとがらすと、口の横幅は L_e より小さくなる。

20

【 0 0 4 7 】

図 3 は、口位置の検出を説明するためのフローチャートである。

【 0 0 4 8 】

まず、画像の各走査線（ y 座標：垂直方向の座標）について、左目の x 座標（水平方向の座標）と右目の x 座標の間にある画素の平均濃度を計算しプロットする（ステップ S 2 0 2）。

【 0 0 4 9 】

図 4 は、顔の下半分について、各走査線（水平方向）毎に右目の x 座標から左目の x 座標までの間の画素の濃度の平均をプロットした図である。

【 0 0 5 0 】

30

図 4 に示すとおり、このように、幅 L_e の平均濃度をみると口の位置が一番暗くなるので、ここを口の y 座標として特定できる。すなわち、目の下 $0.7 L_e$ から $1.4 L_e$ の間で、平均濃度が一番低い（暗い） y 座標をみつけ、口の y 座標とする。

【 0 0 5 1 】

図 5 は、左口角テンプレートおよび右口角テンプレートを示す概念図である。

【 0 0 5 2 】

上記のように決定された口の y 座標上で、図 5 に示すような左口角テンプレートおよび右口角テンプレートともっともよくマッチする x 座標を探索して、口の左右の端とする。

【 0 0 5 3 】

すなわち、まず、口の y 座標上で、左口角テンプレートともっともマッチする x 座標を口の左端とする（ステップ S 2 0 6）。

40

【 0 0 5 4 】

次に、口の y 座標上で、右口角テンプレートともっともマッチする x 座標を口の右端とする（ステップ S 2 0 8）。

【 0 0 5 5 】

この場合、マッチ度は問わないので、口の開閉に関わらず口の両端を検出することができることになる。

【 0 0 5 6 】

（口の状態検出フロー）

次に、口の状態の検出フローについて説明する。

50

【 0 0 5 7 】

図 6 は、検出された口の形状を示す概念図である。

【 0 0 5 8 】

図 6 において、以下で説明するように、検出した左右の口角の距離を M とする。

【 0 0 5 9 】

図 7 は、口の状態を検出するための処理を説明するためのフローチャートである。

【 0 0 6 0 】

図 7 を参照して、まず、検出した左右の口角の距離を M とし (ステップ S 3 0 2)、左右の口角の中心を口の中心とする (ステップ S 3 0 4)。

【 0 0 6 1 】

次に、口の状態を検出するために、図 6 に示した、例えば、 $0.7M \times 0.2Le$ の矩形内の平均濃度 (明るさ) を計算する (ステップ S 3 0 6)。

【 0 0 6 2 】

この平均濃度は、口を開けた状態では口を閉じた状態より暗くなり、舌を出した状態では口を閉じた状態より明るくなる。そこで、上述したように、システムの起動時に、3つの状態をシステム 1 0 0 に教示して、3つの状態の平均濃度をハードディスク 5 4 に記憶させておく。

【 0 0 6 3 】

次に、計算された平均濃度に基づいて、その値が上記 3つの状態のどれに近いかが比較をし (ステップ S 3 0 8)、現在の口の状態を判定する (ステップ S 3 1 0)。

【 0 0 6 4 】

システム稼働時には、口角を検出する度に、この平均濃度を計算して、その値が3つの状態のどれに近いかが判定して、口の状態を判定する。

【 0 0 6 5 】

(目の状態の検出フロー)

次に、目の状態の検出フローについて説明する。

【 0 0 6 6 】

図 8 は、目の状態を検出するための処理を説明するためのフローチャートである。

【 0 0 6 7 】

図 8 を参照して、システム稼働時には、入力顔画像に対して、まず目間基準パターンのテンプレートマッチングにより目間位置を検出し (ステップ S 4 0 2)、正面視時の目間からの目の相対位置データに基づいて、左右の基準目位置を計算する (ステップ S 4 0 4)。

【 0 0 6 8 】

次に、計算された基準目位置を中心とする、たとえば、 $0.3Le \times 0.2Le$ の矩形領域の平均濃度を計算する (ステップ S 4 0 6)。計算された平均濃度が閉眼時の値に近いかが開眼時の値に近いかが比較し、閉眼状態か開眼状態かを判定する (ステップ S 4 0 8)。

【 0 0 6 9 】

閉眼状態ならば (ステップ S 4 1 0)、閉眼状態と判断して、結果をユーザインタフェースのメインルーチンに返す (ステップ S 4 2 8)。

【 0 0 7 0 】

一方、開眼状態ならば (ステップ S 4 1 0)、検出された目位置と基準目位置とのずれを計算し (ステップ S 4 1 2)、そのずれ量がしきい値以下なら (ステップ S 4 1 4)、正面視状態と判定する (ステップ S 4 1 8)。

【 0 0 7 1 】

特に、限定されないが、目の位置のずれの量は、図 2 で説明したように、正面視の場合の目の部分のテンプレート画像と、目を指定された方向に向けた際の目の部分のテンプレート画像とを、予め獲得してハードディスク 5 4 に格納しておき、これらとの比較により特定することが可能である。

10

20

30

40

50

【 0 0 7 2 】

ずれがしきい値以上ならば（ステップ S 4 1 4）、左右上下視のどのケースにもっとも近いと比較し（ステップ S 4 1 6）、一番近いケースに応じて、それぞれ、開眼状態で上視（上を見た状態）（ステップ S 4 2 0）、開眼状態で下上視（下を見た状態）（ステップ S 4 2 2）、開眼状態で左視（左を見た状態）（ステップ S 4 2 4）、開眼状態で右視（右を見た状態）（ステップ S 4 2 6）と判断して、結果をユーザインタフェースのメインルーチンに返す。

【 0 0 7 3 】

目の 6 状態、口の 3 状態の区別は、U S B (Universal Serial Bus) カメラなど安価な画像入力装置により、得られる画像で判別可能であり、赤外線照射も不要なのでズームレンズさえ使えば、ユーザとの距離も自由に設定でき、非装着で操作できる。

10

【 0 0 7 4 】

このようなユーザインタフェースに基づく処理の他の例としては、たとえば、目の向きで画面上のカーソルを操作し、目をつむることによって、そのカーソル位置のメニューを選択することにする。ただし、舌がでていいる時には操作目的で目を使っていることとし、舌がでていなければ、単に見る目的で目が動いていると判断する。

【 0 0 7 5 】

以上のような処理では、ユーザの視線がディスプレイ 4 2 上のどの位置に向いているか、というような細かい情報は必要ないので、目の状態、口の状態の区別により、簡易にコンピュータとユーザとの間のインタフェースを実現できる。

20

【 0 0 7 6 】

図 9 は、このような口の開閉、視線の方向の組み合わせを変化させた画像の一例を示す図である。

【 0 0 7 7 】

図 9 には、上述した 1 8 通りのうち、8 通りの画像について示している。

【 0 0 7 8 】

次に、以下では、画像中から人間の顔を検出し、顔内から目の位置を検出する方法の一例として、上述した特許出願の特願 2 0 0 2 - 3 3 8 1 7 5 号に記載されている方法を説明する。

【 0 0 7 9 】

[顔画像の抽出の基本的原理]

まず、目の位置の検出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。

30

【 0 0 8 0 】

連続した画素が眉間候補となる場合は、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（上述した目間の標準パターン）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。

40

【 0 0 8 1 】

以下、本発明の顔検出手続きについて、さらに詳しく説明する。

【 0 0 8 2 】

（ 6 分割矩形フィルタ ）

図 1 0 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。

【 0 0 8 3 】

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

50

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。

【0084】

図10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 $S_1 \sim S_6$ に分割する。

【0085】

図11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11(a)は6分割矩形フィルタの形状を示し、図11(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。

【0086】

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12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。

【0087】

以下の説明では、図12に示すような6分割矩形フィルタを用いるものとする。

【0088】

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。

【0089】

図12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 ($1 \leq i \leq 6$) 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。

【0090】

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。

【0091】

【数1】

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

【0092】

図13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。

【0093】

図13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 $M \times N$ 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。

【0094】

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。

【0095】

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。

10

20

30

40

50

【 0 0 9 6 】

【数 2】

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

【 0 0 9 7 】

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 9 8 】

【数 3】

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

10

【 0 0 9 9 】

$s(x, y)$ は行の画素の総和を表す。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めることができることである。

【 0 1 0 0 】

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 1 4 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。

20

【 0 1 0 1 】

インテグラルイメージを用いて、図 1 4 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。

【 0 1 0 2 】

【数 4】

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

30

【 0 1 0 3 】

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うことが可能である。

【 0 1 0 4 】

(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。

【 0 1 0 5 】

図 1 5 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。

40

【 0 1 0 6 】

図 1 5 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 S 1 0 0 0)。

【 0 1 0 7 】

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 S 1 0 2 0)。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 S 1 0 4 0)。

【 0 1 0 8 】

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 S 1 0 6 0)。

50

【0109】

テスト条件を満たす場合は（ステップS1080）、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S1100）。一方、テスト条件を満たさない場合は（ステップS1080）、処理はステップS1120に移行する。

【0110】

ステップS1120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S1140）。フィルタが動ける範囲内であるときは、処理はステップS1020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S1160）。

10

【0111】

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S1180）。フィルタが動ける範囲内であるときは、処理はステップS1020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S1200）。ここで、「中央の画素」とは、特に限定されないが、たとえば、各連結要素の重心位置とすることができる。

【0112】

（目の候補点の抽出および真の眉間候補点の抽出）

20

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。

【0113】

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。

【0114】

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。図16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。

【0115】

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10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。

30

【0116】

図17は、このような目の候補点の抽出を行った上で、真の眉間候補点の抽出を行う処理を説明するためのフローチャートである。

【0117】

図17を参照して、まず、眉間候補抽出フィルターのブロック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S2000）。

【0118】

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S2020）。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S2040）。

40

【0119】

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S2060）。

【0120】

類似度が予め定めたしきい値以上かを判断し（ステップS2080）、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S2100）。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S2120）。

50

【 0 1 2 1 】

このような処理をすべての眉間候補点について行う。

【 0 1 2 2 】

以下では、上述した「目間の標準パターン」を「眉間テンプレート」と呼ぶことにする。

【 0 1 2 3 】

本願では、眉間テンプレートは、上述したように各ユーザごとに設定することとする。

【 0 1 2 4 】

次に、図 1 7 のステップ S 2 0 6 0 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。

【 0 1 2 5 】

図 1 8 は、ステップ S 2 0 6 0 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。

【 0 1 2 6 】

図 1 8 を参照して、まず、眉間候補点を抽出して（ステップ S 4 0 0 0 ）、必要に応じて、眉間候補点を中心に回転を行い、スケール補正を行う（ステップ S 4 0 2 0 ）。

【 0 1 2 7 】

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 S 4 0 4 0 ）。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 S 4 0 6 0 ）。

【 0 1 2 8 】

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。

【 0 1 2 9 】

以上のような手続きで、目の位置を検出することができる。

【 0 1 3 0 】

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 】

【 0 1 3 1 】

【 図 1 】 本発明にかかるシステム 1 0 0 のハードウェア的構成を示すブロック図である。

【 図 2 】 システム起動時の登録処理について説明するためのフローチャートである。

【 図 3 】 口位置の検出を説明するためのフローチャートである。

【 図 4 】 顔の下半分について、各走査線（水平方向）毎に右目の x 座標から左目の x 座標までの間の画素の濃度の平均をプロットした図である。

【 図 5 】 左口角テンプレートおよび右口角テンプレートを示す概念図である。

【 図 6 】 検出された口の形状を示す概念図である。

【 図 7 】 口の状態を検出するための処理を説明するためのフローチャートである。

【 図 8 】 目の状態を検出するための処理を説明するためのフローチャートである。

【 図 9 】 口の開閉、視線の方向の組み合わせを変化させた画像の一例を示す図である。

【 図 1 0 】 6 分割矩形フィルタを示す図である。

【 図 1 1 】 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 1 2 】 6 分割矩形フィルタの他の構成を示す概念図である。

【 図 1 3 】 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 1 4 】 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 1 5 】 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。

10

20

30

40

50

【図16】右目のテンプレートを示す図である。

【図17】目の候補点の抽出を行った上で、真の眉間候補点の抽出を行う処理を説明するためのフローチャートである。

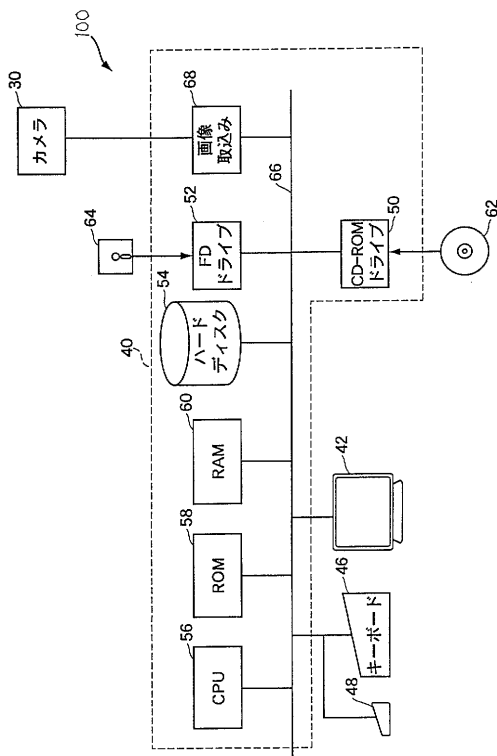
【図18】ステップS2060のテンプレートマッチングの手続きを説明するためのフローチャートである。

【符号の説明】

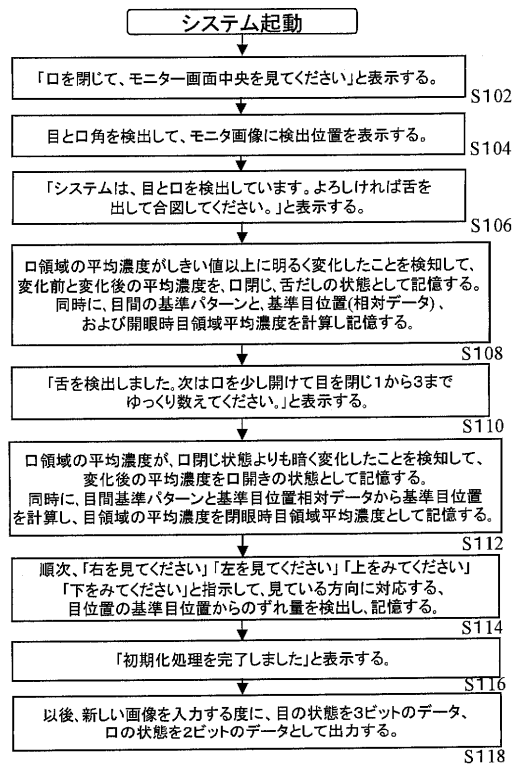
【0132】

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。

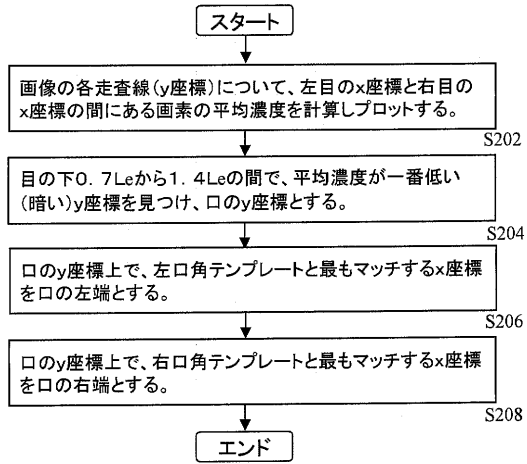
【図1】



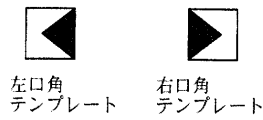
【図2】



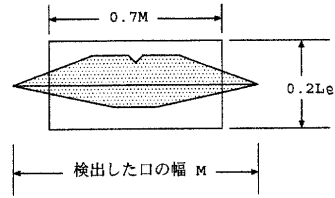
【図3】



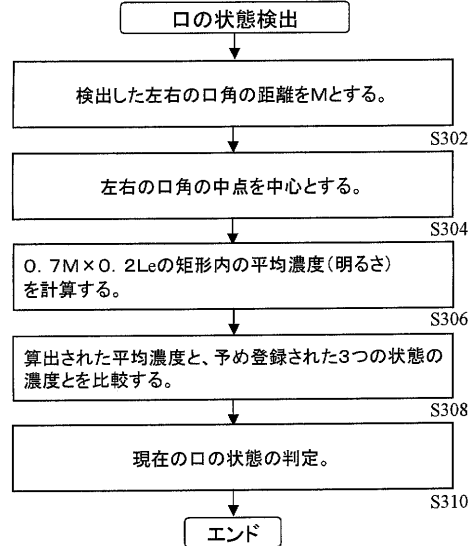
【図5】



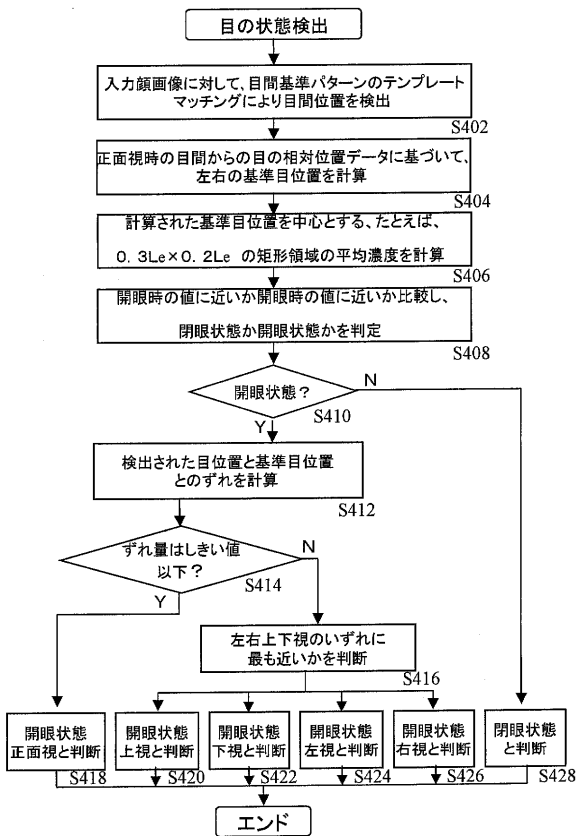
【図6】



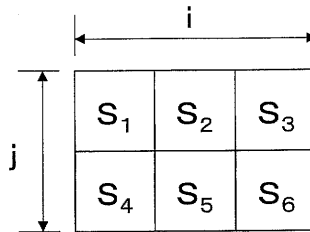
【図7】



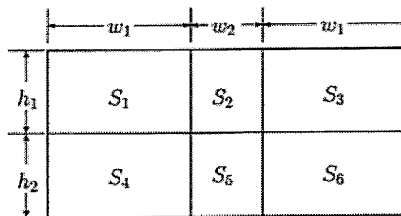
【図8】



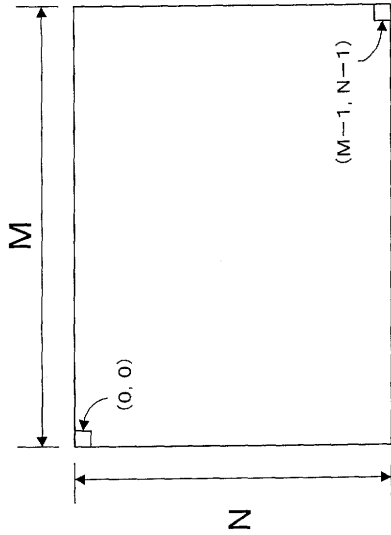
【図10】



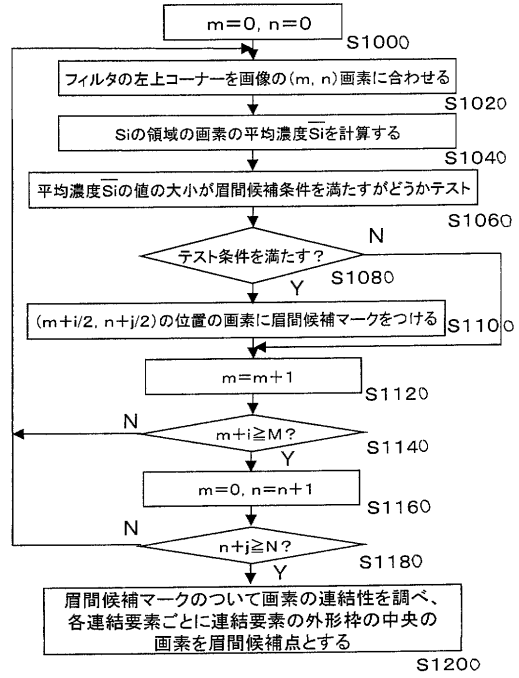
【図12】



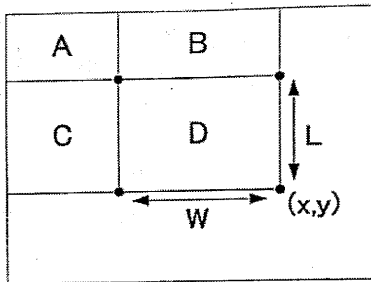
【図13】



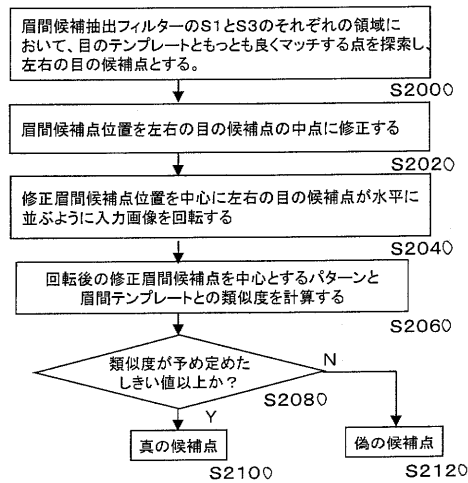
【図15】



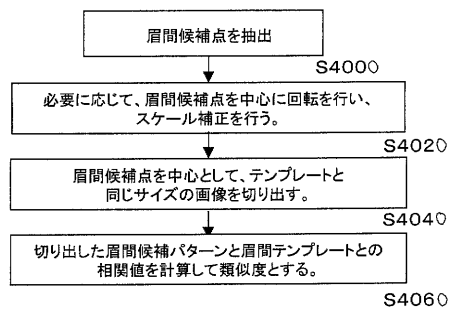
【図14】



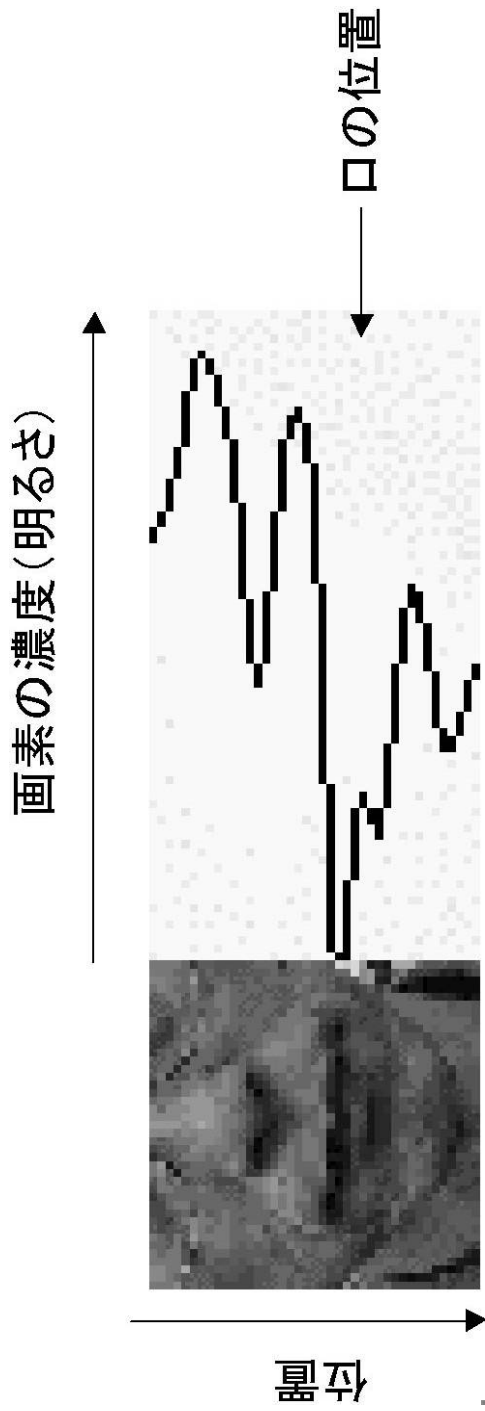
【図17】



【図18】



【 図 4 】



【図9】



口閉、閉眼状態



口閉、開眼、下視状態



口閉、開眼、左視状態



口閉、開眼、右視状態



口閉、開眼、上視状態



口開、開眼、下視状態

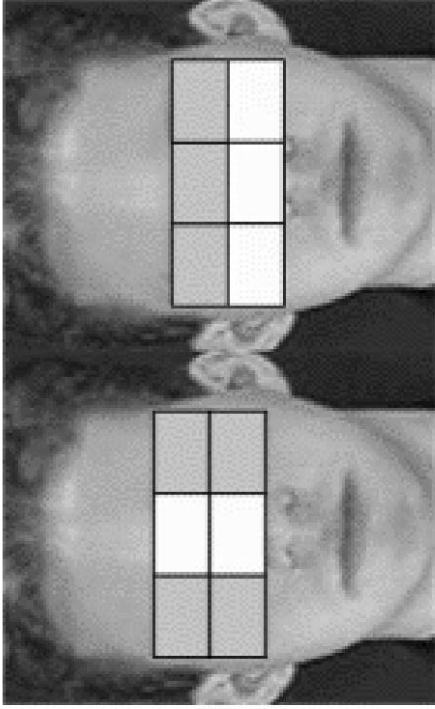


口開、開眼、正面視状態



舌出し、開眼、正面視状態

【 図 1 1 】

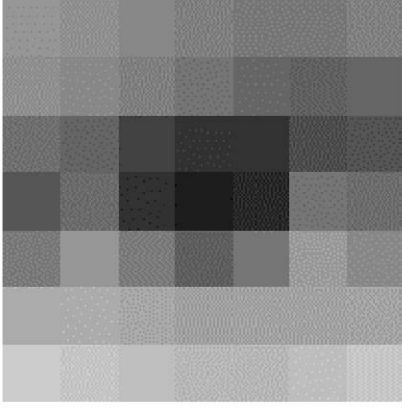


S1	S2	S3
S4	S5	S6

(a)

(b)

【図16】



右目テンプレート

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

審査官 山崎 慎一

(56)参考文献 特開2002-323956(JP,A)

特開2003-085583(JP,A)

特開2002-269544(JP,A)

特開平11-306325(JP,A)

特開2000-020196(JP,A)

特開2001-052176(JP,A)

特開2004-185611(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 3/033

G06T 1/00