

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3980341号  
(P3980341)

(45) 発行日 平成19年9月26日(2007.9.26)

(24) 登録日 平成19年7月6日(2007.7.6)

(51) Int. Cl.	F I
<b>G06T 1/00 (2006.01)</b>	G06T 1/00 340A
<b>A61B 3/113 (2006.01)</b>	A61B 3/10 B
<b>G06T 7/00 (2006.01)</b>	G06T 7/00 300D
<b>G06T 7/20 (2006.01)</b>	G06T 7/20 B
<b>G06T 7/60 (2006.01)</b>	G06T 7/60 150B

請求項の数 15 (全 22 頁)

(21) 出願番号	特願2001-369076 (P2001-369076)	(73) 特許権者	393031586
(22) 出願日	平成13年12月3日(2001.12.3)		株式会社国際電気通信基礎技術研究所
(65) 公開番号	特開2003-168106 (P2003-168106A)		京都府相楽郡精華町光台二丁目2番地2
(43) 公開日	平成15年6月13日(2003.6.13)	(74) 代理人	100064746
審査請求日	平成16年6月4日(2004.6.4)		弁理士 深見 久郎
		(74) 代理人	100085132
			弁理士 森田 俊雄
		(74) 代理人	100091395
			弁理士 吉田 博由
		(74) 代理人	100096781
			弁理士 堀井 豊
		(74) 代理人	100096792
			弁理士 森下 八郎

最終頁に続く

(54) 【発明の名称】 目の位置のトラッキング方法、目の位置のトラッキング装置およびそのためのプログラム

(57) 【特許請求の範囲】

【請求項1】

被験者の顔のビデオ画像を撮像する撮像手段が接続され、記憶装置を有するコンピュータにおいて、前記撮像手段の出力する一連の画像内における前記被験者の目の位置をトラッキングするための方法であって、

前記撮像手段の出力する顔画像内の目の位置を抽出するステップを備え、

前記目の位置を抽出するステップは、

前記撮像手段の出力する二つの画像の差分画像を算出するステップと、

前記差分画像のうち、前記二つの画像の間における前記被験者の顔全体の移動に起因する差分を除いた残りの差分領域を抽出するステップとを含み、

前記残りの差分領域を抽出するステップは、

前記二つの画像の各々を平行移動した画像の画素値と前記差分画像の画素値とが一致する領域を最大にする移動量を算出するステップと、

前記算出した移動量に起因する差分を前記差分画像から削除するステップとを有し、

前記抽出された残りの差分領域から前記二つの画像の画素値の差に応じて二つの目の候補点を選択し、前記二つの目の候補点が予め定められた幾何学的条件を満たす場合に、前記二つの目の候補点を目の位置として抽出するステップをさらに含み、

前記抽出された目の位置によって定められる眉間の位置、前記眉間の位置に対する両目の相対位置、および前記眉間の位置によって定められる眉間の画像パターンを前記記憶装置にセーブするステップと、

10

20

前記目の位置が抽出された画像に後続する画像を取得する画像取得ステップと、  
前記後続する画像内の顔画像の眉間の位置を予測するステップと、  
予測された眉間の位置の近傍で、前記セーブされた眉間の画像パターンと最もよく一致する領域の中心となる点を探索するステップと、

前記探索された点の位置と、前記セーブされた両目の相対位置とに基づいて、前記後続する画像内での両目の位置を予測し、当該予測された領域を中心として予め定められた条件を満足する二つの領域の各々の中心となる点を探索するステップと、

探索された前記二つの中心となる点の中心の位置を新たな眉間の位置とし、当該新たな眉間の位置および前記二つの中心となる点の相対位置によって定まる領域を前記眉間の位置に対する新たな両目の相対位置とし、さらに前記新たな眉間の位置によって定められる領域の画像パターンを新たな眉間の画像パターンとして前記記憶装置に記憶されている眉間の位置、両目の相対位置、および眉間の画像パターンをそれぞれ更新するステップと、

前記後続する画像にさらに後続する画像に対して前記画像取得ステップから処理を実行するステップとをさらに備える、目の位置のトラッキング方法。

【請求項 2】

前記眉間の位置を予測するステップは、前記目の位置が抽出された画像内での眉間の位置と、当該画像に先行する画像内での眉間の位置とから、前記後続する画像の眉間の位置を外挿するステップを含む、請求項 1 に記載の方法。

【請求項 3】

前記二つの領域の各々の中心となる点を探索するステップにおいて、当該点が探索できなかったことに応答して、前記撮像手段の出力する顔画像の眉間の位置、前記眉間の位置に対する両目の相対位置、および前記眉間の位置によって定められる眉間の画像パターンを予め定められた方法によって抽出し、前記画像取得ステップから処理を再開するステップをさらに含む、請求項 1 または請求項 2 に記載の方法。

【請求項 4】

前記二つの領域の各々の中心となる点を探索するステップは、前記探索された点の位置に対し、前記両目の相対位置によって定められる位置の各々を中心とする近傍において、予め定められた形状の領域であって、かつ当該領域内の画素値の平均が最も暗くなる領域の中心を探索して両目の候補点とするステップを含む、請求項 1 ~ 請求項 3 のいずれかに記載の方法。

【請求項 5】

前記二つの領域の各々の中心となる点を探索するステップはさらに、前記最も暗くなる領域の中心を探索するステップで探索された候補点が、

- 1) 候補点間の距離があらかじめ定められた最小値以上であること、
- 2) 候補点間の距離があらかじめ定められた最大値以下であること、および
- 3) 候補点を結ぶ直線と走査線方向とのなす角度が、あらかじめ定められた関係を満足すること、

のすべての条件を満足するか否かを判定し、いずれかの条件が満足されない場合は探索を失敗とするステップを含む、請求項 4 に記載の方法。

【請求項 6】

被験者の顔のビデオ画像を撮像する撮像手段が接続され、記憶装置を有するコンピュータにおいて、前記撮像手段の出力する一連の画像内における前記被験者の目の位置をトラッキングするための装置であって、

前記撮像手段の出力する顔画像内の目の位置を抽出するための手段を備え、

前記目の位置を抽出するための手段は、

前記撮像手段の出力する二つの画像の差分画像を算出するための手段と、

前記差分画像のうち、前記二つの画像の間における前記被験者の顔全体の移動に起因する差分を除いた残りの差分領域を抽出するための手段とを含み、

前記残りの差分領域を抽出するための手段は、

前記二つの画像の各々を平行移動した画像の画素値と前記差分画像の画素値とが一致す

10

20

30

40

50

る領域を最大にする移動量を算出するための手段と、

前記算出した移動量に起因する差分を前記差分画像から削除するための手段とを有し、  
前記抽出された残りの差分領域から前記二つの画像の画素値の差に応じて二つの目の候補点を選択し、前記二つの目の候補点が予め定められた幾何学的条件を満たす場合に、前記二つの目の候補点を目の位置として抽出するための手段をさらに含み、

前記抽出された目の位置によって定められる眉間の位置、前記眉間の位置に対する両目の相対位置、および前記眉間の位置によって定められる眉間の画像パターンを前記記憶装置にセーブするための手段と、

前記目の位置が抽出された画像に後続する画像を取得するための画像取得手段と、

前記後続する画像内の顔画像の眉間の位置を予測するための予測手段と、

予測された眉間の位置の近傍で、前記セーブされた眉間の画像パターンと最もよく一致する領域の中心となる点を探索するための第1の探索手段と、

前記探索された点の位置と、前記セーブされた両目の相対位置とに基づいて、前記後続する画像内での両目の位置を予測し、当該予測された領域を中心として予め定められた条件を満足する二つの領域の各々の中心となる点を探索するための第2の探索手段と、

探索された前記二つの中心となる点の midpoint の位置を新たな眉間の位置とし、当該新たな眉間の位置および前記二つの中心となる点の相対位置によって定まる領域を前記眉間の位置に対する新たな両目の相対位置とし、さらに前記新たな眉間の位置によって定められる領域の画像パターンを新たな眉間の画像パターンとして前記記憶装置に記憶されている眉間の位置、両目の相対位置、および眉間の画像パターンをそれぞれ更新するための更新手段と、

前記後続する画像にさらに後続する画像に対して前記画像取得手段からの処理を繰返すように前記画像取得手段、前記予測手段、前記第1および第2の探索手段、ならびに前記更新手段を制御するための手段とをさらに備える、目の位置のトラッキング装置。

#### 【請求項7】

前記予測手段は、前記目の位置が抽出された画像内での眉間の位置と、当該画像に先行する画像内での眉間の位置とから、前記後続する画像の眉間の位置を外挿するための手段を含む、請求項6に記載の装置。

#### 【請求項8】

前記第2の探索手段が当該点を探索できなかったことに応答して、前記撮像手段の出力する顔画像の眉間の位置、前記眉間の位置に対する両目の相対位置、および前記眉間の位置によって定められる眉間の画像パターンを予め定められた装置によって抽出し、前記画像取得手段から処理を再開するように前記画像取得手段、前記予測手段、前記第1および第2の探索手段、ならびに前記更新手段を制御するための手段をさらに含む、請求項6または請求項7に記載の装置。

#### 【請求項9】

前記第2の探索手段は、

前記探索された点の位置に対し、前記両目の相対位置によって定められる位置の各々を中心とする近傍において、予め定められた形状の領域であって、かつ当該領域内の画素値の平均が最も暗くなる領域の中心を探索して両目の候補点とするための手段を含む、請求項6～請求項8のいずれかに記載の装置。

#### 【請求項10】

前記第2の探索手段はさらに、前記最も暗くなる領域の中心を探索する手段で探索された候補点が、

- 1) 候補点間の距離があらかじめ定められた最小値以上であること、
- 2) 候補点間の距離があらかじめ定められた最大値以下であること、および
- 3) 候補点を結ぶ直線と走査線方向とのなす角度が、あらかじめ定められた関係を満足すること、

のすべての条件を満足するか否かを判定し、いずれかの条件が満足されない場合は探索を失敗とするための手段を含む、請求項9に記載の装置。

10

20

30

40

50

## 【請求項 1 1】

被験者の顔のビデオ画像を撮像する撮像手段が接続され、演算装置と記憶装置とを有するコンピュータにおいて、前記撮像手段の出力する一連の画像内における前記被験者の目の位置をトラッキングするための処理を実行するように前記コンピュータを制御する、コンピュータにより実行可能なプログラムであって、

前記処理は、

前記演算装置が前記撮像手段の出力する顔画像内の目の位置を抽出するステップを備え

、前記目の位置を抽出するステップは、

前記撮像手段の出力する二つの画像の差分画像を算出するステップと、

前記差分画像のうち、前記二つの画像の間における前記被験者の顔全体の移動に起因する差分を除いた残りの差分領域を抽出するステップとを含み、

前記残りの差分領域を抽出するステップは、

前記二つの画像の各々を平行移動した画像の画素値と前記差分画像の画素値とが一致する領域を最大にする移動量を算出するステップと、

前記算出した移動量に起因する差分を前記差分画像から削除するステップとを有し、

前記抽出された残りの差分領域から前記二つの画像の画素値の差に応じて二つの目の候補点を選択し、前記二つの目の候補点が予め定められた幾何学的条件を満たす場合に、前記二つの目の候補点を目の位置として抽出するステップをさらに含み、

前記演算装置が前記抽出された目の位置によって定められる眉間の位置、前記眉間の位置に対する両目の相対位置、および前記眉間の位置によって定められる眉間の画像パターンを前記記憶装置にセーブするステップと、

前記演算装置が前記目の位置が抽出された画像に後続する画像を取得する画像取得ステップと、

前記演算装置が前記後続する画像内の顔画像の眉間の位置を予測するステップと、

前記演算装置が予測された眉間の位置の近傍で、前記セーブされた眉間の画像パターンと最もよく一致する領域の中心となる点を探索するステップと、

前記演算装置が前記探索された点の位置と、前記セーブされた両目の相対位置とに基づいて、前記後続する画像内での両目の位置を予測し、当該予測された領域を中心として予め定められた条件を満足する二つの領域の各々の中心となる点を探索するステップと、

前記演算装置が探索された前記二つの中心となる点の midpoint の位置を新たな眉間の位置とし、当該新たな眉間の位置および前記二つの中心となる点の相対位置によって定まる領域を前記眉間の位置に対する新たな両目の相対位置とし、さらに前記新たな眉間の位置によって定められる領域の画像パターンを新たな眉間の画像パターンとして前記記憶装置に記憶されている眉間の位置、両目の相対位置、および眉間の画像パターンをそれぞれ更新するステップと、

前記演算装置が前記後続する画像にさらに後続する画像に対して前記画像取得ステップから処理を実行するステップとをさらに備える、目の位置のトラッキングプログラム。

## 【請求項 1 2】

前記眉間の位置を予測するステップは、前記目の位置が抽出された画像内での眉間の位置と、当該画像に先行する画像内での眉間の位置とから、前記後続する画像の眉間の位置を外挿するステップを含む、請求項 1 1 に記載のプログラム。

## 【請求項 1 3】

前記処理は、前記二つの領域の各々の中心となる点を探索するステップにおいて、当該点が探索できなかったことに応答して、前記撮像手段の出力する顔画像の眉間の位置、前記眉間の位置に対する両目の相対位置、および前記眉間の位置によって定められる眉間の画像パターンを予め定められた方法によって抽出し、前記画像取得ステップから処理を再開するステップをさらに含む、請求項 1 1 または請求項 1 2 に記載のプログラム。

## 【請求項 1 4】

前記二つの領域の各々の中心となる点を探索するステップは、

10

20

30

40

50

前記探索された点の位置に対し、前記両目の相対位置によって定められる位置の各々を中心とする近傍において、予め定められた形状の領域であって、かつ当該領域内の画素値の平均が最も暗くなる領域の中心を探索して両目の候補点とするステップを含む、請求項 1 1 ~ 請求項 1 3 のいずれかに記載のプログラム。

【請求項 1 5】

前記二つの領域の各々の中心となる点を探索するステップはさらに、前記最も暗くなる領域の中心を探索するステップで探索された候補点が、

- 1) 候補点間の距離があらかじめ定められた最小値以上であること、
- 2) 候補点間の距離があらかじめ定められた最大値以下であること、および
- 3) 候補点を結ぶ直線と走査線方向とのなす角度が、あらかじめ定められた関係を満足すること、

10

のすべての条件を満足するか否かを判定し、いずれかの条件が満足されない場合は探索を失敗とするステップを含む、請求項 1 4 に記載のプログラム。

【発明の詳細な説明】

【0001】

【発明が属する技術分野】

この発明はマン・マシンインタフェース技術に関し、特に、人の視線によってコンピュータなどを操作する際に、人の視線方向を誤りなく検出する目的のために、人の目の位置をトラッキングする方法および装置に関する。

【0002】

20

【従来の技術】

近年のコンピュータ技術の発達により、人々の生活の隅々にまでコンピュータを用いた装置が用いられている。コンピュータを操作する技術なしには、満足な社会生活も営めなくなるおそれさえある。

【0003】

一方で、コンピュータを操作するためには人の意思をコンピュータに伝える必要がある。いかに効率よく、誤りなく、そして簡単に人の意思をコンピュータに伝えるかについては、さまざまな研究がなされ、実用化されている。そうした技術は一般にマン・マシンインタフェースと呼ばれている。

【0004】

30

コンピュータを操作するためだけであれば、テキストベースでコンピュータを操作するためのコマンドをたとえばキーボードなどを介してコンピュータに与えればよい。しかし、それでは人間がそれら多数のコマンドを、そのコマンドの構文および必要なパラメータとともに記憶する必要がある。そこで、一般に G U I (Graphical User Interface) と呼ばれるものが考案され、現在では主流を占めている。

【0005】

G U I では、一般的にはマウスなどのポインティングデバイスを用いて、画面に表示されるアイコンなどをポイントし、クリック、ダブルクリック、ドラッグなどの所定の操作を行なうことにより人間の意思をコンピュータに与えることができる。そのため、多数のコマンドを覚える必要がなく、誤りも生じにくいという特徴がある。

40

【0006】

一方で、G U I ではポインティングデバイスを操作する必要がある。そのため、たとえば手の動作に障害がある人などは、G U I を採用したコンピュータであっても操作が困難である。また応用によっては手を自由に利用できない場合もあり、マウスなどのポインティングデバイスを用いることが困難であったりする場合がある。

【0007】

そこで、人間の視線を用いて人間の意思をコンピュータに伝達する技術が種々考えられている。人間が意識的に視線を操作することにより、その視線をポインティングデバイスとして用いて、人間のその意思をコンピュータに与えることができる。また、人間が無意識のうちに視線を移動させた場合、その視線の移動をコンピュータが検出することにより、

50

人間の意思を推定することもできる。

【0008】

そのための前提となる技術は、人間の眼球を撮像して視線方向を推定する装置である。そのためには一般的にアイカメラが用いられる。アイカメラを用いて人間の視線方向を推定することにより、その人間の意思を推定してコンピュータの操作に役立てることができる。

【0009】

アイカメラを用いて視線方向を推定する場合、推定精度を上げるために高解像度で眼球を撮像する必要がある。ところが、そのためにアイカメラでは撮像範囲が狭くなる。また被写界深度も浅い。その結果、アイカメラを用いた場合には、ユーザは顔を前後左右のいずれにも動かさないよう、ほとんど固定していなければならない。そのため、アイカメラを単独で用いた場合にはその利用形態は非常に限定されてしまう。また、特定の個人だけにしか利用できないようなシステムではなく、見知らぬ顔でも目を検出しトラッキングすることができるようにすることが望ましい。

10

【0010】

こうしたアイカメラの使用に伴う問題点を解決するために、本願発明者は、まず広視野のカメラを使用して利用者の目の位置を検出し、その情報を用いてアイカメラのパン、チルト、フォーカスを制御することが有効であることに想到した。つまり、何らかの方法により人間の目の位置を発見し、トラッキングして、逐次その方向と距離とをアイカメラの制御装置に出力するサブシステムを採用することで、アイカメラによる視線方向の検出の精度を高める方法である。

20

【0011】

通常、利用者はコンピュータを利用する際には、アイカメラから50～100cm程度の距離におり、椅子に座っているものと考えられる。したがって、上記したような目の位置を検出し、その情報を用いてアイカメラを制御することにより、アイカメラによる視線方向検出の精度は非常に高くなると考えられる。

【0012】

想定している状況では、顔の動く範囲がかなり限定されている。そのため、顔が画面の高さいっぱいになる程度に大きく撮像することができる。そこで本願発明の発明者は、このように大きく撮像された顔画像の中の、目蓋の動きを検出することにより、両目の位置を検出することができることに思い至った。人間は無意識のうちに瞬きをしており、自然の瞬きを待つことにより、利用者に意識させずにその目の位置を検出することができる。特に、アイカメラを使用する状況では、利用者の協力を得ることも可能である。

30

【0013】

瞬きは継続する動きではない。そのため、いったん目の位置を検出したなら、何らかの方法で目の位置をトラッキングする必要がある。そのためのひとつの方法は、顔画像のテンプレートを多数用意しておき、撮像された顔の画像とテンプレートとをマッチングすることにより、目の位置を推定することである。

【0014】

しかしこの場合、単純なテンプレートマッチングでは顔の向きの変化に対応できない。テンプレートを更新するようにしたとしても、トラッキング点が徐々に実際の目とずれていくことは避けられない。したがって、こうしたテンプレートマッチングによる弱点を克服し、トラッキング処理のロバスト性を向上させる必要がある。

40

【0015】

なお、アイカメラに与える距離の情報については、目の方向さえわかればたとえば超音波センサなどを用いて測定することで十分な精度が得られる。ただし以下の実施の形態では2眼ステレオ画像を用いて、利用者の顔までの距離を推定することとした。もちろんこの他にも、種々の方法を用いて利用者の顔までの距離を測定することができるが、本発明における目の位置の検出およびトラッキングとは直接の関係はないので、以下の説明では距離の測定については詳細な説明は行なわない。

50

## 【 0 0 1 6 】

目の位置のトラッキング方法としては、すでに幾つかの例がある。その一つは特開 2 0 0 0 - 1 6 3 5 6 4 に開示されたものである。この例では、目を中心とする画像パターンをテンプレートとして、いわゆるテンプレートマッチングの手法を用いて目の位置をトラッキングしている。また他の例は、信学技報 P R M U 9 9 - 1 5 1 ( 1 9 9 9 - 1 1 ) , p p . 9 - 1 4 の「リアルタイム視線検出・動作認識システムの開発」と題された論文に開示されたものである。この例では、2台のビデオカメラを用いて得られた顔画像から、左右の目の両端を含む顔の特徴領域の画像および3次元座標を用いた画像のステレオ処理によって顔トラッキングを行ない、さらにこうして推定された顔位置と、左右の目の両端の中点から眼球へ向かうオフセットベクトルとによって、眼球の中心位置が推定される。

10

## 【 0 0 1 7 】

## 【 発明が解決しようとする課題 】

しかし、特開 2 0 0 0 - 1 6 3 5 6 4 に開示された方法では通常のテンプレートマッチングの手法が用いられており、パターンの回転とスケール変化とに弱いという問題点がある。また信学技報に掲載された例では、2つの画像を用いたステレオ処理のみから目の位置を推定しているため、カメラから目までの距離が限定され、さらにトラッキングするパターンをあらかじめ登録しておく必要があるという問題がある。

## 【 0 0 1 8 】

すなわちこれまでは、一般的な利用者の目の位置を、利用者の顔の位置を拘束することなく正確にトラッキングすることができないという問題点があった。

20

## 【 0 0 1 9 】

それゆえに、この発明の目的は、一般的な顔画像から目の位置をロバスト性高くトラッキングすることが可能で、利用者を不当に拘束することがない目の位置のトラッキング方法、装置およびそのためのプログラムを提供することである。

## 【 0 0 2 0 】

この発明の他の目的は、一般的な顔画像から、顔が回転したり移動したりしても目の位置を精度高くかつロバスト性高く検出することが可能で、利用者を不当に拘束することがない目の位置のトラッキング方法、装置およびそのためのプログラムを提供することである。

## 【 0 0 2 1 】

## 【 課題を解決するための手段 】

本願の第 1 の局面に係る発明は、被験者の顔のビデオ画像を撮像する撮像手段が接続され、記憶装置を有するコンピュータにおいて、撮像手段の出力する一連の画像内における被験者の目の位置をトラッキングするための方法であって、撮像手段の出力する顔画像内の目の位置を抽出するステップを備え、目の位置を抽出するステップは、撮像手段の出力する二つの画像の差分画像を算出するステップと、差分画像のうち、二つの画像の間における被験者の顔全体の移動に起因する差分を除いた残りの差分領域を抽出するステップとを含み、残りの差分領域を抽出するステップは、二つの画像の各々を平行移動した画像の画素値と差分画像の画素値とが一致する領域を最大にする移動量を算出するステップと、算出した移動量に起因する差分を差分画像から削除するステップとを有し、抽出された残りの差分領域から二つの画像の画素値の差に応じて二つの目の候補点を選択し、二つの目の候補点が予め定められた幾何学的条件を満たす場合に、二つの目の候補点を目の位置として抽出するステップをさらに含み、抽出された目の位置によって定められる眉間の位置、眉間の位置に対する両目の相対位置、および眉間の位置によって定められる眉間の画像パターンを記憶装置にセーブするステップと、目の位置が抽出された画像に後続する画像を取得する画像取得ステップと、後続する画像内の顔画像の眉間の位置を予測するステップと、予測された眉間の位置の近傍で、セーブされた眉間の画像パターンと最もよく一致する領域の中心となる点を探索するステップと、探索された点の位置と、セーブされた両目の相対位置とに基づいて、後続する画像内での両目の位置を予測し、当該予測された領域を中心として予め定められた条件を満足する二つの領域の各々の中心となる点を探索す

30

40

50

るステップと、探索された二つの中心となる点の中心の位置を新たな眉間の位置とし、当該新たな眉間の位置および二つの中心となる点の相対位置によって定まる領域を眉間の位置に対する新たな両目の相対位置とし、さらに新たな眉間の位置によって定められる領域の画像パターンを新たな眉間の画像パターンとして記憶装置に記憶されている眉間の位置、両目の相対位置、および眉間の画像パターンをそれぞれ更新するステップと、後続する画像にさらに後続する画像に対して画像取得ステップから処理を実行するステップとをさらに備える。

【0022】

好ましくは、眉間の位置を予測するステップは、目の位置が抽出された画像内での眉間の位置と、当該画像に先行する画像内での眉間の位置とから、後続する画像の眉間の位置を外挿するステップを含む。

10

【0023】

さらに好ましくは、この方法は、二つの領域の各々の中心となる点を探索するステップにおいて、当該点が探索できなかったことに応答して、撮像手段の出力する顔画像の眉間の位置、眉間の位置に対する両目の相対位置、および眉間の位置によって定められる眉間の画像パターンを予め定められた方法によって抽出し、画像取得ステップから処理を再開するステップをさらに含む。

【0024】

また、二つの領域の各々の中心となる点を探索するステップは、探索された点の位置に対し、両目の相対位置によって定められる位置の各々を中心とする近傍において、予め定められた形状の領域であって、かつ当該領域内の画素値の平均が最も暗くなる領域の中心を探索して両目の候補点とするステップを含んでもよい。

20

【0025】

さらに好ましくは、二つの領域の各々の中心となる点を探索するステップはさらに、最も暗くなる領域の中心を探索するステップで探索された候補点が、

- 1) 候補点間の距離があらかじめ定められた最小値以上であること、
- 2) 候補点間の距離があらかじめ定められた最大値以下であること、および
- 3) 候補点を結ぶ直線と走査線方向とのなす角度が、あらかじめ定められた関係を満足すること、

のすべての条件を満足するか否かを判定し、いずれかの条件が満足されない場合は探索を失敗とするステップを含んでもよい。

30

【0026】

本願の第2の局面にかかる発明は、被験者の顔のビデオ画像を撮像する撮像手段が接続され、記憶装置を有するコンピュータにおいて、撮像手段の出力する一連の画像内における被験者の目の位置をトラッキングするための装置であって、撮像手段の出力する顔画像内の目の位置を抽出するための手段を備え、目の位置を抽出するための手段は、撮像手段の出力する二つの画像の差分画像を算出するための手段と、差分画像のうち、二つの画像の間における被験者の顔全体の移動に起因する差分を除いた残りの差分領域を抽出するための手段とを含み、残りの差分領域を抽出するための手段は、二つの画像の各々を平行移動した画像の画素値と差分画像の画素値とが一致する領域を最大にする移動量を算出するための手段と、算出した移動量に起因する差分を差分画像から削除するための手段とを有し、抽出された残りの差分領域から二つの画像の画素値の差に応じて二つの目の候補点を選択し、二つの目の候補点が予め定められた幾何学的条件を満たす場合に、二つの目の候補点を目の位置として抽出するための手段をさらに含み、抽出された目の位置によって定められる眉間の位置、眉間の位置に対する両目の相対位置、および眉間の位置によって定められる眉間の画像パターンを記憶装置にセーブするための手段と、目の位置が抽出された画像に後続する画像を取得するための画像取得手段と、後続する画像内の顔画像の眉間の位置を予測するための予測手段と、予測された眉間の位置の近傍で、セーブされた眉間の画像パターンと最もよく一致する領域の中心となる点を探索するための第1の探索手段と、探索された点の位置と、セーブされた両目の相対位置とに基づいて、後続する画像内

40

50



での両目の位置を予測し、当該予測された領域を中心として予め定められた条件を満足する二つの領域の各々の中心となる点を探査するための第2の探索手段と、探索された二つの中心となる点の中点の位置を新たな眉間の位置とし、当該新たな眉間の位置および二つの中心となる点の相対位置によって定まる領域を眉間の位置に対する新たな両目の相対位置とし、さらに新たな眉間の位置によって定められる領域の画像パターンを新たな眉間の画像パターンとして記憶装置に記憶されている眉間の位置、両目の相対位置、および眉間の画像パターンをそれぞれ更新するための更新手段と、後続する画像にさらに後続する画像に対して画像取得手段からの処理を繰返すように画像取得手段、予測手段、第1および第2の探索手段、および更新手段を制御するための手段とをさらに備える。

【0027】

好ましくは、予測手段は、目の位置が抽出された画像内での眉間の位置と、当該画像に先行する画像内での眉間の位置とから、後続する画像の眉間の位置を外挿するための手段を含む。

【0028】

さらに好ましくは、第2の探索手段が当該点を探査できなかったことに応答して、撮像手段の出力する顔画像の眉間の位置、眉間の位置に対する両目の相対位置、および眉間の位置によって定められる眉間の画像パターンを予め定められた装置によって抽出し、画像取得手段から処理を再開するように画像取得手段、予測手段、第1および第2の探索手段、ならびに更新手段を制御するための手段をさらに含む。

【0029】

また、第2の探索手段は、探索された点の位置に対し、両目の相対位置によって定められる位置の各々を中心とする近傍において、予め定められた形状の領域であって、かつ当該領域内の画素値の平均が最も暗くなる領域の中心を探査して両目の候補点とするための手段を含んでもよい。

【0030】

より好ましくは、第2の探索手段はさらに、最も暗くなる領域の中心を探査する手段で探索された候補点が、

- 1) 候補点間の距離があらかじめ定められた最小値以上であること、
- 2) 候補点間の距離があらかじめ定められた最大値以下であること、および
- 3) 候補点を結ぶ直線と走査線方向とのなす角度が、あらかじめ定められた関係を満足すること、

のすべての条件を満足するか否かを判定し、いずれかの条件が満足されない場合は探索を失敗とするための手段を含んでもよい。

【0031】

【発明の実施の形態】

[ハードウェア構成]

図1に、本願発明にかかる目の位置のトラッキング方法および装置を実現するコンピュータシステム20の外観を示す。なお、コンピュータシステム20は目の位置の検出およびトラッキングを行ない、図示していないコンピュータがコンピュータシステム20から与えられる利用者の顔までの距離および目の方向の情報を用いてアイカメラを制御する。なお、本明細書において「被験者」とは、このシステムが目の位置を検出しトラッキングする利用者のことをいい、人間だけでなく動物など「目」に相当するものを有しているものすべてを含み得るものとする。

【0032】

図1を参照して、このコンピュータシステム20は、CD-ROM (Compact Disc Read Only Memory) のためのCD-ROMドライブ50と、フレキシブルディスク (FD) のためのFDドライブ52とを備えたコンピュータ40と、コンピュータ40に接続された第1のポインティングデバイスであるマウス48と、キーボード46と、モニタ42と、モニタ42とキーボード46との間に配置されたステレオビデオカメラ56, 58と、前述した別のコンピュータにより制御されるアイカメラ54とを含む。

10

20

30

40

50

## 【 0 0 3 3 】

図 2 に、コンピュータ 4 0 の内部構成をコンピュータシステム 2 0 の他の構成要素とともに図示する。図 2 を参照して、コンピュータ 4 0 は、CPU (Central Processing Unit) 7 0 と、CPU 7 0 に接続されたメモリ 7 2 と、CPU 7 0 に接続されたハードディスク 7 4 とを含む。CPU 7 0 は、図示されないネットワークボードを介してネットワーク 8 0 に接続され、このネットワーク 8 0 には、アイカメラ 5 4 を制御する前述したコンピュータ 8 2 が接続されている。コンピュータ 8 2 の構成はコンピュータ 4 0 と同様であるので、ここではその詳細は説明しない。

## 【 0 0 3 4 】

以下に詳細に説明する、本発明にかかる目の位置のトラッキング方法および装置は、図 1 および図 2 を参照して説明したコンピュータハードウェアと、CPU 7 0 によって実行されるソフトウェアとにより実現される。このソフトウェアは、CD-ROM などの記憶媒体に記憶されたり、ネットワークを介したりして市場を流通するものである。このソフトウェアがこのコンピュータ 4 0 にインストールされると、典型的にはソフトウェアはハードディスク 7 4 に格納される。そして、利用者がこのソフトウェアを起動すると、またはコンピュータ 4 0 のオペレーティングシステム (OS) により起動されると、ハードディスク 7 4 からメモリ 7 2 に読み出され、CPU 7 0 によって実行される。

## 【 0 0 3 5 】

なお、このソフトウェアが、もともとコンピュータ 4 0 の OS の一部として組み込まれている機能を利用したり、または OS とは別にインストールされた別のソフトウェアの一部の機能を利用することもあり得る。したがって、本発明を実現するためのソフトウェアとして流通するのは、そうした一部の機能を欠いたものであってもよく、実行時にそうした機能が他のソフトウェアによって実現されることを考慮した構成となっていればよい。もちろん、必要な機能がすべてあらかじめ組み込まれたシステムソフトウェアとして流通してもよい。

## 【 0 0 3 6 】

## 〔原理〕

本発明において、目の位置を検出するための方法の基本は、フレーム画像の差分に基づき、特に瞬きによる明度の差に着目したものである。フレーム間差分から瞬きを検出しようとするとき、顔が動いていると、動いた瞬き部分以外にも明度変化の大きな画素が多くの部分で生じる。そのために、目蓋の動きによる明度変化と、顔全体の動きによる明度変化とを区別する必要がある。なお、口形状の変化や眉の動きなどによっても問題が生じる可能性があるが、それらの明度変化は、目候補点を抽出した後、条件検定によって棄却する。

## 【 0 0 3 7 】

顔全体の動きは、画像平面内の平行移動、画像平面内の回転 (首をかしげる)、画像平面外の回転 (振り向く、頷く) に分けて考えることができる。フレームレートは十分に速くて、フレーム間での動きは小さいものとする。以下、これらについて順にそれらの動きをキャンセルする方法について説明する。

## 【 0 0 3 8 】

顔が画像平面内で平行移動したときの動きをキャンセルする方法は以下のとおりである。図 3 (A) に示される、画像  $f_{t-1}$  に写っている色の付いた円板 9 0 が、右方向に距離  $d$  だけ移動して、次のフレームで図 3 (B) に示される画像  $f_t$  のようになった場合を考える。円板には白い小円マークが付いている。背景は一様でなくともよいが、図示はしていない。

## 【 0 0 3 9 】

図 3 (A) および図 3 (B) の画像の差分画像をとると図 4 に示したとおりの画像となる。図 4 において、斜線で示した領域が画素値の差が大きく現れた部分であり、白い領域は画素値の差が小さい領域である。図 4 に示される差分画像  $D$  の中で、画素値の差が大きく現れた画素位置を、図 3 (A) または (B) に示す元の画像に立ち返って考えてみると、

10

20

30

40

50

これら画素位置はさらに図5の上段の左右に示された領域FおよびBに分類できる。領域FのFはForeground、すなわち円板上の点であり、領域BのBはBackground、すなわち背景に対応する。領域Bは一方の画像で見えていて、他方の画像では見えていない部分である。領域Fは両方の画像で見えていて、互いに対応する画素がある部分である。ただし、実際の画像では、移動物体がどの部分であるかは不明なので、どこが領域Fでどこが領域Bかは区別できない。

#### 【0040】

ここで、図5の上段右に示されている画像 $f_t$ 上の、領域FとBとに相当する画素を $-d$ だけシフトして画像 $f_{t-1}$ に重ねてみる。すると、図5の下段左に示されているように、画素値が一致する領域Mの部分と、一致するもしないともいえない領域Uの部分とがあることがわかる。同様に画像 $f_{t-1}$ 上の領域FとBとに相当する画素を $+d$ だけシフトして画像 $f_t$ に重ねてみると、領域MとUとは図5の下段右に示したようになる。なお、ここでいう移動量 $\pm d$ は二次元ベクトル量である。

10

#### 【0041】

ここで、シフト量が実際のシフト量 $\pm d$ とは異なるシフト量 $\pm d'$ であるとすると、領域Mに属する画素数は、実際のシフト量 $\pm d$ を用いた場合よりも少なくなると考えられる。そこで、領域Mの画素数が最大となるシフト量を探索すれば、それがフレーム間での円板90の移動量 $d$ であるといえる。領域Mの画素数の計数にあたっては、画像 $f_{t-1}$ の画素のシフト量を $d$ とすれば画像 $f_t$ のシフト量が $-d$ となるように、両者が常に逆方向で同じ大きさでなければならない。

20

#### 【0042】

このように、フレーム間の差分画像で抽出された画素のうち、領域Mに属する画素数が最大となるシフト量 $\pm d$ を見出して画像をシフトさせて重ね合わせ、なおかついずれでも画素値が一致しない画素を抽出すると、それは移動量 $d$ での平行移動以外の動きをした部分であると言える。顔全体の移動量を $d$ とすれば、 $d$ 以外の動きをした部分は、瞬き、口形状の変化などに由来する部分である。したがってフレーム間差分画像から顔全体の動きをキャンセルすることにより、瞬きで動く部分、すなわち目位置を抽出することができる。

#### 【0043】

次に振り向く動きについて考える。図6を参照して、カメラ56、58を向いている顔が左に振り向こうとしている場合を考える。この場合、顔の右端(図6における左端)からは隠れていた部分が出現し、顔の左端(図6における右端)では、見えていた部分がセルフオクルージョンで隠れる。このような部分は、フレーム間差分を計算する2枚の画像の一方では隠れているので、平行移動のときとは異なりシフトして画素値の一致を見ることは意味がない。

30

#### 【0044】

しかしこのような部分は、移動領域の周辺部にしか現れない。したがって、フレーム間差分画像(図4)を計算した後、斜線を引いた領域の左右から一定数の画素をそれぞれ削り去ることで、その影響を除去することができる。他方、顔の中心部は、振り向く動作の場合でも平行移動とみなせるので、平行移動に関する前述の手法(2枚の画像を $\pm d$ だけ移動して一致する画素を見る手法)を適用することができる。

40

#### 【0045】

なお、首をかしげる動きと頷く動きとは、回転中心が顔からはずれているため、平行移動とみなすことができる。ただし、いずれの場合も、フレーム間において動きが小さいことが前提である。

#### 【0046】

##### [ソフトウェアの制御構造]

上記した原理に基づく目の検出処理およびトラッキング処理を実現するためのソフトウェアの構造について以下説明する。なお、前述したとおり、以下に述べるソフトウェアの実行時には、OSが持つ機能、またはすでにコンピュータにインストールされている機能を利用することが前提となっている場合があり、その結果、流通するソフトウェアがそうし

50

た機能自体を含まないことがあり得る。しかし、その場合でも、そうした機能を制御構造のどこで利用するかについての情報をもつものである限り、そうしたソフトウェアは本願発明の技術的範囲に属するものである。

【0047】

なお、以下に述べる処理では、すでに述べた目の候補点の検出処理に加えて、目が二つあることを前提に、瞬きは左右の目で同時に起きる、両目間の距離はある範囲に入っている、顔の傾き角度は一定範囲内である、両目の中点近傍の濃淡パターンはほぼ左右対称である、などの条件を用いることで、抽出された目の位置の候補点对が目であるか否かを検出している。

【0048】

また、以下の処理では種々のしきい値が用いられているが、その値は応用により、また求められている精度により経験的に定められるものである。

【0049】

図7を参照して、このソフトウェアの全体構成は以下のようになっている。まず、図8以下を参照して後に詳細に説明する手順を用いて、目の位置を抽出する(ステップ100)。続くステップ102で、抽出された目の位置をアイカメラ54を制御するコンピュータ82に送信する。そして、引き続きステップ104で目の位置のトラッキング処理を行なう。目位置のトラッキングができたか否か、すなわち目の位置を見失ったか否かを判断する(ステップ106)。トラッキングが失敗した場合には制御はステップ100に戻り、再度ステップ100の目の位置の抽出処理から実行する制御が行なわれる。トラッキングができていれば、制御はステップ102に戻り、トラッキングによって得られた目の方向、目までの距離などの情報がコンピュータ82に送信される。以下、ステップ100から106の処理を繰り返し行なうことで、コンピュータ82に対して利用者の目の方向および距離に関する情報を絶えず送信することができる。

【0050】

なお図7のステップ100において目の位置が抽出されない場合もあり得る。その場合は、再度ステップ100の処理を実行して、次のフレームに対して目の位置の抽出処理を実行し、その後再度トラッキング処理を開始する。

【0051】

図8を参照して、図7のステップ100で行なわれる目の位置の抽出処理について説明する。以下の処理は、図4から図6を参照して説明した原理を用いたものである。なおこの処理の前に、前画像がキャプチャされていることが前提である。特に処理の最初では、前画像を現画像と一致させるようにしてもよい。まず、ステップ120でカメラ56, 58からの現画像をキャプチャする。以下の処理は実際には二つのカメラ56の画像に対して行なわれ、目の検出後、カメラ58の画像とあわせてステレオ処理により目までの距離を計算する。カメラ56とカメラ58との役割を交代させてもよい。

【0052】

続いてステップ122で、前画像と現画像とのフレーム間差分2値化画像Dを計算する。より具体的には、各画素ごとに、前画像と現画像との値の差の絶対値を求め、その値があらかじめ定められたしきい値N以上であればその画素値を「1」とし、しきい値N未満であれば画素値を「0」とする。ビデオ信号は、静止物体を撮像している場合でもある程度変動しているので、このようなしきい値Nを用いることが必要である。このしきい値Nはノイズレベルと考えることができる。

【0053】

続いてステップ124で、差分画像D中の画素値「1」の数があらかじめ定められたしきい値G以上か否かを判定する。判定結果が「YES」であれば制御は図9に示すステップ168にとび、目は検出されなかったとされ処理を終了する。この場合は、あまりにも多くの画素に動きが検出されたため、想定以上の動きがあるものとして、次のフレーム入力を待つ。

【0054】

10

20

30

40

50

ステップ124での判定結果が「NO」の場合、ステップ126で差分画像D中の画素値「1」の数がしきい値E以下か否かを判定する。このしきい値Eは前述のしきい値Gよりも小さな値であり、顔が実質的に静止しているか否かを判定するためのものである。ステップ126の判定結果が「YES」の場合、制御は図9に示すステップ152に飛ぶ。この場合には、フレーム間差分で抽出された画素数が少なく、顔が停止状態であると考えられる。

**【0055】**

ステップ126での判定結果が「NO」の場合、制御はステップ128に進む。ステップ128では、差分画像Dの画素値「1」である画素群に対し、その両端の画素を削除する処理が行なわれる。この処理は、図6を参照して説明したように振り向く動作を考慮したものである。

10

**【0056】**

より具体的には、差分画像Dの各走査線ごとに左右の端からそれぞれ右方向および左方向にサーチして、最初に現れる画素値「1」の画素を探す。そしてその画素からそれぞれk個の画素の画素値を「0」に書き換える。この処理により、画像中の画素値「1」の左右領域の所定分が削り去られることになる。

**【0057】**

続いてステップ130で、前画像と現画像とから、差分画像Dの画素値「1」に対応する画素を抽出する。この処理は、図5の中段に示した画像を得る処理である。この後ステップ132で、前述した移動量 $\pm d$ を算出する処理が行なわれる。この処理の詳細については図10を参照して後に詳述する。

20

**【0058】**

移動量 $\pm d$ が算出されると、続いてステップ150(図9)で図5の下段に示す領域Mに相当する画素を差分画像Dから削除する処理が行なわれる。さらにステップ152で、差分画像Dに対して孤立点除去および平滑化処理が実行される。より具体的には、処理対象の画素の8近傍と、中心画素とのうち合計して4画素以上の画素値が「1」であればその処理対象の画素の画素値を「1」とする。さもなければ画素値を「0」とする。この処理により、孤立点の画素値は「0」となって差分画像D上から削除される。

**【0059】**

続いて、ステップ154で、差分画像Dに対してラベリング処理を施す。この処理により、互いに連続した画素値「1」の画素からなる領域の各々についてラベルが付与される。こうして得られた各要素の中心を、目位置の候補点とする(ステップ156)。

30

**【0060】**

得られた候補点の数が2未満か否かを判定する(ステップ158)。目位置は一般的に2個であることが想定されるから、候補点の数が2未満の場合には目は検出されなかったとして、処理はステップ168に進む。候補数が2以上であれば制御はステップ160に進む。

**【0061】**

ステップ160では、各候補点ごとに、候補点を中心とする矩形 $a \times b$ の領域の各画素の、前画像の画素値と現画像の画素値との差の絶対値の総和Sを求める。ここで矩形 $a \times b$ とは、想定される目の大きさに対応する大きさの矩形である。前画像と現画像との間で瞬きがあれば、一方の画像では目の画像であり他方の画像では目蓋の画像となるためその画素値の差の絶対値は大きくなり、かつその大きさは想定される目の大きさに近いはずである。したがって、このように各候補点についてSを求め、その値の大きなものを二つを目の候補点(候補点对)とすることができる。

40

**【0062】**

続いてステップ162で、以上のようにして求められた二つの候補点が以下の諸条件を満足するか否かが判定される。全条件を満足しなければ、目は検出されなかったとして(ステップ168)、処理を終了する。

**【0063】**

50

1) ステップ160で求めた差の絶対値の総和 $S$ が、二つの候補点のいずれにおいても所定のしきい値 $F$ 以上である。

【0064】

2) 候補点間の距離があらかじめ定められた値 $L_{min}$ 以上である。

3) 候補点間の距離があらかじめ定められた値 $L_{max}$ 以下である。

【0065】

4) 候補点を結ぶ直線と走査線方向とのなす角度が、あらかじめ定められたしきい値 $A$ 度以下である。

【0066】

しきい値 $F$ はあらかじめ実験的、または統計的に定められる値であり、目以外の部分が目として抽出されるのを防止するために用いられる。 $L_{min}$ および $L_{max}$ はそれぞれ、目の中心間の距離として想定される最小値および最大値である。これらの値はあらかじめ統計的に定められる。またしきい値角度 $A$ 度も実験的に適切な値が定められるものであるが、通常は30度から45度程度の値が使用される。

10

【0067】

上述の4つの条件がすべて充足された場合、さらにステップ164で、求められた二つの目候補点の midpoint 近傍の濃淡パターンがほぼ左右対称か否かを判定する。ほぼ左右対称であると判定されたら、この二つの目候補点を目の位置とする。左右対称でなければ目を見つけられなかったとして処理を終了する。

【0068】

二つの目候補点の midpoint 近傍のパターンがほぼ左右対称か否かを判定する方法について説明する。図11を参照して、一般的に両目の midpoint (これを以下「眉間」と呼ぶ。)を中心とする領域220において、眉間の左右の濃淡パターンは、二つの目候補点を結ぶ線分に垂直で、眉間(二つの目候補点の midpoint)をとる直線に関してほぼ左右対称となるはずである。そこで、まず現画像を、両目の候補点の midpoint を中心として、両目候補が走査線に水平に並ぶように回転し、候補点の midpoint を中心とする $g \times h$ 画素(領域220とする。)を切出す。図11の下段に示されるように、この領域220の縦 $h$ 画素の画素値の総和を各列ごとに求め、さらに全画素の総和に対する各列の総和を百分率であらわした値を計算することにより、図11に示すグラフ222のような投影プロフィールが得られる。そして、このプロフィールのうち、両候補の midpoint を中心として、左右対称となる互いに隣接する3列ずつの合計を左右双方について計算する。そして、左右の合計の差の総和を、隣接する3列のすべての組み合わせに対して計算し、その値がすべてしきい値 $p$ 以下であれば、二つの目候補点の midpoint 近傍のパターンが、ほぼ左右対称であると判定する。

20

【0069】

具体的には、図11の下段を参照して、たとえば最左端の3列224Lの値の合計と、最右端の3列224Rの値の合計とを計算し、その差 $p_1$ を求める。同様に、左から2列目~4列目の3列226Lの値の合計と、右から2列目~4列目の3列226Rの値の合計とを計算し、その差 $p_2$ を求める。左から3列目~5列目の3列228Lの値の合計と、右から3列目~5列目の3列228Rの値の合計を計算し、その差 $p_3$ を求める。以下同様にして値 $p_4 \sim p_{g-2}$ を計算し、 $p_1 \sim p_{g-2}$ の値の合計を求め、その合計としきい値 $p$ とを比較することで、二つの目候補点の midpoint 近傍の濃淡パターンがほぼ左右対称か否かを判定できる。

30

40

【0070】

なお、上の説明では、二つの目候補点を水平走査線に並ぶようにその midpoint を中心としていったん回転してから濃淡パターンが左右対称か否かを判定した。しかし左右対称の判定方法はこれに限定されるわけではない。たとえば、二つの目候補点を結ぶ線分が水平走査線に対してなす角度がわかれば、その線分に対し垂直な方向で画素値の合計を計算することにより、図11に示すのと同様の処理を実現することができる。また、上の例では各列ごとに合計をとるようにしたが、 midpoint を中心として左右対称な画素ごとに差を計算してその総和を計算し、それがしきい値より小さい場合に左右対称と判定してもよい。

50

## 【0071】

ここで再び図8のステップ132に戻り、移動量 $d$ を算出する処理について、図10を参照して詳細に説明する。移動量 $d$ (二次元ベクトル量) $= (x, y)$ とし、 $x = -m, -m+1, \dots, m, y = -n, -n+1, \dots, n$ の範囲で図5の下段に示すような領域 $M$ に属する画素数の合計の最大値を示すものを探索するものとする。まず、変数 $max$ に初期値0を代入する(ステップ180)。この変数 $max$ は、以下に述べる計算過程で、個々の移動量に対して算出される差分画像のうち、領域 $M$ に属する画素数の合計の最大値を記憶するためのものである。

## 【0072】

変数 $x$ に $-m$ を代入し(ステップ182)、変数 $y$ に $-n$ を代入する(ステップ184) 10  
。以下、 $y$ の各値ごとの繰り返し処理である。

## 【0073】

まず、 $d = (x, y)$ に対する領域 $M$ に属する画素数の合計 $M(d)$ を求める(ステップ186)。そしてこの $M(d)$ が変数 $max$ より大きいか否かを判定する(ステップ188)。 $M(d)$ が変数 $max$ 以下の場合には何もせず制御はステップ192に進む。 $M(d)$ が変数 $max$ より大きい場合には、ステップ190で変数 $max$ に $M(d)$ の値を代入し、 $X, Y$ にそれぞれそのときの $x, y$ の値を代入する。制御はステップ192に進む。

## 【0074】

ステップ192では $y$ に1を加算し、その結果 $y$ が $n$ を超えたか否かを判定する(ステップ194) 20  
。  $y$ が $n$ を超えていない場合には制御はステップ186に戻り、新たな $y$ に対してステップ186~194の処理を繰り返す。 $y$ が $n$ を超えている場合には、制御はステップ196に進む。

## 【0075】

ステップ196では、 $x$ に1を加算する。そして、その結果 $x$ が $m$ を超えたか否かを判定する(ステップ198)。 $x$ が $m$ を超えていないと判定された場合には制御はステップ184に戻り、以下新たな $x$ に対してステップ184~198の処理を繰り返す。 $x$ が $m$ を超えたと判定された場合には、制御はステップ200に進み、変数 $max$ の値と、この変数 $max$ の値が得られたときの $x, y$ の値である $X, Y$ を出力して処理を終了する。

## 【0076】

このようにして、領域 $M$ に含まれる画素数が最大となる移動量 $d$ を求める。なお、移動量 $d$ の算出方法はこれ以外にも種々考えられる。たとえば、 $x, y$ を $m \times n$ の範囲で変化させながらあらかじめ総当りで領域 $M$ に含まれることになる画素数を求めてテーブル化し、そのテーブルの中で最大値を示すセルを求めるようにしてもよい。

## 【0077】

さて、上で述べたようにして検出した目を次のフレームから信頼性高く、かつロバスト性高くトラッキングしていくために、次のようなトラッキング方法をソフトウェアにより実現する。本実施の形態では、目をトラッキングするのに、直接ではなく、目と目の間(眉間)のパターンをトラッキングする。そして、そのようにトラッキングされた眉間からの相対的位置が前フレームの目位置と同じ点の近傍で目を探索する。 40

## 【0078】

眉間パターンは、顔の表情が変化する場合でも比較的安定している。また、額部と鼻筋の明るい部分に両サイドから目と眉とが楔のように割込んだ濃淡パターンを形成しているので、パターンマッチングによって位置を決める際の精度を出しやすい。

## 【0079】

以下に述べるように、眉間のテンプレートを用いてパターンマッチングする際に、テンプレートを更新する必要がある。テンプレートとして本実施の形態では、両目の中心を中点とする矩形パターンを採用し、テンプレートの更新には現フレームで検出された矩形パターンを用いる。

## 【0080】

以下、図12を参照してトラッキング処理の詳細について説明する。なお、前フレームで抽出された両目の中点位置と、眉間パターン（ $s \times t$ 画素）と、中点から見た右目と左目との相対位置  $e_r$ 、 $e_l$ （ $e_r = e_l$ で、いずれも2次元量）とがすでにセーブしてあるものとする。

【0081】

まず、ステップ240でセーブされている中点位置、眉間パターン、および両目の相対位置を初期値として取り込む処理を行なう。続いて、現画像（現フレーム）の取り込みを行なう（ステップ242）。

【0082】

次のステップ244では、前画像と、さらにその前の画像（前々画像）とにおける両目の中点位置から、現画像における両目の中点位置を予測する。すなわち、前フレーム、前々フレームの眉間位置をそれぞれ  $X_{t-1}$ 、 $X_{t-2}$  とする（ただし  $X_{t-1}$ 、 $X_{t-2}$  はいずれも2次元量）と、現フレームでの予測位置  $X_t$  は、 $X_t = 2X_{t-1} - X_{t-2}$  で外挿できる。ただし、最初の検出時は  $X_{t-1} = X_0$  とする。このように現画像における両目の中点位置を予測するのは、顔画像が移動している場合、その移動量はほぼ一定であると考えられるので、移動後の中点位置を予測してから眉間パターンのマッチングをすると効率がよいからである。

【0083】

ステップ246では、ステップ244で予測された中点位置の近傍で、セーブされていた眉間パターンと最もよく一致する眉間パターンを探すマッチング処理を実行する。ここで最もよいマッチが得られた位置を  $X_{t0}$  とする。そして、ステップ248で、 $X_{t0} + e_r$  および  $X_{t0} + e_l$  を中心とする近傍（ $i \times j$ 画素）において右目および左目の位置をそれぞれ探索する（ステップ248）。探索にあたっては、その点を中心とする  $5 \times 5$  画素の平均画素値が最も暗い点を目と判定する。

【0084】

次に、探索結果の目の位置が、図9のステップ162で用いられた条件2)3)4)を満たすか否かを判定する。いずれかひとつの条件でも満足されていなければ、目のトラッキングを誤ったか見失ったものとして、トラッキング失敗の判定をし（ステップ252）処理を終了する。この場合、図7のステップ106から制御はステップ100に戻り、再び目の位置の検出から処理が再実行される。

【0085】

すべての条件が満足されていると判定された場合には、トラッキングが成功したものと判定され、トラッキングにより得られた両目の中点位置と、その点を中心とする眉間パターン（ $s \times t$ 画素）と、中点に対する両目の相対位置  $e_r$ 、 $e_l$  とをセーブ（更新）してトラッキング処理を終了する。

【0086】

なお、このようにしてカメラ56の画像に対して両目の位置が検出されれば、カメラ58の画像とあわせてステレオ処理により目までの距離が求まり、アイカメラ54を制御するためのパラメータ（距離、方向）が計算により決定できる。その処理については周知の技法を適用することができるので、ここでは詳細な説明は繰り返さない。

【0087】

[動作]

上にその構造について説明した本願発明にかかる目の位置の検出装置は以下のように動作する。図1に示すステレオカメラ56、58は、利用者の顔画像を撮影しそれぞれビデオ信号を出力しコンピュータ40に与える。

【0088】

図8に示すように目の位置の検出およびトラッキングのためのソフトウェアを起動すると、カメラ56、58から得られたビデオ画像の各々に対して目の位置が抽出される。具体的には、現画像がキャプチャされ（ステップ120）、差分画像Dの計算（ステップ122）、ステップ124および126の判断が行なわれる。ステップ124または126において判定結果が「YES」となれば目の位置の検出ができなかったとして、次のフレー

10

20

30

40

50



△画像に対して再度目の位置の抽出処理が実行される。

【0089】

ステップ124および126における判定結果がいずれも「NO」となった場合には、ステップ128で画素値「1」の領域の両端の画素が削除される。そして、既に説明した手法にしたがって移動量dを計算した後(ステップ132)、このようにして得られた移動量dを用いて図5に示す移動および重ねあわせを行なって得られた領域Mに相当する画素を、差分画像Dから除去する(ステップ150)。

【0090】

続いて、孤立点除去、平滑化処理(ステップ152)、ラベリング処理(ステップ154)を経て得られた各要素の中心を目位置候補とする(ステップ156)。この候補数が2 10未満の場合には抽出は失敗したものと判定され、最初から処理が再開される。候補数が2以上の場合には、その近傍における画像間の差分の絶対値の総和が最も大きな二つの点が候補点となる(ステップ160)。この候補点が前述した4つの条件のすべてを満足し(ステップ162)、かつ中点近傍のパターンが左右対称であると判定される(ステップ164)と、その二つの候補点を目の位置として、目の位置の抽出処理は終了する。ステップ162、164のテストに失敗すると目の位置の抽出はできなかったものとして、最初から処理が再開される。

【0091】

こうしてステレオカメラ56, 58の双方について、目の方向が決定されると、それらの 20情報を用いてアイカメラ54を制御すべき情報(アイカメラ54から見た目の方向および目までの距離)が得られる。こうした情報をコンピュータ82に送信して(図7のステップ102)、制御は目位置のトラッキング処理に移る(ステップ104)。

【0092】

トラッキング処理の最初には、目の位置の抽出処理で得られた両目の中点位置(すなわち眉間の位置)、眉間の画像パターン、および中点位置から見た両目の相対位置に関する情報がセーブされている。まず最初に、このようにセーブされている情報を初期値として取り込む(ステップ240)。続いて現画像を取込み(ステップ242)、前画像における目の中点位置(眉間の位置)および前々画像における目の中点位置に基づいて現画像での目の中点位置を予測する(ステップ244)。この予測は、たとえば前々画像と前画像との中点位置を外挿することにより行なわれる。予測された中点位置近傍で、既にセーブされ 30ていた前画像での眉間パターンと最もよくマッチするパターンの中心となる位置を探索し、最もよいマッチを示した位置を中心として、両目の位置を探索する(ステップ248)。得られた両目位置が所定の条件を満足していれば(ステップ250でYES)、その両目位置に基づいて両目の中点位置を算出し、その値と、その中点位置を中心とする眉間パターンと、中点位置から両目までの相対位置とをセーブして、トラッキングを一旦終了する。ただしこの場合、図7においてステップ106からステップ102の経路を通り、次のトラッキング処理が行なわれる。

【0093】

トラッキングに失敗すると、図7のステップ106からステップ100の経路を通り、再度目の位置の抽出処理から実行される。 40

【0094】

こうして、ステレオカメラ56, 58を用いて目の位置(方向)を抽出し、トラッキングし、それらの値を用いることによりアイカメラ54を制御するためのパラメータ(目の方向、距離)を計算してアイカメラ54を制御することができる。

【0095】

図13~図15に、本実施の形態による目の位置の抽出およびトラッキング処理の具体例について、画面表示を示す。

【0096】

図13には、入力された画像例を示す。図14に示すのは、この画像と前画像との差分画像である。画面の左上には左右の対称性を判定するために切出したパターン270が、左 50

下には入力画像の眉間パターン 272 が、それぞれ表示されている。差分画像の中で、目位置候補として 274 および 276 が抽出され白丸で表示されている。

【0097】

図15には、両目位置 292, 294 および中心位置 290 を画像に重ねて表示した例を示す。図15に示すように、眉間を中心として、両目の位置が正しくトラッキングされている。

【0098】

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。

10

【図面の簡単な説明】

【図1】 本願発明にかかる目の位置のトラッキング方法および装置を実現するコンピュータシステム 20 の外観図である。

【図2】 図1に示すコンピュータシステム 20 および周辺装置のブロック図である。

【図3】 本発明の一実施の形態における目の位置の抽出の原理を示すための図である。

【図4】 差分画像の一例を示す図である。

【図5】 本発明の一実施の形態における目の位置の抽出の原理を示すための図である。

【図6】 本発明の一実施の形態において、利用者が振り向く動きをキャンセルする原理を示す図である。

20

【図7】 本発明の一実施の形態における目の位置の抽出方法および装置を実現するソフトウェアの全体の制御構造を示すフローチャートである。

【図8】 本発明の一実施の形態における目の位置の抽出処理を実現するソフトウェアのフローチャートである。

【図9】 本発明の一実施の形態における目の位置の抽出処理を実現するソフトウェアのフローチャートである。

【図10】 本発明の一実施の形態における目の位置の抽出処理において、移動量  $d$  の算出処理を実現するソフトウェアのフローチャートである。

【図11】 本発明の一実施の形態における、眉間の近傍の左右対称性を判定する処理を説明するための図である。

30

【図12】 本発明の一実施の形態における目のトラッキング処理を実現するソフトウェアのフローチャートである。

【図13】 本発明の一実施の形態における現画像の表示例を示す図である。

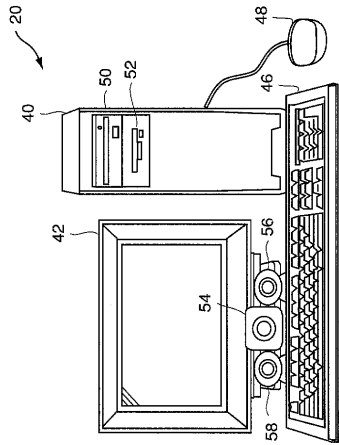
【図14】 本発明の一実施の形態における差分画像の表示例を示す図である。

【図15】 本発明の一実施の形態における目の位置のトラッキングの表示例を示す図である。

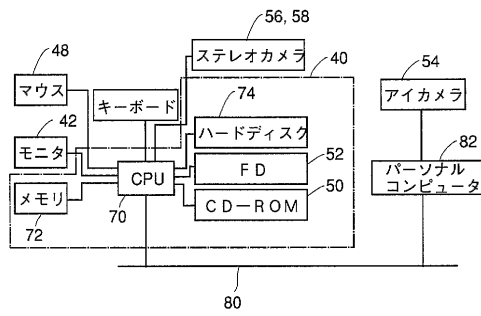
【符号の説明】

20 コンピュータシステム、40, 82 コンピュータ、42 モニタ、46 キーボード、48 マウス、54 アイカメラ、56, 58 ステレオカメラ。

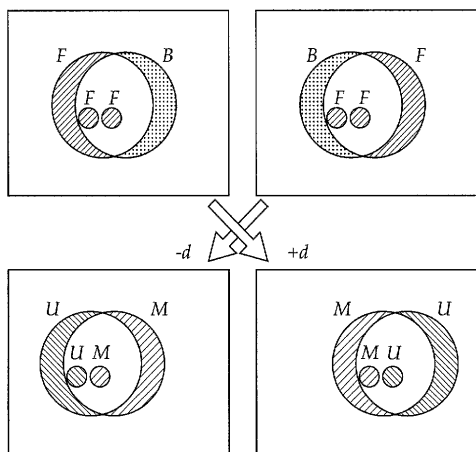
【図1】



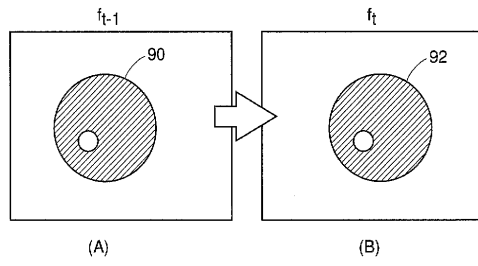
【図2】



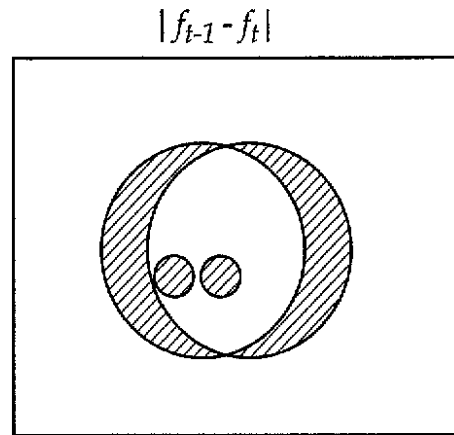
【図5】



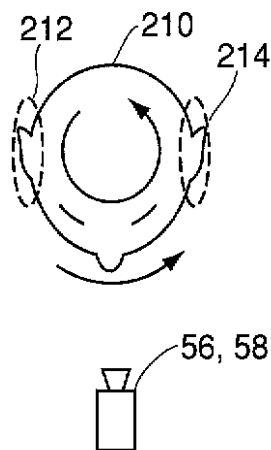
【図3】



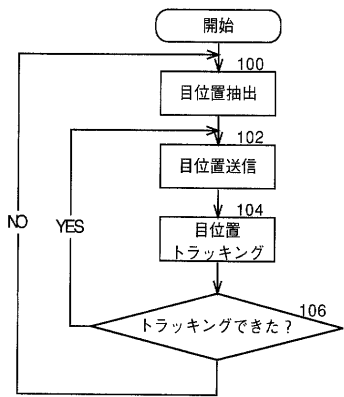
【図4】



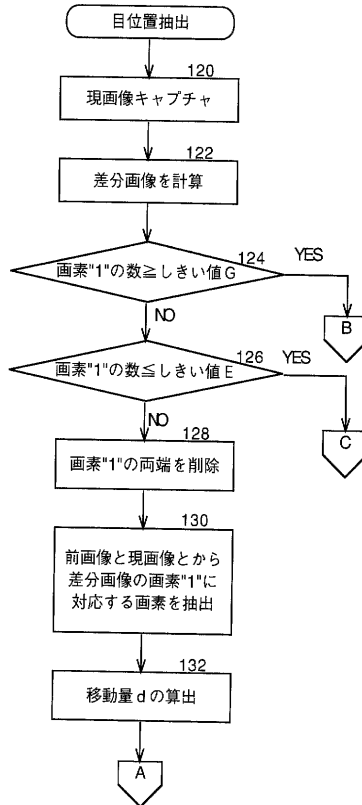
【図6】



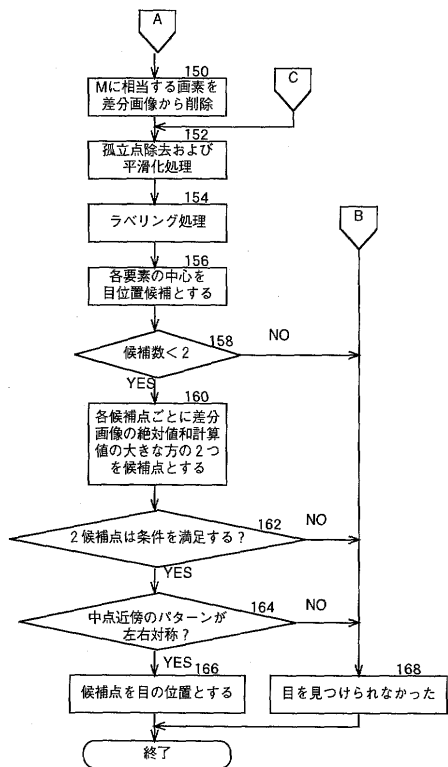
【 図 7 】



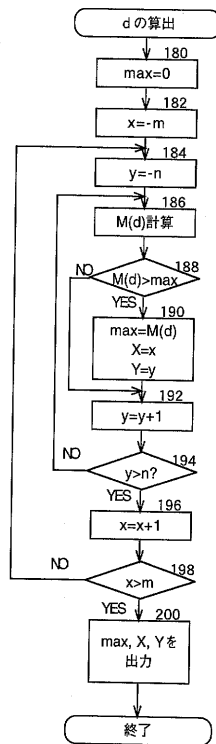
【 図 8 】



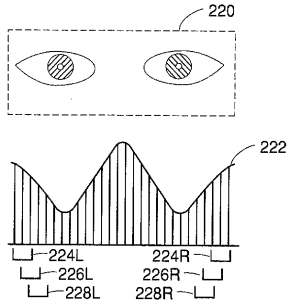
【 図 9 】



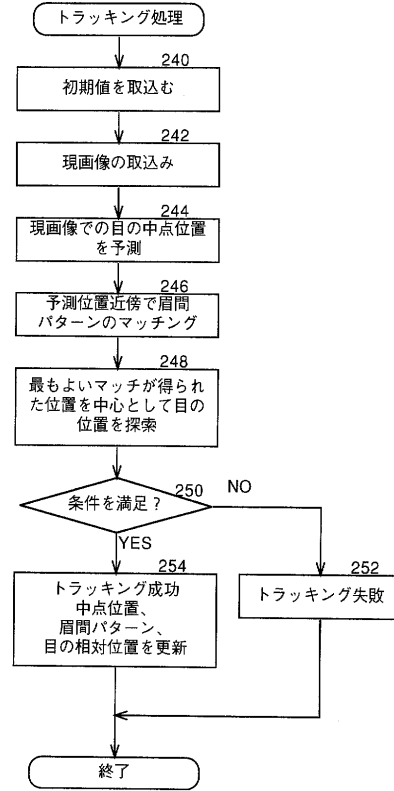
【 図 10 】



【 図 1 1 】



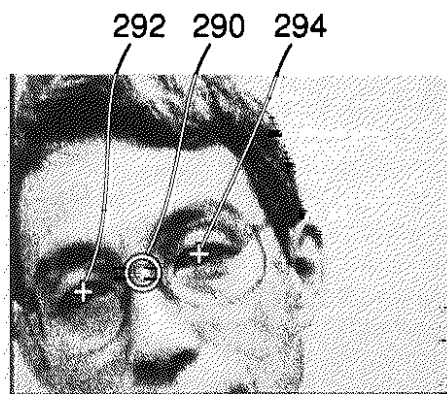
【 図 1 2 】



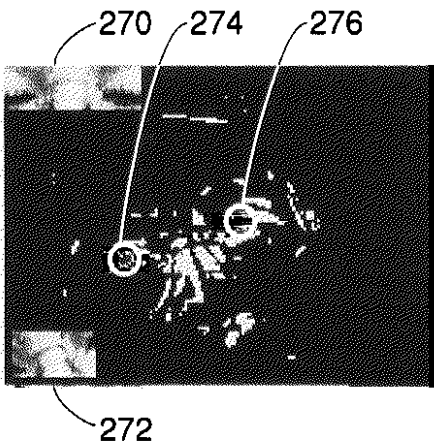
【 図 1 3 】



【 図 1 5 】



【 図 1 4 】



## 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 鉄谷 信二  
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 広 島 明芳

- (56)参考文献 特開平06-078901(JP,A)  
特開平11-286227(JP,A)  
川戸慎二郎、鉄谷信二、目のリアルタイム検出と追跡、電子情報通信学会技術研究報告、日本、社団法人電子情報通信学会、2000年9月14日、Vol.100、No.311、p15-22  
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、電子情報通信学会誌、日本、社団法人電子情報通信学会、2001年12月1日、Vol.J84-D-2、No.12、p2577-2584

## (58)調査した分野(Int.Cl., DB名)

G06T 1/00  
A61B 3/113  
G06T 7/00  
G06T 7/20  
G06T 7/60  
JSTPlus(JDream2)