

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第4692526号
(P4692526)

(45) 発行日 平成23年6月1日(2011.6.1)

(24) 登録日 平成23年3月4日(2011.3.4)

(51) Int. Cl. F I
G06T 7/60 (2006.01) G O 6 T 7/60 I 5 O P
A 6 1 B 3/113 (2006.01) A 6 1 B 3/10 B

請求項の数 12 (全 46 頁)

| | | | |
|--------------|-------------------------------|-----------|--|
| (21) 出願番号 | 特願2007-185996 (P2007-185996) | (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 |
| (22) 出願日 | 平成19年7月17日(2007.7.17) | (74) 代理人 | 110001195 特許業務法人深見特許事務所 |
| (65) 公開番号 | 特開2008-102902 (P2008-102902A) | (72) 発明者 | 内海 章 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 |
| (43) 公開日 | 平成20年5月1日(2008.5.1) | (72) 発明者 | 山添 大丈 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 |
| 審査請求日 | 平成19年7月17日(2007.7.17) | 審査官 | 鹿野 博嗣 |
| (31) 優先権主張番号 | 特願2006-195871 (P2006-195871) | | |
| (32) 優先日 | 平成18年7月18日(2006.7.18) | | |
| (33) 優先権主張国 | 日本国(JP) | | |
| (31) 優先権主張番号 | 特願2006-254531 (P2006-254531) | | |
| (32) 優先日 | 平成18年9月20日(2006.9.20) | | |
| (33) 優先権主張国 | 日本国(JP) | | |

最終頁に続く

(54) 【発明の名称】 視線方向の推定装置、視線方向の推定方法およびコンピュータに当該視線方向の推定方法を実行させるためのプログラム

(57) 【特許請求の範囲】

【請求項1】

画像を獲得するための撮影手段と、
 前記撮影手段により撮影された人間の顔領域を含むフレームの画像に基づき、前記顔領域内の複数の特徴点の間の相対的な3次元の位置関係を特定する相対関係特定手段と、
 前記人間の眼球中心の3次元の位置である眼球中心位置の推定処理を実行する眼球中心推定手段とを備え、
 前記眼球中心推定手段は、予め特定された前記特徴点と前記眼球中心の位置との相対的な3次元の位置関係を用いて、当該推定処理の対象の前記フレームにおいて前記相対関係特定手段により特定された前記複数の特徴点間の相対的な3次元の位置関係に基づいて、

10

前記推定処理の対象の前記フレームの画像領域内において、虹彩の領域と前記虹彩の形状のモデルとを照合することで、虹彩中心が前記画像に投影される虹彩中心位置を抽出する虹彩中心抽出手段と、

前記推定処理の対象の前記フレームの画像において抽出された前記虹彩中心位置と前記眼球中心の位置とに基づき、前記眼球中心と前記虹彩中心とを結ぶ3次元直線の方向として、視線方向を推定する視線推定手段とをさらに備える、視線方向の推定装置。

【請求項2】

前記撮影手段は、人間の顔領域を含む対象画像領域内の各画素に対応する画像データを撮影して獲得するための単眼の撮影手段である、請求項1に記載の視線方向の推定装置。

20

【請求項 3】

前記相対関係特定手段は、較正時において、前記人間が前記撮影手段を見ている状態で前記撮影手段により撮影された複数の較正用画像を予め取得し、前記特徴点と前記眼球中心との間の相対的な 3 次元の位置関係を特定し、

前記眼球中心推定手段は、前記撮影手段により撮影された前記人間の顔領域を含む対象画像領域内において検出された前記複数の特徴点の投影位置により、特定された前記相対的な 3 次元の位置関係に基づいて、前記人間の眼球中心の投影位置を推定する、請求項 1 または 2 記載の視線方向の推定装置。

【請求項 4】

前記相対関係特定手段は、

複数の前記較正用画像内の前記複数の特徴点の投影位置を抽出し、前記複数の特徴点の投影位置を要素として並べた計測行列を算出する計測行列算出手段と、

前記計測行列を因子分解により、前記撮影手段の姿勢に関する情報を要素とする撮影姿勢行列と、前記複数の特徴点間の相対的な 3 次元の位置関係に関する情報を要素とする相対位置関係行列とに分解する因子分解手段とを含む、請求項 3 記載の視線方向の推定装置。

【請求項 5】

前記眼球中心推定手段は、

撮影された画像フレーム内において観測された前記特徴点と前記較正用画像における特徴点との対応付けを行なう特徴点特定手段と、

前記撮影された画像フレーム内において観測された前記特徴点についての前記相対位置関係行列の部分行列と、前記観測された特徴点とから前記眼球中心の投影位置を推定する、請求項 4 記載の視線方向の推定装置。

【請求項 6】

前記相対関係特定手段は、較正時において、前記撮影手段により撮影された複数の較正用画像を予め取得し、前記較正用画像を正規化した上で、前記顔領域内の複数の特徴点と前記眼球中心との間の相対的な 3 次元の位置関係を特定し、

前記眼球中心推定手段は、前記撮影手段により撮影された前記人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、前記人間の眼球中心の投影位置を推定する、請求項 1 または 2 記載の視線方向の推定装置

【請求項 7】

前記相対関係特定手段は、前記撮影手段により撮影された画像を正規化した上で、前記顔領域内の複数の特徴点間の相対的な 3 次元の位置関係を特定し、

前記眼球中心推定手段は、前記撮影手段により撮影された前記人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、前記人間の眼球中心の位置を推定する、請求項 1 または 2 記載の視線方向の推定装置。

【請求項 8】

前記虹彩中心抽出手段は、前記撮影手段により撮影された前記人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された構成のモデル領域とを照合することで、前記虹彩中心位置を抽出する、請求項 1 記載の視線方向の推定装置。

【請求項 9】

人間の顔領域を含む複数のフレームの対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、

前記人間が前記撮影手段を見ている状態で前記撮影手段により撮影された較正用画像を予め取得し、前記顔領域内の複数の特徴点と眼球中心との間の相対的な 3 次元の位置関係を特定するステップと、

前記撮影手段により撮影された前記対象画像領域内において前記複数の特徴点の投影位

10

20

30

40

50

置を検出し、予め較正時に特定されている前記人間の眼球中心の位置と前記特徴点との相対的な3次元の位置関係により、前記フレームにおいて特定された前記複数の特徴点間の相対的な3次元の位置関係に基づいて、前記人間の眼球中心の3次元の位置を推定するステップと、

前記フレームの画像領域内において、虹彩の領域と前記虹彩の形状のモデルとを照合することで、虹彩中心が前記画像に投影される虹彩中心位置を抽出するステップと、

抽出された前記虹彩中心位置と前記眼球中心の投影位置とに基づき、前記眼球中心と前記虹彩中心とを結ぶ3次元直線方向として、視線方向を推定するステップとを備える、視線方向の推定方法。

【請求項10】

10

フレームの画像の人間の顔領域を含む対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、

前記撮影手段により撮影された画像を取得し、前記画像を正規化した上で、前記顔領域内の複数の特徴点の間の相対的な3次元の位置関係を特定するステップと、

前記撮影手段により撮影された前記対象画像領域内において前記複数の特徴点の投影位置を検出し、対象となる前記フレーム以前に撮影されたフレームで特定されている前記人間の眼球中心と前記特徴点との相対的な3次元の位置関係により、前記フレームにおいて特定された前記複数の特徴点間の相対的な3次元の位置関係に基づいて、前記人間の眼球中心の3次元の位置を推定するステップとを備え、

前記推定するステップは、前記撮影手段により撮影された前記人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、前記人間の眼球中心の位置を推定するステップを含み、

20

各前記フレームの前記対象画像領域内において、虹彩の領域と前記虹彩の形状のモデルとを照合することで、虹彩中心が前記画像に投影される虹彩中心位置を抽出するステップと、

抽出された前記虹彩中心位置と前記眼球中心の位置とに基づき、前記眼球中心と前記虹彩中心とを結ぶ3次元直線方向として、視線方向を推定するステップとを備える、視線方向の推定方法。

【請求項11】

30

演算処理手段を有するコンピュータに、対象となる画像領域内の顔について視線方向の推定処理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複数のフレームの対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、

前記人間が前記撮影手段を見ている状態で前記撮影手段により撮影された較正用画像を予め取得し、前記顔領域内の複数の特徴点と眼球中心との間の相対的な3次元の位置関係を特定するステップと、

前記撮影手段により撮影された前記対象画像領域内において前記複数の特徴点の投影位置を検出し、予め較正時に特定されている前記人間の眼球中心の位置と前記特徴点との相対的な3次元の位置関係により、前記フレームにおいて特定された前記複数の特徴点間の相対的な3次元の位置関係に基づいて、前記人間の眼球中心の3次元の位置を推定するステップと、

40

前記フレームの画像領域内において、虹彩の領域と前記虹彩の形状のモデルとを照合することで、虹彩中心が前記画像に投影される虹彩中心位置を抽出するステップと、

抽出された前記虹彩中心位置と前記眼球中心の投影位置とに基づき、前記眼球中心と前記虹彩中心とを結ぶ3次元直線方向として、視線方向を推定するステップとをコンピュータに実行させる、プログラム。

【請求項12】

演算処理手段を有するコンピュータに、対象となる画像領域内の顔について視線方向の推定処理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、

50

フレームの画像の人間の顔領域を含む対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、

前記撮影手段により撮影された画像を取得し、前記画像を正規化した上で、前記顔領域内の複数の特徴点の間の相対的な3次元の位置関係を特定するステップと、

前記撮影手段により撮影された前記対象画像領域内において前記複数の特徴点の投影位置を検出し、対象となる前記フレーム以前に撮影されたフレームで特定されている前記人間の眼球中心と前記特徴点との相対的な3次元の位置関係により、前記フレームにおいて特定された前記複数の特徴点間の相対的な3次元の位置関係に基づいて、前記人間の眼球中心の3次元の位置を推定するステップとを備え、

前記推定するステップは、前記撮影手段により撮影された前記人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、前記人間の眼球中心の位置を推定するステップを含み、

各前記フレームの前記対象画像領域内において、虹彩の領域と前記虹彩の形状のモデルとを照合することで、虹彩中心が前記画像に投影される虹彩中心位置を抽出するステップと、

抽出された前記虹彩中心位置と前記眼球中心の位置とに基づき、前記眼球中心と前記虹彩中心とを結ぶ3次元直線方向として、視線方向を推定するステップとをコンピュータに実行させる、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、画像中の人物の視線方向を推定および検出するための画像認識の分野に関する。

【背景技術】

【0002】

人物の視線方向の推定は、たとえば、マンマシンインタフェースの1つの方法として従来研究されてきた。

【0003】

視線計測について、従来のカメラを利用した手法では、カメラの設置位置によって「頭部装着型」と「非装着型」に分類できる。一般的に「頭部装着型」は精度は高いが、ユーザの負担が大きい。またカメラの座標系が頭部の動きに連動するため、注視対象を判別するためには、外界の座標系と結び付ける工夫が必要となる。

【0004】

具体的な手法としては、近赤外の点光源(Light Emitting Diode:LED)を目に照射し、角膜で反射された光源像と瞳孔の位置から視線を推定する瞳孔角膜反射法が良く知られており、「頭部装着型」または「非装着型」の両方に利用されている。赤外照明によって瞳孔と虹彩のコントラストは可視光照明より高くなり、瞳孔を検出しやすくなるが、その直径は数ミリで、また赤外光源の反射像もごく小さなスポットとして映るため、解像度の高い画像が要求される。そのため、片目をできるだけ大きく撮像することになる。その結果、「非装着型」の場合、顔が少し動くと目がカメラ視界から外れるという問題がある。

【0005】

角膜反射を利用しない「非装着型」による視線推定手法については、ステレオカメラ方式と単眼カメラ方式の大きく2種類に分けられる。

【0006】

ステレオカメラ方式では、まず顔の特徴点(人為的に貼付したマーカや目尻などの自然特徴点など)の3次元位置を2眼ステレオにより推定し、それをもとに眼球中心位置を推定する手法である。視線方向は、求めた眼球中心位置と画像中の瞳孔位置・虹彩位置を結ぶ直線として推定される。しかし、事前にカメラ間のキャリブレーションが必要であるため、カメラの観測範囲の変更は容易ではないという問題がある。

10

20

30

40

50

【0007】

一方、単眼カメラによる方式では、カメラと顔の距離が離れても、ズームなどにより必要な画像解像度が得られれば視線が推定できる。

【0008】

単眼カメラによる方式では、眼の画像パターンから視線を推定するニューラルネットワークによる手法や、観測時の虹彩の楕円形状から視線方向（虹彩の法線方向）を推定する手法、ステレオカメラ方式と同じように特徴点を抽出して虹彩位置との幾何学的関係から視線を推定する手法がある。

【0009】

これらのうち、顔の上の特徴点と虹彩との相対的位置関係から視線を推定する方法は直感的に理解しやすいことから、早くから検討されてきた。

10

【0010】

青山らは顔の向きの変化にも対応するため、左右の目尻と口の両端（口角）から形成される台形を利用して顔の向きを推定すると同時に、両目尻の midpoint と左右の虹彩の midpoint の差から正面視からの目の片寄り量を推定し、両方合わせて視線方向を推定する原理を示した（たとえば、非特許文献1を参照）。

【0011】

伊藤らは目の近くに指標点となるサージカルテープなどを貼付し、その指標マークと虹彩の相対位置変化から視線方向を推定するシステムを肢体不自由者のコミュニケーションの手段として検討している（たとえば、非特許文献2を参照）。この場合、顔をキャリブレーション時と同じ姿勢に保つ必要がある。

20

【0012】

三宅らは眼球の幾何学的モデルから、2つの眼球の中心を結ぶ直線が顔表面と交差する左右の点を参照点とすれば、画像上のその midpoint と左右の虹彩の midpoint から、顔の向きに関係なく視線方向が計算できることを示した（たとえば、非特許文献3を参照）。

【0013】

川戸らは、4つの参照点を利用する顔の向きに影響されない視線方向推定方法を提案した（たとえば、非特許文献4を参照）。4つの参照点は原理的には同一平面上でなければ位置的な制限はなく、4つの参照点から眼球中心を推定し、眼球中心と虹彩中心を結ぶ直線として視線方向を推定するものである。

30

【0014】

なお、以下に説明する本発明の視線方向の推定方法においては、画像中からまず人物の顔を検出する。そこで、従来の画像中からの顔の検出手法の従来技術については、以下のようなものがある。

【0015】

つまり、これまでに、肌色情報を用いた顔検出システムや、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については報告が数多くなされている。

【0016】

たとえば、安定性が高く、かつ実時間での顔の追跡が可能な手法として、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、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るとの手法が提案されている（たとえば、非特許文献5、特許文献1を参照）。

40

【0017】

さらに、他の手法として、たとえば、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して、順次、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し、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、というような顔を検

50

出す方法も提案されている（たとえば、特許文献 2 を参照）。

【 0 0 1 8 】

さらに、顔画像中から鼻位置をリアルタイムで追跡する手法についても報告されている（たとえば、非特許文献 6 を参照）

【非特許文献 1】青山宏、河越正弘：「顔の面对称性を利用した視線感知法」情処研報 89 - CV - 61、pp.1-8（1989）。

【非特許文献 2】伊藤，数藤：「画像センサを用いた眼球運動による意思伝達システム」信学技報 WIT99 - 39（2000）。

【非特許文献 3】三宅哲夫、春田誠司、堀畑聡：「顔の向きに依存しない特徴量を用いた注視判定法」信学論（D - I I）、Vol . J86 - D-II、No . 12、 pp . 1737 - 1744（2003）

10

【非特許文献 4】川戸、内海、安部：「4つの参照点と3枚のキャリブレーション画像に基づく単眼カメラからの視線推定」画像の認識・理解シンポジウム（MIRU2005），pp . 1337 - 1342（2005）。

【非特許文献 5】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 . J 84 - D - I I，no 12，pp . 2577 - 2584，Dec . 2001。

【非特許文献 6】川戸慎二郎、鉄谷信二：鼻位置の検出とリアルタイム追跡：信学技報 IE 2002 - 263、pp . 25 - 29（2003）。

【特許文献 1】特開 2001 - 52176 号公報明細書

20

【特許文献 2】特開 2004 - 185611 号公報明細書

【発明の開示】

【発明が解決しようとする課題】

【 0 0 1 9 】

しかしながら、従来の単眼カメラによる方式の視線検出のための手法では、参照点として人為的なマークが必要であるという問題があった。

【 0 0 2 0 】

それゆえに本発明の目的は、カメラにより撮影された画像情報に基づいて、人為的なマーカを用いることなく、リアルタイムに視線を追跡する視線方向の推定装置、視線方向の推定方法およびコンピュータに当該視線方向の推定方法を実行させるためのプログラムを提供することである。

30

【 0 0 2 1 】

この発明のさらに他の目的は、1つのカメラ（単眼カメラ）により撮影された画像情報に基づいて、リアルタイムに視線を追跡する視線方向の推定装置、視線方向の推定方法およびコンピュータに当該視線方向の推定方法を実行させるためのプログラムを提供することである。

【課題を解決するための手段】

【 0 0 2 2 】

この発明のある局面に従うと、視線方向の推定装置であって、画像を獲得するための撮影手段と、撮影手段により撮影された人間の顔領域を含むフレームの画像に基づき、顔領域内の複数の特徴点の間の相対的な3次元の位置関係を特定する相対関係特定手段と、人間の眼球中心の3次元の位置である眼球中心位置の推定処理を実行する眼球中心推定手段とを備え、眼球中心推定手段は、予め特定された特徴点と眼球中心の位置との相対的な3次元の位置関係を用いて、当該推定処理の対象のフレームにおいて相対関係特定手段により特定された複数の特徴点間の相対的な3次元の位置関係に基づいて、推定処理を実行し、推定処理の対象のフレームの画像領域内において、虹彩の領域と虹彩の形状のモデルとを照合することで、虹彩中心が画像に投影される虹彩中心位置を抽出する虹彩中心抽出手段と、推定処理の対象のフレームの画像において抽出された虹彩中心位置と眼球中心の位置とに基づき、眼球中心と虹彩中心とを結ぶ3次元直線の方向として、視線方向を推定する視線推定手段とをさらに備える。

40

50

【0023】

好ましくは、撮影手段は、人間の顔領域を含む対象画像領域内の各画素に対応する画像データを撮影して獲得するための単眼の撮影手段である。

【0024】

好ましくは、相対関係特定手段は、較正時において、人間が撮影手段を見ている状態で撮影手段により撮影された複数の較正用画像を予め取得し、特徴点と眼球中心との間の相対的な3次元の位置関係を特定し、眼球中心推定手段は、撮影手段により撮影された人間の顔領域を含む対象画像領域内において検出された複数の特徴点の投影位置により、特定された相対的な3次元の位置関係に基づいて、人間の眼球中心の投影位置を推定する。

【0025】

好ましくは、相対関係特定手段は、複数の較正用画像内の複数の特徴点の投影位置を抽出し、複数の特徴点の投影位置を要素として並べた計測行列を算出する計測行列算出手段と、計測行列を因子分解により、撮影手段の姿勢に関する情報を要素とする撮影姿勢行列と、複数の特徴点間の相対的な3次元の位置関係に関する情報を要素とする相対位置関係行列とに分解する因子分解手段とを含む。

【0026】

好ましくは、眼球中心推定手段は、撮影された画像フレーム内において観測された特徴点と較正用画像における特徴点との対応付けを行なう特徴点特定手段と、撮影された画像フレーム内において観測された特徴点についての相対位置関係行列の部分行列と、観測された特徴点とから眼球中心の投影位置を推定する。

【0027】

好ましくは、相対関係特定手段は、較正時において、撮影手段により撮影された複数の較正用画像を予め取得し、較正用画像を正規化した上で、顔領域内の複数の特徴点と眼球中心との間の相対的な3次元の位置関係を特定し、眼球中心推定手段は、撮影手段により撮影された人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、人間の眼球中心の投影位置を推定する。

【0028】

好ましくは、相対関係特定手段は、撮影手段により撮影された画像を正規化した上で、顔領域内の複数の特徴点間の相対的な3次元の位置関係を特定し、眼球中心推定手段は、撮影手段により撮影された人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、人間の眼球中心の位置を推定する。

【0029】

好ましくは、虹彩中心抽出手段は、撮影手段により撮影された人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心位置および眼球半径を用いて算出され投影された構成のモデル領域とを照合することで、虹彩中心位置を抽出する。

【0030】

この発明の他の局面に従うと、視線方向の推定方法であって、人間の顔領域を含む複数のフレームの対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、人間が撮影手段を見ている状態で撮影手段により撮影された較正用画像を予め取得し、顔領域内の複数の特徴点と眼球中心との間の相対的な3次元の位置関係を特定するステップと、撮影手段により撮影された対象画像領域内において複数の特徴点の投影位置を検出し、予め較正時に特定されている人間の眼球中心の位置と特徴点との相対的な3次元の位置関係により、フレームにおいて特定された複数の特徴点間の相対的な3次元の位置関係に基づいて、人間の眼球中心の3次元の位置を推定するステップと、フレームの画像領域内において、虹彩の領域と虹彩の形状のモデルとを照合することで、虹彩中心が画像に投影される虹彩中心位置を抽出するステップと、抽出された虹彩中心位置と眼球中心の投影位置とに基づき、眼球中心と虹彩中心とを結ぶ3次元直線の方向

10

20

30

40

50

として、視線方向を推定するステップとを備える。

この発明のさらに他の局面に従うと、視線方向の推定方法であって、フレームの画像の人間の顔領域を含む対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、撮影手段により撮影された画像を取得し、画像を正規化した上で、顔領域内の複数の特徴点の間の相対的な3次元の位置関係を特定するステップと、撮影手段により撮影された対象画像領域内において複数の特徴点の投影位置を検出し、対象となるフレーム以前に撮影されたフレームで特定されている人間の眼球中心と特徴点との相対的な3次元の位置関係により、フレームにおいて特定された複数の特徴点間の相対的な3次元の位置関係に基づいて、人間の眼球中心の3次元の位置を推定するステップとを備え、推定するステップは、撮影手段により撮影された人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、人間の眼球中心の位置を推定するステップを含み、各フレームの対象画像領域内において、虹彩の領域と虹彩の形状のモデルとを照合することで、虹彩中心が画像に投影される虹彩中心位置を抽出するステップと、抽出された虹彩中心位置と眼球中心の位置とに基づき、眼球中心と虹彩中心とを結ぶ3次元直線の方向として、視線方向を推定するステップとを備える。

10

【0031】

この発明のさらに他の局面に従うと、演算処理手段を有するコンピュータに、対象となる画像領域内の顔について視線方向の推定処理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複数のフレームの対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、人間が撮影手段を見ている状態で撮影手段により撮影された較正用画像を予め取得し、顔領域内の複数の特徴点と眼球中心との間の相対的な3次元の位置関係を特定するステップと、撮影手段により撮影された対象画像領域内において複数の特徴点の投影位置を検出し、予め較正時に特定されている人間の眼球中心の位置と特徴点との相対的な3次元の位置関係により、フレームにおいて特定された複数の特徴点間の相対的な3次元の位置関係に基づいて、人間の眼球中心の3次元の位置を推定するステップと、フレームの画像領域内において、虹彩の領域と虹彩の形状のモデルとを照合することで、虹彩中心が画像に投影される虹彩中心位置を抽出するステップと、抽出された虹彩中心位置と眼球中心の投影位置とに基づき、眼球中心と虹彩中心とを結ぶ3次元直線の方向として、視線方向を推定するステップとをコンピュータに実行させる。

20

30

この発明のさらに他の局面に従うと、演算処理手段を有するコンピュータに、対象となる画像領域内の顔について視線方向の推定処理を実行させるためのプログラムであって、プログラムは、フレームの画像の人間の顔領域を含む対象画像領域内の各画素に対応する画像データを単眼の撮影手段により撮影して獲得するステップと、撮影手段により撮影された画像を取得し、画像を正規化した上で、顔領域内の複数の特徴点の間の相対的な3次元の位置関係を特定するステップと、撮影手段により撮影された対象画像領域内において複数の特徴点の投影位置を検出し、対象となるフレーム以前に撮影されたフレームで特定されている人間の眼球中心と特徴点との相対的な3次元の位置関係により、フレームにおいて特定された複数の特徴点間の相対的な3次元の位置関係に基づいて、人間の眼球中心の3次元の位置を推定するステップとを備え、推定するステップは、撮影手段により撮影された人間の顔領域を含む対象画像領域内において、撮影された画像内の虹彩の領域と、顔座標系内で仮定した眼球中心の位置および眼球半径を用いて算出され投影された虹彩のモデル領域と照合することで、人間の眼球中心の位置を推定するステップを含み、各フレームの対象画像領域内において、虹彩の領域と虹彩の形状のモデルとを照合することで、虹彩中心が画像に投影される虹彩中心位置を抽出するステップと、抽出された虹彩中心位置と眼球中心の位置とに基づき、眼球中心と虹彩中心とを結ぶ3次元直線の方向として、視線方向を推定するステップとをコンピュータに実行させる。

40

【発明を実施するための最良の形態】

50

【 0 0 3 2 】

[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「視線方向の推定装置」について説明する。この視線方向の推定装置は、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に基づいて、視線方向を推定・検出するためのものである。図 1 に、この視線方向の推定装置の外観を示す。

【 0 0 3 3 】

ただし、以下に説明する「視線方向の推定装置」の各機能の一部または全部は、ハードウェアにより実現されてもよい。

【 0 0 3 4 】

図 1 を参照して、この視線方向の推定装置を構成するシステム 2 0 は、C D - R O M (Compact Disc Read-Only Memory) または D V D - R O M (Digital Versatile Disc Read-Only Memory) ドライブ (以下、「光学ディスクドライブ」と呼ぶ) 5 0、あるいは F D (Flexible Disk) ドライブ 5 2 のような記録媒体からデータを読み取るためのドライブ装置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単眼カメラ 3 0 とを含む。この実施の形態の装置では、単眼カメラ 3 0 としては C C D (Charge Coupled Device) または C M O S (Complementary Metal-Oxide Semiconductor) センサのような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 2 0 を操作する人物の顔の位置および視線を推定・検出する処理を行なうものとする。

【 0 0 3 5 】

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 6 】

なお、カメラ 3 0 は、必ずしも単眼カメラに限定される訳ではなく、複数の視点からの画像に基づいて、以下に説明するような視線方向の推定を行なうことも可能であるが、以下の説明でも明らかとなるように、本発明の視線方向の推定装置では、単眼カメラからの画像情報のみに基づいた場合でも、視線の推定を行なうことが可能である。

【 0 0 3 7 】

図 2 は、カメラ 3 0 により撮影された画像に基づいて、コンピュータ本体 4 0 の処理結果がディスプレイ 4 2 に表示される一例を示す図である。

【 0 0 3 8 】

図 2 に示すように、カメラ 3 0 で撮影された画像は、ディスプレイ 4 2 の撮影画像表示領域 2 0 0 にリアルタイムに動画として表示される。特に限定されないが、たとえば、撮影画像表示領域 2 0 0 上に、視線方向を示す指標として、眉間から視線方向に延びる線分を表示してもよい。

【 0 0 3 9 】

図 3 に、このシステム 2 0 の構成をブロック図形式で示す。図 3 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、光学ディスクドライブ 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。光学ディスクドライブ 5 0 には C D - R O M (または D V 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 4 0 】

既に述べたようにこの視線方向の推定装置の主要部は、コンピュータハードウェアと、

10

20

30

40

50

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 (または D V 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、光学ドライブ 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 4 1 】

図 1 および図 3 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、C D - R O M (または D V D - R O M) 6 2、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。

10

【 0 0 4 2 】

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、当該視線方向の推定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視線方向の推定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。

20

【 0 0 4 3 】

[システムの機能ブロック]

以下に説明するとおり、本発明の視線方向の推定装置では、顔特徴点を検出・追跡することにより、単眼カメラで視線方向を推定する。

【 0 0 4 4 】

本発明の視線方向の推定装置では、眼球中心と虹彩中心を結ぶ 3 次元直線を視線方向として推定する。眼球中心は画像からは直接観測することはできないが、カメラを注視している場合には眼球中心と虹彩中心が同じ位置に観測されることを利用し、眼球中心と顔特徴点との相対関係をモデル化することにより、眼球中心の投影位置を推定する。

30

【 0 0 4 5 】

初期校正(キャリブレーション)では、ユーザがカメラを注視しながら顔の向きが異なる、2 枚以上の画像フレーム列を取得し、顔特徴点と虹彩中心を抽出・追跡することにより、顔特徴点と眼球中心との相対関係をモデル化する。

【 0 0 4 6 】

視線推定時には、撮影画像から顔特徴点を追跡することで、眼球中心の投影位置を推定し、その点と虹彩中心とを結ぶ 3 次元ベクトルを視線方向として推定する。

40

【 0 0 4 7 】

以下では、本発明による視線の検知のための構成の概略について説明する。

図 4 は、図 3 に示した C P U 5 6 が、ハードディスク 5 4 等に格納されたプログラムに基づいて行なうソフトウェア処理を示す機能ブロック図である。なお、上述のとおり、この機能ブロックの全部または一部は、専用のハードウェアにより実行される構成としてもよい。

【 0 0 4 8 】

視線方向の推定装置のシステム 2 0 の処理は、単眼カメラ 3 0 をベースとしており、顔位置の検出が可能で光学的に十分なズーム比が得られれば観測距離の制約を受けない。

【 0 0 4 9 】

50

以下では、何らかの手続き、たとえば、ユーザによるカメラの調整等によりあらかじめ大まかな人物顔位置が与えられることを仮定している。

【 0 0 5 0 】

カメラ 3 0 により撮像された動画に対応する映像信号は、フレームごとに画像キャプチャ処理部 5 6 0 2 により制御されてデジタルデータとしてキャプチャされ、画像データ記録処理部 5 6 0 4 により、たとえば、ハードディスク 5 4 のような記憶装置に格納される。

【 0 0 5 1 】

顔検出部 5 6 0 6 では、キャプチャされたフレーム画像列に対して 6 分割矩形フィルタによる顔候補探索（眉間の候補位置の探索）と、探索された眉間の候補位置のうちから、左右の目についてのテンプレートマッチングなどの処理を用いた識別処理を適用することにより画像中の顔位置（両目位置）を検出する。

10

【 0 0 5 2 】

次に、顔領域抽出部 5 6 0 8 は、求めた両目位置から虹彩中心、鼻位置を抽出する。

さらに、特徴点抽出部 5 6 1 0 は、抽出された顔や鼻、目（虹彩）の位置を利用し、複数の特徴点の抽出および追跡を行なって、顔領域内の画像特徴点を検出する。なお、画像特徴点周辺のテクスチャは、キャリブレーション時において追跡のためのテンプレートとして保持しておき、さらに適宜更新して、フレーム間の特徴点の対応づけに利用する。一方、虹彩中心抽出部 5 6 1 2 は、後に説明するように、目の周辺領域に対して、ラプシアンにより虹彩のエッジ候補を抽出し、円のハフ変換を適用することにより、虹彩の中心の投影位置を検出する。相対関係特定部 5 6 1 4 は、キャリブレーション時に、眼球中心および特徴点の 3 次元的な相対位置関係を後に説明するようにモデルとして特定する。眼球中心推定部 5 6 1 6 は、特徴点の投影位置と相対位置関係のモデルとを用いて、眼球の中心位置を推定する。

20

【 0 0 5 3 】

視線方向推定部 5 6 1 8 は、抽出された虹彩の中心の投影位置と推定された眼球の中心位置とに基づいて、視線方向を推定する。

【 0 0 5 4 】

また、表示制御部 5 6 3 0 は、以上のようにして推定された視線の方向を、取得された画像フレーム上に表示するための処理を行なう。

30

【 0 0 5 5 】

すなわち、視線方向の推定装置のシステム 2 0 では、特徴点の追跡処理の安定性を確保するため、同一特徴点に関して異なるフレームにおける複数の観測テクスチャを保持している。初期校正過程では、これらの特徴点と虹彩中心の関係から顔特徴点と眼球中心の相対関係を求める。視線推定過程では、校正過程で得られた関係を元に現フレームで得られている特徴点群から眼球中心位置を推定し、その位置と虹彩中心位置から視線方向を決定する。

【 0 0 5 6 】

[視線方向の推定処理の動作]

(顔検出処理)

40

以下では、まず、視線方向の推定の前提として行なわれる顔検出処理について説明する。

【 0 0 5 7 】

(6 分割矩形フィルタによる顔検出処理)

視線方向の推定装置のシステム 2 0 では、特に限定されないが、たとえば、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。

【 0 0 5 8 】

50

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 5 9 】

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。

(6 分割矩形フィルタ)

図 5 は、眉間候補領域を検出するためのフィルタを説明するための概念図である。

【 0 0 6 0 】

図 5 (a) 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。 10

【 0 0 6 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 6 2 】

図 5 (a) 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。

【 0 0 6 3 】

図 5 (b) は、このような 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。 20

【 0 0 6 4 】

図 6 は、6 分割矩形フィルタの他の構成を示す概念図である。

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 6 は、このような場合の 6 分割矩形フィルタの構成を示す。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。

【 0 0 6 5 】

図 6 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「 $\overline{S_i}$ 」(S_i に上付きの “ - ” をつける) を求める。 30

【 0 0 6 6 】

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 6 7 】

【 数 1 】

$$\overline{S_1} < \overline{S_2} \quad \text{and} \quad \overline{S_1} < \overline{S_4}, \quad (1)$$

$$\overline{S_3} < \overline{S_2} \quad \text{and} \quad \overline{S_3} < \overline{S_6} \quad (2)$$

【 0 0 6 8 】

そこで、これらの関係を満たす点を眉間候補 (顔候補) として抽出する。

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) 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ことができる。インテグラルイメージを利用することでフィルタの大きさに依らず高速に実行することができる。多重解像度画像に本手法を適用することにより、画像上の顔の大きさが変化した場合にも顔候補の抽出が可能となる。 40

【 0 0 6 9 】

このようにして得られた眉間候補（顔候補）に対しては、両目の標準パターンとのテンプレートマッチングにより、真の眉間位置（真の顔領域）を特定することができる。

【 0 0 7 0 】

なお、得られた顔候補に対して、サポートベクトルマシン（SVM）による顔モデルによる検証処理を適用し顔領域を決定することもできる。髪型の違いや髭の有無、表情変化による認識率の低下を避けるため、たとえば、図7に示すように、眉間を中心とした画像領域を利用してSVMによるモデル化を行なうことができる。なお、このようなSVMによる真の顔領域の決定については、文献：S. Kawato, N. Tetsutani and K. Hosaka: "Scale-adaptive face detection and tracking in real time with ssr filters and support vector machine", IEICE Trans. on Info. and Sys., E88-D, 12, pp. 2857-2863 (2005) に開示されている。6分割矩形フィルタによる高速候補抽出とSVMによる処理とを組み合わせることで実時間の顔検出が可能である。

10

【 0 0 7 1 】

（目・鼻の検出処理）

続いて、目、鼻や虹彩中心の位置を、上述した非特許文献4や非特許文献6の手法を用いて抽出する

両目の位置については、前節の顔領域検出で眉間のパターンを探索しているため、眉間の両側の暗い領域を再探索することにより、大まかな両目の位置を推定することができる。しかし、視線方向の推定のためには、虹彩中心をより正確に抽出する必要がある。ここでは、上で求めた目の周辺領域に対して、ラプラシアンにより虹彩のエッジ候補を抽出し、円のハフ変換を適用することにより、虹彩および虹彩の中心の投影位置を検出する。

20

【 0 0 7 2 】

鼻の位置は、鼻先が凸曲面であるため周囲に対し明るい点として観測されやすいことと、両目の位置から鼻の存在範囲が限定できることを利用して抽出する。また、両目、鼻の位置を用いて、大体の顔の向きも推定できる。

【 0 0 7 3 】

図8は、顔検出結果の例を示す図である。検出された顔において、虹彩中心や鼻先や口なども検出されている。たとえば、特徴点としては、鼻先や、左右の目の目尻や目頭、口の両端、鼻腔中心などを用いることができる。

30

【 0 0 7 4 】

（視線推定）

（視線推定の原理）

本発明では、視線方向は眼球中心と虹彩中心を結ぶ3次元直線として与えられるとする。

【 0 0 7 5 】

図9は、視線方向を決定するためのモデルを説明する概念図である。

画像上での眼球半径を l 、画像上での眼球中心と虹彩中心との x 軸方向、 y 軸方向の距離を d_x 、 d_y とすると、視線方向とカメラ光軸とのなす角、つまり、視線方向を向くベクトルが x 軸および y 軸との成す角 φ_x 、 φ_y は次式で表される。

40

【 0 0 7 6 】

【数2】

$$\varphi_x = \sin^{-1}\left(\frac{d_x}{l}\right), \quad \varphi_y = \sin^{-1}\left(\frac{d_y}{l}\right) \quad (3)$$

【 0 0 7 7 】

式(3)により、視線方向を推定するためには、画像上での眼球半径と眼球中心・虹彩中心の投影位置が必要となる。ここで、虹彩中心の投影位置については、上述したとおり、ハフ変換を用いた手法により求めることができる。画像上での眼球直径 r は、解剖学的

50

なモデル（標準的な人の眼球直径）を用いてもよいし、別途キャリブレーションにより求めてもよい。

【0078】

図10は、図9に示した状態から、ユーザがカメラを注視する状態に移行した後の虹彩中心、眼球中心および投影点の関係を示す概念図である。

【0079】

眼球中心の投影位置については、一般には、画像から直接観測することはできない。しかし、ユーザがカメラを注視した場合について考えると、図10に示すとおり、カメラ、虹彩中心、眼球中心の3点が1直線上に並ぶため、画像では虹彩中心と眼球中心は同一点に投影されることがわかる。

10

【0080】

そこで、本発明では、ユーザがカメラを注視しながら、顔の姿勢を変化させている画像フレーム列を撮影し、これらの画像列から虹彩位置と顔特徴点を抽出・追跡することにより、眼球中心と顔特徴点間の相対幾何関係を推定する。

【0081】

後により詳しく説明するように、本発明の視線方向の推定装置では、眼球中心と顔特徴点間の相対関係の推定処理と眼球中心の投影位置推定とを行なう。

【0082】

（顔特徴点の追跡による視線推定）

図11は、視線方向の推定装置の初期設定として行なうキャリブレーションを説明するためのフローチャートである。

20

【0083】

まず、キャリブレーション用の画像列として、ユーザがカメラを注視しながら、顔の姿勢を変化させている画像フレーム列を撮影する（ステップS102）。

【0084】

図12は、このようにしてキャリブレーションにおいて撮影された4枚の画像フレームを示す。

【0085】

ここでは、より一般に、 N ($N \geq 2$) 枚の画像列が得られたとする。各画像フレームを、フレーム I_1, \dots, I_N とする。

30

【0086】

次に、得られた各画像フレーム列に対して、上述したような方法によって顔検出部5606が顔検出処理を行ない（ステップS104）、続いて、顔領域抽出部5608が、目や鼻の検出処理を行なう（ステップS106）。

【0087】

さらに、特徴点抽出部5610が特徴点の抽出・追跡を行なう（ステップS108）。なお、特徴点の抽出方法としては、上述したような方法の他に、たとえば、文献：J. Shi and C. Tomasi: "Good features to track", Proc. CVPR94, pp. 593 - 600 (1994) で提案された手法を用いることもできる。

【0088】

ここで、各画像フレーム I_i ($i = 1, \dots, N$) において M ($M \geq 4$) 点の特徴点 p_j ($j = 1, \dots, M$) が検出・追跡できたとする。画像フレーム I_i における特徴点 p_j の2次元観測位置を $x_j^{(i)}$ (太字) = $[x_j^{(i)}, y_j^{(i)}]^t$ ($i = 1, \dots, N$, $j = 1, \dots, M$) とし、両目の虹彩中心の2次元観測位置をそれぞれ $x_r^{(i)}$ (太字) = $[x_r^{(i)}, y_r^{(i)}]^t$, $x_l^{(i)}$ (太字) = $[x_l^{(i)}, y_l^{(i)}]^t$ ($i = 1, \dots, N$) とする。ここで、行列 W を以下のように定義する。

40

【0089】

【数3】

$$\mathbf{W} = \begin{bmatrix} \mathbf{x}_1^{(1)} & \cdots & \mathbf{x}_M^{(1)} & \mathbf{x}_r^{(1)} & \mathbf{x}_l^{(1)} \\ \vdots & & \vdots & & \vdots \\ \mathbf{x}_1^{(N)} & \cdots & \mathbf{x}_M^{(N)} & \mathbf{x}_r^{(N)} & \mathbf{x}_l^{(N)} \end{bmatrix} \quad (4)$$

【0090】

因子分解法により、特徴点の各フレームでの2次元観測位置を縦に並べた行列W（計測行列）は以下のように分解できる。

【0091】

【数4】

$$\mathbf{W} = \begin{bmatrix} m_1^t \\ \vdots \\ m_N^t \end{bmatrix} \begin{bmatrix} s_1 & \cdots & s_M & s_r & s_l \end{bmatrix} \\ = \mathbf{MS} \quad (5)$$

【0092】

ここで、行列M（「撮影姿勢行列」と呼ぶ）にはカメラの姿勢に関する情報のみが、行列S（「相対位置関係行列」と呼ぶ）には観測対象物の形状に関する情報のみが含まれており、顔特徴点と眼球中心との3次元的位置の相対関係は行列Sとして求まる（ステップS110）。すなわち、正射影を仮定すると、行列Mの各要素が画像フレームでのカメラの姿勢を表す単位ベクトルであって、それぞれの大きさが1であり相互には直交するとの拘束条件のもとで、行列Wは、特異値分解により一義的に行列Mと行列Sの積に分解できることが知られている。なお、このような計測行列Wを、因子分解により、カメラの運動の情報を表す行列と対象物の形状情報を表す行列へ分解する点については、文献：金出，ポールマン，森田：「因子分解法による物体形状とカメラ運動の復元」，電子通信学会論文誌D-I I，J76-D-I I，8，pp.1497-1505（1993）に開示がある。

【0093】

（視線方向推定処理）

図13は、リアルタイムの視線方向の推定処理のフローチャートを示す。

【0094】

次に、以上で得られた結果を用いて、視線方向を推定する手順について説明する。

まず、カメラ30から画像フレームを取得すると（ステップS200）、キャリブレーション時と同様にして、顔の検出および目鼻の検出が行なわれ（ステップS202）、取得された画像フレーム中の特徴点が抽出される（ステップS204）。

【0095】

画像フレーム I_k が得られたとする。ここで、眼球中心以外の特徴点のうち m 点 p_j （ $j = j_1, \dots, j_m$ ）が、それぞれ、 $x_j^{(k)}$ （太字） $= [x_j^{(k)}, y_j^{(k)}]^t$ に観測されたとする。このとき、観測された特徴点について、上述したように特徴点近傍のテンプレートをを用いたテンプレートマッチングを実施することで、キャリブレーション時に特定された特徴点と現画像フレーム中で観測された特徴点との対応付けが行なわれ

10

20

30

40

50

て、現画像フレーム中の特徴点が特定される（ステップS206）。

【0096】

なお、上述のとおり、特徴点を特定するためのテンプレートは、キャリブレーションの時のものに限定されず、たとえば、最近の画像フレームの所定枚数について検出された特徴点の近傍の所定の大きさの領域内の画像を所定個数だけ保持しておき、これら所定枚数のテンプレートについてマッチングをした結果、もっとも一致度の高い特徴点に特定することとしてもよい。

【0097】

顔特徴点 p_j の2次元観測位置 $x_j^{(k)}$ （太字） $= [x_j^{(k)}, y_j^{(k)}]^t$ とキャリブレーションより求めた3次元位置 s_j （太字） $= [X_j, Y_j, Z_j]^t$ （ $j = 1, \dots, M$ ）の間には、 M 個の特徴点のうち観測された m 個の特徴点について注目すると、次式の関係が得られる。

【0098】

【数5】

$$\begin{bmatrix} x_{j1}^{(k)} \\ \vdots \\ x_{jm}^{(k)} \end{bmatrix} = \mathbf{P}^{(k)} \begin{bmatrix} s_{j1} \\ \vdots \\ s_{jm} \end{bmatrix} \quad (6)$$

【0099】

ただし、行列 $\mathbf{P}^{(k)}$ は 2×3 の行列である。右辺の第2項の行列 $\mathbf{S}^{(k)}$ は行列 \mathbf{S} のうち、観測された特徴点に対応する要素のみからなる部分行列である。上述の通り、カメラと顔は十分に離れているとし正射影を仮定している。ここで、4点以上の特徴点が観測されれば、行列 $\mathbf{P}^{(k)}$ は以下のように計算できる（ステップS208）。

【0100】

【数6】

$$\mathbf{P}^{(k)} = (\mathbf{X}^{(k)})^t \mathbf{S}^{(k)} \{(\mathbf{S}^{(k)})^t \mathbf{S}^{(k)}\}^{-1}$$

$$\text{ただし、} \quad \mathbf{X}^{(k)} = \begin{bmatrix} x_{j1}^{(k)} & \dots & x_{jm}^{(k)} \end{bmatrix}^t \quad (7)$$

【0101】

画像フレーム I_k における眼球中心の投影位置 $x_r^{(i)}$ （太字）、 $x_l^{(i)}$ （太字）は、行列 $\mathbf{P}^{(k)}$ を用いて以下のように計算できる（ステップS210）。

【0102】

【数7】

$$x_r^{(k)} = \mathbf{P}^{(k)} s_r \quad (8)$$

$$x_l^{(k)} = \mathbf{P}^{(k)} s_l \quad (9)$$

【0103】

したがって、画像フレーム I_k において特徴点として抽出した虹彩中心の投影位置とこの眼球中心の投影位置を用いると、視線の推定を行なうことができる（ステップS212）。

10

20

30

40

50

【 0 1 0 4 】

なお、行列 P を QR 分解により分解することで、顔の姿勢 R が、以下のように計算できる。

【 0 1 0 5 】

【 数 8 】

$$P = A \begin{bmatrix} r_1 \\ r_2 \end{bmatrix} \quad (10)$$

$$R = \begin{bmatrix} r_1 \\ r_2 \\ r_1 \times r_2 \end{bmatrix} \quad (11)$$

10

【 0 1 0 6 】

ただし r_1 、 r_2 はそれぞれ 1×3 のベクトルである。このような顔の姿勢 R の検出については、文献：L. Quan: "Self-calibration of an affine camera from multiple views", Int'l Journal of Computer Vision, 19, pp. 93 - 105 (1996) に開示がある。

20

【 0 1 0 7 】

ユーザ等の指示により追跡が終了していると判断されれば (ステップ S 2 1 4)、処理は終了し、終了が指示されていなければ、処理はステップ S 2 0 2 に復帰する。

【 0 1 0 8 】

(実 験)

以上説明した視線方向の推定装置の有効性を確認するため、実画像を用いた実験を行なった結果について以下に説明する。

【 0 1 0 9 】

カメラはElmo社製PTC - 400Cを用い、被験者から約 1 5 0 [cm] の位置に設置した。

30

まず、50フレームの画像列を用いて、眼球中心と顔特徴点のキャリブレーションを行なった。キャリブレーション用の画像フレーム列と抽出した特徴点の例は、図 1 2 に示したとおりである。

【 0 1 1 0 】

キャリブレーション用画像フレーム列の撮影に要した時間は約 3 秒であった。(+ 印は抽出された虹彩中心 (眼球中心))、× 印は追跡した顔特徴点)。

【 0 1 1 1 】

次に、キャリブレーションにより求めた顔モデル (行列 S) を用いて、視線推定を行なった。ここで、被験者はそれぞれ右上、上、左下の方向を注視しながら、顔の位置・向きを変化させた。

40

【 0 1 1 2 】

図 1 4 ~ 図 1 6 は、視線推定結果を示す。図 1 4 は、右上方注視の状態であり、図 1 5 は、上方注視の状態であり、図 1 6 は、左下方注視の状態である。ここで、視線方向は両目それぞれで計算された視線方向の平均値としている。結果より、顔の位置・向きの変化とは関係なく、視線方向が推定できていることがわかる。

【 0 1 1 3 】

以上説明したとおり、実施の形態 1 の視線方向の推定装置では、単眼カメラの観測に基づいて顔特徴点を検出・追跡することにより視線方向を推定する。本発明の視線方向の推定装置では、まずキャリブレーションとして視線がカメラ方向を向いたまま顔の向きのみが異なる画像列から得られる虹彩位置と顔特徴点を利用することで、眼球中心と顔特徴点

50

の関係をモデル化し（行列 S を特定し）、その後、その関係に基づいて推定された入力画像中の眼球中心位置と虹彩位置の関係から視線方向を決定する。

【 0 1 1 4 】

実施の形態 1 の視線方向の推定装置は、単眼カメラによるパッシブな観測画像から視線方向を推定するため、顔位置が検出可能で光学的に十分なズーム比を得ることができれば、推定処理は観測距離の制約を受けない。これにより従来の視線推定装置では実現が難しかった遠方にいる人物に対する視線方向推定が可能になる。

【 0 1 1 5 】

[実施の形態 2]

実施の形態 1 の視線方向の推定装置においては、図 1 1 において説明したとおり、まず、キャリブレーション用の画像列として、ユーザがカメラを注視しながら、顔の姿勢を変化させている画像フレーム列を撮影していた。

【 0 1 1 6 】

実施の形態 2 の視線方向の推定装置においては、以下に説明するとおり、ユーザが必ずしもカメラを注視していない状態の画像列を用いて、キャリブレーションを行なうことが可能である。したがって、視線方向推定の被験者（被推定者）の、自然な状態での画像の撮影のみで、視線方向の推定を開始でき、被推定者への負担を低減できる。

【 0 1 1 7 】

なお、キャリブレーション後の視線方向の推定処理自体は、実施の形態 2 の視線方向の推定装置の動作は、実施の形態 1 の視線方向の推定装置の動作と同様である。

【 0 1 1 8 】

（キャリブレーション処理）

図 1 7 は、実施の形態 2 の視線方向の推定装置の初期設定として行なうキャリブレーションを説明するためのフローチャートである。

【 0 1 1 9 】

まず、キャリブレーション用の画像列として、ユーザが自由な方向を向き、顔の姿勢を変化させて画像列を取得する（ステップ S 3 0 2）。

【 0 1 2 0 】

（顔 3 次元モデルの生成）

以下の手続きにより、顔 3 次元モデルを生成する。

【 0 1 2 1 】

ここでは、 N ($N \geq 2$) 枚の画像列が得られたとする。各画像フレームを、フレーム I_1, \dots, I_N とする。

【 0 1 2 2 】

次に、得られた各画像フレーム列に対して、実施の形態 1 と同様な方法によって顔検出部 5 6 0 6 が顔検出処理を行ない（ステップ S 3 0 4）、続いて、顔領域抽出部 5 6 0 8 が、目や鼻、虹彩の検出処理を行なう（ステップ S 3 0 6）。

【 0 1 2 3 】

さらに、特徴点抽出部 5 6 1 0 が特徴点の抽出・追跡を行なう（ステップ S 3 0 8）。なお、特徴点の抽出方法も、実施の形態 1 と同様とする。

【 0 1 2 4 】

ここで、各画像フレーム I_i ($i = 1, \dots, N$) において M ($M \geq 4$) 点の特徴点 p_j ($j = 1, \dots, M$) が検出・追跡できたとする。画像フレーム I_i における特徴点 p_j の 2 次元観測位置を $x_j^{(i)}$ (太字) = $[x_j^{(i)}, y_j^{(i)}]^t$ ($i = 1, \dots, N, j = 1, \dots, M$) とする。

【 0 1 2 5 】

画像 I_i における特徴点 p_j の 2 次元観測位置を以下のとおりとする。

【 0 1 2 6 】

10

20

30

40

【数 9】

$$\mathbf{x}_j^{(i)} = [x_j^{(i)}, y_j^{(i)}]^t \quad (i = 1, \dots, N, j = 1, \dots, M)$$

【0 1 2 7】

画像 I_i における特徴点の重心を以下の記号 (1 2) で表す。

【0 1 2 8】

【数 1 0】

$$\bar{\mathbf{x}}^{(i)} \quad (12)$$

10

【0 1 2 9】

同様に、各観測位置を重心からの相対位置として、以下の記号 (1 3) で表す。

【0 1 3 0】

【数 1 1】

$$\hat{\mathbf{x}}_j^{(i)} \quad (13)$$

【0 1 3 1】

したがって、以下の式 (1 4) および式 (1 5) の関係が成り立つ。

20

【0 1 3 2】

【数 1 2】

$$\hat{\mathbf{x}}_j^{(i)} = \mathbf{x}_j^{(i)} - \bar{\mathbf{x}}^{(i)} \quad (14)$$

ただし、

$$\bar{\mathbf{x}}^{(i)} = \sum_j \mathbf{x}_j^{(i)} \quad (15)$$

30

【0 1 3 3】

このとき、相対観測位を並べた行列 W は、実施の形態 1 と同様に、因子分解法により以下の式 (1 6) (1 7) のように分解できる。ここで、カメラと顔は十分に離れているとし、弱透視変換を仮定している。

【0 1 3 4】

【数 1 3】

$$\mathbf{W} = \begin{bmatrix} \hat{x}_1^{(1)} & \cdots & \hat{x}_M^{(1)} \\ \vdots & & \vdots \\ \hat{x}_1^{(N)} & \cdots & \hat{x}_M^{(N)} \end{bmatrix} \quad (16)$$

$$= \begin{bmatrix} m_1^t \\ \vdots \\ m_N^t \end{bmatrix} \begin{bmatrix} s_1 & \cdots & s_M \end{bmatrix}$$

$$= \mathbf{MS} \quad (17)$$

10

【0 1 3 5】

実施の形態 1 と同様に、行列 M (「撮影姿勢行列」) にはカメラの姿勢変化に関する情報が、行列 S (「相対位置関係行列」) には観測対象物の形状に関する情報が含まれており、顔特徴点間の相対関係は行列 S として求められる (ステップ S 3 1 0)。

20

【0 1 3 6】

次に、画像 I_k が得られ、M 個の特徴点のうち p 個がそれぞれ以下のように観測されたとする。

【0 1 3 7】

【数 1 4】

$$x_{j_1}^{(k)}, \dots, x_{j_p}^{(k)}$$

30

【0 1 3 8】

ここで、特徴点の重心位置を以下の記号 (18) で表す。

【0 1 3 9】

【数 1 5】

$$\bar{x}^{(k)} \quad (18)$$

【0 1 4 0】

このとき、カメラの姿勢変化を表す行列 m_k および式 (18) で表される特徴点の重心位置は、以下の式 (19) のようにして求められる。

40

【0 1 4 1】

【数 1 6】

$$\begin{bmatrix} \mathbf{m}_k & \bar{x}^{(k)} \end{bmatrix} = \begin{bmatrix} x_{j_1}^{(k)} & \cdots & x_{j_p}^{(k)} \end{bmatrix} \begin{bmatrix} s_{j_1}^{(k)} & \cdots & s_{j_p}^{(k)} \\ 1 & \cdots & 1 \end{bmatrix}^* \quad (19)$$

【0 1 4 2】

ただし、行列 m_k は 2×3 の行列、 $[\]^*$ は疑似逆行列を示す。

さらに、行列 m_k を式 (20) のように QR 分解することで、画像 I_k における顔の姿

50

勢 R_k 、位置 t_k が以下の式 (21) のように計算できる。

【0143】

【数17】

$$\mathbf{m}_k = \begin{bmatrix} s_{x,k} & 0 \\ a_k & s_{y,k} \end{bmatrix} \begin{bmatrix} r_1^k \\ r_2^k \end{bmatrix} \quad (20)$$

$$\mathbf{R}_k = \begin{bmatrix} r_1^1 \\ r_2^1 \\ r_1^1 \times r_2^1 \end{bmatrix}^t \begin{bmatrix} r_1^k \\ r_2^k \\ r_1^k \times r_2^k \end{bmatrix}, \quad \mathbf{t}_k = \bar{\mathbf{X}}^{(k)} - \bar{\mathbf{X}}^{(1)} \quad (21)$$

【0144】

ただし、以下の式 (22) の s_k はスケールファクタを表す。

【0145】

【数18】

$$s_k = (s_{x,k} + s_{y,k}) / 2 \quad (22)$$

$$(|a_k| \ll |s_k|)$$

【0146】

なお、上記の式 (20) に含まれる a_k は画像面の傾きを表わし、通常は 0 (ゼロ) とみなされる。

【0147】

したがって、以下では、被験者の顔は、正規化された大きさを有するものとして扱うことができ、被験者とカメラの距離による画像中の顔の大きさの変化を考慮する必要がない。

【0148】

さらに、後に詳しく説明する手順により、眼球モデルパラメータ (眼球中心位置、眼球半径、虹彩半径、画像上における虹彩位置) の推定を行なう (ステップ S312)。

【0149】

(眼球モデルパラメータの推定)

以下では、眼球モデルパラメータの推定処理について、さらに詳しく説明する。なお、以下の処理は、眼球中心推定部 5616 が行なうものである。

【0150】

視線方向を推定するには、前節で求められた顔座標系における眼球中心位置 $X_{\{LR\}}$ (なお、以下では、下付文字 $\{LR\}$ は、左を意味する l 、右を意味する r を総称するものとして使用する)、および眼球半径 l 、画像上における虹彩位置 $X_{\{LR\},k}$ を求める必要がある。

【0151】

入力画像において虹彩位置を効率的に決定するためには、さらに予め虹彩半径 r を得ておくことが望ましい。

【0152】

(眼球 3次元モデルの生成)

図 18 は、眼球モデルパラメータの推定処理の手続きを示す図である。

10

20

30

40

50

【 0 1 5 3 】

以下では、図 1 8 を用いて、上述したキャリブレーションで用いた N フレームの画像列を用いてこれらのパラメータを推定する処理について説明する。

【 0 1 5 4 】

(ステップ 1 : 眼球中心位置初期値の算出)

まず、各フレームの顔位置・姿勢 t_k 、 R_k 、位置 t_k 、スケールファクタ s_k およびステップ S 3 0 6 で求めた虹彩位置から、RANSAC を利用して大まかな眼球中心位置を推定する。

【 0 1 5 5 】

(RANSAC : Random sample consensus)

以下で説明する RANSAC 処理は、外れ値を含むデータから安定にモデルパラメータを定めるための処理であり、これについては、たとえば、以下の文献に記載されているので、その処理の概略を説明するにとどめる。

【 0 1 5 6 】

文献 : M.A.Fischler and R.C.Bolles: " Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography, " Comm. Of the ACM, Vol.24, pp.381-395,1981

文献 : 大江統子、佐藤智和、横矢直和 : " 画像徳著点によるランドマークデータベースに基づくカメラ位置・姿勢推定 "、画像の認識・理解シンポジウム (MIRU2005) 2 0 0 5 年 7 月

RANSAC により観測データから眼球中心位置 X_L 、 X_R (初期値) を計算する手順式 (19) より、フレーム k におけるカメラ幾何を表す行列 m_k および式 (18) で表される特徴点の重心位置を得る。

【 0 1 5 7 】

ここで行列 m_k には、式 (20)、(21) で示したように、フレーム k におけるカメラの位置・姿勢および撮影像のスケールファクタの情報が含まれる。

【 0 1 5 8 】

画像から観測される i フレーム目の画像における虹彩中心位置を以下の式 (23) とする。

【 0 1 5 9 】

【数 19】

$$\mathbf{X}_{\{LR\},i} = \begin{bmatrix} x_{\{LR\},i} \\ y_{\{LR\},i} \end{bmatrix} \quad (23)$$

【 0 1 6 0 】

ランダムに q 個のフレーム f_1, f_2, \dots, f_q が選択されたとすると、それらの観測データから眼球中心位置は次式 (24) により計算できる。

【 0 1 6 1 】

【数 20】

$$\mathbf{X}_{\{LR\}} = \left(\begin{bmatrix} m_{f_1} \\ \vdots \\ m_{f_q} \end{bmatrix} \begin{bmatrix} m_{f_1} \\ \vdots \\ m_{f_q} \end{bmatrix} \right)^{-1} \begin{bmatrix} m_{f_1} \\ \vdots \\ m_{f_q} \end{bmatrix} \begin{bmatrix} x_{\{LR\},f_1} - \bar{x}^{(f_1)} \\ \vdots \\ x_{\{LR\},f_q} - \bar{x}^{(f_q)} \end{bmatrix} \quad (24)$$

10

20

30

40

50

【 0 1 6 2 】

得られた眼球中心位置を入力画像 i ($i = 1, \dots, N$) 上に再投影し、以下の式 (25) で表される虹彩中心とのずれ e_i を評価する。

【 0 1 6 3 】

【 数 2 1 】

$$\mathbf{e}_i = (\mathbf{x}_{\{LR\},i} - \bar{\mathbf{x}}^{(i)}) - \mathbf{m}_i \mathbf{X}_{\{LR\}} \quad (25)$$

【 0 1 6 4 】

図 19 は、このような再投影とずれの関係を示す概念図である。

10

眼球中心の再投影位置と虹彩中心位置は、眼球の回転によって最大で $s_i l$ のずれが生じ得るため、ずれ $|e_i|$ が $s_i l$ 以下であれば、以下の式 (26) のとおり誤差ゼロとみなす (l は眼球半径、 s_i はスケールファクタ)。

【 0 1 6 5 】

【 数 2 2 】

$$\hat{e}_i = \begin{cases} 0 & |e_i| < ls_i \\ |e_i| & otherwise \end{cases} \quad (26)$$

20

【 0 1 6 6 】

誤差 e_i が一定値以下となるフレームが既定数以上あれば、それらのデータを利用して式 (24) により眼球中心位置を再度計算し、式 (26) により利用したフレームのデータに関する誤差を評価する。

【 0 1 6 7 】

以上の処理を複数回行ない、誤差の総計がもっとも小さくなる時の眼球中心位置を出力値とする。

【 0 1 6 8 】

(ステップ 2 : 眼球モデルパラメータの推定)

得られた眼球中心位置を初期値として、入力画像群に対して眼球モデルを当てはめ最適なモデルパラメータを推定する。ここで、入力画像から目の周辺領域を切り出し、色および輝度情報をもとに、以下の式 (27) に従って、虹彩 (黒目)、白目、肌領域の 3 種類にラベル付けを行なう。

30

【 0 1 6 9 】

【 数 2 3 】

$$u_{i,j} = \begin{cases} 0 & \min_k |h_{i,j} - h_{s,k}| < \sigma_h \\ 1 & \min_k |h_{i,j} - h_{s,k}| \geq \sigma_h, \quad v_{i,j} < \sigma_v \\ -1 & otherwise \end{cases} \quad (27)$$

40

【 0 1 7 0 】

ここで、 $h_{s,k}$ は、肌領域の k 番目の画素の色相 (hue) の値を表わす。 $h_{i,j}$ は、入力画像中の画素 i, j の色相の値を表わす。 $v_{s,k}$ は、入力画像中の画素 i, j の明度の値を表わす。

【 0 1 7 1 】

図 20 は、このようなラベリング処理例を示す図である。

続いて各画素が虹彩モデルの内側にあるかどうかをチェックし、眼球モデルとの照合度を評価する (非線形最適化)。

50

【 0 1 7 2 】

図 2 1 は、このような虹彩とモデルとの照合処理の概念を示す図である。

このような非線形最適化処理を行なうにあたり、以下の距離 $d_{\{LR\},i,j}$ を導入する。

【 0 1 7 3 】

【 数 2 4 】

$$d_{\{LR\},i,j} = \sqrt{(x_{i,j} - x_{\{LR\},i})^2 + (y_{i,j} - y_{\{LR\},i})^2} \quad (28)$$

【 0 1 7 4 】

10

一方、 $r_{\{LR\},i,j}$ は虹彩中心から画素 i, j (フレーム i の j 番目の画素) 方向の虹彩半径を示すとすると、図 2 1 に示すとおり、画素 i, j が虹彩の外側にあれば、 $d_{\{LR\},i,j}$ は、 $r_{\{LR\},i,j}$ よりも大きな値を示す。

【 0 1 7 5 】

$r_{\{LR\},i,j}$ は、以下の式 (29) に示すように、眼球中心位置、虹彩位置、眼球半径、虹彩半径の関数となる。

【 0 1 7 6 】

【 数 2 5 】

$$r_{\{LR\},i,j} = f(\mathbf{X}_{\{LR\}}, l, r, s_i, \mathbf{R}_i, \mathbf{t}_i, x_{i,j}, y_{i,j}) \quad (29)$$

20

【 0 1 7 7 】

なお、顔の相対座標を正規化して考えているので、本来は、眼球中心位置は、フレームに拘わらず、一定の位置に存在するはずである。

【 0 1 7 8 】

最後に、眼周辺の全画素について $d_{\{LR\},i,j}$ の評価を行ない、入力画像群に尤もよく当てはまる以下の式 (30) のモデルパラメータ を、式 (31) に従って決定する。

【 0 1 7 9 】

【 数 2 6 】

$$\theta = [\mathbf{X}_L, \mathbf{X}_R, l, r, \mathbf{x}_{L,1}, \dots, \mathbf{x}_{L,N}, \mathbf{x}_{R,1}, \dots, \mathbf{x}_{R,N}] \quad (30)$$

30

$$\theta = \arg \min_{\theta} \sum_{i,j} g_{i,j,\{LR\}} \quad (31)$$

【 0 1 8 0 】

ここで、 $g_{i,j,\{LR\}}$ は、フレーム i 、画素 j における $d_{\{LR\},i,j}$ の評価値であり、対象画素が虹彩領域か白目領域かによって、以下の式に従い、符合を反転させる。

【 0 1 8 1 】

40

【数 2 7】

$$g_{i,j,\{LR\}} = u_{i,j} \cdot G_{i,j,\{LR\}},$$

$$G_{i,j,\{LR\}} = \begin{cases} -1 & e_{\{LR\},i,j} < -1 \\ 1 & e_{\{LR\},i,j} > 1 \\ e_{\{LR\},i,j} - r_{\{LR\},i,j} & otherwise \end{cases}$$

10

ただし

$$e_{\{LR\},i,j} = d_{\{LR\},i,j} - r_{\{LR\},i,j}$$

【0182】

ラベリング u_{ij} が撮影された画像内の虹彩領域を反映し、関数 $G_{i,j,\{LR\}}$ は、眼球モデルから算出される虹彩領域を反映している。

【0183】

以上の処理により、ユーザが自由な方向を向いた状態の画像でキャリブレーションを行なうことができ、かつ、眼球モデルパラメータも決定できることになる。

20

【0184】

以下、視線方向の推定処理の手続きについては、実施の形態 1 と同様である。

(視線方向の推定)

すなわち、視線方向は眼球中心と虹彩中心を結ぶ 3 次元直線として与えられるものとしてモデル化する。画像上での眼球半径を l 、画像上での眼球中心と虹彩中心との x 軸方向、 y 軸方向の距離を d_x 、 d_y とすると、視線方向とカメラ光軸とのなす角、つまり、視線方向を向くベクトルが x 軸および y 軸との成す角 φ_x 、 φ_y は次式で表される。

【0185】

【数 2 8】

$$\varphi_x = \sin^{-1}\left(\frac{d_x}{l}\right), \quad \varphi_y = \sin^{-1}\left(\frac{d_y}{l}\right)$$

30

【0186】

以上のような実施の形態 2 の視線方向の推定装置、推定方法によっても、視線方向の推定について実施の形態 1 と同様の効果を奏することが可能である。

【0187】

(実験結果)

以上説明した手法の有効性を確認するため、実画像を用いた実験を行なった。まず、50 フレームの画像列を用いて、顔の位置・姿勢変化および眼球モデルパラメータのキャリブレーションを行なった。キャリブレーション用画像列の撮影に要した時間は約 2 秒であり、このときに得られた眼球半径および虹彩半径はそれぞれ 11.21 画素、 6.27 画素であった。

40

【0188】

図 2 2 は、キャリブレーションにより得られた顔モデルを用いて、視線推定を行なった結果を示す図である。ここで、被験者はそれぞれ右、正面、左の方向を注視しながら、顔の位置・向きを変化させている(白×印は眼球中心の投影位置推定値)。ここで、視線方向は両目それぞれで計算された視線方向の平均値としている。結果より、顔の位置・向き

50

変化がある場合でも提案手法により視線方向が推定できていることがわかる。

【0189】

[実施の形態3]

実施の形態1および実施の形態2においては、まず、眼球モデルパラメータの推定処理を、たとえば、Nフレームの画像列を用いて行なってから、視線方向の推定処理に移っていた。このとき、視線方向の推定処理においては、虹彩中心を求めるために、たとえば、ハフ変換を用いた処理を行なうこととしていた。

【0190】

たとえば、実施の形態2においては、図18において説明したとおり、眼球モデルパラメータの推定処理において、まず、ステップ1として、「眼球中心の初期値の算出」を行ない、次に、ステップ2として、「眼球モデルパラメータの推定」を非線形最適化処理により行なっていた。

10

【0191】

実施の形態3においては、さらに、視線方向の推定処理においても、虹彩中心位置を求めるために、「非線形最適化処理」を行なう例について説明する。

【0192】

そこで、視線方向の推定処理での非線形最適化処理について説明する前提として、まずは、眼球モデルパラメータの推定における「非線形最適化処理」について、もう一度簡単にまとめる。

【0193】

20

図23は、このような眼球モデルパラメータの推定における「非線形最適化処理」を説明するための概念図である。

【0194】

図23を参照して、Nフレームの画像列を用いて、たとえば、実施の形態2と同様にして、ステップ1として、RANSACを用いて、「眼球中心位置初期値の算出」を行なう。なお、眼球中心位置初期値としては、このようにして求めた値を用いることに限定されず、たとえば、解剖学的な知見から得られた平均的な値を用いることも可能である。

【0195】

続いて、ステップ2として、得られた眼球中心位置を初期値として、入力画像群に対して眼球モデルを当てはめ最適なモデルパラメータを、逐次推定する。

30

【0196】

フレーム1～フレームNについて、目の周辺領域を切り出し、色および輝度情報をもとに、前述した式(27)に従って、虹彩(黒目)、白目、肌領域の3種類にラベル付けを行なう。

【0197】

続いて、各画素が虹彩モデルの内側にあるかどうかをチェックし、眼球モデルとの照合度を評価する(非線形最適化)。

【0198】

すなわち、現時点のステップでのモデルパラメータに対応する、虹彩中心から画素 i, j (フレーム i の j 番目の画素)方向の虹彩半径 $r_{\{LR\}, i, j}$ を用いて、式(28)で表される距離 $d_{\{LR\}, i, j}$ を評価する。

40

【0199】

この評価では、上述した式(30)(31)を用いて、眼球モデルパラメータ(眼球半径、虹彩半径、眼球中心位置)を更新しつつ、眼球中心を再投影した画像とラベル付けされたフレームとを照合する処理を繰り返すことで、式(31)を満たすようにモデルパラメータを決定する。

【0200】

(非線形最適化を用いた視線方向推定)

以上のようにして、モデルパラメータが決定された後に、新たに観測される入力画像について、ここまでで得られた顔・眼球の3次元モデルを利用して、視線方向を推定する。

50

【0201】

図24は、このような視線推定のために虹彩中位置を決定する処理を示す概念図である。ここで、このような虹彩中心位置の決定は、図13で説明される視線方向推定処理のうち、ステップS212で行なわれる処理に対応する。

【0202】

時刻 t における画像 I_t (フレーム t)に対し、まず顔特徴点の追跡を行なう。顔特徴点が $x_1^{(t)}$ (太字)、 \dots 、 $x_M^{(t)}$ (太字)の位置にそれぞれ観測されたとする。

【0203】

ここで、上述した式(8)(9)の行列 $P^{(k)}$ (太字)は4点以上の顔特徴点が観測されれば計算できる。よって画像 I_t における行列 $P^{(t)}$ (太字)と眼球中心の3次元位置 $X_{\{LR\}}$ (太字)を用いて、画像 I_t における両目の眼球中心位置(2次元) $x_{e\{LR\}}$ (太字)は、以下のようにして計算できる。

【0204】

【数29】

$$\mathbf{x}_{e\{LR\}}^{(t)} = \mathbf{P}^{(t)} \mathbf{X}_{\{LR\}}$$

【0205】

続いて、虹彩中心位置を推定する。上で求めた眼球中心位置をもとに目周辺領域を切り出し、眼球モデルパラメータの推定時と同様に、図20で説明したように以下の式にしたがって、色および輝度情報をもとに虹彩(黒目)、白目、肌領域の3種類にラベル付けを行なう。

【0206】

【数30】

$$u_{t,j} = \begin{cases} 0 & \min_k |h_{t,j} - h_{s,ki}| < \sigma_h \\ 1 & \min_k |h_{t,j} - h_{s,ki}| \geq \sigma_h, \quad v_{t,j} < \sigma_v \\ -1 & \text{otherwise} \end{cases}$$

【0207】

続いて、やはり、図21で説明した眼球モデルパラメータの推定時と同様に、以下の距離 $d_{\{LR\},i,j}$ を用いて、各画素が虹彩モデルの内側にあるかどうかをチェックし、眼球モデルとの照合度を評価する。

【0208】

【数31】

$$d_{\{LR\},t,j} = \sqrt{(x_{t,j} - x_{\{LR\},t})^2 + (y_{t,j} - y_{\{LR\},t})^2}$$

【0209】

ここで、 $r_{\{LR\},t,j}$ は、虹彩中心から画像 I_t の画素 j 方向の虹彩半径を示しており、画像 I_t において画素 j が虹彩の外側にあれば、 $d_{\{LR\},t,j}$ は正の値を示す。

【0210】

$r_{\{LR\},t,j}$ は、眼球中心位置、虹彩位置、眼球半径、虹彩半径の関数となるが、眼球中心位置、眼球半径、虹彩半径は、眼球モデルパラメータとして既に得られている。

【0211】

したがって、 $r_{\{LR\},t,j}$ は、虹彩位置 $x_{\{LR\},t}$ (太字)の関数と見なすこと

10

20

30

40

50

ができるので、以下のように表現できる。

【0212】

【数32】

$$r_{\{LR\},t,j} = f(\mathbf{x}_{\{LR\},t})$$

【0213】

よって、眼周辺の全画素について $d_{\{LR\},t,j}$ の評価を行ない、画像 It に尤もよく当てはまるパラメータ $\theta = [x_{L,t} \text{ (太字)}, x_{R,t} \text{ (太字)}]$ を、非線形最適化の手続きで決定することで、虹彩中心位置が計算できる。この非線形最適化の手続きにあたっては、以下の式について最適化を行なう。

10

【0214】

【数33】

$$\theta = \arg \min_{\theta} \sum_j g_{t,j,\{LR\}}$$

【0215】

ここで、 $g_{t,j,\{LR\}}$ は、画像 It 、画素 j における $d_{\{LR\},t,j}$ の評価値であり、対象画素が虹彩領域か白目領域かによって、以下の式に従い、符合を反転させる。

20

【0216】

【数34】

$$g_{t,j,\{LR\}} = u_{t,j} \cdot G_{t,j,\{LR\}},$$

$$G_{t,j,\{LR\}} = \begin{cases} -1 & e_{\{LR\},t,j} < -1 \\ 1 & e_{\{LR\},t,j} > 1 \\ e_{\{LR\},t,j} - r_{\{LR\},t,j} & otherwise \end{cases}$$

30

ただし、

$$e_{\{LR\},t,j} = d_{\{LR\},t,j} - r_{\{LR\},t,j}$$

【0217】

最後に、以上で求めた眼球中心位置と虹彩中心位置より視線方向を計算する。画像上での眼球半径を l 、画像上での眼球中心と虹彩中心との x 軸方向、 y 軸方向の距離を dx 、 dy とすると、視線方向とカメラ光軸とのなす角、つまり、視線方向を向くベクトルが x 軸および y 軸との成す角 φ_x 、 φ_y は次式で表される。

40

【0218】

【数35】

$$\varphi_x = \sin^{-1}\left(\frac{dx}{l}\right), \quad \varphi_y = \sin^{-1}\left(\frac{dy}{l}\right)$$

50

【0219】

ただし、画像上での眼球半径 l は3次元眼球半径と行列 $P(t)$ から計算できる。

以上のような処理により、視線方向の推定処理において、虹彩位置を各フレームにおいて、非線形最適化により求めることとしたので、より高精度に視線方向の推定を行なうことが可能となる。

【0220】

[実施の形態4]

実施の形態2および実施の形態3においては、眼球モデルパラメータの推定処理については、既に述べた複数フレーム画像を入力としたRANSACによる初期値推定と非線形最適化の組合せを用いていた。

【0221】

しかし、実施の形態4では、実施の形態2または実施の形態3の構成において、「眼球モデルパラメータの推定処理」を、以下に説明するような「逐次型眼球モデル推定」の処理に置き換える。

【0222】

図25は、このような眼球モデルパラメータの推定処理を「逐次型眼球モデル推定」の処理に置き換えた場合の処理の流れを説明する概念図である。なお、実施の形態4でも、キャリブレーション処理において、顔3次元モデルの生成処理が、複数枚の画像に基づいてなされているものとする。

【0223】

すなわち、実施の形態4の眼球モデルパラメータの推定処理については、既に述べた複数フレーム画像を入力としたRANSACによる初期値推定と非線形最適化の組合せではなく、平均的なモデルパラメータを初期値とした逐次型のアルゴリズムを用いることもできる。

【0224】

図25を参照して、このアルゴリズムの実装例について説明している。まずアルゴリズムの開始時点では、被験者実験により対象ユーザの平均値を求めておく等の方法で得た眼球位置 X^0_L (太字)、 X^0_R (太字)、大きさ眼球半径 l^0 、虹彩半径 r^0 を初期パラメータとして、眼球モデルパラメータをたとえばハードディスク54に保持している。

【0225】

CPU56は、第1フレームに対するラベリング結果および顔姿勢を入力として、初期パラメータを出発点として、実施の形態3で説明した視線方向推定の処理と同様の非線形最適化処理によって眼球モデルパラメータ X^1_L (太字)、 X^1_R (太字)、大きさ眼球半径 l^1 、虹彩半径 r^1 および第1フレームにおける虹彩中心位置 $x_{L,1}$ (太字)、 $x_{R,1}$ (太字) を得て、たとえば、ハードディスク54に格納する。得られた虹彩中心位置および眼球中心位置から第1フレームにおける視線方向を計算することができる。

【0226】

次フレーム以降の処理においては、前フレームで得られたモデルパラメータを初期値とし、新たに得られる入力データを加えて非線形最適化処理を行なうことでモデルパラメータの更新および当該フレームにおける虹彩中心位置の推定を行なうことができる。

【0227】

このような処理を行なうと、Nフレームの画像を取得して、キャリブレーション処理により、顔3次元モデルの生成した後、さらに、眼球モデルパラメータの推定処理の終了を待ってから、視線の推定処理を開始する場合に比べて、視線方向の推定処理を短時間で開始できるという利点がある。

【0228】

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

20

30

40

50

【図面の簡単な説明】

【0229】

【図1】本発明の実施の形態にかかるシステムの外觀図である。

【図2】視線検出装置の外觀を示す図である。

【図3】本発明の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。

【図4】CPU56が、視線方向推定のプログラムに基づいて行なうソフトウェア処理を示す機能ブロック図である。

【図5】眉間候補領域を検出するためのフィルタを説明するための概念図である。

【図6】6分割矩形フィルタの他の構成を示す概念図である。

【図7】眉間を中心とした画像領域を利用してSVMによるモデル化を説明する図である。

【図8】顔検出結果の例を示す図である。

【図9】視線方向を決定するためのモデルを説明する概念図である。

【図10】ユーザがカメラを注視する状態に移行した後の虹彩中心、眼球中心および投影点の関係を示す概念図である。

【図11】視線方向の推定装置の初期設定の処理のフローを説明するためのフローチャートである。

【図12】キャリブレーションにおいて撮影された4枚の画像フレームを示す図である。

【図13】視線方向の推定装置が実行するリアルタイム視線検出の処理のフローを説明するためのフローチャートである。

【図14】右上方注視の状態での視線推定結果を示す図である。

【図15】上方注視の状態での視線推定結果を示す図である。

【図16】左下方向注視の状態での視線推定結果を示す図である。

【図17】実施の形態2の視線方向の推定装置の初期設定として行なうキャリブレーションを説明するためのフローチャートである。

【図18】眼球モデルパラメータの推定処理の手続きを示す図である。

【図19】再投影とずれの関係を示す概念図である。

【図20】ラベリング処理例を示す図である。

【図21】虹彩とモデルとの照合処理の概念を示す図である。

【図22】キャリブレーションにより得られた顔モデルを用いて、視線推定を行なった結果を示す図である。

【図23】眼球モデルパラメータの推定における「非線形最適化処理」を説明するための概念図である。

【図24】視線推定のために虹彩中位置を決定する処理を示す概念図である。

【図25】眼球モデルパラメータの推定処理を「逐次型眼球モデル推定」の処理に置き換えた場合の処理の流れを説明する概念図である。

【符号の説明】

【0230】

20 視線方向の推定装置、30 カメラ、40 コンピュータ本体、42 モニタ、
 5602 画像キャプチャ処理部、5604 画像データ記録処理部、5606 顔検出
 部、5608 顔領域抽出部、5610 特徴点抽出部、5612 虹彩中心抽出部、5
 614 相対関係特定部、5616 眼球中心推定部、5618 視線方向推定部、56
 30 表示制御部。

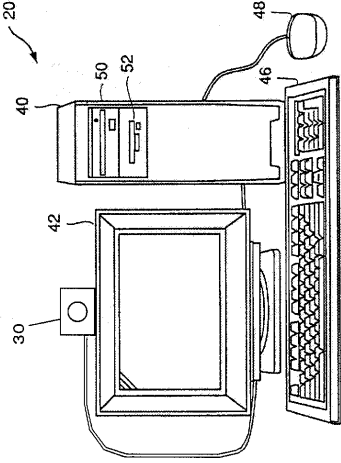
10

20

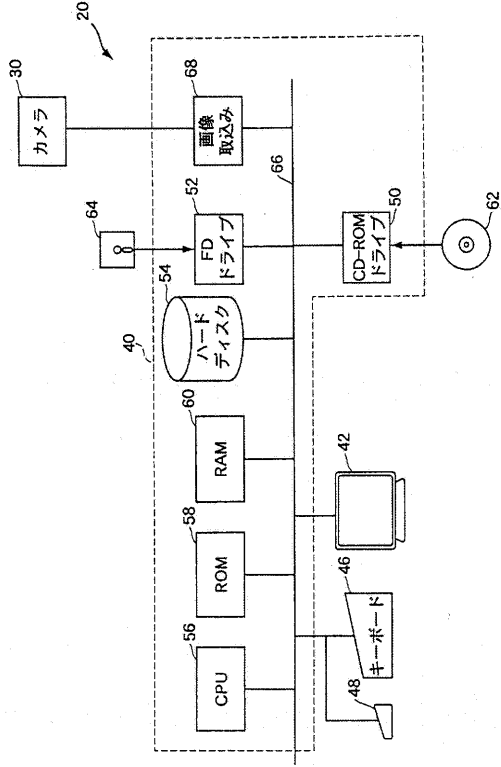
30

40

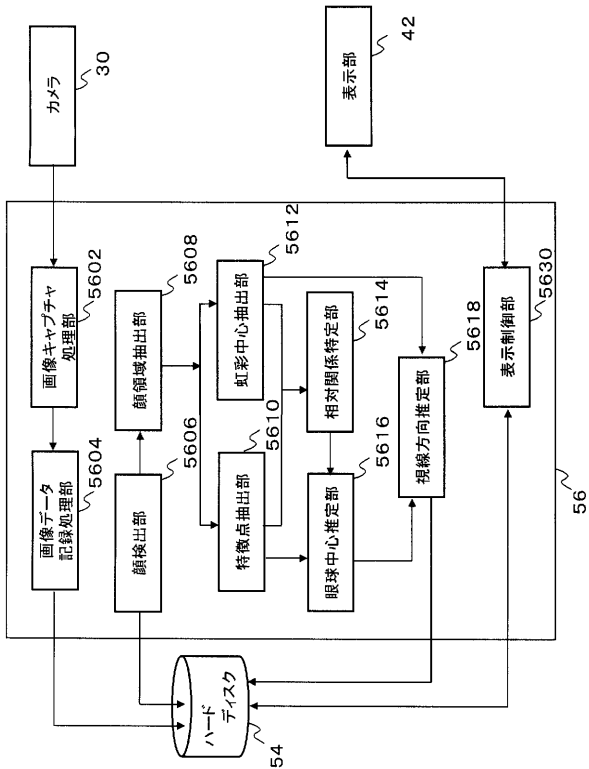
【図1】



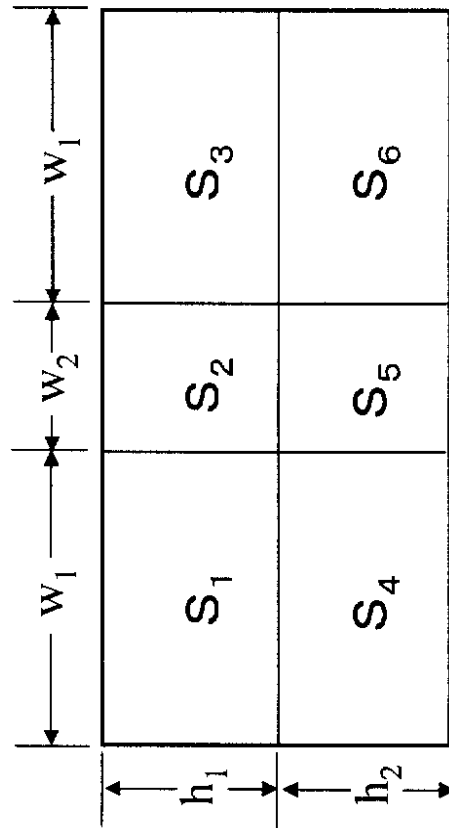
【図3】



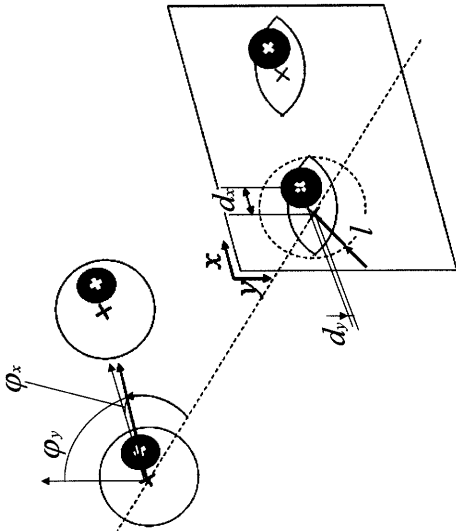
【図4】



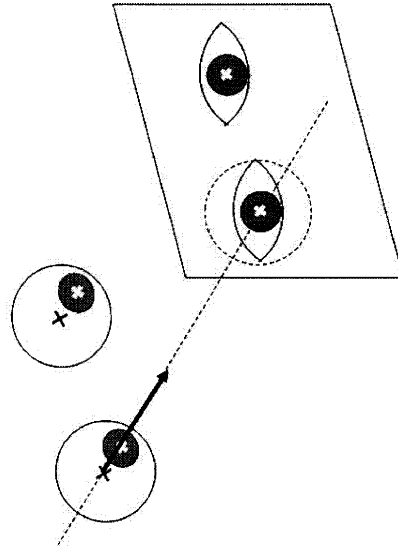
【図6】



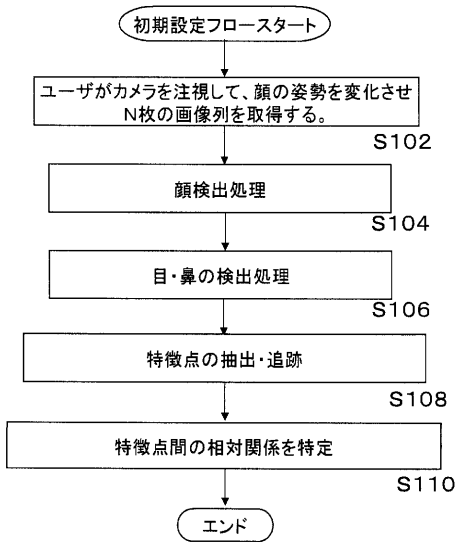
【図9】



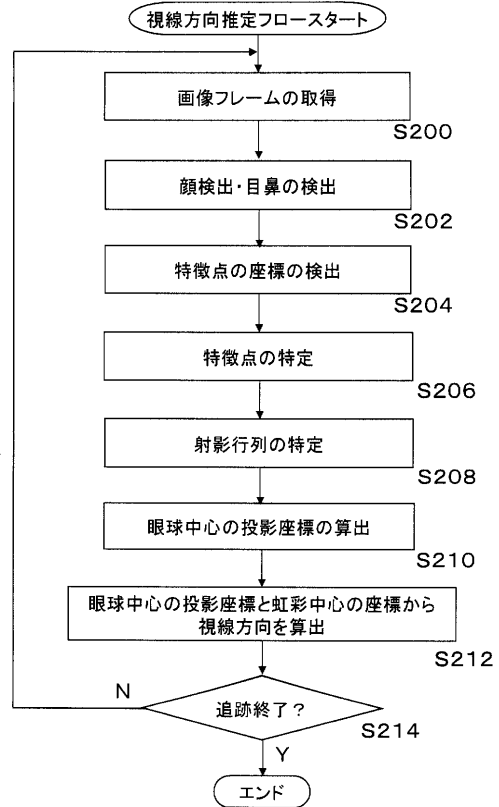
【図10】



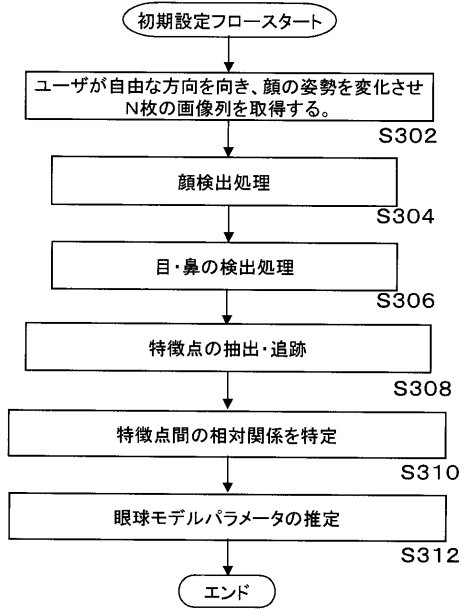
【図11】



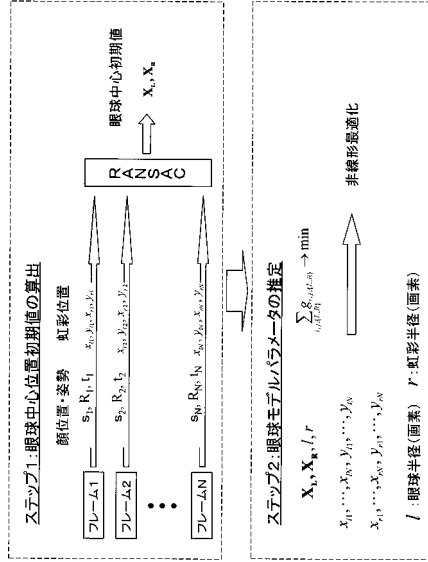
【図13】



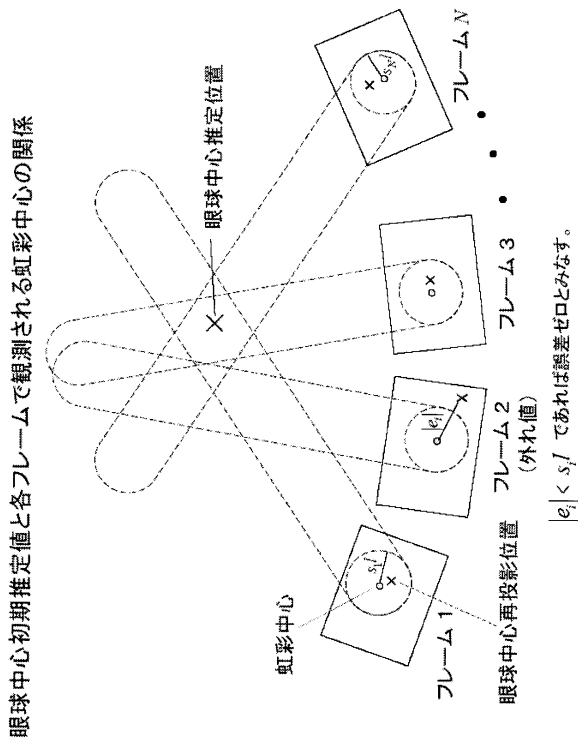
【図17】



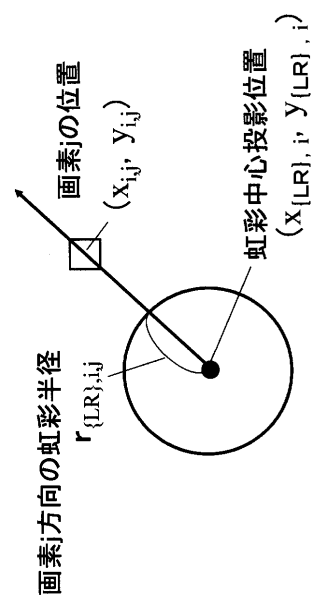
【図18】



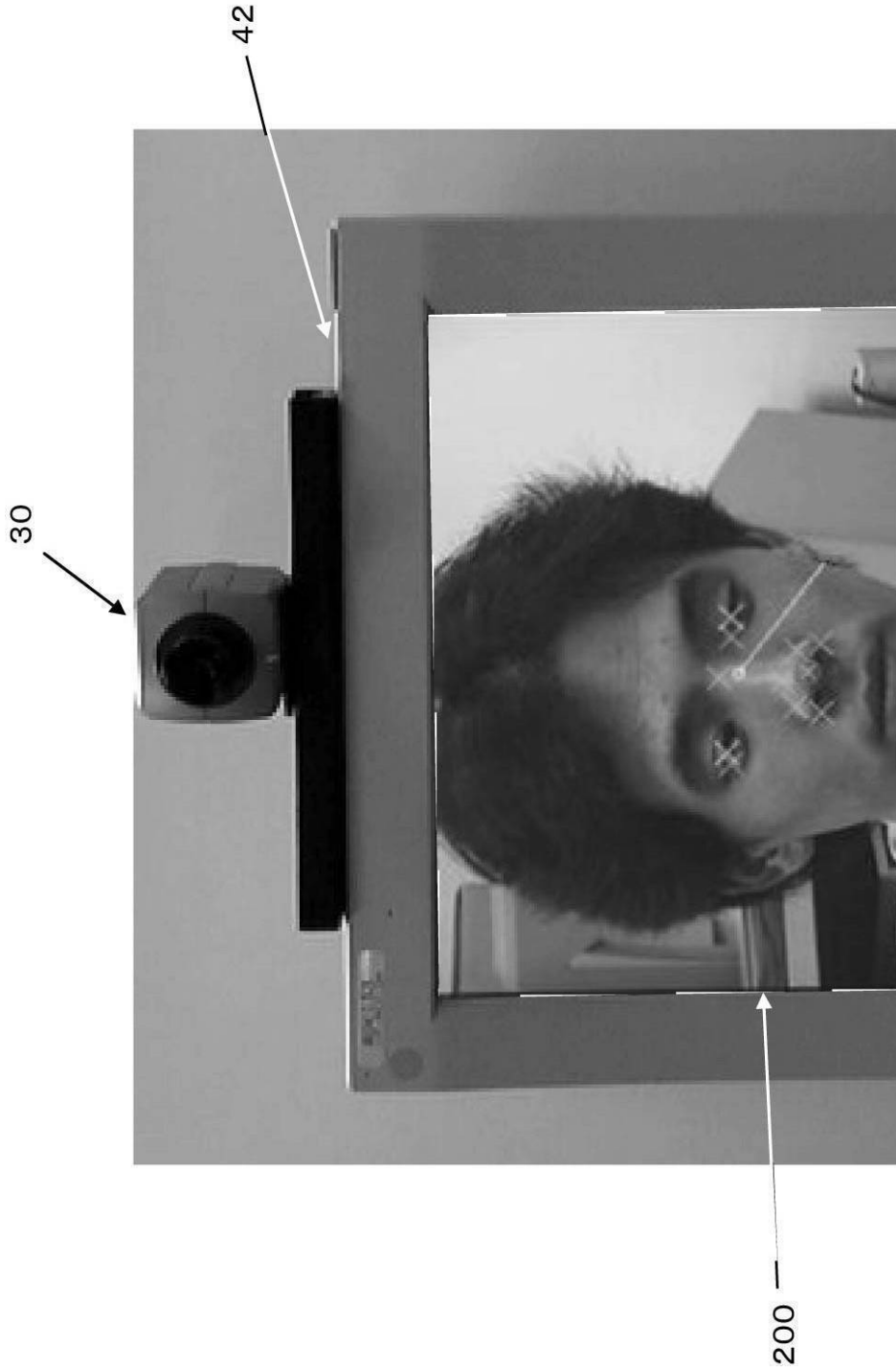
【図19】



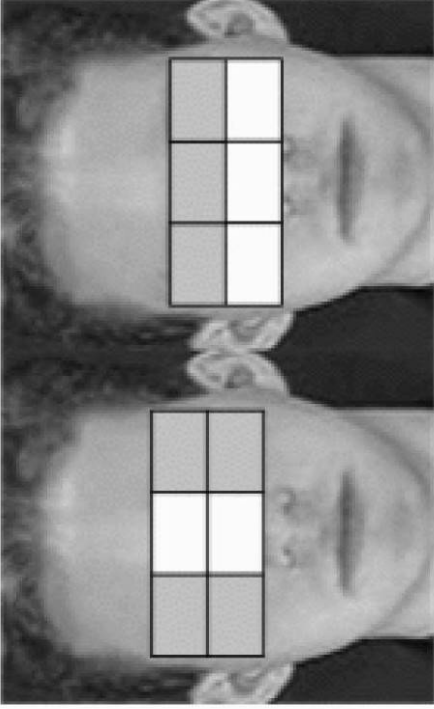
【図21】



【 図 2 】



【 図 5 】

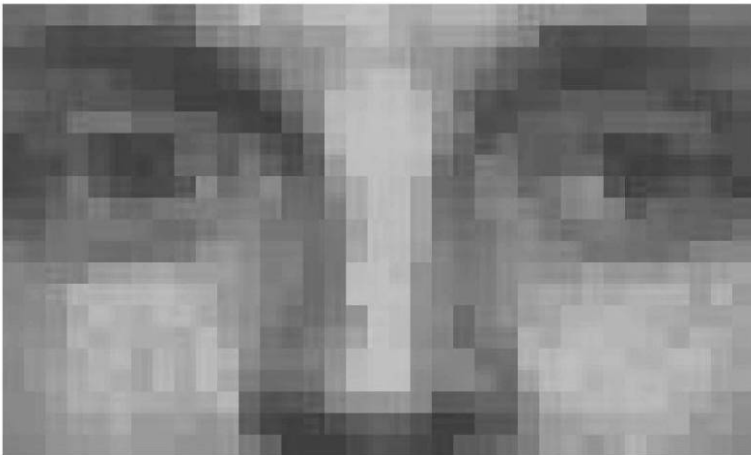


(b)

| | | |
|----|----|----|
| S1 | S2 | S3 |
| S4 | S5 | S6 |

(a)

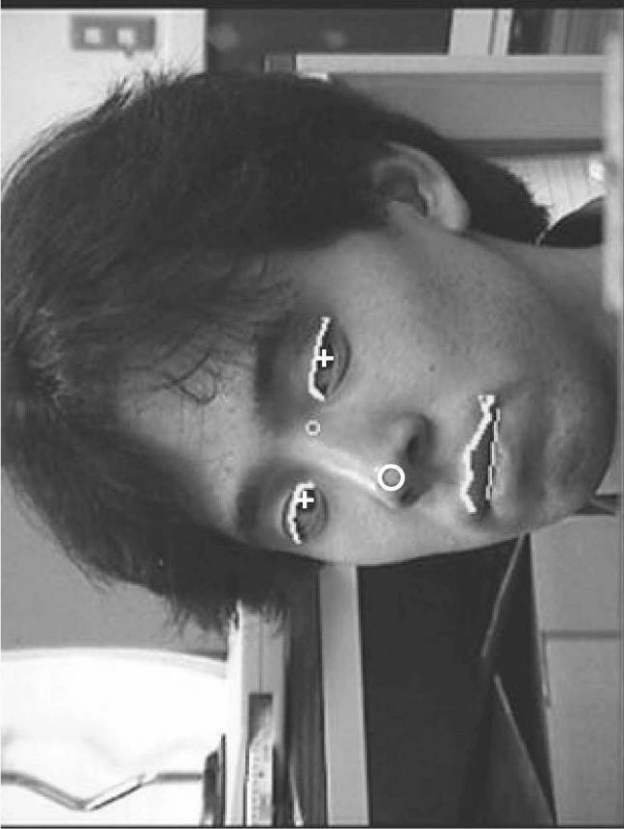
【 図 7 】



【 図 8 】



(b)



(a)

キャラブレージョン

【 1 2 】



(a)



(b)



(c)



(d)

【 図 1 4 】

右上方注視



(a)

(b)

(c)

(d)

上方注視

【 図 1 5 】



(a)

(b)

(c)

(d)

【 図 16 】

左下方注視



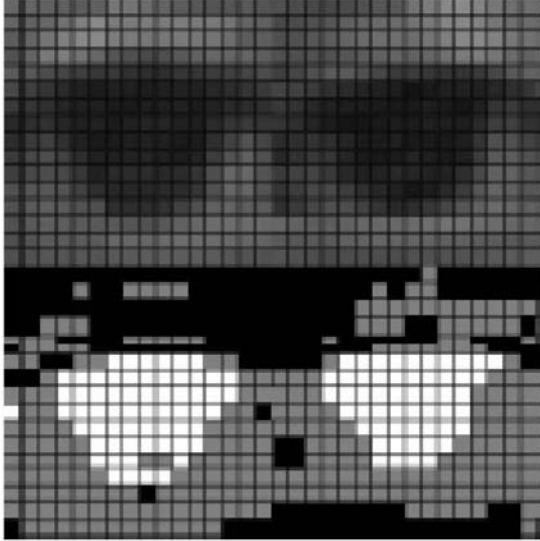
(a)

(b)

(c)

(d)

【図20】



目領域ラベリング例

(黒:肌領域($u=0$), 灰:白目($u=-1$), 白:黒目($u=1$))

【 図 2 2 】

視線推定結果



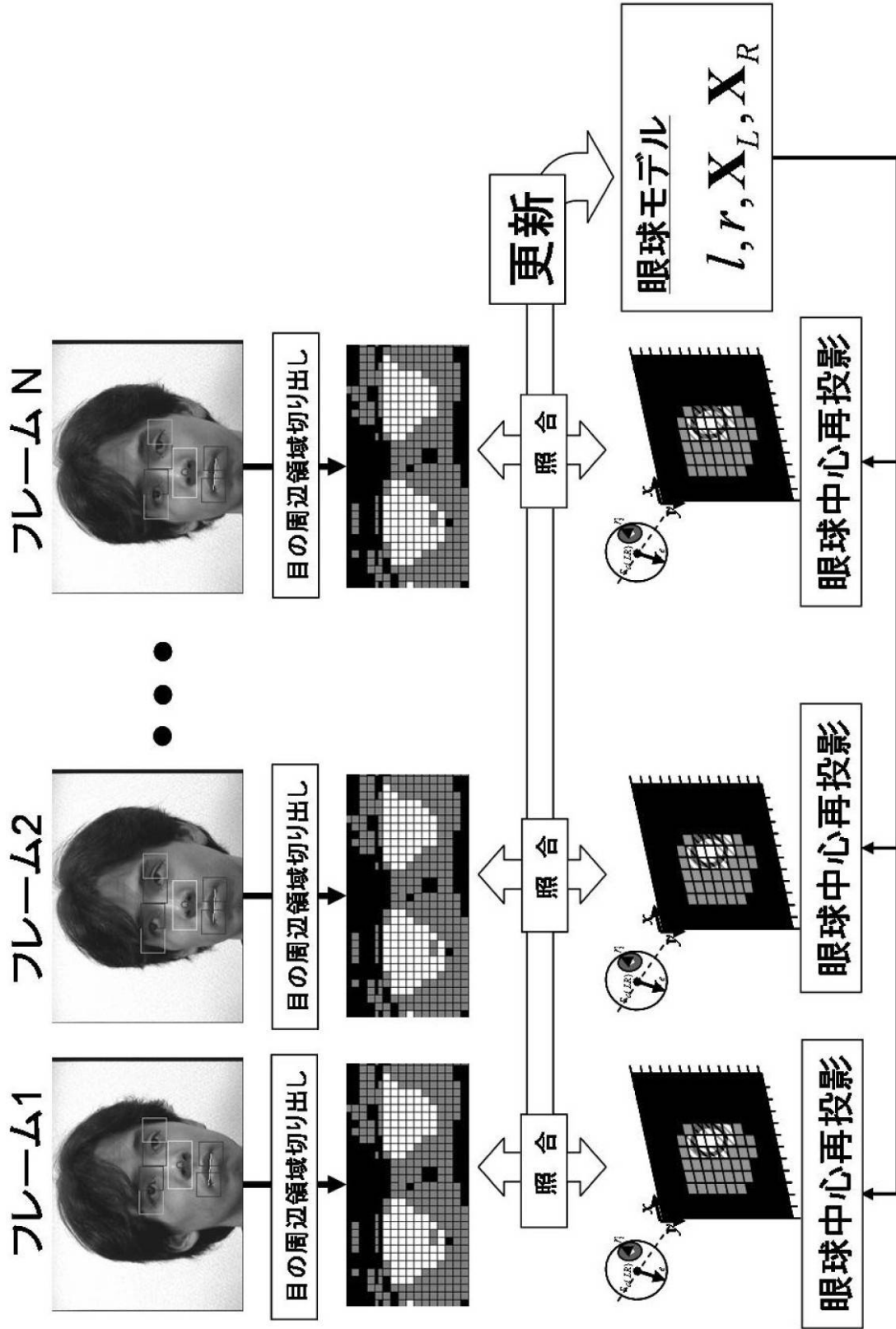
注視方向

右方

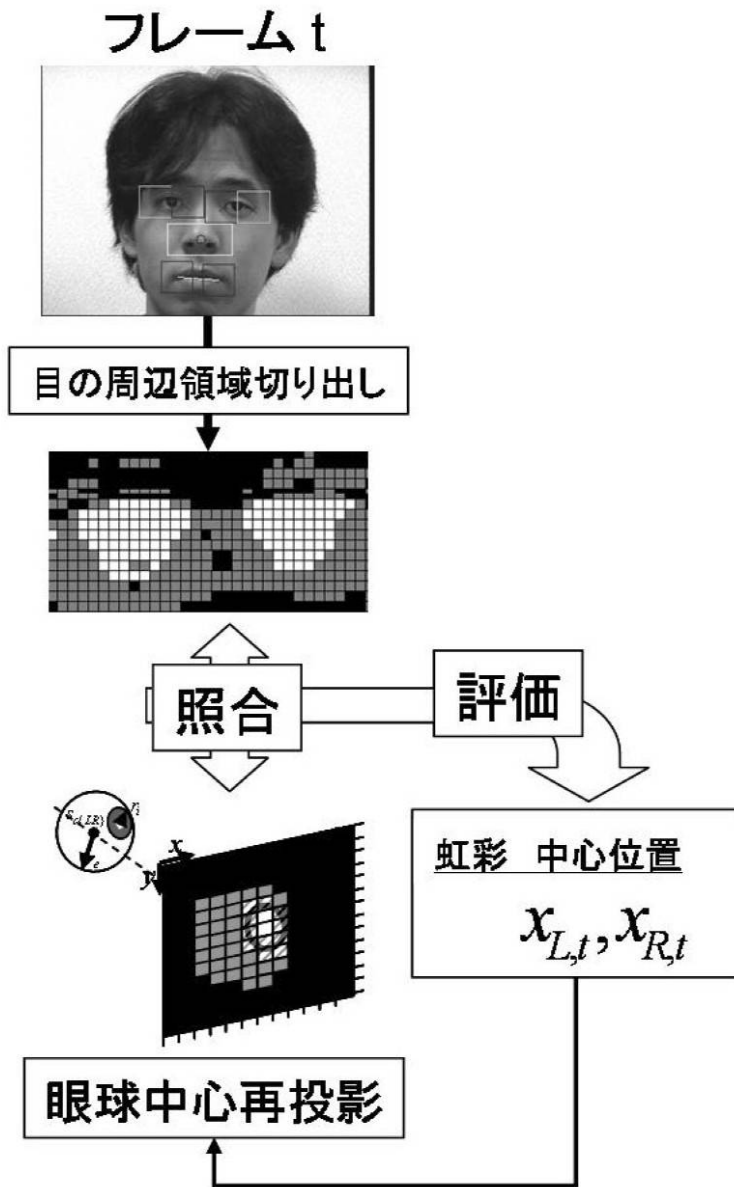
下方

上方

【 図 23 】

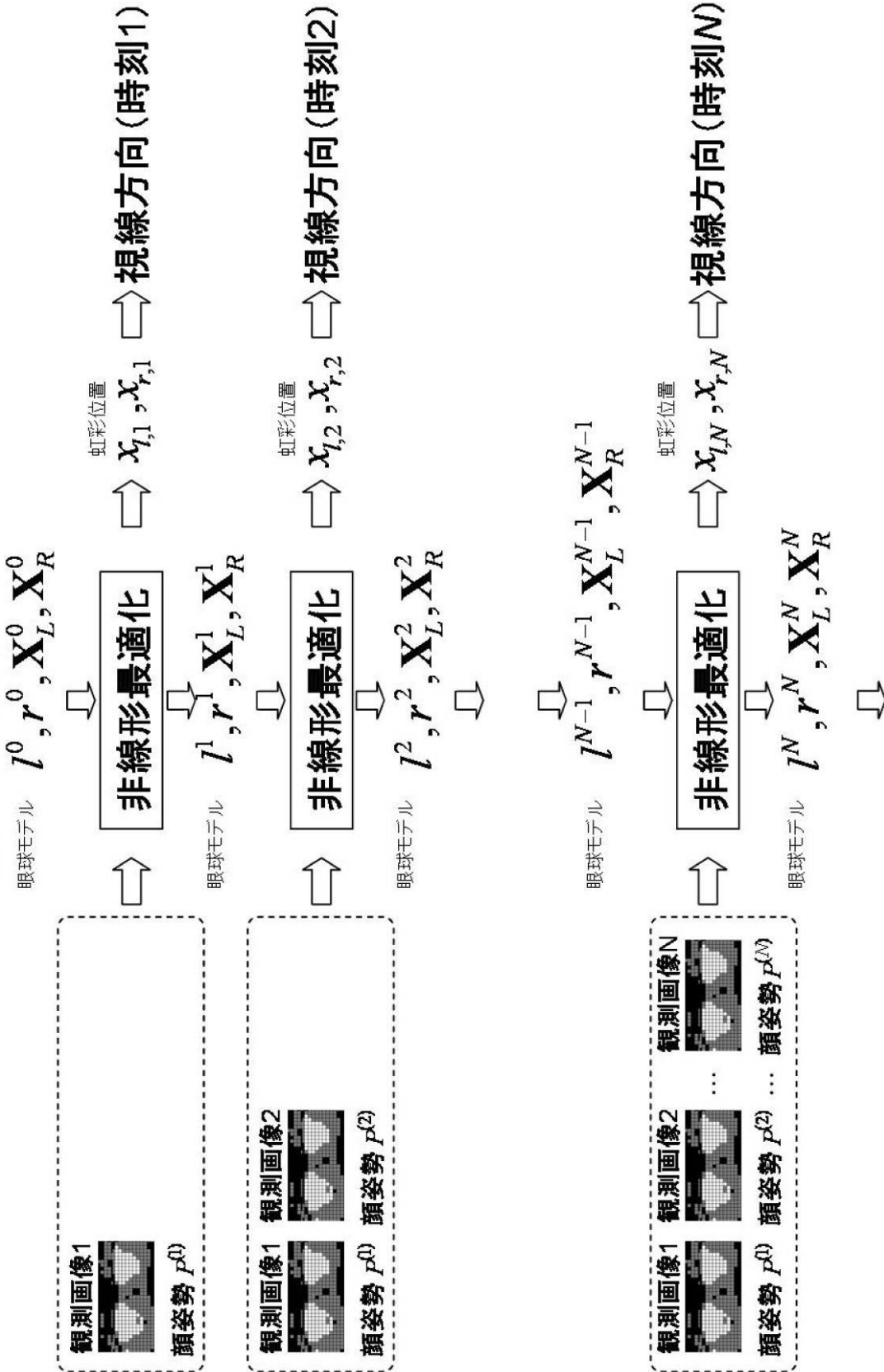


【図24】



【 図 2 5 】

(初期値(解剖学的知見等))



フロントページの続き

(出願人による申告)平成19年度独立行政法人情報通信研究機構、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

特許権者において、実施許諾の用意がある。

前置審査

(56)参考文献 特開2005-196519(JP,A)
特表2004-504684(JP,A)
国際公開第2005/006251(WO,A1)
特開2000-259814(JP,A)
特開2006-285531(JP,A)

(58)調査した分野(Int.Cl.,DB名)

G06T 7/60
A61B 3/113