

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4725949号
(P4725949)

(45) 発行日 平成23年7月13日(2011.7.13)

(24) 登録日 平成23年4月22日(2011.4.22)

(51) Int.Cl.

F I

G 0 6 T 1/00 (2006.01)

G 0 6 T 1/00 3 4 0 A

G 0 6 T 7/00 (2006.01)

G 0 6 T 7/00 2 0 0 Z

請求項の数 6 外国語出願 (全 17 頁)

(21) 出願番号 特願2004-335970 (P2004-335970)
 (22) 出願日 平成16年11月19日(2004.11.19)
 (65) 公開番号 特開2006-146571 (P2006-146571A)
 (43) 公開日 平成18年6月8日(2006.6.8)
 審査請求日 平成19年10月26日(2007.10.26)

(出願人による申告)平成16年度独立行政法人情報通信研究機構、研究テーマ「人間情報コミュニケーションの研究開発」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

(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 100099933
 弁理士 清水 敏
 (72) 発明者 サブリ・グループズ
 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 木下 敏介
 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 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

最終頁に続く

(54) 【発明の名称】 鼻梁及び鼻先端を特定しトラッキングする方法

(57) 【特許請求の範囲】

【請求項1】

顔画像の鼻梁の位置を特定するための、コンピュータで実現される方法であって、
 ビデオ画像生成装置から画像フレームを受信するステップと、
 画像フレーム中の画像の2つの眼の位置を特定するステップと、
 前記2つの眼の位置とその間隔とに基づいて、顔の、鼻と口とを含む顔画像の位置とサイズとを推定するステップと、

前記顔画像の光強度分布に基づいて前記顔画像中の鼻梁区域を決定するステップとを含み、

前記決定するステップは、前記顔画像中で前記2つの目の位置を結ぶ線の両側のうち、
 前記鼻と口とを含む側の領域を、前記2つの目の位置を結ぶ線と平行な複数の行セグメントに分割するステップと、

前記複数の行セグメントの各々について水平光強度プロファイルを計算するステップと、

前記顔画像の前記複数の行セグメントの各々について、前記水平光強度プロファイル中の鼻梁候補を見出すステップとを含み、

前記分割するステップにおいて分割される複数の行セグメントの各々は、単一の行画像ではなく、複数の行画像を含み、

前記見出すステップが、

前記顔画像の前記行セグメントの各々について、前記水平光強度プロファイルに所定の

10

20

最大フィルタを適用するステップを含み、前記最大フィルタは3個のセグメントを含んでおり、さらに、

前記顔画像の前記行セグメントの各々について、 S_1 S_2 S_3 を満たす鼻梁候補を見出すステップを含み、ここで S_1 、 S_2 、 S_3 は最大フィルタ中のそれぞれ最も左のセグメント、中央のセグメント、最も右のセグメントの光強度値をそれぞれ示し、前記最大フィルタはさらに、 A 、 B 、 C をそれぞれ正の定数として、重付け和 $A S_1 + B S_2 + C S_3$ が行セグメント中で最大となるようにする、方法。

【請求項2】

画像シーケンス中の顔画像の鼻梁をトラッキングするための、コンピュータで実現される方法であって、前記顔画像の各々について、請求項1に記載の方法を利用して顔画像の鼻梁の位置を特定するステップを含む、方法。

10

【請求項3】

顔画像から鼻の先端を見出すための、コンピュータで実現される方法であって、

請求項1に記載の全てのステップを実行するステップと、

前記決定するステップで決定された鼻梁区域中において、当該鼻梁区域中で顔画像の光強度値が最も高くなる最大強度点を特定するステップと、

前記最大強度点から顔画像の下端まで延びる顔画像中の列セグメントを規定するステップと、

前記列セグメントの列方向に沿って、前記列セグメントの光強度プロファイルを作成するステップと、

20

前記光強度プロファイルの第1次導関数と第2次導関数とを調査することにより、鼻先端のレベルを特定するステップと、

前記鼻梁候補の回帰直線を求めることにより鼻梁線を形成するステップと、

前記鼻先端のレベルによって規定される線と前記鼻梁線との交点として、鼻先端を見出すステップとを含み、

前記鼻先端のレベルを特定するステップは、

前記列セグメントの前記光強度プロファイルを平滑化するステップと、

前記平滑化するステップにおいて平滑化された前記列セグメントの前記光強度プロファイルの第1次導関数を求め、当該第1次導関数が最大値を有する点を特定することによって、前記列セグメントの前記光強度プロファイル中で鼻梁区域の先端に対応する位置を特定するステップと、

30

前記列セグメントの前記光強度プロファイルの第2次導関数を求め、当該第2次導関数がゼロとなる点を、列セグメントの光強度プロファイル中における鼻孔線の位置として特定するステップと、

前記鼻梁区域の先端に対応する位置と前記鼻孔線の位置との間の予め定められた点として、鼻先端の位置を決定するステップとを含む、顔画像から鼻の先端を見出すための方法。

【請求項4】

前記鼻先端の位置を決定するステップが、前記列セグメントにおいて、前記鼻梁区域の先端の位置と前記鼻孔線の位置との平均点を計算するステップを含む、請求項3に記載の方法。

40

【請求項5】

画像シーケンス中の顔画像の鼻先端をトラッキングするための、コンピュータで実現される方法であって、前記顔画像の各々について、請求項3～請求項4のいずれかに記載の方法を利用して顔画像の鼻先端の位置を特定するステップを含む、画像シーケンス中の顔画像の鼻先端をトラッキングするための方法。

【請求項6】

コンピュータ上で実行されると、コンピュータに、請求項1～請求項5のいずれかに記載の全てのステップを行なわせる、コンピュータプログラム。

【発明の詳細な説明】

50

【技術分野】

【0001】

この発明はリアルタイムの人の鼻梁及び鼻先端のトラッキング技術に関し、視覚による人物認識、2D/3D仮想現実、先進的な人-コンピュータインタフェース、及びロボティクス応用等の様々な応用において他の顔の特徴と共に用いることができる。

【背景技術】

【0002】

伝統的に、コンピュータ及びロボットは人間の命令により割当てられた仕事を行なうことのできるツールであると考えられてきた。最近、本発明者らは、機械が人間と同様なやり方でこの世界で動き、特にそれらが他の人間と相互に有益な関係で仲間として相互作用するような、共生システムを考えるようになった。

10

【0003】

このようなマン-マシンの交流では、複数の方法による人と機械との相互作用が不可欠である。特に、顔の表情のトラッキングは必須である。

【0004】

先行技術の顔特徴トラッキングアルゴリズムは、テンプレート又はエッジに基づく特徴の集まりに基づくものであり、初期設定と特徴のトラッキングとが困難であった。特徴の学習が必要とされること、トラッキングの過程で一致の可能性のある候補が多く生じることがその理由である[非特許文献1、2を参照]。さらにこのようなアプローチはより多くの計算時間を必要とする。実用的にするためには、顔特徴トラッキング方法はリアルタイムで動作し、どこでも利用でき、正確で頑健である必要がある。特に、システムはトレーニングの必要性なしに、任意のユーザに対し自然な条件のもとで動作し、眼鏡やひげがある場合など様々な顔の特徴に対処できなければならない。

20

【0005】

最も支配的な顔の特徴は眼、鼻(鼻梁、鼻先端、鼻孔)、口であると言われている。顔、眼及び口のトラッキングに関しては多くの研究がある[非特許文献3、4、5、6、7を参照]が、鼻先端及び鼻孔に関する自動化されたトラッキングには見るべきものがほとんどない。本発明者らの知る限りでは、鼻梁のトラッキングについての研究は皆無である。

【0006】

パタジャンは彼の著作において、口のトラッキングの参照点として鼻孔を用いている[非特許文献2を参照]。鼻孔の検出は、それらが見えている場合は非常に安定している。しかし、顔の向きによっては、それらはさえぎられてしまう。従って、鼻孔を参照点として用いるのは一般に実用的なアプローチとは言えない。

30

【0007】

カワトは特許文献1において、顔がどのような向きを向いていても鼻先端を検出する方法を記載している。ここでは、二つの眼の位置から推定された検索領域内で光強度の最も高い領域を検索している。

【0008】

照明の設定によっては、この方法は鼻先端領域でなく、頬領域の最も高い光強度の領域を検出してしまうと報告されている。また、画像フレーム中で見える形状によっては、鼻梁領域の一部が鼻先端より高い光強度を有する場合があります、このため検出精度が乱れることがある。

40

【0009】

ゴロドニキ及びロスの鼻先端トラッキング方法は、テンプレート検索に基づくものであり、テンプレートベクトルが、鼻表面の極値付近に選ばれる[非特許文献8を参照]。その後、学習セッション時に記録された画素パターンを、後続の画像フレームでアルゴリズムによりトラッキングする。

【0010】

この方法には、以下の三つの点で限界がある。

50

【 0 0 1 1 】

1) テンプレート記録セッションが必要であること。

【 0 0 1 2 】

2) 照明条件に敏感であること。このため、ユーザの頭の動きを或る範囲内に制限している。

【 0 0 1 3 】

3) 先行するトラッキングの結果に依存して後続のフレームのトラッキングを行なうこと。

【特許文献1】特開2004-157778号

【特許文献2】米国特許第4,975,960号

【非特許文献1】ドミトリ・O・ゴロドニキ、ゲルハルト・ロス、「ノウス『あなたの鼻を口として使う』、ハンズフリーのゲーム及びインタフェース用知覚的視野技術」、画像及び視覚コンピューティング、第22巻、2004年。(Dmitry O. Gorodnichy, and Gerhard Roth: Nouse 'use your nose as a mouse' perceptual vision technology for hands-free games and interfaces, Image and Vision Computing, vol. 22, 2004.)

【非特許文献2】マルグリット・ベッケ、ジェームズ・ジップス、及びピーター・フレミング、「カメラマウス：重度障害者にコンピュータアクセスを提供するための、身体的特徴の視覚的トラッキング」、IEEE 神経及びリハビリテーション工學論文集、第10巻、第1号、2002年3月。(Margrit Betke, James Gips, and Peter Fleming: The Camera Mouse: Visual Tracking of Body Features to Provide Computer Access for People with Severe Disabilities, IEEE Transactions on Neural and Rehabilitation Engineering, Vol. 10, No. 1, March 2002.)

【非特許文献3】シンジロウ・カワト及びノブジ・テツタニ、「SSRフィルタ及びサポートベクトルマシンによる、リアルタイムスケール適応顔検出及びトラッキング」、ACC V 予稿集、第1巻、2004年。(Shinjiro Kawato, and Nobuji Tetsutani: Scale Adaptive Face Detection and Tracking in Real Time with SSR Filter and Support Vector Machine, Proc. of ACCV, vol. 1, 2004.)

【非特許文献4】チェン・チン・チャン、ウェン・カイ・タイ、マウ・ツェン・ヤン、イー・ティン・ハン及びチ・ジャン・ハン、「リアルタイムで唇、眼及び顔を検出するための新方法」、リアルタイム画像処理、第9巻、2003年。(Cheng-Chin Chiang, Wen-Kai Tai, Mau-Tsuen Yang, Yi-Ting Huang, and Chi-Juang Huang: A Novel Method for Detecting Lips, Eyes and Faces in Real-Time, Real-Time Imaging Vol. 9, 2003.)

【非特許文献5】イェン・リ・ティアン、タケオ・カナデ及びジェフリー・F・コーン、「形状、色及び動きを組合せた頑健な唇検出」、ACC V '2000 予稿集、台北、台湾、2000年1月。(Ying-li Tian, Takeo Kanade, and Jeffrey F. Cohn: Robust Lip Tracking by Combining Shape, Color and Motion, in Proc. of ACCV'2000, Taipei, Taiwan, January 2000.)

【非特許文献6】サブリ・グルブズ、ケイスケ・キノシタ及びスミオ・ヤノ、「トレーニング可能多変数ガウス分類器を用いたビデオシーケンスからの口トラッキング」、PRMU 2003、仙台、日本、2003年12月。(Sabri Gurbuz, Keisuke Kinoshita, and Sumio Yano: Mouth Tracking from Video Sequences using Trainable Multivariate Gaussian Classifiers, PRMU 2003, Sendai, Japan, December 2003.)

【非特許文献7】ヌリア・オリバー、アレックス・ペントランド及びフランソワ・ベラード、「ラフター：顔表情認識を伴うリアルタイムの顔及び唇トラッカ」、パターン認識、33:1369-1382、2000年。(Nuria Oliver, Alex Pentland, and Francois Berard: Lafter: A Real-time Face and Lips Tracker with Facial Expression Recognition, Pattern Recognition, 33:1369-1382, 2000)

10

20

30

40

50

【非特許文献 8】ドミトリ・O・ゴロドニキ、「顔トラッキングにおける鼻の重要性について」、IEEE 自動顔及びジェスチャ認識に関する国際会議予稿集 (FG'02)、ワシントン DC、2002 年 5 月 20 - 21 日。(Dmitry O. Gorodnichy: On Importance of Nose for Face Tracking, In Proc. IEEE International Conference on Automatic Face and Gesture Recognition (FG'02), Washington DC, May 20-21, 2002.)

【非特許文献 9】ポール・バイオラ及びマイケル・ジョーンズ、「頑健なリアルタイム物体検出」、視覚モデリング、学習、コンピューティング及びサンプリングに関する統計的及びコンピュータ関連理論に関する第二回国際ワークショップ、カナダ、バンクーバー、2001 年 7 月 13 日。(Paul Viola, and Micheal Jones: Robust Real-time Object Detection, Second International Workshop on Statistical and Computational Theories of Vision- Modeling, Learning, Computing, and Sampling, Vancouver, Canada, July 13, 2001.)

【発明の開示】

【発明が解決しようとする課題】

【0014】

この発明の目的は、リアルタイムのトラッキングに計算上好適な、人の鼻梁トラッキング装置を提供することである。

【0015】

この発明の別の目的は、陰及び照明の変化に対して頑健な鼻梁トラッキング装置を提供することである。

【0016】

この発明のさらなる目的は、正面から見た顔の向き及びスケールの変化に対し不変である、鼻梁トラッキング装置を提供することである。

【0017】

この発明の別の目的は、肌の色及び顔の寸法に対し不変である、鼻梁トラッキング装置を提供することである。

【0018】

従って、この発明の別の目的は、人物に対し何ら仮定を用いることなく、またその人物に関する何らかの種類のトレーニングデータを必要とすることのない、鼻梁の測光的な属性に基づいた、鼻梁トラッキング装置を提供することである。

【0019】

この発明の他の目的、特徴及び利点は、以下の要約、図面及びこの発明並びに好ましい実施の形態の詳細な説明から、パターン分析及び認識分野の当業者には明らかになるであろう。

【課題を解決するための手段】

【0020】

この発明の方法は、鼻先端位置の検出率及び検出精度を改良し、さらに鼻梁線も検出する。これは眼に近い側の鼻梁端部から始めて鼻先端を見出すまで下方に検索を行なう。こうして、鼻梁と鼻先端位置とをリアルタイムで出力することができる。

【0021】

幾何学的形状及び照度の変化がある場合に人の鼻梁をリアルタイムでトラッキングするための方法は、カラー又はモノクロのビデオ装置からの画像フレームを入力画像として得るステップと、現在入力されている画像を処理して眼の位置を特定するステップと、両眼間の距離に基づいて画像中の顔の位置とサイズとを推定するステップと、鼻及びおそらくは口領域を含む、一般的な顔の対象領域 (region of interest: ROI) を切出すステップと、鼻梁の測光的属性に基づいて鼻梁線と鼻先端とをリアルタイムで出力するステップとを含む。

【0022】

より特定的には、この発明の一局面によれば、顔画像の鼻梁の位置を特定するための、

10

20

30

40

50

コンピュータで実現される方法は、ビデオ画像生成装置から画像フレームを受信するステップと、画像フレーム中の画像の眼の位置を特定するステップと、眼の位置とその間隔とに基づいて、ある顔の鼻と口とを含む顔画像の位置とサイズとを推定するステップと、顔画像の光強度分布に基づいて顔画像中の鼻梁区域を決定するステップとを含む。

【0023】

好ましくは、決定するステップは、顔画像の行セグメントの各々について水平光強度プロファイルを計算するステップと、顔画像の行セグメントの各々について水平光強度プロファイル中で鼻梁候補を見出すステップとを含み、水平光強度プロファイルは、鼻梁候補の位置において所定の基準を満たしている。

【0024】

さらに好ましくは、見出すステップは、顔画像の行セグメントの各々について水平光強度プロファイルに所定の最大フィルタを適用するステップを含み、最大フィルタは3個のセグメントを含んでおり、さらに、顔画像の行セグメントの各々について、 S_1 S_2 S_3 を満たす鼻梁候補を見出すステップを含み、ここで S_1 、 S_2 、 S_3 は最大フィルタ中のそれぞれ最も左のセグメント、中央のセグメント、最も右のセグメントの光強度値をそれぞれ示し、最大フィルタはさらに、 A 、 B 、 C をそれぞれ正の定数として、重付け和、 $AS_1 + BS_2 + CS_3$ が行セグメント中で最大となるようにする。

【0025】

定数 B は定数 A 及び C より大きくてもよい。

【0026】

定数 A は定数 C と等しくてもよい。

【0027】

この発明の別の局面に従えば、画像シーケンス中の顔画像の鼻梁をトラッキングするための、コンピュータで実現される方法は、顔画像の各々について、上述のいずれかの方法を利用して顔画像の鼻梁の位置を特定するステップを含む。

【0028】

この発明のさらに別の局面に従えば、顔画像から鼻の先端を見出すための、コンピュータで実現される方法は、上述のいずれかの方法の全てのステップを実行するステップと、決定するステップで決定された鼻梁区域中において、当該鼻梁区域中で顔画像の光強度値が最も高くなる最大強度点を特定するステップと、最大強度点から顔画像の下端まで延びる顔画像中の列セグメントを規定するステップと、列セグメントの列方向に沿って、列セグメントの光強度プロファイルを作成するステップと、光強度プロファイルの第1次導関数と第2次導関数とを調査することにより、鼻先端のレベルを特定するステップと、鼻梁候補の回帰直線を求めることにより鼻梁線を形成するステップと、鼻先端のレベルによって規定される線と鼻梁線との交点として、鼻先端を見出すステップとを含む。

【0029】

好ましくは、鼻先端レベルを特定するステップは、列セグメントの光強度プロファイルの第1次導関数が最大値を有する点を特定することによって、列セグメントの光強度プロファイル中で鼻梁区域の先端に対応するレベルを特定するステップと、列セグメントの光強度プロファイルの第2次導関数が予め定められた値を有する点を特定することにより、列セグメントの光強度プロファイル中における鼻孔レベルを特定するステップと、鼻梁区域の先端レベルと鼻孔レベルとの間の予め定められた点として、鼻先端のレベルを決定するステップと、を含む。

【0030】

鼻先端のレベルを決定するステップは、列セグメントにおいて、鼻梁区域の先端と鼻孔レベルとの平均点を計算するステップを含む。

【0031】

予め定められた値はゼロであってもよい。

【0032】

この発明の別の局面に従えば、画像シーケンス中の顔画像の鼻先端をトラッキングする

10

20

30

40

50

ための、コンピュータで実現される方法は、顔画像の各々について、上述のいずれかの方法を利用して顔画像の鼻先端の位置を特定するステップを含む。

【 0 0 3 3 】

この発明のコンピュータプログラムは、コンピュータ上で実行されると、コンピュータに、上述の方法のいずれかの全てのステップを行なわせる。

【発明を実施するための最良の形態】

【 0 0 3 4 】

[実施の形態の概観]

以下に説明するこの発明の実施の形態は、人間の鼻の曲線が持つ測光的属性の利点を眼トラッキング技術と組合せて、鼻トラッキングに関して従来から認識されていた問題点を克服するものである。初めに、眼の位置を見出して顔のROIを推定し、その後この方法は鼻梁と鼻先端との両方に沿った鼻の曲線の光強度プロファイルを利用する。提案されたアプローチでは、この鼻梁トラッキング装置は照明条件の変化、肌の色合い、及び正面から見た顔の幾何学的形状並びに顔の動きに対し頑健となる。

【 0 0 3 5 】

図1は鼻梁及び鼻先端をトラッキングするための、この実施の形態に従ったトラッキング装置30の概観ブロック図である。図1を参照して、トラッキング装置30は、予め定められた速度で人の顔の画像フレームシーケンスを撮影するためのカラー（又はモノクロ）ビデオカメラ40と、ビデオカメラによって撮影された各フレームを記憶するためのフレームメモリ42と、フレームメモリ42に記憶された各フレーム内で眼の位置を検出するための眼検出モジュール44と、カメラ40によって撮影された人の顔の鼻梁を検出するとともに鼻梁線及び鼻先端位置データを出力するための、鼻梁検出モジュール46とを含む。

【 0 0 3 6 】

各フレームに対し、この実施の形態のトラッキング装置30はビデオカメラ40からの画像フレームを得る。眼検出モジュール44は光強度平面を用いて各フレームの眼の位置を検出する。すなわち、鼻梁検出モジュール46は人の鼻の曲線が有する相対的光強度情報（光強度分布）を、「眼の検出及びトラッキング」と題したセクションで簡単に説明する眼のトラッキング技術と組合せて利用する。この実施の形態では、撮影された画像は赤、緑、青（red-green-blue: RGB）の画像である。従って、鼻梁検出モジュール46は色平面の一つを光強度情報の尺度として用いることによって計算時間を節約する。この実施の形態では、光強度情報を代表するものとして、緑平面を用いる。

【 0 0 3 7 】

こうして、顔の支配的な特徴である眼がまず眼検出モジュール44によって基準点としてトラッキングされる。

【 0 0 3 8 】

その後、鼻梁と鼻先端とが、各画像フレームにおいてリアルタイムで、過去のトラッキング結果に依存することなく、鼻梁検出モジュール46によって検出される。眼検出モジュール44及び鼻梁検出モジュール46で用いられるアルゴリズムは後述する。

【 0 0 3 9 】

- 眼の検出及びトラッキング -

この実施の形態のトラッキング装置30はコンピュータハードウェアとコンピュータ上で実行される特定のコンピュータプログラムとで実現される。図2はコンピュータプログラムの全体の制御構造を示す。図2の全体構造を説明するに先だって、図2の各ステップの詳細を図3から図12を参照して説明する。

【 0 0 4 0 】

図3を参照して、撮影された画像120中の顔の眼の位置122及び124が最初に検出される。このステップで用いられる眼の検出及びトラッキングアルゴリズムは非特許文献3に詳述されている。

【 0 0 4 1 】

次に、顔の対象領域 (R O I) 1 2 6 が画像 1 2 0 から抽出される。簡単に言えば、両眼の間のパターンを検出し更新したテンプレートマッチングでトラッキングする。顔の大きさの変化に対処するため、検出のために様々な縮尺の画像を考慮し、両眼間の距離に従って、適切な縮尺を選択する。

【 0 0 4 2 】

アルゴリズムは、非特許文献 9 に記載の「統合画像」 (I n t e g r a l i m a g e) と呼ばれる入力画像の中間表現を計算する。その後、6 セグメント直交 (s i x - s e g m e n t e d r e c t a n g u l a r : S S R) フィルタを用いて、画像中の眼の領域の明暗関係をフィルタリングする。結果として得られる顔候補をさらに、サポートベクトルマシン (s u p p o r t v e c t o r m a c h i n e : S V M) によって検証する。

10

【 0 0 4 3 】

- 鼻梁及び鼻先端の検出 -

「眼の検出及びトラッキング」と題されたセクションで説明されたように、顔の R O I 1 4 0 が選択される。図 3 を参照して、撮影された画像 1 2 0 中の眼の位置 1 2 2 及び 1 2 4 を利用した顔の R O I 選択処理は、およそ $1.2 \times$ 両眼間の距離の幅と、 $1.8 \times$ 両眼間の距離の高さの寸法を有する、回転補正された矩形領域 1 2 6 を選択することを含む。結果として得られる R O I 候補 1 2 6 は切出され、回転されて顔 R O I 1 4 0 の画像が得られる。顔 R O I 1 4 0 は眼の位置 1 4 2 及び 1 4 4 によって規定される眼のレベルから始まり、典型的には首上部までである。

20

【 0 0 4 4 】

人の鼻は、凸形状を有し、眼のレベルから鼻先端までの鼻梁は一線上に延びる。一般に、或る点での光の強度の測定値は、その点と光源との距離が増加するにつれて小さくなる。これは、光がその光源の位置から離れるにつれて広がるからである。従って、曲率に関する情報は光強度分布で表現される。図 4 は右側に、顔 R O I 1 4 0 から抽出された行セグメント 1 6 0 の光強度プロファイル 1 7 0 を示す。

【 0 0 4 5 】

この実施の形態では、この物理的現象をノイズ減少及びパターン検索と関連させて利用する。

【 0 0 4 6 】

鼻は凸形状を有する。従って、鼻梁には側面に比べより多くの光強度が集まっている。光強度プロファイルのために顔の R O I 画像の単一の行を用いる代わりに、複数の行を用いて垂直の矩形 1 6 2 によって示されるピクセルの垂直線の光強度値の和をとることにより、ビデオノイズの影響を大幅に削減できる。従って、ノイズの影響が相殺されるのと同時に、光強度値が鼻の側部よりも鼻梁においてより速い速度で累積される。眼の線より下の、互いに重複する部分を持つセグメントに対しこの処理を繰返すことにより、光強度プロファイル 1 7 0 の 3 D パターン が 生成される。

30

【 0 0 4 7 】

すなわち、凸曲面上の点は側部より高い光強度を有することとなる。ここで、この実施の形態では行セグメント 1 6 0 の光強度分布を用いて鼻梁のトレースを行なうために高速フィルタリングのアプローチを利用している。

40

【 0 0 4 8 】

非特許文献 3 に記載された両眼間の検出フィルタと同様に、この実施の形態では最大フィルタと呼ばれる、3 セグメントの畳込みフィルタを用い、光強度の積分値を用いて鼻梁候補の位置のトレーシングを行なう。ここで中央のセグメントは側部のセグメントと等しいかまたは大きく、3 個のセグメント全ての定数を乗じた光強度の合計は、鼻梁の位置で最大値となる。図 4 及び図 5 は眼の線から始めて各行の最大光強度パターンをトレースする、3 個のセグメント 1 8 2、1 8 4 及び 1 8 6 を備えた最大フィルタ 1 8 0 を示す。基準は以下の通りである。

50

【 0 0 4 9 】

【 数 1 】

$$S_1 \leq S_2 \geq S_3 \quad \text{かつ}$$

$$\arg \max (AS_1 + BS_2 + CS_3) \quad (1)$$

ここで S_i は図 5 に示す最大フィルタのセグメント 1 8 2、1 8 4 及び 1 8 6 の光強度の積分値を示し、A、B 及び C は全て予め定められた定数である。好ましくは、定数 B は定数 A 及び C より大きい。定数 A は定数 C と等しくてもよい。例えば、 $B = 3$ であり $A = C = 1$ である。

10

【 0 0 5 0 】

フィルタは顔の ROI 画像の全ての行セグメントと畳み込まれる。各行で、式 1 の基準を満たす画素の位置が、さらに処理すべき鼻梁候補であると考えられる。

【 0 0 5 1 】

最大フィルタと画像の行セグメントとの直接畳み込みは、計算上のコストが大きい処理である。このため、図 6 を参照して、行セグメントの光強度値の累積和 2 0 0 を用いて、 S_i を S_i の 2 個の参照指標を用いて計算できる。各行セグメントの累積和は以下のように計算される。

$$CSUM(0) = rowArr(0),$$

$$CSUM(j) = CSUM(j-1) + rowArr(j); j = 1, 2, \dots, N \quad (2)$$

20

ここで $rowArr$ と $CSUM$ とはそれぞれ行アレイの光強度値とその累積和である。従って、指標 k_1 と k_2 との間の S_i は以下のように計算することができる。

$$S_i(k_1; k_2) = CSUM(k_2) - CSUM(k_1) \quad (3)$$

各行セグメントは、式 1 を用いて鼻梁位置を見出すために、顔 ROI のいずれかの端部に達するか、又は式 1 が満たされなくなるまで、処理される。すなわち、図 7 に示すように、鼻梁候補点 2 2 0 が見出されると、それらのうち幾つか 2 2 2 は鼻梁上にはなく、鼻先端よりも低い場合がある。このため、図 8 のボックス 2 4 0 で示されるように、鼻梁候補が鼻梁上に位置するよう、修正処理が必要となる。

【 0 0 5 2 】

$T(r)$ は、以下の式で規定される、最大フィルタ 1 8 0 から出力される顔 ROI 画像の r 番目の行セグメントの j 番目の位置での最大累積光強度合計を示すものとする。

30

【 0 0 5 3 】

【 数 2 】

$$T(r) = \max_i (AS_1 + BS_2 + CS_3) \quad (4)$$

ここで $T(r)$ は鼻梁上の 3 セグメントフィルタの区域下の光強度の累算に相当する。従って、拡散光の条件下で、或る点と光源との距離が増加すれば、又はその点が陰になれば、その点の $T(r)$ 値は小さくなる。このことから、鼻先端は鼻孔線上の点よりも高い $T(r)$ を有することが分る。なぜなら、鼻孔線上の点は、図 8 の右側の $T(r)$ プロファイル線 2 4 2 で示されるように、鼻曲線上にはないからである。このため、鼻先端と口ひげの線との間では $T(r)$ の勾配が高いことが期待され、さらに、鼻孔線上方の全ての候補点は鼻孔線上のものより高い $T(r)$ を有するものと期待される。

40

【 0 0 5 4 】

- 鼻梁線及び鼻先端を得るためのアルゴリズム -

鼻梁線及び鼻先端を得るためのアルゴリズムを以下で説明する。

【 0 0 5 5 】

1. $T(r)$ 線 2 4 2 の勾配 $T(r)$ を決定する。この実施の形態では、 $T(r)$ は図 9 に示すように決定される。図 9 を参照して、 $T(r)$ は、(例えば点 2 6 0 の) $T(r)$ と (例えば点 2 6 2 の) $T(r+1)$ との差を Δ で除したものに等しい。ここで Δ は予め定められた自然数であって、例えば、3 である。この定義によれば、 $T(r)$

50

は $T(r)$ の値が減少していれば正となる。絶対値は曲線の急峻さを示す。もちろん、他の定義を用いてもよい。

【0056】

2. $T(r)$ の指標を勾配値の降順にソートする。

【0057】

3. 行の指標が目のレベルから始まるものとして、 $T(r)$ の最も高い勾配の位置から始めて、 $T(r) > T(k)$ $r < k$ となる k 番目の行を見出す。これは、 $T(r)$ の値が急峻に減少する点が最も早く検討されることを意味する。

【0058】

4. 図10を参照して、 $T(i)$ が $i < k$ について最大となる行の指標 i を見出し、 i 番目の行から開始する列セグメントを選択する。図10に列セグメントを示す。図10を参照して、列セグメント270は i 番目の行から開始し典型的には顔のROIの下端まで伸びる、顔ROIの垂直セグメントである。

10

【0059】

5. 図11(a)を参照して、垂直方向の、列セグメント270の水平光強度プロファイル(hproj)を形成する。

【0060】

6. 図11(b)を参照して、hproj280上のビデオノイズの影響を除去して、平滑化された光強度プロファイル(shproj)282を得る。例えば、hprojに高速フーリエ変換(FFT)を行ない、 0.15π デジタル周波数より上の全てのFFT係数をゼロに設定し、結果として得られる係数の配列の逆高速フーリエ変換(IFFT)をとる。平滑化されたプロファイル上で、最初の3個の極値(2個の極大284及び288と1個の極小286)を探す。

20

【0061】

7. 図11(c)を参照して、鼻先端の第1次導関数294の絶対値($|shproj|$)は、第一の極大284と第一の極小286との間で最大値296を有する。

【0062】

8. さらに図11(C)を参照して、鼻孔線における第1次導関数はゼロか定数かいずれかである。すなわち、鼻先端の後、第2次導関数がゼロ($d^2shproj(p) = 0$)となる最初の点298が鼻孔線である。

30

【0063】

9. さらに、鼻梁から鼻先端まで、図12の線312で示すように、式1を用いて抽出された全ての点の回帰直線を求めて鼻梁を表すものとする。図12では、鼻梁点310Aから310Nまでの回帰直線を求めている。線312は鼻梁を表すものと仮定される。

【0064】

10. 図12を参照して、鼻先端は鼻梁線312上の鼻梁の先端から、先のステップで見出された鼻孔線との間のどこかの点である。従って、鼻先端は、鼻梁の先端と鼻孔線との間のある線上にあると仮定される。後述するようにこの実施の形態では、鼻先端は鼻梁の先端と鼻孔線との間の線上にある。図12の線314がこの線を表すものとする。鼻先端は、鼻梁線312と線314との交点316である。

40

【0065】

- コンピュータによる実現 -

この実施の形態の鼻梁及び鼻先端トラッキング装置30は、肌の色の違い及び照明条件によってパラメータを微調整する必要がない。どのユーザがいつ、カメラの前に現れても、トラッキング装置30は自動的にその人の眼、鼻梁及び鼻先端のトラッキングを、学習なしで開始する。

【0066】

図2に戻って、この実施の形態の装置30を実現するプログラムの全体の制御構造を説明する。プログラムはステップ60で始まる。各フレームについて、装置30はステップ60で画像フレームを撮影する。ステップ62で、ステップ60で撮影された画像内で眼

50

が検出される。ステップ 64 で、顔の ROI が抽出される。

【0067】

次に、抽出された顔 ROI の各行セグメントについて、ステップ 66 から 72 のステップが繰返される。具体的には、各行セグメントについて、行セグメントの全ての画素列の累積和をステップ 68 で計算し、ステップ 70 で対象の行セグメントについて計算された累積和に最大フィルタ（図 5 を参照）を適用して、行セグメントの水平方向の $T(r)$ プロファイル（図 4 を参照）を得る。このようにして得た $T(r)$ プロファイルの例を図 4 の右側に示す。

【0068】

行セグメントの各々について、ステップ 68 及び 70 で $T(r)$ プロファイルが得られた後、ステップ 74 で口ひげの線を推定する。口ひげの線をどのように推定するかを図 8 及び図 9 に示す。ステップ 76 で、口ひげの線の上方で最大の $T(r)$ の行が選択される。ステップ 78 で、ステップ 76 で選択された行の下で顔 ROI の下端まで伸びる列セグメントを抽出する。列セグメントの例を図 10 に列セグメント 270 として示す。

【0069】

ステップ 80 で、ステップ 78 で抽出された列セグメントの水平射影が形成される。この射影の一例を図 11 (a) に示す。ステップ 82 で、この射影に対し FFT を適用し、所定のデジタル周波数より上の FFT 係数をゼロに設定し、結果として得られる FFT 係数のアレイの IFFT をとることによって、この射影は平滑化される。平滑化された射影の例を図 11 (b) に示す。

【0070】

ステップ 84 で、プロファイルの最初の 3 個の極値（2 個の極大及びその間の極小）が図 11 (b) に示すように点 284、286 及び 288 として見出される。ステップ 86 で、点 284 と 286 との間のプロファイルの第 1 次導関数がとられ、第 1 次導関数の最大点が、図 11 (c) の点 296 である鼻梁の先端を示す。

【0071】

ステップ 88 で、プロファイルの第 2 次導関数がとられ、点 296 の後で第 2 次導関数がゼロとなる点が鼻孔線を示す。鼻先端は、鼻梁の先端と鼻孔線との間のどこかにある。この実施の形態では、鼻先端レベルは鼻梁の先端と鼻孔線位置との平均（中心点）として、ステップ 90 で計算される。

【0072】

ステップ 92 では、鼻梁上の全ての点の回帰直線（鼻梁線）を求める。ステップ 94 で、鼻先端線と鼻線との交点を探す。この点が鼻先端である。

【0073】

図 13 は上述のプログラムを実行してこの実施の形態の装置 30 を実現するコンピュータシステム 330 の概観を示す。上述の実施の形態はコンピュータハードウェア及びその上で実行されるコンピュータプログラムで実現される。図 13 はこの実施の形態のコンピュータシステム 330 の外観図であり、図 14 はシステム 330 をブロック図で示す。

【0074】

図 13 を参照して、コンピュータシステム 330 は、FD (Flexible Disk) ドライブ 352 及び CD-ROM (Compact Disc Read Only Memory) ドライブ 350 を含むコンピュータ 340 と、キーボード 346 と、マウス 348 と、モニタ 342 と、ビデオカメラ 40 とを含む。

【0075】

図 14 を参照して、コンピュータ 340 は、FD ドライブ 352 及び CD-ROM ドライブ 350 に加えて、CPU (Central Processing Unit) 356 と、CPU 356、CD-ROM ドライブ 350 及び FD ドライブ 352 に接続されたバス 366 と、ブートアッププログラム等のプログラムを記憶するための Read-Only Memory (ROM) 358 と、CPU 356 に接続され、アプリケーションプログラムの命令を一時的に記憶するとともに一時記憶空間を提供するための Random

10

20

30

40

50

Access Memory (RAM) 360と、アプリケーションプログラム、システムプログラム、及びデータを記憶するためのハードディスク354とを含む。ここでは図示しないが、コンピュータ340はさらに、ローカルエリアネットワーク(LAN)への接続を提供するネットワークアダプタボードを含んでもよい。

【0076】

コンピュータシステム330にこの実施の形態のトラッキング装置30の機能を実行させるプログラムは、CD-ROM362又はFD364に記憶されて、CD-ROMドライブ350又はFDドライブ352に挿入され、さらにハードディスク354に転送されてもよい。これに代えて、プログラムは、図示しないネットワークを介してコンピュータ340に送信され、ハードディスク354に記憶されてもよい。プログラムは実行の際にRAM360にロードされる。プログラムはCD-ROM362、FD364又はネットワークから直接ロードされてもよい。

10

【0077】

プログラムは、コンピュータ340にこの実施の形態のトラッキング装置30の機能を実行させるための幾つかの命令を含む。必要とされる基本的機能の幾つかは、コンピュータ340上で実行されるオペレーティングシステム(OS)又はサードパーティプログラム、もしくはコンピュータ340にインストールされたモジュールによって提供されるので、プログラムはこの実施の形態のトラッキング装置30を実現するための基本的機能全てを必ずしも含まなくてもよい。プログラムは、制御された状態で適切な機能と呼出し、所望の結果が得られるようにする命令の部分のみを含んでいればよい。コンピュータシステム330がどのように動作するかは周知であり、従ってここでは繰返さない。

20

【0078】

- トラッキングの結果 -

図15及び16は自動的にトラッキングされた鼻梁及び先端の幾つかの例を示す。図15に示すように、鼻先端380、382及び384は異なる向きと異なる照明設定との下で正確にトラッキングされている。さらに、図16で示すように、鼻先端390はサンガラスをかけている場合でもトラッキングできている。

【0079】

提案されたアルゴリズムはC++で実現され、通常のOSの下で実行される。実験結果により、コンピュータハードウェア及びソフトウェアで実現された装置30はリアルタイムで、完全に自動的に、毎秒30フレーム(fps)で実行されたことが検証された。装置30は、(1)異なる肌の色、(2)正面から見た顔の向き及びスケールの変化、並びに(3)照明の変化によらず個人を成功裏にトラッキングした。

30

【0080】

この実施の形態では、鼻梁のトラッキングのために、凸曲面トラッキングフィルタを用いる。上述の説明はまた鼻梁先端の光強度プロファイルの第1次導関数が最大値を有することを示している。同様に、光強度プロファイルの第2次導関数は鼻孔線でゼロである。鼻先端はこれら2個のレベルの間にある。

【0081】

この発明の装置と方法とが、鼻梁トラッキングに関連して説明された。しかし、他の顔の特徴点をトラッキングするために、様々な修正や変更が可能である。

40

【0082】

上述の実施の形態では、画像はカラーであったが、この発明はそのような実施の形態に限定されず、画像は光強度画像のみであってもよい。その場合、光強度情報そのものを直接、光強度情報の表現として用いてもよい。

【0083】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味および範囲内のすべての変更を含む。

50

【図面の簡単な説明】

【 0 0 8 4 】

【図 1】この発明の一実施の形態のトラッキング装置 3 0 のブロック図である。

【図 2】この実施の形態のトラッキング装置 3 0 を実現するコンピュータ上で実行されるプログラムのフローチャートである。

【図 3】顔 R O I の抽出を示す図である。

【図 4】行セグメント 1 6 0 が顔 R O I からどのように抽出されるかを示す図である。

【図 5】この実施の形態で用いられる、3 セグメントの畳込みフィルタ 1 8 0 を示す図である。

【図 6】行セグメントの光強度値の累積和 2 0 0 を示す図である。

10

【図 7】成功裏に抽出された鼻梁候補と、鼻梁及び鼻先端の更なる処理により正確に見出された、幾つかの誤った鼻梁候補とを示す図である。

【図 8】成功裏に抽出された鼻梁候補（その幾つかは鼻先端点より下にあり従って正確でない。）と、その光強度プロファイルとを示す図である。

【図 9】 $T(r)$ の定義を示す図である。

【図 1 0】列セグメントを示す図である。

【図 1 1】列セグメントのプロファイルを示す図である。

【図 1 2】鼻先端をどのように決定するかを示す図である。

【図 1 3】上述のプログラムを実行し、この実施の形態の装置 3 0 を実現するコンピュータシステム 3 3 0 の外観を示す図である。

20

【図 1 4】図 1 3 に示すコンピュータ 3 4 0 の構造を示す図である。

【図 1 5】この実施の形態の実験結果を示す図である。

【図 1 6】この実施の形態の実験結果を示す図である。

【符号の説明】

【 0 0 8 5 】

3 0 トラッキング装置

4 0 ビデオカメラ

4 2 フレームメモリ

4 4 眼検出モジュール

4 6 鼻梁検出モジュール

30

1 2 2、1 2 4、1 4 2、1 4 4 眼の位置

1 2 6、1 4 0 顔の対象領域 (R O I)

1 6 0 行セグメント

1 7 0 光強度プロファイル

1 8 0 最大フィルタ

2 7 0 列セグメント

2 8 0 光強度プロファイル (h p r o j)

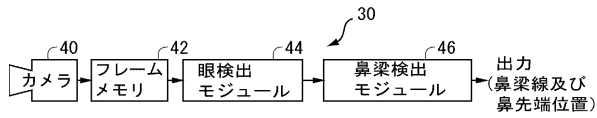
2 8 2 平滑化された光強度プロファイル (s h p r o j)

3 3 0 コンピュータシステム

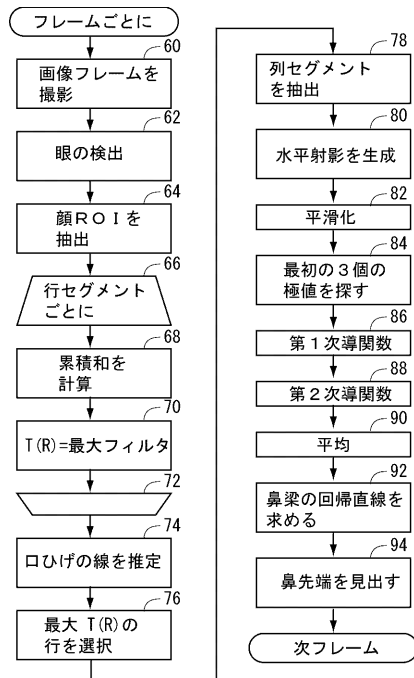
3 4 0 コンピュータ

40

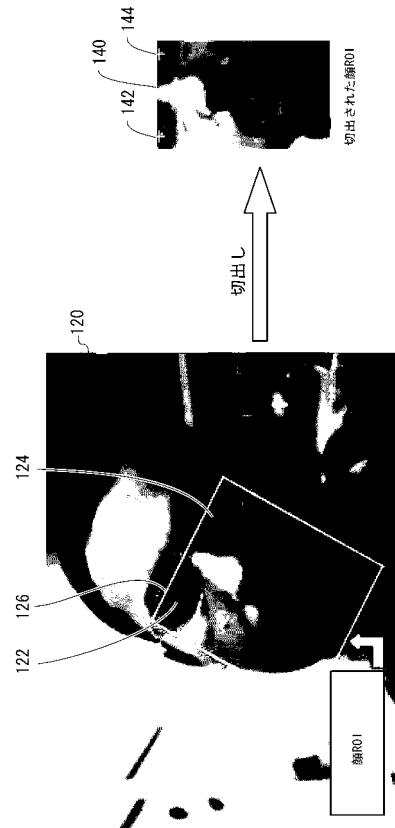
【図 1】



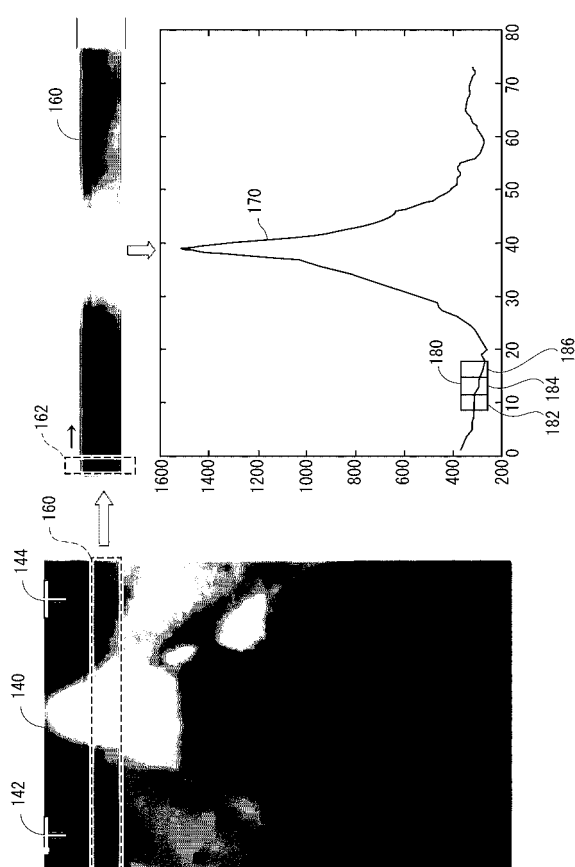
【図 2】



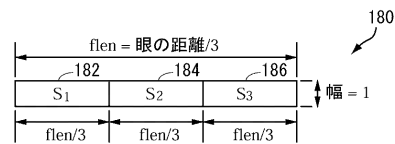
【図 3】



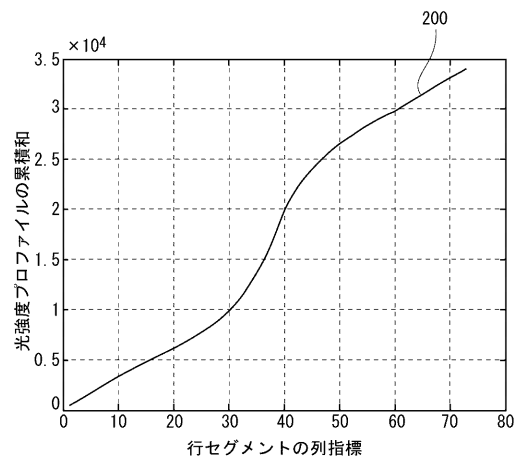
【図 4】



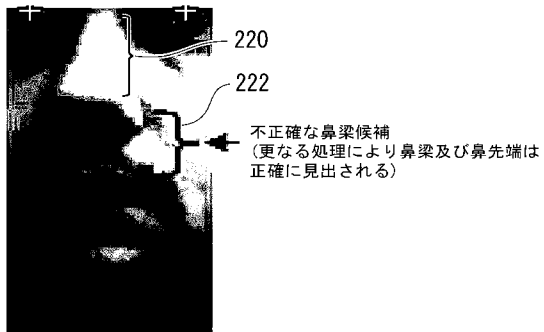
【図 5】



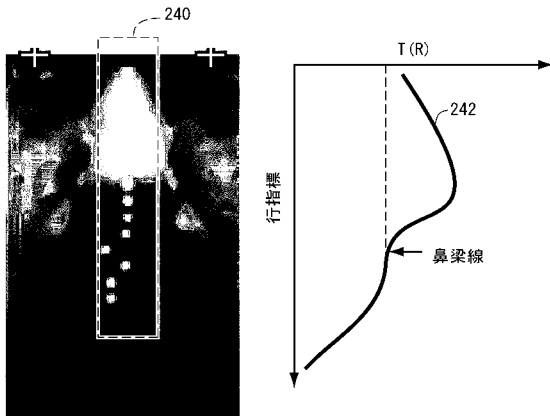
【図 6】



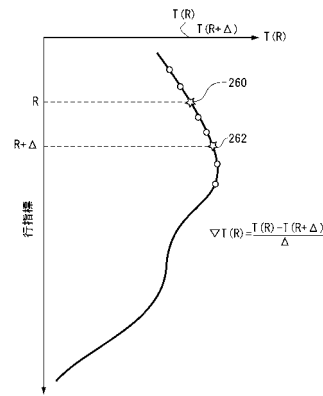
【図 7】



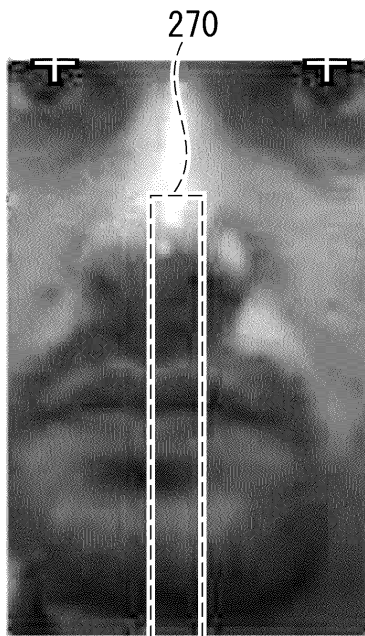
【図 8】



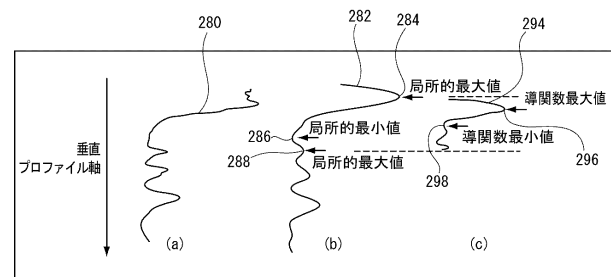
【図 9】



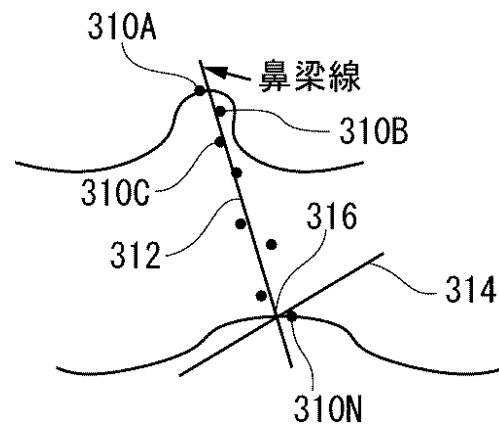
【図 10】



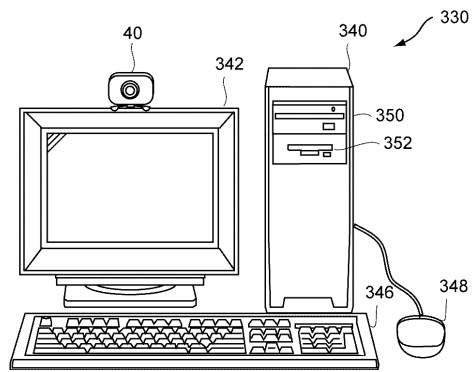
【図 11】



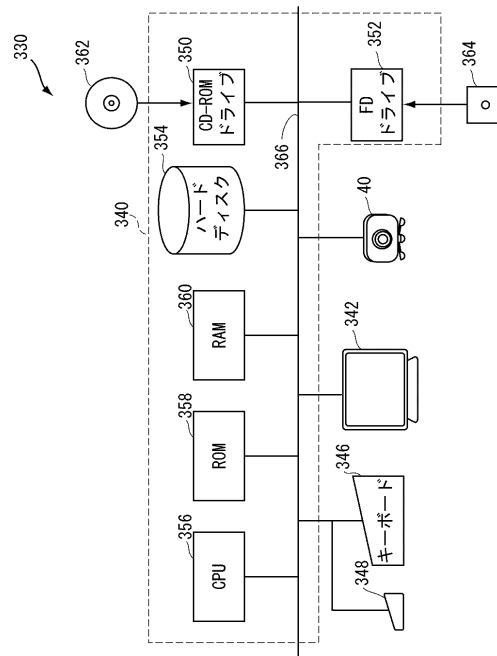
【図 12】



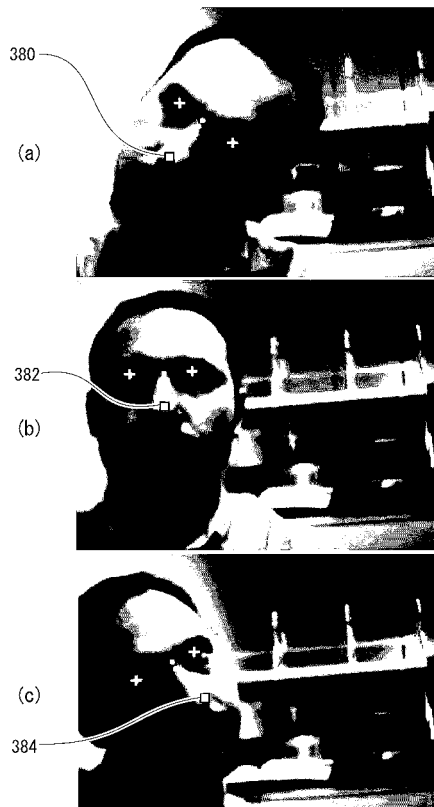
【図 13】



【図 14】



【図 15】



【図 16】



フロントページの続き

審査官 秦野 孝一郎

(56)参考文献 特開 2 0 0 4 - 1 5 7 7 7 8 (J P , A)
特開 2 0 0 4 - 2 3 4 3 6 7 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
G 0 6 T 1 / 0 0
G 0 6 T 7 / 0 0