

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4831750号
(P4831750)

(45) 発行日 平成23年12月7日(2011.12.7)

(24) 登録日 平成23年9月30日(2011.9.30)

(51) Int.Cl. F I
A 6 1 M 21/00 (2006.01) A 6 1 M 21/00
A 6 3 B 69/00 (2006.01) A 6 3 B 69/00 A

請求項の数 7 (全 25 頁)

<p>(21) 出願番号 特願2006-300403 (P2006-300403)</p> <p>(22) 出願日 平成18年11月6日(2006.11.6)</p> <p>(65) 公開番号 特開2008-113875 (P2008-113875A)</p> <p>(43) 公開日 平成20年5月22日(2008.5.22)</p> <p>審査請求日 平成21年9月9日(2009.9.9)</p> <p>(出願人による申告)平成18年度独立行政法人情報通信研究機構「民間基盤技術研究促進制度／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」、産業活力再生特別措置法第30条の適用を受ける特許出願</p>	<p>(73) 特許権者 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2</p> <p>(74) 代理人 100090181 弁理士 山田 義人</p> <p>(72) 発明者 米澤 朋子 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</p> <p>(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</p> <p>(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</p> <p style="text-align: right;">最終頁に続く</p>
---	---

(54) 【発明の名称】 コミュニケーション誘発システム

(57) 【特許請求の範囲】

【請求項1】

それぞれが視線方向を変更できる2体の擬人的媒体を用い、ユーザが視認できる位置に間隔を隔てて前記2体の擬人的媒体を配置し、前記2体の擬人的媒体のそれぞれに対して隣接する隣接区画と前記隣接区画の外側の周辺区画とを設定し、前記ユーザの視線が前記隣接区画および周辺区画のどこに向けられているかに応じて異なる態様で前記ユーザからのコミュニケーションを誘発するコミュニケーション誘発システムであって、

前記ユーザの視線の位置を検出する視線位置検出手段、

前記擬人的媒体の位置に対する前記視線の位置を判定する位置判定手段、および

前記ユーザの視線が前記周辺区画に向けられているとき、前記2体の擬人的媒体のそれぞれの視線方向を、前記ユーザの視線を追うように、制御する視線制御手段を備える、 コミュニケーション誘発システム。

【請求項2】

前記視線が前記2体の擬人的媒体のどちらかの位置にあるとき、前記2体の擬人的媒体が連動した動作をして前記ユーザからのコミュニケーションを誘発するように前記2体の擬人的媒体を制御する連動動作手段をさらに備える、 請求項1記載のコミュニケーション誘発システム。

【請求項3】

前記2体の擬人的媒体は可動の腕を有し、前記連動動作手段は互いに近いほうの腕を動かす第1腕動作手段を含む、 請求項1または2記載のコミュニケーション誘発システム。

10

20

【請求項 4】

前記 2 体の擬人的媒体は発音手段を有し、さらに
前記ユーザの視線が前記隣接区画に向けられているとき、前記 2 体の擬人的媒体が前記ユーザに対して呼びかけ音声を出力するように、それぞれの発音手段を制御する発音制御手段を備える、請求項 1 ないし 3 のいずれかに記載のコミュニケーション誘発システム。

【請求項 5】

前記 2 体の擬人的媒体は可動の腕を有し、さらに
前記ユーザの視線に近い腕を動かす第 2 腕動作手段を含む、請求項 4 記載のコミュニケーション誘発システム。

【請求項 6】

前記ユーザを撮影可能なカメラを有するテレビ電話機、
 前記ユーザとのコミュニケーションがあったとき、前記テレビ電話機をオンして予め登録している電話番号に発呼する発呼制御手段をさらに備える、請求項 1 ないし 5 のいずれかに記載のコミュニケーション誘発システム。

【請求項 7】

ユーザと他の人との間で通信する通信手段、および
 前記視線位置検出手段によって検出した前記ユーザの視線の位置が前記通信手段の位置に該当するとき前記通信手段を制御して前記ユーザと前記他の人との通信を可能にする手段をさらに備える、請求項 1 ないし 6 のいずれかに記載のコミュニケーション誘発システム。

【発明の詳細な説明】

【技術分野】

【0001】

この発明はコミュニケーション誘発システムに関し、特に、たとえばロボットなどの擬人的媒体を認知症患者などの軽度脳障害を持つユーザが視認できる位置に配置し、その擬人的媒体によってユーザからのコミュニケーションを誘発する、新規なコミュニケーション誘発システムに関する。

【背景技術】

【0002】

特許文献 1 などで、認知症患者などに右脳刺激を与えることによって、その機能回復が図れることが知られている。

【特許文献 1】特開 2005 - 160806 号公報 [A63B 24/00 23/035]

【発明の開示】

【発明が解決しようとする課題】

【0003】

しかしながら、このような機能回復訓練を受けられるのは少なくとも他人とのコミュニケーションが図れることが条件である。そこで、このような患者ではまずコミュニケーション能力を回復させる必要がある。その場合、患者からコミュニケーションを引出す（誘発する）ように作用するシステムがあれば好都合である。

【0004】

それゆえに、この発明の主たる目的は、新規な、コミュニケーション誘発システムを提供することである。

【0005】

この発明の他の目的は、ユーザからのコミュニケーションを積極的に引き出すことができる、コミュニケーション誘発システムを提供することである。

【課題を解決するための手段】

【0006】

請求項 1 の発明は、それぞれが視線方向を変更できる 2 体の擬人的媒体を用い、ユーザが視認できる位置に間隔を隔てて 2 体の擬人的媒体を配置し、2 体の擬人的媒体のそれぞれに対して隣接する隣接区画と隣接区画の外側の周辺区画とを設定し、ユーザの視線が隣

10

20

30

40

50

接区画および周辺区画のどこに向けられているかに応じて異なる態様でユーザからのコミュニケーションを誘発するコミュニケーション誘発システムであって、ユーザの視線の位置を検出する視線位置検出手段、擬人的媒体の位置に対する視線の位置を判定する位置判定手段、およびユーザの視線が周辺区画に向けられているとき、2体の擬人的媒体のそれぞれの視線方向を、ユーザの視線を追うように、制御する視線制御手段を備える、コミュニケーション誘発システムである。

【0007】

請求項1の発明において、2体の擬人的媒体(14A, 14B:実施例で相当する部分を例示する参照符号。以下同様。)がユーザ(12)の前方の、ユーザの視線が届く位置に互いに間隔を隔てて配置される。2体の擬人的媒体のそれぞれに対して隣接する隣接区画と隣接区画の外側の周辺区画とを設定し、コミュニケーション誘発システムはユーザの視線が隣接区画および周辺区画のどこに向けられているかに応じて異なる態様でユーザからのコミュニケーションを誘発する。視線位置検出手段(16, 18)は、たとえば、カメラ(16)によって撮影したユーザの顔画像を視線サーバ(18)で処理することによって、ユーザの視線(12A)の方向または位置をリアルタイムに推定または検出する。位置判定手段(32)は、ぬいぐるみの動作を制御するコンピュータの一部の機能であり、この位置判定手段が、そのときのユーザの視線の位置がぬいぐるみの位置に対してどのような位置関係にあるか判定する。たとえば、ユーザの視界にぬいぐるみがあるか、視線の位置はぬいぐるみの位置とかなり離れている(周辺区画)か、視線の位置はぬいぐるみの位置に隣接する位置(隣接区画)か、視線の位置はぬいぐるみの位置か、などを判定する。視線制御手段(32, S11)は、ユーザの視線が周辺区画に向けられているとき、2体の擬人的媒体のそれぞれの視線方向を、ユーザの視線を追うように、制御する。つまり、2体の擬人的媒体はユーザの視線ユーザの視線を共同注視するように、制御される。

【0008】

請求項1の発明によれば、ユーザの視線の位置の擬人的媒体の位置に対する判定結果に応じて異なる動作をするように擬人的媒体を制御することができ、ユーザの視線位置が周辺区画にあるとき2体の擬人的媒体がユーザの視線を共同注視させるので、ユーザからのコミュニケーションを誘発することができる。

【0009】

請求項2の発明は、視線が2体の擬人的媒体のどちらかの位置にあるとき、2体の擬人的媒体が連動した動作をしてユーザからのコミュニケーションを誘発するように2体の擬人的媒体を制御する連動動作手段をさらに備える、請求項1記載のコミュニケーション誘発システムである。

【0010】

請求項2の発明では、ユーザの視線が擬人的媒体(ぬいぐるみ)に向けられたとき、その擬人的媒体とユーザとの間でコミュニケーションを確立することができる。

【0011】

請求項3の発明は、2体の擬人的媒体は可動の腕を有し、連動動作手段は互いに近いほうの腕を動かす第1腕動作手段を含む、請求項1または2記載のコミュニケーション誘発システムである。

請求項4の発明は、2体の擬人的媒体は発音手段を有し、さらにユーザの視線が隣接区画に向けられているとき、2体の擬人的媒体がユーザに対して呼びかけ音声を出力するように、それぞれの発音手段を制御する発音制御手段を備える、請求項1ないし3のいずれかに記載のコミュニケーション誘発システムである。

請求項5の発明は、2体の擬人的媒体は可動の腕を有し、さらにユーザの視線に近い腕を動かす第2腕動作手段を含む、請求項4記載のコミュニケーション誘発システムである。

請求項6の発明は、ユーザを撮影可能なカメラを有するテレビ電話機、ユーザとのコミュニケーションが所定時間継続したとき、テレビ電話機をオンして予め登録している電話番号に発呼する発呼制御手段をさらに備える、請求項1ないし5のいずれかに記載のコミ

10

20

30

40

50

コミュニケーション誘発システムである。

【0012】

請求項3、請求項4および請求項5のそれぞれの発明においても、最適のコミュニケーション誘発動作を擬人的媒体に行なわせることができる。

請求項6の発明では、ユーザと擬人的媒体との間のコミュニケーションが一定時間以上継続的に行なわれたとき、つまり、ユーザがぬいぐるみに関心をもったとき、ユーザを撮影するTV電話機(60)をオンして自動的に発呼させるので、たとえば家族や介護士などが、その状態でのユーザを遠隔地で確認することができる。

【0013】

請求項7の発明は、ユーザと他の人との間で通信する通信手段、および視線位置検出手段によって検出したユーザの視線の位置が通信手段の位置に該当するとき通信手段を制御してユーザと他の人との通信を可能にする手段をさらに備える、請求項1ないし6のいずれかに記載のコミュニケーション誘発システムである。

10

【0014】

請求項7の発明では、ユーザが通信手段を注視したとき、その通信手段を用いてユーザと他の人との通信を開始させることができる。

【発明の効果】

【0015】

この発明によれば、ユーザの視線位置がどこにあるかによって最適のコミュニケーション誘発動作を擬人的媒体に行なわせることができるので、ユーザからのコミュニケーションを積極的に引出すことができる。

20

【0016】

この発明の上述の目的、その他の目的、特徴、および利点は、図面を参照して行う以下の実施例の詳細な説明から一層明らかとなろう。

【発明を実施するための最良の形態】

【0017】

図1に示すこの発明の一実施例のコミュニケーション誘発システム10は、たとえば認知症患者のような軽度脳障害を持つ被験者またはユーザ12からのコミュニケーションを積極的に引き出すように、このユーザ12に働きかけるぬいぐるみ14Aおよび14Bを含む。このぬいぐるみ14Aおよび14Bが擬人的媒体である。擬人的媒体とは、人間のように発話できたり、あるいは人間のような動作ができたりする媒体のことであり、典型的にはぬいぐるみや、ヒューマノイドやコミュニケーションロボットなども、この擬人的媒体として十分機能できる。

30

【0018】

この実施例のシステム10では、ユーザ12の主として顔を撮影するカメラ16を用いて、ユーザ12の視線の方向や位置をリアルタイムで検出し、その視線の方向や位置に応じてぬいぐるみ14Aおよび14Bの発話や動作を制御することによって、ユーザ12からのぬいぐるみ14Aおよび14Bに対するコミュニケーションを誘発する。なお、ぬいぐるみ14Aおよび14Bを区別なく説明する必要があるときには、参照番号「14」として表示することがあることに留意されたい。

40

【0019】

ユーザ12の視線方向または位置は、後に詳しく説明するように、視線サーバ18がカメラ16からの顔画像データまたは信号を処理することによって、リアルタイムで検出または判定する。

【0020】

図2は図1実施例を俯瞰した状態を図解する図解図であり、図3はそれぞれを側面から見た図解図である。これらの図からわかるように、患者または被験者ないしユーザ12は部屋のような空間10Aの一方に椅子に腰掛けて存在していて、その前方の、この空間10A内の他方に、2体のぬいぐるみ14Aおよび14Bが互いの間に適宜の間隔を隔てて、一直線上に配置されている。ただし、ぬいぐるみ14Aおよび14Bの配置は、両者の

50

間に間隔がある状態なら、直線上の配置でなくても問題ない。

【 0 0 2 1 】

そして、カメラ 1 6 は空間 1 0 A の一隅からユーザ 1 2 の顔前面を撮影できるように設置されている。

【 0 0 2 2 】

ユーザ 1 2 の視線 1 2 A が、ぬいぐるみ 1 4 A (1 4 B) に正対する方向の線 vline に対してずれている角度が旋回角 であり、視線 1 2 A が水平線 hline に対してずれている角度が俯仰角 として図 1 の視線サーバ 1 8 によってリアルタイムに計算される。そして、各ぬいぐるみ 1 4 A および 1 4 B がその検出角度 および からその視線 1 2 A の方向または位置を特定し、その視線 1 2 A の方向または位置に応じて、異なる動作および / または発話をする事によって、ユーザ 1 2 からコミュニケーションを誘発しようとするのである。

10

【 0 0 2 3 】

図 4 にはぬいぐるみ 1 4 が図示される。このぬいぐるみ 1 4 は、頭部 2 0 とそれを支える胴体 2 2 とを含む。胴体 2 2 の上部に左右の左腕 2 4 L および右腕 2 4 R が設けられ、頭部 2 0 には、前面に口 2 6 が配置され、その口 2 6 の上方には眼球 2 8 が設けられる。頭部 2 0 の上部側面は耳 3 0 が取り付けられている。

【 0 0 2 4 】

胴部 2 0 は胴体 2 2 によって、旋回・俯仰可能に支持され、また、眼球 2 8 も可動的に保持されている。口 2 6 にはスピーカ 4 8 (図 5) が内蔵されていて、耳 3 0 にはマイク 5 0 (図 5) が内蔵されている。なお、マイク 5 0 を両方の耳 3 0 にそれぞれ内蔵すれば、ステレオマイクとして機能し、それによって、そのステレオマイクに入力された音声の位置を必要に応じて特定することができる。

20

【 0 0 2 5 】

図 5 に示すように、実施例のぬいぐるみ 1 4 にはコンピュータ 3 2 が内蔵されていて、このコンピュータ 3 2 が、通信路の一例であるバス 3 4 を通して、図 1 に示す視線サーバ 1 8 に結合される。したがって、コンピュータ 3 2 は、視線サーバ 1 8 が特定または検出したユーザ 1 2 の視線方向や位置を示すデータをこのバス 3 4 を通して刻々受け取ることができる。ただし、通信路はバス 3 4 であっても、その他の形式の通信路であっても、さらには無線であっても、有線であってもよい。

30

【 0 0 2 6 】

コンピュータ 3 2 には、バス 3 4 を介してメモリ 3 6 が結合される。このメモリ 3 6 には図示しない ROM や RAM が組み込まれていて、ROM には主として、後述のフローチャート (図 2 0 - 2 1 , 2 3) で表現されるプログラムが予め記憶されているとともに、ぬいぐるみ 1 4 の各フェーズ (内部状態) の動作と発話 (音声) とを対応的に設定している、動き-音声反応対応テーブル 3 6 A が予め設定されている。この動き-音声反応対応テーブル 3 6 A は、後述するぬいぐるみ 1 4 の内部状態「 0 」 「 4 」のそれぞれにおいてぬいぐるみ 1 4 が発話する音声と実行できる動作とを対応させて設定したものである。RAM は、一時記憶メモリとして用いられるとともに、ワーキングメモリとして利用される。

40

【 0 0 2 7 】

モータ制御ボード 3 8 は、たとえば DSP (Digital Signal Processor) で構成され、図 2 に示すぬいぐるみ 1 4 の各腕や頭部の各軸モータを制御する。すなわち、モータ制御ボード 3 8 は、コンピュータ 3 2 からの制御データを受け、右腕 2 4 R (図 4) を前後や左右に動かすことができるように、X , Y および Z 軸のそれぞれの角度を制御する 3 つのモータ (図 5 ではまとめて、「右腕モータ」として示す。) 4 0 R の回転角度を調節する。また、モータ制御ボード 3 8 は、左腕 2 4 L の 3 つのモータ (図 5 ではまとめて、「左腕モータ」として示す。) 4 0 L の回転角度を調節する。モータ制御ボード 3 8 は、また、頭部 2 0 の旋回角や俯仰角を制御する 3 のモータ (図 5 ではまとめて、「頭部モータ」として示す。) 4 2 の回転角度を調節する。モータ制御ボード 3 8 は、また、眼球 2 8 を動

50

かす眼球モータ 4 4 も制御する。

【 0 0 2 8 】

なお、上述のモータは、制御を簡単化するためにそれぞれステッピングモータまたはパルスモータであるが、直流モータであってよい。

【 0 0 2 9 】

スピーカ 4 8 には音声入力 / 出力ボード 4 6 を介して、コンピュータ 3 2 から、合成音声データが与えられ、それに応じて、スピーカ 4 8 からはそのデータに従った音声または声が出力される。そして、マイク 5 0 からの音声入力が、音声入力 / 出力ボード 4 6 を介して、コンピュータ 3 2 に取り込まれる。

【 0 0 3 0 】

センサ入力 / 出力ボード 5 2 も、同様に、DSP で構成され、各センサやカメラからの信号を取り込んでコンピュータ 3 2 に与えるが、実施例ではあまり関係がないので、ここでは、そのセンサなどの詳細な説明は省略する。

【 0 0 3 1 】

なお、図 1 に示す実施例において、ユーザ 1 2 の前方すなわちぬいぐるみ 1 4 A および 1 4 B の後方は、図 6 に示すように区画されている。ユーザ 1 2 の視線位置が区画 A I のいずれにあるかによって、ぬいぐるみ 1 4 A および 1 4 B の発話や動作を異ならせ、それによってユーザ 1 2 の積極的な関与つまりコミュニケーションを誘発しようとするものである。

【 0 0 3 2 】

実施例では、ぬいぐるみ 1 4 A および 1 4 B は、図 6 に示す区画 X および Y に固定的に配置されているものとする。そして、そのぬいぐるみ 1 4 A および 1 4 B すなわち区画 X および Y の上方に 3 つの区画 A , B および C が設定され、区画 X を挟んで左右に区画 D および E が設定され、その区画 E とともに区画 Y を挟むように区画 F が設定されている。区画 X および Y の下方に 3 つの区画 G , H および I が設定される。

【 0 0 3 3 】

ただし、区画 A I および区画 X および Y 以外にユーザ 1 2 の視線が位置する場合には、「範囲外」となる。

【 0 0 3 4 】

また、これらの区画 A I や X および Y のそれぞれに包含される空間座標は空間 1 0 A (図 2 , 3) に固有に決まっているので、図 5 に示すメモリ 3 6 の所定の記憶場所に、これら全ての区画の座標データが予め記憶されている。

【 0 0 3 5 】

なお、実施例では、ぬいぐるみ 1 4 A および 1 4 B はユーザ 1 2 との間の距離、すなわち空間座標の Z 軸位置は一定なので、各区画の座標データとしては、X 軸と Y 軸だけでよい。たとえば、区画 A についてみると、この区画 A は、 $(x_{a1} - x_{a2}, y_{a1} - y_{a2})$ のように、X 軸 Y 軸とも範囲を有して設定される。ここで、 x_{a1} は、区画 A の X 軸の開始 (左端) 座標で、 x_{a2} は、区画 A の X 軸の終了 (右端) 座標を示す。同様に、 y_{a1} は、区画 A の Y 軸の開始 (下端) 座標で、 y_{a2} は、区画 A の Y 軸の終了 (上端) 座標を示す。残りの各区画 B - I , X , Y も同様に座標範囲のデータとして設定される。

【 0 0 3 6 】

また、この実施例では、ぬいぐるみ 1 4 A および 1 4 B が自律的に自己の動作や発話を制御する図 5 のような制御回路を持つものとして説明するが、それぞれのぬいぐるみ 1 4 A および 1 4 B を一括して、または個別的に制御するための 1 台またはそれ以上のコンピュータをぬいぐるみ 1 4 A および 1 4 B とは別に設けるようにしてもよい。

【 0 0 3 7 】

このような実施例において、まず、ユーザ 1 2 の視線を推定する方法について説明する。この実施例では、以下に説明するように、ユーザ 1 2 の視線方向を推定または検出するために 1 つのカメラ (単眼カメラ) を用いるユニークな方法を採用している。しかしながら、ユーザ 1 2 の視線 1 2 A (図 2 , 図 3) を検出する方法としては、2 以上のカメラを

10

20

30

40

50

利用する従来からの一般的な方法が採用されてもよいことは、勿論である。つまり、この発明では、ユーザ 12 の視線を推定したり検出したりする必要はあるが、その具体的な方法は重要な意味を持たず、どのような公知の方法が用いられてもよい。

【0038】

図 1 に示すように、ユーザ 12 の前方に、たとえば 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 16 が設置され、このカメラ 16 からの顔画像信号が視線サーバ 18 に取り込まれ、視線サーバ 18 が画像処理することによって、視線 12 A の角度 およびを推定する。

【0039】

図 7 に示すように、カメラ 16 により撮影された画像は、視線サーバ 18 に附属して設けたディスプレイ 54 (図 1 では図示せず) の撮影画像表示領域 56 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 56 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。

【0040】

視線サーバ 18 は、一般的なコンピュータであり、特に変わったハードウェア構成ではないので、ハードウェア自体は説明しないが、視線方向や視線位置の推定は、いかに説明するソフトウェアにより実現される。

【0041】

視線サーバ 18 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。

【0042】

視線方向の推定処理の動作の前提として、まず、たとえば 6 分割矩形フィルタを利用して、顔検出処理が実行される。

【0043】

視線サーバ 18 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。

【0044】

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。以下、さらに詳しく説明する。

【0045】

図 8 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り、図 8 (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。

【0046】

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。たとえば、1 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。そして、図 8 (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 S1 ~ S6 に分割する。

【0047】

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てみると、図

10

20

30

40

50

8 (b) のようになる。

【 0 0 4 8 】

ただし、図 8 の 6 分割フィルタは書く矩形領域が等分されたものであったが、このフィルタは図 9 に示すように変形されてもよい。

【 0 0 4 9 】

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 9 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。

10

【 0 0 5 0 】

図 9 に示す 6 分割矩形フィルタにおいて、それぞれのブロック S i (1 ≤ i ≤ 6) について、画素の輝度の平均値「バー S i 」 (S i に上付きの “ - ” をつける) を求める。

【 0 0 5 1 】

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) および (2) が成り立つ。

【 0 0 5 2 】

【数 1】

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

20

【 0 0 5 3 】

【数 2】

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

【 0 0 5 4 】

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。

【 0 0 5 5 】

矩形枠内の画素の総和を求める処理には、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。

30

【 0 0 5 6 】

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 5 7 】

40

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 10 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。

【 0 0 5 8 】

50

続いて、目、鼻や虹彩中心の位置を、公知の文献、たとえば『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)』あるいは、『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IE2002 - 263, pp. 25 - 29 (2003)』などの手法を用いて抽出する。

【0059】

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、ラブ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。

10

【0060】

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。

【0061】

図11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。

【0062】

視線の推定においては、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ものとする。

20

【0063】

図12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 r 、画像上での眼球中心と虹彩中心との距離を d とすると、視線方向とカメラ光軸とのなす角 θ は次式(3)で表される。

【0064】

【数3】

$$\theta = \sin^{-1}\left(\frac{d}{r}\right) \quad (3)$$

30

【0065】

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。

【0066】

図13は、図12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。

40

【0067】

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12がカメラ16を注視した場合について考えると、図13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。

【0068】

そこで、この実施例での視線推定では、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。

【0069】

50

後により詳しく説明するように、この実施例の視線方向の推定では、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。

【0070】

視線方向の推定のための初期設定として、視線サーバ18は、図14に示すフローチャートで表現されるキャリブレーションを実行する。

【0071】

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)。図15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。

【0072】

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。

【0073】

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処理を行い(ステップS104)、続いて、目や鼻の検出処理を行なう(ステップS106)。

【0074】

さらに、視線サーバ18は、特徴点の抽出、追跡を行う(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。

【0075】

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。

【0076】

【数4】

$$W = \begin{bmatrix} x_1^{(1)} & \dots & x_M^{(1)} & x_r^{(1)} & x_l^{(1)} \\ \vdots & & \vdots & & \\ x_1^{(N)} & \dots & x_M^{(N)} & x_r^{(N)} & x_l^{(N)} \end{bmatrix} \quad (4)$$

【0077】

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。

【0078】

【数5】

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

【0079】

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

10

20

30

40

50

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-II,8,pp.1497-1505(1993)に開示がある。

【0080】

図16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。

10

【0081】

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。

【0082】

まず、カメラ16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)。

【0083】

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て、現画像フレーム中の特徴点が特定される(ステップS206)。

20

【0084】

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。

【0085】

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。

30

【0086】

【数6】

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

40

【0087】

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 $S^{(k)}$ は行列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)。

【0088】

【数 7】

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

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

【0089】

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 S 2 1 0)。

【0090】

【数 8】

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

【0091】

【数 9】

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

【0092】

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 S 2 1 2)。

【0093】

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

【0094】

【数 10】

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

【0095】

【数 11】

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

【0096】

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 \mathbf{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。

【0097】

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いければ、処理はステップ S 2 0 2 に復帰する。

【0098】

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った結果について以下に説明する。

【0099】

10

20

30

40

50

カメラはElmo社製PTC - 400Cを用い、被験者から約150 [cm]の位置に設置した。

【0100】

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15に示したとおりである。

【0101】

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)。

【0102】

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。

【0103】

図17~図19は、視線推定結果を示す。図17は、右上方注視の状態であり、図18は、上方注視の状態であり、図19は、左下方向注視の状態である。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、顔の位置や向きの変化とは関係なく、視線方向が推定できた。

【0104】

以上説明したとおり、この実施例の視線方向の推定方法では、単眼カメラの観測に基づいて顔特徴点を検出し、追跡することにより視線方向を推定する。つまり、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の角度、を決定する。

【0105】

視線サーバ18が視線方向の角度、を検出した後、ぬいぐるみ14のコンピュータ32は図20-21に示すフローチャートを実行して、ぬいぐるみ14の発話や動作を制御する。ただし、この図20-21に示すフローチャートは、視線サーバ18からの視線角度データが入力される都度、すなわちフレーム毎に繰り返し実行される。

【0106】

図20の最初のステップS1で、このようにして視線サーバ18が推定した視線角度およびのデータが、ぬいぐるみ14のバス34(図5)を通してコンピュータ32に与えられる。

【0107】

コンピュータ32では、次のステップS3で、その角度データから視線12A(図2,3)の空間10Aにおける空間座標(x, y, z)を計算する。

【0108】

一方、ぬいぐるみ14Aおよび14Bはそれぞれ、図6に示す区画XおよびYに固定的に配置されている。したがって、このぬいぐるみ14Aおよび14Bの存在する空間10A内の座標は既に計算されて、たとえばメモリ36(図5)に記憶されている。

【0109】

そこで、次のステップS5で、コンピュータ32は、ステップS3で計算した視線の空間座標とメモリ36内に蓄積されている各ぬいぐるみの座標とを比較し、ユーザ12から見える範囲にぬいぐるみ14Aおよび/または14Bが存在するかどうか、つまり、そのときユーザ12にぬいぐるみ14Aおよび/または14Bが見えているかどうかを判断する。

【0110】

視線の座標が図6に外枠として示す「範囲外」の領域に含まれる座標に相当する場合には、このステップS5で“NO”が判断され、次のステップS7に進む。ただし、ステップS5で“YES”なら、つまり、そのときユーザ12はぬいぐるみ14A, 14Bが見

10

20

30

40

50

えているなら、ステップS 9に進む。

【0111】

ステップS 7で、コンピュータ3 2は、ぬいぐるみの内部状態を「0」に設定する。ただし、この「内部状態0」というのは「何もしない」ということであり、コンピュータ3 2は図5に示すモータ制御ボード3 8や音声入出力ボード4 6に何も制御データを与えない。

【0112】

ステップS 5で“YES”となってステップS 9に進んだとき、コンピュータ3 2は、視線位置の座標(x, y, z)データが、先に説明した図6の各区画A CおよびG Iのそれぞれの範囲座標データのいずれかに該当するかどうか判断する。ただし、ぬいぐるみ1 4 Aについては、さらに、区画Fに該当するかどうか、また、ぬいぐるみ1 4 Bに対しては区画Eに該当するかどうかを、判断する。各ぬいぐるみ1 4 Aおよび1 4 Bにとって、区画A CおよびG Iはかなり離れたところであり、ぬいぐるみ1 4 Aにとっての区画F、ぬいぐるみ1 4 Bにとっての区画Eも同様に離れた場所であると認識できる。このように、ユーザ1 2の視線位置がぬいぐるみ1 4 A、1 4 Bから離れた場所にあるときには、ぬいぐるみ1 4 Aおよび1 4 Bに「共同注視」という動作を行なわせる。

【0113】

したがって、このステップS 9で“YES”が判断されたときには、次のステップS 1 1で、コンピュータ3 2は、ぬいぐるみ1 4 Aおよび1 4 Bを、共同注視のモード、つまり内部状態「1」に設定する。共同注視とは、ユーザ1 2のコミュニケーションを幾分消極的に誘発する行為で、このモードでは、コンピュータ3 2は、ユーザ1 2の視線1 2 A(図2, 3)の方向を追うように、ぬいぐるみ1 4 Aおよび/または1 4 Bの頭部2 0(図4)および必要なら眼球2 8を動かすように、モータ制御ボード3 8に制御データを与える。このときの制御データは、メモリ3 6の動き-音声反応対応テーブル3 6 Aに予めこのモード(内部状態1)のために設定されているものを読み出して、モータ制御ボード3 8に与える。そして、モータ制御ボード3 8から頭部モータ4 2および必要に応じて眼球モータ4 4に制御信号を与える。

【0114】

なお、この実施例では、内部状態1のモードではぬいぐるみ1 4 Aおよび1 4 Bからは、発話するようには設定されていない。したがって、このときぬいぐるみ1 4 Aおよび1 4 Bのいずれも、ユーザ1 2に対する呼びかけのような音声を出すことはない。しかしながら、「共同注視」によってユーザ1 2の視線を追うような仕草をぬいぐるみ1 4 A, 1 4 Bがするので、ユーザの視線がぬいぐるみに向かっていなくても、そのようなぬいぐるみの動きに気づき、ぬいぐるみに関心を持ち始め、そこからコミュニケーションが引き出される可能性がある。

【0115】

ステップS 9で“NO”と判断したとき、すなわち、ユーザ1 2の視線がぬいぐるみ1 4 Aおよび/または1 4 Bの位置からそんなに離れていないと判断したとき、コンピュータ3 2は、次のステップS 1 3で、ユーザ1 2の視線1 2 Aは、ぬいぐるみ1 4 Aおよび/または1 4 Bの隣の区画に向けられているかどうか判断する。隣の区画とは、ぬいぐるみ1 4 Aについては区画DまたはEであり、ぬいぐるみ1 4 Bについては区画EまたはFである。したがって、コンピュータ3 2は、ステップS 1 3で、視線位置の座標(x, y, z)データが、区画DまたはE、もしくはEまたはFのそれぞれの範囲座標データのいずれかに該当するかどうか判断する。

【0116】

このステップS 1 3で“YES”が判断されるということは、ユーザ1 2の視線位置がぬいぐるみ1 4 Aおよび/または1 4 Bに隣接する区画に向けられていることを意味するので、コンピュータ3 2は、ユーザ1 2からコミュニケーションを誘発するように積極的に働きかけるように、各ぬいぐるみを制御する。

【0117】

10

20

30

40

50

具体的には、コンピュータ32は、ぬいぐるみ14Aおよび/または14Bの内部状態を「2」に設定する。このモードでは、コンピュータ32は、メモリ36から読み出した当該モード用の制御コマンドに従って、ユーザの視線のある位置(区画)に近い方の腕を動かしたりする動作を行わせる。

【0118】

たとえば、視線位置が区画Dであるときには、ぬいぐるみ14Aの右腕24Rが動くように、コンピュータ32は、ぬいぐるみ14Aのモータ制御ボード38から右腕モータ40Rに制御信号を与える。視線位置が区画Fであるときには、ぬいぐるみ14Bの左腕24Lが動くように、コンピュータ32はぬいぐるみ14Bのモータ制御ボード38から左腕モータ40Lに制御信号を与える。視線位置が区画Eであるなら、そのときコンピュータ32は、ぬいぐるみ14Aのモータ制御ボード38に制御コマンドを出力し、それによってぬいぐるみ14Aの左腕モータ40Lに制御信号を与えるとともに、ぬいぐるみ14Aのモータ制御ボード38にも制御コマンドを出力してぬいぐるみ14Bの右腕モータ40Rに制御信号を与える。したがって、この場合には、ぬいぐるみ14Aの左腕24Lおよびぬいぐるみ14Bの右腕24Rがともに動かされる。

10

【0119】

この内部状態2のモードでは、さらに、ぬいぐるみ14Aおよび/または14Bからユーザ12に対して発話による呼びかけを行う。たとえば、コンピュータ32は、動き-音声反対応テーブル36A(図5)から、「ねえねえ、さん」とか「おーい」とかの音声データを読み出し、その音声データを音声入出力ボード46に与える。そうすると、この音声入出力ボードからスピーカ48に音声信号が与えられ、該当のぬいぐるみ14Aおよび/または14Bのスピーカ48からユーザ12に対して呼びかけが行われる。

20

【0120】

このように、この実施例の内部状態2のモードでは、ユーザの視線に近いほうのぬいぐるみの腕が動かされるだけでなく、ぬいぐるみから呼びかけがあるので、ユーザ12は、先の内部状態1のときよりもさらに強く、ぬいぐるみの動きや声に気づき、ぬいぐるみに強く関心を持つ。したがって、ユーザ12からのコミュニケーションがさらに強く引き出される可能性がある。

【0121】

先のステップS13で“NO”が判断されたとき、つまり、ユーザ12の視線12Aの位置がぬいぐるみ14Aおよび14Bの周辺の区画にはなくかつそれらに隣接する区画にもないときには、つまり、その視線12Aは区画XおよびYのどちらかに向けられていることを意味する。すなわち、ステップS9で“NO”、ステップS13で“NO”のときは、ユーザの視線は、ぬいぐるみ14Aおよび14Bのどちらかに向けられている。したがって、次のステップS17では、視線位置は区画XおよびYつまり、ぬいぐるみ14Aおよび14Bのいずれにあるかを判断する。ここでは、ぬいぐるみ14Aに対しては区画Yにあるか、ぬいぐるみ14Bに対しては区画Xにあるか、という判断をしている。

30

【0122】

ステップS17で“NO”が判断されたとき、つまり、ぬいぐるみ14Aに対して視線位置が区画Xであるか、またはぬいぐるみ14Bに対して視線位置が区画Yであるとき、図21に示すステップS19に進み、ユーザ12によるコミュニケーションを確認するとともに、ユーザ12とのコミュニケーションを実行する。

40

【0123】

ステップS19では、まず、該当するぬいぐるみ14Aおよび14Bのどちらかのコンピュータ32は、該当するぬいぐるみの内部状態を「31」に設定する。そのため、このステップS19では、コンピュータ32は、メモリ36の動き-音声反対応テーブル36Aから内部状態3-1のための制御コマンドを読み出して、該当のぬいぐるみの動作や発話を制御する。

【0124】

具体的には、コンピュータ32はまず、テーブル36Aから読み出した音声メッセージ

50

データを音声入出力ボード 46 に設定することによって、このぬいぐるみ 14A または 14B のスピーカ 48 から「どうしたの」とか「お話したい？」などを発話し、ユーザ 12 に呼びかけまたは問いかけを行なう。と同時に、コンピュータ 32 は、テーブル 36A から読み出した動作制御データをモータ制御ボード 38 に設定することによって、たとえば頭部モータ 42 を制御して、ぬいぐるみ 14A または 14B の首をかしげる動作を行なわせる。このときの動作は、ユーザ 12 のコミュニケーションを確認する動作であり、先の発話と協働して、ユーザ 12 に対して、ユーザからコミュニケーションするかどうかの確認の機能を果たす。

【0125】

そして、次のステップ S21 で、まず、該当するぬいぐるみ 14A および 14B のどちらかのコンピュータ 32 は、該当するぬいぐるみの内部状態を「32」に設定する。そのため、このステップ S21 では、コンピュータ 32 は、メモリ 36 の動き-音声反応対応テーブル 36A から内部状態 3-2 のための制御コマンドを読み出して、該当のぬいぐるみの動作や発話を制御する。

【0126】

具体的には、コンピュータ 32 はまず、テーブル 36A から読み出した動作制御データをモータ制御ボード 38 に設定することによって、たとえば頭部モータ 42 や右腕モータ 40R もしくは左腕モータ 40L を制御して、該当のぬいぐるみ 14A または 14B にならず動作や、右腕 24R または左腕 24L を前に出して振るなどの動作を行なわせる。このときの動作は、ユーザ 12 とのコミュニケーションを開始する動作である。それとともに、テーブル 36A から読み出した音声メッセージデータを音声入出力ボード 46 に設定することによって、このぬいぐるみ 14A または 14B のスピーカ 48 から「うんうん」などの呼びかけ音声を発話し、ユーザ 12 に呼びかけを行なう。このように、内部状態 32 のときのぬいぐるみ 14A または 14B の動作や発話によって、ユーザ 12 とのコミュニケーションを開始することができる。

【0127】

つまり、内部状態 31 でユーザに対してコミュニケーションを開始させる働きかけを行ない、内部状態 32 で実際にコミュニケーションを実行することができる。

【0128】

なお、ステップ S17 での“YES”を判断するときには、ユーザ 12 の視線は反対側のぬいぐるみに向けられていることを意味し、この場合には、該当するぬいぐるみ 14A または 14B のコンピュータ 32 は、該当するぬいぐるみの内部状態を「4」に設定する。具体的には、コンピュータ 32 は、動き-音声反応対応テーブル 36A から読み出した動作制御データをモータ制御ボード 38 に設定することによって、たとえば右腕モータ 40R または左腕モータ 40L を制御して、反対側のぬいぐるみ 14B または 14A に近い方の腕を動かす動作を行なわせる。このときの動作は、ユーザ 12 とのコミュニケーションを実行しているぬいぐるみ 14A または 14B との連動した動作であり、そうさせることによってコミュニケーションをさらに継続させる効果が期待できる。

【0129】

図 22 はこの発明の他の実施例に従った制御回路の一例を示すブロック図で、先の実施例の図 5 に対応するものである。

【0130】

この実施例は、テレビ電話機 60 を設け、それをコンピュータ 32 によって制御データきるようにした点を除いて、図 5 と同じである。そして、TV 電話機 60 は、ユーザ 12 の近傍に配置され、ユーザ 12 の全身または部分（たとえば顔など）が撮影可能なカメラ（図示せず）を有する。TV 電話機 60 はよく知られているように、そのようなカメラで撮影した動画をたとえば光回線のような電話回線 62 を通して、相手方の同様の TV 電話機に送信するとともに、相手方からの動画をディスプレイ（図示せず）に表示できるようにされている。

【0131】

10

20

30

40

50

この実施例において、コンピュータ32がTV電話機60の動作を制御できる。つまり、図23に示すように、ステップS21でぬいぐるみ14Aまたは14Bの動作モードを32に設定してユーザ12とのコミュニケーションを開始した後、ステップS25でその状態が一定時間、たとえば3分間継続したかどうか判断し、“YES”のとき、次のステップS27で、TV電話機60をオンして、そのTV電話機60から、予め登録している電話番号に発呼する。ここで、登録している電話番号の例としては、ユーザと離れた家族の家のTV電話機の電話番号、ユーザの介護または介助を担当している介護士などが見えるTV電話機の電話番号などが考えられる。

【0132】

そして、この実施例では、ユーザ12とぬいぐるみ14Aまたは14Bとの間のコミュニケーションが一定時間以上継続的に行なわれたとき、つまり、ユーザがぬいぐるみに関心をもったとき、ユーザを撮影するTV電話機60から自動的に発呼するようにすれば、上記家族や介護士などが、その状態でのユーザをTV電話機で確認することができるので、たとえば認知症患者の行動を遠隔地で観察できる。

10

【0133】

この図22の実施例においても先の実施例と同様に、視線サーバ18によってユーザ12の視線12Aの方向または位置を推定し、その視線位置がどの区画にあるかによって異なる動作（発話および/または動き）をぬいぐるみ（擬人的媒体）14に行わせるようにするので、ユーザからのコミュニケーションを誘発することができる。

【0134】

なお、実施例で説明した、擬人的媒体すなわちぬいぐるみの動作（発話、動き）は単なる例示であり、ユーザの視線位置とそのぬいぐるみの位置との関係に応じて、任意の動作、たとえば、ユーザの関心を引くような動作、ユーザのコミュニケーションを継続させるための動作などを実行させることができる。

20

【0135】

さらに、上述の実施例では、ユーザがぬいぐるみと一定時間コミュニケーションをしたときにTV電話機60を自動発呼させるようにしたが、ユーザの視線位置がぬいぐるみの位置になったとき、時間経過なしに直ちにTV電話機から発呼させるようにしてもよい。

【0136】

また、ユーザの視線位置がぬいぐるみの位置になったことを条件とするのではなく、たとえば図6の区画A Iの適宜の区画にTV電話機60を設置しておき、ユーザの視線がその区画に向けられたら、そのことをトリガとしてTV電話機の発呼処理を行なうようにしてもよい。

30

【0137】

さらに、通信手段の例としてTV電話機を挙げたが、通信手段はTV電話機に限ることはなく、通常の音声電話機や携帯電話機、さらにはインタホンなど、ユーザと他の人との間で通信できる任意のものを含むことを理解されたい。したがって、たとえば、或る区画（図6）にインタホンを置いておき、ユーザの視線位置がその区画すなわちインタホンの位置に向けられたとき、つまり、ユーザがインタホンを注視したとき、そのインタホンを介してたとえば別の部屋の家族と会話できるようにすることなども容易に考えられる。このような制御は、視線サーバ18が行なってもよいし、各ぬいぐるみの内蔵コンピュータ32が行なってもよい。さらには、そのための別のコンピュータを設置することも可能である。

40

【図面の簡単な説明】

【0138】

【図1】この発明の一実施例のコミュニケーション誘発システムの概念を示す図解図である。

【図2】図1実施例におけるユーザとぬいぐるみとの平面的な位置関係およびユーザの視線角度を示す図解図である。

【図3】図1実施例におけるユーザとぬいぐるみとの側面的な位置関係およびユーザの視

50

線角度を示す図解図である。

【図 4】図 1 実施例において用いられるぬいぐるみの一例を示す図解図である。

【図 5】図 1 実施例におけるぬいぐるみの制御回路の一例を示すブロック図である。

【図 6】図 1 実施例におけるユニットの視線の位置を判定する区画の一例を示す図解図である。

【図 7】図 1 実施例における視線サーバのディスプレイに表示されているユーザの顔画像の一例を示す図解図である。

【図 8】図 8 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。

【図 9】図 9 は 6 分割矩形フィルタの他の構成を示す概念図である。

【図 10】図 10 は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解図である。 10

【図 11】図 11 は顔検出結果の例を示す図解図である。

【図 12】図 12 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。

【図 13】図 13 は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 14】図 14 は視線サーバによる初期設定の処理動作を示すフロー図である。

【図 15】図 15 は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解図である。

【図 16】図 16 は視線サーバが実行するリアルタイム視線検出の処理動作を示すフロー図である。 20

【図 17】図 17 は右上方注視の状態での視線推定結果を示す図解図である。

【図 18】図 18 は上方注視の状態での視線推定結果を示す図解図である。

【図 19】図 19 は左下方注視の状態での視線推定結果を示す図である。

【図 20】図 1 実施例におけるぬいぐるみのコンピュータの動作を示すフロー図である。

【図 21】図 1 実施例におけるぬいぐるみのコンピュータの動作の図 20 の続きを示すフロー図である。

【図 22】この発明の他の実施例におけるぬいぐるみの制御回路の一例を示すブロック図である。

【図 23】図 2 実施例におけるぬいぐるみのコンピュータの動作の図 20 の続きを示すフロー図である。 30

【符号の説明】

【 0 1 3 9 】

1 0 ... コミュニケーション誘発システム

1 4 , 1 4 A , 1 4 B ... ぬいぐるみ

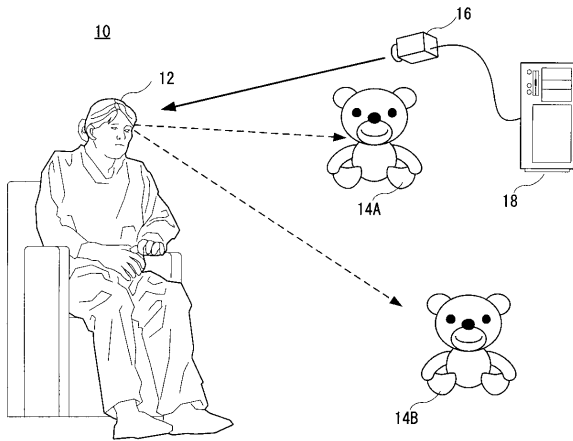
1 6 ... カメラ

1 8 ... 視線サーバ

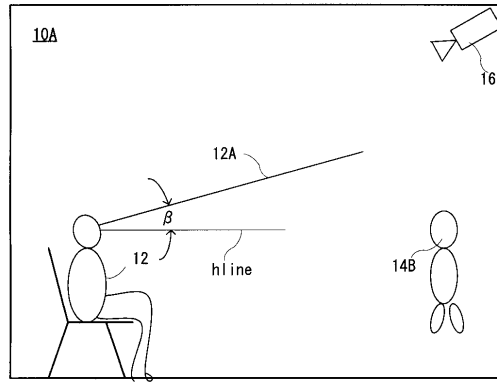
3 2 ... コンピュータ

3 6 A ... 動き 音声反応対応テーブル

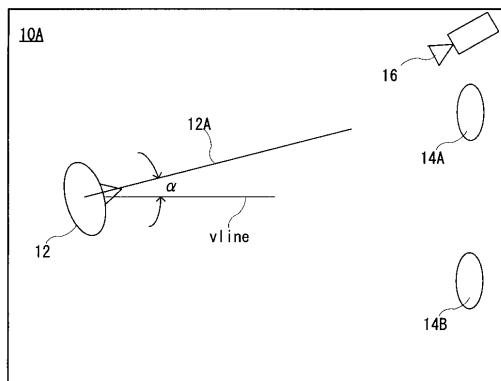
【図1】



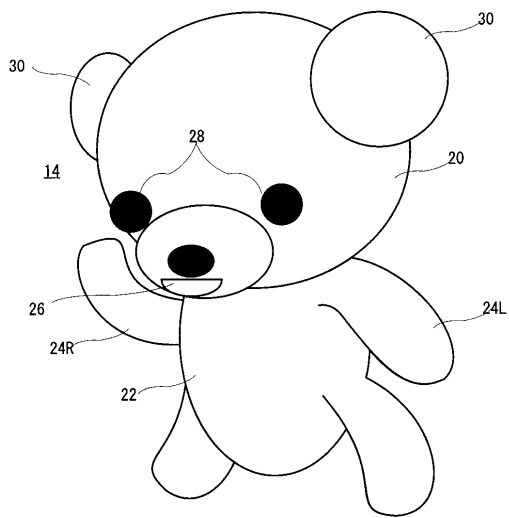
【図3】



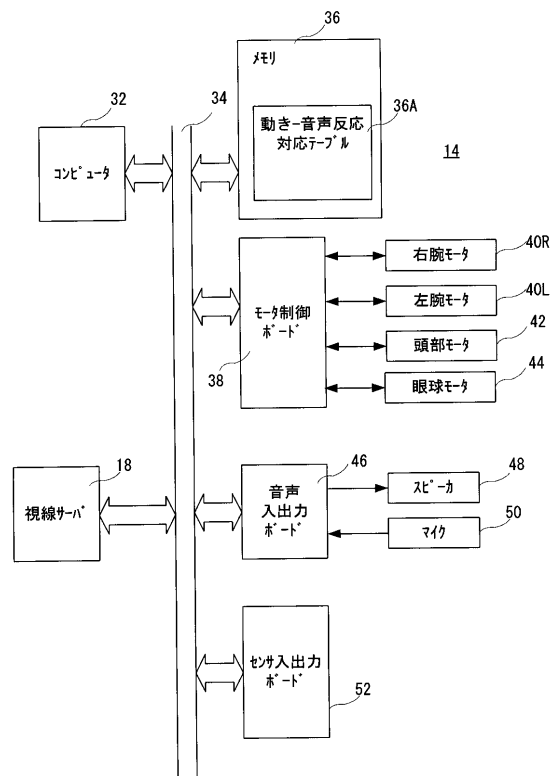
【図2】



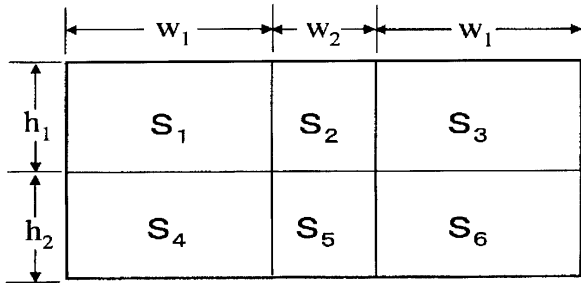
【図4】



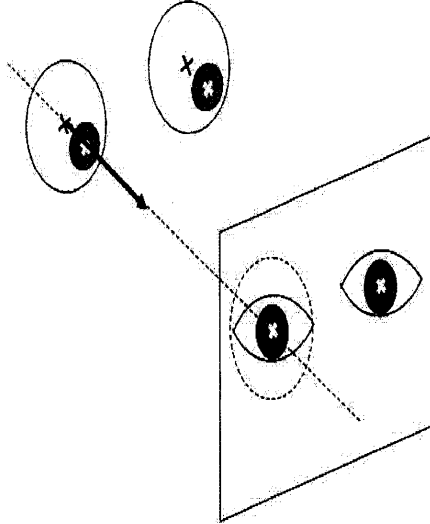
【図5】



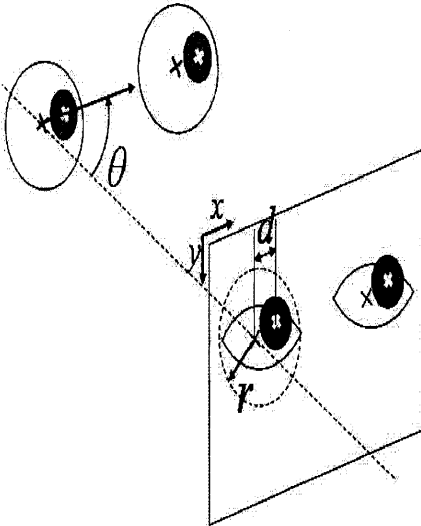
【図9】



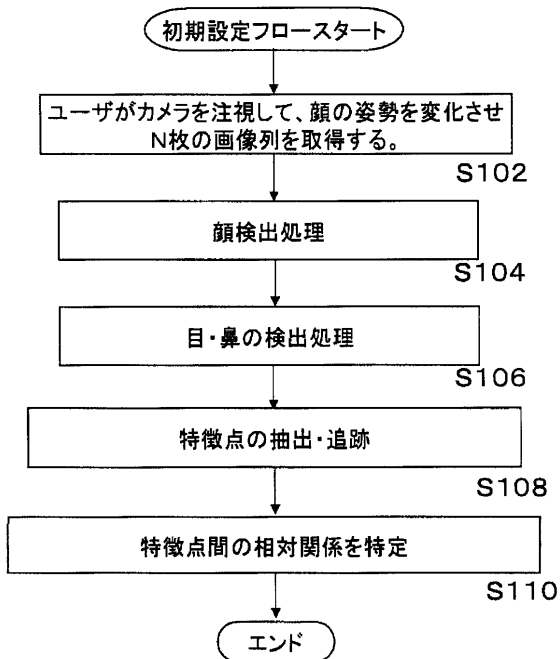
【図13】



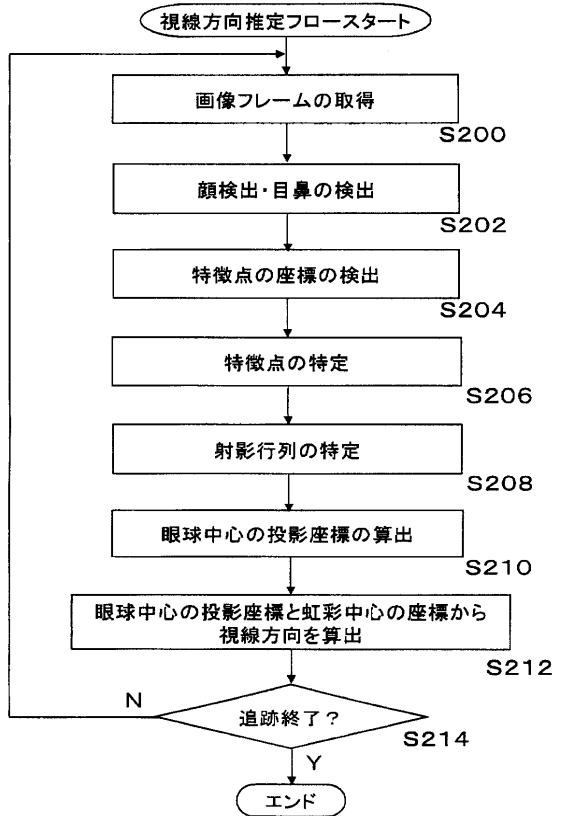
【図12】



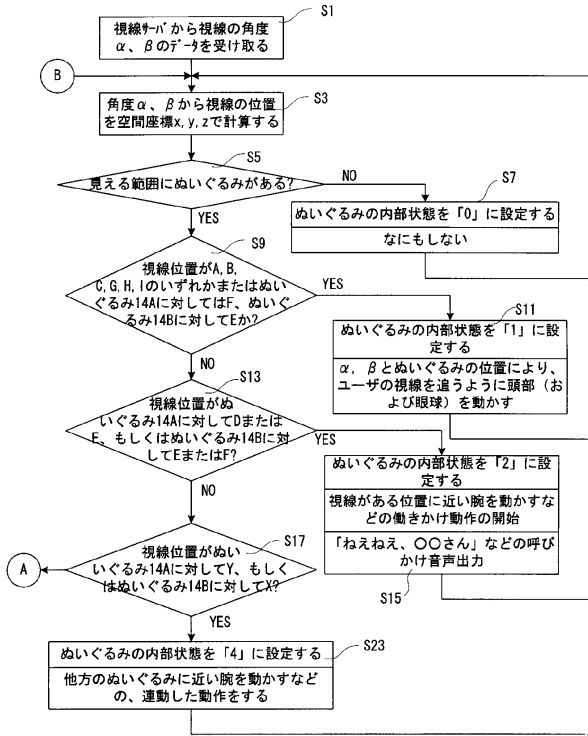
【図14】



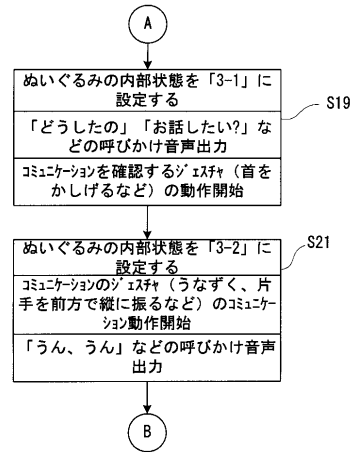
【図16】



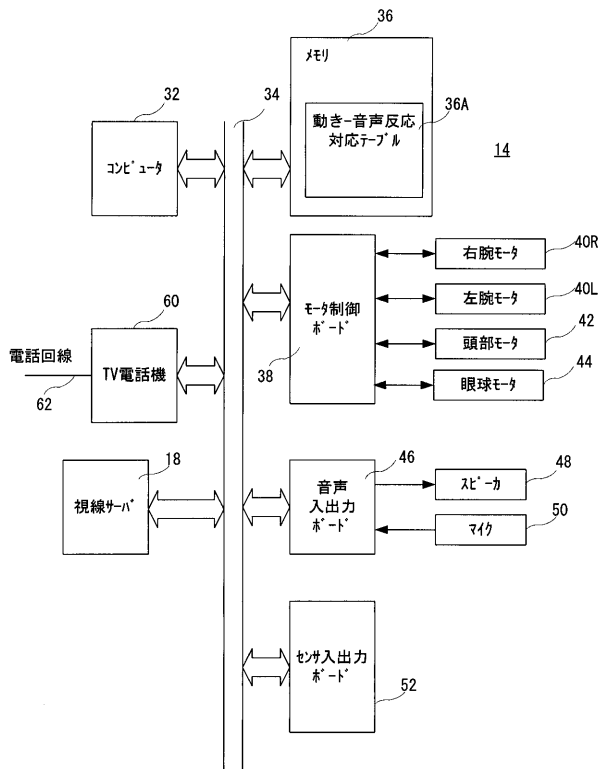
【図20】



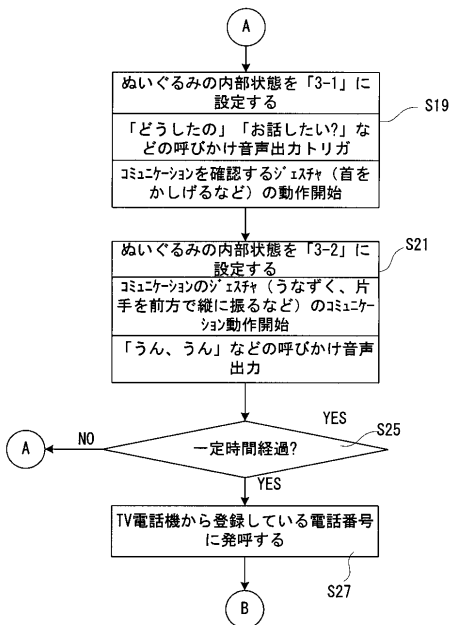
【図21】



【図22】



【図23】



【 図 6 】

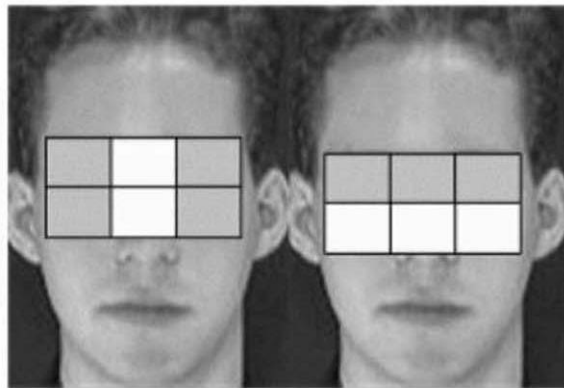


【 図 7 】



【 図 8 】

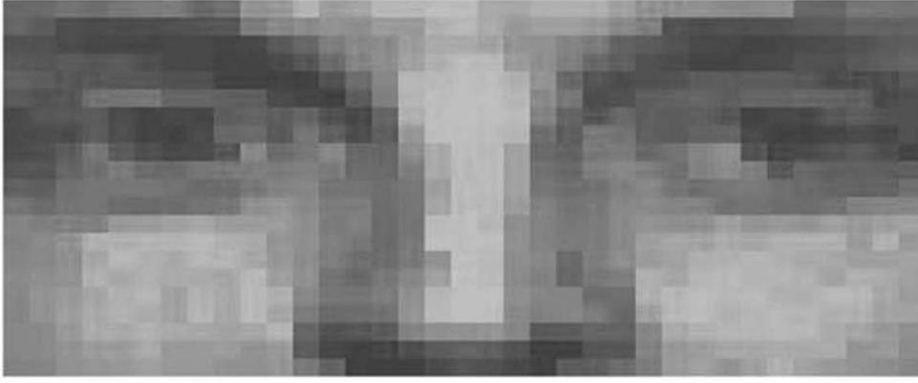
S1	S2	S3
S4	S5	S6



(a)

(b)

【図10】



【図11】



(a)

(b)

【図15】

キャリブレーション



(a)

(b)

(c)

(d)

【 図 1 7 】

右上方注視



(a)

(b)

(c)

(d)

【 図 1 8 】

上方注視



(a)

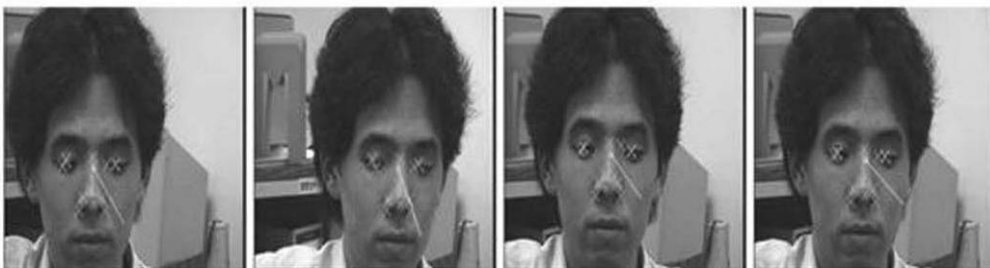
(b)

(c)

(d)

【 図 1 9 】

左下方注視



(a)

(b)

(c)

(d)

フロントページの続き

(72)発明者 安部 伸治

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 岩田 洋一

(56)参考文献 特開2005-238382(JP,A)

特開2004-177315(JP,A)

特開平10-234681(JP,A)

特開2005-130501(JP,A)

(58)調査した分野(Int.Cl., DB名)

A61M 21/00

A63B 69/00