

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第5407069号
(P5407069)

(45) 発行日 平成26年2月5日(2014.2.5)

(24) 登録日 平成25年11月15日(2013.11.15)

(51) Int.Cl.	F I
H04N 7/14 (2006.01)	H04N 7/14
H04M 1/00 (2006.01)	H04M 1/00 R
B25J 13/00 (2006.01)	B25J 13/00 Z

請求項の数 14 (全 49 頁)

(21) 出願番号 特願2009-252276 (P2009-252276)
 (22) 出願日 平成21年11月2日(2009.11.2)
 (65) 公開番号 特開2011-97531 (P2011-97531A)
 (43) 公開日 平成23年5月12日(2011.5.12)
 審査請求日 平成24年9月28日(2012.9.28)

(出願人による申告)平成21年6月19日付け、支出
 負担行為担当官 総務省大臣官房会計課企画官、研究テ
 ーマ「高齢者・障害者のためのユビキタスネットワーク
 ロボット技術の研究開発」に関する委託研究、産業技術
 力強化法第19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 100090181
 弁理士 山田 義人
 (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
 (72) 発明者 神山 祐一
 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 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

最終頁に続く

(54) 【発明の名称】 傾聴対話持続システム

(57) 【特許請求の範囲】

【請求項1】

第1カメラおよびマイクを含むテレビ電話機およびロボットを含む、傾聴対話持続システムであって、

前記第1カメラによって撮影された画像および前記マイクによって集音された音声に基づいて、ユーザの行動を判定する判定手段、

前記判定手段によって判定された第1所定時間分の行動から、前記ユーザの状態を認識する認識手段、および

前記認識手段によって認識された前記ユーザの状態に基づいて、対話を持続させるように前記ロボットを動作させる動作付与手段を備える、傾聴対話持続システム。

【請求項2】

前記認識手段は、対話に対する積極的および非積極的を認識する積極性認識手段を含み、

前記動作付与手段は、前記積極性認識手段の認識結果に基づき、前記ユーザに対する疑似傾聴を前記ロボットに行わせる傾聴動作付与手段を含む、請求項1記載の傾聴対話持続システム。

【請求項3】

前記認識手段は、対話における聴取側状態および話手側状態を認識する話者状態認識手段をさらに含み、

前記ユーザまたは相手の発声時間を計測する計測手段をさらに備え、

10

20

前記動作付与手段は、前記計測手段によって計測された発声時間に基づき、前記ユーザの発話を制御するように、前記ロボットを動作させる発話制御動作付与手段をさらに含む、請求項 2 記載の傾聴対話持続システム。

【請求項 4】

前記発話制御動作付与手段は、前記積極性認識手段によって積極的と認識され、かつ話者状態認識手段によって話手側状態と認識される積極的発話状態であり、かつ発声時間が閾値以上になったとき、前記ユーザの発話が抑制されるように、前記ロボットを動作させる発話抑制動作付与手段を含む、請求項 3 記載の傾聴対話持続システム。

【請求項 5】

前記発話抑制動作付与手段は、前記ユーザの注意を誘導するように、前記ロボットを動作させる注意誘導手段を含む、請求項 4 記載の傾聴対話持続システム。

10

【請求項 6】

前記発話制御動作付与手段は、前記積極性認識手段によって積極的と認識され、かつ話者状態認識手段によって聴取側状態と認識される積極的聴取側状態であるとき、前記ユーザの発話が促進されるように、前記ロボットを動作させる発話促進動作付与手段をさらに含む、請求項 3 ないし 5 のいずれかに記載の傾聴対話持続システム。

【請求項 7】

前記動作付与手段は、前記積極性認識手段によって非積極的と認識されたとき、前記ユーザを対話に参加させるように、前記ロボットを動作させる参加動作付与手段をさらに含む、請求項 2 ないし 5 のいずれかに記載の傾聴対話持続システム。

20

【請求項 8】

前記参加動作付与手段は、前記ユーザの注意を引きつけるように前記ロボットを動作させる注意引付手段を含む、請求項 7 記載の傾聴対話持続システム。

【請求項 9】

前記参加動作付与手段は、前記ユーザの発話を促すように前記ロボットを動作させる、発話促し手段をさらに含む、請求項 7 または 8 記載の傾聴対話持続システム。

【請求項 10】

前記ロボットが接続されるネットワーク、
前記ネットワークに接続されるサーバ、
前記判定手段によって判定されたユーザの行動を前記サーバに送信する送信手段、および

30

前記サーバから相手ユーザの行動を取得する取得手段をさらに備え、

前記動作付与手段は、前記取得手段によって取得された相手ユーザの行動とユーザの行動とに基づき、前記相手ユーザに対する疑似傾聴を前記ロボットに行わせる相手傾聴動作付与手段をさらに含む、請求項 1 ないし 9 のいずれかに記載の傾聴対話持続システム。

【請求項 11】

前記認識手段は、前記ユーザの興味対象を認識する興味対象認識手段をさらに含み、

前記動作付与手段は、前記興味対象認識手段による認識結果に基づき、前記ユーザが対話に興味を持つように、前記ロボットを動作させる興味動作付与手段をさらに含む、請求項 1 ないし 10 のいずれかに記載の対話傾聴持続システム。

40

【請求項 12】

前記ロボットは、第 2 カメラを含み、

前記第 1 カメラおよび第 2 カメラによるそれぞれの画像に対して顔認識処理を実行する顔認識手段をさらに備え、

前記興味対象認識手段は、前記顔認識手段による第 2 所定時間分の顔認識結果から第 1 カメラの第 1 認識率を算出する第 1 認識率算出手段、前記顔認識手段による第 2 所定時間分の顔認識結果から第 2 カメラの第 2 認識率を算出する第 2 認識率算出手段および前記第 1 認識率と前記第 2 認識率とに基づいて認識結果を設定する設定手段を含む、請求項 11 記載の対話傾聴持続システム。

【請求項 13】

50

前記設定手段によって設定された興味対象に基づいて前記第 1 カメラまたは前記第 2 カメラの画像のどちらか一方を送信する画像送信手段をさらに備える、請求項 1 2 記載の対話傾聴持続システム。

【請求項 1 4】

前記判定手段は、前記ユーザの姿勢を判定する姿勢判定手段、前記ユーザの発話の有無を判定する発話判定手段、前記ユーザの頭部方向を判定する頭部方向判定手段、前記ユーザの視線方向を判定する視線方向判定手段、前記ユーザの頷きの有無を判定する頷き判定手段および前記ユーザの相槌を判定する相槌判定手段を含み、

前記ユーザの行動は、姿勢判定結果、発話判定結果、頭部方向判定結果、視線方向判定結果、頷き判定結果および相槌判定結果に基づいて決定される、請求項 1 ないし 1 3 のいずれかに記載の対話傾聴持続システム。

10

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、傾聴対話持続システムに関し、特にたとえば、テレビ電話を利用して対話を行う、傾聴対話システムに関する。

【背景技術】

【0002】

特許文献 1 に開示されている、対話型アノテーションシステムでは、認知症患者である被介護者が利用するテレビ電話が、ネットワークを介して介護者が利用するコンピュータと接続される。ネットワークにはコンピュータおよびテレビ電話の間で、コマンドの転送を行うサーバが接続され、さらにサーバには写真の画像データが蓄えられるデータベースが結合される。また、介護者は、コンピュータによってサーバの URL を開いて写真を指定することで、相手（被介護者）が利用するテレビ電話に表示させることができる。これにより、介護の一環として、テレビ電話を利用して介護者と被介護者とが会話をすることができる。

20

【特許文献 1】特開 2 0 0 7 - 1 5 0 9 5 5 号公報 [H04N 7/14]

【発明の概要】

【発明が解決しようとする課題】

【0003】

30

近年、軽度認知症患者や、高齢者および言語障害のある人の話を聴き、情緒的なサポートを行う傾聴活動が注目されている。そして、傾聴活動の一環として、傾聴ボランティアが特許文献 1 の介護者として、軽度認知症患者などの被介護者と会話するようになっている。しかし、傾聴ボランティアの人数は、話し相手を求める介護者の人数に対して少ない。そこで、軽度認知症患者などは対話を行うことも可能であるため、話し相手を求める軽度認知症患者同士で対話を行う活動が進められている。

【0004】

ところが、軽度認知症患者同士の対話では、軽度認知症患者が対話に対して集中力を欠如してしまい、対話が持続しないことが多々ある。

【0005】

40

それゆえに、この発明の主たる目的は、新規な、傾聴対話持続システムを提供することである。

【0006】

この発明の他の目的は、コミュニケーション障害のある者同士の対話を持続させる、傾聴対話持続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】

【0007】

この発明は、上記の課題を解決するために、以下の構成を採用した。なお、括弧内の参照符号および補足説明等は、この発明の理解を助けるために記述する実施形態との対応関係を示したものであって、この発明を何ら限定するものではない。

50

【 0 0 0 8 】

第 1 の発明は、第 1 カメラおよびマイクを含むテレビ電話機およびロボットを含む、傾聴対話持続システムであって、第 1 カメラによって撮影された画像およびマイクによって集音された音声に基づいて、ユーザの行動を判定する判定手段、判定手段によって判定された第 1 所定時間分の行動から、ユーザの状態を認識する認識手段、および認識手段によって認識されたユーザの状態に基づいて、対話を持続させるようにロボットを動作させる動作付与手段を備える、傾聴対話持続システムである。

【 0 0 0 9 】

第 1 の発明では、傾聴対話持続システム (1 0 0) は、モニタ (1 6 a , 1 6 b) の近くに設けられる第 1 カメラ (2 2 a , 2 2 b) およびユーザが発話する音声を集音するマイク (2 0 a , 2 0 b) と接続されるテレビ電話機 (1 4 a , 1 4 b) と、ぬいぐるみ型のロボット (1 0 a , 1 0 b , 1 4 a , 1 4 b) とを含む。

10

【 0 0 1 0 】

判定手段 (6 4 , S 6 5 , S 6 7) は、第 1 カメラによって撮影された画像に対して、所定のテンプレートマッチング処理や顔認識処理を加えた結果と、マイクによって集音された音声の音声レベルとに基づいて、ユーザの行動を判定する。認識手段 (6 4 , S 7 5 - S 8 1) は、たとえば 3 0 秒である第 1 所定時間分の行動に基づいて、ユーザの状態を認識する。動作付与手段 (6 4 , S 1 7 1 , S 1 7 3 , S 1 7 7 , S 1 7 9) は、ユーザの状態に基づいて、たとえばユーザの発話を疑似傾聴したり、ユーザの発話を調節したり、ユーザの注意を引きつけたりするようにロボットを動作させることで、対話を持続させる。

20

【 0 0 1 1 】

第 1 の発明によれば、ユーザの状態に応じてロボットが対話を持続させるように動作するため、コミュニケーション障害のある者同士の対話を持続させることができる。

【 0 0 1 2 】

第 2 の発明は、第 1 の発明に従属し、認識手段は、対話に対する積極的および非積極的を認識する積極性認識手段を含み、動作付与手段は、積極性認識手段の認識結果に基づき、ユーザに対する疑似傾聴をロボットに行わせる傾聴動作付与手段を含む。

【 0 0 1 3 】

第 2 の発明では、積極性認識手段 (6 4 , S 7 5) は、たとえばユーザが相手を見る時間 (W T) 、前傾視線の時間 (F T) および相槌の頻度 (R F) に基づいて、対話に対する積極的 (アクティブ状態) および非積極的 (パッシブ状態) を認識する。傾聴動作付与手段 (6 4 , S 1 7 1 , S 1 7 3 , S 2 0 1 , S 2 2 5 , S 2 6 5 , S 2 8 5 , S 3 0 5 , S 3 6 7 , S 4 3 5 , S 4 4 1 , S 4 4 3) は、たとえばユーザが対話に対して積極的である場合に、ユーザの発話を疑似傾聴するようにロボットを動作させる。

30

【 0 0 1 4 】

第 2 の発明によれば、ユーザは、ロボットの疑似傾聴によって、自身の話を聴いてもらっているように感じることが出来るため、対話が持続するようになる。

【 0 0 1 5 】

第 3 の発明は、第 2 の発明に従属し、認識手段は、対話における聴取側状態および話手側状態を認識する話者状態認識手段をさらに含み、ユーザまたは相手の発声時間を計測する計測手段をさらに備え、動作付与手段は、計測手段によって計測された発声時間に基づき、ユーザの発話を制御するように、ロボットを動作させる発話制御動作付与手段をさらに含む。

40

【 0 0 1 6 】

第 3 の発明では、話者状態認識手段 (6 4 , S 7 7) は、ユーザの発話の有無に基づいて、対話における傾聴 (リッスン状態) および発話 (トーク状態) を認識する。計測手段 (6 4 , S 1 4 1 , S 1 4 3 , S 1 4 5) は、たとえば音声レベルが所定値以上であるとき、ユーザまたは相手の発声時間を計測する。発話制御動作付与手段 (6 4 , S 2 0 7 , S 2 2 7 , S 2 9 7 , S 3 1 9) は、たとえばとちらかのユーザが一方的に話している状

50

態で発声時間が閾値以上になったとき、ユーザの発話を制御するように、ロボットを動作させる。

【 0 0 1 7 】

第 3 の発明によれば、発話が制御されることでお互いがバランス良く発話するため、対話が持続するようになる。

【 0 0 1 8 】

第 4 の発明は、第 3 の発明に従属し、発話制御動作付与手段は、積極性認識手段によって積極的と認識され、かつ話者状態認識手段によって話手状態と認識される積極的発話状態であり、かつ発声時間が閾値以上になったとき、ユーザの発話が抑制されるように、ロボットを動作させる発話抑制動作付与手段を含む。

10

【 0 0 1 9 】

第 4 の発明では、発話抑制動作付与手段（ 6 4 , S 2 9 7 , S 4 0 5 , S 4 1 1 , S 4 1 7 , S 4 2 3 , S 4 2 5 ）は、ユーザが対話に対して積極的に発話している場合に、発声時間が閾値以上になると、ユーザの発話を抑制するように、ロボットを動作させる。

【 0 0 2 0 】

第 4 の発明によれば、ユーザが一方的に話しすぎている場合には、発話が抑制される。

【 0 0 2 1 】

第 5 の発明は、第 4 の発明に従属し、発話抑制動作付与手段は、ユーザの注意を誘導するように、ロボットを動作させる注意誘導手段を含む。

【 0 0 2 2 】

20

第 5 の発明では、注意誘導手段（ 6 4 , S 4 2 5 , S 4 5 5 , S 4 6 1 ）は、たとえばユーザと相手が表示されるモニタ（ 1 6 ）とを交互に見るように、ロボットを動作させることでユーザの注意を誘導する。

【 0 0 2 3 】

第 5 の発明によれば、一方的に話すユーザの注意が誘導されるため、発話が抑制される。また、ユーザの注意が相手に誘導される場合には、相手が発話する機会を得ることができ、対話がより持続するようになる。

【 0 0 2 4 】

第 6 の発明は、第 3 の発明ないし第 5 の発明のいずれかに従属し、発話制御動作付与手段は、積極性認識手段によって積極的と認識され、かつ話者状態認識手段によって聴取側状態と認識される積極的聴取側状態であるとき、ユーザの発話が促進されるように、ロボットを動作させる発話促進動作付与手段をさらに含む。

30

【 0 0 2 5 】

第 6 の発明では、発話促進動作付与手段（ 6 4 , S 2 2 7 , S 3 1 9 , S 4 7 1 ）は、ユーザが積極的に相手の話を聴き取っている（傾聴している）場合に、たとえば相手の発声時間が閾値以上になれば、ユーザの発話が促進されるように、ロボットを動作させる。

【 0 0 2 6 】

第 6 の発明によれば、ユーザが相手の話を積極的に傾聴している状態が長く続けば、発話が促進される。

【 0 0 2 7 】

40

第 7 の発明は、第 2 の発明ないし第 5 の発明のいずれかに従属し、動作付与手段は、積極性認識手段によって非積極的と認識されたとき、ユーザを対話に参加させるように、ロボットを動作させる参加動作付与手段をさらに含む。

【 0 0 2 8 】

第 7 の発明では、参加動作付与手段（ 6 4 , S 1 7 7 , S 2 3 1 , S 3 1 9 , S 3 2 1 ）は、ユーザが対話に非積極的であれば、ユーザを対話に参加させるように、ロボットを動作させる。

【 0 0 2 9 】

第 7 の発明によれば、対話に非積極的なユーザを対話に参加させることで、対話を持続させる。

50

【 0 0 3 0 】

第 8 の発明は、第 7 の発明に従属し、参加動作付与手段は、ユーザの注意を引きつけるようにロボットを動作させる注意引付手段を含む。

【 0 0 3 1 】

第 8 の発明では、注意引付手段（ 6 4 , S 3 2 1 , S 3 3 5 , S 3 4 1 , S 3 4 3 ）は、たとえば、ユーザを見たり、話しかけたりするようにロボットを動作させることで、ユーザの注意を引きつける。

【 0 0 3 2 】

第 8 の発明によれば、対話に非積極的なユーザの注意を引きつけることで、対話に参加させる。

10

【 0 0 3 3 】

第 9 の発明は、第 7 の発明または第 8 の発明に従属し、参加動作付与手段は、ユーザの発話を促すようにロボットを動作させる、発話促し手段をさらに含む。

【 0 0 3 4 】

第 9 の発明では、発話促し手段（ 6 4 , S 3 1 9 , S 4 7 1 ）は、たとえばユーザに対して質問をするようにロボットを動作させることで、ユーザの発話を促す。

【 0 0 3 5 】

第 9 の発明によれば、対話に非積極的なユーザに発話させることで、対話に参加させる。

【 0 0 3 6 】

第 1 0 の発明は、第 1 の発明ないし第 9 の発明のいずれかに従属し、ロボットが接続されるネットワーク、ネットワークに接続されるサーバ、判定手段によって判定されたユーザの行動をサーバに送信する送信手段、およびサーバから相手ユーザの行動を取得する取得手段をさらに備え、動作付与手段は、取得手段によって取得された相手ユーザの行動とユーザの行動とに基づき、相手ユーザに対する疑似傾聴をロボットに行わせる相手傾聴動作付与手段をさらに含む。

20

【 0 0 3 7 】

第 1 0 の発明では、ネットワーク（ 2 0 0 ）には、たとえば無線 LAN などによってロボットおよびサーバ（ 2 4 ）が接続される。送信手段（ 6 4 , S 6 9 ）は、ユーザの行動を一定時間毎にサーバに送信する。取得手段（ 6 4 , S 2 6 9 , S 2 8 9 , S 3 0 9 ）は、サーバに送信された相手ユーザの行動を取得する。相手傾聴動作付与手段（ 6 4 , S 2 7 5 , S 2 9 5 , S 3 1 5 , S 3 8 5 , S 3 9 1 , S 3 9 3 ）は、相手ユーザが発話していれば、たとえば相手ユーザが表示されるモニタに対して疑似傾聴動作を行うように、ロボットを動作させる。

30

【 0 0 3 8 】

第 1 0 の発明によれば、たとえユーザが相手ユーザの話を傾聴していなくても、ロボットが相手の話を傾聴することで、相手ユーザは自身の話を聴いてもらっているように感じることができる。

【 0 0 3 9 】

第 1 1 の発明は、第 1 の発明ないし第 1 0 の発明のいずれかに従属し、認識手段は、ユーザの興味対象を認識する興味対象認識手段をさらに含み、動作付与手段は、興味対象認識手段による認識結果に基づき、ユーザが対話に興味を持つように、ロボットを動作させる興味動作付与手段をさらに含む。

40

【 0 0 4 0 】

第 1 1 の発明では、興味対象認識手段（ 6 4 , S 7 9 ）は、ユーザが見ているものをユーザの興味対象として認識する。興味動作付与手段（ 6 4 , S 1 7 9 , S 2 4 1 , S 3 2 1 , S 3 3 5 , S 3 4 1 , S 3 4 3 , S 4 0 5 ）は、ユーザの興味が、たとえばロボットでなければ、ユーザが対話に興味を持つように、ロボットを動作させる。

【 0 0 4 1 】

第 1 1 の発明によれば、ユーザに対話に興味を持たせることで、対話が持続するように

50

なる。

【 0 0 4 2 】

第 1 2 の発明は、第 1 1 の発明に従属し、ロボットは、第 2 カメラを含み、第 1 カメラおよび第 2 カメラによるそれぞれの画像に対して顔認識処理を実行する顔認識手段をさらに備え、興味対象認識手段は、顔認識手段による第 2 所定時間分の顔認識結果から第 1 カメラの第 1 認識率を算出する第 1 認識率算出手段、顔認識手段による第 2 所定時間分の顔認識結果から第 2 カメラの第 2 認識率を算出する第 2 認識率算出手段および第 1 認識率と第 2 認識率とに基づいて認識結果を設定する設定手段を含む。

【 0 0 4 3 】

第 1 2 の発明では、ロボットは、たとえば腹部に第 2 カメラ (1 2) が設けられる。顔認識手段 (6 4 , S 4 1) は、第 1 カメラおよび第 2 カメラによって撮影されたユーザの顔を認識する。第 1 認識率算出手段 (6 4 , S 1 2 3) は、第 1 カメラによって撮影されたユーザの顔の第 1 認識率 (M) を算出する。第 2 認識率算出手段 (6 4 , S 1 2 5) は、第 2 カメラによって撮影されたユーザの顔の第 2 認識率 (S) を算出する。そして、設定手段 (6 4 , S 1 3 1 , S 1 3 3 , S 1 3 5) は、第 1 認識率および第 2 認識率に基づいて、ユーザの興味対象を設定する。

【 0 0 4 4 】

第 1 2 の発明によれば、興味対象がユーザの顔の認識率 M , S に基づいて設定されるため、ユーザが見ている対象を正確に認識することができる。

【 0 0 4 5 】

第 1 3 の発明は、第 1 2 の発明に従属し、設定手段によって設定された興味対象に基づいて第 1 カメラまたは第 2 カメラの画像のどちらか一方を送信する画像送信手段をさらに備える。

【 0 0 4 6 】

第 1 3 の発明では、画像送信手段 (6 4 , S 4 8 9 , S 4 9 1) は、たとえばユーザの興味対象がテレビ電話機であれば第 1 カメラの画像を送信し、ロボットであれば第 2 カメラの画像を送信する。

【 0 0 4 7 】

第 1 3 の発明によれば、相手は、たとえユーザがロボットに話しかけている状態であったとしても、相手にはユーザの顔が正面に写る画像が送られる。そのため、相手は自身に話しかけられているように感じるため、対話が持続するようになる。

【 0 0 4 8 】

第 1 4 の発明は、第 1 の発明ないし第 1 3 の発明のいずれかに従属し、判定手段は、ユーザの姿勢を判定する姿勢判定手段、ユーザの発話の有無を判定する発話判定手段、ユーザの頭部方向を判定する頭部方向判定手段、ユーザの視線方向を判定する視線方向判定手段、ユーザの頷きの有無を判定する頷き判定手段およびユーザの相槌を判定する相槌判定手段を含み、ユーザの行動は、姿勢判定結果、発話判定結果、頭部方向判定結果、視線方向判定結果、頷き判定結果および相槌判定結果に基づいて決定される。

【 0 0 4 9 】

第 1 4 の発明では、姿勢判定手段 (6 4 , S 9)、頭部方向判定手段 (6 4 , S 2 5)、視線方向判定手段 (6 4 , S 3 3) および頷き判定手段 (6 4 , S 4 9) は、ユーザが写る画像に基づいて判定する。発話判定手段 (6 4 , S 1 7) はユーザの音声に基づいて判定する。さらに、相槌判定手段 (6 4 , S 5 9) は、ユーザの発話と頷きの判定結果に基づいて判定する。そして、ユーザの行動は、各判定結果に基づいて決められる。

【 0 0 5 0 】

第 1 4 の発明によれば、ユーザの行動データに複数の判定結果が含まれることで、ユーザの状態を適確に認識できるようになる。

【 発明の効果 】

【 0 0 5 1 】

この発明によれば、ユーザの状態に応じてロボットが対話を持続させるように動作する

10

20

30

40

50

ため、コミュニケーション障害のある者同士の対話を持続させることができる。

【 0 0 5 2 】

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。

【図面の簡単な説明】

【 0 0 5 3 】

【図 1】図 1 はこの発明の一実施例の傾聴対話持続システムの概要を示す図解図である。

【図 2】図 2 は図 1 に示すモニタカメラとモニタとロボットとユーザとの側面的な位置関係およびそのモニタカメラと腹部カメラとの撮影範囲の一例を示す図解図である。

【図 3】図 3 は図 1 に示すロボットの外観を正面から見た図解図である。

10

【図 4】図 4 は図 1 に示すロボットの電氣的な構成の一例を示すブロック図である。

【図 5】図 5 は図 1 に示す P C の電氣的な構成の一例を示すブロック図である。

【図 6】図 6 は図 4 に示すメモリに記憶される行動テーブルの一例を示す図解図である。

【図 7】図 7 は図 1 に示すモニタカメラおよび腹部カメラによる顔認識結果の一例を示す図解図である。

【図 8】図 8 は図 1 に示すサーバの電氣的な構成の一例を示すブロック図である。

【図 9】図 9 は図 4 に示す P C のメモリのメモリマップの一例を示す図解図である。

【図 1 0】図 1 0 は図 9 に示すデータ記憶領域の一例を示す図解図である。

【図 1 1】図 1 1 は図 9 に示す状況認識プログラムの構成の一例を示す図解図である。

【図 1 2】図 1 2 は図 9 に示すロボット制御プログラムの構成の一例を示す図解図である

20

。

【図 1 3】図 1 3 は図 5 に示す P C のプロセッサの画像 / 音声取得処理を示すフロー図である。

【図 1 4】図 1 4 は図 5 に示す P C のプロセッサの姿勢判定処理を示すフロー図である。

【図 1 5】図 1 5 は図 5 に示す P C のプロセッサの発話判定処理を示すフロー図である。

【図 1 6】図 1 6 は図 5 に示す P C のプロセッサの頭部方向判定処理を示すフロー図である。

【図 1 7】図 1 7 は図 5 に示す P C のプロセッサの視線方向判定処理を示すフロー図である。

【図 1 8】図 1 8 は図 5 に示す P C のプロセッサの顔認識処理を示すフロー図である。

30

【図 1 9】図 1 9 は図 5 に示す P C のプロセッサの頷き判定処理を示すフロー図である。

【図 2 0】図 2 0 は図 5 に示す P C のプロセッサの相槌判定処理を示すフロー図である。

【図 2 1】図 2 1 は図 5 に示す P C のプロセッサの同期処理を示すフロー図である。

【図 2 2】図 2 2 は図 5 に示す P C のプロセッサの状態認識処理を示すフロー図である。

【図 2 3】図 2 3 は図 5 に示す P C のプロセッサの A c / P a 認識処理を示すフロー図である。

【図 2 4】図 2 4 は図 5 に示す P C のプロセッサの T a / L i 認識処理を示すフロー図である。

【図 2 5】図 2 5 は図 5 に示す P C のプロセッサの興味対象認識処理を示すフロー図である。

40

【図 2 6】図 2 6 は図 5 に示す P C のプロセッサの発声時間計測処理を示すフロー図である。

【図 2 7】図 2 7 は図 5 に示す P C のプロセッサの全体処理を示すフロー図である。

【図 2 8】図 2 8 は図 5 に示す P C のプロセッサのアクティブトーク処理を示すフロー図である。

【図 2 9】図 2 9 は図 5 に示す P C のプロセッサのアクティブリスン処理を示すフロー図である。

【図 3 0】図 3 0 は図 5 に示す P C のプロセッサの非アクティブ処理を示すフロー図である。

【図 3 1】図 3 1 は図 5 に示す P C のプロセッサのアザー処理を示すフロー図である。

50

【図 3 2】図 3 2 は図 5 に示す P C の発話継続処理を示すフロー図である。

【図 3 3】図 3 3 は図 5 に示す P C のプロセッサの発話抑制処理を示すフロー図である。

【図 3 4】図 3 4 は図 5 に示す P C のプロセッサの発話促進処理を示すフロー図である。

【図 3 5】図 3 5 は図 5 に示す P C のプロセッサの注意引きつけ処理を示すフロー図である。

【図 3 6】図 3 6 は図 5 に示す P C のプロセッサの傍参与者的疑似傾聴処理を示すフロー図である。

【図 3 7】図 3 7 は図 5 に示す P C のプロセッサの第 1 積極的疑似傾聴処理を示すフロー図である。

【図 3 8】図 3 8 は図 5 に示す P C のプロセッサのユーザ発話抑制処理を示すフロー図である。

10

【図 3 9】図 3 9 は図 5 に示す P C のプロセッサの第 2 積極的疑似傾聴処理を示すフロー図である。

【図 4 0】図 4 0 は図 5 に示す P C のプロセッサの注意誘導処理を示すフロー図である。

【図 4 1】図 4 1 は図 5 に示す P C のプロセッサのユーザ発話促進処理を示すフロー図である。

【図 4 2】図 4 2 は図 5 に示す P C のプロセッサのカメラ制御処理を示すフロー図である。

【発明を実施するための形態】

【 0 0 5 4 】

20

図 1 を参照して、この実施例の傾聴対話持続システム 1 0 0 は、たとえば認知症患者のような軽度脳障害を持つユーザ A と遠隔地に居るユーザ B との対話に利用される。そのため、傾聴対話持続システム 1 0 0 には、ユーザ A が居る部屋 1 に設置される腹部カメラ 1 2 a (第 2 カメラ) を含むぬいぐるみ型ロボット (以下、単に「ロボット」と言う。) 1 0 a、P C 1 4 a、モニタ 1 6 a、スピーカ 1 8 a、マイク 2 0 a およびモニタカメラ 2 2 a (第 1 カメラ) と、ユーザ B が居る部屋 2 (遠隔地) に設置される腹部カメラ 1 2 b を含むロボット 1 0 b、P C 1 4 b、モニタ 1 6 b、スピーカ 1 8 b、マイク 2 0 b およびモニタカメラ 2 2 b と、ネットワーク 2 0 0 に接続されるサーバ 2 4 とを備える。なお、本明細書では、部屋 1 および部屋 2 において対応する機器および人間を、区別なく説明する必要がある場合には、符号にアルファベットを添えずに参照番号だけで表示されることがあることに留意されたい。

30

【 0 0 5 5 】

ロボット 1 0 は P C 1 4 による制御信号に基づいて傾聴動作や、発話を行う。ロボット 1 0 の腹部に設けられた腹部カメラ 1 2 はユーザを撮影し、ロボット 1 0 を介して画像を P C 1 4 に出力する。P C 1 4 は、ロボット 1 0 に対して制御信号を出力するとともに、腹部カメラ 1 2 およびモニタカメラ 2 2 によって撮影された画像と、マイク 2 0 によって集音される音声とが入力される。そして、P C 1 4 は、入力された画像と音声とに基づいてユーザの行動および状態を判定および認識し、その結果をネットワーク 2 0 0 を介してサーバ 2 4 に送信する。

【 0 0 5 6 】

40

P C 1 4、モニタ 1 6、スピーカ 1 8、マイク 2 0 およびモニタカメラ 2 2 はテレビ電話機として機能する。たとえば、P C 1 4 a は、ユーザ B 側の P C 1 4 b から送信されたユーザ B の画像および音声を受信する。そのため、モニタ 1 6 a はユーザ B の画像を表示し、スピーカ 1 8 はユーザ B の音声を出力する。さらに、マイク 2 0 はユーザ A の音声を集音して P C 1 4 に出力し、モニタカメラ 2 2 はユーザ A の画像を撮影して P C 1 4 に出力する。そして、P C 1 4 は、ユーザ A の画像と音声とを、ネットワーク 2 0 0 を介して P C 1 4 b に送信する。

【 0 0 5 7 】

サーバ 2 4 は、P C 1 4 a および P C 1 4 b から送信される、ユーザ A およびユーザ B の行動や状態のデータを受信すると、データベース (D B) に蓄積させる。そして、P C

50

14から行動および状態のデータを取得する要求がある場合に、その要求に応じてデータをPC14に送信する。

【0058】

なお、他の実施例では、ロボット10とPC14とが有線接続ではなく、無線接続であってもよい。また、PC14およびサーバ24のネットワーク200との接続も、有線接続であってもよいし、無線接続であってもよい。

【0059】

図2は図1に示す実施例を側面から見た実施例である。図2から分かるように、モニタカメラ22はモニタ16の上に置かれ、ロボット12とモニタ16とは机の上に置かれる。ユーザは、机の上に置かれるモニタ16およびモニタカメラ22に対面する状態で、腹部カメラ12およびモニタカメラ22によって撮影される。さらに、ロボット10は、ユーザとモニタ16との間に配置されるため、モニタカメラ22はロボット10とユーザとを同時に撮影する。これにより、ロボット10は、ユーザAに対して疑似的な傾聴動作（疑似傾聴動作）を行ったり、ユーザBが表示されるモニタ16aに対して疑似傾聴動作を行ったりする。

【0060】

なお、ロボット10は、モニタカメラ22によって撮影され、かつユーザを撮影可能な位置であれば、机の上に置かれていなくてもよい。

【0061】

図3にはロボット10の外観が図示される。このロボット10は、頭部26とそれを支える胴体28とを含む。胴体28の上部（人間の肩に相当）の左右に左腕30Lおよび右腕30Rが設けられ、胴体28の腹部には腹部カメラ12が設けられる。この腹部カメラ12には、たとえばCCDやCMOSのような固体撮像素子を用いるカメラを採用することができる。また、頭部26には、前面に口32が配置され、その口32の上方には眼球34が設けられる。そして、頭部26の上部側面には耳36が取り付けられている。

【0062】

頭部26は、胴体28によって旋回・俯仰可能に支持され、眼球34も稼働的に保持されている。また、胴体28は、腰の部分を中心として左右方向に傾くことが可能である。さらに、口32にはスピーカ56（図4）が内蔵され、耳36にはマイク58（図4）が内蔵される。

【0063】

なお、マイク58を両方の耳36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。また、ロボット10の外見は、熊だけに限らず、他の動物や、人型であってもよい。

【0064】

図4にはロボット10の電氣的な構成を示すブロック図が示される。ロボット10には、マイクロコンピュータ或いはCPUとも呼ばれる、プロセッサ38が内蔵されており、通信路の一例であるバス40を介して、腹部カメラ12、メモリ42、モータ制御ボード44、音声入力/出力ボード54、センサ入力/出力ボード60およびI/O62に接続される。

【0065】

メモリ42は、図示しないROMやRAMが組み込まれており、ROMには主として、ロボット10による傾聴動作や、発話を行うためのプログラムや、発話を行う際にスピーカ56から出力される音声データなどが予め記憶されている。また、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。

【0066】

モータ制御ボード44は、たとえばDSP(Digital Signal Processor)で構成され、図3に示すロボット10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44は、プロセッサ38からの制御データを受け、右腕30R（図3）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4ではまとめて、「右腕モータ」として示す。）46Rの回転角度を調節する。また、モータ制御ボード44は、左腕30Lの3つのモータ（図4ではまとめて、「左腕モータ」として示す。）46Lの回転角度を調節する。モータ制御ボード44は、また、頭部26の旋回角や俯仰角を制御する3つのモータ（図4ではまとめて、「頭部モータ」として示す。）48の回転角度を調節する。モータ制御ボード44は、また、眼球34を動かす眼球モータ50および胴体28を傾ける腰モータ52も制御する。

【0067】

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もよい。

10

【0068】

スピーカ56には音声入出力ボード54を介して、プロセッサ38から合成音声データが与えられ、それに応じて、スピーカ56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58によって集音された音声は、音声入出力ボード54を介して、プロセッサ38に取り込まれる。

【0069】

センサ入力／出力ボード60は、モータ制御ボード44と同様に、DSPで構成され、腹部カメラ12からの信号を取り込んで、プロセッサ38に与える。腹部カメラ12からの映像信号が、必要に応じてセンサ入力／出力ボード60で所定の処理を施してからプロセッサ38に入力される。

20

【0070】

I/O62は、各々入力／出力の制御が可能なデジタルポートであり、出力ポートからは映像信号が出力され、PC14に与えられる。一方、PC14からは、制御信号が出力され、入力ポートに与えられる。

【0071】

図5にはPC14の電氣的な構成を示すブロック図が示される。PC14には、ロボット10と同様、マイクロコンピュータ或いはCPUとも呼ばれる、プロセッサ64が内蔵されており、バス66を介して、メモリ68、視線サーバ70、音声入力／出力ボード72およびI/O74に接続される。なお、プロセッサ64には、日時情報を出力するRTC (Real Time Clock) 64aが内蔵されている。

30

【0072】

メモリ68は、図示しないROM、RAMおよびHDDが組み込まれており、ROMには主として、電話機能を実現するためのプログラムや、後述のフローチャート（図13 - 図42）で表現されるプログラムが記憶される。また、RAMには主として、腹部カメラ12およびモニタカメラ22によって撮影された画像や、マイク20によって集音された音声などが一時的に記憶されるバッファなどが設定されている。そして、HDDには主として、ユーザの行動を判断した結果や、状態を認識した結果などが随時記憶される。

【0073】

視線サーバ70は、腹部カメラ12およびモニタカメラ22によって撮影されたユーザの顔の画像から、ユーザの視線方向や位置をリアルタイムで検出する。そして、プロセッサ20は、視線サーバ70が特定または検出したユーザの視線方向や位置を示すデータを、バス66を通して刻々受け取ることで、ユーザが注視する対象を判定する。

40

【0074】

なお、視線サーバ70によるユーザの視線方向や位置の検出方法については、本件出願人が先に出願し既に公開された、特開2008-113875号公報に開示されているので、ここでは記述を省略する。

【0075】

スピーカ18には、音声入力／出力ボード72を介して、プロセッサ64から相手ユーザの音声データが与えられ、それに応じて、スピーカ18からはそのデータに従った音声が出力される。そして、マイク20によって集音された相手ユーザの音声は、音声入力／

50

出力ボード 7 2 を介して、プロセッサ 6 4 に取り込まれる。

【 0 0 7 6 】

I / O 7 4 は、ロボット 1 0 の I / O 6 2 と同様に、各々入力 / 出力の制御が可能なデジタルポートであり、出力ポートからは、制御信号がロボット 1 0 に出力され、画像信号がモニタ 1 6 に出力される。また、ロボット 1 0 およびモニタカメラ 2 2 からは、映像信号が出力され、入力ポートに与えられる。なお、スピーカ 1 8 が出力する音声データおよびマイク 2 0 から入力される音声データは、I / O 7 4 を介して音声入力 / 出力ボード 7 2 に入出力されるようにされてもよい。

【 0 0 7 7 】

また、プロセッサ 6 4 は、バス 6 6 を介して通信 LAN ボード 7 6 に接続される。この通信 LAN ボード 7 6 は、たとえば DSP で構成され、プロセッサ 6 4 から与えられた送信データを無線通信装置 7 8 に与える。無線通信装置 7 8 は送信データを、ネットワーク 2 0 0 を介して外部のコンピュータ (サーバ 2 4 および相手の PC 1 4) に送信する。また、通信 LAN ボード 7 6 は、無線通信装置 7 8 を介してデータを受信し、受信したデータをプロセッサ 6 4 に与える。

10

【 0 0 7 8 】

たとえば、送信データとしては、テレビ電話機として必要なコマンド、画像データおよび音声データや、ユーザの行動を判定した結果およびユーザの状態を認識した結果であったりする。また、受信データとしては、テレビ電話機として得られる相手の画像データおよび音声データや、相手ユーザの行動を判定した結果および状態を認識した結果であったりする。

20

【 0 0 7 9 】

図 6 には PC 1 4 のメモリ 6 8 に記憶される、行動テーブルが示される。この行動テーブルとは、ユーザの行動が判定された結果が行動データにされ、その行動データが一定時間 (たとえば、1 秒) 毎に刻々と記録されるテーブルである。

【 0 0 8 0 】

図 6 を参照して、行動テーブルは、左側から「時刻」、「前傾姿勢」、「発話」、「頭部方向 (腹部)」、「視線方向 (腹部)」、「顔認識 (腹部)」、「頭部方向 (モニタ)」、「視線方向 (モニタ)」、「顔認識 (モニタ)」、「頷き」および「相槌」の列で構成されている。そして、各行動データは、「時刻」の列に同期して、各欄に記録される判定結果から構成される。

30

【 0 0 8 1 】

「時刻」の列に記録される数値は、RTC 6 4 a が出力する日時情報であり、たとえば「10:00:30」は「1 0 時 0 0 分 3 0 秒」を表す。

【 0 0 8 2 】

「前傾姿勢」の列には、ユーザがモニタ 1 6 に対して前傾姿勢を取っているか否かを示す判定結果が記録される。たとえば、「前傾姿勢」の列に「あり」が記録されていればユーザがモニタ 1 6 に対して前傾姿勢を取っており、「なし」が記録されていれば、ユーザがモニタ 1 6 に対して前傾姿勢を取っていないことを示す。そして、前傾姿勢の「あり」/「なし」は、腹部カメラ 1 2 およびモニタカメラ 2 2 によって撮影された画像に対して、テンプレートマッチング処理を加えて判定される。たとえば、ユーザの姿勢が前傾視線のテンプレートと一致する場合に前傾姿勢が「あり」と判定され、テンプレートと一致しない場合に前傾姿勢が「なし」と判定される。なお、ロボット 1 0 に超音波センサなどの距離を計測可能なセンサを取りつけて、ユーザの姿勢を判定するようにしてもよい。

40

【 0 0 8 3 】

「発話」の列には、ユーザが発話しているか否かを示す判定結果が記録される。たとえば、「発話」の列に「あり」が記録されていれば、ユーザが発話していることを示し、「なし」が記録されていれば、ユーザが発話していないことを示す。そして、発話の「あり」/「なし」は、マイク 2 0 によって集音された音声データの音声レベルから判定される。たとえば、音声データの音声レベルが決められた値以上であれば「あり」と判定され、

50

決められた値未満であれば「なし」と判定される。

【 0 0 8 4 】

「頭部方向（腹部）」および「頭部方向（モニタ）」の列には、ユーザの頭部が向いている方向に有る物が記録される。たとえば、「モニタ」が記録されていれば、ユーザの頭部が向く方向にはモニタ 1 6 が有ることを示し、「ロボット」が記録されていれば、ユーザの頭部が向く方向にはロボット 1 0 が有ることを示し、「アザー」が記録されていれば、ユーザの頭部が向く方向にはロボット 1 0 およびモニタ 1 6 が無いことを示す。そして、「頭部方向（腹部）」の列については、腹部モニタ 1 2 によって撮影された画像に対してテンプレートマッチング処理を加えて判定される。一方、「頭部方向（モニタ）」の列については、モニタカメラ 2 2 によって撮影された画像に対してテンプレートマッチング処理を加えて判定される。たとえば、ユーザの頭部の形がモニタ 1 6 の方向を向くテンプレートと一致していれば「モニタ」と判定され、ユーザの頭部の形がロボット 1 0 の方向を向くテンプレートと一致していれば「ロボット」と判定され、ユーザの頭部の形がいずれのテンプレートとも一致しなければ「アザー」と判定される。

10

【 0 0 8 5 】

「視線方向（腹部）」および「視線方向（モニタ）」の列には、ユーザの視線が向いている方向に有る物が記録される。また、「視線方向（腹部）」および「視線方向（モニタ）」の列には、「頭部方向（腹部）」の列と同様に、「モニタ」、「ロボット」および「アザー」が記録される。そして、「視線方向（腹部）」および「視線方向（モニタ）」の列における、「モニタ」、「ロボット」および「アザー」は視線サーバ 7 0 の出力に基づいて判定される。たとえば、ユーザの視線方向がモニタ 1 6 に向いていれば「モニタ」と判定され、ユーザの視線方向がロボット 1 0 に向いていれば「ロボット」と判定され、ユーザの視線方向がロボット 1 0 およびモニタ 1 6 のいずれの方向にも向いていなければ、「アザー」と判定される。

20

【 0 0 8 6 】

「顔認識（腹部）」および「顔認識（モニタ）」の列には、ユーザの顔の認識結果が記録される。たとえば、ユーザの顔が認識されていれば「成功」が記録され、ユーザの顔が認識されていなければ「失敗」が記録される。そして、「顔認識（腹部）」の列については、腹部モニタ 1 2 によって撮影された画像に対して顔認識処理が実施されることで判定される。一方、「顔認識（モニタ）」の列については、モニタカメラ 2 2 によって撮影された画像に対して所定の顔認識処理が実施されることで判定される。

30

【 0 0 8 7 】

図 7 (A) にはモニタカメラ 2 2 による顔認識結果の成功列が示され、図 7 (B) には腹部カメラ 1 2 による顔認識結果の成功例が示される。まず、図 7 (A) を参照して、左側が腹部カメラ 1 2 による画像であり、右側がモニタカメラ 2 2 による画像であり、どちらの画像も同じ時刻に撮影された画像である。このとき、ユーザはモニタカメラ 2 2 を注視している状態である。そのため、モニタカメラ 2 2 による画像では、ユーザの顔が正面に写っているため、顔認識が成功している。一方、腹部カメラ 1 2 による画像では、ユーザの顔は傾いて写っているため、顔認識が失敗している。

40

【 0 0 8 8 】

次に、図 7 (B) を参照して、図 7 (A) と同様に、左側が腹部カメラ 1 2 による画像であり、右側がモニタカメラ 2 2 による画像であり、同じ時刻に撮影された画像である。このとき、ユーザは腹部カメラ 1 2 を注視しているため、腹部カメラ 1 2 による画像では顔認識が成功している。一方、モニタカメラ 2 2 による画像では顔認識が失敗している。

【 0 0 8 9 】

図 6 に戻って、「頷き」の列には、ユーザによる頷きの有無が記録される。たとえば、「頷き」の列に「あり」が記録されていればユーザが頷いたことを示し、「なし」が記録されていればユーザが頷かなかったことを示す。そして、ユーザの頷きの有無は、腹部カメラ 1 2 およびモニタカメラ 2 2 によって撮影された画像に対して、所定のテンプレートマッチング処理を加えることで判定する。たとえば、ユーザの頭部の形が頷いている状態

50

のテンプレートと一致していれば「あり」と判定され、テンプレートと一致していなければ「なし」と判定される。

【0090】

「相槌」の列には、ユーザによる相槌の有無が記録される。たとえば、「相槌」の列に「あり」と記録されていればユーザが頷いたことを示し、「なし」と記録されていればユーザが頷かなかったことを示す。そして、相槌の有無については、頷きの判定結果と、音声データにおける「あー」や「うん」などの相槌らしい音を判定可能な、音声プロソディの判定結果とに基づいて判断する。たとえば、ユーザの頷きが「あり」ときに相槌らしい音があれば、相槌の判定結果は「あり」と判定される。

【0091】

なお、上述したテンプレートマッチング処理および顔認識処理は、広く一般的な手法が用いられているため、詳細な説明は省略する。

【0092】

そして、PC14では、これらの処理によって一定時間毎に記録された行動データから、ユーザの状態を第1所定時間（たとえば、30秒）毎に認識する。具体的には、第1所定時間毎の行動データに基づいて、Ac/Pa認識、Ta/Li認識および興味対象認識の処理を実行し、各処理の認識結果からユーザの状態を決定する。

【0093】

まず、Ac/Pa認識とは、ユーザが対話に積極的に参加し、対話に集中している「アクティブ：Active（Ac）」状態か、ユーザが対話に非積極的であり、他のものに集中または注意力が散漫な「パッシブ：Passive（Pa）」状態かを認識する処理である。そのため、本実施例のAc/Pa認識では、第1所定時間分のユーザの行動データのうち、視線方向、前傾姿勢、頷きおよび相槌に基づいて、アクティブ状態らしさを示すAC値を算出することで、アクティブ状態またはパッシブ状態を認識する。

【0094】

たとえば、相手を見る時間をWT、体を倒して近づく前傾姿勢の時間をAT、相槌の頻度をRFで示す場合に、相手を見る時間WTは、行動テーブルにおける「視線方向（腹部）」および「視線方向（モニタ）」の列で、第1所定時間分の行のうち、「あり」と判定された回数をカウントすることで求めることができる。また、前傾姿勢の時間FTは、「前傾姿勢」の列で、第1所定時間分の行のうち、「あり」と判定された回数をカウントすることで求めることができる。さらに、相槌の頻度RFは、「相槌」の列で「あり」と判定された回数を、「頷き」の列で「あり」と判定された回数で割ることで求めることができる。

【0095】

また、算出した相手を見る時間WT、体を倒して近づく時間ATおよび相槌の頻度RFから、数1に示す式に基づいてAC値を算出する。

【0096】

[数1]

$$C1 \times WT + C2 \times FT + C3 \times RF = AC \text{ 値}$$

C1, C2, C3：定数

そして、このように算出されたAC値を閾値Aに基づいて判断することで、アクティブ状態またはパッシブ状態を認識することができる。つまり、数1に示す式で算出されたAC値が閾値Aより大きければアクティブ状態と認識され、AC値が閾値以下であればパッシブ状態と認識される。なお、定数C1, C2およびC3の値を変化させることで、各パラメータに重みを付けることができる。

【0097】

次に、Ta/Li認識とは、ユーザが話し手となり発話している「トーク：Talk（Ta）」状態（話手側状態）であるか、相手ユーザの話を傾聴している「リスン：Listen（Li）」状態（聴取側状態）であるかを認識する処理である。そのため、Ta/Li認識では、第1所定時間分のユーザの行動のうち、ユーザの発話に基づいて認識す

10

20

30

40

50

る。この実施例では、ユーザの発声時間を発話量とし、この発話量が閾値 T に基づいて、トーク状態またはリッスン状態を認識する。

【0098】

たとえば、発話量を T_a とする場合に、発話量 T_a は、行動テーブルにおける「発話」の列で、第1所定時間分の行のうち、「あり」と判定された回数をカウントすることで求めることができる。そして、このように算出された発話量 T_a を閾値 T と比較することで、トーク状態またはパッシブ状態を認識することができる。つまり、発話量 T_a が閾値 T より大きければトーク状態と認識され、発話量 T_a が閾値 T 以下であれば、リッスン状態と認識される。

【0099】

そして、興味対象認識とは、ユーザの興味が「ロボット」、「モニタ」および「アザー」のいずれであるかを認識する処理である。そのため、興味対象認識では、第2所定時間（たとえば、10秒）分のユーザの行動データのうち、ユーザの顔認識および視線方向に基づいて、モニタカメラ22側の顔の認識率 M と、腹部カメラ12側の顔の認識率 S とを算出することで、ユーザの興味が有る物を認識する。

【0100】

たとえば、認識率 M は、行動テーブルにおける「顔認識（モニタ）」および「視線方向（モニタ）」の列で、第1所定時間分の行のうち、「成功」および「モニタ」と判定された回数をカウントすることで求めることができる。一方、認識率 S は、行動テーブルにおける「顔認識（腹部）」および「視線方向」の列で、「成功」および「ロボット」と判定された回数をカウントすることで求めることができる。そして、このように算出された認識率 M および認識率 S と閾値 I_n とを比較することでユーザの興味対象を認識することができる。つまり、認識率 M および認識率 S が共に閾値 I_n 以下であれば、興味対象が「アザー」と認識される。また、認識率 M および認識率 S が共に閾値 I_n より大きく、かつ認識率 M が認識率 S より大きければ、興味対象が「モニタ」と認識され、認識率 S が認識率 M より大きければ、興味対象が「ロボット」と認識される。

【0101】

なお、認識率 S および認識率 M は、顔認識の結果だけで求められてもよい。また、顔認識の結果が百分率で示される場合には、1秒毎の認識結果の総積を認識率としてもよい。さらに、興味対象認識では、認識率の代わりに、前傾姿勢の時間割合と視線方向の時間割合と頭部方向の時間割合とから求められる数値に基づいて認識されてもよい。そして、第1所定時間と第2所定時間とは同じ長さであってもよい。

【0102】

このようにして、 A_c/P_a 認識、 T_a/L_i 認識および興味対象認識の処理が行われた後に、ユーザの状態は各認識結果を組み合わせることで決定する。具体的なユーザの状態としては、ユーザの興味対象がモニタ16である「アクティブ・トーク・モニタ」、「アクティブ・リッスン・モニタ」、「パッシブ・トーク・モニタ」および「パッシブ・リッスン・モニタ」の4種類と、ユーザの興味対象がロボット10である「アクティブ・トーク・ロボット」、「アクティブ・リッスン・ロボット」、「パッシブ・トーク・ロボット」、「パッシブ・リッスン・ロボット」の4種類と、ユーザの興味対象がロボット10でもモニタ16でもない「アクティブ・トーク・アザー」、「アクティブ・リッスン・アザー」、「パッシブ・トーク・アザー」および「パッシブ・リッスン・アザー」の4種類とから成る、合計12種類である。

【0103】

ここで、一定時間が1秒、第1所定時間を30秒とした場合に、「アクティブ・トーク・モニタ」、「アクティブ・リッスン・モニタ」、「パッシブ・トーク・モニタ」および「パッシブ・リッスン・モニタ」と認識されるユーザの状態の一例について説明する。なお、「ロボット」および「アザー」におけるユーザの状態については、ユーザの興味対象が異なるだけで「モニタ」の場合と同じであるため、詳細な説明は省略する。

【0104】

10

20

30

40

50

まず、ユーザの興味対象が「モニタ」と認識されている状態で、発話の継続時間が15秒(50%)以上であり、視線方向の判定において「モニタ」が30回のうち18回(60%)以上判定され、さらに頷きが2回以上あれば「アクティブ・トーク・モニタ」と認識される。

【0105】

次に、ユーザの興味対象が「モニタ」と認識されている状態で、発話の継続時間が9秒(30%)以下であり、視線方向の判定において「モニタ」が30回のうち15回(50%)以上判定され、さらに前傾姿勢の判定において「あり」が30回のうち21回(70%)以上判定されていれば、「アクティブ・リッスン・モニタ」と認識される。

【0106】

次に、ユーザの興味対象が「モニタ」と認識されている状態で、発話の継続時間が12秒(40%)以下であり、視線方向の判定において「モニタ」が30回のうち6回(20%)以上判定され、さらに頷きが一度のなければ、「パッシブ・トーク・モニタ」と認識される。

【0107】

そして、ユーザの興味対象が「モニタ」と認識されている状態で、発話の継続時間が9秒(30%)以下であり、頷きが一度も無く、さらに前傾姿勢の判定において「あり」が30回のうち6回(20%)以下であれば、「パッシブ・リッスン・モニタ」と認識される。

【0108】

なお、このようにして認識されたユーザの状態は、状態データ354(図10)としてメモリ64に記憶されると共に、サーバ24に送信される。また、ユーザの状態の認識には、SVM(Support Vector Machine)が利用されてもよい。

【0109】

図8にはサーバ24の電氣的な構成を示すブロック図が示される。サーバ24は、プロセッサ38、64と同様に、マイクロコンピュータ或いはCPUとも呼ばれる、プロセッサ80が内蔵されている。また、プロセッサ80は、バス82を介して、メモリ84、第1ロボット情報DB86、第2ロボット情報DB88および通信LANボード90に接続されている。

【0110】

メモリ84は、図示しないROMやRAMが組み込まれており、ROMには主として、サーバ24とPC14a、14bなどとのデータ通信を行うためのプログラムなどが予め記憶されている。また、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。

【0111】

第1ロボット情報DB86は、PC14aから送信されるユーザAの行動データおよび状態データを蓄積するためのデータベースである。また、第2ロボットDB88は、PC14bから送信されるユーザBの行動データおよび状態データを蓄積するためのデータベースである。そして、第1ロボット情報DB86および第2ロボット情報DB88は、HDDやSSDのような記憶媒体を用いて構成される。

【0112】

通信LANボード90は、PC14の通信LANボード76と同様に、たとえばDSPで構成され、プロセッサ80から与えられた送信データを無線通信装置92に与える。無線通信装置92は送信データを、ネットワーク200を介して外部のコンピュータ(PC14a、14b)に送信する。また、通信LANボード90は、無線通信装置92を介してデータを受信し、受信データをプロセッサ80に与える。

【0113】

たとえば、受信データはPC14aから送信されるユーザAの行動データであり、プロセッサ80はユーザAの行動データを第1ロボット情報DB86に保存する。さらに、受信データとして、PC14bからユーザAの行動データ取得要求がプロセッサ80に与え

10

20

30

40

50

られると、プロセッサ 80 は、ユーザ A の行動データを送信データとして、通信 LAN ボード 90 に与える。

【0114】

ここで、本願発明の傾聴対話持続システムでは、ユーザの行動データおよび状態データや、相手ユーザの行動データや状態データに基づいて、2人の対話が持続するように、ロボット 10 の動作を制御する。そして、ロボット 10 は、ユーザ A とユーザ B との対話に対して、「疑似傾聴動作」、「発話制御動作」および「注意の引きつけの動作」の3種類の動作を行い、対話を持続させる。

【0115】

また、疑似傾聴動作とは、ユーザ A とユーザ B とが積極的に対話している場合には、どちらか一方の発話を傾聴しているかのように振る舞う動作のことである。さらに、発話制御動作とは、どちらかのユーザが一方向的に話している場合に、2人の発話のバランスを取るため、ユーザを見ることで発話を抑制したり、ユーザに話しかけたりすることで発話を促進したりする動作のことである。そして、注意の誘導や引きつけの動作は、ユーザが対話に対して集中していない場合に、ユーザに話しかけることでユーザの注意を引きつける動作の事である。なお、具体的なロボット 10 の動作については、図 27 - 図 41 に示すフローチャートを用いて説明する。

【0116】

図 9 は、図 5 に示す PC 14 におけるメモリ 68 のメモリマップ 300 の一例を示す図解図である。図 9 に示すようにメモリ 68 はプログラム記憶領域 302 およびデータ記憶領域 304 を含む。プログラム記憶領域 302 には、PC 14 を動作させるためのプログラムとして、データ通信プログラム 312、状況認識プログラム 314、ロボット制御プログラム 316、カメラ制御プログラム 318、発声時間計測プログラム 320 および乱数生成プログラム 322 などが記憶される。

【0117】

データ通信プログラム 312 は、サーバ 24 とデータ通信を行うためのプログラムである。状況認識プログラム 314 は、ユーザの行動を判定し、状態を認識するためのプログラムである。ロボット制御プログラム 316 は、ロボット 10 の動作を決定するためのプログラムである。カメラ制御プログラム 318 は、相手に送信するカメラ画像を決定するためのプログラムである。発声時間計測プログラム 320 は、ユーザ A およびユーザ B の発声時間を計測するためのプログラムである。乱数生成プログラム 322 は、ロボット 10 の動作をランダムに決定する際に実行される処理である。

【0118】

なお、図示は省略するが、PC 14 を動作させるためのプログラムとしては、テレビ電話機能を実現するためのプログラムなどを含む。

【0119】

また、図 10 を参照して、データ記憶領域 304 には、時刻バッファ 330、モニタカメラバッファ 332、腹部カメラバッファ 334、音声バッファ 336、判定結果バッファ 338、顔認識結果バッファ 340、興味対象認識結果バッファ 342、データ通信バッファ 344、相手行動データバッファ 346、状態データバッファ 348 および乱数バッファ 350 が設けられる。さらに、データ記憶領域 304 には、行動テーブルデータ 352 および状態データ 354 が記憶されるとともに、Ac / Pa フラグ 356、Ta / Li フラグ 358、状態カウンタ 360 および発声カウンタ 362 がさらに設けられる。

【0120】

時刻バッファ 330 は、RTC 64a が出力する日時情報が一時的に記憶されるバッファである。モニタカメラバッファ 332 は、モニタカメラ 22 によって撮影された画像が一時的に記憶されるバッファである。腹部カメラバッファ 334 は、腹部カメラ 12 によって撮影された画像が一時的に記憶されるバッファである。音声バッファ 336 は、マイク 20 によって集音された音声が一時的に記憶されるバッファである。

【0121】

判定結果バッファ 338 は、ユーザの前傾姿勢の有無を判定する姿勢判定、ユーザの発話の有無を判定する発話判定、ユーザの頭部方向を判定する頭部方向判定、ユーザの視線方向を判定する視線方向判定、ユーザの頷きを判定する頷き判定およびユーザの相槌を判定する相槌判定の各判定結果を一時的に記憶するためのバッファである。顔認識結果バッファ 340 は、モニタカメラ 22 および腹部カメラ 12 によって撮影された画像に対して行われる顔認識の結果が一時的に記憶されるバッファである。興味対象認識結果バッファ 342 は、ユーザが興味を持っていると判定された結果が一時的に記憶されるバッファであり、たとえば「ロボット」、「モニタ」および「アザー」を示すデータが一時的に記憶される。

【0122】

データ通信バッファ 344 は、サーバ 24 とのデータ通信によって得られた相手の行動データや、状態データなどが一時的に記憶されるバッファである。相手動作データバッファ 346 は、相手の行動データを一時的に記憶するためのバッファである。状態データバッファ 348 は、相手の状態データや、こちら側のユーザの状態データを一時的に記憶するためのバッファである。乱数データバッファ 350 は、乱数生成プログラム 322 によって生成された乱数が一時的に記憶されるバッファである。

【0123】

行動テーブルデータ 352 は、図 6 に示す行動テーブルであり、一定時間毎に最新の行動データが追記される。状態データ 354 は、第 1 所定時間分の行動データから認識されるユーザの状態を示すデータであり、たとえば「アクティブ・トーク・ロボット」を示す文字列で構成される。

【0124】

Ac / Pa フラグ 356 は、Ac / Pa 認識結果を示すフラグである。たとえば Ac / Pa フラグ 356 は 1 ビットのレジスタで構成される。Ac / Pa フラグ 356 がオン（成立）されると、レジスタにはデータ値「1」が設定される。一方、Ac / Pa フラグ 356 がオフ（不成立）されると、レジスタにはデータ値「0」が設定される。また、Ac / Pa フラグ 356 は、アクティブ状態と認識されるとオンになり、パッシブ状態と認識されるとオフになる。Ta / Li フラグ 358 は、Ta / Li 認識結果を示すフラグである。つまり、Ta / Li フラグ 358 はトーク状態と認識されるとオンになり、リスン状態と認識されるとオフになる。なお、Ta / Li フラグ 358 の構成については、Ac / Pa フラグ 356 と同様であるため、詳細な説明は省略する。

【0125】

状態カウンタ 360 は、ユーザの状態を認識する際に、第 1 所定時間分の行動データを取得するために用いられるカウンタである。たとえば、状態カウンタ 360 は、PC14 の電源がオンにされるとカウントされ始め、第 1 所定時間分の行動データが取得される毎にリセットされる。発声カウンタ 362 は、発声時間計測プログラム 320 によって計測される発声時間をカウントするためのカウンタである。

【0126】

なお、図示は省略するが、データ記憶領域 304 には、各判定に利用されるテンプレートのデータや、様々な計算の結果を一時的に格納するバッファなどが設けられると共に、PC14 の動作に必要な他のカウンタやフラグなども設けられる。

【0127】

図 11 には状況認識プログラム 314 を構成するプログラムが示される。図 11 を参照して、状況認識プログラム 314 は、画像 / 音声取得プログラム 314 a、視線判定プログラム 314 b、発話判定プログラム 314 c、頭部方向判定プログラム 314 d、視線方向判定プログラム 314 e、顔認識プログラム 314 f、頷き判定プログラム 314 g、相槌判定プログラム 314 h、同期プログラム 314 j、状態認識プログラム 314 k、Ac / Pa 認識プログラム 314 m、Ta / Li 認識プログラム 314 n および興味対象認識プログラム 314 p から構成される。

【0128】

画像／音声取得プログラム 3 1 4 a は、モニタカメラ 2 2 および腹部カメラ 1 2 によって撮影された画像と、マイク 2 0 によって集音された音声とを、バッファに取り込むためのプログラムである。姿勢判定プログラム 3 1 4 b は、ユーザの姿勢が前傾姿勢であるかを判定するためのプログラムである。発話判定プログラム 3 1 4 c は、ユーザが発話しているかを判定するためのプログラムである。頭部方向判定プログラム 3 1 4 d は、ユーザの頭部方向を判定するためのプログラムである。視線方向判定プログラム 3 1 4 e は、ユーザの視線方向を判定するためのプログラムである。顔認識プログラム 3 1 4 f は、モニタカメラ 2 2 および腹部カメラ 1 2 によって撮影された画像における顔領域を認識するためのプログラムである。頷き判定プログラム 3 1 4 g は、ユーザが頷いたかを判定するためのプログラムである。相槌判定プログラム 3 1 4 h は、ユーザが相槌をしたかを判定するためのプログラムである。同期プログラム 3 1 4 j は、姿勢判定結果、発話判定結果、頭部方向判定結果、視線方向判定結果、顔認識結果、頷き判定結果および相槌判定結果を同期して、行動データとするためのプログラムである。

10

【 0 1 2 9 】

状態認識プログラム 3 1 4 k は、ユーザの状態を第 1 所定時間毎に認識するためのプログラムである。A c / P a 認識プログラム 3 1 4 m は、ユーザがアクティブ状態かパッシブ状態かを認識するためのプログラムである。T a / L i 認識プログラム 3 1 4 n は、ユーザがトーク状態かリスン状態かを認識するためのプログラムである。興味対象認識プログラム 3 1 4 p は、ユーザの興味対象を認識するためのプログラムである。

20

【 0 1 3 0 】

図 1 2 にはロボット制御プログラム 3 1 6 を構成するプログラムが示される。図 1 2 を参照して、全体プログラム 3 1 6 a、アクティブトークプログラム 3 1 6 b、アクティブリスンプログラム 3 1 6 c、非アクティブプログラム 3 1 6 d、アザープログラム 3 1 6 e、発話継続プログラム 3 1 6 f、発話抑制プログラム 3 1 6 g、発話促進プログラム 3 1 6 h、注意引きつけプログラム 3 1 6 j、傍参与者の疑似傾聴プログラム 3 1 6 k、第 1 積極的疑似傾聴プログラム 3 1 6 m、ユーザ発話抑制プログラム 3 1 6 m、第 2 積極的疑似傾聴プログラム 3 1 6 p、注意誘導プログラム 3 1 6 q およびユーザ発話促進プログラム 3 1 6 r から構成される。

【 0 1 3 1 】

全体プログラム 3 1 6 a は、メインルーチンとも呼ばれ、状態データに応じて、対話が持続するようにロボット 1 0 を制御するための全体的な処理を行うプログラムである。アクティブトークプログラム 3 1 6 b は、状態データが「アクティブ・トーク・ロボット」である場合に実行されるプログラムである。アクティブリスンプログラム 3 1 6 c は、状態データが「アクティブ・リスン・ロボット」である場合に実行されるプログラムである。非アクティブプログラム 3 1 6 d は、ユーザがパッシブ状態であるか、興味対象が「ロボット」と認識されている場合に実行されるプログラムである。アザープログラム 3 1 6 e は、ユーザの興味対象が「アザー」である場合に実行されるプログラムである。

30

【 0 1 3 2 】

発話継続プログラム 3 1 6 f は、2 人のユーザの対話が継続されるようにロボット 1 0 を動作させるためのプログラムである。発話抑制プログラム 3 1 6 g は、ユーザが一方的に発話している場合に発話を抑制するようにロボット 1 0 を動作させるためのプログラムである。発話促進プログラム 3 1 6 h は、ユーザが発話をせずに、相手ユーザの発話を一方的に傾聴している場合に、発話を促すようロボット 1 0 を動作させるプログラムである。注意引きつけプログラム 3 1 6 j は、対話に興味を失っているユーザの注意を引けつめるようにロボット 1 0 を動作させるためのプログラムである。

40

【 0 1 3 3 】

傍参与者の疑似傾聴プログラム 3 1 6 k は、2 人のユーザの対話が継続している場合に、どちらか一方のユーザの話を傾聴しているかのように、ロボット 1 0 を動作させるためのプログラムである。なお、「傍参与者」とは、会話に参加しているが「話し手」でも「聴き手」でもない人物を示す。

50

【 0 1 3 4 】

第 1 積極的疑似傾聴プログラム 3 1 6 m は、相手ユーザの話を傾聴しているかのように、ロボット 1 0 を動作させるためのプログラムである。ユーザ発話抑制プログラム 3 1 6 n は、たとえばロボット 1 0 a の傍に居るユーザ A の発話を抑制するように、ロボット 1 0 a を動作させるためのプログラムである。第 2 積極的疑似傾聴プログラム 3 1 6 p は、たとえばロボット 1 0 a の傍に居るユーザ A の発話を傾聴しているかのように、ロボット 1 0 a を動作させるためのプログラムである。注意誘導プログラム 3 1 6 q は、ユーザの注意を誘導するように、ロボット 1 0 を動作させるためのプログラムである。ユーザ発話促進プログラム 3 1 6 r は、たとえばロボット 1 0 a の傍に居るユーザ A の発話を促すように、ロボット 1 0 を動作させるためのプログラムである。

10

【 0 1 3 5 】

以下、P C 1 4 によって実行される本願発明のフロー図について説明する。また、図 1 3 - 図 2 5 のフロー図は状況認識プログラム 3 1 4 を構成する各プログラムの処理を示し、図 2 6 は発声時間計測プログラム 3 2 0 による処理を示し、図 2 7 - 図 4 1 はロボット制御プログラム 3 1 6 を構成する各プログラムの処理を示し、図 4 2 はカメラ制御プログラム 3 1 8 による処理を示す。

【 0 1 3 6 】

図 1 3 には画像 / 音声取得プログラム 3 1 4 a の処理を示すフロー図が示される。たとえば、P C 1 4 のプロセッサ 6 4 は、ユーザによって P C 1 4 の電源がオンにされると、ステップ S 1 で腹部カメラ 1 2 による画像データを取得する。つまり、ロボット 1 0 から入力される映像信号から画像データを取得し、腹部カメラバッファ 3 3 4 に一旦記憶させる。続いて、ステップ S 3 では、モニタカメラ 2 2 による画像データを取得する。つまり、モニタカメラ 2 2 から入力される映像信号から画像データを取得し、モニタカメラバッファ 3 3 2 に一旦記憶させる。続いて、ステップ S 5 では、音声データを取得し、ステップ S 1 に戻る。つまり、マイク 2 0 によって集音された音声から音声データを抽出し、音声バッファ 3 3 6 に一旦記憶させる。なお、ステップ S 1 - S 5 の処理は、約 1 秒毎に繰り返される。

20

【 0 1 3 7 】

図 1 4 には姿勢判定プログラム 3 1 4 b の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 7 で画像データを取得したか否かを判断する。つまり、モニタカメラ バッファ 3 3 2 および腹部カメラバッファ 3 3 4 に新たな画像データが記憶されたか否かを判断する。ステップ S 7 で“ N O ”であれば、つまり画像データが取得されていなければステップ S 7 の処理を繰り返し実行する。一方、ステップ S 7 で“ Y E S ”であれば、つまり画像データが取得されていれば、ステップ S 9 でユーザの姿勢を判定する。たとえば、モニタカメラバッファ 3 3 2 および腹部カメラバッファ 3 3 4 に記憶される画像データに対してユーザの前傾姿勢の有無を判定する。そして、どちらの画像でも前傾姿勢が「あり」と判定された場合に、姿勢判定結果を「あり」とする。なお、ステップ S 9 の処理を実行するプロセッサ 6 4 は姿勢判定手段として機能する。

30

【 0 1 3 8 】

続いて、ステップ S 1 1 では、現在時刻を取得する。つまり、時刻バッファ 3 3 0 に記憶される日時情報を取得する。続いて、ステップ S 1 3 では、姿勢の判定結果に現在時刻を対応付ける。つまり、複数の判定結果を同期させるために、現在時刻を対応付ける。そして、現在時刻が対応付けられた姿勢の判定結果は判定結果バッファ 3 3 8 に一時的に記憶される。

40

【 0 1 3 9 】

なお、他の実施例では、モニタカメラ 2 2 または腹部カメラ 1 2 のどちらか一方の画像だけで姿勢判定を行ってもよい。また、図 1 5 - 図 2 0 に示す他の判定処理でも、ステップ S 1 1 と同様に日時情報を取得し、ステップ S 1 3 と同様に日時情報を対応付ける処理が存在するが、処理内容は全て同じであるため、他のフロー図では詳細な説明は省略する。

50

【 0 1 4 0 】

図 1 5 には発話判定プログラム 3 1 4 c の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 1 5 で音声データを取得したか否かを判断する。つまり、音声バッファ 3 3 6 に新たな音声データが記憶されたか否かを判断する。ステップ S 1 5 で “ N O ” であれば、つまり音声データが取得されていなければステップ S 1 5 の処理を繰り返し実行する。一方、ステップ S 1 5 で “ Y E S ” であれば、つまり音声データが取得されていればステップ S 1 7 でユーザの発話を判定する。たとえば、音声バッファ 3 3 6 に記憶される音声データの音声レベルが一定値以上であるか否かを判定する。ステップ S 1 9 では現在時刻を取得し、ステップ S 2 1 では発話の判定結果に現在時刻を対応付ける。そして、ステップ S 1 9 の処理が終了するとステップ S 1 5 に戻る。なお、ステップ S 1 9 では、現在時刻が対応付けられた発話の判定結果は、判定結果バッファ 3 3 8 に一旦記憶される。なお、ステップ S 1 7 の処理を実行するプロセッサ 6 4 は発話判定手段として機能する。

10

【 0 1 4 1 】

図 1 6 には頭部方向判定プログラム 3 1 4 d の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 2 3 で画像データを取得したか否かを判断する。つまり、モニタカメラバッファ 3 3 2 および腹部カメラバッファ 3 3 4 に新たな画像データが記憶されたか否かを判断する。ステップ S 2 3 で “ N O ” であれば、つまり画像データが新たに取得されていなければステップ S 2 3 の処理を繰り返す。一方、ステップ S 2 3 で “ Y E S ” であれば、つまり新たな画像データが取得されていればステップ S 2 5 でユーザの頭部方向を判定する。たとえば、モニタカメラバッファ 3 3 2 に記憶される画像に対してテンプレートマッチング処理を実行することで、「モニタ」、「ロボット」および「アザー」を判定する。さらに、腹部カメラバッファ 3 3 4 に記憶される画像に対しても同様に判定する。なお、ステップ S 2 5 の処理を実行するプロセッサ 6 4 は頭部方向判定手段として機能する。

20

【 0 1 4 2 】

続いて、ステップ S 2 7 では現在時刻を取得し、ステップ S 2 9 では頭部方向の判定結果に現在時刻を対応付ける。そして、ステップ S 2 9 の処理が終了するとステップ S 2 3 に戻る。なお、ステップ S 2 9 では、モニタカメラ 2 2 によって撮影された画像の頭部方向の判定結果と、腹部カメラ 1 2 によって撮影された画像の頭部方向の判定結果とをそれぞれ判定結果バッファ 3 3 8 に記憶させる。

30

【 0 1 4 3 】

図 1 7 には視線方向判定結果プログラム 3 1 4 e の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 3 1 で画像データを取得したか否かを判断する。つまり、モニタカメラバッファ 3 3 2 および腹部カメラバッファ 3 3 4 に新しい画像が記憶されたか否かを判断する。ステップ S 3 1 で “ N O ” であれば、つまり新しい画像データが取得されていなければステップ S 3 1 の処理を繰り返し実行する。一方、ステップ S 3 1 で “ Y E S ” であれば、つまり画像データが取得されれば、ステップ S 3 3 でユーザの視線方向を判定する。たとえば、ステップ S 3 3 では、まず視線サーバ 7 0 にモニタカメラバッファ 3 3 2 に記憶される画像データを入力することで、ユーザの視線方向を特定する。次に、視線サーバ 7 0 によって特定された視線方向に基づいて「モニタ」、「ロボット」および「アザー」を判定する。そして、腹部カメラバッファ 3 3 4 に記憶される画像に対しても、同様の判定を行う。なお、ステップ S 3 3 の処理を実行するプロセッサ 6 4 は視線方向判定手段として機能する。

40

【 0 1 4 4 】

続いて、ステップ S 3 5 では現在時刻を取得し、ステップ S 3 7 で視線方向の判定結果に現在時刻を対応付ける。そして、ステップ S 3 7 の処理が終了すれば、ステップ S 3 1 に戻る。なお、ステップ S 3 7 の処理では、ステップ S 2 9 (図 1 6) と同様に、モニタカメラ 2 2 によって撮影された画像の視線方向の判定結果と、腹部カメラ 1 2 によって撮影された画像の視線方向の判定結果とをそれぞれ判定結果バッファ 3 3 8 に記憶させる。

50

【 0 1 4 5 】

図 1 8 には顔認識処理プログラム 3 1 4 f の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 3 9 で画像データを取得したか否かを判断する。つまり、モニタカメラバッファ 3 3 2 および腹部カメラバッファ 3 3 4 に新しい画像が記憶され、更新されたか否かを判断する。ステップ S 3 9 で “ N O ” であれば、つまり画像データが取得されていなければ、ステップ S 3 9 の処理を繰り返す。一方、ステップ S 3 9 で “ Y E S ” であれば、つまり画像データが取得されれば、ステップ S 4 1 でユーザの顔を認識する。たとえば、モニタカメラバッファ 3 3 2 に記憶される画像データに対して所定の顔認識処理を加えることで、顔領域を認識できたか否かを判断する。そして、認識できた場合には顔認識結果を「成功」と判定し、認識できなかった場合には顔認識結果を「失敗」と判定する。そして、腹部カメラバッファ 3 3 4 に記憶される画像データに対しても同様の判定を行う。なお、ステップ S 4 1 の処理を実行するプロセッサ 6 4 は顔認識手段として機能する。

10

【 0 1 4 6 】

続いて、ステップ S 4 3 では現在時刻を取得し、ステップ S 4 5 では顔認識結果に現在時刻を対応付ける。そして、ステップ S 4 5 の処理が終了すれば、ステップ S 3 9 に戻る。また、ステップ S 3 9 の処理では、現在時刻が対応付けられた各認識結果を顔認識結果バッファ 3 4 0 に一旦記憶させる。

【 0 1 4 7 】

図 1 9 には顔き判定プログラム 3 1 4 g の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 4 7 で画像データを取得したか否かを判断する。つまり、モニタカメラバッファ 3 3 2 および腹部カメラバッファ 3 3 4 が更新されたか否かを判断する。ステップ S 4 7 で “ N O ” であれば、つまり画像データが取得されていなければ、ステップ S 4 7 を繰り返す。一方、ステップ S 4 7 で “ Y E S ” であれば、つまり画像データが取得されれば、ステップ S 4 9 でユーザの顔きを判定する。たとえば、モニタカメラバッファ 3 3 2 および腹部カメラバッファ 3 3 4 に記憶される画像データに対して、テンプレートマッチング処理を加えて、ユーザの顔きの有無をそれぞれ判定する。そして、2つの判定結果で顔きが「あり」と判定される場合に、ユーザの顔きが「あり」と判定される。なお、ステップ S 4 9 の処理を実行するプロセッサ 6 4 は顔き判定手段として機能する。

20

【 0 1 4 8 】

続いて、ステップ S 5 1 で現在時刻を取得し、ステップ S 5 3 で顔きの判定結果に現在時刻を対応付ける。そして、ステップ S 5 3 の処理が終了すれば、ステップ S 4 7 に戻る。また、日時情報が対応付けられた顔きの判定結果は、判定結果バッファ 3 3 8 に記憶される。

30

【 0 1 4 9 】

図 2 0 には相槌判定プログラム 3 1 4 h の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 5 5 で顔き判定が終了したか否かを判断する。たとえば、顔きの判定結果が判定結果バッファ 3 3 8 に記憶されているか否かを判断する。ステップ S 5 5 で “ N O ” であれば、つまり顔き判定が終了していなければ、ステップ S 5 5 の処理を繰り返す。一方、ステップ S 5 5 で “ Y E S ” であれば、つまり顔き判定が終了していれば、ステップ S 5 7 で発話判定が終了したか否かを判断する。たとえば、発話の判定結果が判定結果バッファ 3 3 8 に記憶されているか否かを判断する。ステップ S 5 7 で “ N O ” であれば、つまり発話判定が終了していなければステップ S 5 5 に戻る。一方、ステップ S 5 7 で “ Y E S ” であれば、つまり発話判定が終了していれば、ステップ S 5 9 で日時情報に基づいてユーザの相槌を判定する。たとえば、発話判定結果と顔き判定結果とに対応付けられている日時情報に基づいて、それぞれを同期する。そして、顔きの判定結果が「あり」である場合に、音声バッファに記憶される音声データに対して、上記した音声プロソディを判定することで、相槌の判定をする。なお、ステップ S 5 9 の処理を実行するプロセッサ 6 4 は相槌判定手段として機能する。

40

【 0 1 5 0 】

50

続いて、ステップ S 6 1 では、相槌の判定結果に時刻を対応付けて、ステップ S 5 5 に戻る。つまり、ステップ S 6 1 では、発話または頷きの判定結果に対応付けられている日時情報を、相槌の判定結果に対応付ける。そして、時刻が対応付けられた相槌の判定結果は、判定結果バッファ 3 4 0 に記憶される。

【 0 1 5 1 】

図 2 1 には同期プログラム 3 1 4 j の処理を示すフロー図が示される。プロセッサ 6 4 は、ステップ S 6 3 で各判定が終了したか否かを判断する。たとえば、判定結果バッファ 3 3 8 に姿勢、発話、頭部方向、視線方向、頷きおよび相槌の判定結果が記憶され、かつ顔認識結果バッファ 3 4 0 に顔認識結果が記憶されているか否かを判定する。ステップ S 6 3 で “ N O ” であれば、つまり各判定が終了していなければ、ステップ S 6 3 の処理を繰り返して実行する。一方、ステップ S 6 3 で “ Y E S ” であれば、つまり各判定が終了していれば、ステップ S 6 5 で各判定結果および顔認識結果を時刻に基づいて同期する。つまり、各判定結果および顔認識結果に対応付けられた時刻に基づいて同期する。

【 0 1 5 2 】

続いて、ステップ S 6 7 では、同期した各判定結果を行動データとし、行動テーブルに記録する。つまり、図 6 に示す行動テーブルにおいて、新たな行に各判定結果および顔認識結果を記録する。続いて、ステップ S 6 9 では、現在の行動データをサーバ 2 4 に送信する。そして、ステップ S 6 9 の処理が終了すればステップ S 6 3 に戻る。つまり、ステップ S 6 9 では、行動テーブルにおいて、新たに追加された行に対応する行動データをサーバ 2 4 に送信する。

【 0 1 5 3 】

このように、図 1 3 - 図 2 1 の処理が一定時間毎に並列的に実行されることで、ユーザの行動データが判定されるとともに、サーバ 2 4 に送信される。そして、ユーザの行動データに複数の判定結果が含まれるため、ユーザの状態が適確に認識される。

【 0 1 5 4 】

なお、ステップ S 6 5 , S 6 7 の処理を実行するプロセッサ 6 4 は判定手段として機能する。また、ステップ S 6 9 の処理を実行するプロセッサ 6 4 は送信手段として機能する。

【 0 1 5 5 】

図 2 2 には状態認識プログラム 3 1 4 k の処理を示すフロー図が示される。たとえば、P C 1 4 のプロセッサ 6 4 は、ユーザによって P C 1 4 の電源がオンにされると、ステップ S 7 1 で第 1 所定時間が経過したか否かを判断する。たとえば、状態カウンタ 3 6 0 の値が第 1 所定時間を示す値と超えたか否かを判断する。ステップ S 7 1 で “ N O ” であれば、つまり第 1 所定時間が経過していなければ、ステップ S 7 1 の処理を繰り返して実行する。一方、ステップ S 7 1 で “ Y E S ” であれば、つまり第 1 所定時間が経過してれば、ステップ S 7 3 で第 1 所定時間分の行動データを取得する。つまり、行動テーブルから、現在時刻から第 1 所定時間前までの行動データを取得する。なお、ステップ S 7 3 の処理が終了すると、状態カウンタ 3 6 0 はリセットされる。

【 0 1 5 6 】

続いて、ステップ S 7 5 では A c / P a 認識処理を実行し、ステップ S 7 7 では T a / L i 認識処理を実行し、ステップ S 7 9 では興味対象認識処理を実行する。なお、ステップ S 7 5 , S 7 7 および S 7 9 の処理については、図 2 3 、図 2 4 および図 2 5 に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【 0 1 5 7 】

続いて、ステップ S 8 1 では、各認識結果を状態データ 3 5 4 として記憶する。たとえば、A c / P a 認識結果がアクティブ状態であり、T a / L i 認識結果がトークであり、興味対象認識結果が「ロボット」であれば、状態データ 3 5 4 は、「アクティブ・トーク・ロボット」としてメモリ 6 8 に記憶される。続いて、ステップ S 8 3 では、状態データ 3 5 4 をサーバ 2 4 に送信し、ステップ S 7 1 に戻る。

【 0 1 5 8 】

なお、ステップS 7 5の処理を実行するプロセッサ6 4は積極性認識手段として機能し、ステップS 7 7の処理を実行するプロセッサ6 4は話者状態認識手段として機能し、ステップS 7 9の処理を実行するプロセッサ6 4は興味対象認識手段として機能する。そして、ステップS 7 5 - 8 1の処理を実行するプロセッサ6 4は認識手段として機能する。
【0 1 5 9】

図2 3にはA c / P a 認識プログラム3 1 4 mの処理を示すフロー図が示される。プロセッサ6 4はステップS 9 1で相手を見る時間W Tを算出し、ステップS 9 3では前傾姿勢の時間F Tを算出し、ステップS 9 5では相槌の頻度R Fを算出する。つまり、ステップS 9 1, S 9 3およびS 9 5では、上位ルーチンのステップS 7 3で取得された行動データに基づいて、相手を見る時間W T、前傾姿勢の時間F Tおよび相槌の頻度R Fを算出する。そして、ステップS 9 7では、各算出結果からA C 値を算出する。つまり、上述の数1に示す式に基づいてA C 値を算出する。続いて、ステップS 9 9では、算出したA C 値が閾値Aより大きいか否かを判断する。

10

【0 1 6 0】

ステップS 9 9で“ Y E S ”であれば、つまりA C 値が閾値Aよりも大きければ、ステップS 1 0 1でアクティブ状態に設定する。つまり、A c / P a フラグ3 5 6をオンに設定する。一方、ステップS 9 9で“ N O ”であれば、つまりA C 値が閾値A以下であれば、ステップS 1 0 3でパッシブ状態に設定する。つまり、A c / P a フラグ3 5 6をオフに設定する。なお、ステップS 1 0 1またはステップS 1 0 3の処理が終了すれば、A c / P a 認識処理を終了して状態認識処理に戻る。

20

【0 1 6 1】

図2 4にはT a / L i 認識プログラム3 1 4 nの処理を示すフロー図が示される。プロセッサ6 4はステップS 1 1 1で第1所定時間分の発話判定結果から発話量T aを算出する。つまり、上位ルーチンのステップS 7 3で取得された行動データに基づいて、発話量T aを算出する。続いて、ステップS 1 1 3では算出された発話量T aが閾値Tより大きいか否かを判断する。ステップS 1 1 3で“ Y E S ”であれば、つまり発話量T aが閾値Tよりも大きければ、ステップS 1 1 5でトーク状態に設定する。つまり、ステップS 1 1 5ではT a / L i フラグ3 5 8をオンにする。一方、ステップS 1 1 3で“ N O ”であれば、つまり発話量T aが閾値T a以下であれば、ステップS 1 1 7でリッスン状態に設定する。つまり、ステップS 1 1 7ではT a / L i フラグ3 5 8をオフにする。

30

【0 1 6 2】

そして、ステップS 1 1 5またはステップS 1 1 7の処理が終了すれば、T a / L i 認識処理を終了して、状態認識処理に戻る。

【0 1 6 3】

図2 5には興味対象認識プログラム3 1 4 pの処理を示すフロー図が示される。プロセッサ6 4はステップS 1 2 1で第2所定時間分の行動データを取得する。つまり、上位ルーチンである、ステップS 7 3で取得された行動データから、第2所定時間分の行動データを取得する。続いて、ステップS 1 2 3では、モニタカメラ1 2 3の顔認識結果および視線判定結果から認識率Mを算出し、ステップS 1 2 5では、腹部カメラ1 2の顔認識の結果および視線判定結果から認識率Sを算出する。つまり、ステップS 1 2 3, S 1 2 5では、ステップS 1 2 1で取得した第2所定時間分の行動データに基づいて、認識率Mおよび認識率Sを算出する。なお、ステップS 1 2 3の処理を実行するプロセッサ6 4は第1認識率算出手段として機能し、ステップS 1 2 5の処理を実行するプロセッサ6 4は第2認識率算出手段として機能する。

40

【0 1 6 4】

続いて、ステップS 1 2 7では、各認識率は閾値I nより大きいか否かを判断する。つまり、認識率Mおよび認識率Sが閾値I nより大きいか否かを判断する。ステップS 1 2 7で“ N O ”であれば、つまり認識率Mおよび認識率Sが共に閾値I n以下であれば、ステップS 1 3 5に進む。ステップS 1 2 7で“ Y E S ”であれば、つまり認識率Mおよび認識率Sが閾値I nより大きければ、ステップS 1 2 9で、認識率Mが認識率Sよりも大

50

きいか否かを判断する。

【0165】

ステップS129で“YES”であれば、つまり認識率Mが認識率Sよりも大きければ、ステップS131でモニタ状態に設定する。つまり、ユーザの興味対象がモニタ16であるため、興味対象結果バッファ342に「モニタ」を示すデータを一時記憶させる。一方、ステップS129で“NO”であれば、つまり認識率Mが認識率S以下であれば、ステップS133でロボット状態に設定する。つまり、ユーザの認識対象がロボット12であるため、興味対象認識結果バッファ342に「ロボット」を示すデータを一時的に記憶させる。そして、認識率Mおよび認識率Sが閾値In以下である場合、ステップS135でアザー状態に設定する。つまり、ステップS135では、ユーザの興味対象がロボット12およびモニタ16ではないため、興味対象認識結果バッファ342に「アザー」を示すデータを一時的に記憶させる。

10

【0166】

なお、ステップS131、S133およびステップS135の処理が終了すれば、興味対象認識処理を終了して、状態認識処理に戻る。また、本実施例では、興味対象がユーザの顔の認識率M、Sに基づいて設定されるため、ユーザが見る対象を正確に認識することができる。

【0167】

また、ステップS131、S133およびS135の処理を実行するプロセッサ64は設定手段として機能する。

20

【0168】

このように、図22 - 図25に示す処理が実行されることで、ユーザの状態データがメモリ68に記憶されるとともに、サーバ24に送信される。また、他の実施例ではAc/Pa認識処理、Ta/Li認識処理および興味対象認識処理が並列的に処理されてもよく、この場合は、興味対象認識処理に限り、行動データを第2所定時間毎に取得することで、ユーザの興味対象を認識する。

【0169】

図26には発声時間計測プログラム320の処理を示すフロー図が示される。プロセッサ20はステップS141で音声レベルが所定値以上か否かを判断する。たとえば、マイク20によって集音された音声の音声レベルが、人間の発話と判断できる所定値以上であるか否かを判断する。ステップS141で“NO”であれば、つまり音声レベルが所定値未満であれば、ステップS141の処理を繰り返す。一方、ステップS141で“YES”であれば、つまり音声レベルが所定値以上であれば、ステップS143で発声カウンタ362をインクリメントする。つまり、ステップS143では発声時間をカウントするために、発声カウンタ362をインクリメントする。なお、ステップS141 - S145の処理を実行するプロセッサ64は計測手段として機能する。

30

【0170】

続いて、ステップS145では、音声レベルが所定値未満になったか否かを判断する。たとえば、マイク20によって集音された音声の音声レベルが所定値未満になったか否かを判断する。つまり、ステップS145で“NO”であれば、つまり音声レベルが所定値以上であれば、ステップS143に戻る。一方、ステップS145で“YES”であれば、つまり音声レベルが所定値未満であれば、ステップS147で発声カウンタ362を初期化し、ステップS141に戻る。つまり、ユーザの発話が終了したため、発声時間を計測する発声カウンタ362を初期化する。

40

【0171】

なお、ステップS141およびステップS145では、マイク20に入力される音声の音声レベルだけに限らず、スピーカ18から出力される音声の音声レベルに基づいて判断する。これにより、相手ユーザの発声時間を計測することも可能になる。

【0172】

また、他の実施例では、発声カウンタ362を利用せず、行動テーブルにおける発話の

50

判定結果に基づいて発声時間が計測されてもよい。つまり、発話判定において連続して「あり」と判定される回数をカウントすることで、発声時間を測定することができる。また、この場合、たとえばユーザ A だけに特化して、発声時間を計測することが可能になり、ユーザ A とユーザ B とが同時に発話しているときには、正確に各ユーザの発声時間を計測できるようになる。さらに、対話中の全ての発話判定結果が行動テーブルに記録されるようにすれば、対話中の総合発声時間や、各発声時間をそれぞれ算出することができる。

【0173】

図 27 には、ロボット制御プログラム 316 に含まれる全体プログラム 316a の処理を示すフロー図が示される。たとえば、PC14 のプロセッサ 64 は、テレビ電話機能による通話が開始されると、ステップ S161 で終了操作か否かを判断する。たとえば、テレビ電話機能による通話を終了する操作がされたか否かを判断する。ステップ S161 で“YES”であれば、つまり終了操作が行われると、全体処理を終了する。一方、ステップ S161 で“NO”であれば、つまり終了操作が行われなければ、ステップ S163 で状態データ 354 を参照する。

【0174】

続いて、ステップ S165 ではモニタ状態か否かを判断する。つまり、状態データ 354 に、ユーザの興味対象がモニタ 16 であること示す「モニタ」が含まれているか否かを判断する。ステップ S165 で“NO”であれば、つまりユーザの興味対象がモニタ 16 でなければ、ステップ S175 に進む。一方、ステップ S165 で“YES”であれば、つまりユーザの興味対象がモニタ 16 であれば、ステップ S167 でアクティブか否かを判断する。つまり、状態データ 354 に、ユーザがアクティブ状態であることを示す「アクティブ」が含まれているか否かを判断する。ステップ S167 で“NO”であれば、つまりユーザがパッシブ状態であれば、ステップ S177 に進む。

【0175】

一方、ステップ S167 で“YES”であれば、つまりユーザがアクティブ状態であれば、ステップ S169 でトークであるか否かを判断する。つまり、状態データ 354 に、ユーザがトーク状態であることを示す「トーク」が含まれるか否かを判断する。ステップ S169 で“YES”であれば、つまり、ユーザがトーク状態であれば状態データ 354 は「アクティブ・トーク・モニタ」であるため、ステップ S171 でアクティブトーク処理を実行し、ステップ S161 に戻る。なお、このアクティブトーク処理については、図 28 に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0176】

ステップ S169 で“NO”であれば、つまりユーザがリスン状態であれば状態データ 354 は「アクティブ・リスン・モニタ」であるため、ステップ S173 でアクティブリスン処理を実行し、ステップ S161 に戻る。なお、このアクティブリスン処理については、図 29 に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0177】

また、ユーザの興味対象がモニタ 16 以外である場合、ステップ S175 で、ロボット状態か否かを判断する。つまり、状態データ 354 に、ユーザの興味対象がロボット 10 であることを示す「ロボット」が含まれているか否かを判断する。ステップ S175 で“YES”であれば、つまりユーザの興味対象がロボット 10 であれば、ステップ S177 で非アクティブ処理を実行し、ステップ S161 に戻る。なお、この非アクティブ処理については図 30 に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。一方、ステップ S175 で“NO”であれば、つまりユーザの興味対象がロボット 10 でもモニタ 16 でもなければ、ステップ S179 でアザー処理を実行し、ステップ S161 に戻る。なお、アザー処理については図 31 に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0178】

なお、他の実施例では、ステップ S165、S175 および S179 の処理を省略し、

10

20

30

40

50

ステップS 1 6 7で、アクティブ・モニタか否かを判断するようにしてもよい。また、ステップS 1 7 1, S 1 7 3, S 1 7 7およびS 1 7 9の処理を実行するプロセッサ6 4は動作付与手段として機能する。

【0 1 7 9】

図2 8にはアクティブトークプログラム3 1 6 bの処理を示すフロー図が示される。なお、図2 8 - 図4 2までのフロー図に示される処理は、P C 1 4 aによって実行されるものとして説明する。そのため、「ユーザ」はユーザAを示し、「相手」はユーザB、つまり相手ユーザを示す。

【0 1 8 0】

P C 1 4 aのプロセッサ6 4は、ステップS 1 7 1の処理が実行されると、ステップS 1 9 1でサーバ2 4とのデータ通信を確立する。つまり、データ通信処理を実行することで、サーバ2 4とのデータ通信を確立する。続いて、ステップS 1 9 3では、相手の状態データを取得する。つまり、サーバ2 4に記憶されるユーザBの状態データを、サーバ2 4とのデータ通信によって取得する。そして、取得された相手の状態データは、データ通信バッファ3 4 4に一時的に記憶される。

10

【0 1 8 1】

続いて、ステップS 1 9 5では、相手がモニタ状態か否かを判断する。つまり、データ通信バッファ3 4 4に一時的に記憶されたユーザBの状態データに基づいて、ユーザBの興味対象がモニタ1 6 bであるか否かを判断する。ステップS 1 9 5で“ N O ”であれば、つまりユーザBの興味対象がモニタ1 6 bでなければ、ステップS 2 0 7に進む。一方、ステップS 1 9 5で“ Y E S ”であれば、つまりユーザBの興味対象がモニタ1 6 bであれば、ステップS 1 9 7で相手がアクティブか否かを判断する。つまり、取得されたユーザBの状態データに基づいて、ユーザBがアクティブ状態であるか否かを判断する。ステップS 1 9 7で“ N O ”であれば、つまりユーザBがパッシブ状態であれば、ステップS 2 0 7に進む。

20

【0 1 8 2】

一方、ステップS 1 9 7で“ Y E S ”であれば、つまりユーザBがアクティブ状態であれば、ステップS 1 9 9で相手がトークであるか否かを判断する。つまり、ユーザBの状態データに基づいて、ユーザBがトーク状態であるか否かを判断する。ステップS 1 9 9で“ Y E S ”であれば、つまりユーザBがトーク状態であればステップS 2 0 1で発話継続処理を実行する。そして、発話継続処理が終了すれば、アクティブトーク処理を終了して、全体処理に戻る。なお、この発話継続処理については、図3 2に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

30

【0 1 8 3】

また、ステップS 1 9 9で“ N O ”であれば、つまりユーザBがリッスン状態であれば、ステップS 2 0 3でユーザと相手との状態データを一時記憶する。つまり、ユーザAの状態を示す状態データ3 5 4と、データ通信バッファ3 4 4に記憶されるユーザBの状態データとを状態データバッファ3 4 8に一旦格納する。続いて、ステップS 2 0 5では、発声時間が閾値Lより短いかなんかを判断する。つまり、発声カウンタ3 6 2によってカウントされたユーザAの発声時間が閾値Lより短いかなんかを判断する。

40

【0 1 8 4】

ステップS 2 0 5で“ Y E S ”であれば、つまりユーザAの発声時間が閾値Lよりも短ければ、ステップS 2 0 1に進む。一方、ステップS 2 0 5で“ N O ”であれば、つまりユーザAの発声時間が閾値L以上であれば、ユーザAが一方向的に話しすぎている状態のため、ステップS 2 0 7で発話抑制処理を実行する。そして、ステップS 2 0 7の処理が終了すれば、アクティブトーク処理を終了して、全体処理に戻る。なお、発話抑制処理については、図3 3に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0 1 8 5】

以下、図2 9 - 図3 1に示すフロー図において、図2 8のフローと重複するフローについては、詳細な説明を省略する。

50

【0186】

図29にはアクティブリスンプログラム316cの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS173の処理が実行されると、ステップS211でサーバ24とのデータ通信を確立し、ステップS213では相手の状態データを取得する。続いて、ステップS215では、相手がモニタ状態であるか否かを判断する。ステップS215で“NO”であれば、つまりユーザBの興味対象がモニタ16b以外であれば、ステップS227に進む。

【0187】

ステップS215で“YES”であれば、つまりユーザBの興味対象がモニタ16bであれば、ステップS217で相手がアクティブか否かを判断する。ステップS217で“NO”であれば、つまりユーザBがパッシブ状態であれば、ステップS227に進む。一方、ステップS217で“YES”であれば、つまりユーザBがアクティブ状態であれば、ステップS219で相手がトークであるか否かを判断する。ステップS219で“NO”であれば、つまりユーザBがリスン状態であれば、ステップS227に進む。一方、ステップS219で“YES”であれば、つまりユーザBがトーク状態であれば、ステップS221でユーザと相手との状態データを一時記憶する。

【0188】

続いて、ステップS223では発声時間が閾値Lよりも短いかなかを判断する。つまり、発声カウンタ362によってカウントされたユーザBの発声時間が、閾値Lよりも短いかなかを判断する。ステップS223で“YES”であれば、つまりユーザBの発声時間が閾値Lよりも短ければ、ステップS225で発話継続処理を実行する。そして、ステップS225の処理が終了すれば、アクティブリスン処理を終了して、全体処理に戻る。一方、ステップS223で“NO”であれば、つまりユーザBの発声時間が閾値L以上であれば、相手ユーザであるユーザBが一方向的に話しすぎている状態のため、ステップS227で発話促進処理を実行する。また、ステップS227の処理が終了すれば、アクティブリスン処理を終了して、全体処理に戻る。なお、発話促進処理については、図34に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0189】

そして、ステップS207の発話抑制処理およびステップS227の発話促進処理が実行されると、ロボット10aはユーザAの発話を制御するように動作する。したがって、ユーザAとユーザBとがバランス良く発話するようになるため、対話が持続するようになる。

【0190】

図30には非アクティブプログラム316dの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS177の処理が実行されると、ステップS231で発話促進処理を実行する。そして、ステップS231の処理が終了すれば、非アクティブ処理を終了して、全体処理に戻る。たとえば、対話に非積極的なユーザの発話を促進することで、ユーザを対話に参加させ、対話を持続させる。

【0191】

図31にはアザープログラム316eの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS179の処理が実行されると、ステップS241で注意引きつけ処理を実行する。そして、ステップS241の処理が終了すると、アザー処理を終了して、全体処理に戻る。

【0192】

なお、注意引きつけ処理の詳細な説明は省略するが、この処理が実行されると、ロボット10aは、ユーザAの注意を引きつけるように動作する。つまり、ユーザAの興味対象がモニタ16（ユーザB）でも、ロボット10aでなければ、ロボット10aは、ユーザAが対話に興味を持つように注意を引きつける。このように、アクティブ状態ではないユーザに、対話へ興味を持たせることで、対話が持続するようになる。

【0193】

図32には発話継続プログラム316fの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS201またはステップS225の処理が実行されると、ステップS261で行動データを参照する。つまり、ステップS261では、ユーザAの行動データを、行動テーブルデータ352から参照する。続いて、ステップS263では、発話しているか否かを判断する。つまり、ユーザAの行動データにおいて、発話判定結果が「あり」であるか否かを判断する。ステップS263で“YES”であれば、つまりユーザAが発話していれば、ステップS265で傍参与者的疑似傾聴処理を実行する。また、ステップS265の処理が終了すると、発話継続処理を終了して、上位ルーチンに戻る。なお、傍参与者的疑似傾聴処理については、図36のフロー図を用いて後述するため、ここでの詳細な説明は省略する。

10

【0194】

また、ステップS263で“NO”であれば、つまりユーザAが発話していなければ、ステップS267で、ステップS191と同様に、サーバ24とのデータ通信を確立する。続いて、ステップS269では相手の行動データを取得する。つまり、ユーザBの行動データをサーバ24から取得する。続いて、ステップS271では、相手の行動データを一時記憶する。つまり、取得されたユーザBの行動データを相手行動データバッファ346に一旦格納する。

【0195】

続いて、ステップS273では、相手が発話しているか否かを判断する。つまり、相手行動データバッファ346に記憶されている、ユーザBの行動データを読み出し、発話判定結果が「あり」となっているか否かを判断する。

20

【0196】

ステップS273で“YES”であれば、つまりユーザBが発話していれば、ステップS275で第1積極的疑似傾聴処理を実行する。また、ステップS275の処理が終了すれば、発話継続処理を終了して上位ルーチンに戻る。なお、第1積極的疑似傾聴処理については、図37に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。一方、ステップS273で“NO”であれば、つまりユーザBが発話していなければ、ステップS277で直前の動作を継続し、処理が終了すれば、発話継続処理を終了して上位ルーチンに戻る。たとえば、ステップS277では、前回の処理で「ユーザAを見る」の動作命令が発行されていれば、今回の処理でも同じ動作命令を発行する。

30

【0197】

なお、傍参与者的疑似傾聴処理および第1積極的疑似傾聴処理では、ユーザAまたはユーザBに対して、ロボット10が疑似傾聴を行うように動作命令が付与される。つまり、ユーザAおよびユーザBは、ロボット10の疑似傾聴によって、自身の話を聴いてもらっているように感じることができるよう、対話が持続するようになる。

【0198】

以下、図33、図34に示すフロー図において、図32のフローと重複するフローについては、詳細な説明を省略する。

【0199】

図33には発話抑制プログラム316gの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS207の処理が実行されると、ステップS281で行動データを取得する。続いて、ステップS283では発話しているか否かを判断する。ステップS283で“YES”であれば、つまりユーザAが発話していれば、ステップS285で傍参与者的疑似傾聴処理を実行し、ステップS285の処理が終了すれば、発話抑制処理を終了してアクティブトーク処理に戻る。

40

【0200】

一方、ステップS283で“NO”であれば、つまりユーザAが発話していなければ、ステップS287でサーバ24とのデータ通信を確立し、ステップS289で相手の行動データを取得する。そして、ステップS291では相手の行動データを一時記憶する。続いて、ステップS293では、相手が発話しているか否かを判断する。ステップS293

50

で“YES”であれば、つまりユーザBが発話すれば、ステップS295で第1積極的疑似傾聴処理を実行する。また、ステップS295の処理が終了すれば、発話抑制処理を終了してアクティブトーク処理に戻る。

【0201】

また、ステップS293で“NO”であれば、つまりユーザBが発話していなければ、ステップS297でユーザ発話抑制処理を実行する。そして、ステップS297の処理が終了すれば、発話抑制処理を終了してアクティブトーク処理に戻る。

【0202】

なお、ユーザ発話抑制処理の詳細な説明は後述するが、この処理が実行されると、ロボット10はユーザの発話を抑制するように動作する。つまり、ユーザAが一方的に発話している場合には、ユーザAの発話が抑制される。

10

【0203】

図34には発話促進プログラム316hの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS227またはステップS231の処理が実行されると、ステップS301で行動データを参照する。続いて、ステップS303では発話しているか否かを判断する。ステップS303で“YES”であれば、つまりユーザAが発話していれば、ステップS305で傍参与者的疑似傾聴処理を実行する。一方、ステップS303で“NO”であれば、つまりユーザAが発話していなければ、ステップS307でサーバ24とのデータ通信を確立し、ステップS309で相手の行動データを取得する。そして、ステップS311では、相手の行動データを一時記憶する。続いて、ステップS313では、相手が発話しているか否かを判断する。ステップS313で“YES”であれば、つまりユーザBが発話していれば、ステップS315で第1積極的疑似傾聴処理を実行する。

20

【0204】

また、ステップS313で“NO”であれば、つまりユーザBが発話していなければステップS317でユーザがロボット10を見ているか否かを判断する。つまり、ユーザAの行動データで、視線方向判定結果が「ロボット」であるか否かを判断する。

【0205】

ステップS317で“YES”であれば、つまりユーザAがロボット10aを見ていれば、ステップS319でユーザ発話促進処理を実行する。そして、ステップS319の処理が終了すれば、発話促進処理を終了して、上位ルーチンに戻る。

30

【0206】

なお、ユーザ発話促進処理の詳細な説明は後述するが、この処理が実行されると、ロボット10aはユーザAが発話するように動作する。たとえば、上位ルーチンが非アクティブ処理であれば、ユーザAが会話に参加できるように発話を促進、つまり促すようにロボット10が動作する。これにより、対話に非積極的なユーザが発話するため、ユーザAが対話に参加するようになる。

【0207】

また、ステップS207、S227、S297およびS319の処理を実行するプロセッサ64は、発話制御動作付与手段として機能する。さらに、ステップS269、S289およびS309の処理を実行するプロセッサ64は取得手段として機能する。

40

【0208】

一方、ステップS317で“NO”であれば、つまりユーザAがロボット10aを見ていなければ、ステップS321で注意引きつけ処理を実行する。そして、ステップS321の処理が終了すれば、発話促進処理を終了して、上位ルーチンに戻る。

【0209】

なお、注意引きつけ処理の詳細な説明は後述するが、この処理が実行されると、ロボット10aはユーザAの注意を引きつけるように動作する。たとえば、上位ルーチンが非アクティブ処理であれば、ユーザAの注意を引きつけることで、ユーザAが会話に参加できるようにロボット10が動作する。これにより、対話に非積極的なユーザは注意を引きつ

50

けられ、対話に参加する。

【0210】

なお、ステップS177, S231, S319およびS321の処理を実行するプロセッサ64は参加動作付与手段として機能する。

【0211】

ここで、発話促進処理の上位ルーチンがアクティブリスン処理である場合に、ステップS319またはステップS321の処理でユーザの発話を促進したり、注意を引きつけたりするようにロボット10が動作することで、ユーザAが発話するようになる。つまり、ユーザが相手の話を積極的に傾聴している状態が長く続けば、ユーザAが発話が求められる。

10

【0212】

図35には注意引きつけプログラム316jの処理を示すフロー図が示される。PC14aのプロセッサ64は、たとえば、ステップS251の処理が実行されると、ステップS331で乱数生成処理を実行する。たとえば、ステップS331では乱数生成処理では桁数の少ない疑似乱数を1つ生成し、生成した乱数を乱数バッファ350に格納する。

【0213】

続いて、ステップS333では、乱数が奇数であるか否かを判断する。つまり、乱数バッファ350に格納される乱数を「2」で割ったときの余りが「1」であるか否かを判断する。ステップS333で“YES”であれば、つまり乱数が奇数であれば、ステップS335で「ユーザを見る」の動作命令を発行する。つまり、ユーザAを見るような動作命令をロボット10aに付与する。そして、ステップS335の処理が終了すれば、注意引きつけ処理を終了して上位ルーチンに戻る。

20

【0214】

また、ステップS333で“NO”であれば、つまり乱数が偶数であればステップS337で、再び乱数生成処理を実行する。また、ステップS337の処理で作成された乱数は、ステップS331で作成された乱数を消去した後に記憶される。つまり、ステップS337の処理が実行されると、乱数バッファ350に記憶される乱数が更新される。

【0215】

続いて、ステップS339では、乱数が奇数か否かを判断する。つまり、ステップS337で生成された乱数が奇数であるか否かを判断する。ステップS339で“YES”であれば、つまり乱数が奇数であれば、ステップS341で「ユーザに声をかける」の動作命令を発行する。たとえば、発行される動作命令は、「ねえねえ」などの合成音声を、口32に設けられたスピーカ56から出力させる。そして、ユーザAは、合成音声を聴くことで、ロボット10aに話しかけられたように感じる。

30

【0216】

また、ステップS339で“NO”であれば、つまり乱数が偶数であれば、ステップS343で「ユーザと同じものを見る」の動作命令を発行する。つまり、ユーザAの行動データにおける、視線判定結果に基づいて、ロボット10aの頭部モータ48および眼球モータ50を駆動させる動作命令を、ロボット10aに付与する。

【0217】

たとえば、ユーザAの視線判定結果が「ロボット」であれば、ロボット10aが自身の胴体28を見るように、頭部モータ48および眼球モータ50を駆動させる動作命令を発行する。また、視線判定結果が「モニタ」であれば、ロボット10aがモニタ16aを見るように、頭部28を動作させる動作命令を発行する。そして、視線判定結果が「アザー」であれば、ロボット10aが周囲をきょろきょろするように、頭部28を動作させる動作命令を発行する。なお、視線判定結果が4パターン以上ある場合には、さらに動作命令が多様化する。また、行動データにおける視線判定結果が空間座標で示される場合には、ロボット10aの空間座標位置と視線判定結果の空間座標位置とに基づいて、ユーザAが注視する対象をより正確に判断して、見るようにしてもよい。

40

【0218】

50

このように、注意引きつけ処理では、ユーザを見たり、ユーザに声を掛けたり、ユーザと同じものを見たりするようにロボット10を動作させる動作命令がランダムに発行される。そして、このようにロボット10が動作することで、ユーザの注意が引きつけられる。

【0219】

なお、右腕30Lまたは左腕30Lを動かしてユーザAの注意を引きつけてもよい。また、ステップS321、S335、S341およびS343の処理を実行するプロセッサ64は注意引付手段として機能する。

【0220】

以下、図36 - 図40に示すフロー図において、図35のフローと重複するフローについては、詳細な説明を省略する。

【0221】

図36には傍参与者的疑似傾聴プログラム316kの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS361で行動データを参照する。続いて、ステップS363では、話し始めかを判断する。たとえば、発声カウンタ362で発声時間がカウントされる場合には、発声カウンタ362の値が所定の値（たとえば、2秒を示す値）よりも小さいか否かを判断する。また、行動テーブルにおける発話判定結果に基づいて発声時間が計測されている場合には、発話判定結果が「なし」から「あり」となってから所定回数（たとえば、2回）以内であるか否かを判断する。

【0222】

ステップS363で“YES”であれば、つまりユーザAが話し始めていれば、ステップS367で第2積極疑似傾聴処理を実行する。そして、ステップS367の処理が終了すれば、傍参与者的疑似傾聴処理を終了して、上位ルーチンに戻る。なお、第2積極的疑似傾聴処理については、図39に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0223】

また、ステップS363で“NO”であれば、ステップS365でロボット10aを見ているか否かを判断する。つまり、参照されたユーザAの行動データで、視線方向判定が「ロボット」であるか否かを判断する。ステップS365で“YES”であれば、つまりユーザAがロボット10aを見ていれば、ステップS367に進む。一方、ステップS365で“NO”であれば、つまりユーザAがロボット10aを見ていなければ、ステップS369で乱数生成処理を実行し、ステップS371で生成された乱数が奇数であるか否かを判断する。ステップS371で“NO”であれば、つまり乱数が偶数であれば、ステップS367に進む。一方、ステップS371で“YES”であれば、つまり乱数が奇数であれば、ステップS373で第1積極的疑似傾聴処理を実行する。そして、ステップS373の処理が終了すると、傍参与者的疑似傾聴処理を終了して、上位ルーチンに戻る。

【0224】

このように、傍参与者的疑似傾聴処理では、ユーザBに対して疑似傾聴を行う第1疑似傾聴処理と、ユーザAに対して疑似傾聴を行う第2疑似傾聴処理とが、ユーザAの発話や、視線に基づいて選択される。

【0225】

図37には第1積極的疑似傾聴プログラム316mの処理を示すフロー図が示される。PC14aのプロセッサ64は、たとえばステップS373の処理が実行されると、ステップS381で、乱数生成処理を実行し、ステップS383では生成された乱数が奇数であるか否かを判断する。ステップS383で“YES”であれば、つまり乱数が奇数であれば、ステップS385で「モニタ16aに顔を向ける」の動作命令を発行する。つまり、ユーザBの映像が表示されるモニタ16aの方向に、頭部28の頭部方向および眼球34の視線方向が向くように、頭部モータ48および眼球モータ50を駆動させる動作命令を、ロボット10aに付与する。そして、ステップS385の処理が終了すれば、第1積極的疑似傾聴処理を終了して、上位ルーチンに戻る。

【 0 2 2 6 】

また、ステップ S 3 8 3 で “ N O ” であれば、つまり乱数が偶数であれば、ステップ S 3 8 7 で再び乱数生成処理を実行し、ステップ S 3 8 9 で乱数が奇数であるか否かを判断する。ステップ S 3 8 9 で “ Y E S ” であれば、つまり乱数が奇数であればステップ S 3 9 1 で「モニタ 1 6 a 側に体を傾ける」の動作命令を発行する。つまり、体（頭部 2 6 および胴体 2 8 など）がモニタ 1 6 a に傾くように、腰モータ 5 2 を駆動させる動作命令を、ロボット 1 0 a に付与する。そして、ステップ S 3 9 1 の処理が終了すれば、第 1 積極的疑似傾聴処理を終了して、上位ルーチンに戻る。

【 0 2 2 7 】

さらに、ステップ S 3 8 9 で “ N O ” であれば、つまり乱数が偶数であれば、ステップ S 3 9 3 で「相槌をうつ」の動作命令を発行する。たとえば、人間の相槌のように、スピーカ 5 6 から「あー」などの合成音声を出力するとともに、頭部 2 6 がモニタ 1 6 a 側に傾くように頭部モータ 4 8 を駆動させる動作命令を、ロボット 1 0 a に付与する。そして、ステップ S 3 9 3 の処理が終了すれば、第 1 積極的疑似傾聴処理を終了して、上位ルーチンに戻る。

【 0 2 2 8 】

このように、第 1 積極的疑似傾聴処理では、相手ユーザの映像が映るモニタ 1 6 に対して、ロボット 1 0 の顔（頭部 2 6 および眼球 3 4 の方向）を向けたり、ロボット 1 0 の体を傾けたり、相槌をうったりすることで、相手ユーザに対して傾聴しているかのようにロボット 1 0 が動作する。そのため、たとえユーザ A がユーザ B の話を傾聴していなくても、ロボット 1 0 a がユーザ B の話を疑似傾聴することで、ユーザ B は自身の話を聴いてもらっているように感じることができる。

【 0 2 2 9 】

なお、ステップ S 2 7 5 , S 2 9 5 , S 3 1 5 , S 3 7 3 , S 3 8 5 , S 3 9 1 および S 3 9 3 の処理を実行するプロセッサ 6 4 は相手傾聴動作付与手段として機能する。

【 0 2 3 0 】

図 3 8 にはユーザ発話抑制プログラム 3 1 6 n の処理を示すフロー図が示される。P C 1 4 a のプロセッサ 6 4 は、ステップ S 2 9 7 の処理が実行されると、ステップ S 4 0 1 で、行動データを参照する。続いて、ステップ S 4 0 3 では、ステップ S 3 6 5 と同様に、ロボット 1 0 a を見ているか否かを判断する。ステップ 4 0 3 で “ N O ” であれば、つまりユーザ A がロボット 1 0 a を見ていなければ、ステップ S 4 0 5 で注意引きつけ処理を実行する。そして、ステップ S 4 0 5 の処理が終了すれば、ユーザ発話抑制処理を終了して、発話抑制処理に戻る。なお、ステップ S 1 7 9 , S 2 4 1 , S 3 2 1 , S 3 3 5 , S 3 4 1 , S 3 4 3 および S 4 0 5 の処理を実行するプロセッサ 6 4 は興味動作付与手段として機能する。

【 0 2 3 1 】

また、ステップ S 4 0 3 で “ Y E S ” であれば、つまりユーザ A がロボット 1 0 a を見ていなければ、ステップ S 4 0 7 で乱数生成処理を実行し、ステップ S 4 0 9 で乱数が奇数か否かを判断する。ステップ S 4 0 9 で “ Y E S ” であれば、つまり乱数が奇数であれば、ステップ S 4 1 1 で「ユーザを見る」の動作命令を発行する。つまり、ユーザ A の方向に、頭部 2 8 および眼球 3 4 の方向が向くように、頭部モータ 4 8 および眼球モータ 5 0 を駆動させる動作命令を、ロボット 1 0 a に付与する。そして、ステップ S 4 1 1 の処理が終了すれば、ユーザ発話抑制処理を終了して、発話抑制処理に戻る。

【 0 2 3 2 】

一方、ステップ S 4 0 9 で “ N O ” であれば、ステップ S 4 1 3 で乱数生成処理を再度実行し、ステップ S 4 1 5 で乱数が奇数であるか否かを判断する。ステップ S 4 1 5 で “ Y E S ” であれば、つまり乱数が奇数であれば、ステップ S 4 1 7 で「発話を止める」の動作命令を発行する。つまり、ステップ S 4 1 7 では、発話を止めさせるような合成音声をスピーカ 5 6 から出力させる動作命令を、ロボット 1 0 a に付与する。また、発話を止めさせるような合成音声とは「ちょっと待って」などである。そして、ステップ S 4 1 7

の処理が終了すれば、ユーザ発話抑制処理を終了して、発話抑制処理に戻る。

【0233】

さらに、ステップS415で“NO”であれば、つまり乱数が偶数であれば、ステップS419で再び乱数生成処理を実行し、ステップS421で乱数が奇数であるか否かを判断する。ステップS421で“YES”であれば、つまり生成された乱数が奇数であればステップS423で「モニタ16aを指し示す」の動作命令を発行する。つまり、左腕30Lまたは右腕30Rの先端がモニタ16aの方を指し示すように、右腕モータ46Rまたは左腕モータ46Lを駆動させる動作命令を、ロボット10aに付与する。そして、ステップS423の処理が終了すれば、ユーザ発話抑制処理を終了して、発話抑制処理に戻る。

10

【0234】

そして、ステップS421で“NO”であれば、つまり生成された乱数が偶数であれば、ステップS425で注意誘導処理を実行する。また、ステップS425の処理が終了すれば、ユーザ発話抑制処理を終了して、発話抑制処理に戻る。なお、注意誘導処理については、図40に示すフロー図を用いて後述するため、ここでの詳細な説明は省略する。

【0235】

このように、ユーザ発話抑制処理では、積極的に発話するユーザAに対して、ロボット10aがユーザAを見たり、発話を止める音声を発したり、モニタ16aを指示したり、注意を引きつけたり、注意を誘導したりすることで、発話を抑制する。したがって、一方的に話すユーザの注意が誘導されるため、発話が抑制される。また、ユーザの注意が相手に誘導される場合には、相手が発話する機会を得ることができるため、対話がより持続するようになる。

20

【0236】

なお、ステップS297、S405、S411、S417、S423およびS425の処理を実行するプロセッサ64は発話抑制動作付与手段として機能する。

【0237】

図39には第2積極的疑似傾聴プログラム316pの処理を示すフロー図が示される。PC14aのプロセッサ64は、ステップS367の処理が実行されると、ステップS431で乱数生成処理を実行し、ステップS433で乱数が奇数か否かを判断する。ステップS433で“YES”であれば、つまり乱数が奇数であれば、ステップS435で「ユーザに顔を向ける」の動作命令を発行する。つまり、つまりユーザAの方向に、顔（頭部26および眼球34）の方向が向くように、頭部モータ48および眼球モータ50を駆動させる動作命令を、ロボット10aに付与する。また、ステップS435の処理が終了すれば、第2積極的疑似傾聴処理を終了して、傍参与者的疑似傾聴処理に戻る。

30

【0238】

一方、ステップS433で“NO”であれば、つまり乱数が偶数であれば、ステップS437で乱数生成処理を実行し、ステップS439で乱数が奇数か否かを再び判断する。ステップS439で“YES”であれば、つまり乱数が奇数であれば、ステップS441で「ユーザA側に体を傾ける」の動作命令を発行する。つまり、ロボット10aの体（頭部26および胴体28）がユーザA側に傾くように、腰モータ52を駆動させる動作命令を、ロボット10aに付与する。

40

【0239】

また、ステップS439で“NO”であれば、つまり乱数が偶数であれば、ステップS443で「相槌をうつ」の動作命令を発行する。つまり、頭部26がユーザA側に傾くように頭部モータ48を駆動させ、人間の相槌のような合成音声をスピーカ56から出力させる動作命令を、ロボット10aに付与する。そして、ステップS443の処理が終了すれば、第2積極的疑似傾聴処理を終了して、傍参与者的疑似傾聴処理に戻る。

【0240】

このように、ロボット10aは、ユーザAに対して顔を向けたり、体を傾けたり、相槌をうったりすることで、ユーザAの発話を傾聴しているかのような動作を行う。

50

【 0 2 4 1 】

なお、ステップ S 1 7 1 , S 1 7 3 , S 2 0 1 , S 2 2 5 , S 2 6 5 , S 2 8 5 , S 3 0 5 , S 3 6 7 , S 4 3 5 , S 4 4 1 および S 4 4 3 の処理を実行するプロセッサ 6 4 は傾聴動作付与手段として機能する。

【 0 2 4 2 】

図 4 0 には注意誘導プログラム 3 1 6 q の処理を示すフロー図が示される。P C 1 4 a のプロセッサ 6 4 は、ステップ S 4 2 5 の処理が実行されると、ステップ S 4 5 1 で乱数生成処理を実行し、ステップ S 4 5 3 で乱数が奇数であるか否かを判断する。そして、ステップ S 4 5 3 で “ Y E S ” であれば、つまり乱数が奇数であれば、ステップ S 4 5 5 で「モニタ 1 6 a を見る」の動作命令を発行する。たとえば、ステップ S 4 5 5 では、モニタ 1 6 a の方向に、頭部 2 8 の頭部方向および眼球 3 4 の視線方向が向くように、頭部モータ 4 8 および眼球モータ 5 0 を駆動させる動作命令を、ロボット 1 0 a に付与する。そして、ステップ S 4 5 5 の処理が終了すれば、注意誘導処理を終了して、ユーザ発話抑制処理に戻る。

10

【 0 2 4 3 】

一方、ステップ S 4 5 3 で “ N O ” であれば、つまり乱数が偶数であれば、ステップ S 4 5 7 で乱数生成処理を再度実行し、ステップ S 4 5 9 で乱数が奇数であるか否かを判断する。そして、ステップ S 4 5 9 で “ Y E S ” であれば、つまり乱数が奇数であれば、ステップ S 4 5 5 に進む。一方、ステップ S 4 5 9 で “ N O ” であれば、つまり生成された乱数が偶数であれば、ステップ S 4 6 1 で、ステップ S 4 1 1 と同様に、「ユーザ A を見る」の動作命令を発行する。そして、ステップ S 4 6 1 の処理が終了すれば、注意誘導処理を終了して、ユーザ発話抑制処理に戻る。

20

【 0 2 4 4 】

このように、注意誘導処理が何度か実行されると、ユーザ A とモニタ 1 6 a とを交互に見るような動作を行うようになるため、ユーザ A の注意が誘導される。特に本実施例では、ロボット 1 0 a は、ユーザ A よりもモニタ 1 6 a を見る回数が増えるように設定されているため、ユーザ A の注意はモニタ 1 6 a 、つまりユーザ B に誘導される。また、他の実施例では、注意誘導処理を複数回繰り返すことで、より効果的にユーザ A の注意を誘導してもよい。

【 0 2 4 5 】

なお、ステップ S 4 2 5 , S 4 5 5 および S 4 6 1 の処理を実行するプロセッサ 6 4 は注意誘導手段として機能する。

30

【 0 2 4 6 】

図 4 1 にはユーザ発話促進プログラム 3 1 6 r の処理を示すフロー図が示される。P C 1 4 a のプロセッサ 6 4 は、ステップ S 3 1 9 の処理が実行されると、ステップ S 4 7 1 で「ユーザに質問する」の動作命令を発行する。たとえば、ステップ S 4 7 1 では、「今日の食事のことにについて聞いてみたら？」などの合成音声を、口 3 2 に設けられたスピーカ 5 6 から出力されるような動作命令を、ロボット 1 0 a に付与する。そして、ステップ S 4 7 1 の処理が終了すれば、ユーザ発話促進処理を終了して、発話促進処理に戻る。

【 0 2 4 7 】

このように、ユーザ A がユーザ B に対して質問をさせるようなことを、ロボット 1 0 a が発話することで、ユーザ B が発話するようになり、対話が持続するようになる。

40

【 0 2 4 8 】

なお、ステップ S 2 2 7 , S 3 1 9 および S 4 7 1 の処理を実行するプロセッサ 6 4 は発話促進動作付与手段として機能する。また、ステップ S 3 1 9 および S 4 7 1 の処理を実行するプロセッサ 6 4 は発話促し手段として機能する。

【 0 2 4 9 】

図 4 2 にはカメラ制御プログラム 3 1 8 の処理を示すフロー図が示される。P C 1 4 a のプロセッサ 6 4 は、テレビ電話機能によって通話が始めると、ステップ S 4 8 1 でモニタカメラ 2 2 a の画像を送信する。つまり、テレビ電話機能において利用される画像デー

50

タを、ネットワーク 200 を介して PC 14 b に送信する。続いて、ステップ S 483 では、興味対象認識結果を参照する。つまり、状態データ 354 に含まれる、興味対象の認識結果を参照する。

【0250】

続いて、ステップ S 485 では、アザー状態か否かを判断する。つまり、ユーザ A の興味対象の認識結果が「アザー」であるか否かを判断する。ステップ S 485 で“YES”であれば、つまりユーザ A がロボット 10 a またはモニタ 16 a を見ていなければ、ステップ S 483 に戻る。一方、ステップ S 485 で“NO”であれば、つまりユーザ A がロボット 10 a またはモニタ 16 a を見ていれば、ステップ S 487 でモニタ状態か否かを判断する。つまり、ユーザ A の興味対象がモニタ 16 であるか否かを判断する。

10

【0251】

ステップ S 487 で“YES”であれば、つまりユーザ A の興味対象がモニタ 16 a であれば、ステップ S 489 でモニタカメラ 22 a の画像を送信する。つまり、ユーザ A がモニタ 16 を見ることで、ユーザ A の顔がモニタカメラ 22 a によって正面から撮影されていれば、ステップ S 489 で、モニタカメラ 22 a の画像を PC 14 b に送信する。一方、ステップ S 487 で“NO”であれば、つまりユーザ A の興味対象がロボット 10 a であれば、ステップ S 491 で腹部カメラ 12 a の画像を送信する。つまり、ユーザ A がロボット 10 a を見ることで、ユーザ A の顔が腹部カメラ 12 a によって正面から撮影されていれば、ステップ S 491 で、腹部カメラ 12 a の画像を PC 14 b に送信する。

【0252】

20

なお、ステップ S 489 またはステップ S 491 の処理が終了すれば、ステップ S 483 に戻る。また、ステップ S 489 および S 491 の処理を実行するプロセッサ 64 は画像送信手段として機能する。

【0253】

これにより、たとえばユーザ A の顔が認識されている画像が PC 14 b に送信されるため、モニタ 16 b に表示される画像には必ずユーザ A の顔が表示されるようになる。そのため、ユーザ B は、たとえユーザ A がロボット 10 a に話しかけている状態であったとしても、ユーザ B 自身に話しかけられているように感じるため、対話が持続するようになる。

【0254】

30

この実施例によれば、傾聴対話持続システム 100 は、相手ユーザが表示されるモニタ 16、ユーザの音声が集音されるマイク 20 およびユーザを撮影するモニタカメラ 22 と接続される PC 14 と、腹部カメラ 12 が設けられるロボット 10 とを含む。

【0255】

PC 14 では、モニタカメラ 22 および腹部カメラ 12 によって撮影されたユーザの画像と、マイク 20 によって集音された音声とに基づいてユーザの行動が判定され、さらにメモリ 68 に記憶される。また、PC 14 では、第 1 所定時間分の行動データからユーザの状態を認識する。そして、たとえばユーザの状態が「アクティブ・トーク・モニタ」と認識されていれば、ロボット 10 は、ユーザに対して疑似傾聴を行うように、動作命令が PC 14 から付与される。また、ユーザの状態が「パッシブ・リッスン・モニタ」と認識されていれば、ロボット 10 はユーザに対して話しかけるなどして注意を引きつけ、ユーザを対話に参加させる。

40

【0256】

このように、ユーザの状態に応じてロボット 10 が対話を持続させるように動作するため、コミュニケーション障害のある者同士の対話を持続させることができる。

【0257】

なお、腹部カメラ 12 およびモニタカメラ 22 以外に、モニタ 16 およびロボット 10 以外の位置でユーザの顔を撮影する第 3 カメラを設置し、その第 3 カメラによってユーザの興味対象が「アザー」であるか否かの判定を補完的に行うようにしてもよい。つまり、第 3 カメラによって、ユーザの顔認識が失敗した状態と、ユーザの興味対象が「アザー」

50

である状態とを区別できるようにする。そして、たとえば顔認識が失敗したときの行動データを興味対象認識処理に反映させないようにすることで、傾聴対話持続システム 100 がユーザの興味対象を精度よく認識できるようにしてもよい。

【0258】

さらに、ユーザの視線方向を特定する第4カメラを設置し、その第4カメラによってユーザの視界にロボット10が入っているか否かの判定を補完的に行うようにしてもよい。そして、傾聴対話持続システム100では、上記判定結果に基づいて、ロボット10による発話や、ユーザの注意の引きつけの必要性が判断されてもよい。たとえば、PC14は、ユーザの視界にロボット10が入っていなければ、ロボット10を必ず発話するように動作させる。一方、PC14は、ユーザの視界にロボット10が入っていれば、ロボット10を必ずユーザの注意の引きつけるように動作させる。

10

【0259】

このように、カメラを3台以上設置することで、興味対象の認識精度を向上させたり、ロボット10をより効果的に動作させたりすることができる。なお、第3カメラおよび第4カメラの代わりに、1台のカメラで上記2つの判定が行われてもよい。また、より多くのカメラを設置することで、ユーザの行動の判定および状態の認識の精度を向上させたり、ロボット10に付与される動作命令の判断をより高度なものにしたりしてもよい。

【0260】

また、他の実施例は、PC14がロボット10に内蔵されていてもよく、この場合にロボット10は自律的に傾聴動作や発話機能を実行する。また、この場合、ロボット10は、ユーザのようなコミュニケーションの対象（コミュニケーション対象）との間で、身振り手振りのような身体動作および音声の少なくとも一方を含むコミュニケーション行動を実行する機能を有する相互作用指向のロボット（コミュニケーションロボット）であってもよい。

20

【0261】

また、PC14およびネットワーク200を利用せずに電話網などを介して、ユーザの画像と音声とがPC14に送受信されてもよい。また、電話機能をモニタ16などが有していれば、PC14を利用せずに、モニタ16、スピーカ18、マイク20およびモニタカメラ22のみでテレビ電話機の機能が実現されてもよい。そして、モニタ16、スピーカ18、マイク20およびモニタカメラ22が同一の筐体に組み込まれてもよい。さらに、この場合、ロボット10とテレビ電話機とが接続された状態で、テレビ電話の通話が開始される。

30

【0262】

また、PC14に代えて、サーバ24によってユーザの行動および状態が判定および認識され、ロボット10に対して動作命令を付与するようにしてもよい。この場合、腹部カメラ12およびモニタカメラ22によって撮影された画像と、マイク20によって集音された音声とはサーバ24に直接送信され、ロボット10はサーバ24からの動作命令に従って動作する。さらに、ユーザの状態を認識する処理だけが、サーバ24で実行されてもよい。つまり、図22 - 図25に示す処理がサーバ24のプロセッサ80によって実行され、PC14は第1所定時間毎に状態認識結果を取得する。

40

【0263】

また、傾聴対話持続システム100の他の実施例としては、ユーザAおよびユーザBの状態認識結果によらず、ユーザAの状態データのみに基づいてロボット10aが「疑似傾聴動作」、「発話制御動作」および「注意の引きつけの動作」を行ってもよい。たとえば、ユーザAが発話すれば、ロボット10aが疑似傾聴動作を行い、ユーザAの発声時間が長くなれば、ロボット10aがユーザAを見ることで発話を抑制する。そして、ユーザAが発話していなかったり、モニタ16aおよびロボット10aを見ていなかったりすれば、ロボット10aはユーザAの注意を引きつける動作を行う。

【0264】

そして、状態データは行動データと同様、図6に示すテーブルのような形式で記憶され

50

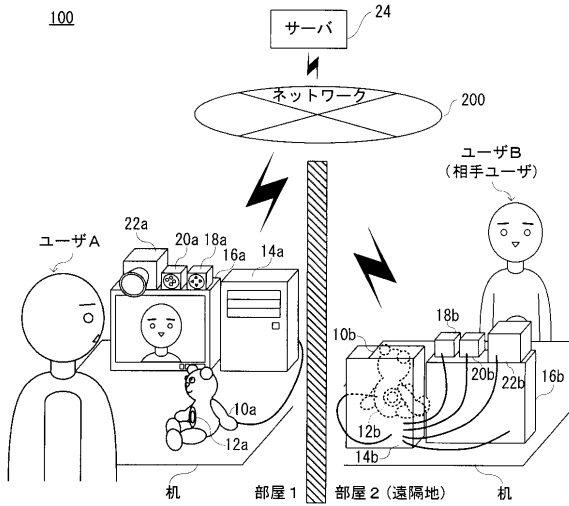
てもよい。また、状態データがテーブル形式で記憶される際には、ユーザの状態変化を示すログデータとして利用されてもよい。そして、状態データのテーブルに基づいて作成されるグラフが、モニタ 16 に表示されてもよい。

【符号の説明】

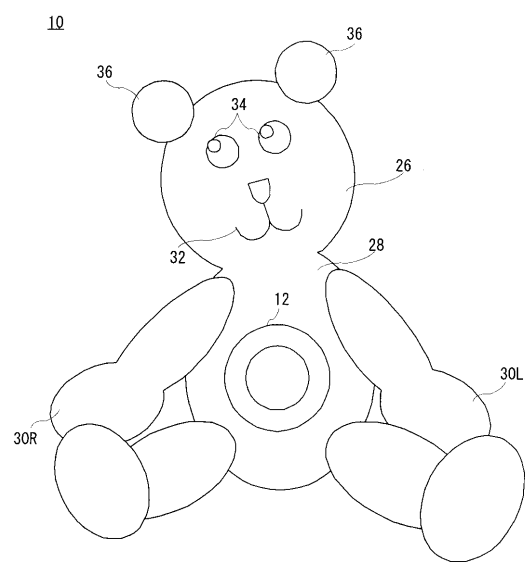
【0265】

10a, 10b	... ロボット	
12a, 12b	... 腹部カメラ	
14a, 14b	... PC	
16a, 16b	... モニタ	
20a, 20b	... マイク	10
22a, 22b	... モニタカメラ	
24	... サーバ	
26	... 頭部	
28	... 胴体	
30R	... 右腕	
30L	... 左腕	
32	... 口	
34	... 眼球	
46R	... 右腕モータ	
46L	... 左腕モータ	20
48	... 頭部モータ	
50	... 眼球モータ	
52	... 腰モータ	
56	... スピーカ	
64	... プロセッサ	
68	... メモリ	
70	... 視線サーバ	
72	... 通信LANボード	
70	... 無線通信装置	
100	... 傾聴対話持続システム	30
200	... ネットワーク	

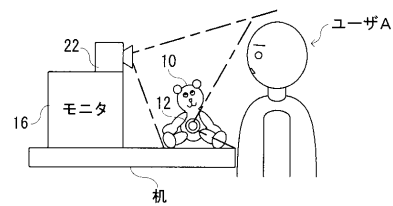
【図 1】



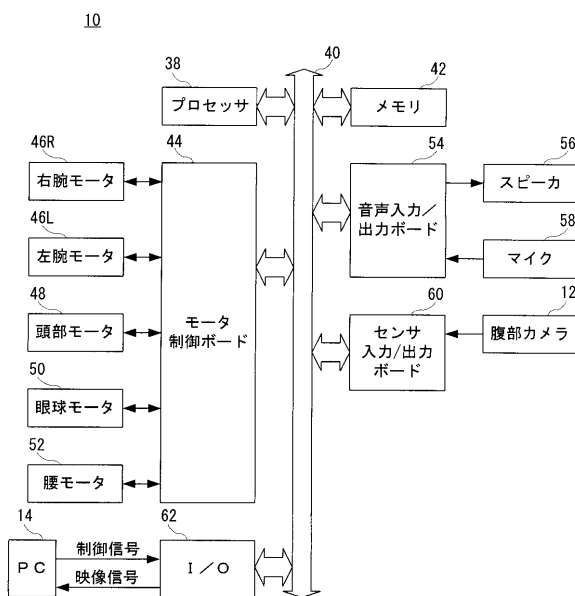
【図 3】



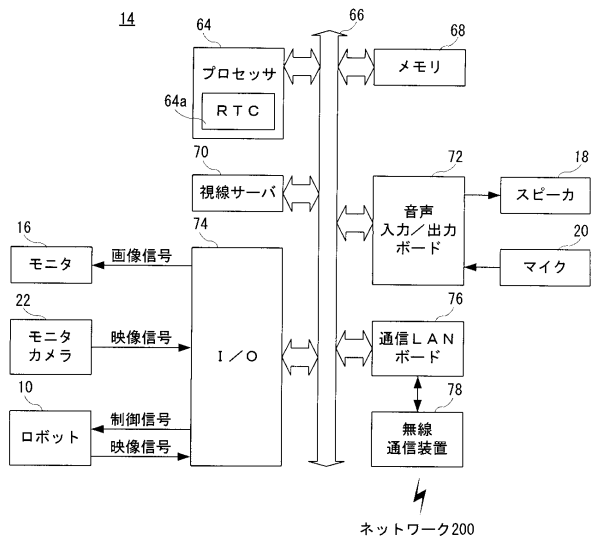
【図 2】



【図 4】



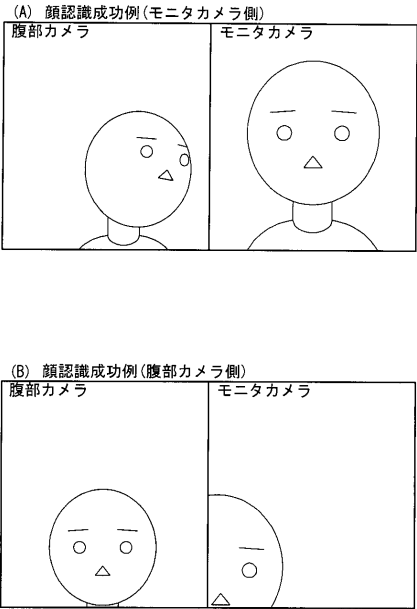
【図 5】



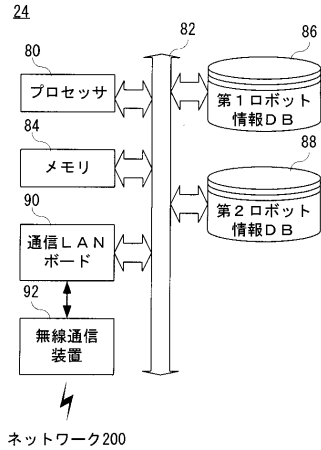
【図 6】

行動テーブル	時刻	前傾姿勢	発話	頭部方向 (腹部)	視線方向 (腹部)	顔認識 (腹部)	頭部方向 (モニタ)	視線方向 (モニタ)	傾き	相違
	10:00:30	なし	なし	アザー	アザー	失敗	アザー	アザー	あり	なし
	...									
	10:02:00	なし	なし	ロボット	ロボット	成功	ロボット	ロボット	あり	あり
	...									
	10:04:45	あり	あり	モニタ	モニタ	失敗	モニタ	モニタ	あり	あり
	...									
	...									
	...									
	...									

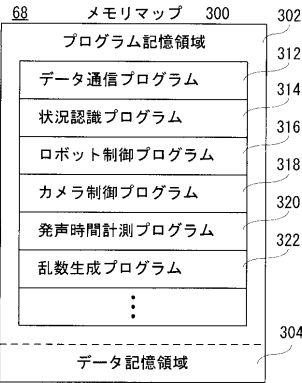
【図 7】



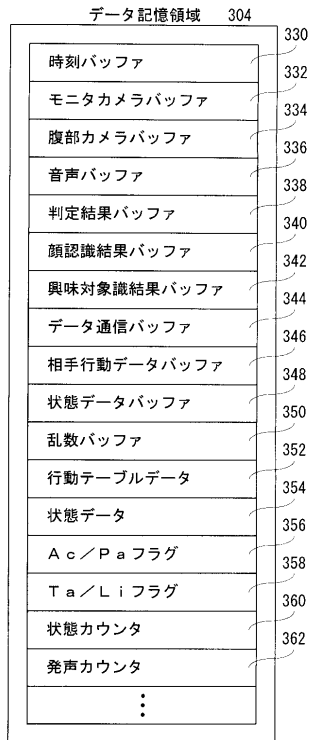
【図 8】



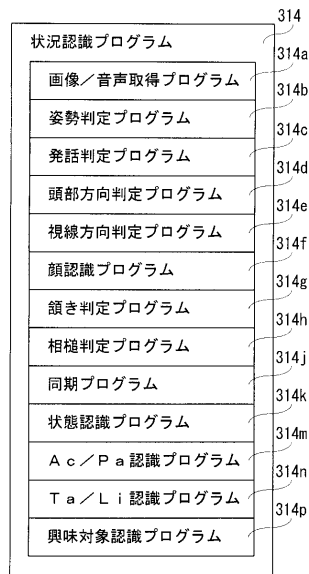
【図 9】



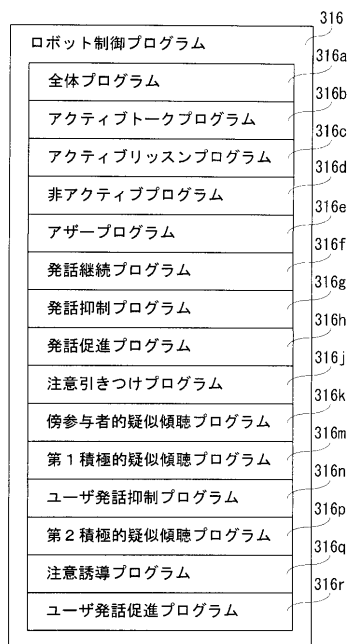
【図 10】



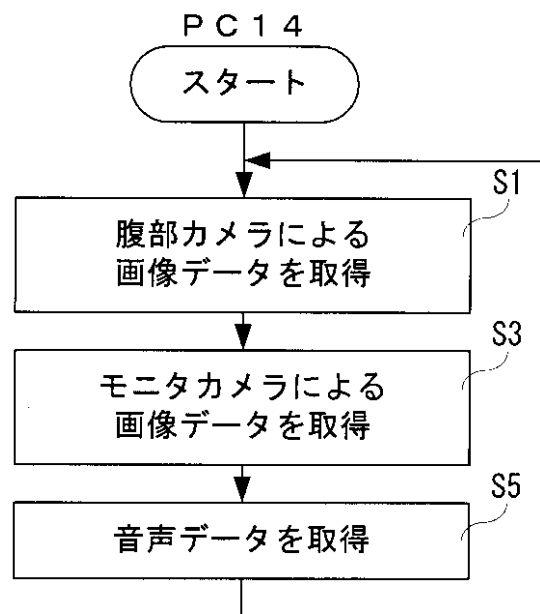
【図 11】



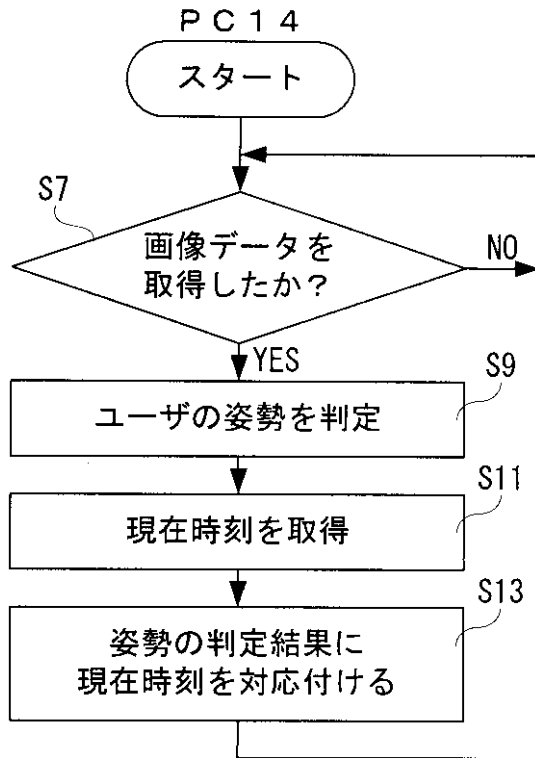
【図 12】



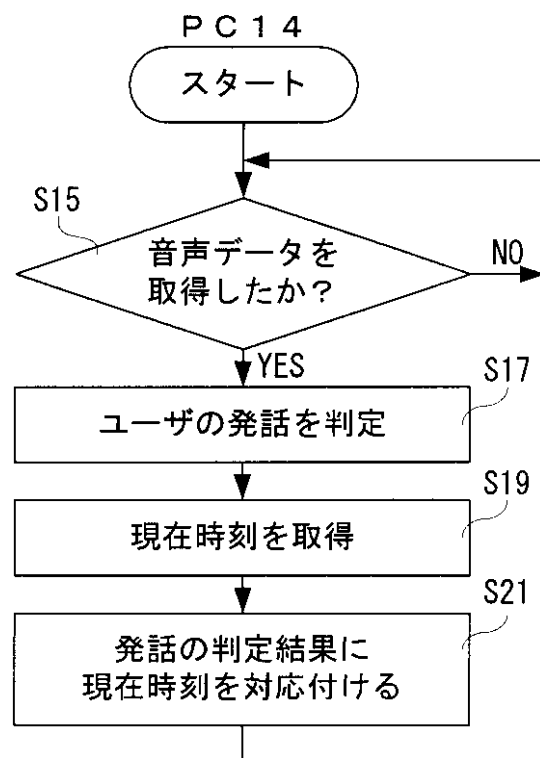
【図 13】



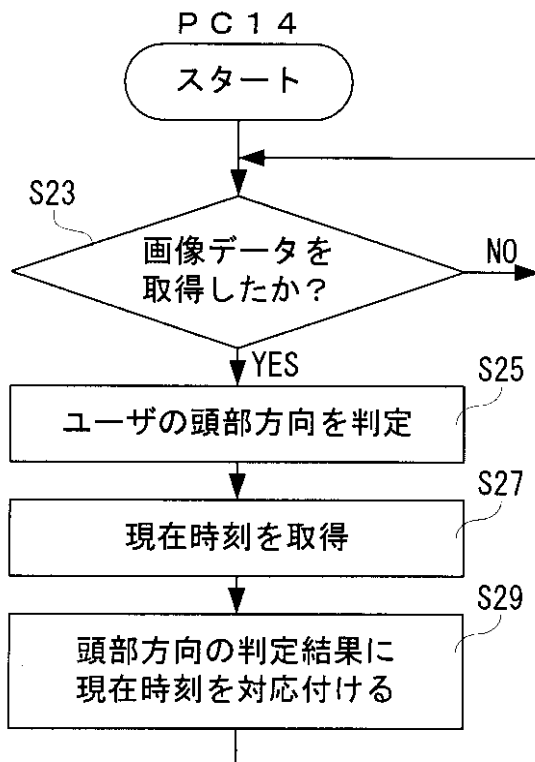
【図14】



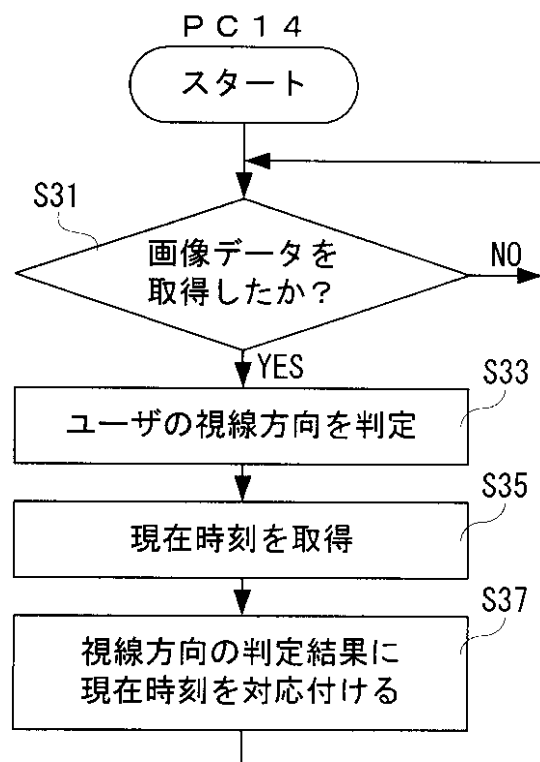
【図15】



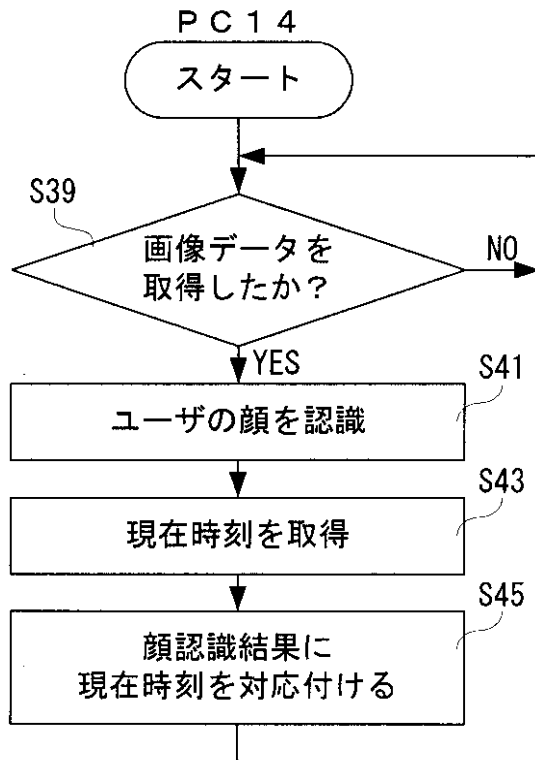
【図16】



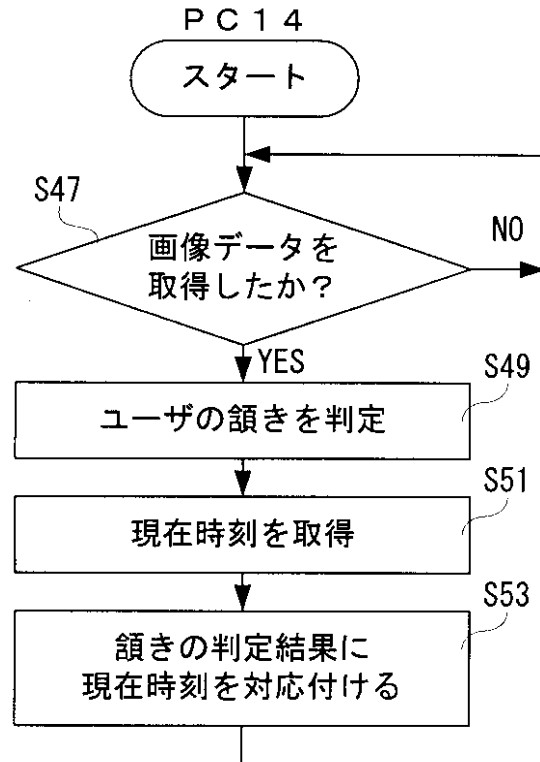
【図17】



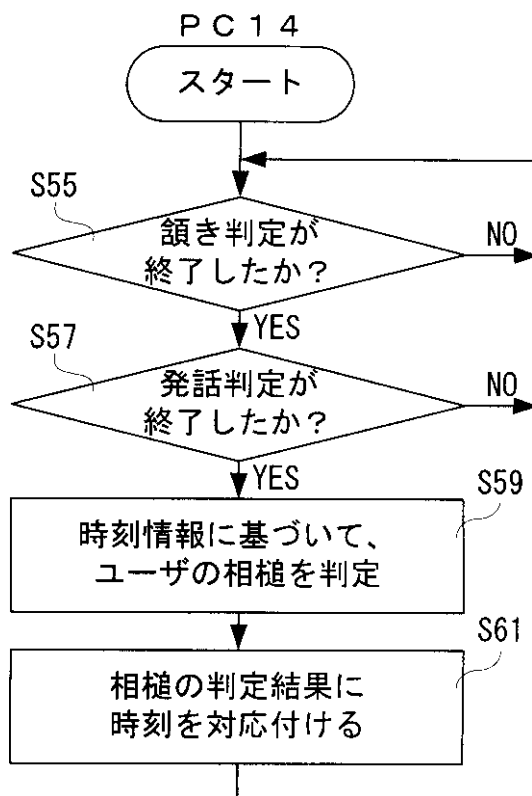
【図18】



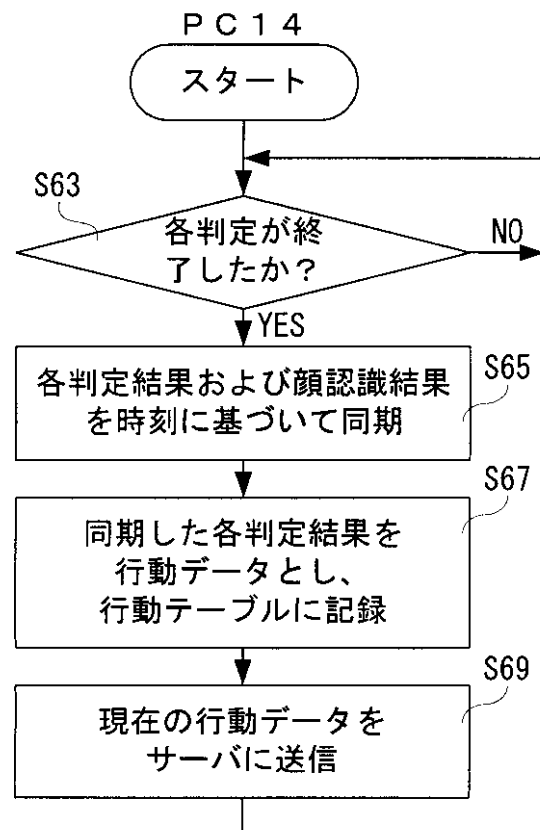
【図19】



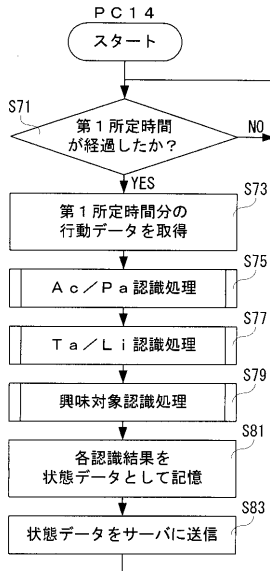
【図20】



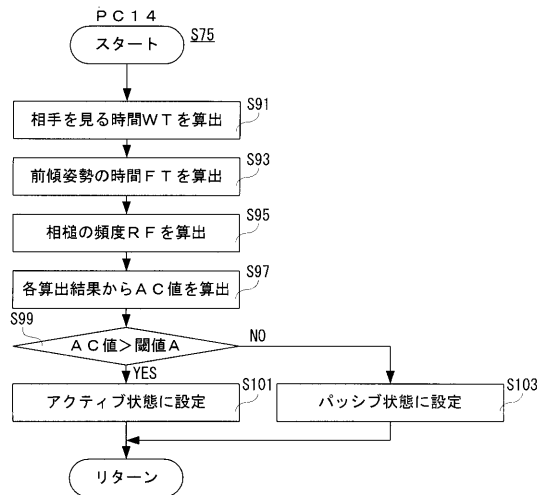
【図21】



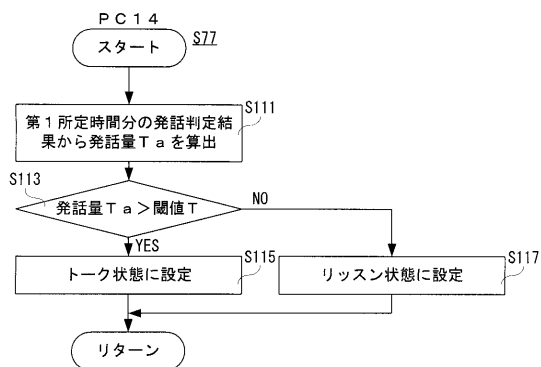
【図 22】



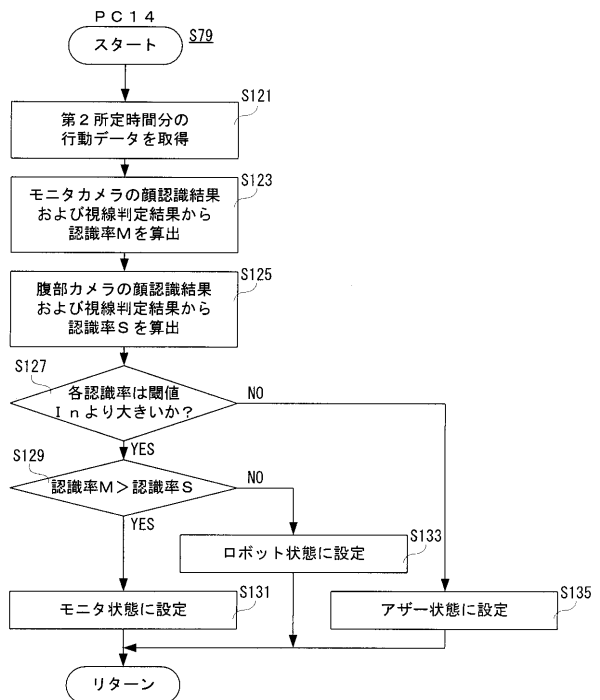
【図 23】



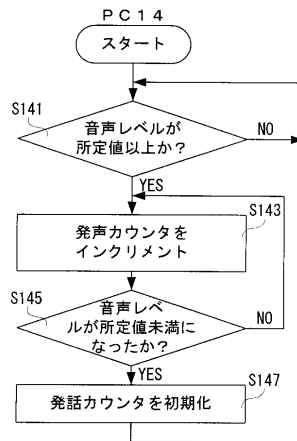
【図 24】



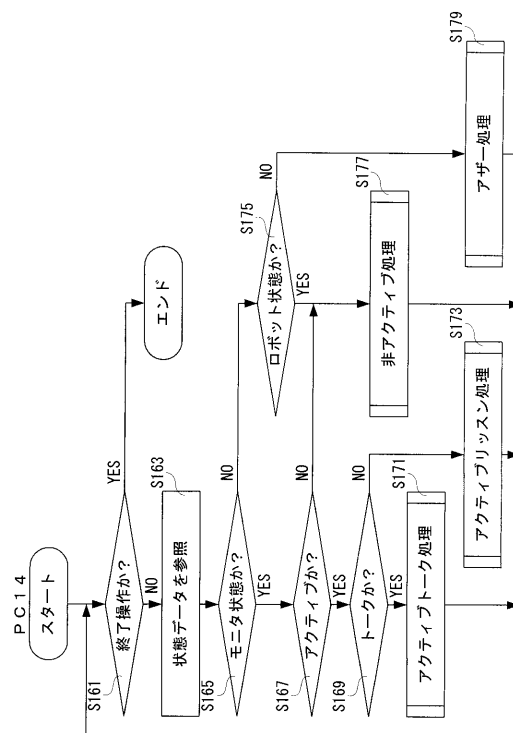
【図 25】



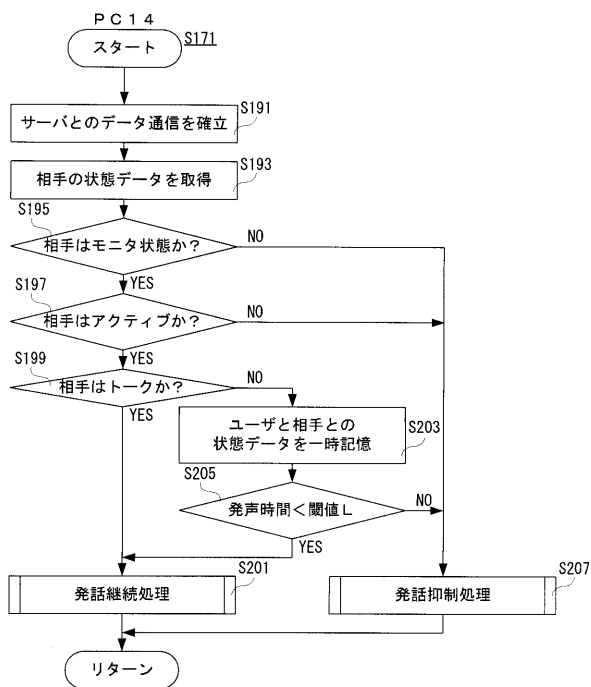
【図 26】



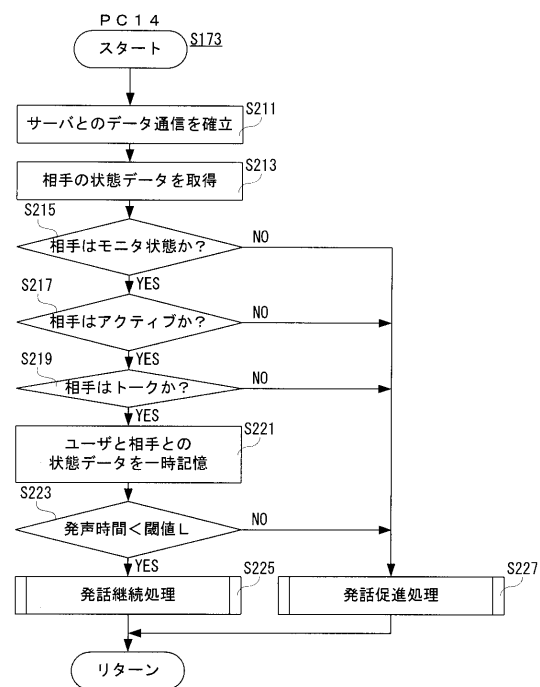
【図 27】



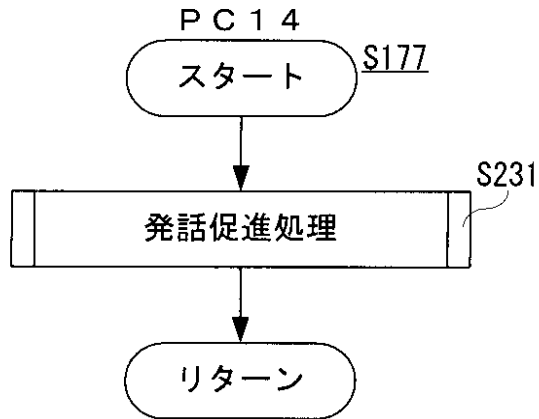
【図 28】



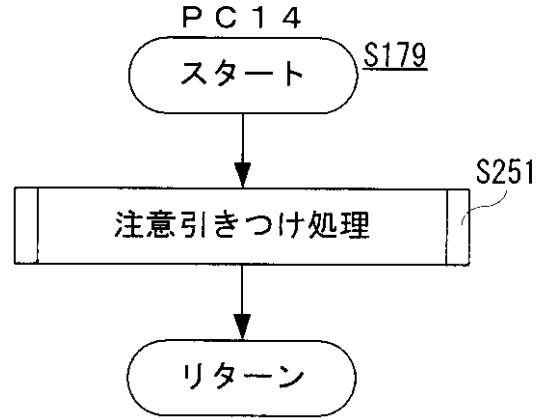
【図 29】



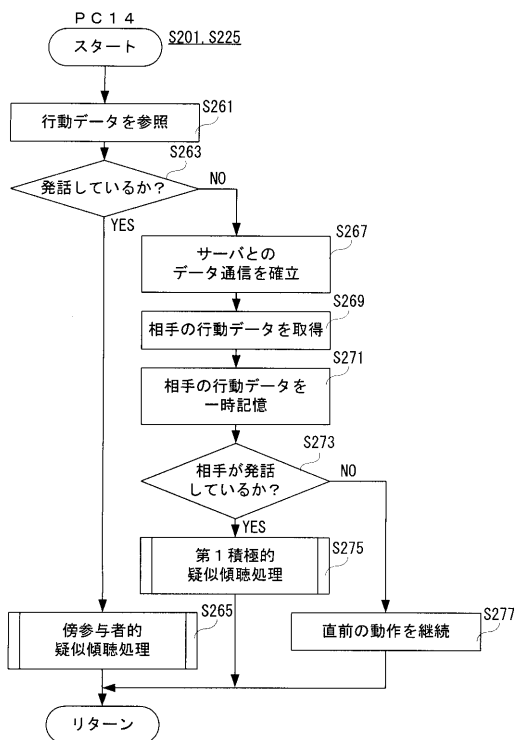
【図 30】



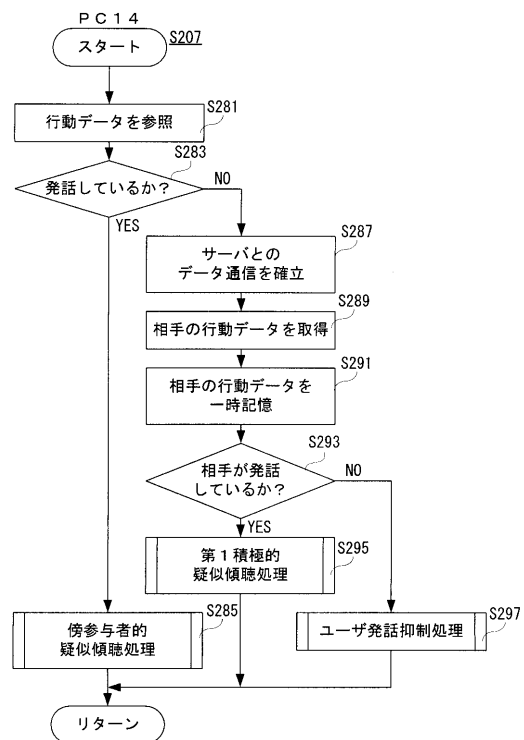
【図 31】



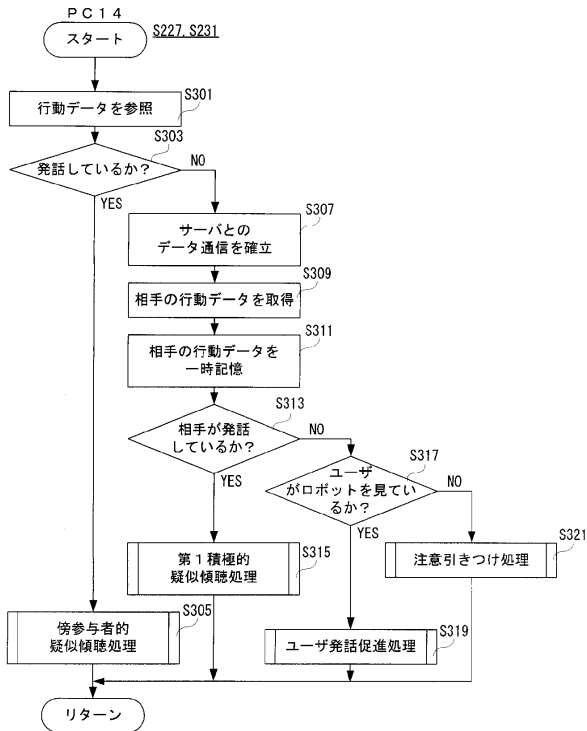
【図 32】



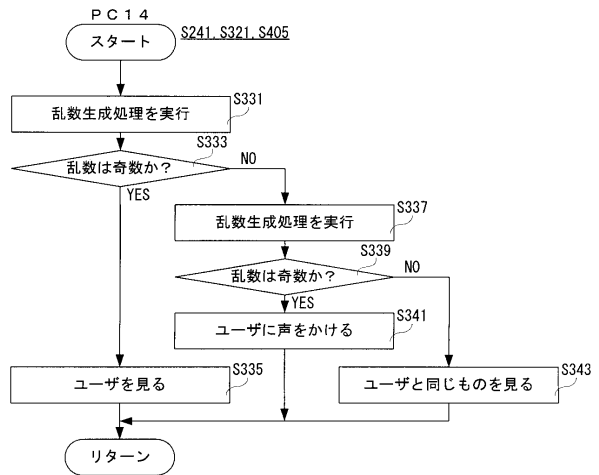
【図 33】



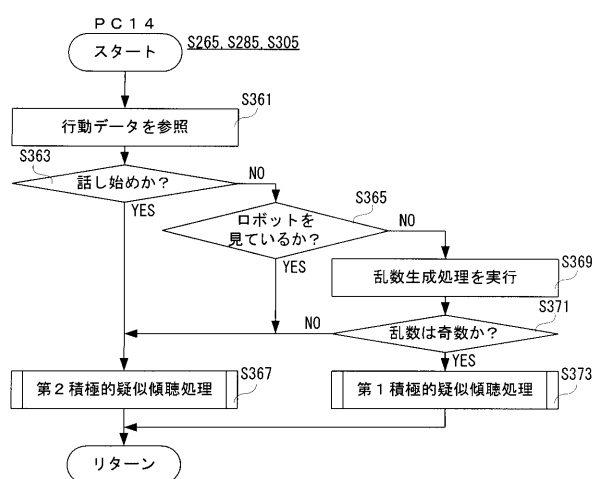
【図 34】



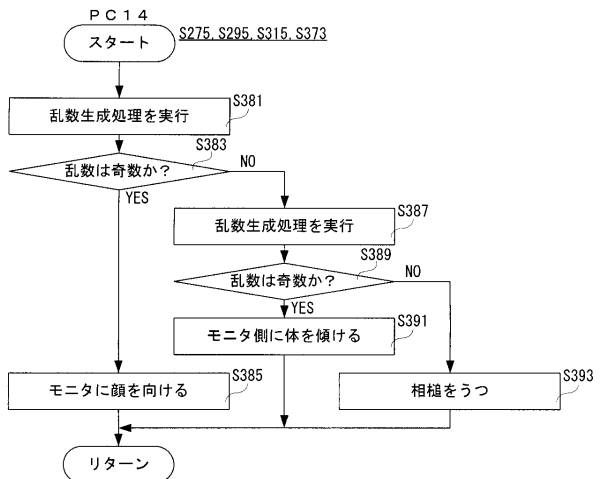
【図 35】



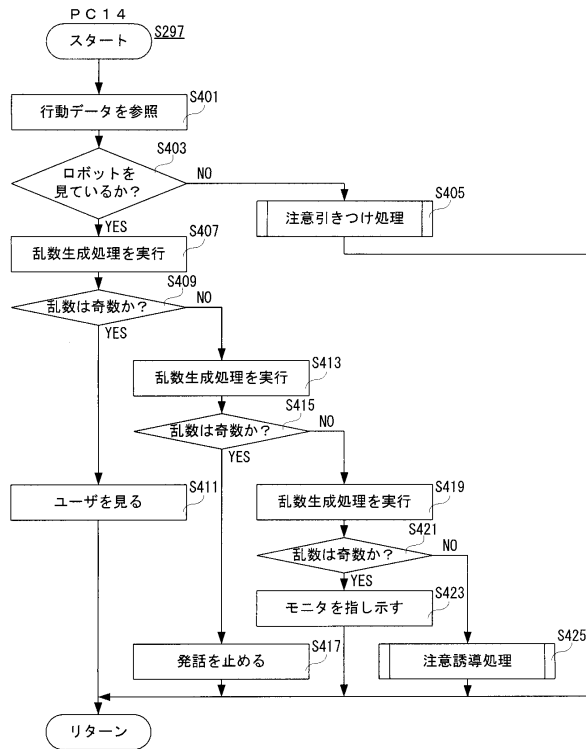
【図 36】



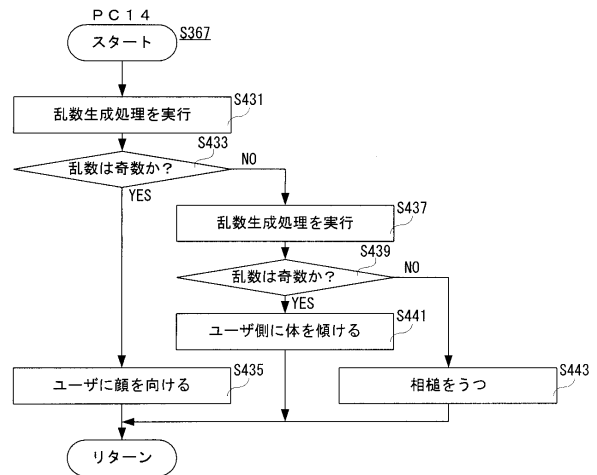
【図 37】



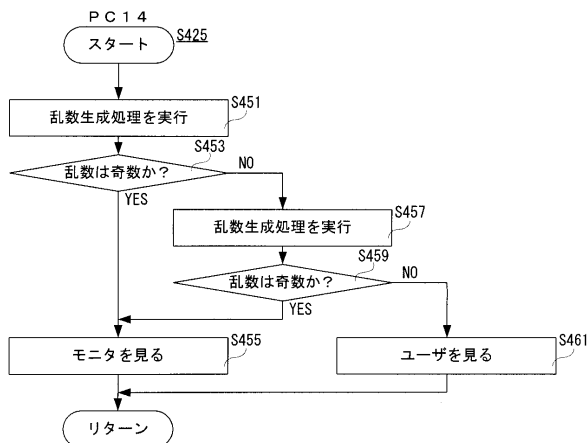
【図 38】



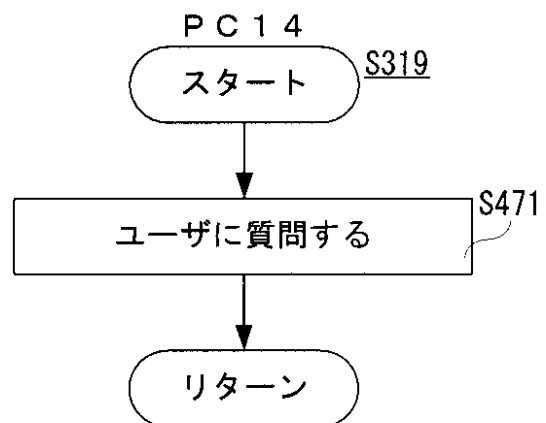
【図 39】



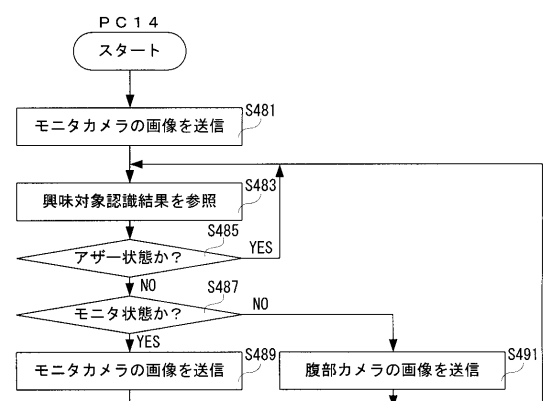
【図 40】



【図 41】



【図 42】



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 福西 章人

(56)参考文献 特開平11-289523(JP,A)
特開2009-106325(JP,A)
特開2002-261966(JP,A)
特開2007-336445(JP,A)
特開2001-162573(JP,A)
特開平09-185488(JP,A)