

明 細 書

発明の名称：対話ユーザの感情情報の提供装置

技術分野

[0001] 本発明は、遠隔に離れたユーザ同士の対話における、対話ユーザの感情情報の提供装置に関わる。

背景技術

[0002] 昨今、ビデオ会議や電話会議が普及するようになり、遠隔に離れたユーザ間の対話において、円滑なコミュニケーションを図るための技術が提供されている。

[0003] 例えば、特許文献1において、ビデオ会議装置の表示部の近傍に備えられた撮像部によって撮像された画像からユーザの視線方向を解析し、ユーザが注目する画面領域を拡大してユーザに配信する技術が開示されている。

先行技術文献

特許文献

[0004] 特許文献1：特開2014-050018号

発明の概要

発明が解決しようとする課題

[0005] しかしながら、特許文献1には、遠隔に離れた対話ユーザの感情を伝えることでコミュニケーションの向上を図る技術は開示されていない。

[0006] そこで、本発明は、遠隔に離れた対話ユーザのコミュニケーションの向上を図ることを目的とする。

課題を解決するための手段

[0007] 本発明の一態様における、相互に遠隔に位置する、第1のユーザの入出力端末と第2ユーザの入出力端末とを用いた、第1ユーザと第2ユーザとのビデオ対話を支援する装置であって、前記装置は、第1のユーザの、前記第1のユーザの入出力端末上の視点情報を受信する入力受付部と、前記視点情報を解析する解析部と、前記解析した視点情報を基に、感情情報を生成する感情

情報生成部と、を有する。

発明の効果

[0008] 本発明によれば、遠隔に離れた対話ユーザのコミュニケーションの向上を図ることができる。

図面の簡単な説明

[0009] [図1]本発明の第一実施形態に係る、遠隔対話システムを示すブロック構成図である。

[図2]図1のサーバ端末100を示す機能ブロック構成図である。

[図3]図1の対話装置200を示す機能ブロック構成図である。

[図4]対話装置の一例として、撮像部を説明する図である。

[図5]サーバ100に格納されるユーザデータの一例を示す図である。

[図6]サーバ100に格納される解析データの一例を示す図である。

[図7]サーバ100に格納される感情情報の一例を示す図である。

[図8]感情情報を時系列で表現した図である。

[図9]サーバ100に格納される感情情報の他の一例を示す図である。

[図10]本発明の第一実施形態に係る、感情情報の生成方法を示すフローチャートである。

発明を実施するための形態

[0010] 以下、本発明の実施形態について図面を参照して説明する。なお、以下に説明する実施形態は、特許請求の範囲に記載された本発明の内容を不当に限定するものではない。また、実施形態に示される構成要素のすべてが、本発明の必須の構成要素であるとは限らない。

[0011] <構成>

図1は、本発明の第一実施形態に係る、本発明の第一実施形態に係る、遠隔対話システムを示すブロック構成図である。本システム1は、視点情報を格納し、解析し、感情情報を生成するためのサーバ端末100と、ユーザ同士の対話に用いられ、カメラ等の撮像部を内蔵し、ユーザの視点情報を取得する対話装置200A、200Bと、を含む。なお、説明の便宜上、サーバ

端末を単一のものとして、また、対話装置を2台記載しているが、複数のサーバ端末、また、1台または2台より多くの対話装置で構成されてもよい。

[0012] サーバ端末100及び対話装置200A、200Bは各々、ネットワークNWを介して接続される。ネットワークNWは、インターネット、イントラネット、無線LAN (Local Area Network) やWAN (Wide Area Network) 等により構成される。

[0013] サーバ端末100は、例えば、ワークステーションやパーソナルコンピュータのような汎用コンピュータとしてもよいし、或いはクラウド・コンピューティングによって論理的に実現されてもよい。

[0014] 対話装置200は、例えば、ビデオ会議装置のほか、パーソナルコンピュータやタブレット端末等の情報処理装置、または、スマートフォンや携帯電話、PDA等により構成しても良い。また、例えば、対話装置として、パーソナルコンピュータやスマートフォンと液晶表示装置とを近距離無線通信等で接続し、対話を行う自ユーザ及び他ユーザの画像を液晶表示装置に表示しながら、必要な操作を、パーソナルコンピュータやスマートフォンを介して行うことを可能とする構成としてもよい。

[0015] 図2は、図1のサーバ端末100の機能ブロック構成図である。サーバ端末100は、通信部110と、記憶部120と、制御部130とを備える。

[0016] 通信部110は、ネットワークNWを介して対話装置200と通信を行うための通信インターフェースであり、例えばTCP/IP (Transmission Control Protocol/Internet Protocol) 等の通信規約により通信が行われる。

[0017] 記憶部120は、各種制御処理や制御部130内の各機能、また、遠隔対話アプリケーションを実行するためのプログラム、入力データ等を記憶するものであり、RAM (Random Access Memory)、ROM (Read Only Memory) 等から構成される。また、記憶部120は、ユーザに関連する各種データを格納するユーザデータ格納部121、及びユーザから視点情報を解析した解析データ及び解析結果を基に生成

した感情情報を格納する、解析データ格納部 122 を有する。なお、各種データを格納したデータベース（図示せず）が記憶部 120 またはサーバ端末 100 外に構築されていてもよい。

[0018] 制御部 130 は、記憶部 120 に記憶されているプログラムを実行することにより、サーバ端末 100 の全体の動作を制御するものであり、CPU (Central Processing Unit) や GPU (Graphics Processing Unit) 等から構成される。制御部 130 の機能として、各装置からの視点情報等の情報を受け付ける入力受付部 131 と、視点情報を解析する、解析部 132 と、視点情報の解析結果を基に感情情報を生成する、感情情報生成部 133 とを有する。この入力受付部 131、解析部 132、及び感情情報生成部 133 は、記憶部 120 に記憶されているプログラムにより起動されてコンピュータ（電子計算機）であるサーバ端末 100 により実行される。

[0019] 入力受付部 131 は、対話装置 200 において取得した、ユーザの視点情報を受信し、ビデオ通話の場合、ユーザから音声情報、画像情報等を受信することができる。受信したユーザの視点情報を、記憶部 120 の、ユーザデータ格納部 121 及び／または解析データ格納部 122 に格納することができる。

[0020] 解析部 132 は、受信した視点情報を解析し、解析した視点情報を、ユーザデータ格納部 121 及び／または解析データ格納部 122 に格納することができる。

[0021] 感情情報生成部 133 は、解析した視点情報を基に感情情報を生成し、感情情報を、ユーザデータ格納部 121 及び／または解析データ格納部 122 に格納することができる。

[0022] また、制御部 130 は、図示しない、感情情報報知制御部を有することもでき、対話装置 200 に備えられた報知部を介して、感情情報を報知できるよう、例えば、報知部がスマートフォン端末を振動させる振動モータ等である場合、対話ユーザの感情に基づいて振動をアクティベートさせるよう制御

する制御信号を生成し、制御信号を対話ユーザとは異なる対話装置に送信することができる。

[0023] また、制御部 130 は、図示しない、画面生成部を有することもでき、対話装置 200 のユーザインターフェースを介して表示される画面情報を生成する。例えば、記憶部 120 に格納された（図示しない）画像及びテキストデータを素材として、所定のレイアウト規則に基づいて、各種画像及びテキストをユーザインターフェースの所定の領域に配置することで、（例えば、広告主に対して広告効果を可視化して示すためのダッシュボード等の）ユーザインターフェースを生成する。画像生成部に関連する処理は、GPU（Graphics Processing Unit）によって実行することもできる。特に、上記生成した感情情報を可視化して対話装置 200 に表示したい場合、画面生成部は、感情情報を色や文字等で識別化して可視化した画面情報を生成することができる。

[0024] また、制御部 130 は、その他、複数のユーザ間の、ビデオによる遠隔対話を実現するための遠隔対話アプリケーションに含まれる各種処理を実行することができる。

[0025] 図 3 は、図 1 の対話装置 200 を示す機能ブロック構成図である。対話装置 200 は、通信部 210 と、表示操作部 220 と、記憶部 230 と、制御部 240、撮像部 250、報知部 260 とを備える。

[0026] 通信部 210 は、ネットワーク NW を介してサーバ端末 100、他の対話装置 200 と通信を行うための通信インターフェースであり、例えば TCP/IP 等の通信規約により通信が行われる。

[0027] 表示操作部 220 は、ユーザが指示を入力し、制御部 240 からの入力データに応じてテキスト、画像等を表示するために用いられるユーザインターフェースであり、対話装置 200 がパーソナルコンピュータで構成されている場合はディスプレイとキーボードやマウスにより構成され、対話装置 200 がスマートフォンまたはタブレット端末で構成されている場合はタッチパネル等から構成される。この表示操作部 220 は、記憶部 230 に記憶され

ている制御プログラムにより起動されてコンピュータ（電子計算機）である対話装置２００により実行される。

[0028] 記憶部２３０は、各種制御処理や制御部４４０内の各機能を実行するためのプログラム、入力データ等を記憶するものであり、ＲＡＭやＲＯＭ等から構成される。また、記憶部２３０は、サーバ端末１００との通信内容を一時的に記憶している。

[0029] 制御部２４０は、記憶部２３０に記憶されているプログラム（遠隔対話アプリケーションに含まれるプログラムを含む）を実行することにより、対話装置２００の全体の動作を制御するものであり、ＣＰＵやＧＰＵ等から構成される。

[0030] また、対話装置２００が、パーソナルコンピュータやスマートフォン、タブレット端末等で構成されている場合は、ユーザの眼球を赤外線により撮像し、ユーザの液晶表示画面上の視点位置をトラッキングすることが可能な、内蔵カメラ等の撮像部２５０を有し、また、スマートフォン等で構成されている場合は、振動を発生させる振動モータ等、感情情報をユーザに対して報知するための報知部を有することができる。

[0031] 図４は、対話装置２００の他の一例として、撮像部を説明する図である。

[0032] 図４に示す対話装置２００は、液晶表示装置２１０を含み、液晶表示部２２０の中央部に貫通穴２３０が設けられ、ＣＣＤカメラ２４０が貫通穴２３０に嵌合されるように備えられる。本例の対話装置２００は、液晶表示装置２１０と近距離無線通信または有線により接続する（図示しない）スマートフォンをさらに備え、スマートフォンは、遠隔対話アプリケーションに含まれる、ビデオ通話、画面共有等の各種処理を実行し、遠隔対話するユーザの対話装置２００Ａから、サーバ端末１００、ネットワークＮＷを介して伝送される画像情報から生成される画面を、液晶表示装置２１０の液晶表示部２１０に表示させることができる。ＣＣＤカメラ２４０は、対話装置２００を用いるユーザの眼球を赤外線により撮像し、液晶表示装置上のユーザの視点位置をトラッキングすることができる。液晶表示部の中央部に撮像部（ＣＣ

Dカメラ)を設けることで、その液晶表示装置を用いて対話を行うユーザは自然な形で液晶表示部に表示される相手の対話ユーザと会話をすることができる。本例において、このような自然な形での対話方法を実現するために、撮像部が位置する領域に相手ユーザの顔の位置(より好ましくは眼の位置)が合わさるよう表示されることが好ましく、相手ユーザが移動する場合には、常に顔が中心に位置するよう相手ユーザの対話装置に備えられるカメラが追従されることが好ましい。

[0033] 図5は、サーバ100に格納されるユーザデータの一例を示す図である。

[0034] ユーザデータ1000は、ユーザに関連する各種データを格納する。図5において、説明の便宜上、一ユーザ(ユーザID「10001」で識別される予定)の例を示すが、複数のユーザに関連する情報を格納することができる。ユーザに関連する各種データとして、例えば、ユーザの基本情報(例えば、「氏名、住所、年齢、性別、職業」等のユーザとしての属性情報として利用される情報)、視点情報(例えば、撮像された画像に基づいて解析された、ユーザID「10001」で識別されるユーザの液晶表示画面上の視点位置情報)、及び感情情報(例えば、視点位置情報に基づいて生成された、ユーザID「10001」で識別されるユーザの感情情報)を含むことができる。

[0035] 図6は、サーバ100に格納される解析データの一例を示す図である。

[0036] 解析データとして、(例えば、撮像された画像に基づいて解析された、各ユーザの液晶表示画面上の視点位置情報)、及び感情情報(例えば、視点位置情報に基づいて生成された、各ユーザの感情情報)を含むことができる。

[0037] 図7は、サーバ100に格納される感情情報の一例を示す図である。

[0038] 図7に示す感情情報のテーブルにおいて、例えば、ユーザが、液晶表示部(液晶表示画面)の中心部の座標をx軸、y軸方向に(0、0)として定義するときに、あるユーザの視点位置を、テーブルの上から下方向に向けてトラッキングし、対応する感情情報を含むように構成されたものである。例えば、あるユーザと対話する対話ユーザの映像が、画面中央に表示される液晶表

示画面において、ユーザが、視点位置（0、0）、すなわち、画面中央に視点を合わせるときは、そのユーザは、対話ユーザとのコミュニケーションについて非常にポジティブである（高い関心を示している）ことが推定できる。一方で、ユーザの視点が、画面中央から離れていくに従って、そのユーザは、コミュニケーションについてネガティブとなる（関心が低くなる）ことが推定できる。ここで、ユーザの視点位置（座標）と対応する感情情報については、中心部の座標を中心として、座標の範囲に対応するようにルールを予め設定することもできるし、一のユーザの過去の視点情報と感情情報の組み合わせ、及び／または、複数のユーザの、過去の視点情報と感情情報の組み合わせを学習モデルとして、機械学習により、視点情報の入力から感情情報を出力することもできる。学習モデルの生成に際しては、ユーザから感情情報のフィードバックを、サーベイや音声情報等の追加的な情報によって取得することもできる。音声情報を用いる場合、例えば、音声情報からユーザの感情を検出したり、音声情報から自然言語解析を行い、会話内容から感情情報を検出し、入力情報（視点情報）に対する出力として評価することができる。

[0039] 図8は、感情情報を時系列で表現した図である。

[0040] 図8において、縦軸は、ユーザの感情を5段階（1: Very Negative、2: Negative、3: Neutral、4: Positive、5: Very Positive）で示しており、横軸は時間軸として示している。図8に示すように、ユーザの視点情報を基に感情情報を導出し、これを時系列で表現することができる。図8においては、ユーザが対話の冒頭、コミュニケーションについて高い関心を示しており、途中関心が低くなり、その後関心の高まりを次第に示す様子が可視化される。このような可視化された感情情報の遷移を、上述のようにサーバ端末100の画面生成部により画面情報として生成し、対話装置200に送信し、表示させることで、ユーザは対話ユーザの感情情報の遷移を参照しながらコミュニケーションを図ることができる。

[0041] 図9は、サーバ100に格納される感情情報の他の一例を示す図である。

[0042] 図9に示すように、ユーザの視点情報を位置別に回数をカウントし、及び／または、その注視時間の累計を格納することで、そのユーザが、対話ユーザとのコミュニケーション全体（途中経過を含む）として、どのような感情を抱いているか、測ることができる。例えば、図9に示す情報から、ユーザは、コミュニケーション全体を通じて、視点位置が、座標（0、0）、すなわち、画面中央に最も注がれていることが理解でき、そのユーザが、コミュニケーションに対して、Very Positive（関心が非常に高い）な感情を頂いていることがわかる。

[0043] <処理の流れ>

図10を参照しながら、本実施形態のシステム1が実行する感情情報生成処理の流れについて説明する。図10は、本発明の第一実施形態に係る、感情情報の生成方法を示すフローチャートである。

[0044] ここで、本システム1を利用するために、ユーザは、各対話装置のウェブブラウザまたはアプリケーション等を利用してサーバ端末100にアクセスし、初めてサービスを利用する場合は、前述のユーザ基本情報等を入力し、既にユーザアカウントを取得済の場合は、例えばIDとパスワードを入力する等の所定の認証を受けてログインすることで、サービスが利用可能となる。この認証後、ウェブサイト、アプリケーション等を介して所定のユーザインターフェースが提供され、ビデオ通話サービスを利用することができ、図10に示すステップS101へ進む。

[0045] まず、ステップS101の処理として、サーバ端末100の制御部130の入力受付部131は、通信部110を介して、対話装置200Aから、視点情報を受信する。視点情報については、例えば、図4に示す対話装置の液晶表示部220に備えられた、CCDカメラ240により、ユーザの画像を撮像することで、視点位置の情報を取得することができる。図4に示す対話装置を用いる場合は、液晶表示部220の中心部（カメラ240が備えられる位置）に対話ユーザの映像が表示されることが好ましい。ここで、対話装

置200Aにおいて、撮像された画像に基づいてユーザの視点位置を算出したうえで、視点位置に係る情報を、対話装置200Aからサーバ端末100に送信することもできるし、画像情報をサーバ端末100に送信したうえで、サーバ端末100の制御部130の解析部132によって、受信した画像に基づいて視点位置を算出することもできる。

[0046] 次に、ステップS102の処理として、サーバ端末100の制御部130の解析部132は、視点情報を解析する。また、解析部132は、視点情報として、ユーザの、液晶表示部（画面）上の視点位置を、継続的に、または、所定時間間隔にて、視点情報を取得した都度、特定のユーザに紐づけてユーザデータ格納部121、及び／または、解析データ格納部122に格納する。また、解析部132は、時系列でユーザの視点情報をトラッキングし、格納することができる。さらに、解析部132は、視点情報を基に、ユーザの視点位置が所定の座標に置かれた回数をカウントしたり、または、所定の座標に置かれた時間を都度計測したり、時間の累計を算出したりすることができる。また、上述のように、解析部132は、対話装置200Aから受信した、対話ユーザを含む画像に基づいて視点位置を算出することもできる。

[0047] 続いて、ステップS103の処理として、サーバ端末100の制御部130の感情情報生成部133は、解析した視点情報に基づいて、感情情報を生成する。例えば、感情情報生成部133は、図7に示すように、ユーザの視点位置が、液晶表示部の中心部を中心とした座標からどの範囲にあるか、という所定のルールに基づいて、感情情報を生成することができる。例えば、ユーザの視点位置が座標（0、0）、すなわち、画面中央にあるときは、そのユーザは、対話ユーザとのコミュニケーションについて非常にポジティブである（高い関心を示している）という感情情報が生成され、一方で、ユーザの視点が、画面中央から離れており、座標（-500、500）にあるときは、そのユーザは、コミュニケーションについて非常にネガティブである（関心が非常に低い）、という感情情報を生成することができる。または、上述のように、ユーザの視点情報と感情情報とで構成される学習モデルから、

機械学習により、入力された視点情報を基に感情情報を生成することもできる。

[0048] また、図8に示すように、感情情報の遷移を時系列で変化することを可視化する情報を生成したり、図9に示すように、ユーザの視点が置かれた座標の回数及び／または累計時間によって、そのユーザの、コミュニケーション全体における感情を評価した情報を生成することもできる。生成した感情情報を可視化した情報として、対話装置200Bに送信し、対話装置200Bの表示部に表示させたり、感情情報を、対話装置200Bを用いるユーザに報知させるために、感情情報の度合い（上述の5段階評価）に基づいてアイコン等により識別表示させたり、ユーザに感情情報を感覚的に伝達するために、対話装置200Bの振動モータ等の報知部を駆動させるための制御信号を生成し、送信することができる。

[0049] 以上により、ユーザの視点位置に基づいて感情情報を生成することにより、遠隔のユーザのコミュニケーションにおいて、相互のユーザに感情情報を共有することが可能となり、コミュニケーションの質の向上を図ることができる。

[0050] 以上、発明に係る実施形態について説明したが、これらはその他の様々な形態で実施することが可能であり、種々の省略、置換および変更を行なって実施することが出来る。これらの実施形態および変形例ならびに省略、置換および変更を行なったものは、特許請求の範囲の技術的範囲とその均等の範囲に含まれる。

符号の説明

[0051] 1 システム 100 サーバ端末、110 通信部、120 記憶部、130 制御部、200 対話装置、NW ネットワーク