

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7465013号
(P7465013)

(45)発行日 令和6年4月10日(2024.4.10)

(24)登録日 令和6年4月2日(2024.4.2)

(51)国際特許分類 F I
H 0 4 N 7 / 1 5 (2006.01) H 0 4 N 7 / 1 5

請求項の数 1 (全20頁)

(21)出願番号	特願2022-517945(P2022-517945)	(73)特許権者	520408744 株式会社 I ' m b e s i d e y o u 東京都港区虎ノ門1-17-1 虎ノ門 ヒルズビジネスタワー15階
(86)(22)出願日	令和2年12月31日(2020.12.31)	(74)代理人	110002790 O n e i p 弁理士法人
(86)国際出願番号	PCT/JP2020/049297	(72)発明者	神谷 渉三 東京都世田谷区成城2丁目34番2号村 上ハイツ301 株式会社 I ' m b e s i d e y o u 内
(87)国際公開番号	WO2022/145042	審査官	松元 伸次
(87)国際公開日	令和4年7月7日(2022.7.7)		
審査請求日	令和5年12月27日(2023.12.27)		
早期審査対象出願			

最終頁に続く

(54)【発明の名称】 ビデオミーティング評価端末、ビデオミーティング評価システム及びビデオミーティング評価プログラム

(57)【特許請求の範囲】

【請求項1】

他の端末との間で行われるビデオミーティングの動画像を取得する取得手段と、
前記動画像内に含まれる対象者の少なくとも顔画像を所定のフレーム単位ごとに認識する顔認識手段と、
前記動画像内に含まれる前記対象者の少なくとも音声認識する音声認識手段と
認識した前記顔画像及び前記音声の双方に基づいて複数の観点による評価値を算出する評価手段と、

前記複数の観点による評価値の相対的な乖離が一定以上になった場合に通知を行う通知手段と、を備え、

前記評価手段は、前記顔画像から評価される第1の感情と、前記音声から得られる第2の感情とが相反する場合に、前記第1及び第2の感情を包含する第3の感情を推定する、
ビデオミーティング評価端末。

【発明の詳細な説明】

【技術分野】

【0001】

本開示は、ビデオミーティング評価端末、ビデオミーティング評価システム及びビデオミーティング評価プログラムに関する。

【背景技術】

【0002】

従来、オンラインで何らかの知識を教授したり、説明等を行うためのシステムが知られている（例えば、特許文献 1 参照）。

【 0 0 0 3 】

また、このようなオンラインで行われるビデオミーティングの効果測定の方法として、例えば、ビデオミーティング後にアンケート等を行う方法も提案されている（例えば、特許文献 2 参照）。

【先行技術文献】

【特許文献】

【 0 0 0 4 】

【文献】特開 2 0 1 9 - 5 8 6 2 5 号公報

10

【発明の概要】

【発明が解決しようとする課題】

【 0 0 0 5 】

上述したアンケートによる効果測定の方法は、主観的になりがちであり、ビデオミーティングの内容に関して客観的な効果測定を行う方法としては不十分である。

【 0 0 0 6 】

また、第三者によってビデオミーティングを監視し、第三者による客観的な評価を取得する方法も考えられるが、手間と時間がかかりすぎて現実的ではない。

【 0 0 0 7 】

そこで、本発明は、ビデオミーティングの特に内容に関する評価を客観的に行うことを目的とする。

20

【課題を解決するための手段】

【 0 0 0 8 】

本発明によれば、他の端末との間で行われるビデオミーティングの動画像を取得する取得手段と、前記動画像内に含まれる対象者の少なくとも顔画像を所定のフレーム単位ごとに認識する顔認識手段と、

前記動画像内に含まれる前記対象者の少なくとも音声を認識する音声認識手段と

認識した前記顔画像及び前記音声の双方に基づいて複数の観点による評価値を算出する評価手段と、

30

前記複数の観点による評価値の相対的な乖離が一定以上になった場合に通知を行う通知手段と、を備える

ビデオミーティング評価端末が得られる。

【発明の効果】

【 0 0 0 9 】

本開示によれば、ビデオミーティングの動画像を評価することにより、特に内容に関する評価を客観的に行うことができる。

【 0 0 1 0 】

特に、本発明によれば、取得した動画像は当該端末に保存されることを前提としていることから、端末上で分析評価され、その結果が当該端末のユーザに提供される。従って、例えば個人情報を含むビデオミーティングや機密情報を含むビデオミーティングであっても、その動画自体を外部の評価事業者等に提供することなく分析評価できる

40

【 0 0 1 1 】

また、分析・評価した結果をリアルタイムで又は事後的に動画像に重畳して表示することとすれば、より利便性が向上できる。

【図面の簡単な説明】

【 0 0 1 2 】

【図 1】本発明の実施の形態によるシステム全体図を示す図である。

【図 2】本発明の実施の形態による端末の構成例を示す図である。

【図 3】本発明の第 1 の実施の形態による評価端末の機能ブロック図の一例である。

50

【図 4】本発明の第 1 の実施の形態による機能ブロック図である。

【図 5】本発明の第 1 の実施の形態による機能ブロック図である。

【図 6】本発明の第 1 の実施の形態による機能ブロック図である。

【図 7】本発明の第 1 の実施の形態による機能ブロック図である。

【図 8】本発明の第 1 の実施の形態による機能ブロック図である。

【図 9】本発明の第 2 の実施の形態による機能ブロック図である。

【図 10】本発明の第 2 の実施の形態による機能ブロック図である。

【図 11】本発明の第 3 の実施の形態による実装例である。

【図 12】本発明の第 4 の実施の形態による実装例である。

【図 13】本発明の第 5 の実施の形態による実装例である。

【図 14】本発明の第 6 の実施の形態による実装例である。

【発明を実施するための形態】

【0013】

本開示の実施形態の内容を列記して説明する。本開示は、以下のような構成を備える。

[項目 1]

他の端末との間で行われるビデオミーティングの動画像を取得する取得手段と、
前記動画像内に含まれる対象者の少なくとも顔画像を所定のフレーム単位ごとに認識する顔認識手段と、

前記動画像内に含まれる前記対象者の少なくとも音声認識する音声認識手段と
認識した前記顔画像及び前記音声の双方に基づいて複数の観点による評価値を算出する
評価手段と、

前記複数の観点による評価値の相対的な乖離が一定以上になった場合に通知を行う通知手段と、を備える

ビデオミーティング評価端末。

[項目 2]

項目 1 に記載のビデオミーティング評価端末であって、

前記通知手段は、前記顔画像及び前記音声の夫々から得られる評価値が一定以上乖離していた場合に通知する、

ビデオミーティング評価端末。

[項目 3]

項目 1 に記載のビデオミーティング評価端末であって、

前記通知手段は、前記顔画像から取得できる顔の動き及び前記音声の夫々から得られる評価値が一定以上乖離していた場合に通知する、

ビデオミーティング評価端末。

[項目 4]

項目 1 に記載のビデオミーティング評価端末であって、

前記通知手段は、前記顔画像から評価される感情の度合いと、前記音声とから得られる評価値が一定以上乖離していた場合に通知する、

ビデオミーティング評価端末。

[項目 5]

コンピュータを、

他の端末との間で行われるビデオミーティングの動画像を取得する取得手段、
前記動画像内に含まれる対象者の少なくとも顔画像を所定のフレーム単位ごとに認識する顔認識手段、

前記動画像内に含まれる前記対象者の少なくとも音声認識する音声認識手段、
認識した前記顔画像及び前記音声の双方に基づいて複数の観点による評価値を算出する
評価手段、

前記複数の観点による評価値の相対的な乖離が一定以上になった場合に通知を行う通知手段として機能させる

ビデオミーティング評価プログラム。

10

20

30

40

50

【0014】

以下に添付図面を参照しながら、本開示の好適な実施の形態について詳細に説明する。なお、本明細書及び図面において、実質的に同一の機能構成を有する構成要素については、同一の符号を付することにより重複説明を省略する。

【0015】

<基本機能>

本実施形態のビデオミーティング評価システム（以下、単に「本システム」という場合がある）は、複数人でビデオミーティング（以下、一方向及び双方向含めてオンラインセッションという）が行われる環境において、当該複数人の中の解析対象者について他者とは異なる特異的な感情（自分または他人の言動に対して起こる気持ち。快・不快またはその程度など）を解析し評価するシステムである。

10

【0016】

オンラインセッションは、例えばオンライン会議、オンライン授業、オンラインチャットなどであり、複数の場所に設置された端末をインターネットなどの通信ネットワークを介してサーバに接続し、当該サーバを通じて複数の端末間で動画像をやり取りできるようにしたものである。

【0017】

オンラインセッションで扱う動画像には、端末を使用するユーザの顔画像や音声が含まれる。また、動画像には、複数のユーザが共有して閲覧する資料などの画像も含まれる。各端末の画面上に顔画像と資料画像とを切り替えて何れか一方のみを表示させたり、表示領域を分けて顔画像と資料画像とを同時に表示させたりすることが可能である。また、複数人のうち1人の画像を全画面表示させたり、一部または全部のユーザの画像を小画面に分割して表示させたりすることが可能である。

20

【0018】

端末を使用してオンラインセッションに参加する複数のユーザのうち、何れか1人または複数人を解析対象者として指定することが可能である。例えば、オンラインセッションの主導者、進行者または管理者（以下、まとめて主催者という）が何れかのユーザを解析対象者として指定する。オンラインセッションの主催者は、例えばオンライン授業の講師、オンライン会議の議長やファシリテータ、コーチングを目的としたセッションのコーチなどである。オンラインセッションの主催者は、オンラインセッションに参加する複数のユーザの中の一人であるのが普通であるが、オンラインセッションに参加しない別人であってもよい。なお、解析対象者を指定せず全ての参加者を解析対象としてもよい。

30

【0019】

また、オンラインセッションの主導者、進行者または管理者（以下、まとめて主催者という）が何れかのユーザを解析対象者として指定することも可能である。オンラインセッションの主催者は、例えばオンライン授業の講師、オンライン会議の議長やファシリテータ、コーチングを目的としたセッションのコーチなどである。オンラインセッションの主催者は、オンラインセッションに参加する複数のユーザの中の一人であるのが普通であるが、オンラインセッションに参加しない別人であってもよい。

【0020】

本実施の形態によるビデオミーティング評価システムは、複数の端末間においてビデオミーティングセッションが確立された場合に、当該ビデオミーティングから取得される少なくとも動画像を表示される。表示された動画像は、端末によって取得され、動画像内に含まれる少なくとも顔画像を所定のフレーム単位ごとに識別される。その後、識別された顔画像に関する評価値が算出される。当該評価値は必要に応じて共有される。

40

【0021】

特に、本実施の形態においては、取得した動画像は当該端末に保存され、端末上で分析評価され、その結果が当該端末のユーザに提供される。従って、例えば個人情報を含むビデオミーティングや機密情報を含むビデオミーティングであっても、その動画自体を外部の評価機関等に提供することなく分析評価できる。また、必要に応じて、当該評価結果（

50

評価値)だけを外部端末に提供することによって、結果を可視化したり、クロス分析等行うことができる。

【0022】

図1に示されるように、本実施の形態によるビデオミーティング評価システムは、少なくともカメラ部及びマイク部等の入力部と、ディスプレイ等の表示部とスピーカー等の出力部とを有するユーザ端末10、20と、ユーザ端末10、20に双方向のビデオミーティングを提供するビデオミーティングサービス端末30と、ビデオミーティングに関する評価の一部を行う評価端末40とを備えている。

【0023】

<ハードウェア構成例>

図2は、本実施形態に係る各端末10乃至40を実現するコンピュータのハードウェア構成例を示す図である。コンピュータは、少なくとも、制御部110、メモリ120、ストレージ130、通信部140および入出力部150等を備える。これらはバス160を通じて相互に電氣的に接続される。

【0024】

制御部110は、各端末全体の動作を制御し、各要素間におけるデータの送受信の制御、及びアプリケーションの実行及び認証処理に必要な情報処理等を行う演算装置である。例えば制御部110は、CPU等のプロセッサであり、ストレージ130に格納されメモリ120に展開されたプログラム等を実行して各情報処理を実施する。

【0025】

メモリ120は、DRAM等の揮発性記憶装置で構成される主記憶と、フラッシュメモリまたはHDD等の不揮発性記憶装置で構成される補助記憶と、を含む。メモリ120は、制御部110のワークエリア等として使用され、また、各端末の起動時に実行されるBIOS、及び各種設定情報等を格納する。

【0026】

ストレージ130は、アプリケーション・プログラム等の各種プログラムを格納する。各処理に用いられるデータを格納したデータベースがストレージ130に構築されていてもよい。特に本実施の形態においては、ビデオミーティングサービス端末30のストレージ130にはオンラインセッションにおける動画は記録されず、ユーザ端末10のストレージ130に格納される。また、評価端末40は、ユーザ端末10上において取得された動画を評価するために必要なアプリケーションその他のプログラムを格納し、ユーザ端末10が利用可能に適宜提供する。なお、評価端末40の管理するストレージ13には、例えば、ユーザ端末10によって解析された結果、評価された結果のみが共有されることとしてもよい。

【0027】

通信部140は、端末をネットワークに接続する。通信部140は、例えば、有線LAN、無線LAN、Wi-Fi(登録商標)、赤外線通信、Bluetooth(登録商標)、近距離または非接触通信等の方式で、外部機器と直接またはネットワークアクセスポイントを介して通信する。

【0028】

入出力部150は、例えば、キーボード、マウス、タッチパネル等の情報入力機器、及びディスプレイ等の出力機器である。

【0029】

バス160は、上記各要素に共通に接続され、例えば、アドレス信号、データ信号及び各種制御信号を伝達する。

【0030】

特に、本実施の形態による評価端末は、ビデオミーティングサービス端末から動画を取得し、当該動画内に含まれる少なくとも顔画像を所定のフレーム単位ごとに識別すると共に、顔画像に関する評価値を算出する(詳しくは後述する)。

<動画の取得方法>

10

20

30

40

50

図3に示されるように、ビデオミーティングサービス端末が提供するビデオミーティングサービス（以下、単に「本サービス」と言うことがある）は、ユーザ端末10、20に対して双方向に画像および音声によって通信が可能となるものである。本サービスは、ユーザ端末のディスプレイに相手のユーザ端末のカメラ部で取得した動画像を表示し、相手のユーザ端末のマイク部で取得した音声をスピーカーから出力可能となっている。

【0031】

また、本サービスは双方の又はいずれかのユーザ端末によって、動画像及び音声（これらを合わせて「動画像等」という）を少なくともいずれかのユーザ端末上の記憶部に記録（レコーディング）することが可能に構成されている。記録された動画像情報Vs（以下「記録情報」という）は、記録を開始したユーザ端末にキャッシュされつついずれかのユーザ端末のローカルのみ記録されることとなる。ユーザは、必要があれば当該記録情報を本サービスの利用の範囲内で自分で視聴、他者に共有等行うこともできる。

10

【0032】

ユーザ端末10は、以上のようにして取得した動画を以下のような分析によって評価を行う。

【0033】

以下、本発明の一実施形態を図面に基づいて説明する。図4は、本実施形態による構成例を示すブロック図である。図4に示すように、本実施形態のビデオミーティング評価システムは、ユーザ端末10が有する機能構成として実現される。すなわち、ユーザ端末10はその機能として、動画像取得部11、生体反応解析部12、特異判定部13、関連事象特定部14、クラスタリング部15および解析結果通知部16を備えている。

20

【0034】

上記各機能ブロック11～16は、例えばユーザ端末10に備えられたハードウェア、DSP（Digital Signal Processor）、ソフトウェアの何れによっても構成することが可能である。例えばソフトウェアによって構成する場合、上記各機能ブロック11～16は、実際にはコンピュータのCPU、RAM、ROMなどを備えて構成され、RAMやROM、ハードディスクまたは半導体メモリ等の記録媒体に記憶されたプログラムが動作することによって実現される。

【0035】

動画像取得部11は、オンラインセッション中に各端末が備えるカメラにより複数人（複数のユーザ）を撮影することによって得られる動画像を各端末から取得する。各端末から取得する動画像は、各端末の画面上に表示されるように設定されているものか否かは問わない。すなわち、動画像取得部11は、各端末に表示中の動画像および非表示中の動画像を含めて、動画像を各端末から取得する。

30

【0036】

生体反応解析部12は、動画像取得部11により取得された動画像（画面上に表示中のものか否かは問わない）に基づいて、複数人のそれぞれについて生体反応の変化を解析する。本実施形態において生体反応解析部12は、動画像取得部11により取得された動画像を画像のセット（フレーム画像の集まり）と音声とに分離し、それぞれから生体反応の変化を解析する。

40

【0037】

例えば、生体反応解析部12は、動画像取得部11により取得された動画像から分離したフレーム画像を用いてユーザの顔画像を解析することにより、表情、目線、脈拍、顔の動きの少なくとも1つに関する生体反応の変化を解析する。また、生体反応解析部12は、動画像取得部11により取得された動画像から分離した音声を解析することにより、ユーザの発言内容、声質の少なくとも1つに関する生体反応の変化を解析する。

【0038】

人は感情が変化すると、それが表情、目線、脈拍、顔の動き、発言内容、声質などの生体反応の変化となって現れる。本実施形態では、ユーザの生体反応の変化を解析することを通じて、ユーザの感情の変化を解析する。本実施形態において解析する感情は、一例と

50

して、快/不快の程度である。本実施形態において生体反応解析部12は、生体反応の変化を所定の基準に従って数値化することにより、生体反応の変化の内容を反映させた生体反応指標値を算出する。

【0039】

表情の変化の解析は、例えば以下のようにして行う。すなわち、フレーム画像ごとに、フレーム画像の中から顔の領域を特定し、事前に機械学習させた画像解析モデルに従って特定した顔の表情を複数に分類する。そして、その分類結果に基づいて、連続するフレーム画像間でポジティブな表情変化が起きているか、ネガティブな表情変化が起きているか、およびどの程度の大きさの表情変化が起きているかを解析し、その解析結果に応じた表情変化指標値を出力する。

10

【0040】

目線の変化の解析は、例えば以下のようにして行う。すなわち、フレーム画像ごとに、フレーム画像の中から目の領域を特定し、両目の向きを解析することにより、ユーザがどこを見ているかを解析する。例えば、表示中の話者の顔を見ているか、表示中の共有資料を見ているか、画面の外を見ているかなどを解析する。また、目線の動きが大きい小さいか、動きの頻度が多いか少ないかなどを解析するようにしてもよい。目線の変化はユーザの集中度にも関連する。生体反応解析部12は、目線の変化の解析結果に応じた目線変化指標値を出力する。

【0041】

脈拍の変化の解析は、例えば以下のようにして行う。すなわち、フレーム画像ごとに、フレーム画像の中から顔の領域を特定する。そして、顔の色情報(RGBのG)の数値を捉える学習済みの画像解析モデルを用いて、顔表面のG色の变化を解析する。その結果を時間軸に合わせて並べることによって色情報の変化を表した波形を形成し、この波形から脈拍を特定する。人は緊張すると脈拍が速くなり、気持ちが落ち着くと脈拍が遅くなる。生体反応解析部12は、脈拍の変化の解析結果に応じた脈拍変化指標値を出力する。

20

【0042】

顔の動きの変化の解析は、例えば以下のようにして行う。すなわち、フレーム画像ごとに、フレーム画像の中から顔の領域を特定し、顔の向きを解析することにより、ユーザがどこを見ているかを解析する。例えば、表示中の話者の顔を見ているか、表示中の共有資料を見ているか、画面の外を見ているかなどを解析する。また、顔の動きが大きい小さいか、動きの頻度が多いか少ないかなどを解析するようにしてもよい。顔の動きと目線の動きとを合わせて解析するようにしてもよい。例えば、表示中の話者の顔をまっすぐ見ているか、上目遣いまたは下目使いに見ているか、斜めから見ているかなどを解析するようにしてもよい。生体反応解析部12は、顔の向きの変化の解析結果に応じた顔向き変化指標値を出力する。

30

【0043】

発言内容の解析は、例えば以下のようにして行う。すなわち、生体反応解析部12は、指定した時間(例えば、30~150秒程度の時間)の音声について公知の音声認識処理を行うことによって音声を文字列に変換し、当該文字列を形態素解析することにより、助詞、冠詞などの会話を表す上で不要なワードを取り除く。そして、残ったワードをベクトル化し、ポジティブな感情変化が起きているか、ネガティブな感情変化が起きているか、およびどの程度の大きさの感情変化が起きているかを解析し、その解析結果に応じた発言内容指標値を出力する。

40

【0044】

声質の解析は、例えば以下のようにして行う。すなわち、生体反応解析部12は、指定した時間(例えば、30~150秒程度の時間)の音声について公知の音声解析処理を行うことによって音声の音響的特徴を特定する。そして、その音響的特徴に基づいて、ポジティブな声質変化が起きているか、ネガティブな声質変化が起きているか、およびどの程度の大きさの声質変化が起きているかを解析し、その解析結果に応じた声質変化指標値を出力する。

50

【 0 0 4 5 】

生体反応解析部 1 2 は、以上のようにして算出した表情変化指標値、目線変化指標値、脈拍変化指標値、顔向き変化指標値、発言内容指標値、声質変化指標値の少なくとも 1 つを用いて生体反応指標値を算出する。例えば、表情変化指標値、目線変化指標値、脈拍変化指標値、顔向き変化指標値、発言内容指標値および声質変化指標値を重み付け計算することにより、生体反応指標値を算出する。

【 0 0 4 6 】

特異判定部 1 3 は、解析対象者について解析された生体反応の変化が、解析対象者以外の他者について解析された生体反応の変化と比べて特異的か否かを判定する。本実施形態において、特異判定部 1 3 は、生体反応解析部 1 2 により複数のユーザのそれぞれについて算出された生体反応指標値に基づいて、解析対象者について解析された生体反応の変化が他者と比べて特異的か否かを判定する。

10

【 0 0 4 7 】

例えば、特異判定部 1 3 は、生体反応解析部 1 2 により複数人のそれぞれについて算出された生体反応指標値の分散を算出し、解析対象者について算出された生体反応指標値と分散との対比により、解析対象者について解析された生体反応の変化が他者と比べて特異的か否かを判定する。

【 0 0 4 8 】

解析対象者について解析された生体反応の変化が他者と比べて特異的である場合として、次の 3 パターンが考えられる。1 つ目は、他者については特に大きな生体反応の変化が起きていないが、解析対象者について比較的大きな生体反応の変化が起きた場合である。2 つ目は、解析対象者については特に大きな生体反応の変化が起きていないが、他者について比較的大きな生体反応の変化が起きた場合である。3 つ目は、解析対象者についても他者についても比較的大きな生体反応の変化が起きているが、変化の内容が解析対象者と他者とで異なる場合である。

20

【 0 0 4 9 】

関連事象特定部 1 4 は、特異判定部 1 3 により特異的であると判定された生体反応の変化が起きたときに解析対象者、他者および環境の少なくとも 1 つに関して発生している事象を特定する。例えば、関連事象特定部 1 4 は、解析対象者について特異的な生体反応の変化が起きたときにおける解析対象者自身の言動を動画像から特定する。また、関連事象特定部 1 4 は、解析対象者について特異的な生体反応の変化が起きたときにおける他者の言動を動画像から特定する。また、関連事象特定部 1 4 は、解析対象者について特異的な生体反応の変化が起きたときにおける環境を動画像から特定する。環境は、例えば画面に表示中の共有資料、解析対象者の背景に写っているものなどである。

30

【 0 0 5 0 】

クラスタリング部 1 5 は、特異判定部 1 3 により特異的であると判定された生体反応の変化（例えば、目線、脈拍、顔の動き、発言内容、声質のうち 1 つまたは複数の組み合わせ）と、当該特異的な生体反応の変化が起きたときに発生している事象（関連事象特定部 1 4 により特定された事象）との相関の程度を解析し、相関が一定レベル以上であると判定された場合に、その相関の解析結果に基づいて解析対象者または事象をクラスタリングする。

40

【 0 0 5 1 】

例えば、特異的な生体反応の変化がネガティブな感情変化に相当するものであり、当該特異的な生体反応の変化が起きたときに発生している事象もネガティブな事象である場合には一定レベル以上の相関が検出される。クラスタリング部 1 5 は、その事象の内容やネガティブな度合い、相関の大きさなどに応じて、あらかじめセグメント化した複数の分類の何れかに解析対象者または事象をクラスタリングする。

【 0 0 5 2 】

同様に、特異的な生体反応の変化がポジティブな感情変化に相当するものであり、当該特異的な生体反応の変化が起きたときに発生している事象もポジティブな事象である場合

50

には一定レベル以上の相関が検出される。クラスタリング部 15 は、その事象の内容やポジティブな度合い、相関の大きさなどに応じて、あらかじめセグメント化した複数の分類の何れかに解析対象者または事象をクラスタリングする。

【0053】

解析結果通知部 16 は、特異判定部 13 により特異的であると判定された生体反応の変化、関連事象特定部 14 により特定された事象、およびクラスタリング部 15 によりクラスタリングされた分類の少なくとも 1 つを、解析対象者の指定者（解析対象者またはオンラインセッションの主催者）に通知する。

【0054】

例えば、解析結果通知部 16 は、解析対象者について他者とは異なる特異的な生体反応の変化が起きたとき（上述した 3 パターンの何れか。以下同様）に発生している事象として解析対象者自身の言動を解析対象者自身に通知する。これにより、解析対象者は、自分がある言動を行ったときに他者とは違う感情を持っていることを把握することができる。このとき、解析対象者について特定された特異的な生体反応の変化も併せて解析対象者に通知するようにしてもよい。さらに、対比される他者の生体反応の変化を更に解析対象者に通知するようにしてもよい。

10

【0055】

例えば、解析対象者が普段どおりの感情で特に意識せずに行った言動、または、解析対象者がある感情を伴って特に意識して行った言動に対して他者が受けた感情と、言動の際に解析対象者自身が抱いていた感情とが相違している場合に、そのときの解析対象者自身の言動が解析対象者に通知される。これにより、自分の意識に反して他者の受けが良い言動や他者の受けが良くない言動などを発見することも可能である。

20

【0056】

また、解析結果通知部 16 は、解析対象者について他者とは異なる特異的な生体反応の変化が起きたときに発生している事象を、特異的な生体反応の変化と共にオンラインセッションの主催者に通知する。これにより、オンラインセッションの主催者は、指定した解析対象者に特有の現象として、どのような事象がどのような感情の変化に影響を与えているのかを知ることができる。そして、その把握した内容に応じて適切な処置を解析対象者に対して行うことが可能となる。

【0057】

また、解析結果通知部 16 は、解析対象者について他者とは異なる特異的な生体反応の変化が起きたときに発生している事象または解析対象者のクラスタリング結果をオンラインセッションの主催者に通知する。これにより、オンラインセッションの主催者は、指定した解析対象者がどの分類にクラスタリングされたかによって、解析対象者に特有の行動の傾向を把握したり、今後起こり得る行動や状態などを予測したりすることができる。そして、それに対して適切な処置を解析対象者に対して行うことが可能となる。

30

【0058】

なお、上記実施形態では、生体反応の変化を所定の基準に従って数値化することによって生体反応指標値を算出し、複数人のそれぞれについて算出された生体反応指標値に基づいて、解析対象者について解析された生体反応の変化が他者と比べて特異的か否かを判定する例について説明したが、この例に限定されない。例えば、以下のようにしてもよい。

40

【0059】

すなわち、生体反応解析部 12 は、複数人のそれぞれについて目線の動きを解析して目線の方向を示すヒートマップを生成する。特異判定部 13 は、生体反応解析部 12 により解析対象者について生成されたヒートマップと他者について生成されたヒートマップとの対比により、解析対象者について解析された生体反応の変化が、他者について解析された生体反応の変化と比べて特異的か否かを判定する。

【0060】

このように、本実施の形態においては、ビデオミーティングの動画像をユーザ端末 10 のローカルストレージに保存し、ユーザ端末 10 上で上述した分析を行うこととしている

50

。ユーザ端末10のマシンスペックに依存する可能性があるとはいえ、動画像の情報を外部に提供することなく分析することが可能となる。

【0061】

生体反応解析部12は、対象者について解析された生体反応の変化に基づいて、複数の対象者間で平準化された評価基準に従って対象者の感情の度合いを評価する感情評価部を有していてもよい。例えば、感情評価部は、生体反応解析部12により対象者について解析された生体反応の変化（生体反応指標値）に基づいて、複数の対象者間で平準化された評価基準に基づく感情反応絶対値を算出する。

【0062】

感情評価部が算出する感情反応絶対値は、例えば、生体反応解析部12により算出された生体反応指標値を、対象者による同じ感情の生起しやすさに応じて調整した値である。例えば、感情評価部は、生体反応解析部12により算出された生体反応指標値に対し、同じ感情を生起する頻度に応じた重み値を乗算することによって感情反応絶対値を算出する。

10

【0063】

例えば、対象者Aについて算出された生体反応指標値と対象者Bについて算出された生体反応指標値とが同じ値であった場合としても、同じ感情の生起しやすさ（同じ感情を生起する頻度）が対象者Aと対象者Bとで異なる場合、感情評価部により算出される感情反応絶対値は対象者Aと対象者Bとで異なる値となる。一例として、感情評価部は、同じ感情を生起しやすいほど重み値が小さくなり、同じ感情を生起しにくいほど重み値が大きくなるような関数に従って感情反応絶対値を算出する。

20

【0064】

このように算出した感情反応絶対値を用いることにより、対象者に関する真の意味での感情の度合いを評価することが可能となり、異なる対象者間で感情の度合いを客観的に対比することができる。

【0065】

なお、感情評価部は、平常時の生体反応に対する現在の生体反応の違いの大きさに基づく感情の程度であって、対象者による同じ感情の生起しやすさに応じて調整された感情の度合いを評価するようにしてもよい。例えば、感情評価部は、生体反応解析部12により算出された生体反応指標値を、平常時の生体反応に対する現在の生体反応の違いの大きさと、対象者による同じ感情の生起しやすさとに応じて調整することによって感情反応絶対値を算出する。このように算出される感情反応絶対値は、平常時の生体反応に対する現在の生体反応の違いの大きさに基づく感情の程度を表す値であって、対象者が同じ感情を生起しやすいまたは生起しにくい度合いに応じて調整された値である。

30

【0066】

また、上記実施形態では、同じ感情の生起しやすさを表す尺度として、同じ感情を生起する頻度を用いる例について説明したが、これに限定されない。例えば、同じ感情を生起する頻度に代えてまたは加えて、対象者の性質または性格を用いるようにしてもよい。

【0067】

なお、図5に示されるように、画面に表示されていない参加者を含めて生体反応解析部12aにより解析された生体反応の変化を示す情報を提示する反応情報提示部13aを備えていてもよい。例えば、反応情報提示部13aは、生体反応の変化を示す情報をオンラインセッションの主導者、進行者または管理者（以下、まとめて主催者という）に提示する。オンラインセッションの主催者は、例えばオンライン授業の講師、オンライン会議の議長やファシリテータ、コーチングを目的としたセッションのコーチなどである。オンラインセッションの主催者は、オンラインセッションに参加する複数のユーザの中の一人であるのが普通であるが、オンラインセッションに参加しない別人であってもよい。

40

【0068】

このようにすることにより、オンラインセッションの主催者は、複数人でオンラインセッションが行われる環境において、画面に表示されていない参加者の様子も把握することができる。

50

【 0 0 6 9 】

< 第 1 の実施の形態 >

図 6 乃至図 8 を参照して、上述した構成に基づく本システムの第 1 の実施の形態を説明する。本実施の形態においては、動画像内に含まれる対象者の顔画像を所定のフレーム単位ごとに認識するとともに、対象者の音声を認識する。認識は複数の対象者に対して行うこととしてもよい。

【 0 0 7 0 】

続いて、認識された顔画像及び音声の双方に基づいて、前記対象者の複数の観点による感情が数値化され評価される。評価された感情はその度合いとともにグラフにプロットされる。

10

【 0 0 7 1 】

グラフは動画の時系列に沿ってプロットされる。図 6 に示されるように、位置の動画像内における一の対象者について、幸福度 (Happy Score) という一の観点で評価された数値をプロットすることとしてもよい。また、図 7 に示されるように、同一の動画像内における複数の対象者について、対象者ごとに複数の観点による感情の度合いをプロットすることとしてもよい。

【 0 0 7 2 】

更には、図 8 に示されるように、ある対象者が含まれる複数の動画像 (例えばユーザがオンラインで受講した複数の授業の動画像や、ユーザが参加した複数のオンライン会議の動画像等) における感情の平均値 (最高値、最低値、最頻値等でもよい) の度合いについて、横軸に動画像のタイトル、縦軸に感情の度合いをとってプロットすることとしてもよい。これにより、当該対象者が複数のビデオミーティングへの参加経験に伴って感情がどのように変化したのかにいてが可視化可能となる。

20

【 0 0 7 3 】

図示されたグラフは、折れ線グラフであるが、棒グラフやヒートマップ等どのようなものであってもよい。また、感情の種類毎に色分けして表示することとしてもよい。

【 0 0 7 4 】

各グラフの縦軸は、例えば、複数の対象者間で平準化された評価基準に従って対象者の感情の度合いを感情の種類毎にプロットすることとしてもよい。これにより、同一の軸の尺度 (例えば 0 ~ 100) でも客観的な評価を行うことが可能になる。また、当該対象者の平常時の生体反応に対する現在の生体反応の違いの大きさに基づく感情の程度であって、同じ感情の生起しやすさに応じて調整された感情の度合いを評価して、当該調整された感情の度合いを、感情の種類毎にプロットすることとしても同様の効果を得られる。

30

【 0 0 7 5 】

< 第 2 の実施の形態 >

図 9 及び図 10 を参照して本発明の第 2 の実施の形態を説明する。図に示されるように本実施の形態による画面には、検索ワードを受け付ける検索ワードボックスが表示されている。検索ワードボックスにワードが入力されると (例えば、「塩基」と入力されると) 入力された検索ワードに対応する音声が含まれる動画像内の所定範囲が抽出され表示される。

40

【 0 0 7 6 】

本実施の形態によるシステムは、動画像内に含まれる対象者の少なくとも顔画像を所定のフレーム単位ごとに認識する顔認識手段と、動画像内に含まれる対象者の少なくとも音声を認識する音声認識手段と、検索ワードの入力を受け付ける検索受付手段とを備えている。かかる構成によれば、図示されているように、動画像ファイル「20201230_Biology_Tanaka」の中から「塩基」と話されている範囲に対応する動画像の一部と、テキスト情報を表示することが可能となる。

【 0 0 7 7 】

抽出された動画像と対応するテキスト情報を同一画面内に表示することにより、当該ワードを話していた際にどのような状況・表情をしていたのかの確認を容易に行うことがで

50

きる。

【 0 0 7 8 】

また、図示される画面例には「塩基」というワードが3か所に抽出されている。表示されている検索ワード（例えば、一番最初の「塩基」）を選択すると、当該ワードを話していたときのフレームを含む動画像が（部分的に）再生される。

【 0 0 7 9 】

なお、動画像内に検索ワードが複数抽出されていた場合に、当該検索ワードを含む複数の部分的動画像をつなぎ合わせたダイジェスト動画像を生成することとしてもよい。これにより、短時間で効率的に検索したワード周囲の動画像を確認することができる。

【 0 0 8 0 】

また、あらかじめ検索ワードの登録を受け付けておき、動画像内に、登録された検索ワードが抽出されていた場合にアラートを発したり、当該検索ワードを含む複数の部分的動画像をつなぎ合わせたダイジェスト動画像を自動生成することとしてもよい。

【 0 0 8 1 】

例えば図10に示されるように、alert valueの列に示されるような単語を登録しておき、当該単語が登場する動画像の情報の所在を示す情報を関連付けることとしてもよい。

【 0 0 8 2 】

更には、ビデオミーティング中に当該単語が発された場合に動画像内の当該地点への再生リンクを生成することとしてもよい。

【 0 0 8 3 】

登録された単語は事前に何らかのタグ付け（alert pattern）をしておくことにより管理が容易にやすくなる。

【 0 0 8 4 】

動画中に当該単語が何回発生したかをカウントすることとしてもよい。例えば、認識された音声の中（又は変換されたテキスト情報）の中における単語の登場回数をカウントすることとすればよい。

【 0 0 8 5 】

< 第3の実施の形態 >

図11を参照して本発明の第3の実施の形態によるシステムを説明する。本実施の形態においては、特に、相手側がカメラをOFF（カメラから得られた視覚的な情報が提供されない状態）にした際であっても、他方に対して、当該ユーザがカメラの範囲内（即ち、コンピュータの正面）に存在していることやその時の表情を提供することが可能になる。

【 0 0 8 6 】

授業や会議のように複数の参加者によりビデオミーティングがされると、カメラから取得された情報を悪意を持って利用されることが想定される。例えば、他人の顔をスクリーンショットで記録したり、背景に映りこんだ部屋等のプライベート空間を見られたりすることがある。このため、参加者の中にはカメラ機能をOFFにして参加するものも少なくない。

【 0 0 8 7 】

しかしながら、カメラをOFFにした場合、ホスト側からは、相手の顔が見えないことから、講義や会議にきちんと参加しているのかどうかを確かめることができない。このように、ホスト側からは、カメラの正面にゲストがきちんと参加しているのかどうかを知りたい一方、ゲスト側からは、きちんと参加していることを伝えたくてもカメラをONにしたくない状況が生じ得る。

【 0 0 8 8 】

そこで本実施の形態においては、ゲストユーザの端末は、ゲストユーザのカメラの動画像を取得、当該動画像内に含まれる少なくとも顔画像を所定のフレーム単位ごとに識別するとともに、識別した顔画像を顔情報に変換してホストユーザ端末に提供する。

【 0 0 8 9 】

顔情報を例示すれば、カメラの正面にいるかいないか、顔の向き、カメラでとらえた表

10

20

30

40

50

情から得られる感情、当該感じ用に基づいて生成されたオブジェクト情報などが挙げられるがこれに限られない。

【0090】

このような構成によれば、ゲストユーザのプライベートな情報（カメラで得られる視覚的情報）はホスト側に共有することなく、ゲストがカメラで検出されていることや、画面を正視していること、表情などの情報を提供することができる。

【0091】

<第4の実施の形態>

図12を参照して本発明の第4の実施の形態によるシステムを説明する。本実施の形態においては、動画像から取得されたテキスト情報に感情に関する情報を付加するものである。

10

【0092】

本実施の形態によるシステムは、テキストに感情に関する評価値を関連付けて出力する。例えば、図12に示されるように、所定以上の大きさの声で話された内容はフォントサイズを大きくしたり、カメラ目線で話された単語にはアンダーラインが引かれたりして表示される。

【0093】

すなわち、本実施の形態によれば、動画像から得られるユーザの直接的な言動や、当該言動を分析した結果等に基づいて、テキストに所定の加工を加えることが可能になる。

【0094】

ユーザの言動としては、身振り、手ぶり、ジェスチャー、顔の向き、顔の動き等の他、音声の抑揚、大きさ、高低、速さ等を事前に登録することができる。また、テキストへ加える加工は、フォントサイズの変更、太さの変更、斜体に変更、文字色の変更、影の付加、フォント種類の変更等などが例示できる。

20

【0095】

また、例えば、音声に関する条件（例えば、一定以上の大きさ）の登録を受け付けておき、登録された条件を満たす音声に対応するテキストを所定の表示形式に変更（大きく表示）することとしてもよい。

【0096】

<第5の実施の形態>

図13を参照して本発明の第5の実施の形態によるシステムを説明する。本実施の形態においては、動画像から取得された複数の観点による分析評価情報それぞれの間における矛盾（評価値の乖離）によって、潜在的な感情の分析を行うものである。

30

【0097】

具体的には、本システムは、顔画像及び音声の双方に基づいて複数の観点による評価値を算出する。複数の観点による評価値の相対的な乖離が一定以上になった場合に通知手段による通知が行われる。

【0098】

乖離のパターンとしては、例えば、顔画像及び音声の夫々から得られる評価値が一定以上乖離していた場合、顔画像から取得できる顔の動き及び音声の夫々から得られる評価値が一定以上乖離していた場合、顔画像から評価される感情の度合いと、音声とから得られる評価値が一定以上乖離していた場合などが例示できる。

40

【0099】

例えば、図13示されるように、ある地点において、幸福度を表すグラフの点（H01）と、怒り度を表すグラフの点（A01）とを比較すると、幸福度の方が優位に高い値を示しており、当該ユーザの感情としては幸福である（よいコミュニケーションである）と感じていることが推定される。一方、別の地点における幸福度を表すグラフの点（H02）と、怒り度を表すグラフの点（A02）とを比較すると、いずれも同じ程度の値を示しており、感情的には矛盾することとなる。このような場合、例えば、顔は笑っているけれども、内心は疑っていたり、不審に思っていたりすることなどが考えられる。

50

【0100】

本実施の形態によれば、評価値の乖離に基づく真の感情を推定することが可能となる。乖離の判断としては、所定の相関関係に基づいて判断すればよく、機械学習的な判断を行うこととしてもよい。

【0101】

<第6の実施の形態>

図14を参照して本発明の第6の実施の形態によるシステムを説明する。本実施の形態においては、同時に参加している多数のユーザによる動画像から分析評価された評価値を当該動画像のそれぞれに関連付けて表示する。これにより、参加者全体の状態が、ヒートマップのように一括して理解できるようになる。

10

【0102】

本システムは、同時にビデオミーティングに参加している複数のユーザのカメラから得られた動画像を一覧に表示する表示手段と、複数の動画像の夫々について、評価値に関連付けられたオブジェクトを関連付けて表示する関連付け手段を備えている。

【0103】

例えば、関連付け手段は、評価値に応じたヒートマップを生成し、対応する色を動画像の夫々にオーバーレイして表示することとしてもよい。図14に示されるように、怒り度が高いユーザをグレースアウトしたりすることとしてもよい。

【0104】

また、関連付け手段は、評価値に応じた喜怒哀楽アイコンを生成し、対応するアイコンを動画像の夫々に併記して表示することとしてもよい。

20

【0105】

<第7の実施の形態>

再度図4を参照して、本発明の第7の実施の形態を説明する。本実施の形態においては、上述した特異判定部13によって、判定された同一のユーザについて以前とは異なる特異的な反応を解析した場合に所定の端末(画面等)にアラートを出す通知手段を備えている。

【0106】

通知手段は、反応が閾値の範囲を超えた場合に通知を行う。例えば、普段は怒らないユーザの怒り度数が以前よりも極めて高くなった場合などに通知を行うこととしてもよいし、普段なかなか笑わないユーザが笑った際に通知を行うこととしてもよい。どのような条件のときにどのような通知を行うかはあらかじめ登録することができる。

30

【0107】

また、通知手段は、同一の時間帯(1度の会議、1度の授業等)に行われたオンラインミーティング内において、ユーザの特定の反応が所定の回数を超えた場合に前記通知を行うこととしてもよい。

【0108】

<ハードウェア構成の補足>

本明細書において説明した装置による一連の処理は、ソフトウェア、ハードウェア、及びソフトウェアとハードウェアとの組合せのいずれを用いて実現されてもよい。本実施形態に係る情報共有支援装置10の各機能を実現するためのコンピュータプログラムを作製し、PC等を実装することが可能である。また、このようなコンピュータプログラムが格納された、コンピュータで読み取り可能な記録媒体も提供することが可能である。記録媒体は、例えば、磁気ディスク、光ディスク、光磁気ディスク、フラッシュメモリ等である。また、上記のコンピュータプログラムは、記録媒体を用いずに、例えばネットワークを介して配信されてもよい。

40

【0109】

また、本明細書においてフローチャート図を用いて説明した処理は、必ずしも図示された順序で実行されなくてもよい。いくつかの処理ステップは、並列的に実行されてもよい。また、追加的な処理ステップが採用されてもよく、一部の処理ステップが省略されても

50

よい。

【 0 1 1 0 】

また、本明細書に記載された効果は、あくまで説明的または例示的なものであって限定的ではない。つまり、本開示に係る技術は、上記の効果とともに、または上記の効果に代えて、本明細書の記載から当業者には明らかな他の効果を奏しうる。

【 符号の説明 】

【 0 1 1 1 】

- 1 0、2 0 ユーザ端末
- 3 0 ビデオミーティングサービス端末
- 4 0 評価端末

10

20

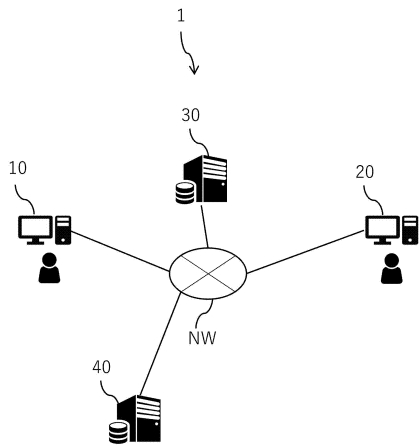
30

40

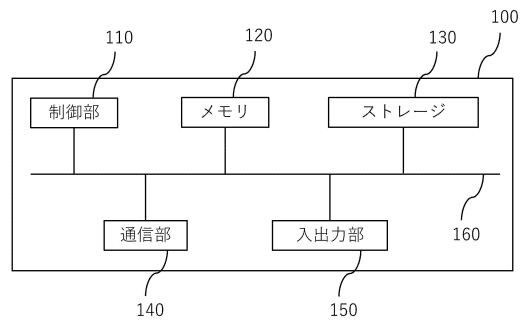
50

【図面】

【図 1】

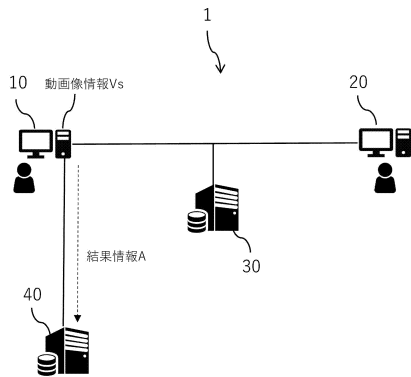


【図 2】

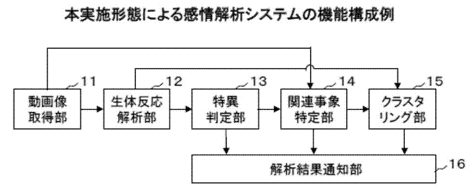


10

【図 3】



【図 4】



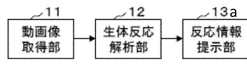
20

30

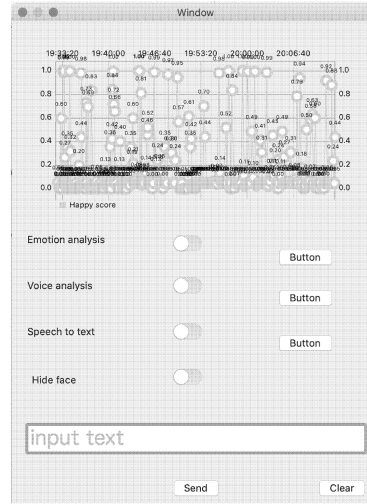
40

50

【 図 5 】

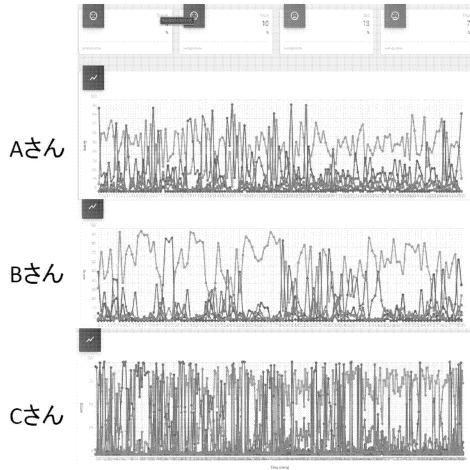


【 図 6 】



10

【 図 7 】



【 図 8 】



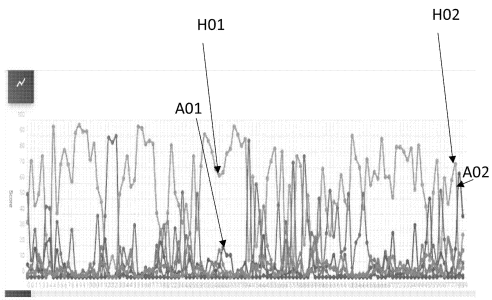
20

30

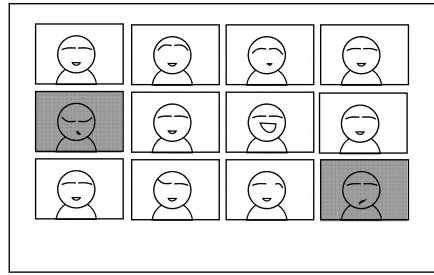
40

50

【 1 3 】



【 1 4 】



10

20

30

40

50

フロントページの続き

- (56)参考文献 特開2018-068618(JP,A)
特開2019-058625(JP,A)
特開2017-215468(JP,A)
- (58)調査した分野 (Int.Cl., DB名)
- A61B5/06-5/22
 - G06T1/00-1/40
 - 3/00-7/90
 - G06V10/00-20/90
 - 30/418
 - 40/16
 - 40/20
 - H04M3/00
 - 3/16-3/20
 - 3/38-3/58
 - 7/00-7/16
 - 11/00-11/10
 - H04N7/10
 - 7/14-7/173
 - 7/20-7/56
 - 21/00-21/858