

(19)日本国特許庁(JP)

## (12)特許公報(B1)

(11)特許番号  
特許第7100938号  
(P7100938)

(45)発行日 令和4年7月14日(2022.7.14)

(24)登録日 令和4年7月6日(2022.7.6)

(51)国際特許分類 F I  
H 0 4 N 7 / 1 5 ( 2 0 0 6 . 0 1 ) H 0 4 N 7 / 1 5

請求項の数 2 (全17頁)

(21)出願番号	特願2022-518706(P2022-518706)	(73)特許権者	520408744 株式会社 I ' m b e s i d e y o u 東京都世田谷区成城2丁目34番2号 村上ハイツ301
(86)(22)出願日	令和3年3月22日(2021.3.22)	(74)代理人	110002790 One i p特許業務法人
(86)国際出願番号	PCT/JP2021/011816	(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
審査請求日	令和4年5月11日(2022.5.11)	審査官	松元 伸次
早期審査対象出願			

最終頁に続く

(54)【発明の名称】 動画像分析プログラム

## (57)【特許請求の範囲】

## 【請求項1】

複数人の参加者でオンラインセッションが行われる環境において、オンラインセッション中に参加者が画面に表示されているか否かによらず、前記参加者を撮影することによって得られる動画像をもとに前記参加者の反応を解析する動画像分析システムであって、前記オンラインセッション中に前記参加者を撮影することによって得られる動画像を取得する動画像取得部と、前記動画像取得部により取得された動画像に基づいて、前記参加者について生体反応の変化を解析する解析部と、前記動画像内に含まれる単語を分析して所定の文書ベクトルに変換する単語分析部と、変換した文書ベクトルを所定の次元に当該単語と共にプロットするプロット生成部と、オンラインセッションの開始前に用意された事前情報に含まれる前記文書ベクトルと、前記オンラインセッションの前記動画像の前記文書ベクトルとを比較して重なり具合を評価する評価部と、を備える、動画像分析システム。

## 【請求項2】

請求項1に記載の動画像分析システムであって、前記プロットを操作された際に対応する単語が含まれる動画像を当該単語を含むフレームから再生する再生部を更に備える、動画像分析システム。

## 【発明の詳細な説明】

## 【技術分野】

## 【0001】

本発明は、複数人の参加者でオンラインセッションが行われる環境において、オンラインセッション中に参加者が画面に表示されているか否かによらず、参加者を撮影することによって得られる動画像をもとに参加者の反応を解析する動画像分析システムに関する。

## 【背景技術】

## 【0002】

発言者の発言に対して他者が受ける感情を解析する技術が知られている（例えば、特許文献1参照）。また、対象者の表情の変化を長期間にわたり時系列的に解析し、その間に抱いた感情を推定する技術も知られている（例えば、特許文献2参照）。さらに、感情の変化に最も影響を与えた要素を特定する技術も知られている（例えば、特許文献3～5参照）。さらにまた、対象者の普段の表情と現在の表情とを比較して、表情が暗い場合にアラートを発する技術も知られている（例えば、特許文献6参照）。また、対象者の平常時（無表情時）の表情と現在の表情とを比較して、対象者の感情の度合いを判定するようにした技術も知られている（例えば、特許文献7～9参照）。更に、また、組織としての感情や、個人が感じるグループ内の雰囲気进行分析する技術も知られている（例えば、特許文献10、11参照）。

10

## 【先行技術文献】

## 【特許文献】

20

## 【0003】

【文献】特開2019-58625号公報

特開2016-149063号公報

特開2020-86559号公報

特開2000-76421号公報

特開2017-201499号公報

特開2018-112831号公報

特開2011-154665号公報

特開2012-8949号公報

特開2013-300号公報

30

特開2011-186521号公報

WO15/174426号公報

## 【発明の概要】

## 【発明が解決しようとする課題】

## 【0004】

上述したすべての技術は、現実空間におけるコミュニケーションが主である状況におけるサブ的な機能にすぎない。即ち、昨今の業務のDX(Digital Transformation)化や、世界的な感染症の流行等を受け、業務や授業等のコミュニケーションがオンラインで行われることが主とされる状況に生まれたものではない。

## 【0005】

40

本発明は、会議や講義等、オンラインコミュニケーションが主となる状況において、より効率的なコミュニケーションを行うために、これらのコミュニケーションを客観的に評価することを目的とする。

## 【課題を解決するための手段】

## 【0006】

本発明によれば、

複数人の参加者でオンラインセッションが行われる環境において、オンラインセッション中に参加者が画面に表示されているか否かによらず、前記参加者を撮影することによって得られる動画像をもとに前記参加者の反応を解析する動画像分析システムであって、前記オンラインセッション中に前記参加者を撮影することによって得られる動画像を取得

50

する動画像取得部と、  
前記動画像取得部により取得された動画像に基づいて、前記参加者について生体反応の変化を解析する解析部と、  
前記動画像内に含まれる単語を分析して所定の文書ベクトルに変換する単語分析部と、  
変換した文書ベクトルを所定の次元に当該単語と共にプロットするプロット生成部と、  
動画像分析システムが得られる。

【発明の効果】

【0007】

本開示によれば、ビデオセッションの動画像を分析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。

10

【0008】

特に、本発明によれば、オンラインコミュニケーションが主となる状況において、より効率的なコミュニケーションを行うために、交わされたコミュニケーションを客観的に評価することができる。

【図面の簡単な説明】

【0009】

【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評価端末の機能ブロック図の一例である。

【図3】本発明の実施の形態による評価端末の機能構成例1を示す図である。

【図4】本発明の実施の形態による評価端末の機能構成例2を示す図である。

20

【図5】本発明の実施の形態による評価端末の機能構成例3を示す図である。

【図6】図6の機能構成例3による画面表示例である。

【図7】図6の機能構成例3による他の画面表示例である。

【図8】本発明の実施の形態による評価端末の機能構成例3の他の構成を示す図である。

【図9】本発明の実施の形態による評価端末の機能構成例3の他の構成を示す図である。

【図10】本発明の実施の形態によるシステムの機能ブロック図である。

【図11】本発明の実施の形態によるシステムによって生成されるプロットのイメージ図である。

【図12】本発明の実施の形態によるシステムによるシステムの処理のイメージを示す図である。

【図13】本発明の実施の形態によるシステムによって生成されるプロットのイメージ図である。

30

【発明を実施するための形態】

【0010】

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]

複数人の参加者でオンラインセッションが行われる環境において、オンラインセッション中に参加者が画面に表示されているか否かによらず、前記参加者を撮影することによって得られる動画像をもとに前記参加者の反応を解析する動画像分析システムであって、  
前記オンラインセッション中に前記参加者を撮影することによって得られる動画像を取得する動画像取得部と、

40

前記動画像取得部により取得された動画像に基づいて、前記参加者について生体反応の変化を解析する解析部と、

前記動画像内に含まれる単語を分析して所定の文書ベクトルに変換する単語分析部と、  
変換した文書ベクトルを所定の次元に当該単語と共にプロットするプロット生成部と、  
動画像分析システム。

[項目2]

項目1に記載の動画像分析システムであって、

前記プロットを操作された際に対応する単語が含まれる動画像を当該単語を含むフレームから再生する再生部を更に備える、

動画像分析システム。

50

## [ 項目 3 ]

項目 1 又は項目 2 に記載の動画像分析システムであって、  
 オンラインセッションの開始前に用意された事前情報に含まれる前記文書ベクトルと、前記オンラインセッションの前記動画像の前記文書ベクトルとを比較して重なり具合を評価する評価部を更に備える、  
 動画像分析システム。

## [ 項目 4 ]

項目 1 乃至項目 3 のいずれかに記載の動画像分析システムの構成を備えた動画像分析装置。

## [ 項目 5 ]

項目 1 乃至項目 3 のいずれかに記載の動画像分析システムの構成を動画像分析装置に機能させる動画像分析プログラム。

10

## [ 項目 6 ]

項目 1 乃至項目 3 のいずれかに記載の動画像分析システムの構成をステップとして実行する動画像分析方法。

## 【 0 0 1 1 】

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。

## 【 0 0 1 2 】

## &lt; 基本機能 &gt;

20

本実施形態のビデオセッション評価システムは、複数人でビデオセッション（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。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。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。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。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。

30

40

## 【 0 0 1 3 】

本実施の形態によるビデオセッション評価システムは、複数の端末間においてビデオセッションセッションが確立された場合に、当該ビデオセッションから取得される少なくとも

50

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。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セッションや機密情報を含むビデオセッションであっても、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（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。

#### 【0014】

図1に示されるように、本実施の形態によるビデオセッション評価システムは、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に双方向のビデオセッションを提供するビデオセッションサービス端末30と、ビデオセッションに関する評価の一部を行う評価端末40とを備えている。

10

#### 【0015】

##### <ハードウェア構成例>

以下に説明する各機能ブロック、機能単位、機能モジュールは、例えばコンピュータ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、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。本明細書において説明するシステム及び端末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。

20

#### 【0016】

本実施の形態による評価端末は、ビデオセッションサービス端末から動画像を取得し、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)。

30

#### 【0017】

##### <動画の取得方法>

図3に示されるように、ビデオセッションサービス端末が提供するビデオセッションサービス(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から出力可能となっている。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。

40

#### 【0018】

##### <機能構成例1>

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セッション評価システムは、ユーザ端末10が有する機能構成として実現され

50

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。

【0019】

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。

【0020】

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。

【0021】

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。

【0022】

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。

【0023】

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。

【0024】

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。

【0025】

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。

10

20

30

40

50

## 【 0 0 2 6 】

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 1 2 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。

10

## 【 0 0 2 7 】

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 1 2 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。

## 【 0 0 2 8 】

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 1 2 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。

20

## 【 0 0 2 9 】

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。

30

## 【 0 0 3 0 】

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。

## 【 0 0 3 1 】

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。

40

## 【 0 0 3 2 】

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。

50

## 【 0 0 3 3 】

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。

## 【 0 0 3 4 】

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。

## 【 0 0 3 5 】

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 3 6 】

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。

## 【 0 0 3 7 】

解析結果通知部 1 6 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、関連事象特定部 1 4 により特定された事象、およびクラスタリング部 1 5 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。

## 【 0 0 3 8 】

例えば、解析結果通知部 1 6 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。

## 【 0 0 3 9 】

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。

## 【 0 0 4 0 】

また、解析結果通知部 1 6 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

10

20

30

40

50

ヨ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。

【 0 0 4 1 】

また、解析結果通知部 1 6 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。

10

【 0 0 4 2 】

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。

【 0 0 4 3 】

すなわち、生体反応解析部 1 2 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 1 3 は、生体反応解析部 1 2 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。

20

【 0 0 4 4 】

このように、本実施の形態においては、ビデオセッションの動画をユーザ端末 1 0 のローカルストレージに保存し、ユーザ端末 1 0 上で上述した分析を行うこととしている。ユーザ端末 1 0 のマシンスペックに依存する可能性があるとはいえ、動画の情報に外部に提供することなく分析することが可能となる。

【 0 0 4 5 】

< 機能構成例 2 >

図 5 に示すように、本実施形態のビデオセッション評価システムは、機能構成として、動画取得部 1 1、生体反応解析部 1 2 および反応情報提示部 1 3 a を備えていてもよい。

30

【 0 0 4 6 】

反応情報提示部 1 3 a は、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 1 2 a 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。例えば、反応情報提示部 1 3 a 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催者、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

【 0 0 4 7 】

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。

【 0 0 4 8 】

< 機能構成例 3 >

図 6 は、本実施形態による構成例を示すブロック図である。図 6 に示すように、本実施形態のビデオセッション評価システムは、機能構成として、上述した実施の形態 1 と類似する機能については同一の参照符号を付して説明を省略することができる。

【 0 0 4 9 】

50

本実施の形態によるシステムは、ビデオセッションの映像を取得するカメラ部及び音声を取得するマイク部と、動画像を分析及び評価する解析部、取得した動画像を評価することによって得られた情報に基づいて表示オブジェクト（後述する）を生成するオブジェクト生成部、前記ビデオセッション実行中にビデオセッションの動画像と表示オブジェクトの両方を表示する表示部と、を備えている。

【0050】

解析部は、上述した説明と同様に、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。各要素の機能については上述したとおりである。

【0051】

図7に示されるように、オブジェクト生成部は、解析部によってビデオセッションから取得される動画像を解析した結果に基づいて、必要に応じて、当該認識した顔の部分を示すオブジェクト50と、上述した分析・評価した内容を示す情報100を当該動画像に重畳して表示する。当該オブジェクト50は、複数人の顔が動画像内に移っている場合には、複数人全員の顔を識別し、表示することとしてもよい。

【0052】

また、オブジェクト50は、例えば、相手側の端末において、ビデオセッションのカメラ機能を停止している場合（即ち、物理的にカメラを覆う等ではなく、ビデオセッションのアプリケーション内においてソフトウェア的に停止している場合）であっても、相手側のカメラで相手の顔を認識していた場合には、相手の顔が位置している部分にオブジェクト50やオブジェクト100を表示することとしてもよい。これにより、カメラ機能がオフになっていたとしても、相手側が端末の前にいることがお互い確認することが可能となる。この場合、例えば、ビデオセッションのアプリケーションにおいては、カメラから取得した情報を非表示にする一方、解析部によって認識された顔に対応するオブジェクト50やオブジェクト100のみを表示することとしてもよい。また、ビデオセッションから取得される映像情報と、解析部によって認識され得られた情報とを異なる表示レイヤーに分け、前者の情報に関するレイヤーを非表示にすることとしてもよい。

【0053】

オブジェクト50やオブジェクト100は、複数の動画像を表示する領域がある場合には、すべての領域又は一部の領域のみに表示することとしてもよい。例えば、図8に示されるように、ゲスト側の動画像のみに表示することとしてもよい。

【0054】

以上説明した基本構成例1乃至基本構成例3において説明した発明の実施の形態は、単独の装置として実現されてもよく、一部または全部がネットワークで接続された複数の装置（例えばクラウドサーバ）等により実現されてもよい。例えば、各端末10の制御部110およびストレージ130は、互いにネットワークで接続された異なるサーバにより実現されてもよい。即ち、本システムは、ユーザ端末10、20と、ユーザ端末10、20に双方向のビデオセッションを提供するビデオセッションサービス端末30と、ビデオセッションに関する評価を行う評価端末40とを含んでいるところ、以下のような構成のバリエーション組み合わせが考えられる。

（1）すべてをユーザ端末のみで処理

図9に示されるように、解析部による処理をビデオセッションを行っている端末で行うことにより、（一定の処理能力は必要なものの）ビデオセッションを行っている時間と同時に（リアルタイムに）分析・評価結果を得ることができる。

（2）ユーザ端末と評価端末とで処理

図10に示されるように、ネットワーク等で接続された評価端末に解析部を備えさせることとしてもよい。この場合、ユーザ端末で取得された動画像は、ビデオセッションと同時に又は事後的に評価端末に共有され、評価端末における解析部によって分析・評価されたのちに、オブジェクト50及びオブジェクト100の情報がユーザ端末に動画像データと共に又は別に（即ち、少なくとも解析データを含む情報が）共有され表示部に表示される。

10

20

30

40

50

## 【 0 0 5 5 】

上述した機能構成例 1 乃至機能構成例 3 の各構成又はそれらの組み合わせを用いて、以下のシステムが実現する。

## 【 0 0 5 6 】

## &lt; 実施の形態 &gt;

本発明の実施の形態による動画像分析システム（以下、単に「システム」という）は、複数人の参加者でオンラインセッションが行われる環境において、当該参加者の全員又は特定の参加者のみを撮影することによって得られる動画像をもとに参加者の反応を解析・分析するものである。分析は、オンラインセッション中に参加者が画面に表示されているか否かによらず行われるものとしてもよい。

10

## 【 0 0 5 7 】

図 10 に示されるように、本実施の形態によるシステムは、動画像取得部と、解析部と、テキスト抽出部と、単語分析部と、プロット生成部とを備えている。動画像取得部は、オンラインセッション中に参加者を撮影することによって得られる動画像を取得する。解析部は、動画像取得部により取得された動画像に基づいて、参加者について生体反応の変化を解析する（図 3 乃至図 5 等も併せて参照）。

## 【 0 0 5 8 】

図 11 に示されるように、単語分析部は、動画像内に含まれる単語を分析して所定の文書ベクトルに変換する。プロット生成部は、変換した文書ベクトルを所定の次元に当該単語と共にプロットする。図示される例は、講師と学生とで行われた「講義」を解析したものである。図示されるように、縦軸付近には動画像内で用いられている単語として学業試験に関する「進路」「合格」「当日」という単語がプロットされており、横軸付近には動画像内で用いられている単語として特定の学科（数学）学業試験に関する「数学」「方程式」「解」という単語がプロットされている。

20

## 【 0 0 5 9 】

図 12 に示されるように、例えば、図 11 に示されるプロットのうち「進路」の点を選択すると、「進路」という単語が含まれる動画が抽出され、（順次又は選択させて上で）再生される。動画像は、当該単語を含むフレームから再生される。

## 【 0 0 6 0 】

なお、図 13 に示されるように、オンラインセッションの開始前に用意された事前情報に含まれる文書ベクトル（事前ベクトル）と、オンラインセッションで得られた動画像の文書ベクトル（事後ベクトル）とを比較して重なり具合を評価する評価部を更に備えていてもよい。図示されたように、黒塗りのポイントは事前ベクトルを表し、斜線のポイントは事後ベクトルを表す。

30

## 【 0 0 6 1 】

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よい。

## 【 0 0 6 2 】

以上説明した実施の形態を適宜組み合わせる実施することとしてもよい。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。

40

## 【 符号の説明 】

## 【 0 0 6 3 】

- 10、20 ユーザ端末
- 30 ビデオセッションサービス端末
- 40 評価端末

50

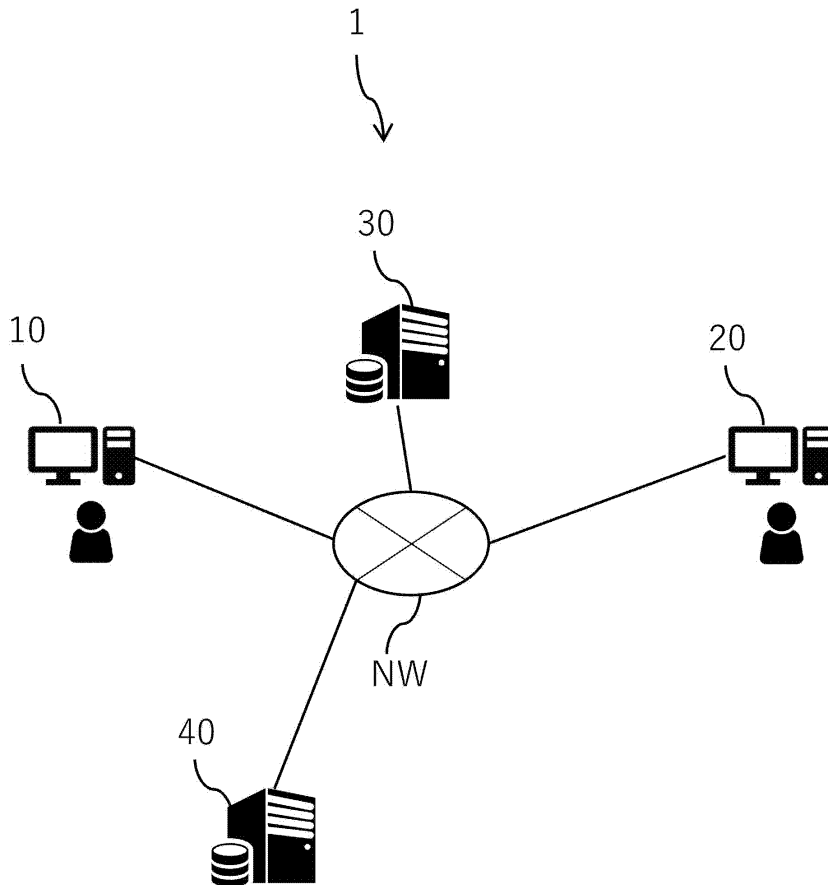
【要約】

【課題】会議や講義等、オンラインコミュニケーションが主となる状況において、より効率的なコミュニケーションを行うために、これらのコミュニケーションを客観的に評価すること。

【解決手段】本開示のシステムは、オンラインセッション中に参加者を撮影することによって得られる動画像を取得する動画像取得部と、動画像取得部により取得された動画像に基づいて、参加者について生体反応の変化を解析する解析部と、動画像内に含まれる単語を分析して所定の文書ベクトルに変換する単語分析部と、変換した文書ベクトルを所定の次元に当該単語と共にプロットするプロット生成部と、を備える。

【選択図】図 1

10



20

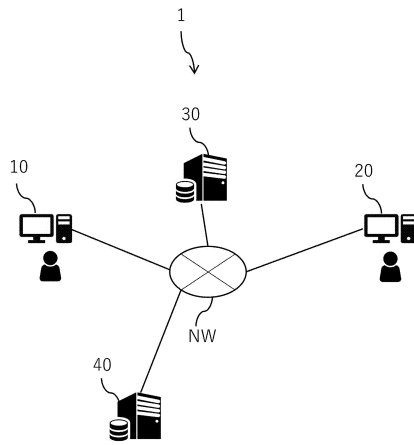
30

40

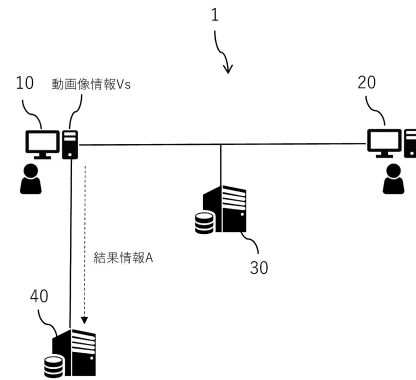
50

【図面】

【図 1】



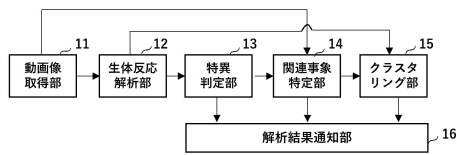
【図 2】



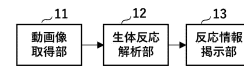
10

【図 3】

解析部の構成例



【図 4】



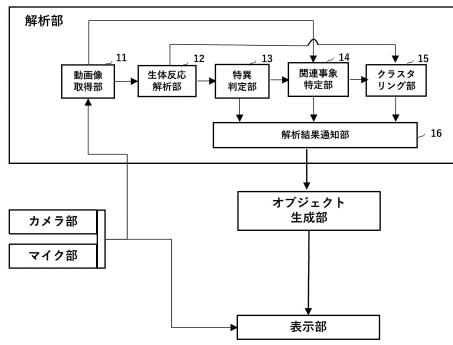
20

30

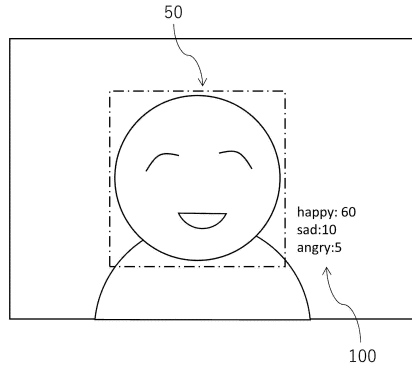
40

50

【図5】

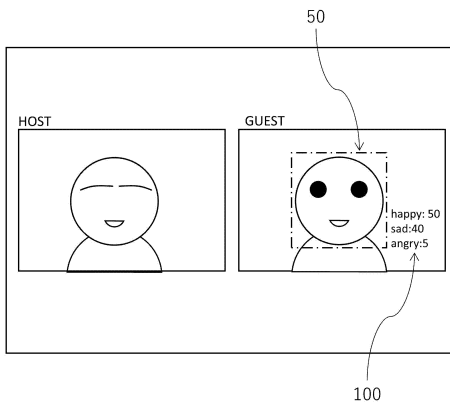


【図6】

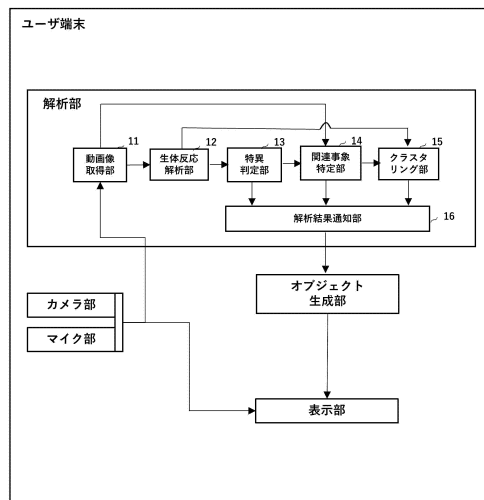


10

【図7】



【図8】



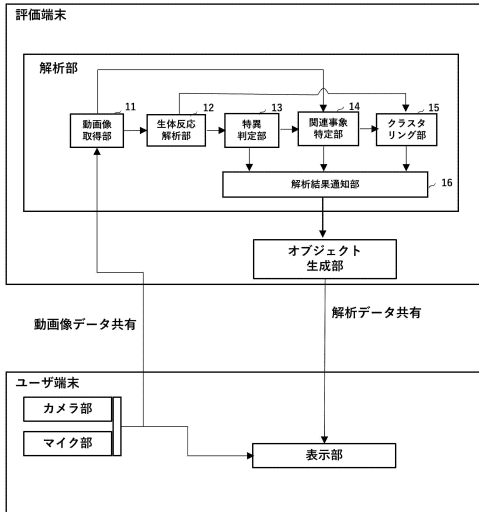
20

30

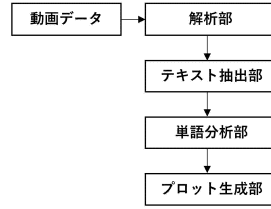
40

50

【 図 9 】

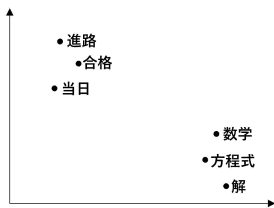


【 図 10 】

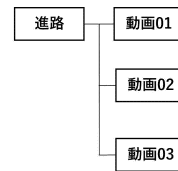


10

【 図 11 】



【 図 12 】



20

30

40

50



## フロントページの続き

(56)参考文献 米国特許出願公開第2003/0154072 (US, A1)

特開2015-186127 (JP, A)

特開2000-124901 (JP, A)

特開2021-22909 (JP, A)

(58)調査した分野 (Int.Cl., DB名)

G06T1/00 - 1/40

3/00 - 7/90

G06V10/00 - 20/90

30/418

40/16

40/20

G07C1/00 - 15/00

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