

明 細 書

発明の名称：

注力判定システム、コミュニケーション解析システム、及び、注力判定方法

技術分野

[0001] 本発明は、注力判定システム、コミュニケーション解析システム、及び、注力判定方法に関する。

背景技術

[0002] 会社などの組織においては、従業員同士がコミュニケーションを密にとって各自のタスクに取り組むことが重要である。このようなコミュニケーションに関する技術として、特許文献1には、会議や打ち合わせなどの場における知的活動に対する評価をするために対象者の心理状態を情報として提供する情報提供装置が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2004-112518号公報

発明の概要

発明が解決しようとする課題

[0004] 本発明は、対象者がコミュニケーションに対してどの程度注力しているかを判定することができる注力判定システム等を提供する。

課題を解決するための手段

[0005] 本発明の一態様に係る注力判定システムは、ディスプレイ及びホワイトボードが設置された空間に滞在する複数の人が映る映像の映像情報を取得し、取得した前記映像情報に基づいて前記空間における前記複数の人の位置である第一位置を検知する人検知ユニットと、前記映像情報を取得し、取得した前記映像情報に基づいて前記複数の人のうちの対象者の頭部の向きを検知す

る頭部向き検知ユニットと、検知された前記第一位置、及び、検知された前記対象者の頭部の向きに基づいて、前記空間において前記複数の人によって行われるコミュニケーションに対する前記対象者の注力を判定する注力判定ユニットとを備え、前記注力判定ユニットは、前記複数の人のうち前記対象者以外の人が発話しているときに、検知された前記対象者の頭部の向きが、前記複数の人のうち前記対象者以外の人、前記ディスプレイ、及び、前記ホワイトボードのいずれかを向いている期間を、前記対象者が前記コミュニケーションに対して注力している期間であると判定する。

[0006] 本発明の一態様に係るコミュニケーション解析システムは、前記注力判定システムと、判定された前記対象者の注力に基づいて、前記空間において前記複数の人によって行われるコミュニケーションの品質を解析するコミュニケーション解析ユニットとを備える。

[0007] 本発明の一態様に係る注力判定方法は、ディスプレイ及びホワイトボードが設置された空間に滞在する複数の人が映る映像の映像情報を取得し、取得した前記映像情報に基づいて前記空間における前記複数の人の位置である第一位置を検知する第一検知ステップと、前記映像情報を取得し、取得した前記映像情報に基づいて前記複数の人のうちの対象者の頭部の向きを検知する第二検知ステップと、検知された前記第一位置、及び、検知された前記対象者の頭部の向きに基づいて、前記空間において前記複数の人によって行われるコミュニケーションに対する前記対象者の注力を判定する注力判定ステップとを含み、前記注力判定ステップにおいては、前記複数の人のうち前記対象者以外の人が発話しているときに、検知された前記対象者の頭部の向きが、前記複数の人のうち前記対象者以外の人、前記ディスプレイ、及び、前記ホワイトボードのいずれかを向いている期間を、前記人が前記コミュニケーションに対して注力している期間であると判定する。

[0008] 本発明の一態様に係るプログラムは、前記注力判定方法をコンピュータに実行させるためのプログラムである。

発明の効果

[0009] 本発明の注力判定システム等は、対象者がコミュニケーションに対してどの程度注力しているかを判定することができる。

図面の簡単な説明

[0010] [図1]図1は、実施の形態に係るコミュニケーション解析システムの機能構成を示すブロック図である。

[図2]図2は、実施の形態に係るコミュニケーション解析システムが適用される空間の一例を示す図である。

[図3]図3は、実施の形態に係るセンシング装置の外観図である。

[図4]図4は、実施の形態に係るコミュニケーション解析システムが備える音源検知ユニットの動作のフローチャートである。

[図5]図5は、音源の位置の検知結果を模式的に示す図である。

[図6]図6は、実施の形態に係るコミュニケーション解析システムが備える人検知ユニットの動作のフローチャートである。

[図7]図7は、人の位置の検知結果を模式的に示す図である。

[図8]図8は、実施の形態に係るコミュニケーション解析システムが備える発話量推定ユニットの動作のフローチャートである。

[図9]図9は、発話者の位置の検知結果を模式的に示す図である。

[図10]図10は、発話量を示す情報の一例を示す図である。

[図11]図11は、実施の形態に係るコミュニケーション解析システムが備える頭部向き検知ユニットの動作のフローチャートである。

[図12]図12は、実施の形態に係るコミュニケーション解析システムが備える頭部向き検知ユニットの動作を説明するための三次元座標空間を示す図である。

[図13]図13は、頭部の向きの角度の補正を説明するための図である。

[図14]図14は、実施の形態に係るコミュニケーション解析システムが備える注力判定ユニットの動作のフローチャートである。

[図15]図15は、対象者の目的方向を説明するための図である。

[図16]図16は、注力期間を示す情報の一例を示す図である。

[図17]図17は、実施の形態に係るコミュニケーション解析システムが備えるコミュニケーション解析ユニットの動作のフローチャートである。

[図18]図18は、情報端末に表示される、コミュニケーションの品質を示すスコアの表示画面の一例を示す図である。

発明を実施するための形態

[0011] 以下、実施の形態について、図面を参照しながら具体的に説明する。なお、以下で説明する実施の形態は、いずれも包括的または具体的な例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序などは、一例であり、本発明を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、独立請求項に記載されていない構成要素については、任意の構成要素として説明される。

[0012] なお、各図は模式図であり、必ずしも厳密に図示されたものではない。また、各図において、実質的に同一の構成に対しては同一の符号を付し、重複する説明は省略または簡略化される場合がある。

[0013] (実施の形態)

[構成]

まず、実施の形態に係るコミュニケーション解析システムの構成について説明する。図1は、実施の形態に係るコミュニケーション解析システムの機能構成を示すブロック図である。図2は、コミュニケーション解析システムが適用される空間の一例を示す図である。

[0014] 図1及び図2に示されるように、コミュニケーション解析システム10は、会議室などの空間70を有するオフィスなどで使用される、空間70内に位置する複数の人のコミュニケーションの品質を解析するためのシステムである。空間70は、例えば、閉空間であるが、開放空間であってもよい。空間70としては、会議室のほかに、オフィス空間内の開放的休憩所（オフィス空間の一部に椅子及びテーブルが置かれている場所）が例示される。また、空間70は、物理的に区切られている必要は無く、全体空間のうち照明光

または気流などで区切られた場所であってもよい。例えば、色温度5000Kの昼光色で照明されたオフィス空間の一角に、色温度3000Kの暖色の領域が設けられ、この領域が空間70とされてもよい。

[0015] コミュニケーション解析システム10は、センシング装置20と、情報処理システム30とを備える。まず、センシング装置20について図1及び図2に加えて図3を参照しながら説明する。図3は、センシング装置20の外観図である。図3の(a)は、センシング装置20の上面図であり、図3の(b)は、センシング装置20の側面図である。なお、図3では測距センサ23は図示されていない。

[0016] センシング装置20は、空間70に設置された机40の上に設置され、空間70における音及び映像などをセンシングする。センシング装置20は、具体的には、机40の上の中央部に設置される。センシング装置20は、マイクロフォンアレイ21と、複数のカメラ22と、測距センサ23とを備える。

[0017] マイクロフォンアレイ21は、空間70における音を取得し、取得した音の音情報(複数の音信号)を出力する。マイクロフォンアレイ21は、具体的には、複数のマイクロフォン素子を含み、複数のマイクロフォン素子のそれぞれは、空間70における音を取得し、取得した音の音信号を出力する。

[0018] 複数のカメラ22のそれぞれは、空間70に滞在する人が映る映像(言い換えれば、動画像)を撮影し、当該映像の映像情報を出力する。カメラ22は、CMOSイメージセンサなどによって実現される一般的なカメラであるが、魚眼カメラなどであってもよい。センシング装置20は、机40の上からセンシング装置20の周囲の全体を撮影できるように4つのカメラを備えているが、空間70に滞在する人の全員を撮影することができる少なくとも1つのカメラを備えていればよい。

[0019] 測距センサ23は、センシング装置20(カメラ22)から対象物までの距離を計測し、計測した対象物までの距離を示す距離情報を出力する。対象物は、空間70に滞在する人などである。測距センサ23は、例えば、TO

F (Time Of Flight) 方式のLiDAR (Light Detection and Ranging) であるが、距離画像センサなどであってもよい。センシング装置20は、少なくとも1つの測距センサ23を備えていればよいが、カメラ22と対応して複数の測距センサ23を備えてもよい。

[0020] 次に、情報処理システム30について説明する。情報処理システム30は、センシング装置20と有線または無線の通信を行い、当該通信によってセンシング装置20から取得したセンシング情報（具体的には、音情報、映像情報、及び、距離情報など）に基づいて、コミュニケーションの品質を解析する。情報処理システム30は、例えば、空間70を有する施設に設置されるエッジコンピュータであるが、当該施設の外に設置されるクラウドコンピュータであってもよい。情報処理システム30がエッジコンピュータである場合、センシング装置20及び情報処理システム30は、一体的な1つの装置として実現されてもよい。また、情報処理システム30の一部の機能がエッジコンピュータとして実現され、他の一部の機能がクラウドコンピュータによって実現されてもよい。

[0021] 情報処理システム30は、具体的には、音源検知ユニット31、人検知ユニット32、頭部向き検知ユニット33、発話量推定ユニット34、注力判定ユニット35、コミュニケーション解析ユニット36、及び、記憶ユニット37を備える。

[0022] 音源検知ユニット31は、空間70において取得された音の音情報をセンシング装置20から取得し、取得した音情報に基づいて空間70における音源の位置である第二位置を検知する。

[0023] 人検知ユニット32は、空間70に滞在する人が映る映像の映像情報をセンシング装置20から取得し、取得した映像情報に基づいて空間70における人の位置である第一位置を検知する。

[0024] 頭部向き検知ユニット33は、空間70に滞在する人が映る映像の映像情報をセンシング装置20から取得し、取得した映像情報に基づいて人の頭部

の向き（言い換えれば、顔の向き）を検知する。頭部向き検知ユニット３３は、取得した映像情報に基づいて、人の目線の向きを検知してもよい。

[0025] 発話量推定ユニット３４は、人検知ユニット３２によって検知された第一位置、及び、音源検知ユニット３１によって検知された第二位置に基づいて、人の発話量を推定する。

[0026] 注力判定ユニット３５は、人検知ユニット３２によって検知された第一位置、及び、頭部向き検知ユニット３３によって検知された人の頭部の向きに基づいて、空間７０において当該人を含む複数の人によって行われるコミュニケーションに対する人の注力を判定する。

[0027] コミュニケーション解析ユニット３６は、発話量推定ユニット３４によって推定された人の発話量、及び、注力判定ユニット３５によって判定された人の注力の少なくとも一方に基づいて、コミュニケーションの品質を解析する。また、コミュニケーション解析ユニット３６は、解析結果を出力する。

[0028] 以上説明した、音源検知ユニット３１、人検知ユニット３２、頭部向き検知ユニット３３、発話量推定ユニット３４、注力判定ユニット３５、及び、コミュニケーション解析ユニット３６のそれぞれは、マイクロコンピュータまたはプロセッサによって実現される。音源検知ユニット３１、人検知ユニット３２、頭部向き検知ユニット３３、発話量推定ユニット３４、注力判定ユニット３５、及び、コミュニケーション解析ユニット３６の機能は、例えば、上記マイクロコンピュータまたはプロセッサが記憶ユニット３７に記憶されたコンピュータプログラムを実行することによって実現される。

[0029] 記憶ユニット３７は、上記コンピュータプログラム、及び、上記各構成要素の機能を実現するために必要な情報などが記憶される記憶装置である。記憶ユニット３７は、例えば、HDD（Hard Disk Drive）によって実現されるが、半導体メモリによって実現されてもよい。

[0030] なお、音源検知ユニット３１、人検知ユニット３２、及び、発話量推定ユニット３４を含むシステムは、話者ダイアライゼーションシステム３８とも記載される。つまり、話者ダイアライゼーションシステム３８は、音源検知

ユニット 31、人検知ユニット 32、及び、発話量推定ユニット 34 を備える。話者ダイアライゼーションシステム 38 は、さらに、頭部向き検知ユニット 33 または注力判定ユニット 35 を備えてもよい。

[0031] また、人検知ユニット 32、頭部向き検知ユニット 33、及び、注力判定ユニット 35 を含むシステムは、注力判定システム 39 とともに記載される。つまり、注力判定システム 39 は、人検知ユニット 32、頭部向き検知ユニット 33、及び、注力判定ユニット 35 を備える。注力判定システム 39 は、さらに、音源検知ユニット 31 または発話量推定ユニット 34 を備えてもよい。

[0032] [音源検知ユニットの動作]

次に、音源検知ユニット 31 の動作についてより具体的に説明する。図 4 は、音源検知ユニット 31 の動作のフローチャートである。

[0033] まず、音源検知ユニット 31 は、空間 70 において取得された音の音情報をセンシング装置 20 のマイクロフォンアレイ 21 から取得する (S11)。音源検知ユニット 31 は、具体的には、マイクロフォンアレイ 21 に含まれる複数のマイクロフォン素子が出力する複数の音信号を取得する。

[0034] ここで、取得された複数の音信号のそれぞれは、時間領域の信号である。音源検知ユニット 31 は、複数の音信号のそれぞれをフーリエ変換することにより、時間領域の信号から周波数領域の信号に変換する (S12)。

[0035] 次に、音源検知ユニット 31 は、周波数領域に変換された後の複数の音信号に基づいて定まる入力ベクトルから、空間相関行列を算出する (S13)。ここで、マイクロフォンアレイ 21 がマイクロフォン素子を M 個備え、このうち m 番目のフーリエ変換後の音信号を $X_m(\omega, t)$ とすると、入力ベクトル $x(\omega, t)$ は以下の式で表される。T は転置を意味する。

[0036] [数1]

$$x(\omega, t) = [X_1(\omega, t) \cdots, X_M(\omega, t)]^T$$

[0037] また、空間相関行列 R は、以下の式で表される。H は、共役転置を意味す

る。なお、以降は説明の簡略化のため周波数のインデックス ω は省略される。

[0038] [数2]

$$R = E[x(t)x^H(t)]$$

[0039] 次に、音源検知ユニット31は、上記の空間相関行列を固有値分解することにより、固有ベクトルを算出する（S14）。音源検知ユニット31は、具体的には、以下の式に基づいて上記の空間相関行列を固有値分解することにより、固有値ベクトル $e_1, \dots, e_M$ と、固有値 $\lambda_1, \dots, \lambda_M$ とを算出することができる。

[0040] [数3]

$$Re_n = \lambda_n e_n$$

[0041] 次に、音源検知ユニット31は、固有ベクトルから音源の位置を検知する（S15）。音源検知ユニット31は、具体的には、固有ベクトルによって音の大きさとその音が到来する方向を特定することができ、比較的大きい音が到来する方向を音源の方向（位置）として検知することができる。

[0042] この結果、図5に示されるように、音源検知ユニット31は、センシング装置20の位置Oを基準にどの方向（角度）に音源が位置するかを検知することができる。図5は、音源の位置の検知結果を模式的に示す図（机40を上方から見た図）である。図5の例では、音源S1及び音源S2の2つが検知されている。なお、音源検知ユニット31は、少なくとも図5に示されるような二次元的な音源の位置（上面視において角度で表される方向）を検知すればよいが、三次元的な音源の位置を検知してもよい。以下の実施の形態では、音源検知ユニット31によって検知された音源の位置は、第二位置とも記載される。

[0043] 音源検知ユニット31は、図4の動作を単位時間ごとに繰り返すことにより、音源の位置（第二位置）を追跡することができる。なお、音源とは、具体的には、空間70内に滞在する発話者（人）であるが、空間70内に設置

された機器である可能性もある。

[0044] [人検知ユニットの動作]

次に、人検知ユニット32の動作についてより具体的に説明する。図6は、人検知ユニット32の動作のフローチャートである。なお、以下の動作は、実際には4つのカメラ22のそれぞれから取得された映像情報について行われるが、以下では、便宜上、1つのカメラ22から映像情報が取得されるものとして説明が行われる。

[0045] まず、人検知ユニット32は、空間70において取得された映像の映像情報をセンシング装置20のカメラ22から取得する(S21)。

[0046] 次に、人検知ユニット32は、取得された映像情報に基づいて、映像内で人が映っている領域を特定する(S22)。人検知ユニット32は、パターンマッチングを用いた手法、または、機械学習モデルを用いた手法などにより、映像内で人が映っている領域を特定することができる。

[0047] 次に、人検知ユニット32は、特定した領域(つまり、人がいる領域)に識別情報を割り当てる(S23)。例えば、人検知ユニット32は、3つの領域を特定し、特定した3つの領域にA、B、Cの識別情報を割り当てる。以下では、領域Aに対応する人を人A、領域Bに対応する人を人B、領域Cに対応する人を人Cとも記載する。

[0048] 次に、人検知ユニット32は、ステップS23で特定した領域の映像内での位置に基づいて、人がいる方向を特定する(S24)。記憶ユニット37には、カメラ22の設置位置(机40の上の中央)と、カメラ22の撮影範囲(画角)とを示す情報があらかじめ記憶されており、人検知ユニット32は、映像内の位置がどの方向に相当するかを特定することができる。

[0049] 次に、人検知ユニット32は、ステップS23で特定した領域の大きさに基づいて、センシング装置20(カメラ22)から人までの距離を推定する(S25)。この場合、ステップS23で特定した領域が大きいほど、センシング装置20(カメラ22)から人までの距離は近いと推定される。なお、ステップS25では、測距センサ23から取得される距離情報(距離の実

測値)によってセンシング装置20(カメラ22)から人までの距離が特定されてもよい。

[0050] 以上のステップS24及びステップS25の結果、図7に示されるように、人検知ユニット32は、空間70における人の位置を検知することができる。図7は、人の位置の検知結果を模式的に示す図(机40を上方から見た図)である。図7の例では、人A、人B、及び、人Cの三人の位置が検知されている。なお、人検知ユニット32は、三次元的な人の位置(人の位置の三次元座標)を検知するが、少なくとも図7に示されるような二次元的な音源の位置(上面視において角度で表される方向)を検知すればよい。以下の実施の形態では、人検知ユニット32によって検知された人の位置は、第一位置とも記載される。

[0051] 人検知ユニット32は、図6の動作を単位時間ごとに繰り返すことにより、人の位置(第一位置)を追跡することができる。このときステップS23の識別情報の割り当ては、最初の1回のみ行われればよい。

[0052] [発話量推定ユニットの動作]

次に、発話量推定ユニット34の動作についてより具体的に説明する。図8は、発話量推定ユニット34の動作のフローチャートである。

[0053] 発話量推定ユニット34は、人検知ユニット32によって検知された第一位置、及び、音源検知ユニット31によって検知された第二位置を取得する(S31)。このとき取得される第一位置及び第二位置は、実質的に同一のタイミングで検知されたものである。実質的に同一とは、多少のずれを含んでもよいことを意味する。

[0054] 次に、発話量推定ユニット34は、三次元座標で表される第一位置を、第二位置相当の二次元座標(角度)に変換する(S32)。なお、マイクロフォンアレイ21の位置とカメラ22の位置とが異なる場合には、これらの位置の差分に基づいて、二次元座標に変換された後の第一位置が補正される。

[0055] 次に、発話量推定ユニット34は、変換後の第一位置と、第二位置とを照合することにより、発話者(発話者の位置)を検知する(S33)。図9は

、発話者の位置の検知結果を模式的に示す図である。図 9 は、第二位置（図 5）と、第一位置（図 7）とを重畳した図である。

[0056] 発話量推定ユニット 34 は、例えば、音源 S 1 の第二位置と人 A の第一位置との角度差 $\Delta \theta 1$ が所定値以下であるときに、音源 S 1 が人 A であると検知する。つまり、人 A は発話者として検知される。また、発話量推定ユニット 34 は、例えば、音源 S 2 の第二位置と人 C の第一位置との角度差 $\Delta \theta 2$ が所定値以下であるときに、音源 S 2 が人 C であると検知する。つまり、人 C は発話者として検知される。

[0057] 発話量推定ユニット 34 は、図 8 の動作を単位時間ごとに繰り返すことにより、人 A、人 B、及び、人 C のそれぞれを追跡し、人 A、人 B、及び、人 C それぞれの発話量を推定することができる。発話量推定ユニット 34 は、具体的には、人 A、人 B、及び、人 C のそれぞれが発話者として検知された期間を、人 A、人 B、及び、人 C のそれぞれが発話している期間と推定することができる。つまり、発話量推定ユニット 34 は、人 A、人 B、及び、人 C のそれぞれが発話している期間を示す情報（発話量を示す情報）を記憶ユニット 37 に記憶することができる。図 10 は、発話量を示す情報の一例を示す図である。図 10 に示されるように、発話量を示す情報は、上記ステップ S 23 で割り当てられた識別情報のそれぞれに、発話量（発話時間）が紐づけられた情報である。

[0058] このように、発話量推定ユニット 34 は、人検知ユニット 32 によって検知された第一位置、及び、音源検知ユニット 31 によって検知された第二位置に基づいて、複数の人それぞれを追跡し、かつ、空間 70 に滞在する複数の人のそれぞれの発話量（対象者の発話量）を推定することができる。このような発話量推定ユニット 34 による発話量の推定方法は、複数の人の座席の移動を伴う会議等におけるコミュニケーションの品質の解析に有用である。この場合の複数の人の座席の移動とは、例えば、ホワイトボード 60 を使うために移動することなどを意味する。

[0059] また、発話量推定ユニット 34 は、音声認識による個人特定、及び、画像

認識による個人特定を行うのではなく、複数の人の匿名性を維持したまま、複数の人それぞれの発話量を推定することができる。

[0060] 〔頭部向き検知ユニットの動作〕

次に、頭部向き検知ユニット 33 の動作についてより具体的に説明する。図 11 は、頭部向き検知ユニット 33 の動作のフローチャートである。図 12 は、頭部向き検知ユニット 33 の動作を説明するための三次元座標空間を示す図である。なお、図 12 に示されるように、以下の頭部向き検知ユニット 33 の動作の説明においては、カメラ 22 は、机 40 の中心位置 $O(0, 0, 0)$ ではなく、机 40 上の位置 $C(x_0, y_0, z_0)$ に設置されているものとして説明が行われる。なお、図 12 には、 $X-Y-Z$ で示される空間 70 における三次元座標に加えて、 $U-V$ で示される映像中の座標も合わせて図示されている。

[0061] まず、頭部向き検知ユニット 33 は、空間 70 において取得された映像の映像情報をセンシング装置 20 のカメラ 22 から取得する (S41)。

[0062] 次に、頭部向き検知ユニット 33 は、取得された映像情報に基づいて、映像に映る人の頭部の向きを特定する (S42)。頭部向き検知ユニット 33 は、例えば、顔認識処理により、映像における人の頭部に相当する領域を特定し、特定した領域に機械学習モデルを適用することにより、人の頭部の向きを検知することができる。図 12 に示されるように、このときの頭部の向きを示すベクトルは、当該人の位置 P とカメラの位置 C とを結ぶ線分（直線）に対して A° の角度をなすものとする。

[0063] 次に、頭部向き検知ユニット 33 は、人検知ユニット 32 によって検知される上記人の位置（より詳細には、人の頭部の位置）を取得する (S43)。図 12 に示されるように、このときの人の位置は、 $P(x_1, y_1, z_1)$ とされる。

[0064] 例えば、机 40 の実際のサイズによって定まる $z_1 - z_0$ と、カメラ 22 の水平方向の視野角 α と、映像中の映像内の人の位置 (u, v) と、映像の横幅 w とを用いると、 x_1 は、以下の式で表現できる。なお、机 40 のサイ

ズ、カメラ22の水平方向の視野角 α 、及び、映像の横幅 w などの情報は、あらかじめ記憶ユニット37に記憶される。

[0065] $x_1 = x_0 + ((u - w/2) / w) \times (z_1 - z_0) \times \tan(\alpha)$

[0066] 次に、頭部向き検知ユニット33は、ステップS42において特定されたカメラ22の位置を基準とした頭部の向き（角度A）を、ステップS43において取得した人の位置（具体的には、座標 x_1 の値）に基づいて補正する（S44）。図13は、このような角度の補正を説明するための図（机40を上方から見た図）である。

[0067] 頭部向き検知ユニット33は、ステップS43において取得された x_1 の座標に基づいて、図13における $\angle OPC$ を算出することができ、 $A + \angle OPC$ を補正後の角度とすることができる。

[0068] 頭部向き検知ユニット33は、図11の動作を単位時間ごとに繰り返すことにより、人の頭部の向きを追跡することができる。

[0069] [注力判定ユニットの動作]

次に、注力判定ユニット35の動作についてより具体的に説明する。図14は、注力判定ユニット35の動作のフローチャートである。

[0070] 注力判定ユニット35は、人検知ユニット32によって検知された第一位置、及び、頭部向き検知ユニット33によって検知された人の頭部の向きを取得する（S51）。このとき取得される第一位置は、より詳細には、空間70に滞在する複数の人それぞれの第一位置であり、人の頭部の向きは、当該複数の人それぞれの頭部の向きである。

[0071] 次に、注力判定ユニット35は、複数の人の1人を対象者として、対象者の目的方向を決定する（S52）。図15は、対象者の目的方向を説明するための図（机40を上方から見た図）である。

[0072] 注力判定ユニット35は、例えば、位置Pにいる人が対象者である場合、位置Pから他の人がいる位置Gを結ぶ方向を目的方向に決定する。なお、目的方向は、必ずしも人に対して決定される必要はなく、後述のようにディスプレイ50及びホワイトボード60に対して決定されてもよい。

- [0073] 次に、注力判定ユニット35は、目的方向を中心として許容範囲を決定する(S53)。例えば、図15の例では、 $\angle GPO - \beta$ から $\angle GPO + \beta$ までの範囲が許容範囲として決定される。 β は許容範囲を定めるための係数である。なお、 β の値は、目的方向が向かう先が、人であるか、ディスプレイ50であるか、ホワイトボード60であるかによって異なる値とされる。
- [0074] 次に、注力判定ユニット35は、ステップS51で取得した対象者の頭部の向きと、ステップS53で決定した許容範囲とを比較することにより、対象者の注力を判定する(S54)。ここでの注力は、空間70において対象者を含む複数の人によって行われるコミュニケーションに対する注力を意味する。注力判定ユニット35は、対象者の頭部の向きが許容範囲内であるときには、対象者が他の人を見ていると考えられることから、コミュニケーションに注力していると判定する。一方、注力判定ユニット35は、対象者の頭部の向きが許容範囲外であるときには、対象者が他の人を見ていないと考えられることからコミュニケーションに注力していないと判定する。
- [0075] 注力判定ユニット35は、図14の動作を単位時間ごとに複数の人(人A、人B、及び、人Cとする)それぞれを対象者として繰り返すことにより、人A、人B、及び、人Cそれぞれの注力を判定することができる。注力判定ユニット35は、具体的には、人A、人B、及び、人Cのそれぞれが注力していると判定された期間を積算し、複数の人がコミュニケーションに注力している期間を示す情報を記憶ユニット37に記憶することができる。図16は、注力期間を示す情報の一例を示す図である。図16に示されるように、発話量を示す情報は、上記ステップS23で割り当てられた識別情報のそれぞれに、注力期間が紐づけられた情報である。
- [0076] このように、注力判定ユニット35は、人検知ユニット32によって検知された第一位置、及び、頭部向き検知ユニット33によって検知された頭部の向きに基づいて、空間70において複数の人によって行われるコミュニケーションに対する対象者の注力を判定することができる。
- [0077] なお、注力判定ユニット35は、音源検知ユニット31によって検知され

た第二位置をさらに取得して発話者の位置を検知するか、あるいは、発話量推定ユニット34によって検知される発話者の位置を取得してもよい。いずれの場合も、注力判定ユニット35は、複数の人のそれぞれが発話者であるか否かを考慮して注力を判定することができる。例えば、対象者の頭部が発話者の方向を向いているときのみ対象者がコミュニケーションに注力していると判定し、発話していない人の方向を見ているときは対象者がコミュニケーションに注力していないと判定することができる。

[0078] また、複数の人のそれぞれが発話者であるか否かを考慮する場合、注力判定ユニット35は、以下のように注力を判定してもよい。

[0079] 例えば、空間70において複数の人によって会議が行われているような場合、対象者（人Cとする）は、当該対象者の真横に位置している人Aが発話していても、真横に位置している人のほうは向きづらい。したがって、他の人Bのほうを見ると考えられる。また、対象者は、会議の報告資料等が表示されているディスプレイ50（図2に図示）、及び、ホワイトボード60（図2に図示）のほうを向くことも考えられる。このような場合、上記のように対象者の頭部が発話者の方向を向いているときのみ対象者がコミュニケーションに注力していると判定されると、実際にはコミュニケーションに注力している対象者がコミュニケーションに注力していないと判定されてしまうという課題がある。

[0080] そこで、このような場合、人Aが発話している間は、（1）対象者が人Aの方向を向いている期間、（2）対象者がディスプレイ50の方向を向いている期間、（3）対象者が人A以外の人（人B）の方向を向いている期間、（4）対象者がホワイトボード60の方向を向いている期間のいずれも、対象者がコミュニケーションに注力していると判定されてもよい。つまり、注力判定ユニット35は、複数の人のうち対象者以外の人が発話しているときに、検知された対象者の頭部の向きが、複数の人のうち対象者以外の人、ディスプレイ50、及び、ホワイトボード60のいずれかを向いている期間を、人がコミュニケーションに対して注力している期間であると判定してもよ

い。なお、対象者がホワイトボード60のほうを向いている期間は、人Aがホワイトボード付近で発話しているという条件の下で、対象者がコミュニケーションに注力している期間と判定されてもよい。

[0081] [コミュニケーション解析ユニットの動作]

次に、コミュニケーション解析ユニット36の動作についてより具体的に説明する。図17は、コミュニケーション解析ユニット36の動作のフローチャートである。

[0082] コミュニケーション解析ユニット36は、記憶ユニット37に記憶された、発話量を示す情報(図10)、及び、注力期間を示す情報(図16)を記憶ユニット37から読み出す(S61)。

[0083] 次に、コミュニケーション解析ユニット36は、取得した発話量を示す情報、及び、取得した注力期間を示す情報に基づいて、空間70におけるコミュニケーションの品質を解析する。例えば、コミュニケーション解析ユニットは、以下の採点基準でコミュニケーションの品質をスコア化する。

[0084] コミュニケーション解析ユニット36は、例えば、取得した発話量を示す情報に基づいて、人A～人Cの発話量の比率を算出し、この比率の最小値／最大値が1に近い(つまり、人A～人Cがまんべんなく発言している)ほど、発話量に関する第一スコアを大きい値にする。コミュニケーション解析ユニット36は、具体的には、人A～人Cの発話量の比率が1:1、2:1、5である場合、1(最小値)／1.5(最大値)と1との差分に基づいて第一スコアを算出する。

[0085] また、コミュニケーション解析ユニット36は、例えば、取得した注力期間を示す情報に基づいて、人A～人Cの注力期間の平均値が大きいほど、注力に関する第二スコアを大きい値にする。

[0086] そして、コミュニケーション解析ユニット36は、第一スコア及び第二スコアの合計を、空間70におけるコミュニケーションの品質を示す最終スコアとして算出する。人A～人Cのそれぞれは、例えば、スマートフォンまたはパーソナルコンピュータなどの情報端末を用いてコミュニケーション解析

システム１０（情報処理システム３０）にアクセスすることにより、コミュニケーションの品質を示す最終スコア（つまり、コミュニケーションの品質の解析結果）を確認することができる。図１８は、情報端末に表示される、コミュニケーションの品質を示すスコアの表示画面の一例を示す図である。

[0087] このように、コミュニケーション解析ユニット３６は、発話量の推定結果、及び、注力の判定結果に基づいて、空間７０において複数の人によって行われるコミュニケーションの品質を解析することができる。なお、コミュニケーション解析ユニット３６は、発話量の推定結果、及び、注力の判定結果の少なくとも一方に基づいて、空間７０において複数の人によって行われるコミュニケーションの品質を解析すればよい。また、上記のようなスコアの算出基準は一例であり、スコアは空間７０におけるコミュニケーションに求められる内容に応じて適宜定められればよい。

[0088] 〔変形例〕

上記実施の形態では、マイクロフォンアレイ２１、カメラ２２、及び、測距センサ２３は机４０の上に設置された。しかしながら、マイクロフォンアレイ２１、カメラ２２、及び、測距センサ２３は、天井に設置されてもよい。また、マイクロフォンアレイ２１、カメラ２２、及び、測距センサ２３が１か所に集約される必要はなく、マイクロフォンアレイ２１、カメラ２２、及び、測距センサ２３は互いに異なる場所に設置されていてもよい。

[0089] また、上記実施の形態では、発話者を検知するために音源検知ユニット３１と人検知ユニット３２とが併用されたが、人検知ユニット３２のみで発話者を検知することもできる。例えば、人検知ユニット３２は、映像に映る人の動きなどから当該人が発話中であるか否かを検知することができる。

[0090] また、上記実施の形態において、コミュニケーション解析ユニット３６は、コミュニケーションの品質を示す最終スコア（つまり、コミュニケーションの品質の解析結果）を空間７０において会議等が行われているときにリアルタイムに算出し、算出した最終スコアに基づいて空間７０における環境をリアルタイムに制御してもよい。

[0091] 例えば、コミュニケーション解析ユニット36は、算出された最終スコアが所定値未満である（つまり、コミュニケーションが活発でない）と判定すると、空間70に設置された環境制御機器（図示せず）へ制御信号を送信することにより、空間70の環境を制御する。

[0092] 環境制御機器は、具体的には、空間70の温度環境を制御する空気調和機、空間70の光環境を制御する照明機器、空間70のにおいを制御する香り発生機、及び、空間70の音環境を制御する楽曲再生装置などである。

[0093] 例えば、コミュニケーション解析ユニット36は、空気調和機の設定温度を高くする、または、照明機器を現在よりも明るくすることにより、空間70におけるコミュニケーションの活性化を図ることができる。また、コミュニケーション解析ユニット36は、空間70に設置された香り発生機を作動させる、または、空間70に設置された楽曲再生装置に楽曲を再生させることにより、空間70におけるコミュニケーションの活性化を図ってもよい。

[0094] [効果等]

以上説明したように、注力判定システム39は、ディスプレイ50及びホワイトボード60が設置された空間70に滞在する複数の人が映る映像の映像情報を取得し、取得した映像情報に基づいて空間70における複数の人の位置である第一位置を検知する人検知ユニット32と、映像情報を取得し、取得した映像情報に基づいて複数の人のうちの対象者の頭部の向きを検知する頭部向き検知ユニット33と、検知された第一位置、及び、検知された対象者の頭部の向きに基づいて、空間70において複数の人によって行われるコミュニケーションに対する対象者の注力を判定する注力判定ユニット35とを備える。注力判定ユニット35は、複数の人のうち対象者以外の人が発話しているときに、検知された対象者の頭部の向きが、複数の人のうち対象者以外の人、ディスプレイ50、及び、ホワイトボード60のいずれかを向いている期間を、対象者がコミュニケーションに対して注力している期間であると判定する。

[0095] このような注力判定システム39は、対象者がコミュニケーションに対し

てどの程度注力しているかを判定することができる。

- [0096] また、例えば、空間 70 には机 40 が設置され、人検知ユニット 32 は、机 40 の上に設置されたカメラ 22 から映像情報を取得する。
- [0097] このような注力判定システム 39 は、机 40 の上に設置されたカメラ 22 から映像情報を取得することができる。
- [0098] また、例えば、空間 70 には、複数のカメラ 22 が設置され、人検知ユニット 32 によって取得される映像情報には、複数のカメラ 22 のそれぞれによって撮影された映像の映像情報が含まれる。
- [0099] このような注力判定システム 39 は、机 40 の上に設置された複数のカメラ 22 から映像情報を取得することができる。
- [0100] また、例えば、空間 70 には、カメラ 22 から対象者までの距離を計測する測距センサ 23 が設置される。人検知ユニット 32 は、取得した映像情報、及び、測距センサ 23 の検知結果に基づいて第一位置を検知する。
- [0101] このような注力判定システム 39 は、測距センサ 23 の検知結果に基づいて第一位置を検知することができる。
- [0102] また、例えば、人検知ユニット 32 は、第一位置を検知するために、映像における対象者の大きさに基づいて人検知ユニット 32 から対象者までの距離を推定する。
- [0103] このような注力判定システム 39 は、映像における対象者の大きさに基づいて人検知ユニット 32 から対象者までの距離を推定することができる。
- [0104] また、コミュニケーション解析システム 10 は、注力判定システム 39 と、判定された前記対象者の注力に基づいて、空間 70 において複数の人によって行われるコミュニケーションの品質を解析するコミュニケーション解析ユニット 36 とを備える。
- [0105] このようなコミュニケーション解析システム 10 は、対象者の注力に基づいてコミュニケーションの品質を解析することができる。
- [0106] また、例えば、コミュニケーション解析システム 10 は、さらに、空間 70 において取得された音の音情報を取得し、取得した音情報に基づいて空間

70における音源の位置である第二位置を検知する音源検知ユニット31と、検知された第一位置、及び、検知された第二位置に基づいて、対象者を追跡し、かつ、追跡中の対象者の発話量を推定する発話量推定ユニット34とを備える。コミュニケーション解析ユニット36は、判定された対象者の注力、及び、推定された対象者の発話量に基づいて、コミュニケーションの品質を解析する。

[0107] このようなコミュニケーション解析システム10は、対象者の注力、及び、対象者の発話量に基づいてコミュニケーションの品質を解析することができる。

[0108] また、注力判定ユニット35などのコンピュータによって実行される注力判定方法は、ディスプレイ50及びホワイトボード60が設置された空間70に滞在する複数の人が映る映像の映像情報を取得し、取得した映像情報に基づいて空間70における複数の人の位置である第一位置を検知する第一検知ステップと、映像情報を取得し、取得した映像情報に基づいて複数の人のうち対象者の頭部の向きを検知する第二検知ステップと、検知された第一位置、及び、検知された対象者の頭部の向きに基づいて、空間70において複数の人によって行われるコミュニケーションに対する対象者の注力を判定する注力判定ステップとを含む。注力判定ステップにおいては、複数の人のうち対象者以外の人が発話しているときに、検知された対象者の頭部の向きが、複数の人のうち対象者以外の人、ディスプレイ50、及び、ホワイトボード60のいずれかを向いている期間を、人がコミュニケーションに対して注力している期間であると判定する。

[0109] このような注力判定方法は、対象者がコミュニケーションに対してどの程度注力しているかを判定することができる。

[0110] (その他の実施の形態)

以上、実施の形態について説明したが、本発明は、上記実施の形態に限定されるものではない。

[0111] 例えば、上記実施の形態では、コミュニケーション解析システムは、複数

の装置によって実現されたが、単一の装置として実現されてもよい。例えば、コミュニケーション解析システムは、情報処理システム、話者ダイアライゼーションシステム、または、注力判定システムに相当する単一の装置として実現されてもよい。コミュニケーション解析システムが複数の装置によって実現される場合、コミュニケーション解析システムが備える機能的な構成要素は、複数の装置にどのように振り分けられてもよい。

[0112] また、上記実施の形態における装置間の通信方法については特に限定されるものではない。また、装置間の通信においては、図示されない中継装置が介在してもよい。

[0113] また、上記実施の形態において、特定の処理部が実行する処理を別の処理部が実行してもよい。また、複数の処理の順序が変更されてもよいし、複数の処理が並行して実行されてもよい。

[0114] また、上記実施の形態において、各構成要素は、各構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、CPUまたはプロセッサなどのプログラム実行部が、ハードディスクまたは半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。

[0115] また、各構成要素は、ハードウェアによって実現されてもよい。例えば、各構成要素は、回路（または集積回路）でもよい。これらの回路は、全体として1つの回路を構成してもよいし、それぞれ別々の回路でもよい。また、これらの回路は、それぞれ、汎用的な回路でもよいし、専用の回路でもよい。

[0116] また、本発明の全般的または具体的な態様は、システム、装置、方法、集積回路、コンピュータプログラムまたはコンピュータ読み取り可能なCD-ROMなどの記録媒体で実現されてもよい。また、システム、装置、方法、集積回路、コンピュータプログラム及び記録媒体の任意な組み合わせで実現されてもよい。

[0117] 例えば、本発明は、話者ダイアライゼーションシステムなどのコンピュー

タが実行する発話量推定方法として実現されてもよいし、このような発話量推定方法をコンピュータに実行させるためのプログラムとして実現されてもよいし、このようなプログラムが記録されたコンピュータ読み取り可能な非一時的な記録媒体として実現されてもよい。

[0118] また、本発明は、注力判定システムなどのコンピュータが実行する注力判定方法として実現されてもよいし、このような注力判定方法をコンピュータに実行させるためのプログラムとして実現されてもよいし、このようなプログラムが記録されたコンピュータ読み取り可能な非一時的な記録媒体として実現されてもよい。

[0119] その他、各実施の形態に対して当業者が思いつく各種変形を施して得られる形態、または、本発明の趣旨を逸脱しない範囲で各実施の形態における構成要素及び機能を任意に組み合わせることによって実現される形態も本発明に含まれる。

符号の説明

- [0120]
- 10 コミュニケーション解析システム
 - 20 センシング装置
 - 21 マイクロフォンアレイ
 - 22 カメラ
 - 23 測距センサ（センサ）
 - 30 情報処理システム
 - 31 音源検知ユニット
 - 32 人検知ユニット
 - 33 頭部向き検知ユニット
 - 34 発話量推定ユニット
 - 35 注力判定ユニット
 - 36 コミュニケーション解析ユニット
 - 37 記憶ユニット
 - 38 話者ダイアライゼーションシステム

3 9 注力判定システム

4 0 机

5 0 ディスプレイ

6 0 ホワイトボード

7 0 空間

請求の範囲

- [請求項1] ディスプレイ及びホワイトボードが設置された空間に滞在する複数の人が映る映像の映像情報を取得し、取得した前記映像情報に基づいて前記空間における前記複数の人の位置である第一位置を検知する人検知ユニットと、
- 前記映像情報を取得し、取得した前記映像情報に基づいて前記複数の人のうちの対象者の頭部の向きを検知する頭部向き検知ユニットと、
- 、
- 検知された前記第一位置、及び、検知された前記対象者の頭部の向きに基づいて、前記空間において前記複数の人によって行われるコミュニケーションに対する前記対象者の注力を判定する注力判定ユニットとを備え、
- 前記注力判定ユニットは、前記複数の人のうち前記対象者以外の人が発話しているときに、検知された前記対象者の頭部の向きが、前記複数の人のうち前記対象者以外の人、前記ディスプレイ、及び、前記ホワイトボードのいずれかを向いている期間を、前記対象者が前記コミュニケーションに対して注力している期間であると判定する注力判定システム。
- [請求項2] 前記空間には机が設置され、
- 前記人検知ユニットは、前記机の上に設置されたカメラから前記映像情報を取得する
- 請求項1に記載の注力判定システム。
- [請求項3] 前記空間には、複数のカメラが設置され、
- 前記人検知ユニットによって取得される前記映像情報には、前記複数のカメラのそれぞれによって撮影された映像の映像情報が含まれる
- 請求項2に記載の注力判定システム。
- [請求項4] 前記空間には、前記カメラから前記対象者までの距離を計測するセンサが設置され、

前記人検知ユニットは、取得した前記映像情報、及び、前記センサの検知結果に基づいて前記第一位置を検知する

請求項 2 に記載の注力判定システム。

[請求項5]

前記人検知ユニットは、前記第一位置を検知するために、前記映像における前記対象者の大きさに基づいて前記人検知ユニットから前記対象者までの距離を推定する

請求項 1 に記載の注力判定システム。

[請求項6]

請求項 1 ～ 5 のいずれか 1 項に記載の注力判定システムと、
判定された前記対象者の注力に基づいて、前記空間において前記複数の人によって行われるコミュニケーションの品質を解析するコミュニケーション解析ユニットとを備える

コミュニケーション解析システム。

[請求項7]

さらに、前記空間において取得された音の音情報を取得し、取得した前記音情報に基づいて前記空間における音源の位置である第二位置を検知する音源検知ユニットと、

検知された前記第一位置、及び、検知された前記第二位置に基づいて、前記対象者を追跡し、かつ、追跡中の前記対象者の発話量を推定する発話量推定ユニットとを備え、

前記コミュニケーション解析ユニットは、判定された前記対象者の注力、及び、推定された前記対象者の発話量に基づいて、前記コミュニケーションの品質を解析する

請求項 6 に記載のコミュニケーション解析システム。

[請求項8]

ディスプレイ及びホワイトボードが設置された空間に滞在する複数の人が映る映像の映像情報を取得し、取得した前記映像情報に基づいて前記空間における前記複数の人の位置である第一位置を検知する第一検知ステップと、

前記映像情報を取得し、取得した前記映像情報に基づいて前記複数の人のうちの対象者の頭部の向きを検知する第二検知ステップと、

検知された前記第一位置、及び、検知された前記対象者の頭部の向きに基づいて、前記空間において前記複数の人によって行われるコミュニケーションに対する前記対象者の注力を判定する注力判定ステップとを含み、

前記注力判定ステップにおいては、前記複数の人のうち前記対象者以外の人が発話しているときに、検知された前記対象者の頭部の向きが、前記複数の人のうち前記対象者以外の人、前記ディスプレイ、及び、前記ホワイトボードのいずれかを向いている期間を、前記人が前記コミュニケーションに対して注力している期間であると判定する注力判定方法。

[請求項9]

請求項8に記載の注力判定方法をコンピュータに実行させるためのプログラム。