(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年7月23日(23.07.2020)

WIPO | PCT

(10) 国際公開番号
WO 2020/148920 A1

(51) 国際特許分類:
G06F 16/00 (2019.01)

(21) 国際出願番号： PCT/JP2019/009523

(22) 国際出願日： 2019年3月8日(08.03.2019)

(25) 国際出願の言語： 日本語

(26) 国際公開の言語： 日本語

(30) 優先権データ：
特願 2019-005363 2019年1月16日(16.01.2019) JP

(71) 出願人: オムロン株式会社 (**OMRON CORPORATION**) [JP/JP]; 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町８０１番地 Kyoto (JP).

(72) 発明者: 江上 慎(**EGAMI, Shin**); 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町８０１番地 オムロン株式会社内 Kyoto (JP). 笠井 一希(**KASAI, Kazuki**); 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町８０１番地 オムロン株式会社内 Kyoto (JP). 和田 純一(**WADA, Junichi**); 〒6008530 京都府京都市下京区塩小路通堀川東入南不動堂町８０１番地 オムロン株式会社内 Kyoto (JP).

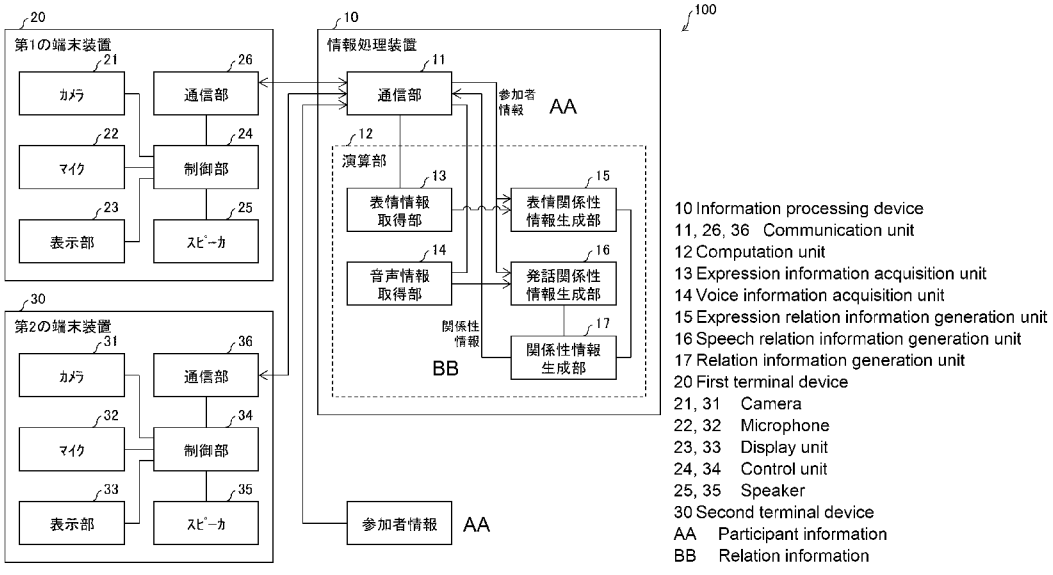
(74) 代理人: 村 上 尚 (**MURAKAMI, Takashi**); 〒5300041 大阪府大阪市北区天神橋２丁目北２番６号 大和南森町ビル４Ｆ Osaka (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ,

(54) **Title:** INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND INFORMATION PROCESSING PROGRAM

(54) 発明の名称： 情報処理装置、情報処理方法、及び情報処理プログラム

[図1]



(57) **Abstract:** An objective of the present invention is to evaluate relations among participants. An information processing device (10) comprises: an expression information acquisition unit (13) which acquires expression information of a plurality of participants; a voice information acquisition unit (14) which acquires speech information of the participants; an expression relation information generation unit (15) which generates expression relation information from the expression information; a speech relation information generation unit (16) which generates speech relation information from

［続葉有］

WO 2020/148920 A1

BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

一 国際調査報告(条約第21条(3))

the speech information; and a relation information generation unit (17) which generates relation information from the expression relation information and the speech relation information.

(57) 要約:参加者間の関係性を評価する。情報処理装置(１０)は、複数の参加者の表情情報を取得する表情情報取得部(１３)と、参加者の発話情報を取得する音声情報取得部(１４)と、表情情報から表情関係性情報を生成する表情関係性情報生成部(１５)と、発話情報から発話関係性情報を生成する発話関係性情報生成部(１６)と、表情関係性情報と発話関係性情報とから関係性情報を生成する関係性情報生成部(１７)とを備えている。

明　細　書

発明の名称 ：

情報処理装置、情報処理方法、及び情報処理プログラム

技術分野

[0001] 本発明は、情報処理装置、情報処理方法、及び情報処理プログラムに関する。

背景技術

[0002] ユーザ間の円滑なコミュニケーションを支援するサービスが従来技術として知られている。特許文献１には、コールセンターのオペレータと顧客との会話における音声に基づいて感情を認識し、音声と感情とを組み合わせて音声分析することで、オペレータのパフォーマンスを評価する電話音声モニタリング評価システムが記載されている。特許文献２には、ユーザが入力したチャット文が示す感情を認識し、ユーザ間の感情の類似度を求める感情マッチング装置が記載されている。

先行技術文献

特許文献

[0003] 特許文献1：日本国公開特許公報「特開２０１７－１３５６４２号公報（２０１７年８月３日公開）」

特許文献2：日本国公開特許公報「特開２００５－２８４８２２号公報（２００５年１０月１３日公開）」

発明の概要

発明が解決しようとする課題

[0004] しかしながら、上述のような従来技術は、会話における音声のみに基づいて、又は、入力されたチャット文のみに基づいて、話者の感情を認識するため、多面的な感情認識ができないという問題がある。

[0005] 本発明の一態様は、会話中の各ユーザの感情を多面的に認識し、認識された感情に基づく会話の評価を通知するコミュニケーション支援技術を提供す

ることを目的とする。

課題を解決するための手段

[0006] 前記の課題を解決するために、本発明の一態様に係る情報処理装置は、複数の参加者のうち第１の参加者の表情に関する第１の表情情報と、前記複数の参加者のうち第２の参加者の表情に関する第２の表情情報とを取得する表情情報取得部と、当該第１の参加者の発話に関する第１の発話情報と、前記複数の参加者のうち第２の参加者の発話に関する第２の発話情報とを取得する音声情報取得部と、前記第１の表情情報と前記第２の表情情報とを参照して、前記第１の参加者と前記第２の参加者との表情に関する関係性を示す表情関係性情報を生成する表情関係性情報生成部と、前記第１の発話情報と前記第２の発話情報とを参照して、前記第１の参加者と前記第２の参加者との発話に関する関係性を示す発話関係性情報を生成する発話関係性情報生成部と、前記表情関係性情報と前記発話関係性情報とを参照して前記第１の参加者と前記第２の参加者との関係を示す情報である関係性情報を生成する関係性情報生成部と、を備えていることを特徴としている。

発明の効果

[0007] 本発明の一態様によれば、各参加者の音声情報及び表情情報に基づいて、会議中の参加者間の関係性を評価することができる。

図面の簡単な説明

[0008] [図1]本発明の一実施形態に係る情報処理装置を含む情報処理システムの構成要素の一例を示すブロック図である。

[図2]本発明の一実施形態に係る情報処理装置を含む情報処理システムの概要を示す図である。

[図3]本発明の一実施形態に係る情報処理装置を含む情報処理システムにおけるデータの流れの概要を示す図である。

[図4]本発明の一実施形態に係る情報処理装置を含む情報処理システムが提示する情報の一例を示す図である。

[図5]本発明の一実施形態に係る情報処理装置を含む情報処理システムが提示

する情報の他の例を示す図である。

発明を実施するための形態

[0009] ［実施形態１］

以下、本発明の一実施形態について、詳細に説明する。図１は、本実施形態の情報処理装置１０を含む情報処理システム１００の概要を示す図である。図１に示すように、情報処理システム１００は、情報処理装置１０、第１の端末装置２０、及び第２の端末装置３０を備えている。なお端末装置の数は、本実施形態を限定するものではなく、３以上であってもよい。

[0010] 図２は、本発明の一実施形態に係る情報処理装置１０を含む情報処理システム１００の概要を示す図である。情報処理システム１００においては、図２に示すように、第１の端末装置２０を使用する第１の参加者２００と、第２の端末装置３０を使用する第２の参加者２０１との間の関係性を評価する。

[0011] 情報処理システム１００においては、第１の端末装置２０及び第２の端末装置３０から得られる第１の参加者２００及び第２の参加者２０１の会議中の表情情報及び発話情報を元に、情報処理装置１０が第１の参加者２００と第２の参加者２０１との会議中の関係性を評価する。情報処理システム１００は、関係性を評価した結果を、第１の端末装置２０及び第２の端末装置３０の少なくとも一方に表示させることで、第１の参加者２００及び第２の参加者２０１の少なくとも一方に評価結果をリアルタイムでフィードバックし、会議中のコミュニケーション状態の改善を促す。

[0012] なお、本実実施形態において、「会議」とは、狭義の意味の会議に限定されるものではなく、面談、面接、カウンセリング、問診、接客、接見、相談等が含まれる。一例として、

・上司と部下との面談

・医師による患者への問診

・カウンセラーによる対象者へのカウンセリング

・客に対する店員の接客や窓口相談

・Ｗｅｂ　Ｍｅｅｔｉｎｇ等の遠隔でのコミュニケーション

・ｅ－Ｌｅａｒｎｉｎｇ等の対ビデオ画像に対するコミュニケーション

等が含まれる。

[0013] また、図３は、情報処理システム１００におけるデータの流れの概要を示す図である。

[0014] 〔第１の端末装置２０〕

図１に示すように、第１の端末装置２０は、カメラ２１、マイク２２、表示部２３、制御部２４、スピーカ２５、及び通信部２６を備えている。

[0015] ＜動画取得処理＞

カメラ２１は、第１の参加者を撮像し、撮像画像を制御部２４に供給する。ここで、カメラ２１による撮像画像は、動画像であることが好ましく、当該構成の場合、図３に示すように、カメラ２１は、制御部２４に対して、動画ファイル、及び動画ファイルに含まれる各画像のリストである動画ファイルリストの少なくとも何れかを供給する。カメラ２１は、当該動画ファイルに含まれる各画像の撮像時刻を示すタイムスタンプを制御部２４に供給する。

[0016] なお、第１の端末装置２０は、複数のカメラを備える構成としてもよく、当該構成の場合、制御部２４は、カメラデバイス識別情報を参照することによって、カメラ２１を特定することができる。

[0017] ＜画像認識処理＞

制御部２４は、カメラ２１から供給される動画ファイル、動画ファイルリスト、及びタイムスタンプを参照して、画像認識処理を行う。

[0018] 一例として、図３に示すように、制御部２４は、動画ファイル、動画ファイルリスト、及びタイムスタンプを参照した画像認識処理を行うことによって、時系列表情値、時系列顔パーツ座標、時系列視線座標を算出する。時系列表情値、時系列顔パーツ座標、時系列視線座標は、第１の参加者の表情に関する第１の表情情報の一例である。

[0019] ＜音声取得処理＞

マイク２２は、主として、第１の参加者の発話する音声を集音し、集音した音声を示す音声ファイル、及び、当該音声ファイルにおける発話の時点を特定するためのタイムスタンプを制御部２４に供給する。

[0020] なお、第１の端末装置２０は、複数のマイクを備える構成としてもよく、当該構成の場合、制御部２４は、音声デバイス識別情報を参照することによって、マイク２２を特定することができる。

[0021] ＜発話認識処理＞

制御部２４は、マイク２２から供給される音声ファイル及びタイムスタンプを参照して、発話認識処理を行う。

[0022] 一例として、図３に示すように、制御部２４は、音声ファイル、及びタイムスタンプを参照した発話認識処理を行うことによって、区間時系列テキストデータを生成する。ここで区間時系列テキストデータは、主として第１の参加者が発話した内容を時系列的にテキストデータとして示す情報である。当該区間時系列テキストデータは、第１の参加者の発話を示す第１の発話情報の一例である。

[0023] 〔第２の端末装置３０〕

また、第２の端末装置３０は、カメラ３１、マイク３２、表示部３３、制御部３４、スピーカ３５、及び通信部３６を備えている。

[0024] ＜動画取得処理＞

カメラ３１は、第２の参加者を撮像し、撮像画像を制御部３４に供給する。ここで、カメラ３１による撮像画像は、動画像であることが好ましく、当該構成の場合、図３に示すように、カメラ３１は、制御部３４に対して、動画ファイル、及び動画ファイルに含まれる各画像のリストである動画ファイルリストの少なくとも何れかを供給する。カメラ３１は、当該動画ファイルに含まれる各画像の撮像時刻を示すタイムスタンプを制御部３４に供給する。

[0025] なお、第２の端末装置３０は、複数のカメラを備える構成としてもよく、当該構成の場合、制御部３４は、カメラデバイス識別情報を参照することに

よって、カメラ３１を特定することができる。

[0026] ＜画像認識処理＞

制御部３４は、カメラ３１から供給される動画ファイル、動画ファイルリスト、及びタイムスタンプを参照して、画像認識処理を行う。

[0027] 一例として、図３に示すように、制御部３４は、動画ファイル、動画ファイルリスト、及びタイムスタンプを参照した画像認識処理を行うことによって、時系列表情値、時系列顔パーツ座標、時系列視線座標を算出する。時系列表情値、時系列顔パーツ座標、時系列視線座標は、第２の参加者の表情に関する第２の表情情報の一例である。

[0028] ＜音声取得処理＞

マイク３２は、主として、第２の参加者の発話する音声を集音し、集音した音声を示す音声ファイル、及び、当該音声ファイルにおける発話の時点を特定するためのタイムスタンプを制御部３４に供給する。

[0029] なお、第２の端末装置３０は、複数のマイクを備える構成としてもよく、当該構成の場合、制御部３４は、音声デバイス識別情報を参照することによって、マイク３２を特定することができる。

[0030] ＜発話認識処理＞

制御部３４は、マイク３２から供給される音声ファイル及びタイムスタンプを参照して、発話認識処理を行う。

[0031] 一例として、図３に示すように、制御部３４は、音声ファイル、及びタイムスタンプを参照した発話認識処理を行うことによって、区間時系列テキストデータを生成する。ここで区間時系列テキストデータは、主として第２の参加者が発話した内容を時系列的にテキストデータとして示す情報である。当該区間時系列テキストデータは、第２の参加者の発話を示す第２の発話情報の一例である。

[0032] 〔情報処理装置１０〕

情報処理装置１０は、表情情報取得部１３、音声情報取得部１４、表情関係性情報生成部１５、発話関係性情報生成部１６、及び関係性情報生成部１

７を備えている。情報処理装置１０は、さらに、通信部１１を備えている。表情情報取得部１３、音声情報取得部１４、表情関係性情報生成部１５、発話関係性情報生成部１６、関係性情報生成部１７は、演算部１２が備えている。

[0033] （表情情報取得部１３）

表情情報取得部１３は、通信部１１を介して複数の会議参加者のうち第１の参加者の表情に関する第１の表情情報と、複数の会議参加者のうち第２の参加者の表情に関する第２の表情情報とを取得する。

[0034] ＜数値データ洗浄処理＞

表情情報取得部１３は、一例として、通信部１１を介して、第１の参加者に関する表情情報に含まれる時系列数値データである時系列表情値、時系列顔パーツ座標、及び時系列視線座標を参照し、当該時系列数値データに対して、一例として以下の処理を行うことによって、数値データ洗浄処理を行う。

・無効データ区間を削除する

・有効データ区間におけるデータを平均する

・分散及び項数に変換する

表情情報取得部１３は、上述の数値データ洗浄処理を行うことによって、第１の参加者に関する区間時系列数値データを生成する。当該区間時系列数値データは、有効区間における時系列表情値、時系列顔パーツ座標、及び時系列視線座標を含んでいる。

[0035] 表情情報取得部１３は、第２の参加者に関する表情情報についても同様の処理を行い、第２の参加者に関する区間時系列数値データを生成する。

[0036] ＜表情の検出＞

表情情報取得部１３は、第１の参加者に関する区間時系列数値データを参照して、第１の参加者の表情を表現する複数の第１の指標を算出する。また、表情情報取得部１３は、第２の参加者に関する区間時系列数値データを参照して、第２の参加者の表情を表現する複数の第２の指標を算出する。

[0037] ここで、表情を表現する指標の例には、以下の指標が挙げられる。

・怒り（anger）

・侮辱（contempt）

・嫌悪（disgust）

・恐怖（fear）

・喜び（happiness）

・中立（neutral）

・悲しみ（sadness）

・驚き（surprise）

したがって、表情を表現する指標とは、当該表情が示す感情を表現する指標ということもできる。

[0038] なお、表情情報取得部１３は、第１の参加者に関する区間時系列数値データに含まれる時系列表情値を、そのまま第１の参加者の表情を表現する複数の第１の指標として用いてもよい。同様に、表情情報取得部１３は、第２の参加者に関する区間時系列数値データに含まれる時系列表情値を、そのまま第２の参加者の表情を表現する複数の第２の指標として用いてもよい。

[0039] また、第１の参加者の表情及び第２の参加者の表情は、上記の指標を各成分とするベクトルとして表現することもできる。こられのベクトルを、表情ベクトルと呼ぶこともある。

[0040] なお、各参加者の表情を検出する技術及び検出した表情が示す感情を指標化して表現する技術は、本実施形態を限定するものではなく、例えば、公知の技術を用いることができる。

[0041] ＜視線の検出＞

また、表情情報取得部１３は、通信部１１を介して、第１の端末装置２０及び第２の端末装置３０から、第１の参加者及び第２の参加者の視線方向に関する情報を取得する。具体的には、一例として、表情情報取得部１３は、第１の参加者の視線方向に関する情報として、上述した第１の参加者に関する区間時系列数値データに含まれる時系列視線座標を取得する。同様に、表

情情報取得部１３は、第２の参加者の視線方向に関する情報として、上述した第２の参加者に関する区間時系列数値データに含まれる時系列視線座標を取得する。

[0042] なお、視線座標の取得方法としては、特に限定されないが、第１の端末装置２０及び第２の端末装置３０に、点光源（不図示）を設け、点光源からの光の角膜反射像をカメラ２１及びカメラ３１で所定時間撮影することにより、ユーザの視線座標を取得する方法が挙げられる。点光源の種類は特に限定されず、可視光、赤外光が挙げられるが、例えば赤外線ＬＥＤを用いることで、ユーザに不快感を与えることなく、視線座標を取得することができる。

[0043] ＜距離の検出＞

また、表情情報取得部１３は、当該区間時系列数値データに含まれる第１の参加者の時系列顔パーツ座標を取得し、第１の参加者と撮像手段（カメラ２１）との間の距離を算出してもよい。また、表情情報取得部１３は、当該区間時系列数値データに含まれる第２の参加者の時系列顔パーツ座標を取得し、第２の参加者と撮像手段（カメラ３１）との間の距離を算出してもよい。参加者と撮像手段との間の距離は、例えば、顔パーツ座標から得られる撮像画像中の顔の目尻距離を顔角度補正したものを目尻距離とし、この目尻距離の逆数として算出することができる。

[0044] （音声情報取得部１４）

音声情報取得部１４は、第１の参加者の発話に関する第１の発話情報と、複数の会議参加者のうち第２の参加者の発話に関する第２の発話情報とを取得する。すなわち、音声情報取得部１４は、通信部１１を介して、第１の端末装置２０及び第２の端末装置３０から第１の参加者及び第２の参加者の発話に関する情報を取得する。

[0045] 音声情報取得部１４は、一例として、上述した第１の参加者に関する区間時系列テキストデータに含まれる時系列発話テキストを取得する。同様に、音声情報取得部１４は、一例として、上述した第２の参加者に関する区間時系列テキストデータに含まれる時系列発話テキストを取得する。

[0046] また、一例として、音声情報取得部１４は、第１の参加者に関する時系列発話テキストと、当該時系列発話テキストの発話の時点における時系列顔パーツ座標を取得する。音声情報取得部１４は、時系列顔パーツ座標を参照して、時系列発話テキストの発話の時点において第１の参加者の口が開いていれば、当該時系列発話テキストを第１の参加者の発話に関する第１の発話情報に含める。同様に、音声情報取得部１４は、第２の参加者に関する時系列発話テキストと、当該時系列発話テキストの発話の時点における時系列顔パーツ座標を取得する。音声情報取得部１４は、時系列顔パーツ座標を参照して、時系列発話テキストの発話の時点において第２の参加者の口が開いていれば、当該時系列発話テキストを第２の参加者の発話に関する第２の発話情報に含める。これにより、マイク２２又はマイク３２として、指向性の無い簡易なマイクを用いた場合でも、発話した人物を特定できる。

[0047] （表情関係性情報生成部１５）

表情関係性情報生成部１５は、第１の表情情報と第２の表情情報とを参照して、第１の参加者と第２の参加者との表情に関する関係性を示す表情関係性情報を生成する。

[0048] 会議参加者の会議に対する満足度は、会議の内容及び結論のみならず、参加者間の良好なコミュニケーションの有無にも依存する。参加者間のコミュニケーション状態は、会議中の参加者間の関係性により表され、参加者間の関係性は感情の一致度により評価することができる。表情関係性情報生成部１５は、表情情報取得部１３から第１の表情情報及び第２の表情情報を取得し、これらの表情情報を元に、会議中の参加者間の感情の一致度を参加者双方の表情から評価することで、参加者間のコミュニケーション状態をリアルタイムで評価する。

[0049] 表情関係性情報生成部１５が表情情報取得部１３から取得するそれぞれの表情情報は、区間時系列数値データを元に算出されたものであり、つまり、各参加者のリアルタイム又は経時的な表情に関する情報に基づいて算出されたものである。表情関係性情報生成部１５は、各参加者のリアルタイム又は

経時的な表情情報を元に、表情関係性情報を生成するので、生成された表情関係性情報は、参加者間のリアルタイム又は経時的な表情に関する関係性を表している。

[0050] ＜表情一致率判定＞

上述のように、第１の表情情報には、第１の参加者の表情を表現する複数の第１の指標が含まれており、第２の表情情報には、第２の参加者の表情を表現する複数の第２の指標が含まれている。

[0051] 表情関係性情報生成部１５は、第１の指標と第２の指標との差に関する表情差分情報を生成し、生成した表情差分情報を、表情関係性情報に含めてもよい。

[0052] 一例として、表情関係性情報生成部１５は、第１の参加者の表情を表現する複数の指標を要素とする第１の表情ベクトルと、第２の参加者の表情を表現する複数の指標を要素とする第２の表情ベクトルとの差の絶対値を用いて、表情不一致量を算出する。算出された表情不一致量は、参加者間の会議中の感情の融和状態を表す指標とも言える。また、表情関係性情報生成部１５は、会議開始から現時点までに、表情が一致した割合を示す指標として表情一致率を算出してもよい。表情一致率は、例えば、会議開始から現時点までの時間から表情が不一致であった時間を引いて、会議開始から現時点までの時間で除算することによって得られる。

[0053] ＜視線合致率判定＞

また、第１の表情情報が、第１の参加者の視線方向に関する第１の視線情報を含む構成とし、第２の表情情報が、第２の参加者の視線方向に関する第２の視線情報を含む構成としてもよい。表情関係性情報生成部１５は、第１の視線情報と第２の視線情報とを参照して視線関係性情報を生成し、生成した視線関係性情報を、表情関係性情報に含めてもよい。

[0054] 一例として、表情関係性情報生成部１５は、視線関係性情報として、第１の参加者と第２の参加者との視線合致率を算出する。算出された視線合致率は、会議中に他の参加者の様子を気にかけている状態を表す指標とも言える

。より具体的には、まず、制御部２４又は表情関係性情報生成部１５が、カメラ２１の撮像画像を解析することにより、会議室における第１の参加者の目の位置を特定し、制御部３４又は表情関係性情報生成部１５が、カメラ３１の撮像画像を解析することにより、会議室における第２の参加者の目の位置を特定する。

[0055] そして、表情関係性情報生成部１５は、各時点において、第１の視線情報が示す第１の参加者の視線方向が、第２の参加者の目に向かっているか否かを判定し、第２の視線情報が示す第２の参加者の視線方向が、第１の参加者の目に向かっているか否かを判定することにより、各時点において、第１の参加者の視線と第２の参加者の視線とが合致しているかを判定する。

[0056] 一例として、表情関係性情報生成部１５は、第１の参加者の視線が第２の参加者の目に向かっていると判定した場合に、第１の参加者の視線フラグを１に設定する。また、表情関係性情報生成部１５は、第２の参加者の視線が第１の参加者の目に向かっていると判定した場合に、第２の参加者の視線フラグを１に設定する。そして、表情関係性情報生成部１５は、双方の視線フラグが共に１である場合に、視線が合致していると判定する。

[0057] そして、表情関係性情報生成部１５は、会議開始から現時点までに、視線が合致した割合を示す指標として視線合致率を算出する。視線合致率は、例えば、視線が合致した時間を、会議開始から現時点までの時間で除算することによって得られる。

[0058] なお、視線が互いの目に向かっているか否かの判定には、第１の端末装置２０と第２の端末装置３０との相対的な位置関係を示す位置情報を更に参照する構成としてもよい。

[0059] また、互いの視線が必ずしも相手の目ではなく、相手の顔又は相手の身体の方向を向いている場合に、視線が合致していると判定する構成としてもよい。

[0060] また、参加者がインターネット等を介して会議する場合には、端末装置の画面を通した参加者間の視線合致率を算出する。より具体的には、一例とし

て、第１の端末装置２０の表示画面に表示される第２の参加者の顔の位置を、当該表示画面上の座標として特定し、特定した座標に対して第１の参加者の視線が向けられている場合に、第１の参加者の視線フラグを１に設定する。同様に、第２の端末装置３０の表示画面に表示される第１の参加者の顔の位置を、当該表示画面上の座標として特定し、特定した座標に対して第２の参加者の視線が向けられている場合に、第２の参加者の視線フラグを１に設定する。

[0061] ＜前のめり率判定＞

また、表情関係性情報生成部１５は、第１の参加者と第２の参加者との前のめり率を算出し、算出した前のめり率を表情関係性情報に含めてもよい。算出された前のめり率は、会議中に他の参加者の発話に興味を示している状態を表す指標とも言える。一例として、表情関係性情報生成部１５は、第１の参加者及び第２の参加者の、それぞれの撮像手段からの距離が、予め設定された一定時間内においてしきい値よりも下回った場合に、第１の参加者及び第２の参加者が前のめり状態であると判定する。

[0062] そして、表情関係性情報生成部１５は、会議開始から現時点までの時間において、第１の参加者が前のめりになっている時間の割合を、第１の参加者に関する前のめり率として特定し、第２の参加者が前のめりになっている時間の割合を、第２の参加者に関する前のめり率として特定する。

[0063] また、表情関係性情報生成部１５は、第１の参加者及び第２の参加者それぞれの撮像手段からの距離を元に得られる顔画像サイズについて、予め設定された一定時間内の変化から会議中の参加者の姿勢を算出し、表情関係性情報に含めてもよい。算出された参加者の姿勢は、会議中に他の参加者の発話を聞くにふさわしい態度を表す指標とも言える。

[0064] さらに、表情関係性情報生成部１５は、第１の参加者の姿勢の変化と第２の参加者の第２の表情ベクトルの変化との相関を算出し、その相関を表情関係性情報に含めてもよい。姿勢の変化と表情ベクトルの変化との相関は、一の参加者の姿勢が他の参加者の表情に及ぼす影響を表す指標とも言える。同

様に、表情関係性情報生成部１５は、第２の参加者の姿勢の変化と第１の参加者の第１の表情ベクトルの変化との相関を算出し、その相関を表情関係性情報に含めてもよい。

[0065] また、表情関係性情報生成部１５は、第１の参加者の姿勢と第２の参加者の姿勢とを参照して、第１の参加者と第２の参加者との姿勢状態の類似度を算出し、算出した類似度を表情関係性情報に含めてもよい。姿勢状態の類似度は、ミラーリング状態を表しており、会議中に他の参加者の発話に興味を示している状態を表す指標とも言える。

[0066] なお、表情関係性情報生成部１５は、第１及び第２の参加者の属性を示す参加者情報を更に参照して、表情関係性情報を生成してもよい。参加者の属性を示す参加者情報は、当該参加者の年齢、性別、血液型、性格、出身地、家族関係、役職、勤続年数、転職回数、職務履歴等の少なくとも何れかを含む。また、参加者情報には、当該システムの利用履歴も含まれる。

[0067] 一例として、表情関係性情報生成部１５は、参加者情報を参照し、当該参加者が特定の表情が出やすいと判断した場合には、当該特定の表情に対応する指標に１より小さい重み係数を乗算する補正を行うことによって当該参加者の表情ベクトルを補正し、補正後の表情ベクトルを用いて表情関係性情報を生成してもよい。

例えば、第１の参加者の属性を示す参加者情報が、当該第１の参加者が内気であることを示している場合、表情関係性情報生成部１５は、「中立（neutral）」の指標に対して重み０．８を乗算し、残り０．２の重みを他の指標に比例配分する等の処理を行うことによって、当該第１の参加者の表情ベクトルを補正し、補正後の表情ベクトルを用いて表情関係性情報を生成する構成としてもよい。

[0068] 情報処理装置１０は、参加者の脈波、脳波等の生体情報と、参加者周囲の温度、湿度、二酸化炭素濃度、照度等の環境情報とをさらに取得する構成とし、表情関係性情報生成部１５は、生体情報及び環境情報を更に参照して、表情関係性情報を生成してもよい。

[0069] 一例として、表情関係性情報生成部１５は、第１の参加者の脈波又は呼吸から判定した参加者のストレス状態と、その直前又はその時点における第２の参加者の表情を表現する第２の指標を参照し、第１の参加者にストレスを与える第２の参加者の表情を推定する。そして、表情関係性情報生成部１５は、推定された第２の参加者の表情を、第１の参加者に対するＮＧ表情と認定し、その情報を表情関係性情報に含めてもよい。一の参加者の他の参加者に対するＮＧ表情は、一の参加者の表情が他の参加者のストレス状態に及ぼす影響を表す指標とも言える。同様に、表情関係性情報生成部１５は、第２の参加者にストレスを与える第１の参加者の表情を推定し、第２の参加者に対するＮＧ表情を認定してもよい。

[0070] また、表情関係性情報生成部１５は、予め定められた一定期間内の、参加者の周囲の環境情報の変化と、第１の参加者の第１の表情ベクトル及び第２の参加者の第２の表情ベクトルの平均値の変化との相関を算出し、その相関を表情関係性情報に含めてもよい。環境情報の変化と表情ベクトルの平均値の変化との相関は、参加者の周囲の環境が参加者間のコミュニケーション状態に及ぼす影響を表す指標とも言える。

[0071] ＜対話管理処理＞

表情関係性情報生成部１５は、一例として、通信部１１を介して、第１の参加者及び第２の参加者それぞれのユーザＩＤと、当該ユーザＩＤが表す参加者が会議開始した時刻及び終了した時刻を表すタイムスタンプとを参照し、対話管理処理を行う。表情関係性情報生成部１５は、一の参加者に関する区間時系列数値データのある時点のデータについて、その時点に対話している他の参加者のユーザＩＤを抽出し、どの参加者と対話中に得られたデータであるかを判定して、結果を表情関係性情報に含めてもよい。

[0072] （発話関係性情報生成部１６）

発話関係性情報生成部１６は、第１の発話情報と第２の発話情報とを参照して、第１の参加者と第２の参加者との発話に関する関係性を示す発話関係性情報を生成する。発話関係性情報生成部１６は、音声情報取得部１４から

第１の発話情報及び第２の発話情報を取得し、これらの発話情報を元に、会議中の参加者間の感情の一致度を参加者双方の発話から評価することで、参加者間のコミュニケーション状態を評価する。

[0073] 発話関係性情報生成部１６が音声情報取得部１４から取得するそれぞれの発話情報は、区間時系列テキストデータを元に算出されたものであり、つまり、各参加者のリアルタイム又は経時的な発話に関する情報に基づいて算出されたものである。発話関係性情報生成部１６は、各参加者のリアルタイム又は経時的な発話情報を元に、発話関係性情報を生成するので、生成された発話関係性情報は、参加者間のリアルタイム又は経時的な発話に関する関係性を表している。

[0074] ＜発話比率判定＞

発話関係性情報生成部１６は、第１の発話情報が示す第１の参加者の発話時間と、第２の発話情報が示す第２の参加者の発話時間との関係を示す発話時間関係性情報を生成し、生成した発話時間関係性情報を、発話関係性情報に含めてもよい。

[0075] 一例として、発話関係性情報生成部１６は予め定められた一定時間内の、第１の参加者の発話時間と第２の参加者の発話時間との発話比率を算出し、発話関係性情報に含める。算出された発話比率は、参加者間の関係の対等性を表す指標とも言える。

[0076] ＜発話頻度判定＞

また、発話関係性情報生成部１６は、第１の発話情報及び第２の発話情報の少なくとも何れかに、特定のカテゴリーに含まれる発話内容が含まれているか否かを判定し、判定した結果に応じた情報を前記発話関係性情報に含めてもよい。

[0077] 発話内容に含まれる特定のカテゴリーの例には、オープンクエスチョン、行動促しワード（それで？、なるほど、確かに）、オウム返し、発話の遮り、発話の被り、否定ワード（でも、だけど）が含まれる。一例として、発話関係性情報生成部１６は、予め定められた一定時間内における、このような

特定のカテゴリーに含まれる発話内容が発話された頻度を算出し、算出した頻度に関する情報を発話関係性情報に含める。

[0078] 具体的には、一例として、発話内容に含まれる特定カテゴリーをオープンクエスチョンと設定し、発話関係性情報生成部１６は、第１の参加者の区間時系列テキストデータから、一定時間内に含まれるオープンクエスチョンを表すテキストデータを抽出する。そして、発話関係性情報生成部１６は、抽出されたテキストデータの単語数を上記一定時間内の全テキストデータの単語数で除算することによって、オープンクエスチョンが発話された頻度をオープン質問率として算出する。同様に、発話関係性情報生成部１６は、第２の参加者の区間時系列テキストデータから、オープン質問率を算出する。そして、発話関係性情報生成部１６は、第１の参加者のオープン質問率と第２の参加者のオープン質問率とを比較して、オープン質問比率を算出し、発話関係性情報に含める。算出されたオープン質問比率は、参加者間の関係の対等性を表す指標とも言える。また、発話関係性情報生成部１６は、第１の参加者のオープン質問率及び第２の参加者のオープン質問率を、発話関係性情報に含めてもよい。

[0079] 同様に、発話内容に含まれる特定カテゴリーを行動促しワードと設定し、発話関係性情報生成部１６は、第１の参加者の区間時系列テキストデータから、一定時間内に含まれる行動促しワードを表すテキストデータを抽出する。そして、発話関係性情報生成部１６は、抽出されたテキストデータの単語数を上記一定時間内の全テキストデータの単語数で除算することによって、行動促しワードが発話された頻度を促し質問率として算出する。

[0080] 同様に、発話関係性情報生成部１６は、第２の参加者の区間時系列テキストデータから、促し質問率を算出する。そして、発話関係性情報生成部１６は、第１の参加者の促し質問率と第２の参加者の促し質問率とを比較して、促し質問比率を算出し、発話関係性情報に含める。算出された促し質問比率は、参加者間の関係の対等性を表す指標とも言える。また、発話関係性情報生成部１６は、第１の参加者の促し質問率及び第２の参加者の促し質問率を

、発話関係性情報に含めてもよい。

[0081] ＜単語に基づく評価＞

また、発話関係性情報生成部１６は、第１の発話情報及び第２の発話情報の少なくとも何れかから、所定時間内において相対的に出現頻度の高い単語を抽出し、抽出した単語を発話関係性情報に含めてもよい。

[0082] 一例として、発話関係性情報生成部１６は、第１の参加者及び第２の参加者のそれぞれの区間時系列テキストデータから、予め定められた一定時間内に含まれる各単語の出現数を参加者毎に算出して順位付けし、相対的に出現頻度の高い単語を上位から複数抽出する。そして、発話関係性情報生成部１６は、抽出した出現頻度の上位の単語を頻出単語として発話関係性情報に含める。また、発話関係性情報生成部１６は、第１の参加者及び第２の参加者のそれぞれの頻出単語及びその順位が一致しているかを判定し、その判定結果を発話関係性情報に含めてもよい。

[0083] また、発話関係性情報生成部１６は、第１の参加者及び第２の参加者のそれぞれの区間時系列テキストデータから、予め定められた一定時間内に含まれる単語を参加者毎に抽出し、抽出された単語の一致率を算出して、発話関係性情報に含めてもよい。単語の一致率は、予め定められた一定時間内の区間時系列テキストデータに含まれる全単語中における、第１の参加者と第２の参加者とで一致した単語の比率として算出することができる。算出された単語の一致率は、オウム返しができているかの指標とも言える。

[0084] また、発話関係性情報生成部１６は、第１の参加者及び第２の参加者のそれぞれの区間時系列テキストデータから発話タイミングを抽出し、発話タイミングの時間的なオーバーラップを算出してもよい。そして、発話関係性情報生成部１６は、予め定められた一定時間内のオーバーラップの回数をオーバーラップ頻度として算出し、発話関係性情報に含めてもよい。算出したオーバーラップ頻度は、他の参加者の発話をさえぎる頻度を表す指標とも言える。

[0085] さらに、発話関係性情報生成部１６は、第１及び第２の参加者の属性を示

す参加者情報を更に参照して、発話関係性情報を生成してもよい。参加者の属性を示す参加者情報は、当該参加者の年齢、性別、血液型、性格、出身地、家族関係、役職、勤続年数、転職回数、職務履歴等の少なくとも何れかを含む。また、参加者情報には、当該システムの利用履歴も含まれる。

[0086] 情報処理装置１０は、参加者の脈波、脳波等の生体情報と、参加者周囲の温度、湿度、二酸化炭素濃度、照度等の環境情報とをさらに取得する構成とし、発話関係性情報生成部１６は、生体情報及び環境情報を更に参照して、発話関係性情報を生成してもよい。

[0087] また、一例として、発話関係性情報生成部１６は、第１の参加者の脈波又は呼吸から判定した参加者のストレス状態と、その直前又はその時点における第２の参加者の区間時系列テキストデータとを参照し、第１の参加者にストレスを与える第２の参加者のテキストデータを推定する。そして、発話関係性情報生成部１６は、推定された第２の参加者のテキストデータを、第１の参加者に対するＮＧワードと認定し、その情報を発話関係性情報に含めてもよい。一の参加者の他の参加者に対するＮＧワードは、一の参加者の発言が他の参加者のストレス状態に及ぼす影響を表す指標とも言える。同様に、発話関係性情報生成部１６は、第２の参加者にストレスを与える第１の参加者のテキストデータを推定し、第２の参加者に対するＮＧワードを認定してもよい。

[0088] また、一例として、発話関係性情報生成部１６は、第１の参加者の脳波から判定した参加者の思考の活性度と、その直前又はその時点における第２の参加者の区間時系列テキストデータとを参照し、第１の参加者の思考を活性化させる第２の参加者のテキストデータを推定する。そして、発話関係性情報生成部１６は、推定された第２の参加者のテキストデータを、第１の参加者に対する重要ワードと認定し、その情報を発話関係性情報に含めてもよい。一の参加者の他の参加者に対する重要ワードは、一の参加者の発言が他の参加者の思考の活性化に及ぼす影響を表す指標とも言える。同様に、発話関係性情報生成部１６は、第２の参加者の思考を活性化する第１の参加者のテ

キストデータを推定し、第２の参加者に対する重要ワードを認定してもよい。

[0089] また、発話関係性情報生成部１６は、参加者間の声のトーンの一致度、発話スピードの一致度、音量の一致度等を算出してもよい。

[0090] さらに、発話関係性情報生成部１６は、蓄積された区間時系列テキストデータから、第１の参加者及び第２の参加者それぞれの発話を表すログを抽出し、その形態素解析データを取得して、過去の発話中の頻出単語をリストアップし、発話関係性情報に含めてもよい。過去の発話中の頻出単語は、会議中の参加者双方に提示することで、会議のテーマ決定を支援するために利用され得る。

[0091] ＜対話管理処理＞

発話関係性情報生成部１６は、一例として、通信部１１を介して、第１の参加者及び第２の参加者それぞれのユーザＩＤと、当該ユーザＩＤが表す参加者が会議開始した時刻及び終了した時刻を表すタイムスタンプとを参照し、対話管理処理を行う。発話関係性情報生成部１６は、一の参加者に関する区間時系列テキストデータのある時点のデータについて、その時点に対話している他の参加者のユーザＩＤを抽出し、どの参加者と対話中に得られたデータであるかを判定して、結果を発話関係性情報に含めてもよい。

[0092] （関係性情報生成部１７）

関係性情報生成部１７は、表情関係性情報と前記発話関係性情報とを参照して前記第１の参加者と前記第２の参加者との関係を示すリアルタイム又は経時的な情報である関係性情報を生成する。会議中の参加者双方の表情及び発話の両方を評価することで、参加者間のコミュニケーション状態をより詳細に評価することができる。また、関係性情報生成部１７は、参加者間のリアルタイム又は経時的な表情情報及び発話情報を元に関係性情報を生成するので、参加者間のリアルタイム又は経時的なコミュニケーション状態を評価することができる。

[0093] 関係性情報生成部１７は、第１の参加者及び第２の参加者の少なくとも何

れかに提示する提示情報を生成してもよい。提示情報には、表情関係性情報と発話関係性情報とに基づき総合的に評価した参加者双方の感情の一致度等が含まれていてもよい（例えば、視線合致率が高く、発話比率が対等であれば感情の一致度を高くする等）。

[0094] 関係性情報生成部１７が生成した提示情報を参加者に提示することで、参加者間の関係性を参加者にフィードバックすることができる。提示情報をリアルタイムで参加者に提示すれば、会話中にリアルタイムで関係性を確認することができるので、リアルタイムでコミュニケーションの改善を促すことも可能である。

[0095] 提示情報は、第１の参加者及び第２の参加者の双方に提示するものであってもよいし、いずれか一方に提示するものであってもよい。また、関係性情報は、第１の参加者及び第２の参加者に同じ内容を提示するものであってもよいし、異なる内容を提示するものであってもよい。第１の参加者及び第２の参加者に同じ内容を提示する関係性情報を生成することで、参加者間のフラットな関係性の構築が期待できる。また、提示情報を参加者自身が選択できるようになっていてもよいし、ルール又は参加者間の合意により提示される提示情報が変更されてもよい。

[0096] 関係性情報には、第１の参加者の発話時間と、第２の参加者の発話時間との割合を示す情報、及び、第１の参加者の視線方向と、第２の参加者の視線方向との合致率の経時変化に関する情報が含まれていてもよい。また、関係性情報には、表情一致率又は表情不一致率の経時変化、前のめり率、発話内容のテキスト、頻出単語等に関する情報が含まれていてもよい。さらに、関係性情報には、参加者のＩＤ、参加者自身の顔画像、他の参加者の表情を表すアバター画像、発話内容に基づき蓄積データから抽出した推奨議題又は推奨ワードの表示等が含まれていてもよい。

[0097] また、提示情報に、会議参加者のコミュニケーションスキルを向上させるための評価結果を含めてもよい。一例として、表情一致率と共に、表情一致率を高めることで反射的傾聴スキルが向上させることを促す情報を提示した

り、視線合致率と共に、視線合致率を高めることでコミュニケーションに適した姿勢、態度を取るように促す情報を提示したりしてもよい。また、推奨するワードや質問内容を提示して、対話レベル及び質問レベルの控除を促してもよい。

[0098] 提示情報を提示する方法は、具体的には、一例として、会議参加者の表示部（表示部２３及び表示部３３のそれぞれ）に表示する方法、会議参加者全員が視認できる共通の表示部に表示する方法、ネットワーク配信等により会議参加者以外にも提示する方法、腕時計型デバイスのようなウェアラブルデバイスからの物理的な作用（振動、電気刺激等）により提示する方法、環境設備（証明、空調、音楽等）からの物理的な作用（議論が白熱した場合に部屋を赤く照らす等）により提示する方法、感情を表す指標に対応した画像イメージ（怒りを表す火山の噴火等）により提示する方法、感情を表す指標に対応したアバターの表情により提示する方法等が挙げられる。

[0099] 図４及び５を参照して、表示部２３及び表示部３３の少なくとも一方に提示情報を提示される画面例を説明する。図４は、本発明の一実施形態に係る情報処理装置を含む情報処理システムが提示する情報の一例を示す図であり、図５は、本発明の一実施形態に係る情報処理装置を含む情報処理システムが提示する情報の他の例を示す図である。

[0100] 図４に示すように、画面４００において、領域４０１に会議参加者のユーザＩＤを表示し、領域４０２に会議参加者の顔画像を表示して、提示情報を提示する対象を特定する。領域４０３に発話比率をトーク比率として、例えば円グラフで表示し、領域４０４に対話中の参加者の感情を表す指標に対応したアバターの表情を表示し、また、領域４０５に表情一致率の経時変化をグラフで表示することで、会議中にコミュニケーションの状態の現状を瞬時に確認できるようにする。また、領域４０６に、発話内容を表すテキストをＴａｌｋ　Ｓｔｒｅａｍとして表示し、また、会議中に推奨される会話テーマや単語を表示する。領域４０５に表示する表情一致率は、表情関係性情報について、過去から現在までの経時的な情報に基づき生成された提示情報の

例である。領域４０４に表示するアバターの表情は、表情関係性情報について、リアルタイムの情報に基づき生成された提示情報の例である。

[0101] また、図５に示すように、画面５００において、画面４００と同様に、領域５０１にユーザＩＤを表示し、領域５０２に顔画像を表示し、領域５０３にトーク比率を表示し、領域５０４にアバター表情を表示し、領域５０５に発話内容及び推奨テーマ等を表示すると共に、領域５０５に、表情一致率ではなく視線合致率を表示してもよい。

[0102] ＜付記事項１＞

端末装置の制御部における処理の一部又は全部を、情報処理装置１０の備える演算部１２において行う構成としてもよい。例えば、演算部１２が、通信部１１を介して、カメラ２１の撮像画像を取得し、表情情報取得部１３において、第１の参加者の表情に関する第１の表情情報、及び、第２の参加者の表情に関する第２の表情情報を生成する構成としてもよい。

[0103] ＜付記事項２＞

また、上記の例では、第１の参加者及び第２の参加者の２名による会議を例にしたが、本実施形態はこれに限定されるものではない。当然、Ｎ名（Ｎは３以上）による会議に対しても本明細書に記載の発明を適用することができる。その場合、Ｎ名中の任意の２人のペアに対して、本明細書に記載の構成を個別に適用することができる。例えば、３名（Ａ、Ｂ、Ｃ）による会議に対しては、（Ａ、Ｂ）（Ａ、Ｃ）（Ｂ、Ｃ）の３組に対して本明細書に記載の発明を個別に適用することができる。

このように、本実施形態に記載の発明は、Ｎ人の状態を表すデータ，Ｎ人の状態の履歴データおよびＮ人の環境情報を用いて、当該Ｎ人の内の一部又は全部の参加者間の関係を示す「関係性情報」を生成するものであると表現することもできる。

＜付記事項３＞

また、上記の例では、第１の参加者及び第２の参加者が共に人間である場合を例に挙げたが、これは本実施形態を限定するものではない。

例えば、第２の参加者は、人間ではなく、予め設定されたアバターやＢＯＴのようにコンピュータによって表現される疑似的な人間であってもよい。このような構成の場合、第２の端末装置は、必須ではなく、表情情報取得部１３及び音声情報取得部１４は、予め作成された当該ＢＯＴが表す表情及び発話内容を、第２の参加者の表情情報及び音声情報として取得する構成とすればよい。

なお、ＢＯＴが表す表情及び発話内容は、会議前に事前に作成されたデータを用いてもよいし、会議中の第１の参加者の表情や発話に応じて適応的に変更される構成としてもよい。

[0104] 〔ソフトウェアによる実現例〕

情報処理装置１の制御ブロック（特に演算部１２）は、集積回路（ＩＣチップ）等に形成された論理回路（ハードウェア）によって実現してもよいし、ソフトウェアによって実現してもよい。

[0105] 後者の場合、情報処理装置１は、各機能を実現するソフトウェアであるプログラムの命令を実行するコンピュータを備えている。このコンピュータは、例えば１つ以上のプロセッサを備えていると共に、上記プログラムを記憶したコンピュータ読み取り可能な記録媒体を備えている。そして、上記コンピュータにおいて、上記プロセッサが上記プログラムを上記記録媒体から読み取って実行することにより、本発明の目的が達成される。上記プロセッサとしては、例えばＣＰＵ（Central Processing Unit）を用いることができる。上記記録媒体としては、「一時的でない有形の媒体」、例えば、ＲＯＭ（Read Only Memory）等の他、テープ、ディスク、カード、半導体メモリ、プログラマブルな論理回路などを用いることができる。また、上記プログラムを展開するＲＡＭ（Random Access Memory）などをさらに備えていてもよい。また、上記プログラムは、該プログラムを伝送可能な任意の伝送媒体（通信ネットワークや放送波等）を介して上記コンピュータに供給されてもよい。なお、本発明の一態様は、上記プログラムが電子的な伝送によって具現化された、搬送波に埋め込まれたデータ信号の形態でも実現され得る。

[0106] 〔まとめ〕

本発明の一態様に係る情報処理装置は、複数の参加者のうち第１の参加者の表情に関する第１の表情情報と、前記複数の参加者のうち第２の参加者の表情に関する第２の表情情報とを取得する表情情報取得部と、当該第１の参加者の発話に関する第１の発話情報と、前記複数の参加者のうち第２の参加者の発話に関する第２の発話情報とを取得する音声情報取得部と、前記第１の表情情報と前記第２の表情情報とを参照して、前記第１の参加者と前記第２の参加者との表情に関する関係性を示す表情関係性情報を生成する表情関係性情報生成部と、前記第１の発話情報と前記第２の発話情報とを参照して、前記第１の参加者と前記第２の参加者との発話に関する関係性を示す発話関係性情報を生成する発話関係性情報生成部と、前記表情関係性情報と前記発話関係性情報とを参照して前記第１の参加者と前記第２の参加者との関係を示す情報である関係性情報を生成する関係性情報生成部と、を備えていることを特徴としている。

[0107] これによれば、各参加者の音声情報及び表情情報の両方に基づいて、会議中の参加者間の関係性を評価することができる。

[0108] 前記一態様に係る情報処理装置において、前記関係性情報は、前記第１の参加者と前記第２の参加者との関係を示すリアルタイム又は経時的な情報である。

[0109] これによれば、各参加者の音声情報及び表情情報の両方に基づいて、会議中の参加者間の関係性をリアルタイムで評価することができる。

[0110] 前記一態様に係る情報処理装置において、前記第１の表情情報には、前記第１の参加者の表情を表現する複数の第１の指標が含まれており、前記第２の表情情報には、前記第２の参加者の表情を表現する複数の第２の指標が含まれており、前記表情関係性情報生成部は、前記第１の指標と前記第２の指標との差に関する表情差分情報を生成し、生成した表情差分情報を、前記表情関係性情報に含める。

[0111] これによれば、表情関係性情報を生成するために参照する参加者の表情を

表現するために複数の指標を用いるため、より正確に参加者の表情を表現することができる。

[0112] 前記一態様に係る情報処理装置において、前記第１の表情情報には、前記第１の参加者の視線方向に関する第１の視線情報が含まれており、前記第２の表情情報には、前記第２の参加者の視線方向に関する第２の視線情報が含まれており、前記表情関係性情報生成部は、前記第１の視線情報と前記第２の視線情報とを参照して視線関係性情報を生成し、生成した視線関係性情報を、前記表情関係性情報に含める。

[0113] これによれば、表情関係性情報を生成するために参照する参加者の表情情報に参加者の視線情報も含むため、より正確に参加者の表情を表現することができる。

[0114] 前記一態様に係る情報処理装置において、前記発話関係性情報生成部は、前記第１の発話情報が示す前記第１の参加者の発話時間と、前記第２の発話情報が示す前記第２の参加者の発話時間との関係を示す発話時間関係性情報を生成し、生成した発話時間関係性情報を、前記発話関係性情報に含める。

[0115] これによれば、発話関係性情報を生成するために参照する参加者の発話情報に発話時間関係性情報も含むため、より正確に参加者の発話関係性情報を生成することができる。

[0116] 前記一態様に係る情報処理装置において、前記発話関係性情報生成部は、前記第１の発話情報及び前記第２の発話情報の少なくとも何れかに、特定のカテゴリーに含まれる発話内容が含まれているか否かを判定し、判定した結果に応じた情報を前記発話関係性情報に含める。

[0117] これによれば、発話関係性情報に特定のカテゴリーに含まれる発話内容が含まれているか否かの判定結果に応じた情報も含むため、より正確に参加者間の関係性情報を生成することができる。

[0118] 前記一態様に係る情報処理装置において、前記発話関係性情報生成部は、前記第１の発話情報及び前記第２の発話情報の少なくとも何れかから、所定時間内において相対的に出現頻度の高い単語を抽出し、抽出した単語を前記

発話関係性情報に含める。

[0119] これによれば、発話関係性情報に頻度の高い単語の上方も含むため、より正確に参加者間の関係性情報を生成することができる。

[0120] 前記一態様に係る情報処理装置において、前記関係性情報生成部は、前記関係性情報を参照して、前記第１の参加者及び前記第２の参加者の少なくとも何れかに提示する提示情報を生成する。

[0121] これによれば、提示情報を参加者に提示することで、関係性情報を参加者に認識させることができる。

[0122] 前記提示情報には、前記第１の参加者の発話時間と、前記第２の参加者の発話時間との割合を示す情報、及び、前記第１の参加者の視線方向と、前記第２の参加者の視線方向との合致率の経時変化に関する情報が含まれている。

[0123] これによれば、提示情報を参加者に提示することで、各参加者の発話時間の割合、及び、各参加者の視線方向の合致率の経時変化を参加者に認識させることができる。

[0124] 前記一態様に係る情報処理装置において、前記表情関係性情報生成部、及び前記発話関係性情報生成部は、前記第１及び第２の参加者の属性を示す参加者情報を更に参照して、前記表情関係性情報及び前記発話関係性情報を生成する。

[0125] これによれば、表情関係性情報及び発話関係性情報の生成に参加者の属性も参照するので、より正確な表情関係性情報及び発話関係性情報を生成することができる。

[0126] また、前記課題を解決するために、本発明の一態様に係る情報処理方法は、複数の参加者のうち第１の参加者の表情に関する第１の表情情報と、前記複数の参加者のうち第２の参加者の表情に関する第２の表情情報とを取得する表情情報取得ステップと、当該第１の参加者の発話に関する第１の発話情報と、前記複数の参加者のうち第２の参加者の発話に関する第２の発話情報とを取得する音声情報取得ステップと、前記第１の表情情報と前記第２の表

情情報とを参照して、前記第１の参加者と前記第２の参加者との表情に関する関係性を示す表情関係性情報を生成する表情関係性情報生成ステップと、前記第１の発話情報と前記第２の発話情報とを参照して、前記第１の参加者と前記第２の参加者との発話に関する関係性を示す発話関係性情報を生成する発話関係性情報生成ステップと、前記表情関係性情報と前記発話関係性情報とを参照して前記第１の参加者と前記第２の参加者との関係を示すリアルタイム又は経時的な情報である関係性情報を生成する関係性情報生成ステップと、を含む、ことを特徴としている。

[0127] これによれば、各参加者の音声情報及び表情情報に基づいて、会議中の参加者間の関係性を評価することができる。

[0128] また、前記課題を解決するために、本発明の一態様に係る情報処理プログラムは、前記何れかに記載の情報処理装置としてコンピュータを機能させるための情報処理プログラムであって、前記表情情報取得部、前記音声情報取得部、前記表情関係性情報生成部、前記発話関係性情報生成部、及び前記関係性情報生成部としてコンピュータを機能させる。

[0129] これによれば、各参加者の音声情報及び表情情報に基づいて、会議中の参加者間の関係性を評価することができる。

[0130] 本発明は上述した各実施形態に限定されるものではなく、請求項に示した範囲で種々の変更が可能であり、異なる実施形態にそれぞれ開示された技術的手段を適宜組み合わせて得られる実施形態についても本発明の技術的範囲に含まれる。

符号の説明

[0131] １００ 情報処理システム

１０ 情報処理装置

２０ 第１の端末装置

３０ 第２の端末装置

１３ 表情情報取得部

１４ 音声情報取得部

１５　表情関係性情報生成部

１６　発話関係性情報生成部

１７　関係性情報生成部

請求の範囲

[請求項1] 複数の参加者のうち第１の参加者の表情に関する第１の表情情報と、前記複数の参加者のうち第２の参加者の表情に関する第２の表情情報とを取得する表情情報取得部と、

当該第１の参加者の発話に関する第１の発話情報と、前記複数の参加者のうち第２の参加者の発話に関する第２の発話情報とを取得する音声情報取得部と、

前記第１の表情情報と前記第２の表情情報とを参照して、前記第１の参加者と前記第２の参加者との表情に関する関係性を示す表情関係性情報を生成する表情関係性情報生成部と、

前記第１の発話情報と前記第２の発話情報とを参照して、前記第１の参加者と前記第２の参加者との発話に関する関係性を示す発話関係性情報を生成する発話関係性情報生成部と、

前記表情関係性情報と前記発話関係性情報とを参照して前記第１の参加者と前記第２の参加者との関係を示す情報である関係性情報を生成する関係性情報生成部と、

を備えていることを特徴とする情報処理装置。

[請求項2] 前記関係性情報は、前記第１の参加者と前記第２の参加者との関係を示すリアルタイム又は経時的な情報である

ことを特徴とする請求項１に記載の情報処理装置。

[請求項3] 前記第１の表情情報には、前記第１の参加者の表情を表現する複数の第１の指標が含まれており、

前記第２の表情情報には、前記第２の参加者の表情を表現する複数の第２の指標が含まれており、

前記表情関係性情報生成部は、前記第１の指標と前記第２の指標との差に関する表情差分情報を生成し、生成した表情差分情報を、前記表情関係性情報に含める

ことを特徴とする請求項１又は２に記載の情報処理装置。

[請求項4] 前記第１の表情情報には、前記第１の参加者の視線方向に関する第１の視線情報が含まれており、

前記第２の表情情報には、前記第２の参加者の視線方向に関する第２の視線情報が含まれており、

前記表情関係性情報生成部は、前記第１の視線情報と前記第２の視線情報とを参照して視線関係性情報を生成し、生成した視線関係性情報を、前記表情関係性情報に含める

ことを特徴とする請求項１から３の何れか１項に記載の情報処理装置。

[請求項5] 前記発話関係性情報生成部は、

前記第１の発話情報が示す前記第１の参加者の発話時間と、前記第２の発話情報が示す前記第２の参加者の発話時間との関係を示す発話時間関係性情報を生成し、生成した発話時間関係性情報を、前記発話関係性情報に含める

ことを特徴とする請求項１から４の何れか１項に記載の情報処理装置。

[請求項6] 前記発話関係性情報生成部は、

前記第１の発話情報及び前記第２の発話情報の少なくとも何れかに、特定のカテゴリーに含まれる発話内容が含まれているか否かを判定し、判定した結果に応じた情報を前記発話関係性情報に含める

ことを特徴とする請求項１から５の何れか１項に記載の情報処理装置。

[請求項7] 前記発話関係性情報生成部は、

前記第１の発話情報及び前記第２の発話情報の少なくとも何れかから、所定時間内において相対的に出現頻度の高い単語を抽出し、抽出した単語を前記発話関係性情報に含めることを特徴とする請求項１から６の何れか１項に記載の情報処理装置。

[請求項8] 前記関係性情報生成部は、前記関係性情報を参照して、前記第１の

参加者及び前記第２の参加者の少なくとも何れかに提示する提示情報を生成する

ことを特徴とする請求項１から７の何れか１項に記載の情報処理装置。

[請求項9] 前記提示情報には、

前記第１の参加者の発話時間と、前記第２の参加者の発話時間との割合を示す情報、及び、

前記第１の参加者の視線方向と、前記第２の参加者の視線方向との合致率の経時変化に関する情報

が含まれている

ことを特徴とする請求項８に記載の情報処理装置。

[請求項10] 前記表情関係性情報生成部、及び前記発話関係性情報生成部は、前記第１及び第２の参加者の属性を示す参加者情報を更に参照して、前記表情関係性情報及び前記発話関係性情報を生成する

ことを特徴とする請求項１から９の何れか１項に記載の情報処理装置。

[請求項11] 複数の参加者のうち第１の参加者の表情に関する第１の表情情報と、前記複数の参加者のうち第２の参加者の表情に関する第２の表情情報とを取得する表情情報取得ステップと、

当該第１の参加者の発話に関する第１の発話情報と、前記複数の参加者のうち第２の参加者の発話に関する第２の発話情報とを取得する音声情報取得ステップと、

前記第１の表情情報と前記第２の表情情報とを参照して、前記第１の参加者と前記第２の参加者との表情に関する関係性を示す表情関係性情報を生成する表情関係性情報生成ステップと、

前記第１の発話情報と前記第２の発話情報とを参照して、前記第１の参加者と前記第２の参加者との発話に関する関係性を示す発話関係性情報を生成する発話関係性情報生成ステップと、

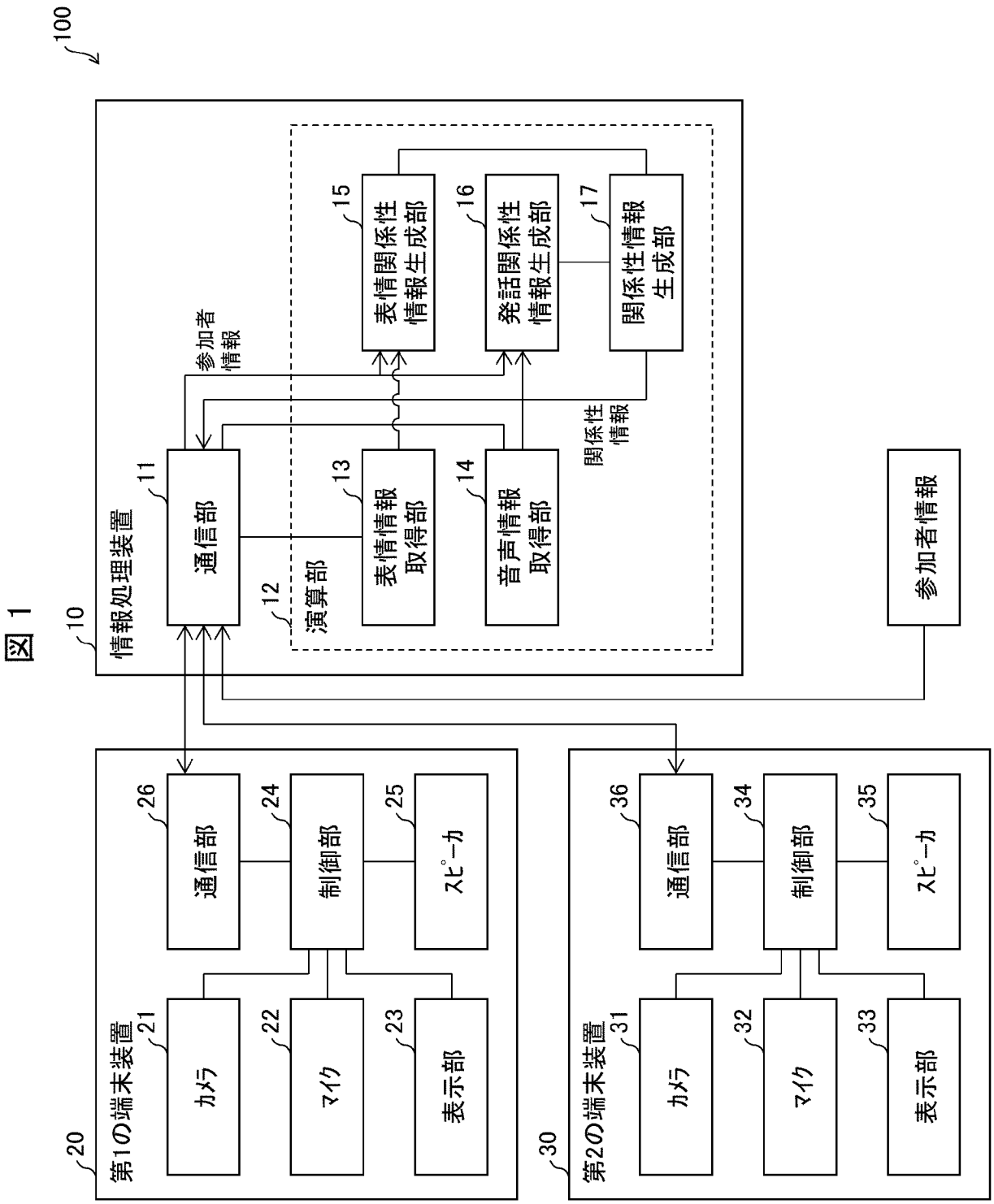
前記表情関係性情報と前記発話関係性情報とを参照して前記第１の参加者と前記第２の参加者との関係を示すリアルタイム又は経時的な情報である関係性情報を生成する関係性情報生成ステップと、

を含むことを特徴とする情報処理方法。

[請求項12] 請求項１～１０の何れか１項に記載の情報処理装置としてコンピュータを機能させるための情報処理プログラムであって、前記表情情報取得部、前記音声情報取得部、前記表情関係性情報生成部、前記発話関係性情報生成部、及び前記関係性情報生成部としてコンピュータを機能させるための情報処理プログラム。

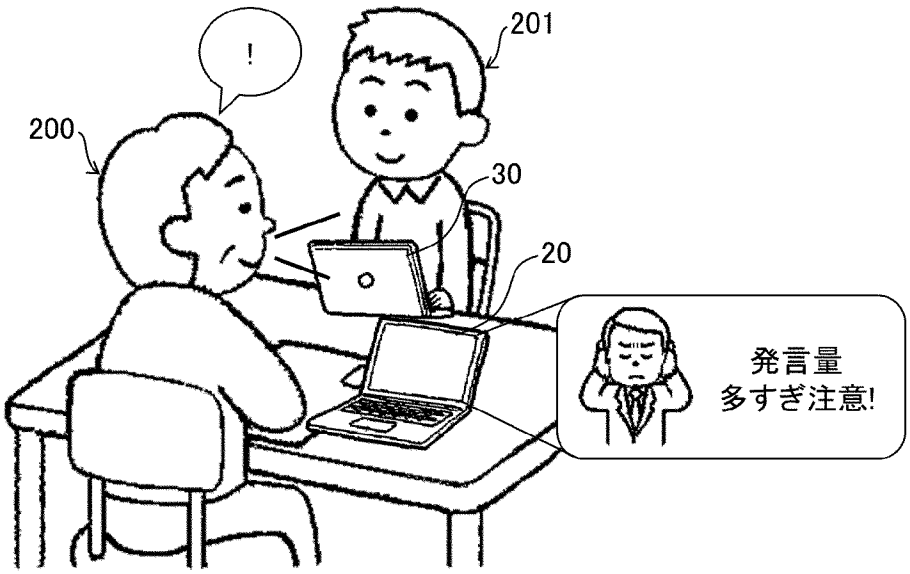
[図1]

図 1


100
情報処理装置
10
通信部
11
演算部
12
表情情報
取得部
13
音声情報
取得部
14
表情関係性
情報生成部
15
発話関係性
情報生成部
16
関係性情報
生成部
17
参加者
情報
関係性
情報
参加者情報
第1の端末装置
20
カメラ
21
マイク
22
スピーカ
23
通信部
26
制御部
24
スピーカ
25
第2の端末装置
30
カメラ
31
マイク
32
スピーカ
33
通信部
36
制御部
34
スピーカ
35

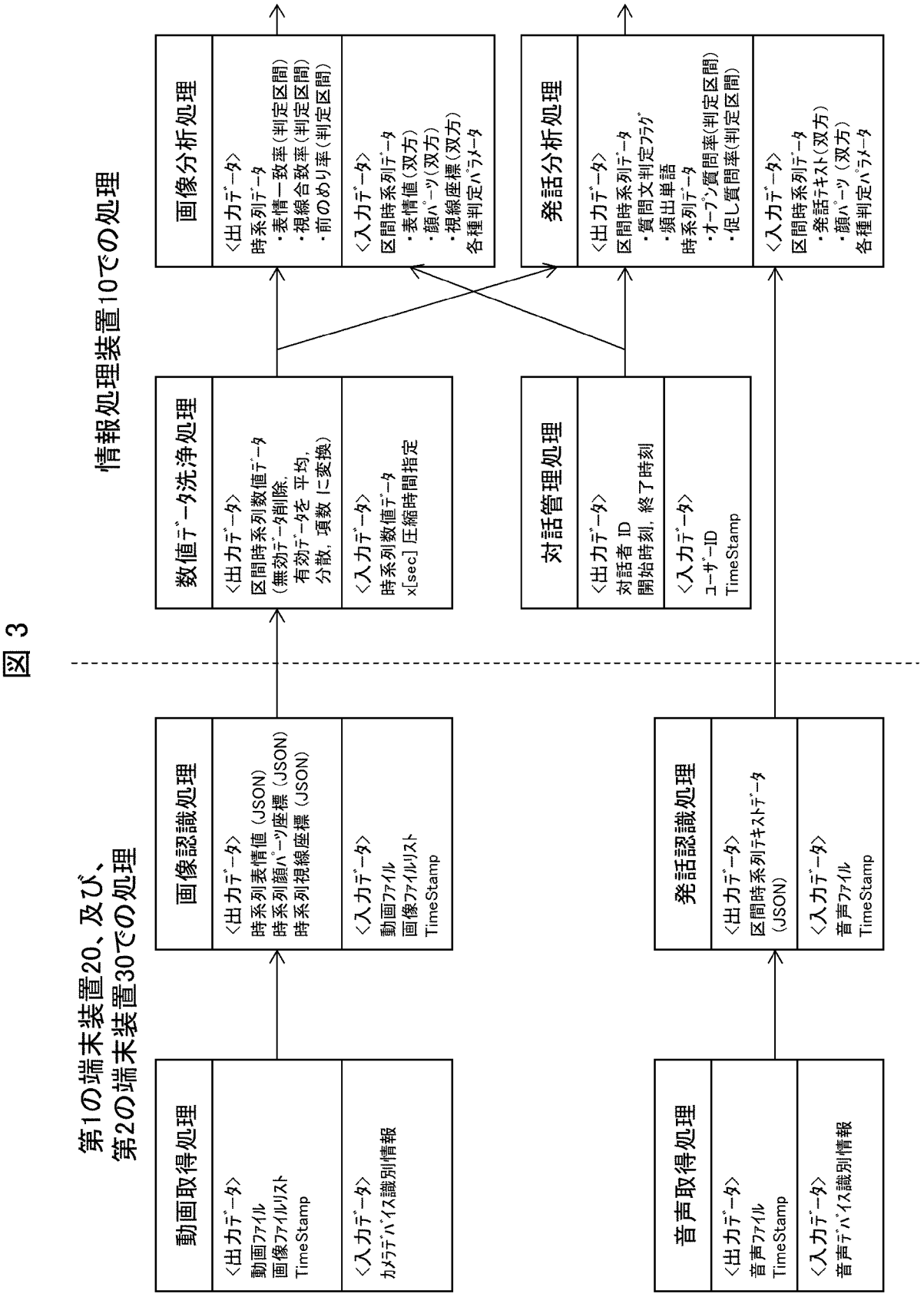
[図2]

図 2



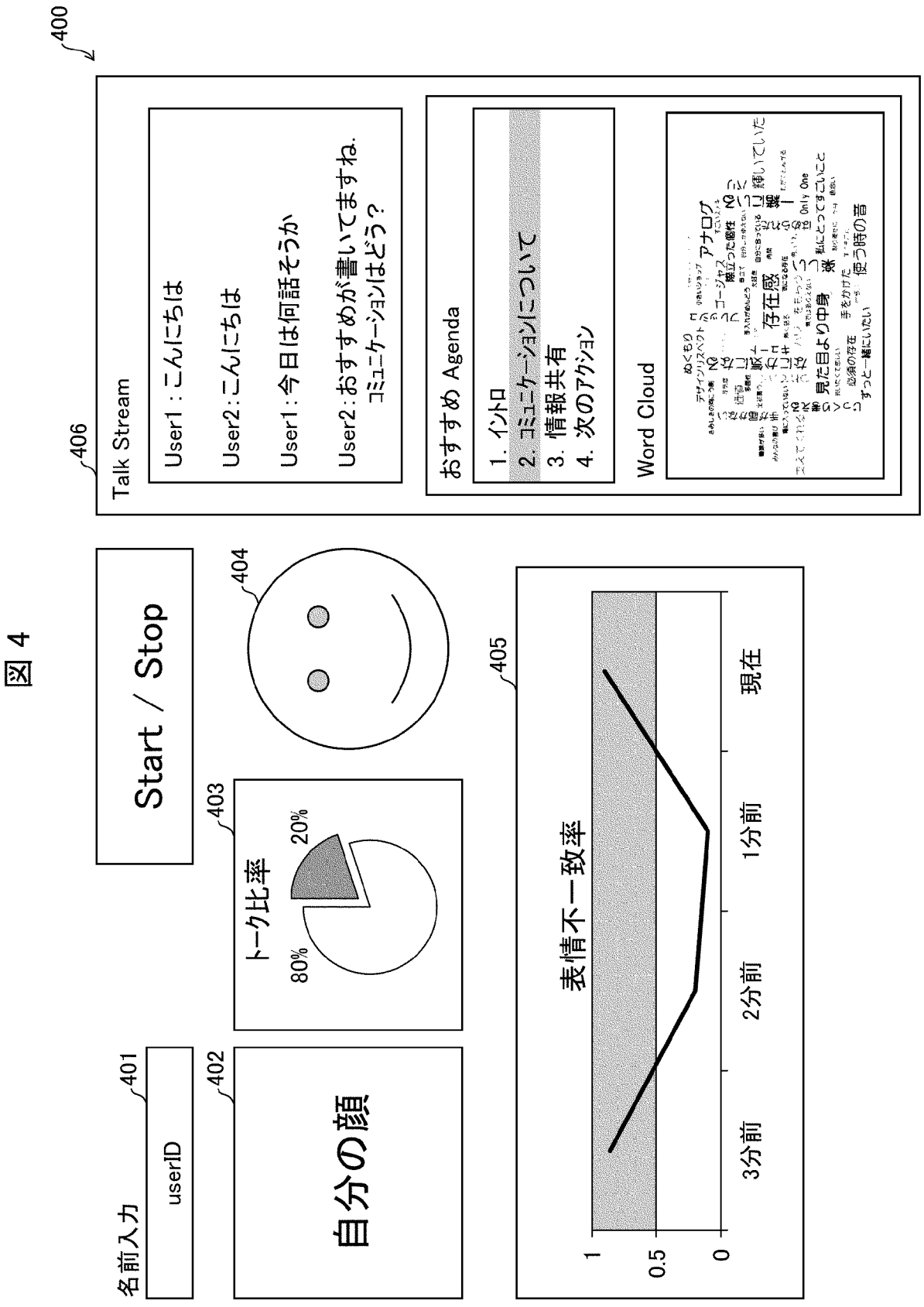
[図3]

図 3


第1の端末装置20、及び、
第2の端末装置30での処理
情報処理装置10での処理
動画取得処理
〈出力データ〉
動画ファイル
画像ファイルリスト
TimeStamp
〈入力データ〉
カメラデバイス識別情報
画像認識処理
〈出力データ〉
時系列表情値（JSON）
時系列顔パーツ座標（JSON）
時系列視線座標（JSON）
〈入力データ〉
動画ファイル
画像ファイルリスト
TimeStamp
数値データ洗浄処理
〈出力データ〉
区間時系列数値データ
（無効データ削除，
有効データを 平均，
分散，項数 に変換）
〈入力データ〉
時系列数値データ
x[sec] 圧縮時間指定
画像分析処理
〈出力データ〉
時系列データ
・表情一致率（判定区間）
・視線合致率（判定区間）
・前のめり率（判定区間）
〈入力データ〉
区間時系列データ
・表情値（双方）
・顔パーツ（双方）
・視線座標（双方）
各種判定パラメータ
音声取得処理
〈出力データ〉
音声ファイル
TimeStamp
〈入力データ〉
音声デバイス識別情報
発話認識処理
〈出力データ〉
区間時系列テキストデータ
（JSON）
〈入力データ〉
音声ファイル
TimeStamp
対話管理処理
〈出力データ〉
対話者 ID
開始時刻，終了時刻
〈入力データ〉
ユーザーID
TimeStamp
発話分析処理
〈出力データ〉
区間時系列データ
・質問文判定フラグ
・頻出単語
時系列データ
・オープン質問率（判定区間）
・促し質問率（判定区間）
〈入力データ〉
区間時系列データ
・発話テキスト（双方）
・顔パーツ（双方）
各種判定パラメータ

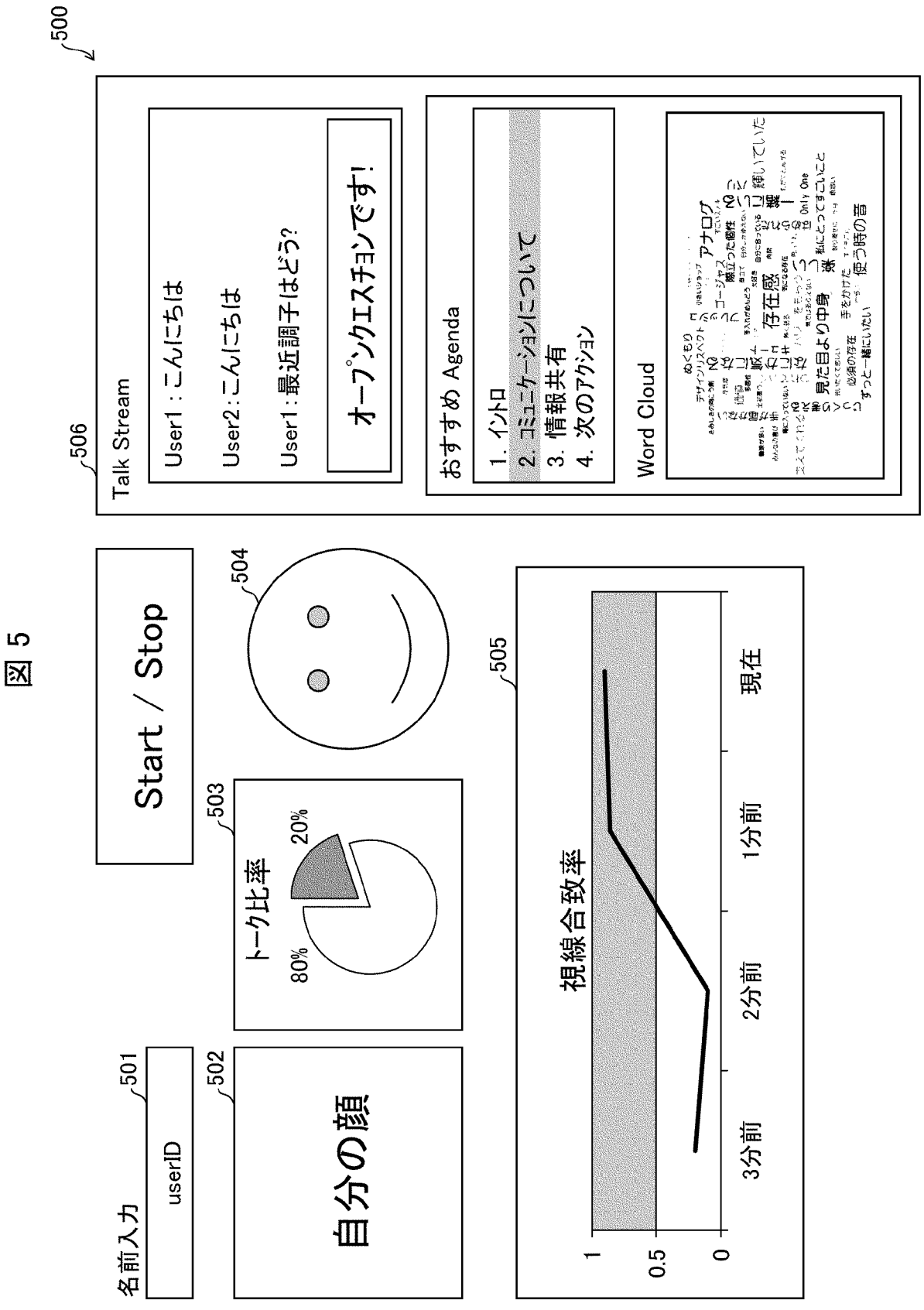
[図4]

図 4


400
406
Talk Stream
User1：こんにちは
User2：こんにちは
User1：今日は何話そうか
User2：おすすめが書いてますね.
コミュニケーションはどう？
おすすめ Agenda
1. イントロ
2. コミュニケーションについて
3. 情報共有
4. 次のアクション
Word Cloud
名前入力
401
userID
402
自分の顔
403
トーク比率
20%
80%
404
405
表情不一致率
1
0.5
0
3分前
2分前
1分前
現在
Start / Stop

[図5]

図 5


500
506
Talk Stream
User1：こんにちは
User2：こんにちは
User1：最近調子はどう?
オープンクエスチョンです!
おすすめ Agenda
1. イントロ
2. コミュニケーションについて
3. 情報共有
4. 次のアクション
Word Cloud
名前入力
501
userID
502
自分の顔
503
トーク比率
20%
80%
504
505
視線合致率
1
0.5
0
3分前
2分前
1分前
現在
Start / Stop

INTERNATIONAL SEARCH REPORT

International application No.
PCT/JP2019/009523

A. CLASSIFICATION OF SUBJECT MATTER
Int.Cl. G06F16/00(2019.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
Int.Cl. G06F16/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan	1922-1996
Published unexamined utility model applications of Japan	1971-2019
Registered utility model specifications of Japan	1996-2019
Published registered utility model applications of Japan	1994-2019

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2016-177483 A (CANON INC.) 06 October 2016, paragraphs [0010]-[0037], fig. 1, 2 (Family: none)	1-12
A	JP 2016-103081 A (KDDI CORPORATION) 02 June 2016, paragraphs [0043]-[0062], fig. 1-3 (Family: none)	1-12
A	JP 2005-284822 A (SEIKO EPSON CORP.) 13 October 2005, paragraphs [0019]-[0023], fig. 1 (Family: none)	1-12

☐ Further documents are listed in the continuation of Box C.

☐ See patent family annex.

* Special categories of cited documents:
"A" document defining the general state of the art which is not considered to be of particular relevance
"E" earlier application or patent but published on or after the international filing date
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
"O" document referring to an oral disclosure, use, exhibition or other means
"P" document published prior to the international filing date but later than the priority date claimed
"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
05 April 2019 (05.04.2019)	23 April 2019 (23.04.2019)

Name and mailing address of the ISA/	Authorized officer
Japan Patent Office 3-4-3, Kasumigaseki, Chiyoda-ku, Tokyo 100-8915, Japan	Telephone No.

国際調査報告

国際出願番号　ＰＣＴ／ＪＰ２０１９／００９５２３

Ａ．　発明の属する分野の分類（国際特許分類（ＩＰＣ））

Int.Cl.　G06F16/00(2019.01)i

Ｂ．　調査を行った分野

調査を行った最小限資料（国際特許分類（ＩＰＣ））

Int.Cl.　G06F16/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	１９２２－１９９６年
日本国公開実用新案公報	１９７１－２０１９年
日本国実用新案登録公報	１９９６－２０１９年
日本国登録実用新案公報	１９９４－２０１９年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

Ｃ．　関連すると認められる文献

引用文献のカテゴリー＊	引用文献名　及び一部の箇所が関連するときは、その関連する箇所の表示	関連する請求項の番号
A	JP 2016-177483 A（キヤノン株式会社）2016.10.06, 段落[0010]-[0037],図1,2（ファミリーなし）	1-12
A	JP 2016-103081 A（ＫＤＤＩ株式会社）2016.06.02, 段落[0043]-[0062]，図1-3（ファミリーなし）	1-12
A	JP 2005-284822 A（セイコーエプソン株式会社）2005.10.13, 段落[0019]-[0023]，図１（ファミリーなし）	1-12

☐　Ｃ欄の続きにも文献が列挙されている。

☐　パテントファミリーに関する別紙を参照。

＊　引用文献のカテゴリー
「Ａ」特に関連のある文献ではなく、一般的技術水準を示すもの
「Ｅ」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「Ｌ」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「Ｏ」口頭による開示、使用、展示等に言及する文献
「Ｐ」国際出願日前で、かつ優先権の主張の基礎となる出願の日の後に公表された文献
「Ｔ」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「Ｘ」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Ｙ」特に関連のある文献であって、当該文献と他の１以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「＆」同一パテントファミリー文献

国際調査を完了した日	国際調査報告の発送日
０５．０４．２０１９	２３．０４．２０１９

国際調査機関の名称及びあて先
日本国特許庁（ＩＳＡ／ＪＰ）
郵便番号１００－８９１５
東京都千代田区霞が関三丁目４番３号

特許庁審査官（権限のある職員）　５Ｎ　４２２６
後藤　彰

電話番号　０３－３５８１－１１０１　内線　３５８６