

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号  
特許第7323098号  
(P7323098)

(45)発行日 令和5年8月8日(2023.8.8)

(24)登録日 令和5年7月31日(2023.7.31)

(51)国際特許分類	F I	
G 0 6 Q 50/10 (2012.01)	G 0 6 Q 50/10	
G 0 6 F 3/01 (2006.01)	G 0 6 F 3/01	5 1 0
G 1 0 L 15/00 (2013.01)	G 1 0 L 15/00	2 0 0 U
G 0 6 F 3/048(2013.01)	G 0 6 F 3/048	
H 0 4 N 21/258(2011.01)	H 0 4 N 21/258	
請求項の数 11 (全22頁) 最終頁に続く		

(21)出願番号	特願2019-76447(P2019-76447)	(73)特許権者	000006013 三菱電機株式会社 東京都千代田区丸の内二丁目7番3号
(22)出願日	平成31年4月12日(2019.4.12)	(74)代理人	100084995 弁理士 加藤 和詳
(65)公開番号	特開2020-173714(P2020-173714 A)	(74)代理人	100099025 弁理士 福田 浩志
(43)公開日	令和2年10月22日(2020.10.22)	(72)発明者	黒木 友裕 千葉県印西市大塚一丁目5番地1 株式 会社竹中工務店 技術研究所内
審査請求日	令和4年2月7日(2022.2.7)	(72)発明者	高橋 幹雄 千葉県印西市大塚一丁目5番地1 株式 会社竹中工務店 技術研究所内
		(72)発明者	高 井 勇志 千葉県印西市大塚一丁目5番地1 株式 最終頁に続く

(54)【発明の名称】 対話支援装置、対話支援システム、及び対話支援プログラム

(57)【特許請求の範囲】

【請求項1】

対話の参加者の前記対話における状況を導出可能な物理量を取得する取得部と、  
前記取得部によって取得された前記物理量を用いて、前記参加者の前記対話における状  
況を導出する導出部と、

前記導出部によって導出された前記状況に対応する状況情報を表示する表示処理、及び  
前記状況情報を記憶する記憶処理の少なくとも一方の処理を行う処理部と、  
を備え、

前記状況情報は、前記参加者の感情を表す情報であり、

前記感情を表す情報は、前記感情を表す画像情報であり、

前記画像情報は、前記感情の度合いが最大となった場合における、対応する前記参加者  
の顔を撮影して得られた顔撮影画像情報である、

対話支援装置。

【請求項2】

前記画像情報は、前記顔撮影画像情報に加えて、前記感情を誇張する情報が含まれる画  
像情報である、

請求項1に記載の対話支援装置。

【請求項3】

前記物理量は、前記参加者を撮影して得られた画像、及び前記参加者の発言を示す音声  
の少なくとも一方である、

請求項 1 に記載の対話支援装置。

【請求項 4】

前記状況は、前記参加者の感情の度合いを表す物理量、及び前記参加者の動作を表す物理量の少なくとも一方である、

請求項 3 に記載の対話支援装置。

【請求項 5】

前記状況は、前記画像から得られる前記参加者のうなずきの頻度を示す物理量、前記画像から得られる前記参加者の表情の度合いを示す物理量、前記画像及び前記音声の少なくとも一方から得られる前記参加者の発言の度合いを示す物理量、の少なくとも 1 つである、

請求項 4 に記載の対話支援装置。

10

【請求項 6】

前記発言の度合いは、直近の予め定められた期間の発言速度に応じた度合いである、

請求項 5 に記載の対話支援装置。

【請求項 7】

前記うなずきの頻度を示す物理量は、当該頻度の低下率を示す情報である、

請求項 5 に記載の対話支援装置。

【請求項 8】

前記感情を表す情報を学習する、

請求項 1 に記載の対話支援装置。

【請求項 9】

前記状況情報は、前記対話の参加者全員の感情を表す情報である、

請求項 1 に記載の対話支援装置。

20

【請求項 10】

請求項 1 から請求項 9 の何れか 1 項に記載の対話支援装置と、

前記対話支援装置の前記取得部に前記対話における状況を導出可能な物理量を送信する送信部、及び前記対話支援装置の前記処理部が前記表示処理を行う場合に、当該表示処理の表示対象となる表示部、を備えた端末と、

を含む対話支援システム。

【請求項 11】

対話の参加者の前記対話における状況を導出可能な物理量を取得し、

取得した前記物理量を用いて、前記参加者の前記対話における状況を導出し、

導出した前記状況に対応する状況情報を表示する表示処理、及び前記状況情報を記憶する記憶処理の少なくとも一方の処理を行う、

処理をコンピュータが実行する対話支援プログラムであって、

前記状況情報は、前記参加者の感情を表す情報であり、

前記感情を表す情報は、前記感情を表す画像情報であり、

前記画像情報は、前記感情の度合いが最大となった場合における、対応する前記参加者の顔を撮影して得られた顔撮影画像情報である、

対話支援プログラム。

30

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、対話支援装置、対話支援システム、及び対話支援プログラムに関する。

【背景技術】

【0002】

日常的な対話において、相手の感情が把握できなかつたり、相手に意図が正確に伝わらなかつたりすることは少なくない。特に、テレビ会議システムや、スカイプ（登録商標）等を用いた遠隔地との対話においては、伝達される情報は、劣化した映像や音声となるため、より問題は顕著になる。

【0003】

50

一方、複数人で行うディスカッション等の場面では、全体を俯瞰しながら対話を適切にコントロールすることは難しい。例えば、発言が偏らない、意見のある人がきちんと発言の機会を与えられる、議論が白熱し過ぎない、といったコントロールである。

【0004】

このような問題を解決するために適用することのできる技術として、特許文献1には、人物の動画像から顔の表情を分析する表情分析手段と、人物の動画像から人物の顔の簡略画像を作成して記憶装置に蓄積する簡略画像作成手段と、表情の分析結果に従って記憶装置に蓄積されている簡略画像から対応する簡略画像を選択する簡略画像選択手段と、選択された簡略画像に対して、表情分析手段により分析した表情に応じて特殊効果を施す特殊効果処理手段とを備えた技術が開示されている。

10

【0005】

また、特許文献2には、第1及び第2の装置が接続されたネットワークシステムであって、ユーザ情報と画像データとを第1の装置で取得する取得手段と、取得した画像データに含まれる顔領域を抽出する抽出手段と、抽出手段で抽出した顔領域の顔の表情を識別する識別手段と、顔の表情毎に、顔の表情を示す表情識別情報と表示情報とを関連付けて記憶する記憶手段と、表示情報の中から、識別手段で識別した顔の表情に対応する表情識別情報と関連付けて記憶されている表示情報を特定する特定手段と、特定された表示情報と取得手段で取得したユーザ情報とを対応付けて第2の装置の表示部に表示させる表示手段とを有する技術が開示されている。

20

【先行技術文献】

【特許文献】

【0006】

【文献】特開2004-64102号公報  
特開2015-165407号公報

【発明の概要】

【発明が解決しようとする課題】

【0007】

しかしながら、特許文献1及び特許文献2の各文献に記載の技術では、対話の参加者の当該対話における全体的な状況については考慮されていないため、必ずしも効果的に対話を活性化することができるとは限らなかった。

30

【0008】

本発明は、以上の事情を鑑みて成されたものであり、対話を効果的に活性化することのできる対話支援装置、対話支援システム、及び対話支援プログラム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】

【0009】

請求項1に記載の本発明に係る対話支援装置は、対話の参加者の前記対話における状況を導出可能な物理量を取得する取得部と、前記取得部によって取得された前記物理量を用いて、前記参加者の前記対話における状況を導出する導出部と、前記導出部によって導出された前記状況に対応する状況情報を表示する表示処理、及び前記状況情報を記憶する記憶処理の少なくとも一方の処理を行う処理部と、を備え、前記状況情報は、前記参加者の感情を表す情報であり、前記感情を表す情報は、前記感情を表す画像情報であり、前記画像情報は、前記感情の度合いが最大となった場合における、対応する前記参加者の顔を撮影して得られた顔撮影画像情報であるものである。

40

【0010】

請求項1に記載の本発明に係る対話支援装置によれば、対話における状況に対応する状況情報の表示及び記憶の少なくとも一方の処理を行うことで、対話を効果的に活性化することができる。

【0012】

また、請求項1に記載の本発明に係る対話支援装置によれば、状況情報を、参加者の感情

50

を表す情報とすることで、より効果的に対話の活性化を促すことができる。

【0018】

また、請求項1に記載の本発明に係る対話支援装置によれば、感情を表す情報を、感情を表す画像情報とすることで、より直感的に参加者の感情を把握することができる。

【0022】

また、請求項1に記載の本発明に係る対話支援装置によれば、画像情報を、感情の度合いが最大となった場合における、対応する参加者の顔を撮影して得られた顔撮影画像情報とすることで、より効果的に参加者の感情を把握することができる。

【0023】

請求項2に記載の本発明に係る対話支援装置は、請求項1に記載の対話支援装置であって、前記画像情報が、前記顔撮影画像情報に加えて、前記感情を誇張する情報が含まれる画像情報とされている。

10

【0024】

請求項2に記載の本発明に係る対話支援装置によれば、画像情報を、顔撮影画像情報に加えて、感情を誇張する情報が含まれる画像情報とすることで、より効果的に参加者の感情を把握することができる。

【0027】

請求項3に記載の本発明に係る対話支援装置は、請求項1に記載の対話支援装置であって、前記物理量が、前記参加者を撮影して得られた画像、及び前記参加者の発言を示す音声の少なくとも一方とされている。

20

【0028】

請求項3に記載の本発明に係る対話支援装置によれば、物理量を、参加者を撮影して得られた画像、及び参加者の発言を示す音声の少なくとも一方とすることで、特殊な装置を用いることなく、対話の活性化を促すことができる。

【0029】

請求項4に記載の本発明に係る対話支援装置は、請求項3に記載の対話支援装置であって、前記状況が、前記参加者の感情の度合いを表す物理量、及び前記参加者の動作を表す物理量の少なくとも一方とされている。

【0030】

請求項4に記載の本発明に係る対話支援装置によれば、状況を、参加者の感情の度合いを表す物理量、及び参加者の動作を表す物理量の少なくとも一方とすることで、より効果的に対話の活性化を促すことができる。

30

【0031】

請求項5に記載の本発明に係る対話支援装置は、請求項4に記載の対話支援装置であって、前記状況が、前記画像から得られる前記参加者のうなずきの頻度を示す物理量、前記画像から得られる前記参加者の表情の度合いを示す物理量、前記画像及び前記音声の少なくとも一方から得られる前記参加者の発言の度合いを示す物理量、の少なくとも1つとされている。

【0032】

請求項5に記載の本発明に係る対話支援装置によれば、状況を、画像から得られる参加者のうなずきの頻度を示す物理量、画像から得られる参加者の表情の度合いを示す物理量、画像及び音声の少なくとも一方から得られる参加者の発言の度合いを示す物理量、の少なくとも1つとすることで、より簡易に対話の活性化を促すことができる。

40

請求項6に記載の本発明に係る対話支援装置は、請求項5に記載の対話支援装置であって、前記発言の度合いが、直近の予め定められた期間の発言速度に応じた度合いであるものである。

請求項7に記載の本発明に係る対話支援装置は、請求項5に記載の対話支援装置であって、前記うなずきの頻度を示す物理量が、当該頻度の低下率を示す情報であるものである。

請求項8に記載の本発明に係る対話支援装置は、請求項1に記載の対話支援装置であって、前記感情を表す情報を学習するものである。

50

請求項 9 に記載の本発明に係る対話支援装置は、請求項 1 に記載の対話支援装置であって、前記状況情報が、前記対話の参加者全員の感情を表す情報であるものである。

【 0 0 3 3 】

請求項 1 0 に記載の本発明に係る対話支援システムは、請求項 1 から請求項 9 の何れか 1 項に記載の対話支援装置と、前記対話支援装置の前記取得部に前記対話における状況を導出可能な物理量を送信する送信部、及び前記対話支援装置の前記処理部が前記表示処理を行う場合に、当該表示処理の表示対象となる表示部、を備えた端末と、を含む。

【 0 0 3 4 】

請求項 1 0 に記載の本発明に係る対話支援システムによれば、対話における状況に対応する状況情報の表示及び記憶の少なくとも一方の処理を行うことで、対話を効果的に活性化することができる。

10

【 0 0 3 5 】

請求項 1 1 に記載の本発明に係る対話支援プログラムは、対話の参加者の前記対話における状況を導出可能な物理量を取得し、取得した前記物理量を用いて、前記参加者の前記対話における状況を導出し、導出した前記状況に対応する状況情報を表示する表示処理、及び前記状況情報を記憶する記憶処理の少なくとも一方の処理を行う、処理をコンピュータが実行する対話支援プログラムであって、前記状況情報は、前記参加者の感情を表す情報であり、前記感情を表す情報は、前記感情を表す画像情報であり、前記画像情報は、前記感情の度合いが最大となった場合における、対応する前記参加者の顔を撮影して得られた顔撮影画像情報であるものである。

20

【 0 0 3 6 】

請求項 1 1 に記載の本発明に係る対話支援プログラムによれば、対話における状況に対応する状況情報の表示及び記憶の少なくとも一方の処理を行うことで、対話を効果的に活性化することができる。

【 発明の効果 】

【 0 0 3 7 】

以上説明したように、本発明によれば、対話を効果的に活性化することができる。

【 図面の簡単な説明 】

【 0 0 3 8 】

【 図 1 】 実施形態に係る対話支援システムのハードウェア構成の一例を示すブロック図である。

30

【 図 2 】 実施形態に係る対話支援システムの機能的な構成の一例を示すブロック図である。

【 図 3 】 実施形態に係る発言度の説明に供するタイムチャートである。

【 図 4 】 実施形態に係る状況対応情報データベースの構成の一例を示す模式図である。

【 図 5 】 実施形態に係る対応情報の学習方法の説明に供する模式図である。

【 図 6 】 実施形態に係る対話情報データベースの構成の一例を示す模式図である。

【 図 7 】 実施形態に係る顔文字・誇張情報データベースの構成の一例を示す模式図である。

【 図 8 】 実施形態に係る対話支援処理の一例を示すフローチャートである。

【 図 9 】 実施形態に係る対話支援画像の構成の一例を示す正面図である。

【 図 1 0 】 実施形態に係る派閥情報表示処理の一例を示すフローチャートである。

40

【 図 1 1 】 実施形態に係る派閥関係表示画像の構成の一例を示す正面図である。

【 図 1 2 】 実施形態に係る対応情報の他の決定方法の説明に供するタイムチャートである。

【 発明を実施するための形態 】

【 0 0 3 9 】

以下、図面を参照して、本発明を実施するための形態例を詳細に説明する。なお、本実施形態では、本発明を、複数人で会議を行う場合における対話（会議での発言）を統括的に支援する対話支援装置と、各々対話の参加者が個別に用いる複数の端末と、を含む対話支援システムに適用した場合について説明する。また、本実施形態では、対話の各参加者が互いに異なる遠隔地に分散している場合について説明する。

【 0 0 4 0 】

50

まず、図 1 及び図 2 を参照して、本実施形態に係る対話支援システム 90 の構成を説明する。図 1 に示すように、本実施形態に係る対話支援システム 90 は、ネットワーク 80 に各々アクセス可能とされた、対話支援装置 10 と、複数の端末 20 と、を含む。なお、対話支援装置 10 の例としては、パーソナルコンピュータ及びサーバコンピュータ等の情報処理装置が挙げられる。また、端末 20 の例としては、据え置き型やノートブック型等のパーソナルコンピュータや、スマートフォン、タブレット端末等の携帯型の端末が挙げられる。

#### 【0041】

本実施形態に係る端末 20 は、対話支援システム 90 を用いた会議での対話の参加者（以下、単に「参加者」という。）に各々割り当てられた端末である。端末 20 は、CPU（Central Processing Unit）21、一時記憶領域としてのメモリ 22、不揮発性の記憶部 23、タッチパネル等の入力部 24、液晶ディスプレイ等の表示部 25 及び媒体読み書き装置（R/W）26 を備えている。また、端末 20 は、カメラ 28、マイク 29 及び無線通信部 27 を備えている。CPU 21、メモリ 22、記憶部 23、入力部 24、表示部 25、媒体読み書き装置 26、カメラ 28、マイク 29 及び無線通信部 27 はバス B1 を介して互いに接続されている。媒体読み書き装置 26 は、記録媒体 96 に書き込まれている情報の読み出し及び記録媒体 96 への情報の書き込みを行う。

10

#### 【0042】

記憶部 23 は、HDD（Hard Disk Drive）、SSD（Solid State Drive）、フラッシュメモリ等によって実現される。なお、本実施形態に係る対話支援システム 90 では、各端末 20 のカメラ 28 の画角内に端末 20 を用いる参加者の顔が収まり、かつ、各端末 20 のマイク 29 による集音範囲内に端末 20 を用いる参加者の発言が入るように、各端末 20 が位置決めされている。

20

#### 【0043】

一方、対話支援装置 10 は、対話支援システム 90 で取り扱う各種情報を統括的に保管して管理する装置である。対話支援装置 10 は、CPU 11、一時記憶領域としてのメモリ 12、不揮発性の記憶部 13、キーボードとマウス等の入力部 14、液晶ディスプレイ等の表示部 15、媒体読み書き装置 16 及び通信インタフェース（I/F）部 18 を備えている。CPU 11、メモリ 12、記憶部 13、入力部 14、表示部 15、媒体読み書き装置 16 及び通信 I/F 部 18 はバス B2 を介して互いに接続されている。媒体読み書き装置 16 は、記録媒体 17 に書き込まれている情報の読み出し及び記録媒体 17 への情報の書き込みを行う。

30

#### 【0044】

記憶部 13 は HDD、SSD、フラッシュメモリ等によって実現される。記憶媒体としての記憶部 13 には、対話支援プログラム 13A が記憶されている。対話支援プログラム 13A は、対話支援プログラム 13A が書き込まれた記録媒体 17 が媒体読み書き装置 16 にセットされ、媒体読み書き装置 16 が記録媒体 17 からの対話支援プログラム 13A の読み出しを行うことで、記憶部 13 へ記憶される。CPU 11 は、対話支援プログラム 13A を記憶部 13 から読み出してメモリ 12 に展開し、対話支援プログラム 13A が有するプロセスを順次実行する。

40

#### 【0045】

また、記憶部 13 には、状況対応情報データベース 13B、対話情報データベース 13C 及び顔文字・誇張情報データベース 13D が記憶される。状況対応情報データベース 13B、対話情報データベース 13C 及び顔文字・誇張情報データベース 13D については、詳細を後述する。

#### 【0046】

次に、図 2 を参照して、本実施形態に係る対話支援装置 10 及び端末 20 の機能的な構成について説明する。図 2 に示すように、対話支援装置 10 は、取得部 11A、導出部 11B 及び処理部 11C を含む。対話支援装置 10 の CPU 11 が対話支援プログラム 13A を実行することで、取得部 11A、導出部 11B 及び処理部 11C として機能する。

50

## 【 0 0 4 7 】

本実施形態に係る取得部 1 1 A は、参加者の対話における状況を導出可能な物理量を取得する。本実施形態に係る取得部 1 1 A では、上記物理量として、参加者を撮影して得られた画像（以下、「撮影画像」という。）及び参加者の発言を示す音声（以下、「発言音声」という。）の 2 種類の物理量を適用しているが、これに限らない。例えば、撮影画像及び発言音声の何れか一方のみを上記物理量として適用する形態としてもよい。

## 【 0 0 4 8 】

また、導出部 1 1 B は、取得部 1 1 A によって取得された物理量を用いて、参加者の対話における状況を導出する。本実施形態に係る導出部 1 1 B では、上記状況として、参加者の感情の度合いを表す物理量（以下、「感情度」という。）及び参加者の動作を表す物理量（以下、「動作量」という。）の 2 種類の物理量を参加者別に導出する。より具体的には、本実施形態に係る導出部 1 1 B は、上記動作量として、撮影画像から得られる受話者の所定期間（本実施形態では、10 秒間）当たりのうなずきの回数（以下、「うなずき頻度」という。） $N$ 、及び発言音声から得られる発話者の発言の度合いを示す物理量（以下、「発言度」という。） $H$ を導出する。また、導出部 1 1 B は、上記感情度として、撮影画像から得られる参加者の表情の度合いを示す物理量（以下、「表情度」という。）を導出する。

## 【 0 0 4 9 】

より具体的に、本実施形態に係る導出部 1 1 B は、上記表情度として、対応する参加者の怒りの度合いを示す怒り度  $I$ 、対応する参加者の嫌悪の度合いを示す嫌悪度  $K$ 、及び対応する参加者の恐れ度  $O$ を導出する。また、本実施形態に係る導出部 1 1 B は、上記表情度として、対応する参加者の喜びの度合いを示す喜び度  $Y$ 、対応する参加者の悲しみの度合いを示す悲しみ度  $S$ 、及び対応する参加者の驚きの度合いを示す驚き度  $B$ を導出する。

## 【 0 0 5 0 】

なお、本実施形態では、これらの 6 種類の感情度を、対応する参加者が用いる端末 2 0 のカメラ 2 8 により得られた撮影画像に基づいて、例えば、“Real-time face detection and emotion/gender classification”、インターネット < URL : [https://github.com/oarriaga/face\\_classification](https://github.com/oarriaga/face_classification) > 等に記載の既知の技術を適用して導出する。

## 【 0 0 5 1 】

この技術では、ニューラルネットワークライブラリである *Keras* をベースとして、CNN (Convolutional Neural Network、畳み込みニューラルネットワーク) により顔の特徴を抽出し、各感情を認識する。例えば、笑顔（喜び度  $Y$ ）であれば笑顔の特徴に関するデータベースが用意されており、対象となる撮影画像から顔の要素（例えば、部分的な目、鼻、口などの形。）から類似度を判定する。本実施形態では、この類似度を感情度として適用する。また、本実施形態では、上記 6 種類の感情度を、共に共通の範囲（本実施形態では、0 から 100 までの範囲）とするように正規化した値として導出する。

## 【 0 0 5 2 】

なお、各感情度の導出は、他にもマイクロソフト社の *Azure*（登録商標）で提供されているサービスである *Emotion API* (Application Programming Interface) 等の多くの既知の技術を適用することにより可能であるため、ここでの、これ以上の説明は省略する。

## 【 0 0 5 3 】

このように、本実施形態では、上記 6 種類の感情度を適用しているが、これに限らず、上記 6 種類のうちの 1 種類、又は 5 種類以下の複数種類の組み合わせを適用する形態としてもよい。

## 【 0 0 5 4 】

一方、本実施形態に係る導出部 1 1 B は、発言度  $H$  を次の式 (1) により算出する。式 (1) における  $s(t)$  は、対象とする発話者の発言速度 (= 発言文字数 / 秒) を表す。

## 【 0 0 5 5 】

10

20

30

40

50

【数 1】

$$H = \int_0^{600} s(t)(600-t)dt \quad (1)$$

【0056】

即ち、式(1)は、直近の10分間(600秒間)の発言速度 $s(t)$ を、算出時点に近い発言ほど重み値を大きくして積算して得られる値を発言度 $H$ として算出する。本実施形態では、発話者の発言速度 $s(t)$ を導出する際に用いる発言文字数として、対応する発言音声を、既知の音声認識技術によって認識し、これによって得られたテキスト情報の文字数を適用するが、これに限るものではない。

10

【0057】

例えば、通常、会議の場における各参加者の発言は、一例として図3に示すように、他者の発言の間に纏めて行われるが、本実施形態では、算出時点に近いタイミングでの発言速度 $s(t)$ ほど重視するものとしている。これにより、発言度 $H$ を、対応する発話者の対話中の話題に対する理解の高さを、よりの確に表すものとして算出できるようにしている。

【0058】

なお、発言度 $H$ を算出する数式は、式(1)には限らない。例えば、式(1)において適用した直近の10分間は一例であり、他の期間としてもよいことは言うまでもない。また、式(1)では、算出時点に近い発言ほど重み値を大きくしているが、この重み付けを行うことなく発言度 $H$ を算出する形態としてもよい。また、本実施形態では、発言度 $H$ の導出に数式を用いる場合について説明したが、この形態に限らず、例えば、テーブル変換により発言度 $H$ を導出する形態としてもよい。更に、本実施形態では、発言度 $H$ を、発言音声をを用いて導出しているが、これに限らない。例えば、撮影画像を用いて、各参加者の口の動きから発言速度 $s(t)$ を導出し、この発言速度 $s(t)$ を式(1)に代入することによって発言度 $H$ を算出する形態等としてもよい。

20

【0059】

また、本実施形態に係る導出部11Bは、うなずき頻度 $N$ を次の式(2)により算出する。式(2)における $n(t)$ は、対象とする受話者の所定時間当たりのうなずき回数(=うなずき回数/秒)を表す。

30

【0060】

【数 2】

$$N = \int_0^{600} n(t)(600-t)dt \quad (2)$$

【0061】

即ち、式(2)は、直近の10分間(600秒間)の所定時間当たりのうなずき回数 $n(t)$ を、算出時点に近いうなずきほど重み値を大きくして積算して得られる値をうなずき頻度 $N$ として算出する。本実施形態では、撮影画像に含まれる受話者の顔画像が、当該受話者から見て前方で、かつ、下方に傾斜したことに引き続いてほぼ元の位置に復帰した場合に、1回うなずいたと判断している。なお、本実施形態では、顔画像の傾斜及び復帰の検出を、顔画像の所定部位の画像(本実施形態では、目の画像)が下方に所定距離(本実施形態では、3mm)以上移動した後に、ほぼ元の位置に戻ったことを検出することにより行っているが、これに限るものではないことは言うまでもない。

40

【0062】

なお、うなずき頻度 $N$ を算出する数式は、式(2)には限らない。例えば、式(2)において適用した直近の10分間は一例であり、他の期間としてもよいことは言うまでもない。また、式(2)では、算出時点に近いうなずき回数 $n(t)$ ほど重み値を大きくしているが、この重み付けを行うことなくうなずき頻度 $N$ を算出する形態としてもよい。また、本実施形態では、うなずき頻度 $N$ の導出に数式を用いる場合について説明したが、この

50

形態に限らず、例えば、テーブル変換によりうなずき頻度  $N$  を導出する形態としてもよい。

【0063】

そして、処理部 11C は、導出部 11B によって導出された上記状況に対応する状況情報を端末 20 の表示部 25 に表示する表示処理、及び上記状況情報を記憶部 13 に記憶する記憶処理の双方の処理を行う。但し、この形態に限らず、上記表示処理及び上記記憶処理の何れか一方の処理を行う形態としてもよい。

【0064】

本実施形態では、上記状況情報として、対応する参加者の感情を表す情報を適用している。より具体的には、本実施形態では、上記感情を表す情報として、当該感情を表すテキスト情報、画像情報、及び各参加者の相互間における感情の関係を示す情報を適用している。

10

【0065】

なお、本実施形態では、上記テキスト情報として、対応する参加者のうちの何れかの発言者による発言に対する他者の感情を表す情報を適用している。また、本実施形態では、上記画像情報として、顔文字を適用している。また、本実施形態では、上記画像情報として、感情の度合いが最大となった場合における、対応する参加者の顔を撮影して得られた顔撮影画像情報及び感情を誇張する情報が含まれる画像情報も適用している。

【0066】

一方、本実施形態に係る端末 20 は、制御部 21A を含む。端末 20 の CPU 21 が、記憶部 23 に予め記憶された図示しない対話支援アプリケーション・プログラムを実行することで、制御部 21A として機能する。

20

【0067】

本実施形態に係る制御部 21A は、送信部としての無線通信部 27 を介して、対話支援装置 10 の取得部 11A に、上記対話における状況を導出可能な物理量を送信する。また、制御部 21A は、対話支援装置 10 の処理部 11C が上記表示処理を行う場合に、当該表示処理の表示対象となる表示部 25 を制御する。

【0068】

次に、図 4 を参照して、本実施形態に係る状況対応情報データベース 13B について説明する。図 4 に示すように、本実施形態に係る状況対応情報データベース 13B は、状況を示す情報と、対応する状況において、対応する参加者の感情を示すものとして当該参加者に対応付けて表示するテキスト情報である対応情報とが関連付けられて記憶されている。上記対応情報が、本発明の感情を表すテキスト情報に相当する。

30

【0069】

ここで、上記状況を示す情報には、図 4 に示すように、発話者による発言度  $H$  及び受話者による 6 種類の感情度に加えて、受話者によるうなずき頻度  $N$  の所定時間（本実施形態では、60 秒）前からの低下率を示す、うなずき頻度低下率  $U$  が含まれる。

【0070】

また、対応する状況に対応する上記対応情報は、一例として以下のように導出する。即ち、まず、一例として図 5 に示すように、会議の場で想定される「IF (状況) THEN (対応情報)」を予め仮説として多数用意する。図 5 に示す例では、発話者の発言度  $H$  が 180 以上であり、かつ、受話者のうなずき頻度低下率  $U$  が 50% 以上であり、かつ、受話者の怒り度  $I$  が 50 以上である状況の場合、受話者の感情を示す対応情報として「ちょっと話についていけないなあ」を仮説としている。

40

【0071】

そして、本実施形態では、仮説として用意した多数の状況及び対応情報の組み合わせを実際の会議の場で適用して、状況の条件を満足する対応情報を端末 20 に表示させ、当該表示が有効であったか否かを繰り返し評価することにより学習する。そして、この学習によって得られた対応情報を状況対応情報データベース 13B に反映させる。なお、ここで行う評価は、受話者の主観による評価でもよいし、対応情報を表示した後の実際の改善効果（例えば、笑顔が増える、発言が増える等）といった客観的な評価でもよい。

50

## 【 0 0 7 2 】

このように、本実施形態では、状況に対応する対応情報を学習させているが、必ずしも学習を行う必要はなく、予め仮説として用意した状況及び対応情報そのものを状況対応情報データベース 1 3 B に選択的に適用する形態としてもよい。

## 【 0 0 7 3 】

次に、図 6 を参照して、本実施形態に係る対話情報データベース 1 3 C について説明する。図 6 に示すように、本実施形態に係る対話情報データベース 1 3 C は、端末 I D ( I D e n t i f i c a t i o n )、画像データ、音声データ、テキストデータ、時刻、及び最大感情度の各情報が関連付けられて記憶される。

## 【 0 0 7 4 】

上記端末 I D は、各参加者が用いる端末 2 0 を識別するために割り振られた情報である。なお、本実施形態では、端末 I D と、当該端末 I D が割り振られた端末 2 0 を用いる参加者を示す情報（本実施形態では、名前）とが対応付けられて記憶部 1 3 に記憶されている。従って、対話支援装置 1 0 は、何れかの端末 2 0 との間で通信を行う場合に、当該端末 2 0 を用いる参加者を特定することができる。

## 【 0 0 7 5 】

また、上記画像データは、対応する端末 2 0 から取得された撮影画像を示す情報であり、上記音声データは、対応する端末 2 0 から取得された発言音声を示す情報であり、上記テキストデータは、対応する発言音声をテキスト化した情報である。なお、本実施形態では、上記テキストデータを、対応する音声データを、既知の音声認識技術を用いてテキストデータに変換することで得ている。

## 【 0 0 7 6 】

また、上記時刻は、対応する画像データ及び音声データが取得された日時を示す情報であり、上記最大感情度は、対応する参加者の、対応する画像データが得られている期間内における最大値となる感情度の種類を示す情報である。

## 【 0 0 7 7 】

なお、本実施形態では、図 6 に示すように、最大感情度における各参加者を示す情報として、当該参加者が用いる端末 2 0 の端末 I D を適用しているが、これに限らないことは言うまでもない。また、図 6 では、最大感情度の種類を符号のみで表しているが、例えば、‘ I ’ は怒り度 I を表し、‘ O ’ は恐れ度 O を表している。更に、図 6 では、最大感情度を発 30 話者のみについて対話情報データベース 1 3 C に記憶している場合を例示しているが、これに限らず、対応する期間における受話者の最大感情度も対話情報データベース 1 3 C に記憶する形態としてもよい。

## 【 0 0 7 8 】

次に、図 7 を参照して、本実施形態に係る顔文字・誇張情報データベース 1 3 D について説明する。図 7 に示すように、本実施形態に係る顔文字・誇張情報データベース 1 3 D は、最大感情度、顔文字、及び誇張情報の各情報が関連付けられて記憶されている。

## 【 0 0 7 9 】

上記最大感情度は上述した対話情報データベース 1 3 C の最大感情度と同一の情報であり、上記顔文字は、対応する最大感情度に対応する顔文字を示すデータであり、上記誇張情報は、対応する最大感情度に対応する誇張の内容を示す情報である。

## 【 0 0 8 0 】

例えば、図 7 に示す顔文字・誇張情報データベース 1 3 D では、最大感情度となる感情度の種類が恐れ度 O である場合に対応する顔文字が「(^\_^;)」であることを示している。また、図 7 に示す例では、最大感情度となる感情度の種類が恐れ度 O である場合に対応する誇張情報が示す誇張の内容が、「ガーン」とのテキスト情報、及び恐れを示す画像であることを示している。なお、上記恐れを示す画像は、例えば、後述する図 9 に示す、対応する参加者の顔画像の額付近に複数の縦線が重畳された画像 2 5 G 等が例示される。

## 【 0 0 8 1 】

次に、図 8 ~ 図 1 1 を参照して、本実施形態に係る対話支援システム 9 0 の作用を説明

10

20

40

50

する。まず、図 8 及び図 9 を参照して、対話支援処理を実行する場合の対話支援装置 10 の作用を説明する。会議の各参加者が用いる端末 20 によって上述した対話支援アプリケーション・プログラムの実行が開始されることに応じて、対話支援装置 10 の CPU 11 が対話支援プログラム 13 A を実行することにより、図 8 に示す対話支援処理が実行される。なお、ここでは、錯綜を回避するために、複数の参加者による対話が時間的に重複することなく進められる場合について説明する。また、ここでは、錯綜を回避するために、状況対応情報データベース 13 B 及び顔文字・誇張情報データベース 13 D が構築済みである場合について説明する。

#### 【 0 0 8 2 】

対話支援アプリケーション・プログラムの実行が開始されると、各参加者が用いる端末 20 は、自身のカメラ 28 による撮影及びマイク 29 の作動を開始し、これによって得られた撮影画像を示す画像データ及び発言音声を示す音声データの対話支援装置 10 への送信を開始する。

10

#### 【 0 0 8 3 】

そこで、図 8 のステップ 200 で、取得部 11 A は、各端末 20 から送信された画像データ及び音声データの受信、及び受信した各データの記憶部 13 への記憶を開始する。なお、取得部 11 A は、受信した各データを記憶部 13 に記憶する際に、対応するデータの送信元の端末 20 に割り振られた端末 ID 及び取得した時点の時刻を関連付けて記憶する。

#### 【 0 0 8 4 】

ステップ 202 で、取得部 11 A は、各端末 20 から受信している音声データによる発言音声が所定期間（本実施形態では、5 秒間）途切れるまで待機することにより、対話の各参加者の一連の発言（以下、「一連発言」という。）が終了するまで待機する。

20

#### 【 0 0 8 5 】

ステップ 204 で、導出部 11 B は、各参加者の直近の一連発言分の画像データ及び音声データを記憶部 13 から読み出す。ステップ 206 で、導出部 11 B は、読み出した音声データを既知の音声認識技術を用いて各参加者別にテキストデータに変換する。

#### 【 0 0 8 6 】

ステップ 208 で、導出部 11 B は、読み出した画像データを用いて、各参加者別に上記 6 種類の感情度（本実施形態では、怒り度 I、嫌悪度 K、恐れ度 O、喜び度 Y、悲しみ度 S、驚き度 B）を上述したように導出する。なお、本実施形態では、感情度を、参加者毎で、かつ、感情度毎に、読み出した直近の一連発言分の画像データにおける最大値を導出する。但し、この形態に限らず、例えば、読み出した直近の一連発言分の画像データにおける時系列順の中央の画像データを用いて導出する形態や、読み出した直近の一連発言分の画像データにおける時系列順の最後の画像データを用いて導出する形態等を適用してもよい。

30

#### 【 0 0 8 7 】

ステップ 210 で、導出部 11 B は、各参加者別の 6 種類の感情度のうち、最大値となった感情度（最大感情度）の導出対象の時点に対応する画像データ（静止画像データ）を各参加者別に特定する。ステップ 212 で、処理部 11 C は、ステップ 210 の処理によって特定した静止画像データが示す撮影画像、及びステップ 206 の処理によって得られたテキストデータを用いて、対話を支援するための画像（以下、「対話支援画像」という。）を構成する。この際、処理部 11 C は、一例として図 9 に示すように、対応する参加者の撮影画像 25 C に対して、テキストデータにより示されるテキスト 25 D を、所謂吹き出しの形態で表示されるように対話支援画像 30 を構成する。

40

#### 【 0 0 8 8 】

ステップ 214 で、導出部 11 B は、ステップ 206 の処理によって得られたテキストデータを用いて、上述したように、式（1）を用いて各参加者別の発言度 H を算出する。ステップ 216 で、導出部 11 B は、ステップ 204 の処理によって読み出した画像データを用いて、上述したように、うなずき頻度低下率 U を算出する。この際、読み出した画像データでは、うなずき頻度低下率 U を算出する際に適用する所定時間（本実施形態では

50

、60秒)前のうなずき頻度N(以下、「起算頻度」という。)が得られない場合がある。この場合、本実施形態では、起算頻度として、各参加者別の過去のうなずき頻度Nの平均値を適用する。但し、この形態に限らず、例えば、各参加者別の直近のうなずき頻度Nを起算頻度として適用する形態等としてもよい。

【0089】

ステップ218で、処理部11Cは、ステップ210の処理において用いた各参加者別の最大感情度に、顔文字・誇張情報データベース13Dに顔文字が登録されている最大感情度が含まれるか否かを判定し、肯定判定となった場合はステップ220に移行する。

【0090】

ステップ220で、処理部11Cは、ステップ218の処理において含まれると判定された最大感情度に対応する顔文字を顔文字・誇張情報データベース13Dから読み出す。ステップ222で、処理部11Cは、一例として図9に示すように、読み出した顔文字25Eが、対応する参加者に対応するテキスト25Dに含めて吹き出し内に表示されるように対話支援画像30を更新し、その後にステップ224に移行する。

10

【0091】

一方、ステップ218において否定判定となった場合は、ステップ220及びステップ222の処理を実行することなくステップ224に移行する。

【0092】

ステップ224で、処理部11Cは、ステップ210の処理において用いた各参加者別の最大感情度に、顔文字・誇張情報データベース13Dに誇張情報が登録されている最大感情度が含まれるか否かを判定し、肯定判定となった場合はステップ226に移行する。

20

【0093】

ステップ226で、処理部11Cは、ステップ224の処理において含まれると判定された最大感情度に対応する誇張情報を顔文字・誇張情報データベース13Dから読み出す。ステップ228で、処理部11Cは、一例として図9に示すように、読み出した誇張情報が示す情報を、対応する参加者に対応されて表示されるように対話支援画像30を更新し、その後にステップ230に移行する。なお、図9に示す対話支援画像30の例では、上記誇張情報が示す情報として、対応する参加者の撮影画像の上部に「ガーン」とのテキスト25Fが表示され、対応する参加者の撮影画像における顔の額付近に複数の縦線が重畳された画像25Gが表示される。

30

【0094】

一方、ステップ224において否定判定となった場合は、ステップ226及びステップ228の処理を実行することなくステップ230に移行する。

【0095】

ステップ230で、処理部11Cは、以上の処理によって算出した発言度H、うなずき頻度低下率U、及び感情度の各参加者別の組み合わせに合致する条件が状況対応情報データベース13Bに含まれるか否かを判定し、肯定判定となった場合はステップ232に移行する。

【0096】

ステップ232で、処理部11Cは、ステップ230の処理において含まれると判定された条件に対応する対応情報を状況対応情報データベース13Bから読み出す。ステップ234で、処理部11Cは、一例として図9に示すように、読み出した対応情報25Hが所定の位置(図9に示す例では、対話支援画像30の上端部近傍の位置)に表示されるように対話支援画像30を更新し、その後にステップ236に移行する。

40

【0097】

一方、ステップ230において否定判定となった場合は、ステップ232及びステップ234の処理を実行することなくステップ236に移行する。

【0098】

ステップ236で、処理部11Cは、対話を支援するための他の支援情報が表示されるように対話支援画像30を更新する。なお、本実施形態では、上記支援情報として、一例

50

として図9に示すように、各参加者の撮影画像、発言度H（図9では「発言」と表記。）  
、うなずき頻度N（図9では「肯定」と表記。）及び顔文字（図9では「気分」と表記。  
）を含む支援情報25Iが表示されるように対話支援画像30を更新する。また、本実施  
形態では、上記他の支援情報として、対応する音声の再生の指示を受け付けるための音声  
ボタン25Jが表示されるように対話支援画像30を更新する。更に、本実施形態では、  
上記他の支援情報として、表示している対話支援画像30の上下方向へのスクロールの指  
示を受け付けるためのスクロールボタン25Kが表示されるように対話支援画像30を更  
新する。なお、その他の支援情報として、図9に示すように、発話者が発言している際の  
受話者の撮影画像25Pを当該発話者における各吹き出しの近傍に並べて表示する形態と  
してもよい。

10

【0099】

ステップ238で、処理部11Cは、以上の処理によって得られた各種情報を対話情報  
データベース13Cに登録（記憶）する。このステップ238の処理により、対話情報デ  
ータベース13Cが逐次構築されることになる。

【0100】

ステップ240で、処理部11Cは、以上の処理によって得られた対話支援画像30を  
示す画像情報を各端末20に送信するように通信I/F部18を制御する。この処理によ  
り、上述した対話支援アプリケーション・プログラムにより、一例として図9に示す対話  
支援画像30が各端末20の表示部25に表示される。各参加者は、自身が用いる端末2  
0に表示された対話支援画像30を参照し、音声を再生させたい場合は対応する音声ボタ  
ン25Jを指定し、対話支援画像30を上下方向にスクロールさせたい場合はスクロール  
ボタン25Kを所望の方向に移動させる。これに応じて、各端末20で実行されている対  
話支援アプリケーション・プログラムは、参加者によって音声ボタン25J及びスクロ  
ールボタン25Kの少なくとも一方が操作された場合に、操作された状態を示す状態情報  
を対話支援装置10に送信する。

20

【0101】

そこで、ステップ242で、処理部11Cは、何れかの端末20から音声ボタン25J  
が指定された旨を示す状態情報が受信されたか否かを判定し、否定判定となった場合はス  
テップ246に移行する一方、肯定判定となった場合はステップ244に移行する。

【0102】

ステップ244で、処理部11Cは、指定された音声ボタン25Jに対応する音声デー  
タを記憶部13から読み出して、対応する状態情報の送信元の端末20に送信し、その後  
にステップ246に移行する。ステップ244の処理により、音声ボタン25Jが指定され  
た旨を示す状態情報を送信した端末20では、対話支援アプリケーション・プログラム  
によって参加者が指定した音声再生される。

30

【0103】

ステップ246で、処理部11Cは、何れかの端末20からスクロールボタン25Kが  
操作された旨を示す状態情報が受信されたか否かを判定し、否定判定となった場合はス  
テップ250に移行する一方、肯定判定となった場合はステップ248に移行する。

【0104】

ステップ248で、処理部11Cは、スクロールボタン25Kが上方向に移動された場  
合には、スクロールボタン25Kの移動量に応じた量だけ対話支援画像30を上方向にス  
クロールさせるための情報を、対応する状態情報の送信元の端末20に送信する。また、  
処理部11Cは、スクロールボタン25Kが下方向に移動された場合には、スクロールボ  
タン25Kの移動量に応じた量だけ対話支援画像30を下方向にスクロールさせるための  
情報を、対応する状態情報の送信元の端末20に送信する。そして、処理部11Cは、以  
上の処理を行った後にステップ250の処理に移行する。ステップ248の処理により、  
スクロールボタン25Kが操作された状態を示す状態情報を送信した端末20では、対話  
支援アプリケーション・プログラムによって、表示部25で表示されている対話支援画  
像30が上記操作に応じてスクロールされる。

40

50

## 【 0 1 0 5 】

ステップ 2 5 0 で、処理部 1 1 C は、本対話支援処理の終了タイミングが到来したか否かを判定し、否定判定となった場合はステップ 2 0 2 に戻る一方、肯定判定となった時点でステップ 2 5 2 に移行する。なお、本実施形態では、対話支援処理の終了タイミングを、本対話支援処理が対象としている会議に参加している全ての参加者の端末 2 0 で実行されている対話支援アプリケーション・プログラムが終了されるタイミングとしているが、これに限らない。例えば、対象としている会議が所定時間（例えば、1 0 分）以上停止したタイミング、対象としている会議に予め設定された時間（例えば、1 時間）が経過したタイミング等を対話支援処理の終了タイミングとしてもよい。

## 【 0 1 0 6 】

ステップ 2 5 2 で、処理部 1 1 C は、ステップ 2 0 0 の処理によって開始した、各端末 2 0 から送信された画像データ及び音声データの受信、及び受信した各データの記憶部 1 3 への記憶を終了した後、本対話支援処理を終了する。

## 【 0 1 0 7 】

一方、本実施形態に係る対話支援システム 9 0 では、何れかの参加者が対話支援画像 3 0 における派閥情報表示ボタン 2 5 A を指定すると、各参加者の相互間における感情の関係をグラフィカルに示す情報である派閥関係表示画像を表示する派閥関係表示機能を有している。

## 【 0 1 0 8 】

次に、図 1 0 ~ 図 1 1 を参照して、派閥関係表示機能の実行時における対話支援システム 9 0 の作用を説明する。なお、図 1 0 は、対象としている会話に参加している何れかの参加者の端末 2 0 から、派閥情報表示ボタン 2 5 A が指定された旨を示す情報が受信された場合に、対話支援装置 1 0 の CPU 1 1 により実行される派閥情報表示処理の流れを示すフローチャートである。

## 【 0 1 0 9 】

図 1 0 のステップ 3 0 0 で、処理部 1 1 C は、その時点から所定時間（本実施形態では、1 0 分間）遡った時間から、その時間までに記憶した画像データを、対応する端末 ID と共に対話情報データベース 1 3 C から読み出す。ステップ 3 0 2 で、処理部 1 1 C は、読み出した画像データを用いて、予め定められた構成とされた派閥関係表示画像を構成する。ステップ 3 0 4 で、処理部 1 1 C は、構成した派閥関係表示画像を示す情報を、派閥情報表示ボタン 2 5 A が指定された旨を示す情報の送信元の端末 2 0 に送信する。派閥関係表示画像を示す情報を受信した端末 2 0 では、一例として図 1 1 に示す派閥関係表示画像 3 2 を表示部 2 5 に表示する。図 1 1 に示すように、本実施形態に係る派閥関係表示画像 3 2 では、対象としている会議の参加者間で相互に抱いている感情がグラフィカルに表示される。

## 【 0 1 1 0 】

なお、本実施形態では、各参加者間で相互に抱いている感情を示す情報として、次の式 ( 3 ) で算出される相互近接度  $SK_{xy}$  を適用している。なお、式 ( 3 ) における  $x$  及び  $y$  は各々異なる参加者を表し、 $n_x$  は参加者  $y$  が発言している際の参加者  $x$  のうなずき回数を表し、 $n_y$  は参加者  $x$  が発言している際の参加者  $y$  のうなずき回数を表す。ここで、うなずき回数  $n_x$  及びうなずき回数  $n_y$  は、読み出した画像データが示す撮影画像を用いて、上述した式 ( 2 ) に用いるうなずき回数  $n(t)$  と同様に導出する。

## 【 0 1 1 1 】

## 【 数 3 】

$$SK_{xy} = n_x + n_y \quad (3)$$

## 【 0 1 1 2 】

そして、本実施形態に係る派閥関係表示画像 3 2 では、算出した相互近接度  $SK_{xy}$  の逆数を離間距離として各参加者の撮影画像を配置する。この際、うなずき回数  $n_x$  とうな

10

20

30

40

50

ずき回数  $n_y$  との差分が所定値より大きな場合、一例として図 1 1 に示すように、うなずき回数が少ない方から多い方に向けて矢印を表示し、かつ、敵対視を示す画像 2 5 L を表示可能とする。また、この場合、うなずき回数が多い方から少ない方に向けて矢印を表示し、かつ、好感を示す画像 2 5 M を表示可能とする。また、上記離間距離が所定距離未満である場合、対応する参加者間を他よりも太い直線で結び、かつ、直線上に好感を示す画像 2 5 M を表示可能とする。更に、上記離間距離が上記所定距離以上である場合、対応する参加者間を直線で結び、かつ、直線上に衝突感を示す画像 2 5 N を表示可能とする。

【 0 1 1 3 】

図 1 1 に示す例では、例えば、Aさんと、他の参加者（Bさん、Cさん、Dさん）との間は相互に衝突感を抱いていることを示しており、また、例えば、CさんはDさんに対して敵対視しているが、DさんはCさんに対して好感を抱いていることを示している。更に、図 1 1 に示す例では、BさんとDさんとは互いに好感を抱いていることを示している。

10

【 0 1 1 4 】

この派閥関係表示画像 3 2 を参照することにより、対話の参加者は、他者が自分に抱いている感情の推定結果を把握することができるため、その場に応じた、より効果的な発言を行ったり、態度をとったりすることができる。また、派閥関係表示画像 3 2 を参照することにより、対話の参加者は、自身の他者に対する感情の推定結果が妥当か否かを判断することができるため、推定結果が誤っている場合に是正することが可能となる。

【 0 1 1 5 】

派閥関係表示画像 3 2 が表示部 2 5 に表示されると、参加者は、当該派閥関係表示画像 3 2 を参照した後、終了ボタン 2 5 B を指定する。これに応じて、対応する端末 2 0 の制御部 2 1 A は、派閥関係表示画像 3 2 の表示を終了する旨を示す情報（以下、「表示終了情報」という。）を対話支援装置 1 0 に無線通信部 2 7 を介して送信する。

20

【 0 1 1 6 】

そこで、次のステップ 3 0 6 で、処理部 1 1 C は、表示終了情報が指定されるまで待機した後、本派閥情報表示処理を終了する。

【 0 1 1 7 】

以上説明したように、本実施形態によれば、対話の参加者の対話における状況を導出可能な物理量を取得する取得部 1 1 A と、取得部 1 1 A によって取得された物理量を用いて、参加者の対話における状況を導出する導出部 1 1 B と、導出部 1 1 B によって導出された状況に対応する状況情報を表示する表示処理、及び状況情報を記憶する記憶処理の双方の処理を行う処理部 1 1 C と、を備えている。従って、対話を効果的に活性化することができる。

30

【 0 1 1 8 】

また、本実施形態によれば、上記状況情報を、参加者の感情を表す情報としている。従って、より効果的に対話の活性化を促すことができる。

【 0 1 1 9 】

また、本実施形態によれば、上記感情を表す情報を、感情を表すテキスト情報としている。従って、より具体的に参加者の感情を把握することができる。

【 0 1 2 0 】

また、本実施形態によれば、上記テキスト情報を、参加者のうちの何れかの発言者による発言に対する他者の感情を表す情報としている。従って、発言を聞いている参加者の感情を把握することができる。

40

【 0 1 2 1 】

また、本実施形態によれば、上記感情を表す情報を、感情を表す画像情報としている。従って、より直感的に参加者の感情を把握することができる。

【 0 1 2 2 】

また、本実施形態によれば、上記画像情報を、顔文字としている。従って、より直感的に参加者の感情を把握することができる。

【 0 1 2 3 】

50

また、本実施形態によれば、上記画像情報を、感情の度合いが最大となった場合における、対応する参加者の顔を撮影して得られた顔撮影画像情報としている。従って、より効果的に参加者の感情を把握することができる。

【0124】

また、本実施形態によれば、上記画像情報を、顔撮影画像情報に加えて、感情を誇張する情報が含まれる画像情報としている。従って、より効果的に参加者の感情を把握することができる。

【0125】

また、本実施形態によれば、上記感情を表す情報を、参加者の相互間における感情の関係を示す情報としている。従って、より効果的に対話の活性化を促すことができる。

10

【0126】

また、本実施形態によれば、上記物理量を、参加者を撮影して得られた画像、及び参加者の発言を示す音声としている。従って、より低コストで対話の活性化を促すことができる。

【0127】

また、本実施形態によれば、上記状況を、参加者の感情の度合いを表す物理量、及び参加者の動作を表す物理量としている。従って、より簡易に対話の活性化を促すことができる。

【0128】

更に、本実施形態によれば、上記状況を、画像から得られる参加者のうなずきの頻度を示す物理量、画像から得られる参加者の表情の度合いを示す物理量、音声から得られる参加者の発言の度合いを示す物理量としている。従って、より簡易に対話の活性化を促すことができる。

20

【0129】

なお、上記実施形態では、各参加者の個別の感情度を用いて、端末20に表示する対応情報を決定する場合について説明したが、これに限定されない。例えば、参加者全員の感情度を用いて対応情報を決定する形態としてもよい。例えば、一例として図12に示すように、各参加者の喜び度 $Y$ が同時に所定値（一例として、50）以上となった場合、参加者全員が一体的に喜んでいと想定できる。この場合、対応情報として、一例として「一体感があり、良い状況です。」といった表示を各端末20で行うことで、より効果的に対話を活性化することができる。

30

【0130】

また、上記実施形態では、本発明を、対話の参加者が互いに異なる場所に分散して会議を行っている形態に適用した場合について説明したが、これに限定されない。例えば、対話の各参加者が同一の会議室等で会議を行う形態に本発明を適用してもよい。この場合、各端末20に設けられたカメラ28及びマイク29に代えて、端末20とは別体として構成された1つ又は複数のカメラ及びマイクを用いて、会議の参加者全員の画像及び音声を収集する形態としてもよい。

【0131】

また、上記実施形態では、対話支援装置10において対話支援処理を実行する場合について説明したが、これに限定されない。例えば、少なくとも1台の端末20によって対話支援処理を実行する形態としてもよい。この形態の場合、本発明の対話支援装置が該当する端末20に含まれることになる。また、例えば、各参加者の発言度 $H$ 、うなずき頻度低下率 $U$ 、及び各感情度の少なくとも1つを、対応する参加者が用いる端末20で導出する形態としてもよい。

40

【0132】

また、上記実施形態では、各端末20において対話支援画像30を表示する場合について説明したが、これに限定されない。例えば、対話支援画像30を対話支援装置10において表示する形態としてもよい。

【0133】

50

また、上記実施形態では、本発明を会議に適用した場合について説明したが、これに限定されない。例えば、人事面接、商談等といった会議以外の複数人で行う対話の場に本発明を適用する形態としてもよい。

【0134】

また、上記実施形態では、本発明の感情を表す画像情報として顔文字を適用した場合について説明したが、これに限定されない。例えば、顔文字に加えて、絵文字、アイコン (Icon) の少なくとも1つを適用する形態としてもよい。

【0135】

また、上記実施形態では、上記状況として、参加者の感情の度合いを表す物理量、及び参加者の動作を表す物理量の双方を適用した場合について説明したが、これに限定されない。例えば、参加者の感情の度合いを表す物理量、及び参加者の動作を表す物理量の何れか一方のみを適用する形態としてもよい。

10

【0136】

また、上記実施形態では、上記状況として、画像から得られる参加者のうなずきの頻度を示す物理量、画像から得られる参加者の表情の度合いを示す物理量、音声から得られる参加者の発言の度合いを示す物理量、の全てを適用した場合について説明したが、これに限定されない。例えば、これらの物理量の1つ、又は全てを除く複数の組み合わせを適用する形態としてもよい。

【0137】

また、上記実施形態では、うなずき頻度低下率Uを用いて対応情報を決定する場合について説明したが、これに限定されない。例えば、うなずき頻度Nそのものを用いて対応情報を決定する形態としてもよい。

20

【0138】

また、上記実施形態では、状況対応情報データベース13Bとして、発話者と受話者の双方に関する情報が混在しているデータベースを適用した場合について説明したが、これに限定されない。例えば、発話者と受話者の各々別に異なるデータベースを構築して適用する形態としてもよい。

【0139】

その他、式(1)~式(3)は何れも一例であり、本発明の主旨を逸脱しない範囲内において、適宜変更して適用することができることは言うまでもない。

30

【0140】

また、上記実施形態において、例えば、取得部11A、導出部11B、処理部11Cの各処理を実行する処理部 (processing unit) のハードウェア的な構造としては、次に示す各種のプロセッサ (processor) を用いることができる。上記各種のプロセッサには、前述したように、ソフトウェア (プログラム) を実行して処理部として機能する汎用的なプロセッサであるCPUに加えて、FPGA (Field-Programmable Gate Array) 等の製造後に回路構成を変更可能なプロセッサであるプログラマブルロジックデバイス (Programmable Logic Device : PLD)、ASIC (Application Specific Integrated Circuit) 等の特定の処理を実行させるために専用に設計された回路構成を有するプロセッサである専用電気回路等が含まれる。

40

【0141】

処理部は、これらの各種のプロセッサのうちの1つで構成されてもよいし、同種又は異種の2つ以上のプロセッサの組み合わせ (例えば、複数のFPGAの組み合わせや、CPUとFPGAとの組み合わせ) で構成されてもよい。また、処理部を1つのプロセッサで構成してもよい。

【0142】

処理部を1つのプロセッサで構成する例としては、第1に、クライアント及びサーバ等のコンピュータに代表されるように、1つ以上のCPUとソフトウェアの組み合わせで1つのプロセッサを構成し、このプロセッサが処理部として機能する形態がある。第2に、システムオンチップ (System On Chip : SoC) 等に代表されるように、処理部を含むシ

50

ステム全体の機能を1つのIC (Integrated Circuit) チップで実現するプロセッサを使用する形態がある。このように、処理部は、ハードウェア的な構造として、上記各種のプロセッサの1つ以上を用いて構成される。

【0143】

更に、これらの各種のプロセッサのハードウェア的な構造としては、より具体的には、半導体素子などの回路素子を組み合わせた電気回路 (circuitry) を用いることができる。

【符号の説明】

【0144】

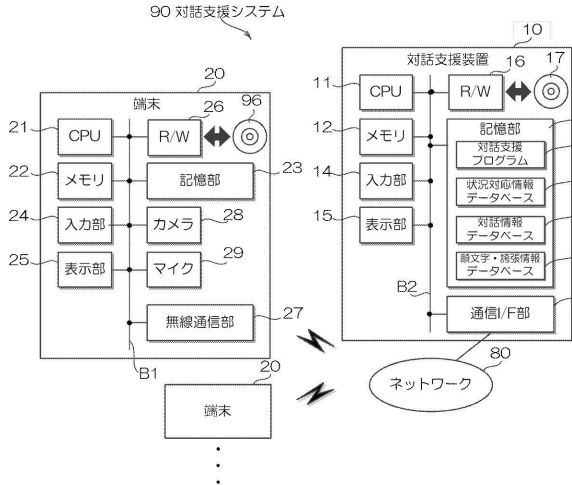
10	対話支援装置	
11	CPU	10
11A	取得部	
11B	導出部	
11C	処理部	
12	メモリ	
13	記憶部	
13A	対話支援プログラム	
13B	状況対応情報データベース	
13C	対話情報データベース	
13D	顔文字・誇張情報データベース	
14	入力部	20
15	表示部	
16	媒体読み書き装置	
17	記録媒体	
18	通信I/F部	
20	端末	
21	CPU	
22	メモリ	
23	記憶部	
24	入力部	
25	表示部	30
25A	派閥情報表示ボタン	
25B	終了ボタン	
25C	撮影画像	
25D	テキスト	
25E	顔文字	
25F	テキスト	
25G	画像	
25H	対応情報	
25I	支援情報	
25J	音声ボタン	40
25K	スクロールボタン	
25L、25M、25N	画像	
25P	撮影画像	
26	媒体読み書き装置	
27	無線通信部	
28	カメラ	
29	マイク	
30	対話支援画像	
32	派閥関係表示画像	
80	ネットワーク	50

9 0 対話支援システム

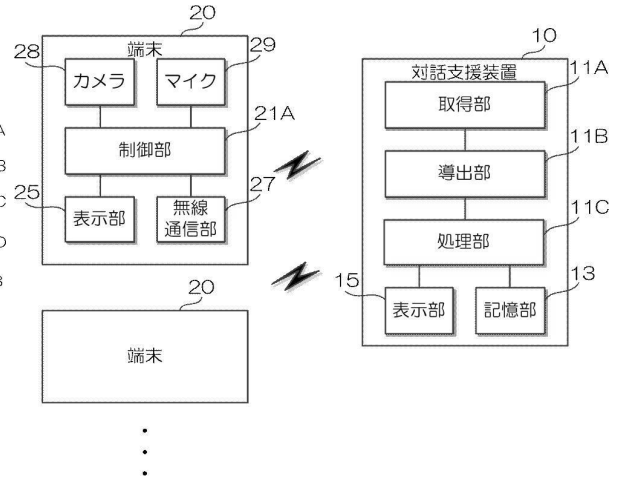
9 6 記録媒体

【図面】

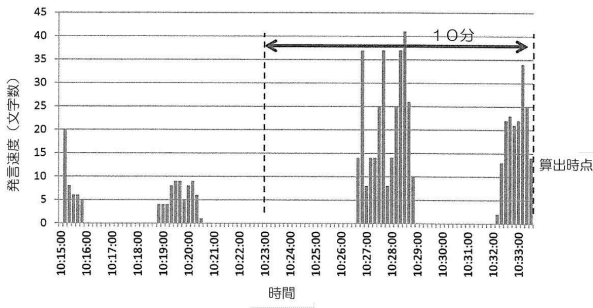
【図 1】



【図 2】



【図 3】



【図 4】

状況対応情報データベース

発言度H	うなずき頻度低下率U(%)	状況					対応情報
		怒り度I	嫌悪度K	恐れ度O	喜び度Y	悲しみ度S	
H ≥ 180	U ≥ 50	I ≥ 50	-	-	-	-	「ちょっと話についていけないなあ」
H ≥ 200	U ≥ 60	I ≥ 60	K ≥ 50	-	-	-	「嬉しい！補足説明を！」
H ≥ 100	U ≥ 70	-	-	O ≥ 70	-	-	「OOさん、何か意見ありますか？」
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

10

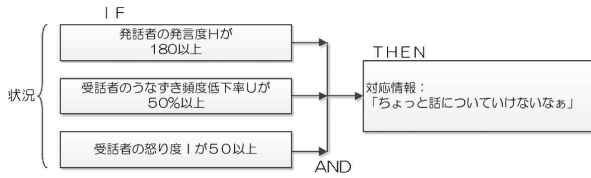
20

30

40

50

【図5】



【図6】

対話情報データベース

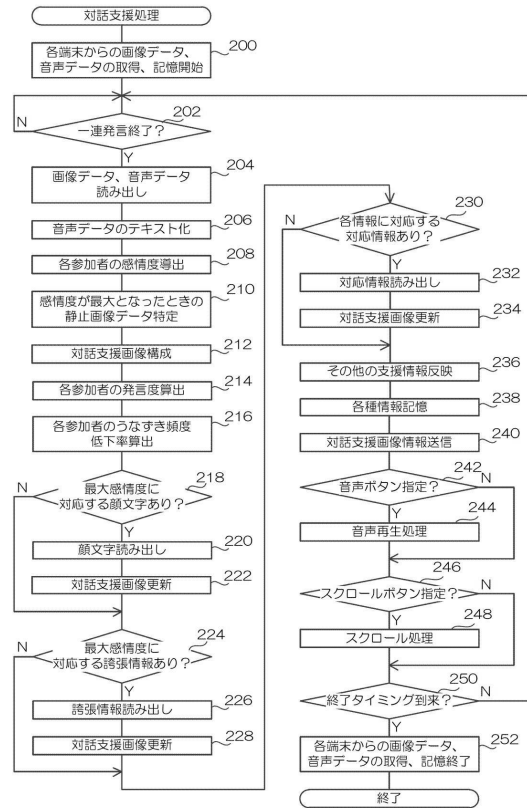
端末ID	画像データ	音声データ	テキストデータ	時刻	最大感情度			
					x1	x2	x3	x4
x1	(画像データA)	(なんとかかんとかがどうのこうの①)	「なんとかかんとか・・・①」	2019/2/10 13:05:02~ 13:05:09	I	-	-	-
x3	(画像データC)	(なんとかかんとかがどうのこうの②)	「なんとかかんとか・・・②」	2019/2/10 13:05:11~ 13:05:18	-	-	O	-
x2	(画像データB)	(なんとかかんとかがどうのこうの③)	「なんとかかんとか・・・③」	2019/2/10 13:05:20~ 13:05:27	-	Y	-	-
x3	(画像データC)	(なんとかかんとかがどうのこうの④)	「なんとかかんとか・・・④」	2019/2/10 13:05:29~ 13:05:40	-	-	O	-
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

【図7】

顔文字・誇張情報データベース

最大感情度	顔文字	誇張情報
Y	(^o^)	-
S	(T T)	-
O	(^-;) )	「ガーン」+(恐れを示す画像)
I	(- -)	-
⋮	⋮	⋮
⋮	⋮	⋮
⋮	⋮	⋮

【図8】



10

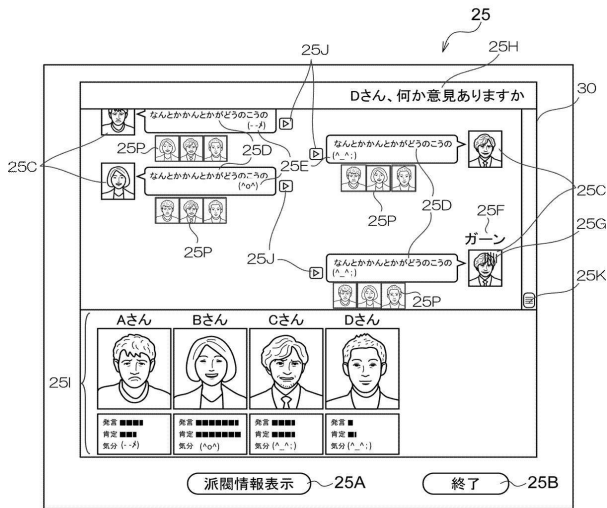
20

30

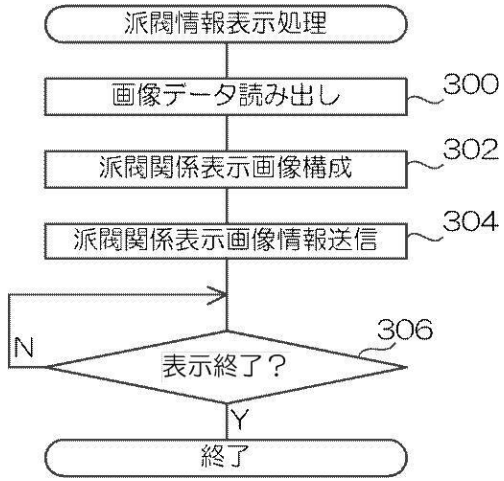
40

50

【図 9】

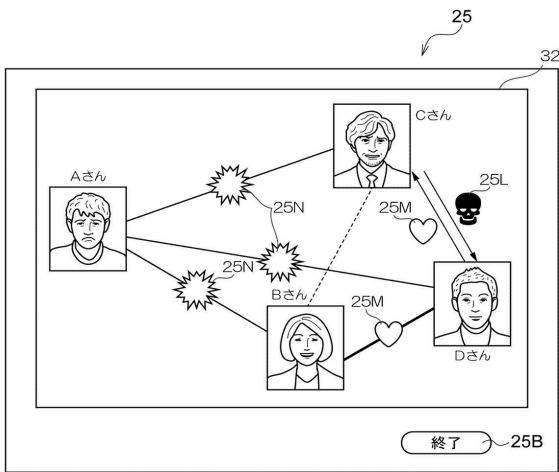


【図 10】

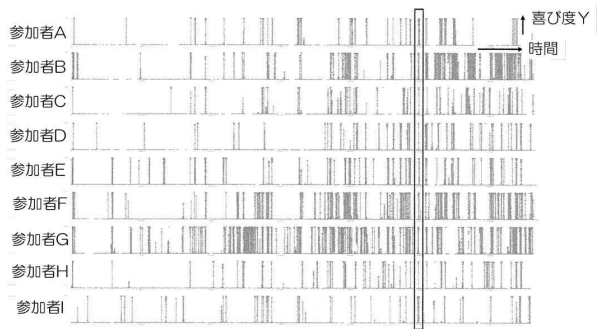


10

【図 11】



【図 12】



20

30

40

50

## フロントページの続き

(51)国際特許分類		F I			
<b>H 0 4 M</b>	<b>3/56 (2006.01)</b>		H 0 4 M	3/56	C
<b>G 0 6 F</b>	<b>3/16 (2006.01)</b>		G 0 6 F	3/16	6 5 0

会社竹中工務店 技術研究所内

- (72)発明者 大塚 貴弘  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 内出 隼人  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 澤田 友哉  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 川島 啓吾  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 津田 由佳  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 志田 哲郎  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 吉田 諒  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 石川 美穂  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内
- (72)発明者 飯田 隆義  
東京都千代田区丸の内二丁目7番3号 三菱電機株式会社内

審査官 毛利 太郎

- (56)参考文献 特開2015-165407(JP,A)  
特開2004-064102(JP,A)  
特開2006-227236(JP,A)

- (58)調査した分野 (Int.Cl., DB名)
- G 0 6 Q 1 0 / 0 0 - 9 9 / 0 0  
G 0 6 F 3 / 0 1  
G 1 0 L 1 5 / 0 0  
G 0 6 F 3 / 0 4 8  
H 0 4 N 2 1 / 2 5 8  
H 0 4 M 3 / 5 6  
G 0 6 F 3 / 1 6