

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2019-211909

(P2019-211909A)

(43) 公開日 令和1年12月12日(2019.12.12)

(51) Int.Cl.			F I	テーマコード (参考)	
G06F	16/00	(2019.01)	G06F	17/30	340A
G06F	16/90	(2019.01)	G06F	17/30	180A
G10L	15/00	(2013.01)	G06F	17/30	210D
G10L	15/22	(2006.01)	G10L	15/00	200T
G10L	13/00	(2006.01)	G10L	15/22	300U

審査請求 未請求 請求項の数 10 OL (全 25 頁) 最終頁に続く

(21) 出願番号 特願2018-106181 (P2018-106181)
 (22) 出願日 平成30年6月1日(2018.6.1)

(71) 出願人 000003193
 凸版印刷株式会社
 東京都台東区台東1丁目5番1号
 (74) 代理人 100149548
 弁理士 松沼 泰史
 (74) 代理人 100139686
 弁理士 鈴木 史朗
 (74) 代理人 100169764
 弁理士 清水 雄一郎
 (74) 代理人 100147267
 弁理士 大槻 真紀子
 (72) 発明者 波多野 亮平
 東京都台東区台東1丁目5番1号 凸版印刷株式会社内

(54) 【発明の名称】 情報提示システム、情報提示方法及びプログラム

(57) 【要約】

【課題】システムが音声により提供する情報を、ユーザの各々が正確に聞き取ることができ、正確に聞き取れるように発話データを変更する変更モデルがユーザ毎に設けられ、ルールベースに比較してデータ量が少ないため変更モデルのメンテナンスが容易に行える情報提示システムを提供する。

【解決手段】本発明の情報提示システムは、ユーザとの対話において、ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、ユーザの聴取志向に対応して変更する変更モデルを生成及び更新する聴取志向推定部と、ユーザ毎に設定されている変更モデルに対応し、対話システムからの音声による回答である発話データを、ユーザの各々の聴取志向となるよう変更する提示制御部とを備える。

【選択図】 図1

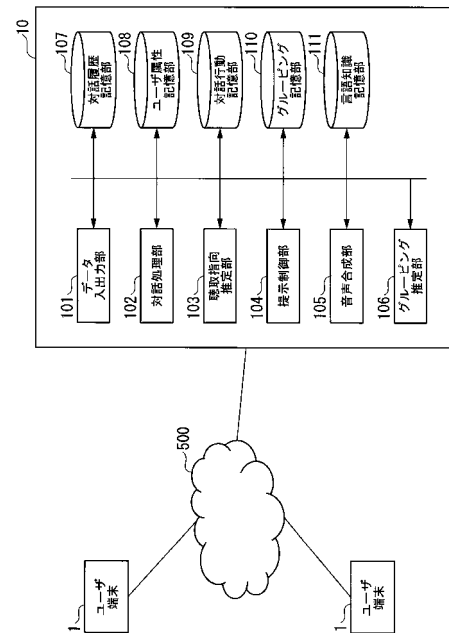


図1

【特許請求の範囲】**【請求項 1】**

ユーザとの対話において、前記ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、前記ユーザの前記聴取志向に対応して変更する変更モデルを生成及び更新する聴取志向推定部と、

前記ユーザ毎に設定されている前記変更モデルに対応し、対話システムからの音声による回答である前記発話データを、前記ユーザの各々の前記聴取志向となるよう変更する提示制御部と

を備えることを特徴とする情報提示システム。

【請求項 2】

前記聴取志向を推定する際に用いる、前記ユーザの各々の前記対話の履歴である対話履歴を対話履歴記憶部に対して、前記ユーザ毎に書き込んで記憶させる、ユーザからの発話に対してルールに基づき応答を決定する対話処理部

をさらに備える

ことを特徴とする請求項 1 に記載の情報提示システム。

【請求項 3】

前記聴取志向推定部が、

前記ユーザの前記対話における前記発話データに対する評価から、当該ユーザの前記聴取志向を抽出して、前記ユーザの属性情報及び当該ユーザの前記聴取志向を示す志向情報の各々を、ユーザ属性記憶部に対して、前記ユーザ毎に書き込んで記憶させる

ことを特徴とする請求項 1 または請求項 2 に記載の情報提示システム。

【請求項 4】

前記ユーザの各々の前記属性情報に対応して、前記ユーザそれぞれを分類するグルーピングを行い、前記分類毎に含まれる前記ユーザに共通する前記聴取志向により、当該分類それぞれの変更モデルであるテンプレート変更モデルを生成するグルーピング推定部

をさらに備えることを特徴とする請求項 3 に記載の情報提示システム。

【請求項 5】

前記聴取志向推定部が、

前記変更モデルが用意されていない前記ユーザに対して、当該ユーザに対応する前記分類の前記テンプレート変更モデルを抽出し、前記対話において抽出される前記聴取志向に対応して、当該ユーザに対応する前記変更モデルを生成する

ことを特徴とする請求項 4 に記載の情報提示システム。

【請求項 6】

前記属性情報が、

少なくとも、前記ユーザの年齢、性別、居住地を含むデモグラフィックデータの各々の組み合わせとして設定される

ことを特徴とする請求項 3 から請求項 5 のいずれか一項に記載の情報提示システム。

【請求項 7】

前記変更モデルが、

少なくとも、前記対話処理部を介して決定した前記発話データにおける単語の置き換え、前記発話データを読み上げる際の音声の周波数及び速度、文節の区切りを変更する処理を示す

ことを特徴とする請求項 1 に記載の情報提示システム。

【請求項 8】

前記提示制御部が、

前記変更モデルによる前記発話データを変更した内容である変更内容を、対話行動記憶部に対して、変更履歴として書き込んで記憶させ、

前記聴取志向推定部が、

前記対話の履歴と前記変更履歴とにより、前記ユーザの前記聴取志向を抽出する

ことを特徴とする請求項 1 から請求項 7 のいずれか一項に記載の情報提示システム。

10

20

30

40

50

【請求項 9】

聴取志向推定部が、ユーザとの対話において、前記ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、前記ユーザの前記聴取志向に対応して変更する変更モデルを生成及び更新する聴取志向推定過程と、

提示制御部が、前記ユーザ毎に設定されている前記変更モデルに対応し、ユーザからの発話に対してルールに基づき応答を決定する対話処理部を介して、対話システムからの音声による回答である前記発話データを、前記ユーザの各々の前記聴取志向となるよう変更する提示制御過程と

を含むことを特徴とする情報提示方法。

10

【請求項 10】

コンピュータを、

ユーザとの対話において、前記ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、前記ユーザの前記聴取志向に対応して変更する変更モデルを生成及び更新する聴取志向推定手段、

前記ユーザ毎に設定されている前記変更モデルに対応し、対話システムからの音声による回答である前記発話データを、前記ユーザの各々の前記聴取志向となるよう変更する提示制御手段

として機能させるためのプログラム。

【発明の詳細な説明】

20

【技術分野】**【0001】**

本発明は、情報提示システム、情報提示方法及びプログラムに関する。

【背景技術】**【0002】**

近年、インターネット環境が充実したことにより、ソーシャル・ネットワーク・サービス（以下、SNSと示す）が普及し、テキストや画像を用いて複数のユーザ間において簡易に意思疎通を行うことが可能となっている。例えば、SNSのアプリケーションとしては、LINE（登録商標）、Facebook（登録商標）メッセージ、Slack（登録商標）などが代表的である。これらのSNSは、一対一のユーザ間の情報のやり取りだけでなく、所定のグループにおける多人数のユーザ間で送受信する情報（複数のユーザ間における対話）を、グループ内の全てのユーザで共有する機能も有している。

30

【0003】

また、マン・マシン対話型のSNSとしては、Google Assistant（登録商標）、Amazon Alexa（登録商標）、Line Clova（登録商標）などがある。

また、上述したアプリケーションの各々が、パーソナルコンピュータ及びスマートデバイスや、Google Home（登録商標）、Amazon Echo（登録商標）、Clova Wave（登録商標）などのスマートスピーカに搭載され、それぞれにおいて音声合成されて、音声を用いた情報提示を主体としたものも広く利用されている。

40

【0004】

SNSにおける情報提示方法には、上記アプリケーション毎に様々な工夫がされている。

例えば、情報提示における提示内容に対して、システム上で定義したキャラクタとともに情報文を提示することを目的として、情報文の言語表現を書き換えて提示する提示方法がある（例えば、特許文献1参照）。

【0005】

また、ユーザに対して音声により情報提示を行う場合、提示に用いる音声の音声合成に関して、システムが合成した情報提示の音声をユーザに対して、スピーカなどを介して発話して情報の提示を行う。このとき、アプリケーションが、提示される発話をユーザが聴

50

いた際に、発話に対して機械的な不自然さを感じさせない処理を行う情報提示の方法がある（例えば、特許文献2及び特許文献3参照）。

【先行技術文献】

【特許文献】

【0006】

【特許文献1】特許第6161656号公報

【特許文献2】特許第5954348号公報

【特許文献3】特許第6232892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、ユーザとシステムとの間において、音声のみによる対話が行われる場合、ユーザの属性が年齢あるいは性別などの多様性を有しているため、ユーザの聴力や単語に対する理解力が異なる。

このため、システムとの対話において、システムが音声により提供する情報が聞き取れない、あるいは、音声における単語を理解できない等により、対話において情報の内容が正確に伝達されない場合がある。

【0008】

また、ユーザとシステムとの対話を行うために、ユーザからの問いかけに対してどのような回答をするかについて、予め対話シナリオを想定して対話を実行するルールベース手法を用いることができる。

しかしながら、ルールベースに設定されたルールの各々に対して、多くのユーザの各々に対応させる必要性から手作業によるメンテナンスの必要が有る。このメンテナンスにおいて、上述した多様性のあるユーザの各々に対応させて、聞き取りにくいあるいは理解が困難な単語についての類義の単語や表現を含む上記ルールを設定し、対話シナリオとして構築することは膨大な作業が必要となる。

【0009】

本発明は、このような状況に鑑みてなされたもので、システムが音声により提供する情報を、ユーザの各々が正確に聞き取ることができ、正確に聞き取れるように発話データを変更する変更モデルがユーザ毎に設けられ、ルールベースに比較してデータ量が少ないため変更モデルのメンテナンス（ユーザに順次対応させていく修正処理）が容易に行える情報提示システム、情報提示方法及びプログラムを提供する。

【課題を解決するための手段】

【0010】

上述した課題を解決するために、本発明の情報提示システムは、ユーザとの対話において、前記ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、前記ユーザの前記聴取志向に対応して変更する変更モデル（実施形態における聴取志向推定モデル及び聴取志向テンプレートモデルの各々を含む）を生成、予測及び更新する聴取志向推定部と、前記ユーザ毎に設定されている前記変更モデルに対応し、対話システムからの音声による回答である前記発話データを、前記ユーザの各々の前記聴取志向となるよう変更する提示制御部とを備えることを特徴とする。

【0011】

本発明の情報提示システムは、前記聴取志向を推定する際に用いる、前記ユーザの各々の前記対話の履歴である対話履歴を対話履歴記憶部に対して、前記ユーザ毎に書き込んで記憶させる、ユーザからの発話に対してルールに基づき応答を決定する対話処理部をさらに備えることを特徴とする。

【0012】

本発明の情報提示システムは、前記聴取志向推定部が、前記ユーザの前記対話における前記発話データに対する評価から、当該ユーザの前記聴取志向を抽出して、前記ユーザの

10

20

30

40

50

属性情報及び当該ユーザの前記聴取志向を示す志向情報の各々を、ユーザ属性記憶部に対して、前記ユーザ毎に書き込んで記憶させることを特徴とする。

【0013】

本発明の情報提示システムは、前記ユーザの各々の前記属性情報に対応して、前記ユーザそれぞれを分類するグルーピングを行い、前記分類毎に含まれる前記ユーザに共通する前記聴取志向により、当該分類それぞれの変更モデルであるテンプレート変更モデルを生成するグルーピング推定部をさらに備えることを特徴とする。

【0014】

本発明の情報提示システムは、前記聴取志向推定部が、前記変更モデルが用意されていない前記ユーザに対して、当該ユーザに対応する前記分類の前記テンプレート変更モデルを抽出し、前記対話において抽出される前記聴取志向に対応して、当該ユーザに対応する前記変更モデルを生成することを特徴とする。

10

【0015】

本発明の情報提示システムは、前記属性情報が、少なくとも、前記ユーザの年齢、性別、居住地を含むデモグラフィックデータの各々の組み合わせとして設定されることを特徴とする。

【0016】

本発明の情報提示システムは、前記変更モデルが、少なくとも、前記対話処理部により決定されたシステム応答の発話データにおける単語の置き換え、前記発話データを読み上げる際の音声の周波数及び速度、文節の区切りを変更する処理を示すことを特徴とする。

20

【0017】

本発明の情報提示システムは、前記提示制御部が、前記変更モデルによる前記発話データを変更した内容である変更内容を、対話行動記憶部に対して、変更履歴として書き込んで記憶させ、前記聴取志向推定部が、前記対話の履歴と前記変更履歴とにより、前記ユーザの前記聴取志向を抽出することを特徴とする。

【0018】

本発明の情報提示方法は、聴取志向推定部が、ユーザとの対話において、前記ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、前記ユーザの前記聴取志向に対応して変更する変更モデルを生成及び更新する聴取志向推定過程と、提示制御部が、前記ユーザ毎に設定されている前記変更モデルに対応し、ユーザからの発話に対してルールに基づき応答を決定する対話処理部を介して、対話システムからの音声による回答である前記発話データを、前記ユーザの各々の前記聴取志向となるよう変更する提示制御過程とを含むことを特徴とする。

30

【0019】

本発明のプログラムは、コンピュータを、ユーザとの対話において、前記ユーザの各々の音声の聞き取り易さを示す聴取志向を推定し、当該ユーザに対して音声により供給される発話データを、前記ユーザの前記聴取志向に対応して変更する変更モデルを生成及び更新する聴取志向推定手段、前記ユーザ毎に設定されている前記変更モデルに対応し、対話システムからの音声による回答である前記発話データを、前記ユーザの各々の前記聴取志向となるよう変更する提示制御手段として機能させるためのプログラムである。

40

【発明の効果】

【0020】

以上説明したように、本発明によれば、システムが音声により提供する情報を、ユーザの各々が正確に聞き取ることができ、正確に聞き取れるように発話データを変更する変更モデルがユーザ毎に設けられ、ルールベースに比較して事前に対話内容を想定して構築すべき対話シナリオをはじめとしたデータ量が少ないため変更モデルのメンテナンス（ユーザに順次対応させていく修正処理）が容易に行える情報提示システム、情報提示方法及びプログラムを提供することができる。

また、グルーピング推定部があることで、変更モデルが存在しない、あるいは発話データなどの蓄積が不十分なユーザの各々に対して、グループ内で一般化された変更モデルで

50

あるテンプレート変更モデルを用いることで、表示情報の最適化を行うことができる。

【図面の簡単な説明】

【0021】

【図1】本発明の一実施形態による、ユーザとシステムとが対話を行う情報提示システムの構成例を示すブロック図である。

【図2】本実施形態における対話処理部102の構成例を示すブロック図である。

【図3】対話履歴記憶部107に記憶されている対話履歴テーブルの構成例を示す図である。

【図4】本実施形態における聴取志向推定部103の構成例を示すブロック図である。

【図5】ユーザ属性記憶部108に記憶されているユーザ属性テーブルの構成例を示す図である。

【図6】対話行動記憶部109に記憶されている対話行動テーブルの構成例を示す図である。

【図7】図6に示したアクションにおける単語の置き換えの処理を説明する概念図である。

【図8】グルーピング記憶部110に記憶されているグルーピングテーブルの構成例を示す図である。

【図9】本実施形態の情報提示システムを用いた対話システムの動作例を示すフローチャートである。

【図10】本発明の一実施形態による、ユーザとシステムとが対話を行う情報提示システムの他の構成例を示す概念図である。

【発明を実施するための形態】

【0022】

本発明は、例えば、ユーザが質問を行うと、システム側がその質問に対応した回答を、発話データを音声合成した音声により通知する、あるいはユーザ同士で対話する対話システムに関するものである。また、ユーザの聴力及び単語の理解力の各々に対応して、システム側が通知する発話データを変更して、ユーザが対話システムからの回答を聞き易く、また内容を理解し易くする構成に関する。

【0023】

ユーザの聴力に対応しては、例えば、発話データを音声合成して音声として出力する際における、この音声の周波数、再生速度、再生する文節の区切り、この区切りの時間幅などの変更を行う。ここで、区切りは、発話データを音声として再生する際に、文節と文節との間あるいは単語と単語との間に挿入される無音である。区切りの時間幅は、文節と文節との間あるいは単語と単語との間に挿入される無音の時間の長さを示している。

また、ユーザの単語の理解力に対応しては、例えば、専門用語を一般的に用いている他の同義語（あるいは類義語、類語）である単語に置き換える変更を行う。

【0024】

以下、本発明の一実施形態について、図面を参照して説明する。図1に対応した以下の説明においては、ユーザとシステムとの対話を例に説明する。

図1は、本発明の一実施形態による、ユーザとシステムとが対話を行う情報提示システムの構成例を示すブロック図である。

図1において、情報提示システム1は、情報提示サーバ10とユーザ端末11との各々を備えている。

情報提示サーバ10とユーザ端末11との各々は、インターネットを含む情報通信網であるネットワーク500を介してデータの送受信を行う。

【0025】

情報提示サーバ10は、ユーザ端末11を介して供給されるユーザの質問等に対応した回答を音声データにより、ユーザ端末11に対して出力する。ここで、音声データは、音声を所定の音声ファイルフォーマット（非圧縮音声フォーマット、非可逆圧縮音声フォーマット及び可逆圧縮音声フォーマットなど）により、デジタル化したデータを示している

10

20

30

40

50

。情報提示サーバ10は、例えば、汎用コンピュータあるいはパーソナルコンピュータである。

【0026】

ユーザ端末11は、Amazon Alexa（登録商標）、Apple Siri（登録商標）、Googleアシスタント（登録商標）などといった仮想パーソナルアシスタント（VPA: Virtual Personal Assistant）を利用するためのプラットフォームとなる、音声によりユーザと情報提示システム1との対話を行うスマートスピーカなどである。また、ユーザ端末11は、スマートフォン及びタブレット端末などの携帯端末、あるいはパーソナルコンピュータでも良く、表示部に画像（テキスト文字も含む）を表示させずに、音声のみによりリクエスト及びレスポンスが行われる対話アプリケーションに対して応用しても良い。

10

【0027】

情報提示サーバ10は、データ入出力部101、対話処理部102、聴取志向推定部103、提示制御部104、音声合成部105、グルーピング推定部106、対話履歴記憶部107、ユーザ属性記憶部108、対話行動記憶部109、グルーピング記憶部110及び言語知識記憶部111の各々を備えている。

【0028】

データ入出力部101は、ユーザがユーザ端末11に対して入力した音声の音声データを含むデータを、ネットワーク500を介して入力する外部入力インターフェースである。

また、データ入出力部101は、ユーザからの音声データの内容に対する回答などの音声データを含むデータを、ネットワーク500を介してユーザ端末11に対して出力する外部出力インターフェースでもある。

20

また、データ入出力部101は、情報提示サーバ10を操作する制御信号などのデータを、マイクロフォン、キーボード、各種センサを含む入力手段から直接に、データ（音声データを含む）を取得する機能を有している。

【0029】

対話処理部102は、ユーザからの音声データを解析し、ユーザの音声データの内容を推定する。そして、対話処理部102は、この推定した内容に対する回答メッセージ（対話における回答）として、テキスト文の発話データを生成する。

【0030】

図2は、本実施形態における対話処理部102の構成例を示すブロック図である。図2において、対話処理部102は、解析部1021、対話管理部1022及び生成部1023の各々を備えている。

30

解析部1021は、データ入出力部101から供給される音声データを、文字化、すなわちテキスト変換し、対話履歴記憶部107に対して書き込んで記憶させる。

【0031】

また、対話処理部102は、テキスト変換したテキスト文の形態素解析を行い、得られた形態素からのキーワード抽出、形態素のベクトル化などの数値化処理を行い、テキスト文の数値データへの変換処理を行う。この変換処理は、自然言語処理技術あるいは機械学習技術を用いて行う。本実施形態において、例えば、 tf (term frequency) - idf (inverse document frequency) 法によるキーワード抽出、 $word2vec$ 及び $doc2vec$ によるベクトル化などの手法を用いても良い。対話処理部102は、抽出されたキーワードあるいはベクトル化された数値データを、対話管理部1022に対して出力する。

40

【0032】

図3は、対話履歴記憶部107に記憶されている対話履歴テーブルの構成例を示す図である。図3において、対話履歴テーブルは、レコード毎に対して、メッセージ時刻、話者ID (Identification)、メッセージ本文、聴取志向フレーズフラグ、メッセージID1及びメッセージID2の各々の欄が設けられている。

【0033】

50

ここで、メッセージ時刻は、メッセージのうち音声データを入力した、あるいはメッセージのうち発話データを出力した時刻を示している。話者IDは、音声データあるいは発話データを出力した話者の識別情報（話者がユーザであれば後述するユーザID）を示している。例えば、話者IDにおいて「U_001」は、話者がユーザであり、このユーザを識別するユーザ識別情報である。また、話者IDにおいて「C_001」は、話者が情報提示サーバ10（システム）であり、このシステムを識別するシステム識別情報である。話者が情報提示サーバ10である場合、ユーザの音声データの内容の種別（天気、健康、科学など）毎に、異なるシステムが回答するため、複数の異なるシステム識別情報がある。

【0034】

また、メッセージ本文は、音声データをテキスト変換した文字データ、あるいは発話データなどのテキストデータが記載されている。聴取志向フレーズフラグは、ユーザからの音声データの内容が質問ではなく、情報提示サーバ10からの発話データに基づく音声理解できないことを示しているメッセージ本文に対して付与するフラグである。ここで、聴取志向フレーズフラグが「0」、すなわちフラグが立っていない場合、対応するメッセージ本文が通常の会話のフレーズであることを示している。

【0035】

一方、聴取志向フレーズフラグが「1」、すなわちフラグが立っている場合、対応するメッセージ本文に対して、ユーザがシステムの音声の内容が判らない（聞き取れない或いは内容が理解できない）ことを示す聴取志向フレーズであることを示している。

例えば、図3において、聴取志向フレーズフラグが「1」とされたメッセージ本文は、「聞き取れなかったから、もう一回御願い」であり、情報提示サーバ10がユーザ端末11から出力する音声聞き取れなかった（周波数、音量、区切りなどの聴取志向におけるパラメータにより）ことを示している。この聴取志向フレーズは、情報提示サーバ10からのユーザのリクエストに対するレスポンスである音声コンテンツに対して、ユーザが音声コンテンツの聞き取り易さの程度を示す単語（後述するポジティブワード及びネガティブワードに対応）、あるいは同義語（類義語、類語）であり、予め識者により所定のフレーズとして登録されている。

【0036】

また、メッセージID1は、同一のレコードにおけるメッセージ本文を指し示す識別情報である。一方、メッセージID2は、ユーザ及びシステムの各々の間の対話において、メッセージID1の示すメッセージ本文のメッセージに対して直前のメッセージの識別情報（すなわち、メッセージID1）である識別情報である。

したがって、ユーザ及びシステムの各々の間の対話における前後の関係を確認する際、確認したいメッセージ本文のレコードにあるメッセージID2を読み出し、このメッセージID2と同一のメッセージID1を検索することにより、確認したいメッセージ本文の直前のメッセージのメッセージ本文を検索することができ、対話の連続したメッセージ本文の各々を容易に確認できる。

【0037】

例えば、話者IDがC_001のメッセージID1：M180101003に対して、メッセージID2：M180101001となっている。このため、「今週の週末はいかがですか？」を回答とする質問が、メッセージID1：M180101001の「いきたい」であることが容易に検索できる。メッセージID1及びメッセージID2の各々が連続した番号でないのは、途中で他の対話が行われる場合があり、一つの対話におけるメッセージが常に連続して入力されないことを示している。

【0038】

図2に戻り、対話管理部1022は、キーワードや数値データの各々とともに供給されるメッセージID1により、対話履歴記憶部107を参照して、メッセージID1及びメッセージID2の各々の関係から、ユーザの状態（会話フレーズか聴取志向フレーズ）であるかを定義する。そして、対話管理部1022は、メッセージ本文が会話フレーズであ

10

20

30

40

50

る場合、情報提示サーバ10によるユーザへのシステム応答の指針（例えば、ジャンル指定、場所指定、天気指定、交通の時刻表指定などの対話行為タイプ）を決定する。

【0039】

すなわち、対話管理部1022は、ユーザからの会話フレーズとしてのリクエスト（問いかけ）に対して、システム側からのレスポンス（応答）をデータベース化した構成としても良いし、機械学習あるいは強化学習などの枠組みを用いて、リクエストに対応するレスポンスの内容を出力する対話モデルを生成して構成として用いても良い。この対話管理部1022の処理については、一般的な公知の技術である対話システムと同様のため、詳細な説明を省略する。

【0040】

また、対話管理部1022は、メッセージID1の各々のメッセージ本文のキーワードや数値データそれぞれから、会話フレーズか聴取志向フレーズのいずれかであることを検出すると、対話履歴記憶部107の対話記憶履歴テーブルの対応するレコードにおける聴取志向フレーズフラグの操作を行う。このとき、対話管理部1022は、メッセージ本文が会話フレーズであると判定した場合、聴取志向フレーズフラグを「0」として、フラグを立てない。一方、対話管理部1022は、メッセージ本文が聴取志向フレーズであると判定した場合、聴取志向フレーズフラグを「1」として、フラグを立てる。

【0041】

生成部1023は、対話管理部1022の出力するシステム応答の指針に基づき、その指針に対応したシステム応答モデルを用いて、テキストデータの応答文である発話データを生成する。すなわち、生成部1023は、リクエストのメッセージ本文から抽出したキーワードやベクトル化した数値データなどを、上記システム応答モデルに対して入力し、リクエストに対応した内容の発話データを生成する。上記システム応答モデルは、公知の技術の教師データ有りの機械学習により、リクエストの内容に対応したレスポンスの発話データが得られるように生成されている。

【0042】

また、生成部1023は、機械学習により得られたシステム応答モデルを用いるのではなく、予め作成しておいた文章の雛形（文章テンプレート）に対して、外部API（Application Programming Interface）を用いて、必要な情報を入手して当てはめて文章を完成させる手法を用いても良い。

例えば、生成部1023は、リクエストが電車の時刻である場合、出発駅と到着駅と、所定の時刻との情報により、時刻表検索の外部APIを用いて、電車の時刻を検索し、検索結果を雛形の所定の位置に挿入して、レスポンスの発話データを生成する。

【0043】

また、本実施形態においては、機械学習によるシステム応答モデルの手法、あるいは、外部APIの情報により、文章の雛形（文章テンプレート）を埋める手法のいずれを用いても良い。

また、本実施形態においては、リクエストと、このリクエストに対応したレスポンスとが予め書き込まれたデータベースを有している構成としても良い。この場合には、対話管理部1022がリクエストに対応するレスポンスを、データベースから抽出するため、生成部1023を備える必要は無い。

生成部1023は、生成した発話データとともに、少なくとも聴取志向フレーズフラグのデータを、聴取志向推定部103に対して出力する。

【0044】

図1に戻り、聴取志向推定部103は、ユーザの属性情報（ユーザ属性情報）や対話履歴などから、このユーザの聴取志向を推定し、発話データを音声としてユーザに供給する際の提示指針を推定する。

図4は、本実施形態における聴取志向推定部103の構成例を示すブロック図である。図4において、モデル構築部1031及び聴取志向管理部1032の各々を備えている。

モデル構築部1031は、ユーザの聴取志向を推定して、提示する際に発話データ及び

10

20

30

40

50

音声の特性を変更する指針（例えば、聴取志向のパラメータである音量、読み上げ速度及び区切りなどの調整量）を決定する処理を行う聴取志向推定モデルを生成する。本実施形態においては、聴取志向のパラメータを単にパラメータと示す場合もある。

【0045】

聴取志向推定モデルは、ユーザの属性情報及び発話データのテキストデータを入力することにより、聴取志向における変更対象のパラメータと、この変更対象のパラメータの変更量、あるいは置き換える他の単語を推定結果として出力する。

また、聴取志向テンプレートモデルは、聴取志向推定モデルと同様に、グループの属性情報及び発話データのテキストデータを入力することにより、聴取志向における変更対象のパラメータと、この変更対象のパラメータの変更量、あるいは置き換える他の単語を推定結果として出力する。

10

【0046】

ここで、モデル構築部1031は、上記聴取志向推定モデルとして、聴取志向を推定するための数式、あるいはルールを生成（構築）し、順次更新していく。本実施形態において、聴取志向は、ユーザの聴力に対応するパラメータとして、ユーザが発話データを音声合成した音声を聞いた際、聞き取り易いやすい音声の周波数、再生速度、再生する文節の区切り、この区切りの時間幅などを示している。

【0047】

モデル構築部1031が聴取志向推定モデルを生成する際、機械学習、教科学習、ニューラルネットワークなどのアルゴリズムを用いて、後述するように、対話履歴記憶部107及び対話行動記憶部109の各々の対話の内容や聴取志向のパラメータの変更における履歴のデータに基づいて、聴取志向推定モデルにおける聴取志向のパラメータやこのパラメータの変更量を推定する推定アルゴリズムにおける最適化を行う。

20

【0048】

すなわち、モデル構築部1031は、推定に必要な数式やルールとして、置き換えの単語の候補、発話データ全体の文章における聴取志向に関するパラメータである周波数、読み上げ速度及び区切りなどを推定するためのパラメータ推定用の基底関数を準備し、対話履歴記憶部107及び対話行動記憶部109の各々の対話の内容や聴取志向のパラメータの変更における履歴のデータを教師データとして、聴取志向推定モデル（あるいは、後述する聴取志向テンプレートモデル）の構築、あるいは更新を行う。

30

【0049】

また、聴取志向には、ユーザが単語の意味を理解できるか否かの知識力も含まれ、ユーザが理解できる一般的な同義語（あるいは類義語、類語）の他の単語に置き換える（変更する）こともパラメータの一つとして含まれる。

また、聴取志向には、上述した発話の周波数に対応して、発話データにおける単語に擦過音あるいは破裂音が含まれている場合、読み上げる際の周波数を低くしても、音声となった場合に高い周波数を含むことになるため、同義語（あるいは類義語、類語）であり、擦過音及び破裂音を含まない他の単語に置き換えることもパラメータの一つとして含まれる。

【0050】

上述した聴取志向推定モデルは、例えば、各ユーザのユーザ属性情報に対応させて、聞き取り易いやすい音声の周波数、再生速度、再生する文節の区切りの頻度、この区切りの時間幅などの各パラメータの調整量を設定した、また専門的或いは難解な単語をユーザに理解可能となるように置き換える、一般的に用いられる同義語であり、理解可能な他の単語の設定、さらに擦過音あるいは破裂音が含まれている単語を、同義語（あるいは類義語、類語）であり、擦過音及び破裂音を含まない他の単語の設定などを、置き換えリストをデータベースとして構成（データベース構成）しても良い。

40

【0051】

また、聴取志向推定モデルは、教師データを用いて機械学習を行う機械学習モデルとして構成（機械学習モデル構成）しても良い。この機械学習モデル構成の場合、モデル構築

50

部 1 0 3 1 は、対話履歴記憶部 1 0 7、ユーザ属性記憶部 1 0 8、対話行動記憶部 1 0 9 及びグルーピング記憶部 1 1 0 の各々に蓄積されたデータを教師データとして用いた機械学習により、ユーザ毎の聴取志向における各パラメータの変更の処理を推定する聴取志向推定モデルを生成する。

【 0 0 5 2 】

聴取志向管理部 1 0 3 2 は、モデル構築部 1 0 3 1 がユーザ毎に生成した（導出した）聴取志向推定モデルを用いて、ユーザの各々に対応した発話データの変更処理の内容を、提示制御部 1 0 4 に対して出力する。ここで、変更処理は、上述した音声の周波数、再生速度、再生する文節の区切りの頻度、この区切りの時間幅などの調整、及び聞き取りやすい発音となる単語への置き換えを行う処理を示している。

10

【 0 0 5 3 】

また、聴取志向管理部 1 0 3 2 は、リクエストの音声データを入力したユーザに対して、このユーザに対応する聴取志向推定モデルが生成されていない場合がある。この場合、聴取志向管理部 1 0 3 2 は、予めテンプレートとして準備されている聴取志向テンプレートモデルを用いて、聴取志向推定モデルが生成されていないユーザに対応した発話データの変更処理の内容を、提示制御部 1 0 4 に対して出力する。

また、聴取志向管理部 1 0 3 2 は、後述するグルーピング情報などを用いて類似したユーザ群の聴取志向テンプレートモデルを用いて、聴取志向推定モデルが生成されていないユーザに対応した発話データの変更処理の内容を、提示制御部 1 0 4 に対して出力する構成としても良い。

20

【 0 0 5 4 】

提示制御部 1 0 4 は、聴取志向推定部 1 0 3 から供給される発話データの変更内容における単語の置き換えに関し、言語知識記憶部 1 1 1 に記憶されている置き換えテーブルにより、対象となる単語を置き換える他の表現の単語を抽出する。この置き換えテーブルは、単語と、この単語に置き換える同義の単語との対応関係を示している。例えば、提示制御部 1 0 4 は、すでに述べたように、「今週」に対して「今度」、「週末」に対して「土曜日或いは日曜日」など、擦過音や破裂音を有する単語を、擦過音や破裂音の無い単語に置き換える処理を、置き換えテーブルを参照して行う。

【 0 0 5 5 】

そして、提示制御部 1 0 4 は、単語の置き換えを行った発話データを、音声合成部 1 0 5 に対して出力する。

30

また、聴取志向推定部 1 0 3 は、単語の置き換えのみでなく、破裂音または擦過音を含む文章（文節）を、破裂音及び擦過音を含まない同義（類義）の文章に置き換えるように構成しても良い。

【 0 0 5 6 】

ここで、言語知識記憶部 1 1 1 には、聴取志向のパラメータとして、聞き取り易さに対する言語的な知見に基づき、理解しにくい単語の同義語であって一般的に用いられて理解し易い他の単語、擦過音または破裂音を含む単語の同義語であって擦過音及び破裂音を含まない他の単語（上述したように文節でも良い）が蓄積されている。

【 0 0 5 7 】

40

例えば、医療従事者や介護士が年齢の高い人間（高齢者）と、対話する際に高齢者に理解させるために用いる単語の言い換えに関する知見、コーパス（テキストや発話を大規模に集めてデータベース化した言語資料）、同義語（類義語、類語）の辞書、シソーラス（言葉の上位概念及び下位概念）などを用いて、所定の単語に対して置き換える他の単語との組み合わせとして、言語知識記憶部 1 1 1 に対して予め、あるいは追加して書き込んで蓄積する。

【 0 0 5 8 】

音声合成部 1 0 5 は、提示制御部 1 0 4 から供給される発話データを、変更するパラメータと、このパラメータの調整量に対応して、発話データのテキストデータを音声合成により、ユーザに対するレスポンスとしての音声コンテンツを生成する。このとき、音声合

50

成部 105 は、例えば、ユーザの聴力に対応する聴取志向のパラメータ、及びその調整量として、ユーザが発話データを音声合成した音声を聞いた際、聞き取り易いやすい音声の周波数、再生する際の読み上げ速度、再生する文節の区切り、この区切りの時間幅などを変更して音声合成を行う。

そして、音声合成部 105 は、音声合成により生成した音声コンテンツを、データ入出力部 101 を介して、ユーザ端末 11 に対して出力する。

【0059】

グルーピング推定部 106 は、新たに履歴の発生したユーザの属性データに対応するグループを、グループ属性テーブルにより検索する。

そして、グルーピング推定部 106 は、グルーピング記憶部 110 において、上記ユーザを検索したグループのグループテーブルに追加して書き込んで記憶させる。

【0060】

図 5 は、ユーザ属性記憶部 108 に記憶されているユーザ属性テーブルの構成例を示す図である。図 5 において、ユーザ属性テーブルは、レコード毎に対して、ユーザ ID、年齢、性別、音量、読み上げ速度、区切り、設置環境、... などのユーザ属性の項目の欄が設けられている。ユーザ ID は、ユーザ端末 11 を用いて情報提示システム 1 を利用しているユーザの各々を識別するための識別情報である。年齢は、対応するユーザ ID で識別されるユーザの年齢を示している。性別は、対応するユーザ ID で識別されるユーザが男性 (male) か女性 (female) であることを示している。

【0061】

また、音量は、対応するユーザ ID で識別されるユーザが、聞き取り易い (聞き取りが可能な) とする音声の音量のレベル (大、中、小) を示している。読み上げ速度は、対応するユーザ ID で識別されるユーザが、発話として聞き取り易いとする音声の速度のレベル (早い、普通、遅い) を示している。区切りは、対応するユーザ ID で識別されるユーザが、聞き取り易いとする発音する文節の区切りを設ける数の量 (多い、普通、少ない) を示している。

【0062】

また、設置環境は、ユーザ端末 11 が設置されている場所、すなわち音声を聞き取る際の環境が、部屋が広くて音声が伝搬し易いか、部屋が小さくて反響し易いか、他の音が混入する可能性が低いかなどのユーザの音声の聞き取り環境を示している。また、ユーザ属性記憶部 108 には、ユーザ毎にユーザ ID に対応して聴取志向推定モデルが書き込まれて記憶されている。

【0063】

図 6 は、対話行動記憶部 109 に記憶されている対話行動テーブルの構成例を示す図である。図 6 において、ユーザ属性テーブルは、レコード毎に対して、時刻、ユーザ ID、アクションタイプ、アクション ID、実施内容、メッセージ ID、... などの項目の欄が設けられている。時刻は、発話データに対する何らかの変更を加える処理 (アクション) が行われた時刻を示している。ユーザ ID は、ユーザ端末 11 を用いて情報提示システム 1 を利用しているユーザの各々を識別するための識別情報である。アクションタイプは、システム側が主導して行ったシステム主導のアクション (active) か、あるいはユーザ側からの要求に対応して行われたユーザ主導のアクション (passive) かのいずれかであることを示している。

【0064】

また、アクション ID は、システム主導のアクションあるいはユーザ主導のアクションの各々の変更の種類を識別する識別情報である。図 6 においては、例えば、アクション ID : A001 が「単語の置き換え」であり、アクション ID : A003 が「読み上げ速度の変更」、アクション ID : A004 が「区切りの変更」を示している。実施内容は、アクションとして実際に発話データに対して実施した変更の内容を示している。図 6 において、アクション ID : A001 の例としては、「今週 今度」が「今週」という単語を「今度」とする類似単語に置き換え、「週末 土曜、日曜」が「週末」という単語を「土曜

10

20

30

40

50

、「日曜」とする類似単語（意味が類似した単語）に置き換えていることを示している。

【0065】

ここで、「周」の「shu」の発音は擦過音であり、高い周波数の成分が含まれるため、高い周波数が聞き取り難いユーザに対しては、擦過音を含まない類似単語に置き換える必要がある。また、高い周波数が聞き取り難いユーザに対しては、破裂音を含む単語も高い周波数を含むことになるので、擦過音の場合と同様に、破裂音を含まない類似単語に置き換える必要がある。

【0066】

また、アクションID：A003の例としては、「速度：-」が、発話の読み上げ速度を低下させた処理を示している。また、アクションID：A004の例としては、「区切り箇所：+」が、発話データにおける文節の間に所定の時間を設け、すなわち読み上げる際に一つの文節を読み上げた後に、所定の時間（間）を置いて次の文節を読み上げる頻度を増加させることを示している。

メッセージIDは、同一のレコードにおける、アクションが行われたメッセージを指し示す識別情報であり、図3におけるメッセージID1と同一の識別情報である。

【0067】

図7は、図6に示したアクションにおける単語の置き換えの処理を説明する概念図である。図7においては、話者であるユーザ及びユーザ端末11の各々が発話する音声を、文字データとして可視化して説明する。

図7(a)は、システム主導のアクションとしての単語の置き換えを示している。ユーザ301がユーザ端末11に対して音声により、ユーザが吹き出し（speech balloon）351の「 は？」というリクエスト（質問）をした際、情報提示サーバ10が吹き出し451の「今週の土曜日...」という発話データをレスポンスとして回答する。このとき、聴取志向管理部1032は、ユーザ301に対応した聴取志向推定モデルを参照しているが、このユーザ301に対して、上記発話データに対して変更の処理を行うことが記載されていないため、対話処理部102が供給する発話データをそのままレスポンス（回答）としている。

【0068】

一方、ユーザ302がユーザ端末11に対して音声により、吹き出し351の「 は？」というリクエスト（質問）をした際、情報提示サーバ10が上記吹き出し451の「今週の土曜日...」という発話データを、吹き出し452の「今度の土曜日...」と変更した後、レスポンスとして回答する。このとき、聴取志向管理部1032は、ユーザ301に対応した聴取志向推定モデルを参照し、このユーザ302に対して、上記発話データに対して変更の処理を行うことが記載されているため、対話処理部102が供給する発話データを、聴取志向推定モデルに対応して変更処理を行っている。

すなわち、属性情報において、ユーザ301（例えば、年齢20代）に比較してユーザ302（例えば、年齢70代）の年齢が高く、ユーザ302は周波数が低い音声の方が聞き易いため、聴取志向推定モデルには破裂音や擦過音を含む単語の置き換えの処理が設定されている。

【0069】

図7(b)は、ユーザ主導のアクションとしての単語の置き換えを示している。図示はしていないが、ユーザ303（例えば、年齢70代）がユーザ端末11に対して音声による「 は？」というリクエスト（質問）をした際、情報提示サーバ10が上記吹き出し453の「今週の土曜日...」という、対話処理部102が供給する発話データをそのままレスポンス（回答）としている。しかしながら、「今週の土曜日...」の音声に対して、ユーザ303が吹き出し353の「えっ？/もう一度」という、聴取志向フレーズのリクエスト（要求）が入力される。

【0070】

このため、聴取志向管理部1032は、ユーザ301に対応した聴取志向推定モデルを参照し、このユーザ302に対して、上記聴取志向フレーズに対応して、対話処理部10

10

20

30

40

50

2 が供給する発話データに変更処理を行っている。これにより、情報提示サーバ 10 は、発話データの吹き出し 453 の「今週の土曜日...」が、吹き出し 454 の「今度の土曜日...」に変更された音声のデータをユーザ端末 11 に対して再度出力する。

【0071】

図 8 は、グルーピング記憶部 110 に記憶されているグルーピングテーブルの構成例を示す図である。図 8 (a) は、グループの属性情報を示すグループ属性情報テーブルの構成例を示している。図 8 (a) において、グループ属性情報テーブルは、一例として、レコード毎に対して、グループ ID、年代、性別及び居住地などの項目の欄が設けられている。グループ ID は、グループの各々を識別するための識別情報である。年代は、グループを構成するユーザの年齢の範囲を示している。

10

【0072】

例えば、グループ ID : G_001 は、少なくとも年齢が 60 歳から 75 歳までの範囲に含まれるユーザの集合体であることを示している。同様に、グループ ID : G_002 は、少なくとも年齢が 10 歳から 20 歳までの範囲に含まれるユーザの集合体であることを示している。性別は、対応するグループ ID で識別されるグループを構成する人間の性別が男性 (male) か女性 (female) であることを示している。居住地は、対応するグループ ID で識別されるグループを構成するユーザの居住地がいずれの地方であるかを示している。

また、このグループ ID で識別されるグループ毎には、それぞれのグループを構成するユーザの上述した属性に対応する聴取志向テンプレートモデルがグルーピング記憶部 110 に対して予め書き込まれて記憶されている。

20

【0073】

図 8 (b) は、グループ ID の各々に属すユーザが割り当てられているグルーピングテーブルの構成例を示している。各レコードには、グループ ID、ユーザ ID、年齢、性別、音量、読み上げ速度、区切り、設置環境の各々の欄が設けられている。グループ ID は、グループの各々を識別するための識別情報である。ユーザ ID は、同一レコードにおけるグループ ID の示すグループに分類されたユーザを示す識別情報であり、図 5 におけるユーザ属性テーブルのユーザ ID と同一の識別情報である。

【0074】

また、年齢は、対応するユーザ ID で識別されるユーザの年齢を示している。性別は、対応するユーザ ID で識別されるユーザが男性 (male) か女性 (female) であることを示している。ここで、性別がグルーピングにおける属性に含まれていない場合、そのグループは男性 (male) か女性 (female) の双方のユーザが存在する。

30

【0075】

また、音量は、対応するユーザ ID で識別されるユーザが、聞き取り易い (聞き取りが可能な) とする音声の音量のレベル (大、中、小) を示している。読み上げ速度は、対応するユーザ ID で識別されるユーザが、発話として聞き取り易いとする音声の速度のレベル (早い、普通、遅い) を示している。区切りは、対応するユーザ ID で識別されるユーザが、聞き取り易いとする発音する文節の区切りを設ける数の量 (多い、普通、少ない) を示している。

40

【0076】

また、設置環境は、ユーザ端末 11 が設置されている場所、すなわち音声を聞き取る際の環境が、部屋が広くて音声が伝搬し易いか、部屋が小さくて反響し易いか、他の音が混入する可能性が低いのか、他の音が混入する可能性が高いかなどのユーザの音声の聞き取り環境を示している。

【0077】

上述したように、グループの各々は、グルーピングテーブルにおいて規定されているグループの属性 (グループ属性) と同様の属性を有するユーザが分類されている。

そして、上述したグルーピングにおける属性の種類は、人間の音声の聞き取り易さに詳しい学者や医者、あるいは介護施設の職員 (看護師や介護士など) の聴取志向に詳しい識

50

者の提示する属性の種類を用いても良い。

また、グルーピングにおける属性の種類は、複数のユーザの属性を特徴量としてクラスタリングなどの処理を行い、最も明確にユーザ動詞を分類できる特徴量の属性の種類を抽出する処理により設定しても良い。

【0078】

上述したいずれの処理により、グルーピングに用いる属性の種類を抽出したとしても、上記識者の治験に対応して、聴取志向テンプレートモデルの聴取志向における音量、読み上げ速度及び区切りなどのパラメータの変更の要否、変更する際のそれぞれのパラメータの調整量を設定しても良い。

本実施形態における情報提示サーバ10の利用を開始した直後のユーザに対し、情報提示サーバ10が上述した聴取志向のパラメータの変更の要否や、変更する際のパラメータの調整量のデータを、音声に対する聴取志向に対するユーザの対応から十分に抽出できていない。

【0079】

このため、聴取志向推定部103は、聴取志向のパラメータのデータが十分に抽出できていないユーザに対し、このユーザの属性に近いグループを上記グルーピングテーブルにおいて検索し、検索して得られたグループの聴取志向テンプレートモデルを用いて、聴取志向のパラメータの要否あるいはパラメータの変更量を推定する。

そして、聴取志向推定部103は、ユーザの属性に用いた聴取志向テンプレートモデルを元に、聴取志向における各パラメータの変更の要否及び変更の際の調整量のデータを、ユーザからの音声に対する変更の要求から取得して、ユーザの各々の聴取志向推定モデルとする処理を行う。このとき、聴取志向推定部103は、すでに述べたように、聴取志向テンプレートモデルに対して、機械学習による最適化の処理を行うことで聴取志向推定モデルを生成しても良い。

【0080】

図9は、本実施形態の情報提示システムを用いた対話システムの動作例を示すフローチャートである。この図9のフローチャートの動作は、例えば、情報提示システム1における情報提示サーバ10に対してアクセスし、ユーザがスマートスピーカなどのユーザ端末11から音声によるリクエストを音声により情報提示サーバ10送信して、情報提示サーバ10との対話を行う際に開始される。以下の図9のフローチャートの動作説明は、グループ毎の聴取志向に対応した聴取志向テンプレートモデルの各々が、聴取志向推定部103において、すでに説明したように生成されて、グルーピング記憶部110に蓄積されている状態において行う。

【0081】

ステップS101:

データ入出力部101は、いずれかのユーザ端末11から音声データが供給されたか否かの判定を行う。そして、データ入出力部101は、いずれかのユーザ端末11から音声データが供給された場合、処理をステップS2へ進める。一方、データ入出力部101は、いずれかのユーザ端末11からも音声データが供給されない場合、ステップS101の処理を繰り返す。

【0082】

このとき、例えば、ユーザがユーザ端末11に対して音声により、コンサート等が行われる日などの予定を問い合わせるリクエストを入力する。そして、ユーザ端末11は、音声データとこの音声を入力したユーザのユーザIDとの各々を、情報提示サーバ10にアクセスして送信する。この場合、データ入出力部101は、いずれかのユーザ端末11から音声データが供給されたことを検出し、処理をステップS102へ進める。

そして、ステップS102に進める際、データ入出力部101は、入力した音声データを対話処理部102に対して出力する。また、データ入出力部101は、入力したユーザIDを聴取志向推定部103に対して出力する。

【0083】

10

20

30

40

50

ステップ S 1 0 2 :

聴取志向推定部 1 0 3 は、データ入出力部 1 0 1 からユーザ ID が供給された場合、このユーザ ID の示すユーザに対話の履歴があるか否かの判定を行う。すなわち、聴取志向推定部 1 0 3 は、ユーザ属性記憶部 1 0 8 を参照して、このユーザ ID に対応して聴取志向推定モデルが記憶されているか否かの判定を行う。すなわち、ユーザに対話の履歴が無ければ、聴取志向テンプレートモデルから聴取志向推定モデルが生成されていない。

このとき、聴取志向推定部 1 0 3 は、ユーザ属性記憶部 1 0 8 にユーザに対応する聴取志向推定モデルが記憶されている場合、処理をステップ S 1 0 3 へ進める。一方、聴取志向推定部 1 0 3 は、ユーザ属性記憶部 1 0 8 にユーザに対応する聴取志向推定モデルが記憶されていない場合、処理をステップ S 1 0 4 へ進める。

10

【 0 0 8 4 】

ステップ S 1 0 3 :

聴取志向推定部 1 0 3 は、ユーザ属性記憶部 1 0 8 からユーザ ID に対応する聴取志向推定モデルを読み出す。

【 0 0 8 5 】

ステップ S 1 0 4 :

聴取志向推定部 1 0 3 は、ユーザ属性記憶部 1 0 8 を参照し、ユーザ ID に対応したユーザの属性情報を読み出す。

そして、聴取志向推定部 1 0 3 は、読み出した属性情報に近い属性情報を有するグループをグルーピング記憶部 1 1 0 のグループ属性情報テーブルから検索し、検索して得られたグループの聴取志向テンプレートモデルを読み出す。

20

また、グルーピング推定部 1 0 6 は、グルーピング記憶部 1 1 0 において、上記ユーザを検索したグループのグループテーブルに追加して書き込んで記憶させる。

【 0 0 8 6 】

ステップ S 1 0 5 :

対話処理部 1 0 2 は、音声データをテキストデータに変換し、形態素解析を行って、得られた単語あるいは文節から、この音声データが会話フレーズであるか、あるいは聴取志向フレーズであるかの判定を行う。音声データが聴取志向フレーズであるということは、ユーザが発話データ（レスポンス）の音声の最適化（自身の聴取志向に合わせる変更）を要求していることを意味している。

30

【 0 0 8 7 】

したがって、対話処理部 1 0 2 は、このステップ S 1 0 5 において、ユーザが発話データの音声の最適化を要求しているか否かの判定を行っている。

そして、対話処理部 1 0 2 は、ユーザが発話データの音声の最適化を要求していない場合、処理をステップ S 1 0 6 へ進める。一方、対話処理部 1 0 2 は、ユーザが発話データの音声の最適化を要求している場合、処理をステップ S 1 0 7 へ進める。

このとき、対話処理部 1 0 2 は、対話履歴記憶部 1 0 7 における対話履歴テーブルに対し、入力された音声データのテキストデータ、聴取志向フレーズの場合に聴取志向フレーズのフラグ、メッセージ ID の各々の書き込みを行う。

40

【 0 0 8 8 】

ステップ S 1 0 6 :

入力された音声データが会話フレーズであるため、対話処理部 1 0 2 は、このリクエストの音声データに対応した発話データの生成を、音声データのテキスト文を形態素解析した単語の各々を用いて行う。

そして、聴取志向推定部 1 0 3 は、聴取志向推定モデルあるいは聴取志向テンプレートモデルにより、システム主導の発話データに対する変更処理の推定、すなわち、ユーザの聴取志向のパラメータのなかから変更対象のパラメータと、変更量（あるいは単語の置き換え）を推定する。

また、聴取志向推定部 1 0 3 は、聴取志向のパラメータのなかから選択した変更対象のパラメータと、このパラメータの変更量（あるいは置き換える単語）とを、提示制御部 1

50

04に対して出力する。

【0089】

ステップS107:

入力された音声データが聴取志向フレーズであるため、この時点においては、このフローチャートにおける前回の会話フレーズのループにおいて、リクエストに対するレスポンスとしての会話フレーズはすでに得られている。

このため、聴取志向推定部103は、聴取志向推定モデルあるいは聴取志向テンプレートモデルにより、聴取志向のパラメータのなかから変更対象のパラメータと、このパラメータの変更量を調整して、提示制御部104に対して出力する。

【0090】

このとき、聴取志向推定部103は、対話行動記憶部109の対話行動テーブルに対して、単語の置き換えを行った処理を書き込んで記憶させる。このとき、聴取志向推定部103は、アクションタイプとしてシステム主導で行ったか、あるいはユーザ主導で行ったかのいずれかを記載する。また、聴取志向推定部103は、予め行動の各々に付されているアクションIDを記載し、アクションIDに対応した実施内容を記載する（記載例としては図6の対話行動テーブルを参照）。実施内容が単語の置き換え（アクションID：A001）の場合、提示制御部104がどの単語をどのような単語に置き換えたかを、対話行動テーブルの実施内容の欄に記載する。

【0091】

ここで、例えば、変更対象のパラメータが音量である場合、予め通常の音量からの変更量と規定されている大きさに対して、より大きい音量を変更量とする（変更量の調整）。また、変更対象のパラメータが読み上げ速度である場合、予め通常の読み上げ速度からの変更量と規定されている遅い速度に対して、より遅い速度を変更量とする。また、変更対象のパラメータが区切りである場合、予め通常の区切りの頻度からの変更量と規定されている区切りの頻度に対して、より多くの区切りの頻度を変更量とする。

また、このパラメータの各々は、一括して変更量を変更してもよいし、フローチャートのループが繰り返される毎に、変更する順番を決めておいて、変更量の調整を行っても良い。

【0092】

ステップS108:

提示制御部104は、聴取志向推定部103から供給される聴取志向における単語の置き換え処理の要求に対応し、発話データのテキストデータに含まれる擦過音及び破裂音を有する単語の各々を抽出する。そして、提示制御部104は、抽出した単語の各々に対応した置き換える単語を、言語知識記憶部111の置き換えテーブルを参照して、それぞれ抽出する。

そして、提示制御部104は、聞き取りやすい単語への置き換えを終了した発話データを、聴取志向のパラメータとそのパラメータの変更量との各々を、音声合成部105に対して出力する。

【0093】

音声合成部105は、提示制御部104から供給される発話データを、変更するパラメータと、このパラメータの調整量に対応して、発話データのテキストデータを音声合成により、ユーザに対するレスポンスとしての音声コンテンツを生成する。

そして、音声合成部105は、音声合成により生成した音声コンテンツを、データ入出力部101を介して、ユーザ端末11に対して出力する。

【0094】

ステップS109:

聴取志向推定部103は、対話行動記憶部109の対話行動テーブルにおけるメッセージIDを参照し、このメッセージIDに連続するメッセージIDを対話履歴記憶部107の対話履歴テーブルから抽出する。

そして、聴取志向推定部103は、抽出したメッセージに対応するメッセージ本文の聴

10

20

30

40

50

取志向フレーズフラグが「0」である場合に、聴取志向のパラメータの変更あるいは単語の置き換えが成功したと判定する。一方、聴取志向推定部103は、抽出したメッセージに対応するメッセージ本文の聴取志向フレーズフラグが「1」である場合に、聴取志向のパラメータの変更あるいは単語の置き換えが、聞き取り易さを向上させるために不十分であると判定する。

【0095】

聴取志向推定部103は、例えば、上述した聴取志向フレーズフラグが「1」であり、かつユーザ主導により変更した聴取志向におけるパラメータと、このパラメータの変更量とにより、ユーザに対応する聴取志向推定モデルを、よりユーザの聴取志向に適合させる修正処理を行う。

また、聴取志向推定部103は、グルーピング記憶部110のグループテーブルを参照し、グループを構成するユーザの各々に共通する変更された聴取志向におけるパラメータと、パラメータの変更量とを抽出し、聴取志向テンプレートモデルを、よりグループに含まれるユーザの聴取志向に適合させる修正処理を行う。

【0096】

このとき、聴取志向推定部103は、例えば、聴取志向フレーズフラグが立っているメッセージIDに対応するメッセージ本文の形態素解析を行い、ポジティブワードあるいはネガティブワードを抽出し、ポジティブワードの場合、変更に対する評価値に「1」を加算（評価値をインクリメント）する処理を行い、一方、ネガティブワードの場合、変更に対する評価値から「1」を減算（評価値をデクリメント）する処理を行う。そして、聴取志向推定部103は、評価値が所定の閾値を超えた場合、変更した聴取志向のパラメータの変更量（あるいは置き換えた単語）を、聴取志向推定モデル及び聴取志向テンプレートモデルに反映させるように構成しても良い。

【0097】

また、聴取志向推定部103は、例えば、ネガティブワードやポジティブワードの抽出を行うのではなく、聴取志向のパラメータを変更して音声コンテンツを出力した後に、「聞き取り易かったですか？ 「はい」 / 「いいえ」でお答え下さい」や、「もう少しゆっくり読み上げましょうか？ 「このまま」 / 「ゆっくり」でお答え下さい」のテキストデータを、音声合成部105により音声合成して確認音声コンテンツに変更する。また、聴取志向推定部103は、この確認音声コンテンツをユーザ端末11に対してデータ入出力部101を介して送信する。このアルゴリズムは、ユーザ主導の聴取志向のパラメータの変更に対応している。

【0098】

そして、聴取志向推定部103は、上述した確認音声コンテンツに対するユーザの回答を入力する。このとき、聴取志向推定部103は、対話処理部102がユーザによる回答の音声データをテキスト変換した回答データを入力する。

そして、聴取志向推定部103は、例えば、「聞き取り易かったですか？」の質問に対する回答データが「はい」の場合、聴取志向のパラメータの変更が成功したと判定する。一方、「聞き取り易かったですか？」の質問に対する回答データが「いいえ」の場合、聴取志向のパラメータの変更が成功しなかったと判定する。

【0099】

これにより、聴取志向推定部103は、成功した場合に成功した聴取志向のパラメータの変更処理を、聴取志向推定モデル及び聴取志向テンプレートモデルに反映させる。

一方、聴取志向推定部103は、変更が失敗した場合、再度、聴取志向の他のパラメータの変更を行った音声コンテンツを生成して、ユーザに対してレスポンスとして出力する。

【0100】

また、聴取志向推定部103は、例えば、「もう少しゆっくり読み上げましょうか？」の質問に対する回答データが「このまま」の場合、聴取志向のパラメータである読み上げ速度の変更が成功したと判定する。一方、「もう少しゆっくり読み上げましょうか？」の

10

20

30

40

50

質問に対する回答データが「ゆっくり」の場合、聴取志向のパラメータである読み上げ速度の変更量が少ないため成功しなかったと判定する。

これにより、聴取志向推定部 103 は、成功した場合に成功した聴取志向のパラメータである読み上げ速度の変更処理を、聴取志向推定モデル及び聴取志向テンプレートモデルに反映させる。

【0101】

一方、聴取志向推定部 103 は、変更が失敗した場合、再度、聴取志向のパラメータである読み上げ速度の変更量を増加させ、すなわちより読み上げ速度を低下させる変更を行った音声コンテンツを生成して、ユーザに対してレスポンスとして出力する。

上述したように、聴取志向のパラメータである周波数、読み上げ速度及び区切りや単語の置き換えなどの変更を行った後に、それぞれの変更が適切であったか否かの質問をユーザに与え、聴取志向のパラメータの変更の成功/不成功の確認を行い、この確認結果を聴取志向推定モデル及び聴取志向テンプレートモデルに反映させる構成としても良い。

【0102】

また、聴取志向推定部 103 は、聴取志向フレーズフラグが「1」となる発生頻度をカウントし、同様の聴取志向のパラメータの変更を行う発生頻度のカウント数が所定の設定値を超えた場合に、発生頻度が所定の設定値を超えたパラメータに基づき、このパラメータ及びパラメータの変更量を、聴取志向推定モデル及び聴取志向テンプレートモデルに反映させるように構成しても良い。

【0103】

ステップ S110 :

提示制御部 104 は、対話履歴記憶部 107 の対話履歴テーブルに対して、発話データのテキストデータを、メッセージ本文に書き込んでメッセージ ID 1 を付与して書き込んで記憶させる。このとき、提示制御部 104 は、話者 ID の欄に対して、レスポンスを行うシステムのシステム識別情報を書き込んで記憶させる。

また、提示制御部 104 は、会話フレーズであるため、聴取志向フレーズフラグを「0」とし、かつ接続されるユーザの音声データのメッセージ本文のメッセージ ID 1 をメッセージ ID 2 の欄に書き込んで記憶させる。

【0104】

上述した構成及び動作により、本実施形態によれば、ユーザのリクエストに対して、レスポンスを行う情報提示サーバ 10 が音声コンテンツにより提供する情報を、ユーザの各々が正確に聞き取ることができるように聴取志向の推定を、ユーザ毎の聴取志向推定モデルまたはグループ毎の聴取志向テンプレートモデルを用いて行うため、従来のようにルールベースで各ユーザあるいは各グループに対して聴取志向の推定を行う構成に比較してデータ量を少なくすることができ、かつデータ量が少ないために聴取志向推定モデル及び聴取志向テンプレートモデルの各々のメンテナンス（ユーザに順次対応させていく修正処理）を容易に行うことができる。

【0105】

また、本実施形態によれば、ユーザの各々の属性情報に対応した聴取志向推定モデルにより、ユーザの聴取志向における聞き取り易さを向上するパラメータの種類と、これらパラメータの変更量（調整量）とが求められ、ユーザのリクエストに対するレスポンスである発話データにおける擦過音あるいは破裂音を含む単語を抽出し、発話データの文脈に対応して同義語（あるいは類義語、類語）である擦過音及び破裂音を含まない他の単語に置き換えるため、発話データを音声合成した音声コンテンツを、ユーザが聞き取り易い音声とすることができる。

【0106】

また、本実施形態によれば、ユーザの各々の属性情報に対応した聴取志向推定モデルにより、ユーザの聴取志向における聞き取り易さを向上するパラメータの種類と、これらパラメータの変更量（調整量）とが求められ、ユーザのリクエストに対するレスポンスである発話データを音声合成する際、発話される音声の周波数、読み上げ速度、区切りなどの

10

20

30

40

50

変更を行うため、音声合成された発話データである音声コンテンツを、ユーザが聞き取り易い状態の音声とすることができる。

【0107】

また、本実施形態によれば、対話の履歴が無いユーザに対して、このユーザと属性情報が類似している他のユーザにより構成されるグループに対応して生成された聴取志向テンプレートモデルを用い、上述した発話データにおける擦過音あるいは破裂音を含む単語を、発話データの文脈に対応して同義語である擦過音及び破裂音を含まない他の単語に置き換えるため、発話データを音声合成した音声コンテンツを、ユーザが聞き取り易い音声とするため、履歴の無いユーザに対しても、レスポンスの音声コンテンツの聞き取り易さを向上させることができる。

10

【0108】

また、本実施形態によれば、対話の履歴が無いユーザに対して、このユーザと属性情報が類似している他のユーザにより構成されるグループに対応して生成された聴取志向テンプレートモデルを用い、ユーザのリクエストに対するレスポンスである発話データを音声合成する際、発話される音声の周波数、読み上げ速度、区切りなどの変更を行うため、音声合成された発話データである音声コンテンツを、ユーザが聞き取り易い状態の音声とするため、履歴の無いユーザに対しても、レスポンスの音声コンテンツの聞き取り易さを向上させることができる。

【0109】

また、本実施形態によれば、上記聴取志向推定モデル及び聴取志向テンプレートモデルの各々を、対話履歴記憶部107及び対話行動記憶部109に記憶されている、リクエスト側(ユーザ)とレスポンス側(情報提示サーバ10)との対話における履歴の各データを用いて順次変更を行うため、ユーザあるいはグループの属性情報に対応した音声の聞き取り易さを向上させていくことができる。

20

【0110】

本実施形態においては、レスポンス側をコンピュータの対話システムとして説明したが、リクエスト側とレスポンス側との各々がユーザ(人間)である場合、対話するユーザ間における相互の聞き取り易さを向上するように、ユーザそれぞれに対応した聴取志向推定モデルにより、対話におけるレスポンス側の発話データにおける単語の置き換えの処理、及び音声合成の際の聞き取り易さを向上するパラメータの変更処理を行う構成としても良い。

30

【0111】

また、図10は、本発明の一実施形態による、ユーザとシステムとが対話を行う情報提示システムの他の構成例を示す概念図である。

情報提示システム1Aは、情報提示サーバ10A、ユーザ端末11_1、ユーザ端末11_2、ユーザ端末11_3、ユーザ端末11_4、対話サーバ12_1、対話サーバ12_2、対話サーバ12_3、対話サーバ12_3の各々がネットワーク500を介して接続されている。

ユーザ端末11_1及びユーザ端末11_2の各々は、すでに説明したスマートスピーカなどであり、ユーザが音声によってリクエストの入力を行い、情報提示サーバ10Aからのレスポンスを音声コンテンツとしてユーザに通知する。

40

【0112】

一方、ユーザ端末11_3はスマートフォンやタブレットコンピュータなどの携帯端末であり、表示画面が備えられている。また、ユーザ端末11_4は、パーソナルコンピュータであり、表示画面が設けられている。

情報提示サーバ10Aは、表示画面を備えているユーザ端末11_3及びユーザ端末11_4の各々に対しては、音声コンテンツではなく、視覚(ビジュアル)的に視認できる画像コンテンツ(文字コンテンツ、動画像あるいはスタンプ画像など)に変更して(出力を切替えて)、リクエストに対するレスポンスとして出力するように構成しても良い。

【0113】

50

また、情報提示サーバ10Aは、すでに説明した図1における情報提示サーバ10と同様の構成であるが、対話処理部102における対話システムの機能を有していない構成である。

対話サーバ12_2、対話サーバ12_3及び対話サーバ12_3の各々は、情報提示サーバ10における対話処理部102の対話システムの機能に換わる装置である。対話サーバ12_2、対話サーバ12_3及び対話サーバ12_3の各々は、例えば、天気予報確認、交通機関の時刻確認、ユーザの計画の確認それぞれを行う対話システムである。

【0114】

この構成の場合、情報提示サーバ10Aは、対話サーバ12_2、対話サーバ12_3及び対話サーバ12_3の各々から、発話データとしてのレスポンスのテキストデータを
10
入力し、すでに述べたように、発話データを音声コンテンツとした際における聞き取り易さを向上する変更を行う。

【0115】

なお、本発明における図1の情報提示サーバ10及び図10の情報提示サーバ10Aの各々の機能を実現するためのプログラムをコンピュータ読み取り可能な記録媒体に記録して、この記録媒体に記録されたプログラムをコンピュータシステムに読み込ませ、実行することにより音声コンテンツをユーザがより聞き取り易いように変更する処理を行ってもよい。なお、ここでいう「コンピュータシステム」とは、OSや周辺機器等のハードウェアを含むものとする。

【0116】

また、「コンピュータシステム」は、ホームページ提供環境（あるいは表示環境）を備えたWWWシステムも含むものとする。また、「コンピュータ読み取り可能な記録媒体」とは、フレキシブルディスク、光磁気ディスク、ROM、CD-ROM等の可搬媒体、コンピュータシステムに内蔵されるハードディスク等の記憶装置のことをいう。さらに「コンピュータ読み取り可能な記録媒体」とは、インターネット等のネットワークや電話回線等の通信回線を介してプログラムが送信された場合のサーバやクライアントとなるコンピュータシステム内部の揮発性メモリ（RAM）のように、一定時間プログラムを保持しているものも含むものとする。
20

【0117】

また、上記プログラムは、このプログラムを記憶装置等に格納したコンピュータシステムから、伝送媒体を介して、あるいは、伝送媒体中の伝送波により他のコンピュータシステムに伝送されてもよい。ここで、プログラムを伝送する「伝送媒体」は、インターネット等のネットワーク（通信網）や電話回線等の通信回線（通信線）のように情報を伝送する機能を有する媒体のことをいう。また、上記プログラムは、前述した機能の一部を実現するためのものであっても良い。さらに、前述した機能をコンピュータシステムにすでに記録されているプログラムとの組み合わせで実現できるもの、いわゆる差分ファイル（差分プログラム）であっても良い。
30

【0118】

また、上記プログラムは、図1に記載のシステムおよびプログラムについて、対話サーバ12を含む情報提示サーバ10とユーザ端末11がネットワークを介して伝送を実現している。しかしながら、これに限らずに、例えば、可能であればネットワークを介することなくユーザ端末11に対話サーバ12を含む情報提示サーバ10の機能が搭載されていてもよい。
40

また、情報提示サーバ10と対話サーバ12との各々が独立した装置として設ける構成ではなく、情報提示サーバ10、対話サーバ12のそれぞれの機能を、1つのサーバによって実現する構成としても良い。

【符号の説明】

【0119】

- 1, 1A ... 情報提示システム
- 10, 10A ... 情報提示サーバ

10

20

30

40

50

- 1 1 , 1 1 _ 1 , 1 1 _ 2 , 1 1 _ 3 , 1 1 _ 4 ... ユーザ端末
- 1 2 _ 1 , 1 2 _ 2 , 1 2 _ 3 ... 対話サーバ
- 1 0 1 ... データ入出力部
- 1 0 2 ... 対話処理部
- 1 0 3 ... 聴取志向推定部
- 1 0 4 ... 提示制御部
- 1 0 5 ... 音声合成部
- 1 0 6 ... グルーピング推定部
- 1 0 7 ... 対話履歴記憶部
- 1 0 8 ... ユーザ属性記憶部
- 1 0 9 ... 対話行動記憶部
- 1 1 0 ... グルーピング記憶部
- 1 1 1 ... 言語知識記憶部
- 5 0 0 ... ネットワーク
- 1 0 2 1 ... 解析部
- 1 0 2 2 ... 対話管理部
- 1 0 2 3 ... 生成部
- 1 0 3 1 ... モデル構築部
- 1 0 3 2 ... 聴取志向管理部
- 1 0 2 3 ... 生成部

10

20

【 図 1 】

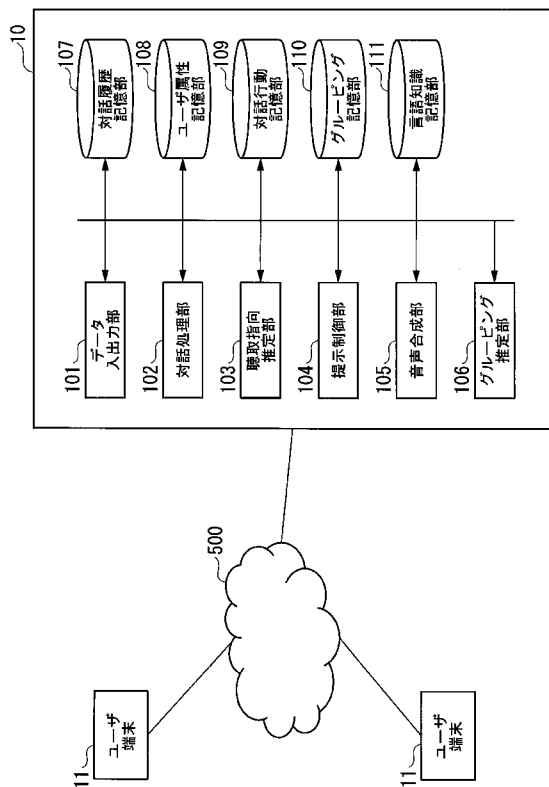


図 1

【 図 2 】

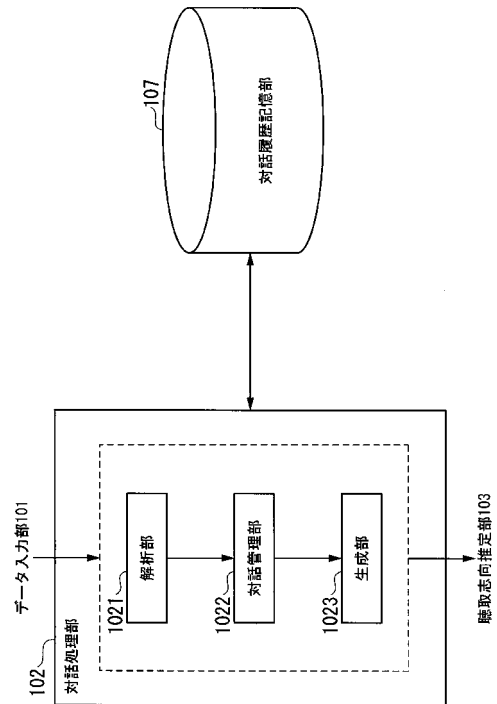


図 2

【 図 3 】

時刻	話者ID	メッセージ本文	聴取志向 フラグ	メッセージ ID1	メッセージ ID2
...
18/01/01 00:00:01	U_001	〇〇に行きたい	0	M1801010001	...
18/01/01 00:00:03	C_001	今週の週末は いかがですか？	0	M1801010003	M1801010001
18/01/01 00:00:10	U_001	聞き取れなかった からもう一回 お願いします	1	M1801010010	M1801010003
18/01/01 00:00:15	C_001	今週の土曜か 日曜に行きま せんか？	0	M1801010015	M1801010010
...

図3

【 図 4 】

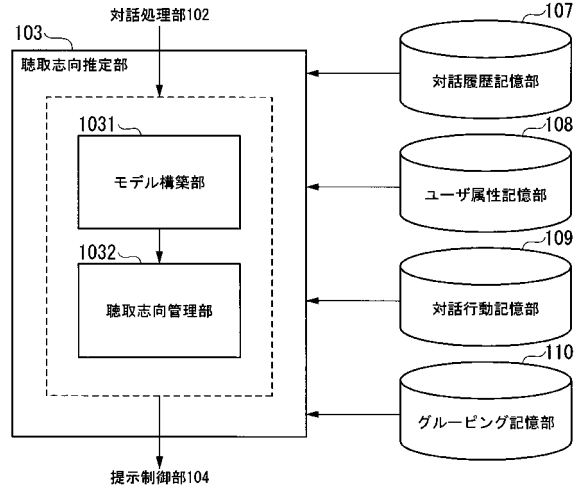


図4

【 図 5 】

ユーザID	年齢	性別	音量	読み上げ 速度	区切り	設置環境
...
U_001	63	male	大	遅い	多	リビング
U_002	73	female	大	遅い	普通	控室
U_003	18	female	小	普通	普通	自室
U_004	12	male	小	普通	普通	自室
...

図5

【 図 6 】

時刻	ユーザID	アクション タイプ	アクションID	実施内容	メッセージID
...
18/01/01 12:00:15	U_001	active (システム主導)	A001 (単語置き換え)	今週→今度	M18010120015
18/01/01 12:00:15	U_001	active	A001 (単語置き換え)	週末→土曜、日曜	M18010120015
18/01/01 15:24:57	U_002	passive (ユーザ主導)	A003 (読み上げ速度の 変更)	速度:-(マイナス)	M180101152457
18/01/01 15:24:57	U_002	passive	A004 (区切りの変更)	区切り箇所:+	M180101152457
...

図6

【図7】

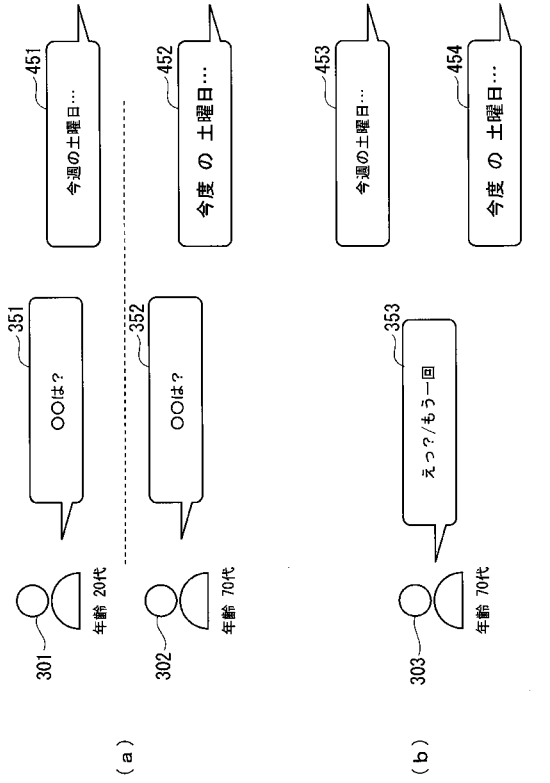


図7

【図8】

グループID	年代	性別	居住地
G_001	60-75	male	関東
G_002	10-20	female	関西

(a)

年齢	性別	音量	読み上げ速度	区切り	設置環境
63	male	大	遅い	多	リビング
73	female	大	遅い	普通	寝室
18	female	小	普通	普通	自室
12	male	小	普通	普通	自室

(b)

図8

【図9】

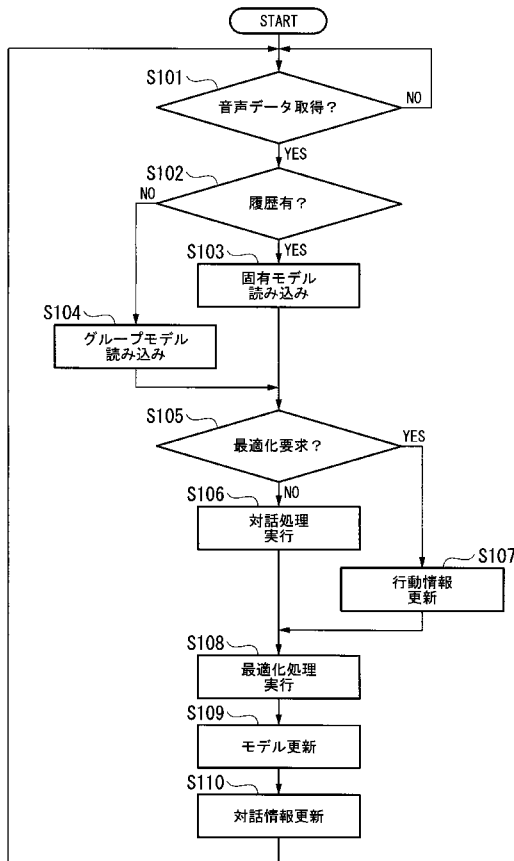


図9

【図10】

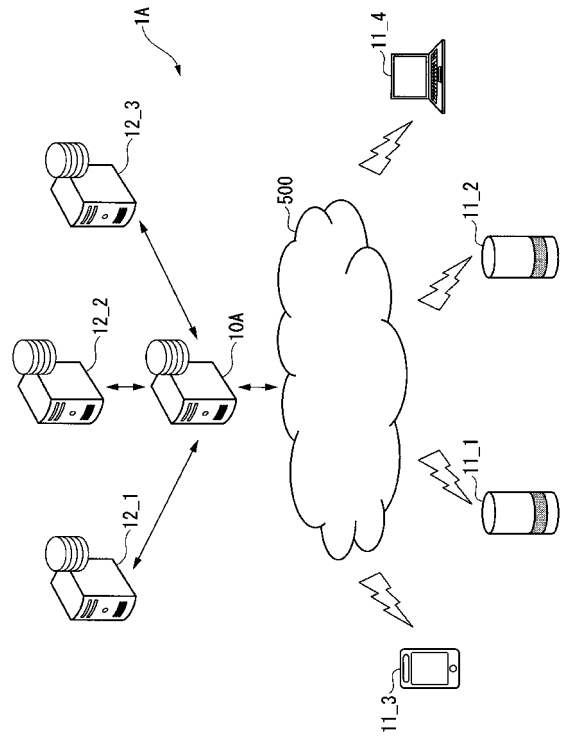


図10

フロントページの続き

(51) Int.Cl.			F I			テーマコード(参考)
G 1 0 L	13/08	(2013.01)	G 1 0 L	13/00	1 0 0 M	
G 1 0 L	13/10	(2013.01)	G 1 0 L	13/08	1 2 4	
			G 1 0 L	13/10	1 1 4	