

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第6962105号
(P6962105)

(45) 発行日 令和3年11月5日 (2021. 11. 5)

(24) 登録日 令和3年10月18日 (2021. 10. 18)

(51) Int. Cl.	F I
G 1 0 L 15/22 (2006. 01)	G 1 0 L 15/22 3 0 0 Z
G 1 0 L 15/00 (2013. 01)	G 1 0 L 15/00 2 0 0 A
G 1 0 L 15/28 (2013. 01)	G 1 0 L 15/28 2 3 0 Z

請求項の数 6 (全 25 頁)

(21) 出願番号	特願2017-186013 (P2017-186013)	(73) 特許権者	000001443
(22) 出願日	平成29年9月27日 (2017. 9. 27)		カシオ計算機株式会社
(65) 公開番号	特開2019-61098 (P2019-61098A)		東京都渋谷区本町 1 丁目 6 番 2 号
(43) 公開日	平成31年4月18日 (2019. 4. 18)	(74) 代理人	100095407
審査請求日	令和2年9月10日 (2020. 9. 10)		弁理士 木村 満
		(72) 発明者	河村 義裕
			東京都羽村市栄町 3-2-1 カシオ計算機株式会社 羽村技術センター内
		審査官	山下 剛史

最終頁に続く

(54) 【発明の名称】 対話装置、サーバ装置、対話方法及びプログラム

(57) 【特許請求の範囲】

【請求項 1】

ユーザが発話した音声に対する応答文を外部のサーバ装置と通信しながら作成する対話装置であって、

自己の位置データを取得する位置取得部と、

ユーザが発話した音声を音声データとして取得する音声取得部と、

前記音声取得部が取得した音声データに基づく音声情報を記録する音声記録部と、

前記サーバ装置と通信する通信部と、

前記通信部による前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録部が記録した音声情報及び通信切断中に前記位置取得部が取得した位置データを前記サーバ装置に送信し、前記音声情報に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得する、応答文情報取得部と、

前記応答文情報取得部が取得した応答文情報を用いて所定の応答文作成ルールに従って作成された応答文に前記応答文情報取得部が取得した場所名を含む前置きを追加した応答文でユーザに応答する応答部と、

を備えることを特徴とする対話装置。

【請求項 2】

前記通信部による前記サーバ装置との通信が切断している間ユーザに聞いている風に見せかける見せかけ部をさらに備える、

ことを特徴とする請求項 1 に記載の対話装置。

【請求項 3】

前記見せかけ部は、前記音声取得部が取得した前記音声データに応じてうなづく、相づちを打つ、つぶやく、の少なくとも1つを実行する、

ことを特徴とする請求項 2 に記載の対話装置。

【請求項 4】

外部のサーバ装置との間で通信する通信部を有し、ユーザが発話した音声に対する応答文を前記通信部を介して前記サーバ装置と通信しながら作成する対話装置が実行する対話方法であって、

自己の位置データを取得する位置取得処理と、

ユーザが発話した音声に基づく音声情報を記録する音声記録処理と、

前記通信部を介した前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録処理で記録した音声情報及び通信切断中に前記位置取得処理で取得した位置データを前記サーバ装置に送信し、前記音声情報に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得する、応答文情報取得処理と、

前記応答文情報取得処理で取得した応答文情報を用いて所定の応答文作成ルールに従って作成された応答文に前記応答文情報取得処理で取得した場所名を含む前置きを追加した応答文でユーザに回答する回答処理と、

を含むことを特徴とする対話方法。

【請求項 5】

ユーザが発話した音声に対する応答文をサーバ装置と通信しながら作成する対話装置と、前記サーバ装置と、を備える対話システムにおけるサーバ装置であって、

前記対話装置と通信する通信部と、

前記ユーザが発話した音声に基づく音声情報を前記対話装置から前記通信部を介して受信する受信部と、

前記受信部が受信した音声情報に含まれる音声データを音声認識してテキストデータを生成する音声認識部と、

前記音声認識部が生成したテキストデータから該テキストデータに含まれる特徴的な単語である特徴単語を抽出する特徴単語抽出部と、

前記特徴単語抽出部が抽出した特徴単語と所定の応答文作成ルールとに基づき、応答文情報を作成する回答作成部と、

前記回答作成部が作成した応答文情報を前記通信部を介して送信する送信部と、

を備え、

前記通信部による前記対話装置との通信が一時的に切断した後に回復した状態において、通信切断中の音声情報及び通信切断中の前記対話装置の位置データを前記対話装置から受信し、前記受信した音声情報に対して作成した応答文情報及び前記受信した位置データに対応する場所名を前記対話装置に送信する、

ことを特徴とするサーバ装置。

【請求項 6】

外部のサーバ装置との間で通信する通信部を有し、ユーザが発話した音声に対する応答文を前記通信部を介して前記サーバ装置と通信しながら作成する対話装置のコンピュータに、

自己の位置データを取得する位置取得処理、

ユーザが発話した音声を音声データとして取得する音声取得処理、

前記音声取得処理で取得した音声データに基づく音声情報を記録する音声記録処理、

前記通信部を介した前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録処理で記録した音声情報及び通信切断中に前記位置取得処理で取得した位置データを前記サーバ装置に送信し、前記音声情報に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得する、応答文情報取得処理、及び、

前記応答文情報取得処理で取得した応答文情報を用いて所定の応答文作成ルールに従って作成された応答文に前記応答文情報取得処理で取得した場所名を含む前置きを追加した応答文でユーザに応答する応答処理、
を実行させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、ロボット等が音声によってユーザと対話を行う技術に関する。

【背景技術】

【0002】

ユーザと対話することができる端末やロボットの開発が進められている。そして、これら端末やロボットがユーザと対話をする際に、例えば音声認識処理、言語理解処理等の負荷の高い処理や、ロボットの記憶手段に記憶されていない情報を検索する処理等を、外部のサーバに行わせるシステムの開発も進められている。例えば、特許文献1には、ユーザとのやり取りに応じて外部のサーバとネットワーク接続して、必要なデータやプログラムを動的に取得し、ユーザとのコミュニケーションに活用するロボット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】特開2003-111981号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

特許文献1に記載のロボット装置は、外部のサーバとの通信状況が悪い場合や通信が切断された場合には、必要なデータを取得することができないため、適当な対話や行動を続けることによって、ユーザとの対話が途切れないようにする。しかし、このロボット装置による適当な対話や行動はその場限りのものであるため、その後のユーザとのコミュニケーションに支障が生じる可能性がある。

【0005】

例えば、外部のサーバとの通信が切断されている時にユーザがこのロボット装置に何らかの問い合わせを行うと、ロボット装置はその問い合わせに対する適当な行動としてうなずき続ける行動をすることが考えられる。しかし、このロボットはそのユーザの問い合わせを聞き流しているだけであるから、その後外部のサーバと通信可能な状態になったとしても、その問い合わせに対する適切な回答を行うことはできない。したがって、ユーザの問い合わせをうなずきながら聞いてくれていたにもかかわらず適切な回答を行うことができないこのロボットに対し、ユーザは不信感を抱く可能性がある。このように、従来の対話装置では、通信状況が悪い場合の受け答え技術に改善の余地がある。

【0006】

本発明は、上記実情に鑑みてなされたものであり、対話装置の通信状況が悪い場合の受け答え技術を改善することを目的とする。

【課題を解決するための手段】

【0007】

上記目的を達成するため、本発明に係る対話装置は、ユーザが発話した音声に対する応答文を外部のサーバ装置と通信しながら作成する対話装置であって、自己の位置データを取得する位置取得部と、ユーザが発話した音声を音声データとして取得する音声取得部と、前記音声取得部が取得した音声データに基づく音声情報を記録する音声記録部と、前記サーバ装置と通信する通信部と、前記通信部による前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録部が記録した音声情報及び通信切断中に前記位置取得部が取得した位置データを前記サーバ装置に送信し、前記音声

10

20

30

40

50

情報に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得する、応答文情報取得部と、前記応答文情報取得部が取得した応答文情報を用いて所定の応答文作成ルールに従って作成された応答文に前記応答文情報取得部が取得した場所名を含む前置きを追加した応答文でユーザに応答する応答部と、を備えることを特徴とする。

また、本発明に係る対話方法は、外部のサーバ装置との間で通信する通信部を有し、ユーザが発話した音声に対する応答文を前記通信部を介して前記サーバ装置と通信しながら作成する対話装置が実行する対話方法であって、自己の位置データを取得する位置取得処理と、ユーザが発話した音声に基づく音声情報を記録する音声記録処理と、前記通信部を介した前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録処理で記録した音声情報及び通信切断中に前記位置取得処理で取得した位置データを前記サーバ装置に送信し、前記音声情報に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得する、応答文情報取得処理と、前記応答文情報取得処理で取得した応答文情報を用いて所定の応答文作成ルールに従って作成された応答文に前記応答文情報取得処理で取得した場所名を含む前置きを追加した応答文でユーザに応答する応答処理と、を含むことを特徴とする。

また、本発明に係るサーバ装置は、ユーザが発話した音声に対する応答文をサーバ装置と通信しながら作成する対話装置と、前記サーバ装置と、を備える対話システムにおけるサーバ装置であって、前記対話装置と通信する通信部と、前記ユーザが発話した音声に基づく音声情報を前記対話装置から前記通信部を介して受信する受信部と、前記受信部が受信した音声情報に含まれる音声データを音声認識してテキストデータを生成する音声認識部と、前記音声認識部が生成したテキストデータから該テキストデータに含まれる特徴的な単語である特徴単語を抽出する特徴単語抽出部と、前記特徴単語抽出部が抽出した特徴単語と所定の応答文作成ルールとに基づき、応答文情報を作成する応答作成部と、前記応答作成部が作成した応答文情報を前記通信部を介して送信する送信部と、を備え、前記通信部による前記対話装置との通信が一時的に切断した後に回復した状態において、通信切断中の音声情報及び通信切断中の前記対話装置の位置データを前記対話装置から受信し、前記受信した音声情報に対して作成した応答文情報及び前記受信した位置データに対応する場所名を前記対話装置に送信する、ことを特徴とする。

また、本発明に係るプログラムは、外部のサーバ装置との間で通信する通信部を有し、ユーザが発話した音声に対する応答文を前記通信部を介して前記サーバ装置と通信しながら作成する対話装置のコンピュータに、自己の位置データを取得する位置取得処理、ユーザが発話した音声を音声データとして取得する音声取得処理、前記音声取得処理で取得した音声データに基づく音声情報を記録する音声記録処理、前記通信部を介した前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録処理で記録した音声情報及び通信切断中に前記位置取得処理で取得した位置データを前記サーバ装置に送信し、前記音声情報に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得する、応答文情報取得処理、及び、前記応答文情報取得処理で取得した応答文情報を用いて所定の応答文作成ルールに従って作成された応答文に前記応答文情報取得処理で取得した場所名を含む前置きを追加した応答文でユーザに応答する応答処理、を実行させることを特徴とする。

【発明の効果】

【 0 0 0 8 】

本発明によれば、対話装置の通信状況が悪い場合の受け答え技術を改善することができる。

【図面の簡単な説明】

【 0 0 0 9 】

【図 1】本発明の第 1 実施形態に係る対話システムの構成を示す図である。

【図 2】第 1 実施形態に係る対話装置の外観を示す図である。

【図 3】第 1 実施形態に係る対話装置の構成を示す図である。

【図４】第１実施形態に係る対話装置が記憶する付加情報付音声情報の一例を示す図である。

【図５】第１実施形態に係るサーバ装置の構成を示す図である。

【図６】第１実施形態に係るサーバ装置が記憶する応答文作成ルールの一例を示す図である。

【図７】第１実施形態に係る対話装置の対話制御処理のフローチャートである。

【図８】第１実施形態に係る対話装置の見せかけスレッドのフローチャートである。

【図９】第１実施形態に係るサーバ装置の応答文作成処理のフローチャートである。

【図１０】本発明の第２実施形態に係る対話装置の構成を示す図である。

【図１１】第２実施形態に係る対話装置が記憶する応答文情報リストの一例を示す図である。 10

【図１２】第２実施形態に係る対話装置の対話制御処理のフローチャートである。

【図１３】第２実施形態に係るサーバ装置の応答文作成処理のフローチャートである。

【図１４】本発明の第３実施形態に係る対話装置の構成を示す図である。

【図１５】第３実施形態に係る対話装置が記憶する位置履歴データの一例を示す図である。

【図１６】第３実施形態に係る対話装置の対話制御処理のフローチャートである。

【図１７】第３実施形態に係るサーバ装置が対話装置に送信する特徴単語、応答文、及び場所名の例を示す図である。

【図１８】第３実施形態に係るサーバ装置の応答文作成処理のフローチャートである。 20

【発明を実施するための形態】

【００１０】

以下、本発明の実施形態について、図表を参照して説明する。なお、図中同一又は相当部分には同一符号を付す。

【００１１】

（第１実施形態）

図１に示すように、本発明の第１実施形態に係る対話システム１０００は、音声でユーザＵと対話するロボットである対話装置１００と、対話装置１００がユーザＵと対話する際に必要な各種処理（例えば音声認識処理、応答文作成処理等）を行うサーバ装置２００と、を備える。対話装置１００はユーザが発話した音声のデータ（音声データ）を外部のサーバ装置２００に送信し、該サーバ装置２００に音声認識処理や応答文情報作成等を行ってもらうことにより、ユーザＵと対話する際の対話装置１００自身の処理負荷を軽くしている。 30

【００１２】

図２に示すように、対話装置１００は頭２０と胴体３０とからなる。そして、対話装置１００の頭２０には、マイク２１と、カメラ２２と、スピーカ２３と、センサ群２４と、が設けられている。

【００１３】

マイク２１は、頭２０の左右、人の顔でいうところの耳の位置に複数設けられており、アレイマイクを構成する。マイク２１は、対話装置１００の周囲にいるユーザＵが発話した音声の音声データとして取得する音声取得部として機能する。 40

【００１４】

カメラ２２は、頭２０の前面の中央部、人の顔でいうところの鼻の位置に設けられている撮像装置である。カメラ２２は、対話装置１００の正面の画像のデータ（画像データ）を取得する画像取得部として機能し、取得した画像データを後述する制御部１１０に入力する。

【００１５】

スピーカ２３は、カメラ２２の下側、人の顔でいうところの口の位置に設けられている。スピーカ２３は、音声を出力する音声出力部として機能する。

【００１６】

センサ群 24 は、人の顔でいうところの目の位置に設けられている。センサ群 24 は、加速度センサ、障害物検知センサ等を含み、各種物理量を検出して、対話装置 100 の姿勢制御、衝突回避、安全性確保等のために使用される。

【0017】

対話装置 100 の頭 20 と胴体 30 とは、図 2 に示すように、破線で示される首関節 31 によって相互に連結されている。首関節 31 は、複数のモータを含む。後述する制御部 110 がこれら複数のモータを駆動することにより、対話装置 100 の頭 20 を上下方向、左右方向及び傾ける方向の 3 軸で回転させることができる。これにより、対話装置 100 は、例えばうなずきの動作をすることができる。

【0018】

図 2 に示すように、対話装置 100 の胴体 30 の下部には足回り部 32 が設けられている。足回り部 32 は、4 つの車輪（ホイール）と駆動モータとを含む。4 つの車輪のうち、2 つが前輪として胴体 30 の前側に、残り 2 つが後輪として胴体 30 の後ろ側に、それぞれ配置されている。車輪としては、例えば、オムニホイール、メカナムホイール等を使用してもよい。後述する制御部 110 が、駆動モータを制御して車輪を回転させると、対話装置 100 は移動する。

【0019】

次に、図 3 を参照して、対話装置 100 の機能構成を説明する。図 3 に示すように、対話装置 100 は、上述の構成に加え、通信部 25 と、操作ボタン 33 と、制御部 110 と、記憶部 120 と、を備える。

【0020】

通信部 25 は、サーバ装置 200 等の外部装置と無線通信するための、アンテナを含む無線モジュールである。例えば、通信部 25 は、無線 LAN (Local Area Network) により無線通信を行うための無線モジュールである。通信部 25 を用いることにより、対話装置 100 は、サーバ装置 200 に音声データ等の音声情報を送信し、また、サーバ装置 200 から後述する応答文情報を受信することができる。対話装置 100 とサーバ装置 200 との無線通信は、ダイレクトな通信でもよいし、基地局、アクセスポイント等を介しての通信でもよい。

【0021】

操作ボタン 33 は、図示しないが、胴体 30 の背中の位置に設けられている。操作ボタン 33 は、対話装置 100 を操作するための各種ボタンである。操作ボタン 33 は、電源ボタン、スピーカ 23 の音量調節ボタン等を含む。

【0022】

制御部 110 は、CPU (Central Processing Unit) 等で構成される。制御部 110 は、記憶部 120 に記憶されたプログラムを実行することにより、後述する音声記録部 111、見せかけ部 112、応答文情報取得部 113 及び応答部 114 として機能する。また、制御部 110 は、時計機能及びタイマー機能を備え、現在時刻（現在日時）や経過時間を取得することができる。

【0023】

記憶部 120 は、ROM (Read Only Memory)、RAM (Random Access Memory) 等で構成され、制御部 110 の CPU が実行するプログラム、各種データ等を記憶する。また、記憶部 120 は、音声取得部（マイク 21）が取得した音声データに発話日時等を付加した付加情報付音声情報 121 も記憶する。

【0024】

付加情報付音声情報 121 は、図 4 に示すように、ユーザが発話した内容を、通信状態及び発話日時とともに記録したデータである。通信状態の値は、通信部 25 がサーバ装置 200 と通信可能な状態なら「接続」、通信不可能な状態なら「切断」となる。図 4 では、通信状態によらずに付加情報付音声情報 121 を記憶しているが、通信状態が「切断」の付加情報付音声情報 121 だけを記憶部 120 に記録するようにしてもよい。また、通信切断の検出をトリガにして付加情報付音声情報 121 の記録を開始するようにしてもよ

10

20

30

40

50

い。また、通信状態の値は付加情報付音声情報 1 2 1 に含めずに、サーバ装置 2 0 0 が発話日時に基づいて通信状態を判断してもよい。

【 0 0 2 5 】

次に、制御部 1 1 0 が実現する各機能について説明する。制御部 1 1 0 は、前述したように、記憶部 1 2 0 に記憶されたプログラムを実行することにより、音声記録部 1 1 1、見せかけ部 1 1 2、応答文情報取得部 1 1 3 及び応答部 1 1 4 として機能する。また、制御部 1 1 0 は、マルチスレッド機能に対応しており、複数のスレッド（異なる処理の流れ）を並行して実行することができる。

【 0 0 2 6 】

音声記録部 1 1 1 は、音声取得部（マイク 2 1）が取得した音声データに発話日時等を付加して付加情報付音声情報 1 2 1 として、記憶部 1 2 0 に記録する。なお、本実施形態では後述するように音声認識処理をサーバ装置 2 0 0 で行うが、音声認識処理を対話装置 1 0 0 で行う実施形態も考えられる。この場合、音声記録部 1 1 1 は、音声データを音声認識したテキストデータを記憶部 1 2 0 に記録してもよい。そこで、対話装置 1 0 0 がサーバ装置 2 0 0 に送信する情報を音声情報と表現することとする。本実施形態では音声情報は音声取得部が取得した音声データであるが、音声情報が音声認識後のテキストデータである実施形態も考えられる。そして、音声情報に発話日時等を付加した情報が、付加情報付音声情報 1 2 1 である。

【 0 0 2 7 】

見せかけ部 1 1 2 は、通信部 2 5 によるサーバ装置 2 0 0 との通信が切断されている場合に、ユーザ U にユーザ U が発話している内容を聞いている風に見せかける動作をするための制御を行う。具体的には、うなづく、相づちを打つ等の動作をするように、首関節 3 1 やスピーカ 2 3 等を制御する。

【 0 0 2 8 】

応答文情報取得部 1 1 3 は、サーバ装置 2 0 0 が作成した応答文に関する情報（応答文情報）を、通信部 2 5 を介して取得する。応答文情報については後述する。

【 0 0 2 9 】

応答部 1 1 4 は、応答文情報取得部 1 1 3 が取得した応答文情報に基づいて作成された応答文で、ユーザ U に対して応答する。具体的には、応答部 1 1 4 は、応答文情報に基づいて作成された応答文を音声合成し、スピーカ 2 3 から該応答文の音声を出力する。なお、音声合成処理をサーバ装置 2 0 0 が行う実施形態も考えられる。このような実施形態では音声合成後のボイスデータが応答文情報としてサーバ装置 2 0 0 から送信されるため、応答部 1 1 4 は、音声合成処理をする必要なく、そのままそのボイスデータをスピーカ 2 3 から出力することができる。

【 0 0 3 0 】

以上、対話装置 1 0 0 の機能構成を説明した。次に、サーバ装置 2 0 0 の機能構成を説明する。図 5 に示すように、サーバ装置 2 0 0 は、制御部 2 1 0 と、記憶部 2 2 0 と、通信部 2 3 0 と、を備える。

【 0 0 3 1 】

制御部 2 1 0 は、CPU 等で構成される。制御部 2 1 0 は、記憶部 2 2 0 に記憶されたプログラムを実行することにより、後述する音声認識部 2 1 1、特徴単語抽出部 2 1 2 及び応答作成部 2 1 3 として機能する。

【 0 0 3 2 】

記憶部 2 2 0 は、ROM、RAM 等で構成され、制御部 2 1 0 の CPU が実行するプログラム、各種データ等を記憶する。また、記憶部 2 2 0 は、後述する応答文作成ルール 2 2 1 も記憶する。

【 0 0 3 3 】

応答文作成ルール 2 2 1 は、図 6 に示すように、特定の単語（特徴単語）毎に応答文を対応させたルールである。なお、図 6 では、応答文作成ルール 2 2 1 は、特徴単語として、「暑い」「映画」「かわいい」のような具体的な単語を割り当てたルールになっている

10

20

30

40

50

が、これに限られない。例えば、特徴単語を「寒暖を表すネガティブな形容詞： い」と定義し、これに対応する応答文を「 い い言っていると余計 くなるよ。」とルール付けしてもよい。また、寒暖を表す形容詞の他の応答文作成ルール例として、例えば、特徴単語を「寒暖を表すポジティブな形容詞： い」と定義し、これに対応する応答文を「最近 は くなってきたのかな。 いと気持ち良いね。」とルール付けしてもよい。ここで、「寒暖を表すネガティブな形容詞」としては、例えば「暑い」「寒い」等が挙げられ、「寒暖を表すポジティブな形容詞」としては、例えば「涼しい」「暖かい」等が挙げられる。

【 0 0 3 4 】

通信部 2 3 0 は、対話装置 1 0 0 等の外部装置と無線通信するための、アンテナを含む無線モジュールである。例えば、通信部 2 3 0 は、無線 LAN (Local Area Network) により無線通信を行うための無線モジュールである。通信部 2 3 0 を用いることにより、サーバ装置 2 0 0 は、対話装置 1 0 0 から音声データ等の音声情報を受信し、また、対話装置 1 0 0 に後述する応答文情報を送信することができる。制御部 2 1 0 は、通信部 2 3 0 を介して対話装置 1 0 0 から音声情報を受信する際には受信部として機能し、通信部 2 3 0 を介して対話装置 1 0 0 に応答文情報を送信する際には送信部として機能する。

【 0 0 3 5 】

次に、制御部 2 1 0 が実現する各機能について説明する。制御部 2 1 0 は、前述したように、記憶部 2 2 0 に記憶されたプログラムを実行することにより、音声認識部 2 1 1、特徴単語抽出部 2 1 2 及び応答作成部 2 1 3 として機能する。

【 0 0 3 6 】

音声認識部 2 1 1 は、対話装置 1 0 0 から送信された付加情報付音声情報 1 2 1 に含まれる音声データを音声認識して、ユーザの発話内容を表すテキストデータを生成する。上述したように、音声認識を対話装置 1 0 0 が行う実施形態においては、音声認識部 2 1 1 は不要であり、この場合は、対話装置 1 0 0 から送信された付加情報付音声情報 1 2 1 に音声認識後のテキストデータが含まれている。

【 0 0 3 7 】

特徴単語抽出部 2 1 2 は、音声認識部 2 1 1 が生成したテキストデータ（又は付加情報付音声情報 1 2 1 に含まれているテキストデータ）から、該テキストデータに含まれる特徴的な単語である特徴単語を抽出する。特徴単語とは、例えば、テキストデータ中に含まれる特定ワード（名詞、動詞、形容詞、形容動詞）のうち、最も多く含まれる特定ワードである。また、テキストデータ中に含まれる特定ワードのうち、強調修飾語（「とても」、「すごく」等）に修飾された特定ワードも、特徴単語になり得る。

【 0 0 3 8 】

応答作成部 2 1 3 は、特徴単語抽出部 2 1 2 が抽出した特徴単語を、記憶部 2 2 0 に記憶されている応答文作成ルール 2 2 1 に適用して、応答文に関する情報（応答文情報）を作成する。なお、本実施形態では、応答作成部 2 1 3 が完成した応答文を作成しているが、これに限られない。対話処理においては、ユーザが発話した音声を音声認識し、構文解析等して、応答文を作成し、音声合成するという一連の処理が存在するが、この中の一部をサーバ装置 2 0 0 が行い、残りの処理を対話装置 1 0 0 が行うようにしてもよい。例えば、音声認識や構文解析等の重い処理をサーバ装置 2 0 0 が行い、応答文を完成させる処理は対話装置 1 0 0 が行ってもよい。これらの処理のどれをどちらの装置で行うかは任意である。そこで、サーバ装置 2 0 0 が対話装置 1 0 0 へ送信する情報を応答文情報と表現し、対話装置 1 0 0 がユーザに発話する情報を応答文と表現するものとする。応答文情報と応答文が同一である（デジタルデータかアナログ音声であるか等の信号形態は異なるとしても、内容として同じ）場合もある。本実施形態では、応答文情報は応答文と同一である。

【 0 0 3 9 】

以上、サーバ装置 2 0 0 の機能構成を説明した。次に、対話装置 1 0 0 の制御部 1 1 0 が行う対話制御処理について、図 7 を参照して説明する。この処理は、対話装置 1 0 0 が

10

20

30

40

50

起動して初期設定が完了すると開始される。

【 0 0 4 0 】

まず、制御部 1 1 0 は、通信部 2 5 によるサーバ装置 2 0 0 との通信が切断されているか否かを判定する（ステップ S 1 0 1）。例えば、通信部 2 5 がアクセスポイントを経由してサーバ装置 2 0 0 と通信している場合、該アクセスポイントの電波が受信できないならサーバ装置 2 0 0 との通信が切断されていると判断する。

【 0 0 4 1 】

サーバ装置 2 0 0 との通信が切断されているなら（ステップ S 1 0 1 ; Y e s）、制御部 1 1 0 は、現在時刻（通信が切断された時刻）を記憶部 1 2 0 に記憶する（ステップ S 1 0 2）。そして、見せかけ部 1 1 2 としての制御部 1 1 0 は、後述する見せかけスレッドを起動し（ステップ S 1 0 3）、見せかけスレッドの処理を並行して行う。

10

【 0 0 4 2 】

そして、音声記録部 1 1 1 としての制御部 1 1 0 は、音声取得部（マイク 2 1）が取得した音声データに、通信状態（切断）及び現在時刻の情報を付加して、付加情報付音声情報 1 2 1 として記憶部 1 2 0 に記録する（ステップ S 1 0 4）。ステップ S 1 0 4 は、音声記録ステップとも呼ばれる。その後、制御部 1 1 0 は、サーバ装置 2 0 0 との通信が回復したか否かを判定する（ステップ S 1 0 5）。サーバ装置 2 0 0 との通信が回復していなければ（ステップ S 1 0 5 ; N o）、制御部 1 1 0 は、ステップ S 1 0 4 に戻って、通信が回復するまで付加情報付音声情報 1 2 1 を記録しながら待機する。サーバ装置 2 0 0 との通信が回復したら（ステップ S 1 0 5 ; Y e s）、制御部 1 1 0 は、見せかけスレッドを終了させる（ステップ S 1 0 6）。

20

【 0 0 4 3 】

そして、制御部 1 1 0 は、ステップ S 1 0 2 で記憶部 2 2 0 に記憶した通信切断時刻から現在時刻までの（通信切断中の）付加情報付音声情報 1 2 1 を、通信部 2 5 を介して、サーバ装置 2 0 0 に送信する（ステップ S 1 0 7）。なお、ここでは通信の回復を対話装置 1 0 0 が検出しているが、サーバ装置 2 0 0 が通信の回復を検出して、対話装置 1 0 0 に付加情報付音声情報 1 2 1 の送信をリクエストしてもよい。対話装置 1 0 0 がステップ S 1 0 7 で送信した付加情報付音声情報 1 2 1 はサーバ装置 2 0 0 で音声認識され、サーバ装置 2 0 0 は対話装置 1 0 0 に応答文情報を送信する。

【 0 0 4 4 】

30

すると、応答文情報取得部 1 1 3 としての制御部 1 1 0 は、通信部 2 5 を介して、サーバ装置 2 0 0 が送信した応答文情報を取得する（ステップ S 1 0 8）。ステップ S 1 0 8 は、応答文情報取得ステップとも呼ばれる。本実施形態では完成文としての応答文を応答文情報として取得するが、これに限らず、サーバ装置 2 0 0 が応答文作成の全部でなく一部を担当する場合は、部分的な情報としての応答文情報（例えば後述する特徴単語の情報）を取得し、対話装置 1 0 0 内で応答文を完成させてもよい。

【 0 0 4 5 】

そして、応答部 1 1 4 としての制御部 1 1 0 は、応答文情報取得部 1 1 3 が取得した応答文情報に基づき、ユーザに応答する（ステップ S 1 0 9）。本実施形態では、応答文情報は応答文そのものなので、具体的には、応答部 1 1 4 は、応答文の内容を音声合成して、スピーカ 2 3 から応答文を発話する。この応答文は、サーバ装置 2 0 0 と対話装置 1 0 0 との連携により、通信切断中の音声に対応する内容の応答文になっているので、ユーザは対話装置 1 0 0 が通信切断中もユーザの発話内容をきちんと聞いてくれたことを確認できる。ステップ S 1 0 9 は、応答ステップとも呼ばれる。そして、制御部 1 1 0 は、処理をステップ S 1 0 1 に戻す。

40

【 0 0 4 6 】

一方、ステップ S 1 0 1 で、サーバ装置 2 0 0 との通信が切断されていないなら（ステップ S 1 0 1 ; N o）、音声記録部 1 1 1 としての制御部 1 1 0 は、マイク 2 1 が取得した音声に、通信状態（接続）及び現在時刻の情報を付加して、付加情報付音声情報 1 2 1 として記憶部 1 2 0 に記録する（ステップ S 1 1 0）。そして、制御部 1 1 0 は、ステッ

50

プ S 1 1 0 で記録した（通信接続中の）付加情報付音声情報 1 2 1 を、通信部 2 5 を介してサーバ装置 2 0 0 に送信する（ステップ S 1 1 1）。

【 0 0 4 7 】

なお、通信状態が「切断」の付加情報付音声情報 1 2 1 だけが記憶部 1 2 0 に記録されるようにした場合は、ステップ S 1 1 0 の処理をスキップし、ステップ S 1 1 1 の処理の代わりに、制御部 1 1 0 は、マイク 2 1 が取得した音声データに、通信状態（接続）及び現在時刻を付加して、付加情報付音声情報 1 2 1 として、通信部 2 5 を介してサーバ装置 2 0 0 に送信する。

【 0 0 4 8 】

本実施形態では、上記いずれの場合も、ここで送信された付加情報付音声情報 1 2 1 に含まれる音声データがサーバ装置 2 0 0 で音声認識され、サーバ装置 2 0 0 は対話装置 1 0 0 に応答文を送信する。このサーバ装置 2 0 0 による処理（応答文作成処理）については後述する。

【 0 0 4 9 】

すると、応答文情報取得部 1 1 3 としての制御部 1 1 0 は、通信部 2 5 を介して、サーバ装置 2 0 0 が送信した応答文情報を取得する（ステップ S 1 1 2）。そして、応答部 1 1 4 としての制御部 1 1 0 は、応答文情報取得部 1 1 3 が取得した応答文情報に基づき、ユーザに応答する（ステップ S 1 1 3）。本実施形態では、応答文情報は応答文そのものなので、具体的には、応答部 1 1 4 は、応答文の内容を音声合成して、スピーカ 2 3 から応答文を発話する。この応答文は、サーバ装置 2 0 0 と対話装置 1 0 0 との連携により、通信接続中の音声に対応する内容の応答文になっているので、従来技術により作成された応答文と同様の内容である。そして、制御部 1 1 0 は、処理をステップ S 1 0 1 に戻す。

【 0 0 5 0 】

次に、ステップ S 1 0 3 で起動される見せかけスレッドの処理について、図 8 を参照して説明する。

【 0 0 5 1 】

まず、制御部 1 1 0 は、制御部 1 1 0 が備えるタイマーを、釈明を行う間隔設定用にするためにリセットする（ステップ S 2 0 1）。これ以降このタイマーを釈明用タイマーと呼ぶことにする。

【 0 0 5 2 】

そして、制御部 1 1 0 は、カメラ 2 2 が取得した画像を認識し（ステップ S 2 0 2）、ユーザに見つめられているか否かを判定する（ステップ S 2 0 3）。ユーザに見つめられているなら（ステップ S 2 0 3 ; Yes）、ユーザに例えば「今、頭がボーッとされていて、きちんとしたお返事ができないのです。ごめんなさい。」等のような釈明をする（ステップ S 2 0 4）。この時は、サーバ装置 2 0 0 との間の通信が切断されており、音声認識や応答文作成ができないからである。

【 0 0 5 3 】

そして、釈明をしたことにより、制御部 1 1 0 は、釈明用タイマーをリセットする（ステップ S 2 0 5）。そして、制御部 1 1 0 は、1 0 秒待ってから（ステップ S 2 0 6）、ステップ S 2 0 2 に戻る。ここで、この 1 0 秒という値は、対話装置 1 0 0 が頻繁に同じ動作を繰り返さないようにするための待ち時間の例であり、1 0 秒に限定する必要はなく、3 秒、1 分等、任意の値に変更可能である。なお、ステップ S 2 0 6 におけるこの待ち時間を、他の待ち時間と区別するために、見せかけ待ち基準時間と呼ぶことにする。

【 0 0 5 4 】

一方、ステップ S 2 0 3 で、ユーザに見つめられていないなら（ステップ S 2 0 3 ; No）、制御部 1 1 0 は、釈明用タイマーの値がリセット後 3 分経過しているか否かを判定する（ステップ S 2 0 7）。なおこの 3 分という値は、対話装置 1 0 0 が頻繁に釈明しないようにするための待ち時間の例であり、3 分に限定する必要はない。例えば、1 分、1 0 分等、任意の値に変更可能である。なお、この待ち時間を他の待ち時間と区別するために、釈明基準時間と呼ぶことにする。

10

20

30

40

50

【 0 0 5 5 】

3分経過しているなら（ステップS207；Yes）、ステップS204に進み、以降の処理は上述した通りとなる。3分経過していないなら（ステップS207；No）、制御部110は、マイク21から取得される音声が入切れたか否かを判定する（ステップS208）。この判定は、例えば、マイク21から取得される音声において、無音期間が基準無音時間（例えば1秒）以上続いたら、制御部110は音声が入切れたと判定する。

【 0 0 5 6 】

音声が入切れていなければ（ステップS208；No）、ステップS202に戻る。音声が入切れていれば（ステップS208；Yes）、制御部110は、「うなずく」、「相づちを打つ」、「つぶやく」の3つの中から1つランダムに選択し、選択した動作を行

10

【 0 0 5 7 】

例えば、「うなずく」を選択したら、制御部110は、首関節31を用いて頭20を縦に振るように動かす。このうなずきの動作について、制御部110は、ステップS209を実行する度に、頭20を振る回数や速度をランダムに変更してもよい。また、「相づちを打つ」を選択したら、制御部110は、首関節31を用いて頭20を縦に振るように動かしながら、スピーカ23から「はい」、「そうですね」、「うん」等を発話する。この相づちの動作についても、制御部110は、ステップS209を実行する度に、制御部110は、頭20を振る回数や速度、スピーカ23から発話する内容をランダムに変更してもよい。

20

【 0 0 5 8 】

また「つぶやく」を選択したら、制御部110は、スピーカ23から適当なつぶやきを発話させる。ここで、適当なつぶやきとは、人間的なつぶやきでもよいが、動物の鳴き声を模倣した音や、ロボットにありがちな人間には理解不能な電子音等でもよい。このつぶやきについても、制御部110は、ステップS209を実行する度に、いくつかの種類の中から制御部110がランダムに選択したつぶやきを発話させるようにしてもよい。

【 0 0 5 9 】

そして、ステップS206に進み、以降の処理は上述した通りとなる。以上、説明した見せかけスレッドの処理により、対話装置100は、サーバ装置200との通信が切断している時でも、ユーザに聞いている風に見せかけることができる。

30

【 0 0 6 0 】

次に、サーバ装置200が行う応答文作成処理について、図9を参照して説明する。なお、サーバ装置200は起動すると応答文作成処理を開始する。

【 0 0 6 1 】

まず、サーバ装置200の通信部230は、対話装置100が送信した付加情報付音声情報121を受信する（ステップS301）。対話装置100から付加情報付音声情報121が送信されていないなら、送信されるまでステップS301で待機する。そして、制御部210は、受信した付加情報付音声情報121が、通信切断中のものであるか否かを判定する（ステップS302）。付加情報付音声情報121には、図4に示すように通信状態を示す情報が含まれているので、この情報を参照することにより、受信した付加情報付音声情報121が通信切断中のものであるか否かを判定することができる。また、サーバ装置200は、対話装置100との通信状況を把握できるので、付加情報付音声情報121に通信状態を示す情報が含まれていなくても、付加情報付音声情報121に含まれている発話日時の情報に基づいて、その付加情報付音声情報121が通信切断中のものであるか否かを判断することもできる。

40

【 0 0 6 2 】

受信した付加情報付音声情報121が、通信切断中のものであるなら（ステップS302；Yes）、音声認識部211としての制御部210は、付加情報付音声情報121に含まれる音声データを音声認識してテキストデータを生成する（ステップS303）。そして、特徴単語抽出部212としての制御部210は、生成されたテキストデータから特

50

徴単語を抽出する（ステップS304）。そして、応答作成部213としての制御部210は、抽出された特徴単語と応答文作成ルール221に基づき、応答文情報（本実施形態では応答文そのもの）を作成する（ステップS305）。そして、応答作成部213は、作成した応答文（応答文情報）を、通信部230を介して対話装置100に送信する（ステップS306）。そして、ステップS301に戻る。

【0063】

一方、受信した付加情報付音声情報121が、通信切断中のものでなければ（ステップS302；No）、音声認識部211としての制御部210は、付加情報付音声情報121に含まれる音声データを音声認識してテキストデータを生成する（ステップS307）。そして、応答作成部213としての制御部210は、生成されたテキストデータに対する応答文情報（本実施形態では応答文そのもの）を、従来の応答文作成技術を用いて作成する（ステップS308）。そして、応答作成部213は、作成した応答文（応答文情報）を、通信部230を介して対話装置100に送信する（ステップS309）。そして、ステップS301に戻る。

10

【0064】

以上説明した応答文作成処理により、通信接続中は通常の応答文情報が生成され、通信切断中は、特徴単語及び応答文作成ルールに基づいて応答文情報が作成される。したがって、サーバ装置200は、対話装置100との通信が切断している間の音声情報に対して、ユーザの発話をきちんと聞いていたかのように思わせるための応答文情報を作成することができる。

20

【0065】

そして、上述の対話装置100の対話制御処理により、サーバ装置200との通信が切断している間の音声情報に対する応答文情報をサーバ装置200から取得することにより、対話装置100は、ユーザの発話をきちんと聞いていたかのように思わせる応答文を発話することができる。

【0066】

例えば、図4のNo1からNo.3に示すユーザの発話内容に対し、対話装置100はその時点では応答文を返答できないが、サーバ装置200との通信が回復した時点で、これらのNo.1からNo.3に示すユーザの発話内容がサーバ装置200に送信される。そして、サーバ装置200の特徴単語抽出部212により、これらのユーザの発話内容から、最も多く使われている特定ワードとして「暑い」が抽出される。この「暑い」を図6に示す応答文作成ルールに適用することにより、応答作成部213は「暑い暑い言っていると余計暑くなるよ。」という応答文情報（本実施形態では、応答文そのもの）を作成する。そして、対話装置100の応答文情報取得部113は、この応答文（応答文情報）を取得し、応答部114により、対話装置100はユーザに対して「暑い暑い言っていると余計暑くなるよ。」と発話することができる。

30

【0067】

このように、対話装置100は、サーバ装置200との通信が切断している時には小まめな応答を行うことができないが、通信が回復した時に、切断中のユーザの発話内容に含まれる特徴単語（最も多く使われている特定ワード等）に基づいた応答文を発話することによって、比較的短い応答文で、通信切断中もきちんとユーザの発話内容を聞いていたことをユーザに示すことができる。このように、対話装置100は、通信状況が悪い場合の受け答え技術を改善することができる。

40

【0068】

（第2実施形態）

上述した第1実施形態では、対話装置100は、サーバ装置200との通信が切断している間にユーザが発話した内容全体の中で最も多く使われている特定ワード等（1つの特徴単語）に対応する応答文で応答する。特徴単語はユーザの印象に残りやすいので、このような応答でもあまり問題は生じないと考えられるが、場合によってはユーザが発話中に話題が変化し、時間の経過とともに複数の特徴単語が同じ位多く使われることもあり得る

50

。このような場合は、話題毎にそれぞれ最も多く使われている特徴単語を抽出して、抽出された複数の特徴単語それぞれに対応する応答文により複数回応答した方が望ましい場合もあると考えられる。そこで、このような複数の応答文により応答可能な第2実施形態について説明する。

【0069】

第2実施形態に係る対話システム1001が対話装置101とサーバ装置201とを備える点は、第1実施形態に係る対話システム1000と同じである。第2実施形態に係る対話装置101は、第1実施形態に係る対話装置100と同じ外観である。対話装置101の機能構成は、図10に示すように、第1実施形態に係る対話装置100と比較して、記憶部120に、応答文情報リスト122を記憶する点が異なる。また、サーバ装置201の機能構成は、第1実施形態に係るサーバ装置200と同じである。

10

【0070】

応答文情報リスト122は、図11に示すように、「発話日時」、「特徴単語」、「ユーザの音声に対する応答文」を含み、これらは、サーバ装置201から送信された情報である。例えば、図11のNo. 1は、ユーザが2017年9月5日10時3分5秒から2017年9月5日10時3分11秒までの間に発話した内容に含まれている特徴単語が「暑い」であり、このユーザの発話に対する応答文が「暑い暑い言っていると余計暑くなるよ。」であることを示している。No. 2以降も同様である。なお、説明のための一例であるが、図11に示される「ユーザの音声に対する応答文」が対応する「ユーザの発話内容」は、図4に示す付加情報付音声情報121に示されているものである。

20

【0071】

次に、対話装置101の制御部110が行う対話制御処理について、図12を参照して説明する。この処理は、第1実施形態に係る対話装置100の対話制御処理（図7）と比べ、一部を除いて同じなので、異なる点を中心に説明する。

【0072】

ステップS101からステップS107及びステップS110からステップS113は、図7を参照して説明した処理と同じである。ステップS107の次のステップであるステップS121では、応答文情報取得部113としての制御部110は、通信部25を介して、サーバ装置201が送信した応答文情報リスト122を取得する。次に、応答文情報リスト122には、1つ以上の応答文情報が含まれているので、応答文情報取得部113としての制御部110は、応答文情報リスト122から応答文情報を1つ取り出す（ステップS122）。

30

【0073】

応答文情報リスト122から取り出した応答文情報は、図11に示すように、「発話日時」が含まれている。制御部110は、「発話日時」の終了時刻が現在時刻より2分以上前であるか否かを判定する（ステップS123）。ここで2分とは、次に述べるステップS124で前置きを追加するか否かを判定するための時間なので、前置き判定基準時間とも言い、2分に限られない。前置き判定基準時間は、例えば3分、10分等、任意の値に変更可能である。

【0074】

40

「発話日時」の終了時刻が現在時刻より2分以上前なら（ステップS123；Yes）、応答部114としての制御部110は、応答文情報に前置きを追加する。ここで前置きとは、例えば「そういえば、暑いと言っていましたけど」というような句である。より一般的には、「そういえば、[特徴単語]と言っていましたけど」と表すことができる。この前置きを追加することにより、「特徴単語」に対応する応答文が唐突に発話されたような印象をユーザに与えることを避けることができる。なお、「発話日時」の終了時刻が現在時刻より2分以上前ということではないなら（ステップS123；No）、前置きを追加すること無く、ステップS125に進む。

【0075】

そして、応答部114としての制御部110は、応答文情報取得部113が取得した応

50

答文情報（ステップS 1 2 4で前置きが追加された場合は、前置き付きの応答文情報）に基づき、ユーザに回答する（ステップS 1 2 5）。本実施形態では、応答文情報は応答文そのものなので、具体的には、応答部1 1 4は、応答文（又は前置き付きの応答文）の内容を音声合成して、スピーカ2 3から応答文を発話する。そして、制御部1 1 0は、応答文情報リスト1 2 2に次の応答文情報（まだ発話の対象になっていない応答文情報）があるかないかを判定する（ステップS 1 2 6）。

【0 0 7 6】

次の応答文情報があるなら（ステップS 1 2 6；Y e s）、ステップS 1 2 2に戻り、応答文情報リストに存在する全ての応答文情報が発話されるまで、ステップS 1 2 2からステップS 1 2 5までの処理を繰り返す。次の応答文情報がないなら（ステップS 1 2 6；N o）、ステップS 1 0 1に戻る。この応答文情報リストには、サーバ装置2 0 1で作成された、通信切断中の音声に対応する内容の、複数の応答文が含まれているので、ユーザは対話装置1 0 1が通信切断中もユーザの発話内容をきちんと聞いてくれたことを確認できる。

【0 0 7 7】

次に、サーバ装置2 0 1が行う応答文作成処理について、図1 3を参照して説明する。この処理は、第1実施形態に係るサーバ装置2 0 0の応答文作成処理（図9）と比べ、一部を除いて同じなので、異なる点を中心に説明する。

【0 0 7 8】

ステップS 3 0 1からステップS 3 0 3及びステップS 3 0 7からステップS 3 0 9は、図9を参照して説明した処理と同じである。ステップS 3 0 3の次のステップであるステップS 3 2 1では、制御部2 1 0は、対話装置1 0 1が送信した音声情報（本実施形態では音声データ）から話の切れ目（話題）を抽出する。これは、ステップS 3 0 3で生成されたテキストデータに基づいて話の切れ目（話題）を抽出してもよいし、音声データに基づいて例えば音声の途切れ等に基づいて話の切れ目（話題）を抽出してもよい。

【0 0 7 9】

次に、特徴単語抽出部2 1 2としての制御部2 1 0は、ステップS 3 2 1で抽出した話の切れ目（話題）毎に、特徴単語を抽出する（ステップS 3 2 2）。例えば、音声データの話の切れ目が発話開始から3分のところと5分のところに抽出された場合を想定する。この場合、発話開始後3分までの部分に最も多く含まれている特定ワードを最初の話題の特徴単語として抽出する。そして、発話開始後3分から5分までの部分に最も多く含まれている特定ワードを2番目の話題の特徴単語として抽出する。そして、発話開始後5分以降の部分に最も多く含まれている特定ワードを3番目の話題の特徴単語として抽出する。

【0 0 8 0】

そして、応答作成部2 1 3としての制御部2 1 0は、話の切れ目（話題）毎に抽出した特徴単語を応答文作成ルール2 2 1に適用して応答文情報（本実施形態では応答文そのもの）を作成し、その応答文に発話日時及び特徴単語を付加して、図1 1に示すような応答文情報リストを作成する（ステップS 3 2 3）。そして、応答作成部2 1 3は、作成した応答文情報リストを、通信部2 3 0を介して対話装置1 0 1に送信する（ステップS 3 2 4）。そして、ステップS 3 0 1に戻る。

【0 0 8 1】

以上説明した応答文作成処理により、通信切断中に複数の話題からなる発話をユーザがしたとしても、各話題にそれぞれ含まれる特徴単語に基づいて応答文情報リストが作成される。したがって、サーバ装置2 0 1は、対話装置1 0 1との通信が切断している間に発話された複数の話題それぞれに対応する応答文情報を作成することができる。

【0 0 8 2】

そして、上述の対話装置1 0 1の対話制御処理により、サーバ装置2 0 1との通信が切断している間の音声情報に対する応答文情報リストをサーバ装置2 0 1から取得することにより、対話装置1 0 1は、複数の応答文による応答を行うことができる。これによって、1つの応答文による応答に比べ、ユーザの発話をよりきちんと聞いていたかのように思

10

20

30

40

50

わせる応答を行うことができる。

【 0 0 8 3 】

例えば、図 4 の N o . 8 から N o . 1 2 に示すユーザの発話内容に対し、対話装置 1 0 1 はその時点では応答文を返答できないが、サーバ装置 2 0 1 との通信が回復した時点で、これらの N o . 8 から N o . 1 2 に示すユーザの発話内容がサーバ装置 2 0 1 に送信される。そして、サーバ装置 2 0 1 の応答文作成処理により、これらのユーザの発話内容から、図 1 1 の N o . 2 及び N o . 3 に示す応答文情報リストが作成される。そして、対話装置 1 0 1 の応答文情報取得部 1 1 3 は、この応答文情報リストを取得し、応答部 1 1 4 により、対話装置 1 0 1 はユーザに対して「そういえば、映画と言ってましたけど、映画って良いよね。私も映画大好き。」、「そういえば、かわいいと言ってましたけど、かわ

10

【 0 0 8 4 】

このように、対話装置 1 0 1 は、サーバ装置 2 0 1 との通信が切断している時には小まめな応答を行うことができないが、通信が回復した時に、切断中のユーザの発話内容に複数の話題が含まれていても、それぞれの話題中の特徴単語（最も多く使われている特定ワード等）に基づいた応答文を発話することができる。したがって、対話装置 1 0 1 は、各話題についてきちんとユーザの発話内容を聞いていたことを示すことができる。このように、対話装置 1 0 1 は、通信状況が悪い場合の受け答え技術をさらに改善することができる。

【 0 0 8 5 】

20

（第 3 実施形態）

対話装置が自己の位置を取得できるようにすると、応答文に位置に関する情報を含めることができるようになり、ユーザの発話内容をどこで聞いていたかということも示すことができるようになる。このような第 3 実施形態について説明する。

【 0 0 8 6 】

第 3 実施形態に係る対話システム 1 0 0 2 が対話装置 1 0 2 とサーバ装置 2 0 2 とを備える点は、第 1 実施形態に係る対話システム 1 0 0 0 と同じである。第 3 実施形態に係る対話装置 1 0 2 は、第 1 実施形態に係る対話装置 1 0 0 と同じ外観である。対話装置 1 0 2 の機能構成は、図 1 4 に示すように、第 1 実施形態に係る対話装置 1 0 0 と比較して、位置取得部 2 6 を備える点と、記憶部 1 2 0 に位置履歴データ 1 2 3 を記憶する点が異なる。また、サーバ装置 2 0 2 の機能構成は、第 1 実施形態に係るサーバ装置 2 0 0 と同じである。

30

【 0 0 8 7 】

位置取得部 2 6 は、GPS (Global Positioning System) 衛星からの電波を受信することによって、自己位置の座標（位置データ）を取得することができる。自己位置の座標の情報は、緯度及び経度で表されている。

【 0 0 8 8 】

位置履歴データ 1 2 3 は、図 1 5 に示すように、自己位置を取得した日時と、自己位置の座標（緯度及び経度）と、のペアの履歴である。

【 0 0 8 9 】

40

次に、対話装置 1 0 2 の制御部 1 1 0 が行う対話制御処理について、図 1 6 を参照して説明する。この処理は、第 1 実施形態に係る対話装置 1 0 0 の対話制御処理（図 7 ）と比べ、一部を除いて同じなので、異なる点を中心に説明する。

【 0 0 9 0 】

ステップ S 1 0 1 からステップ S 1 0 3、ステップ S 1 0 5 からステップ S 1 0 6 及びステップ S 1 1 0 からステップ S 1 1 3 は、図 7 を参照して説明した処理と同じである。ステップ S 1 0 3 の次のステップであるステップ S 1 3 1 では、音声記録部 1 1 1 としての制御部 1 1 0 は、マイク 2 1 が取得した音声データを、通信状態（切断）及び現在時刻とともに、付加情報付音声情報 1 2 1 として記憶部 1 2 0 に記録し、また、制御部 1 1 0 は、位置取得部 2 6 が取得した位置データを取得日時とともに、位置履歴データ 1 2 3 と

50

して記憶部 120 に記憶する。

【0091】

そして、ステップ S106 の次のステップであるステップ S132 では、制御部 110 は、ステップ S102 で記憶部 220 に記憶した通信切断時刻から現在時刻までの（通信切断中の）付加情報付音声情報 121 と位置履歴データ 123 とを、通信部 25 を介して、サーバ装置 202 に送信する。ここで送信した付加情報付音声情報 121 及び位置履歴データ 123 はサーバ装置 202 で音声認識及び場所名検索され、サーバ装置 202 は対話装置 102 に特徴単語、応答文及び位置に対応する場所名を送信する。具体例を示すと、位置に対応する場所名が存在すれば、サーバ装置 202 は、例えば、図 17 の No. 1 に示すように、特徴単語「暑い」、応答文及び場所名「第一公園」を送信する。また、位置に対応する場所名が存在しなければ、サーバ装置 202 は、例えば、図 17 の No. 2 に示すように、特徴単語「映画」、応答文及び場所名が無いことを示すデータ「- - -」を送信する。このサーバ装置 202 による処理（応答文作成処理）については後述する。

10

【0092】

すると、応答文情報取得部 113 としての制御部 110 は、通信部 25 を介して、サーバ装置 202 が送信した特徴単語、応答文情報（本実施形態においては応答文そのもの）及び位置に対応する場所名を取得する（ステップ S133）。そして、応答部 114 としての制御部 110 は、位置に対応する場所名が存在するか否かを判定する（ステップ S134）。位置に対応する場所名が存在するなら（ステップ S134；Yes）、応答文情報取得部 113 は、取得した応答文情報に、場所に関する前置きを追加する（ステップ S135）。場所に関する前置きとは、例えば「そういえば、さっき公園にいたときに、暑いって言ってたけど」というような句である。より一般的には、「そういえば、さっき[位置に対応する場所名]にいたときに、[特徴単語]って言ってたけど」と表すことができる。なお、位置に対応する場所名が存在しないなら（ステップ S134；No）、前置きを追加すること無く、ステップ S136 に進む。

20

【0093】

そして、応答部 114 としての制御部 110 は、応答文情報取得部 113 が取得した応答文情報（ステップ S135 で前置きが追加された場合は、前置き付きの応答文情報）に基づき、ユーザに回答する（ステップ S136）。本実施形態では、応答文情報は応答文そのものなので、具体的には、応答部 114 は、応答文（又は前置き付きの応答文）の内容を音声合成して、スピーカ 23 から応答文を発話する。そして、制御部 110 は、処理をステップ S101 に戻す。

30

【0094】

次に、サーバ装置 202 が行う応答文作成処理について、図 18 を参照して説明する。この処理は、第 1 実施形態に係るサーバ装置 200 の応答文作成処理（図 9）と比べ、一部を除いて同じなので、異なる点を中心に説明する。

【0095】

ステップ S301 からステップ S302、ステップ S303 からステップ S305 及びステップ S307 からステップ S309 は、図 9 を参照して説明した処理と同じである。ステップ S302 の判定が Yes の場合の処理であるステップ S331 では、通信部 230 は、対話装置 102 が送信した位置履歴データ 123 を受信する。そして、制御部 210 は、位置履歴データ 123 に含まれる各座標について、緯度及び経度から場所名を取得するクラウドサービスを利用して、場所名を取得する（ステップ S332）。例えば、google（登録商標）や、ゼンリン（登録商標）等の地図データベースを保有する企業から情報提供を受けることにより、ビル名などのかなり詳細な場所名を取得することができる。ただし、場所名が定義されていない座標も存在するので、場所名が取得できないこともある。

40

【0096】

そして、ステップ S305 の次のステップであるステップ S333 では、制御部 210 は、ステップ S332 で場所名を取得できたか否かを判定する。場所名が取得できたなら

50

(ステップS333; Yes)、応答作成部213は、ステップS304で抽出した特徴単語、ステップS305で作成した応答文情報、及び、ステップS332で取得した場所名を通信部230を介して対話装置102に送信する(ステップS334)。この送信データは、例えば図17のNo. 1やNo. 3に示すようなデータである。

【0097】

場所名が取得できなかったなら(ステップS333; No)、応答作成部213は、ステップS304で抽出した特徴単語、ステップS305で作成した応答文情報、及び、場所名が無いことを示すデータを、通信部230を介して対話装置102に送信する(ステップS335)。この送信データは、例えば図17のNo. 2に示すようなデータである。

10

【0098】

そして、何れの場合(場所名が取得できた場合も取得できない場合も)も、その後、ステップS301に戻る。

【0099】

以上説明した応答文作成処理により、通信切断中の発話内容に対する応答文情報には、特徴単語の情報と場所名の情報を付加して対話装置102に送信することができる。そして、上述の対話装置102の対話制御処理により、サーバ装置202との通信が切断している間の音声情報に対する応答文情報を、サーバ装置202から取得することにより、対話装置102は、ユーザがどの場所でどのような話をしたかをきちんと聞いていたかのように思わせる応答文で応答することができる。このように、対話装置102は、通信状況が悪い場合の受け答え技術をさらに改善することができる。

20

【0100】

(変形例)

上述の各実施形態は任意に組み合わせることができる。例えば第2実施形態と第3実施形態とを組み合わせることにより、複数の話題に対応した応答文を、各話題を発話した場所についての前置きとともに発話させることができるようになる。これにより、例えば、「そういえば、さっき、第一公園にいたときに、暑いって言ってたけど、暑い暑い言っていると余計暑くなるよ。」、「そういえば、映画って言ってましたけど、映画って良いよね。私も映画大好き。」、「そういえば、さっき、第三食堂にいたときに、かわいって言ってたけど、かわいって私のこと?嬉しい。」のような発話を対話装置にさせることができる。これにより、対話装置がサーバ装置と通信できない状態のときのユーザの発話内容の話題の変化や各話題がどの場所で発話されたかに対して、あたかも対話装置がきちんと聞いていたかのように受け答えすることができる。したがって、この対話装置の変形例は、通信状況が悪い場合の受け答え技術をさらに改善することができる。

30

【0101】

また、上述の各実施形態では、サーバ装置と対話装置の通信環境の乱れを想定して説明したが、節電等のために意図的に両装置間の通信を遮断した場合にも適用できるものである。

【0102】

また、上述の各実施形態では、対話装置がユーザ1名に対応しているイメージで説明したが、対話装置が個人認識機能を搭載することにより、複数のユーザにそれぞれ対応した受け答えをすることが可能である。

40

【0103】

なお、対話装置100, 101, 102の各機能は、通常のPC(Personal Computer)等のコンピュータによっても実施することができる。具体的には、上記実施形態では、対話装置100, 101, 102が行う対話制御処理等のプログラムが、記憶部120のROMに予め記憶されているものとして説明した。しかし、プログラムを、フレキシブルディスク、CD-ROM(Compact Disc Read Only Memory)、DVD(Digital Versatile Disc)及びMO(Magneto-Optical Disc)等のコンピュータ読み取り可能な記

50

録媒体に格納して配布し、そのプログラムをコンピュータに読み込んでインストールすることにより、上述の各機能を実現することができるコンピュータを構成してもよい。

【0104】

以上、本発明の好ましい実施形態について説明したが、本発明は係る特定の実施形態に限定されるものではなく、本発明には、特許請求の範囲に記載された発明とその均等の範囲が含まれる。以下に、本願出願の当初の特許請求の範囲に記載された発明を付記する。

【0105】

(付記1)

ユーザが発話した音声に対する応答文を外部のサーバ装置と通信しながら作成する対話装置であって、

ユーザが発話した音声を音声データとして取得する音声取得部と、

前記音声取得部が取得した音声データに基づく音声情報を記録する音声記録部と、

前記サーバ装置と通信する通信部と、

前記通信部による前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録部が記録した音声情報を前記サーバ装置に送信し、前記音声情報に対する応答文情報を前記サーバ装置から取得する、応答文情報取得部と、

前記応答文情報取得部が取得した応答文情報に基づいて作成された応答文でユーザに回答する応答部と、

を備える対話装置。

【0106】

(付記2)

前記通信部による前記サーバ装置との通信が切断している間ユーザに聞いている風に見せかける見せかけ部をさらに備える、

付記1に記載の対話装置。

【0107】

(付記3)

前記見せかけ部は、前記音声取得部が取得した前記音声データに応じてうなづく、相づちを打つ、つぶやく、の少なくとも1つを実行する、

付記2に記載の対話装置。

【0108】

(付記4)

前記見せかけ部は、釈明基準時間が経過すると、ユーザに対し適切な応答ができないことを釈明する、

付記2又は3に記載の対話装置。

【0109】

(付記5)

前記応答文は、前記音声データを音声認識して取得したテキストデータに含まれる特徴単語に基づいて作成される、

付記1から4のいずれか1つに記載の対話装置。

【0110】

(付記6)

前記特徴単語は、前記音声データを音声認識して取得したテキストデータに最も多く含まれる特定ワードである、

付記5に記載の対話装置。

【0111】

(付記7)

前記特徴単語は、前記音声データを音声認識して取得したテキストデータに含まれる特定ワードのうち、強調修飾語に修飾された特定ワードである、

付記5に記載の対話装置。

【0112】

10

20

30

40

50

(付記 8)

前記応答文は、前記特徴単語に応答文作成ルールを適用することによって作成される、
付記 5 から 7 のいずれか 1 つに記載の対話装置。

【 0 1 1 3 】

(付記 9)

前記応答文情報取得部は、通信切断中に前記音声記録部が記録した音声情報の話題毎の
前記音声情報に対する応答文情報を前記サーバ装置から取得し、

前記応答部は、前記応答文情報取得部が取得した話題毎の応答文情報に基づいて作成し
た応答文でユーザに回答する、

付記 1 から 8 のいずれか 1 つに記載の対話装置。

10

【 0 1 1 4 】

(付記 10)

前記応答部は、前記応答文情報取得部が取得した応答文情報に基づいて作成した応答文
に前置きを追加した応答文で、ユーザに回答する、

付記 1 から 9 のいずれか 1 つに記載の対話装置。

【 0 1 1 5 】

(付記 11)

自己の位置データを取得する位置取得部をさらに備え、

前記応答文情報取得部は、前記通信部による前記サーバ装置との通信が一時的に切断し
た後に回復した状態において、通信切断中に前記音声記録部が記録した音声情報及び通信
切断中に前記位置取得部が取得した位置データを前記サーバ装置に送信し、前記音声情報
に対する応答文情報及び前記位置データに対応する場所名を前記サーバ装置から取得し、

20

前記応答部は、前記応答文情報取得部が取得した応答文情報に基づいて作成された応答
文に、前記応答文情報取得部が取得した場所名を含む前置きを追加した応答文でユーザに
回答する、

付記 1 から 10 のいずれか 1 つに記載の対話装置。

【 0 1 1 6 】

(付記 12)

制御部が、ユーザが発話した音声に基づく音声情報を記録し、

外部のサーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中
に記録された前記音声情報に対応する応答文情報を前記サーバ装置に作成させ、

30

前記サーバ装置から受信した前記応答文情報に基づいた応答文でユーザに回答する対話
方法。

【 0 1 1 7 】

(付記 13)

ユーザが発話した音声に対する応答文をサーバ装置と通信しながら作成する対話装置と
、前記サーバ装置と、を備える対話システムにおけるサーバ装置であって、

前記対話装置と通信する通信部と、

前記ユーザが発話した音声に基づく音声情報を前記対話装置から前記通信部を介して受
信する受信部と、

40

前記受信部が受信した音声情報に含まれる音声データを音声認識してテキストデータを
生成する音声認識部と、

前記音声認識部が生成したテキストデータから該テキストデータに含まれる特徴的な単
語である特徴単語を抽出する特徴単語抽出部と、

前記特徴単語抽出部が抽出した特徴単語に基づき、応答文情報を作成する応答作成部と

、

前記応答作成部が作成した応答文情報を前記通信部を介して送信する送信部と、

を備え、

前記通信部による前記対話装置との通信が一時的に切断した後に回復した状態において
、通信切断中の音声情報を前記対話装置から受信し、前記受信した音声情報に対する応答

50

文情報を作成して前記対話装置に送信する、
ことを特徴とするサーバ装置。

【 0 1 1 8 】

(付 記 1 4)

ユーザが発話した音声に対する応答文を外部のサーバ装置と通信しながら作成する対話装置のコンピュータに、

ユーザが発話した音声に基づく音声情報を記録する音声記録ステップ、

前記サーバ装置との通信が一時的に切断した後に回復した状態において、通信切断中に前記音声記録ステップで記録した音声情報を前記サーバ装置に送信し、前記音声情報に対する応答文情報を前記サーバ装置から取得する、応答文情報取得ステップ、及び、

前記応答文情報取得ステップで取得した応答文情報に基づいて作成された応答文でユーザに応答する応答ステップ、

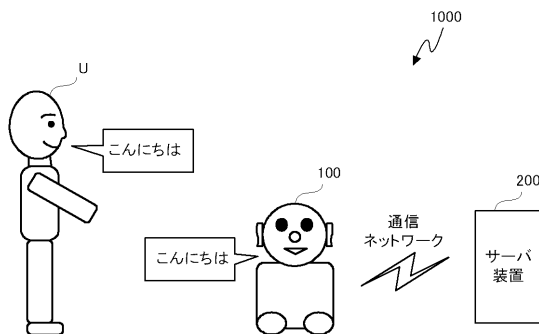
を実行させるためのプログラム。

【 符号の説明 】

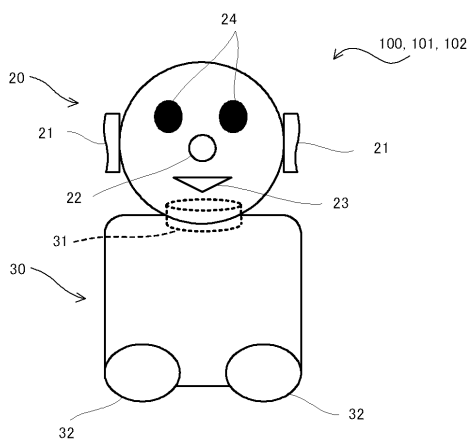
【 0 1 1 9 】

2 0 ... 頭、2 1 ... マイク、2 2 ... カメラ、2 3 ... スピーカ、2 4 ... センサ群、2 5 , 2 3 0 ... 通信部、2 6 ... 位置取得部、3 0 ... 胴体、3 1 ... 首関節、3 2 ... 足回り部、3 3 ... 操作ボタン、1 0 0 , 1 0 1 , 1 0 2 ... 対話装置、1 1 0 , 2 1 0 ... 制御部、1 1 1 ... 音声記録部、1 1 2 ... 見せかけ部、1 1 3 ... 応答文情報取得部、1 1 4 ... 応答部、1 2 0 , 2 2 0 ... 記憶部、1 2 1 ... 付加情報付音声情報、1 2 2 ... 応答文情報リスト、1 2 3 ... 位置履歴データ、2 0 0 , 2 0 1 , 2 0 2 ... サーバ装置、2 1 1 ... 音声認識部、2 1 2 ... 特徴単語抽出部、2 1 3 ... 応答作成部、2 2 1 ... 応答文作成ルール、1 0 0 0 , 1 0 0 1 , 1 0 0 2 ... 対話システム、U ... ユーザ

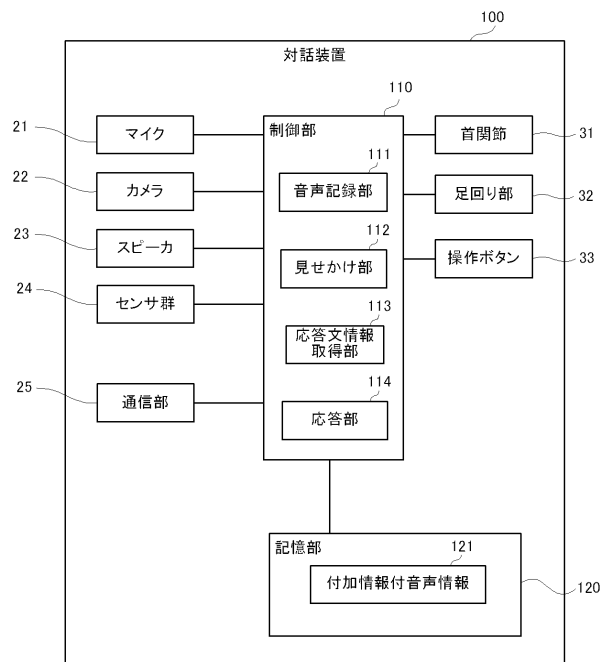
【 図 1 】



【 図 2 】



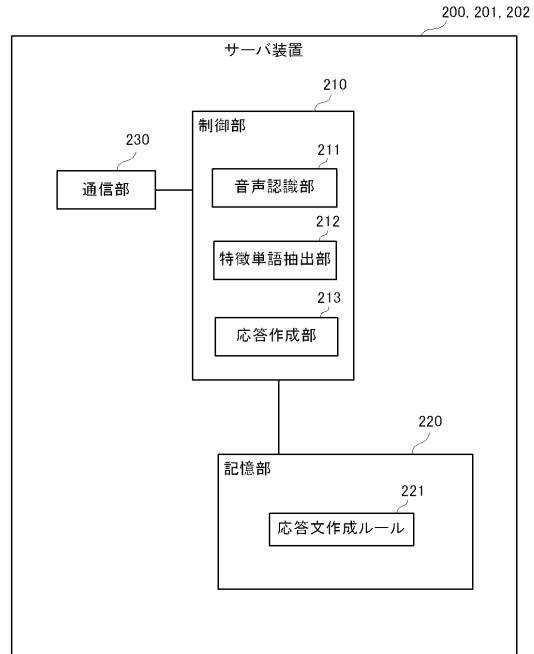
【 図 3 】



【図 4】

No.	通信状態	発話日時	ユーザの発話内容
1	切断	2017.09.05 10:03:05	ねえ、暑くない？
2	切断	2017.09.05 10:03:07	今年の夏はとても暑くなるみたい。
3	切断	2017.09.05 10:03:11	無理しないようにしてね。
4	接続	2017.09.05 10:05:21	ただいま。
5	接続	2017.09.05 10:05:23	家は涼しくて良いね。
6	接続	2017.09.05 10:05:26	エアコンが効いているからかな。
7	接続	2017.09.05 10:05:31	公園は暑かったね。
8	切断	2017.09.05 10:23:15	この映画楽しみな。
9	切断	2017.09.05 10:23:18	今年のお勧めの映画はある？
10	切断	2017.09.05 10:23:21	あなたの好きな映画を教えて。
11	切断	2017.09.05 10:31:31	あのぬいぐるみ超かわいいね。
12	切断	2017.09.05 10:31:37	私かわいいもの大好き。
...

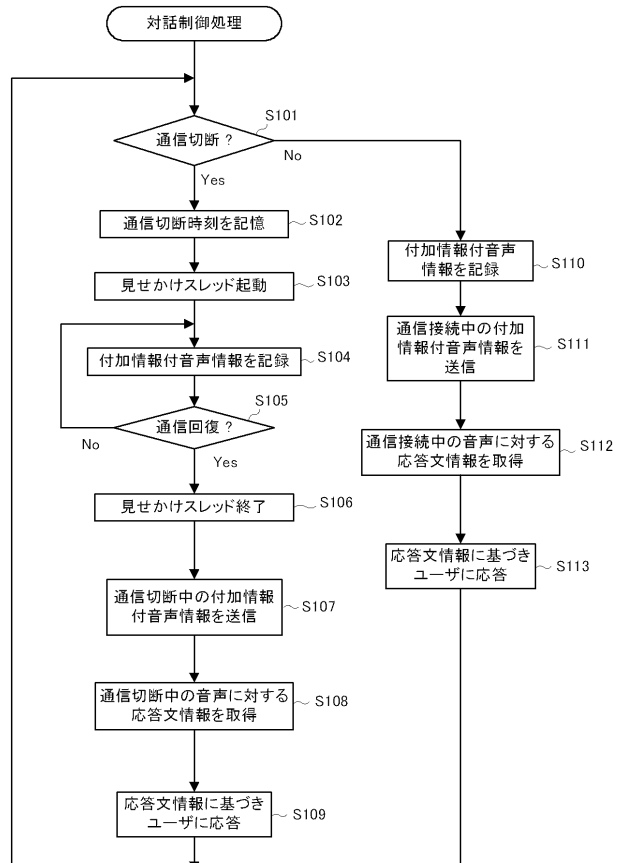
【図 5】



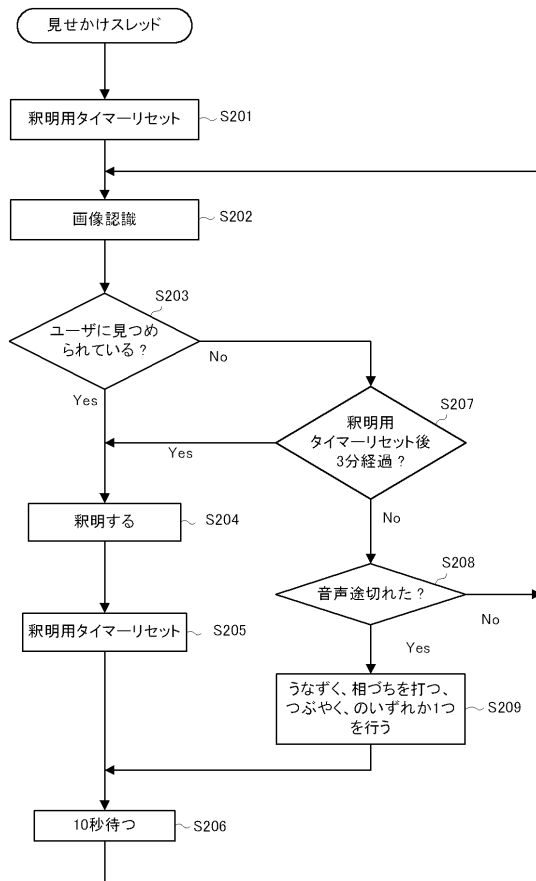
【図 6】

No.	特徴単語	応答文
1	暑い	暑い暑い言っていると余計暑くなるよ。
2	映画	映画って良いよね。私も映画大好き。
3	かわいい	かわいって私のこと？ 嬉しい。
...

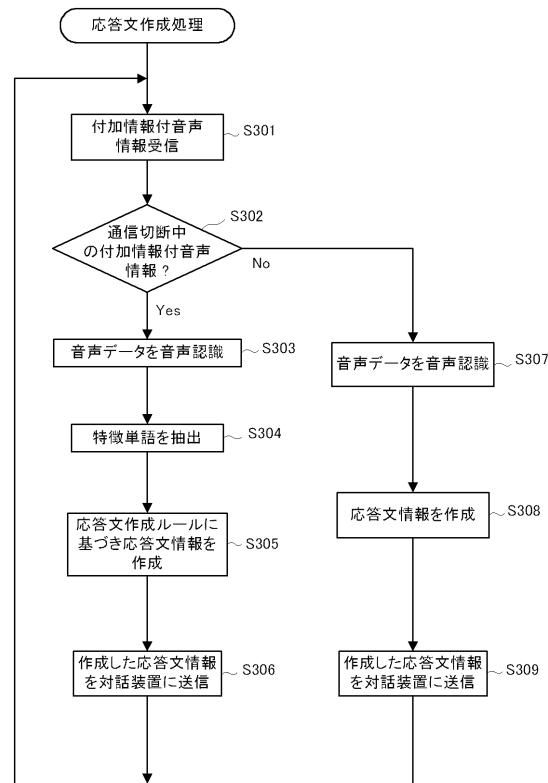
【図 7】



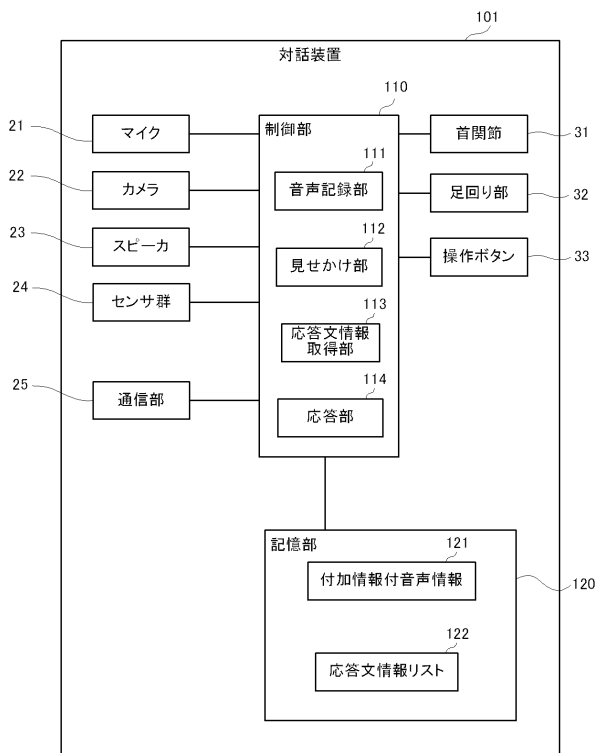
【図 8】



【図 9】



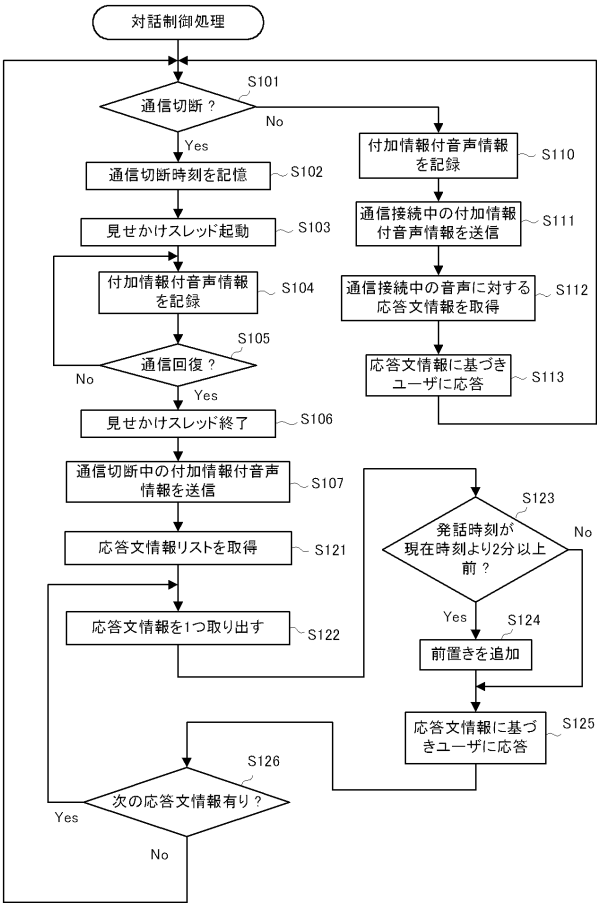
【図 10】



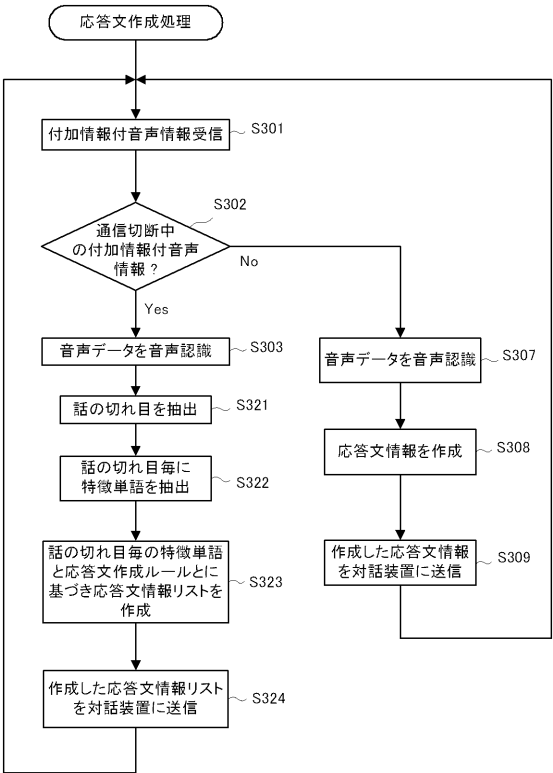
【図 11】

No.	発話日時	特徴単語	ユーザの音声に対する応答文
1	2017.09.05 10:03:05 ~ 2017.09.05 10:03:11	暑い	暑い暑い言っていると余計暑くなるよ。
2	2017.09.05 10:23:15 ~ 2017.09.05 10:23:21	映画	映画って良いよね。私も映画大好き。
3	2017.09.05 10:31:31 ~ 2017.09.05 10:31:37	かわいい	かわいいって私のこと？ 嬉しい。
...

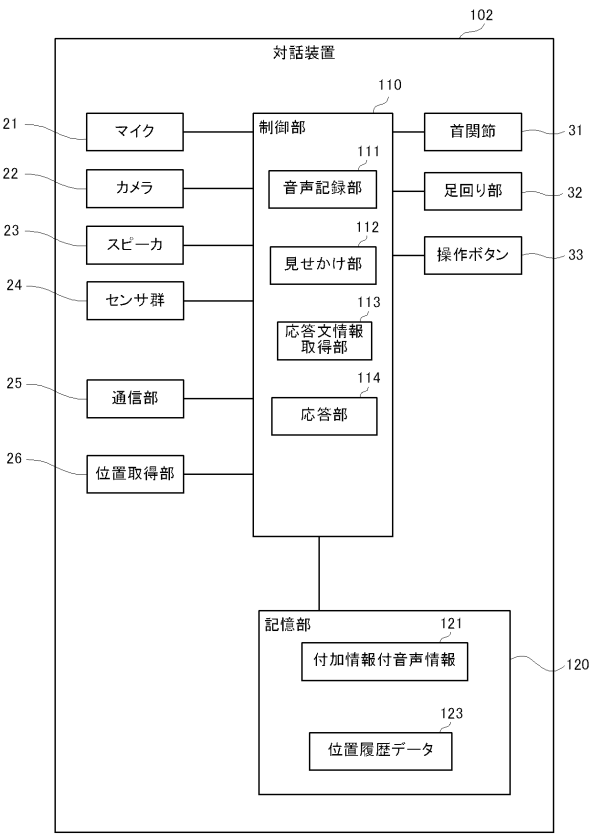
【図 1 2】



【図 1 3】



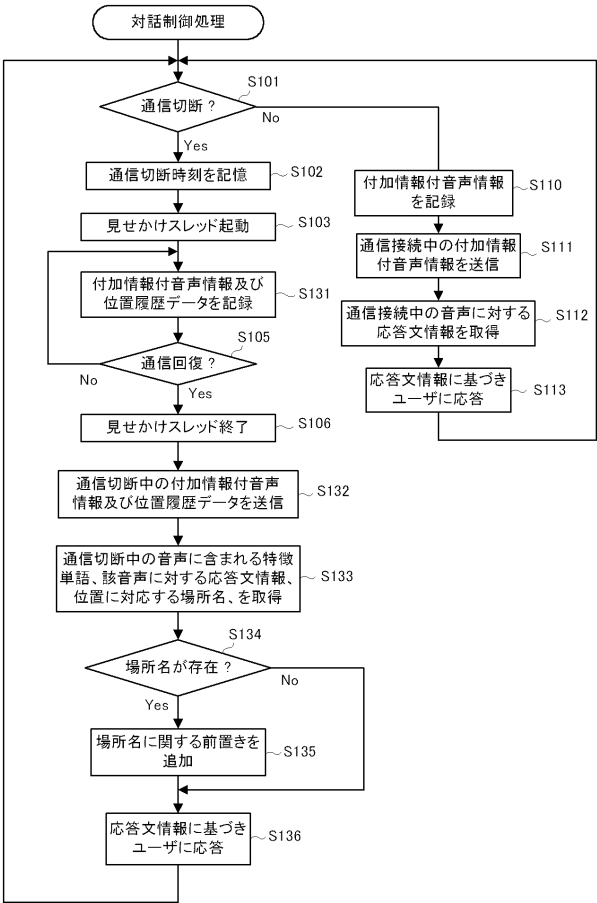
【図 1 4】



【図 1 5】

No.	取得日時	自己位置の座標(緯度及び経度)
1	2017.09.05 10:04:10	(36.0001, 140.0002)
2	2017.09.05 10:24:20	(35.9001, 139.9002)
3	2017.09.05 10:32:40	(35.8902, 139.9001)
...

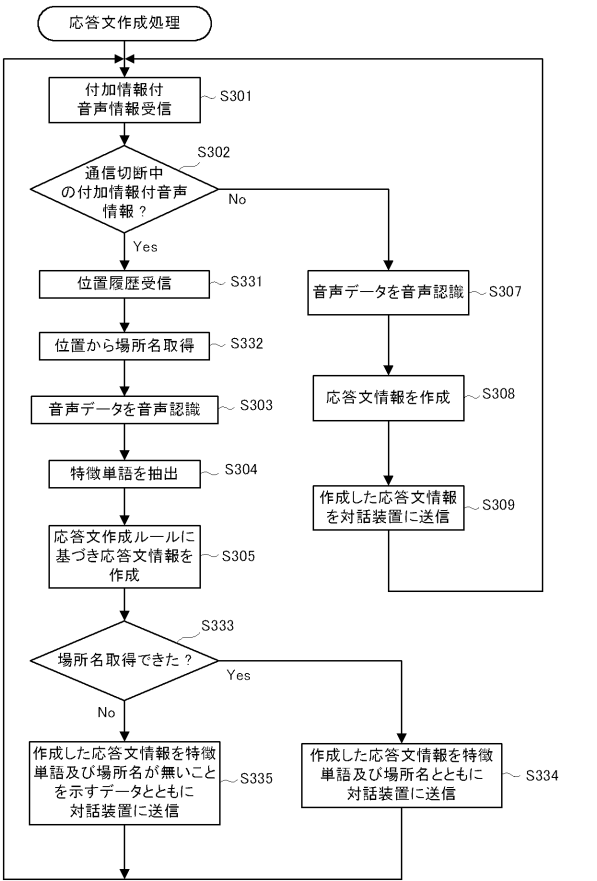
【図 16】



【図 17】

No.	特徴単語	応答文	場所名
1	暑い	暑い暑い言っていると余計暑くなるよ。	第一公園
2	映画	映画って良いよね。私も映画大好き。	---
3	かわいい	かわいいって私のこと？ 嬉しい。	第三食堂
...

【図 18】



フロントページの続き

(56)参考文献 特開 2 0 1 5 - 1 8 3 2 7 (J P , A)
特開 2 0 1 5 - 1 8 4 5 6 3 (J P , A)
特開 2 0 0 9 - 1 9 8 8 7 1 (J P , A)
特開 2 0 0 8 - 8 3 1 0 0 (J P , A)
特開 2 0 1 7 - 4 9 4 7 1 (J P , A)
特開 2 0 1 4 - 1 7 4 4 8 5 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)
G 1 0 L 1 3 / 0 0 - 1 5 / 3 4