(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2019-191490
(P2019-191490A)
(43)公開日　令和1年10月31日(2019.10.31)

(51)Int.Cl.		F I		テーマコード（参考）
G10L 15/20	(2006.01)	G10L	15/20	36OB
G10L 15/10	(2006.01)	G10L	15/10	2OOW
G10L 15/30	(2013.01)	G10L	15/30	

審査請求 未請求　請求項の数 9　OL　（全 19 頁）

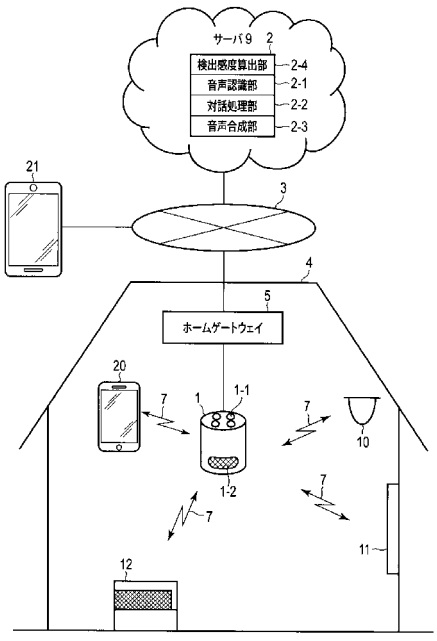
(21)出願番号　特願2018-86985(P2018-86985)
(22)出願日　平成30年4月27日(2018.4.27)

(71)出願人　000214984
東芝映像ソリューション株式会社
青森県三沢市南町三丁目３１番地２７７６号
(74)代理人　110001737
特許業務法人スズエ国際特許事務所
(72)発明者　向出　隆信
青森県三沢市南町三丁目３１番地２７７６号　東芝映像ソリューション株式会社内

(54)【発明の名称】音声対話端末、および音声対話端末制御方法

(57)【要約】
【課題】音声対話端末が特定ワードを認識する感度を、ユーザが調整することが可能な音声対話端末および音声対話端末制御方法を提供することを目的とする。
【解決手段】本発明の実施形態は、マイクと、ネットワークとの通信を制御する通信制御部と、前記マイクから入力された音声を、前記通信制御部を介して前記ネットワークに出力する出力処理部と、前記マイクから入力された前記音声のうち、予め定められたワードである特定ワードを検出する検出部と、前記通信制御部を介して前記ネットワークより入力された、前記検出部が前記特定ワードを検出するための感度の値である感度設定値を、前記検出部に設定する設定部と、を具備する音声対話端末である。
【選択図】図１

図1



【特許請求の範囲】
【請求項１】
マイクと、
ネットワークとの通信を制御する通信制御部と、
前記マイクから入力された音声を、前記通信制御部を介して前記ネットワークに出力する出力処理部と、
前記マイクから入力された前記音声のうち、予め定められたワードである特定ワードを検出する検出部と、
前記通信制御部を介して前記ネットワークより入力された、前記検出部が前記特定ワードを検出するための感度の値である感度設定値を、前記検出部に設定する設定部と、
を具備する音声対話端末。
【請求項２】
前記感度は、前記マイクより入力された音声と、前記特定ワードとの類似度である、請求項１に記載の音声対話端末。
【請求項３】
前記出力処理部は、前記通信制御部を介してネットワークから入力された感度設定開始イベントから感度設定終了イベントまでの間に、前記マイクから入力された第１音声を、前記通信制御部を介して前記ネットワークに出力し、
前記通信制御部を介して前記ネットワークより入力された、前記第１音声に基づいて算出された前記感度設定値を前記設定部が前記検出部に設定する、請求項１に記載の音声対話端末。
【請求項４】
前記出力処理部は、前記マイクから入力された、感度設定開始を示す第１特定ワードから感度設定終了を示す第２特定ワードまでの間に、前記マイクから入力された第２音声を、前記通信制御部を介して前記ネットワークに出力し、
前記通信制御部を介して前記ネットワークより入力された、前記第２音声に基づいて算出された前記感度設定値を前記設定部が前記検出部に設定する、請求項１に記載の音声対話端末。
【請求項５】
前記感度設定値は、予め関連付けられている携帯端末から送られてくる請求項３に記載の音声対話端末。
【請求項６】
マイクと、
前記マイクから入力された音声のうち、予め定められたワードである特定ワードを検出する検出部と、
感度設定開始を示す開始イベントから感度設定終了を示す終了イベントまでの間に前記マイクから入力された第１音声に基づいて、前記検出部が前記特定ワードを検出するための感度の値である感度設定値を算出する感度算出部と、
前記感度算出部が算出した前記感度設定値を、前記検出部に設定する設定部と、
を具備する音声対話端末。
【請求項７】
さらに携帯端末との通信を制御する通信制御部を具備し、前記開始イベントおよび前記終了イベントは、前記通信制御部を介して前記携帯端末から入力される、請求項６に記載の音声対話端末。
【請求項８】
前記感度算出部は、前記マイクから入力された第１特定ワードを、前記感度設定開始を示す前記開始イベントと認識し、前記マイクから入力された第２特定ワードを、前記感度設定終了を示す前記終了イベントと認識する、請求項６に記載の音声対話端末。
【請求項９】
マイクから入力された音声を、ネットワークに出力するステップと、

前記マイクから入力された音声のうち、予め定められたワードである特定ワードを検出するステップと、
前記ネットワークから入力された前記特定ワードを検出する感度の値である感度設定値に更新して前記特定ワードを検出するステップと、
からなる音声対話端末制御方法。
【発明の詳細な説明】
【技術分野】
【０００１】
本実施形態は、音声対話端末および音声対話端末の制御方法に関するものである。
【背景技術】
【０００２】
ユーザが発した音声を例えばマイクで収集し、この収集した音声を音声認識処理により解析することでユーザが発した内容を判別し、この判別した内容に応じた応答をユーザに提供する音声対話システムがある。この音声対話システムは、音声対話サービスの部分と音声対話端末の部分の大きく２つの部分を含む。
【０００３】
音声対話サービスの部分は、ユーザが発した内容を入力として、この入力された内容を音声認識処理により解析し、この解析結果に応じた応答をユーザに提供する機能を持つ。
【０００４】
音声対話端末の部分は、マイクで収集したユーザの発話内容を、音声データとして音声対話サービスに入力したり、音声対話サービスが出力する応答の内容を、ユーザに対して音声として出力したり周辺機器の制御を行ったりする機能を持つ。
【０００５】
さらに、音声対話サービスの部分は、ユーザが発した内容の入力に対応した応答の提供ではなく、音声対話サービスが自発的に音声対話端末に情報を提供する場合もある。
【先行技術文献】
【特許文献】
【０００６】
【特許文献１】特開２０１７－５３７３６１号公報
【発明の概要】
【発明が解決しようとする課題】
【０００７】
音声対話端末は、ユーザの発話内容をマイクで収集する際に、予め決められた特定ワードを認識することで、以降のユーザの発話内容の収集を行う。
【０００８】
音声対話端末が置かれている場所は、ユーザが特定ワードを発話する以外に、日常会話やテレビジョン等のＡＶ機器の音等の室内雑音が発生している場合が一般的である。このため音声対話端末が、これらの室内雑音をユーザが特定ワードを発したと誤認識してしまう場合がある。このため音声対話端末が置かれている場所の室内雑音の状況に応じて、音声対話システムを使うユーザが、音声対話端末が特定ワードを認識する感度を調整できることが望ましい。
【０００９】
しかし現状の音声対話システムでは、音声対話端末が特定ワードを認識する感度を、ユーザが調整する仕組みがない、という課題があった。
【００１０】
そこで本発明の本実施形態では、音声対話端末が特定ワードを認識する感度を、ユーザが調整することが可能な音声対話端末および音声対話端末制御方法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】
【００１１】

マイクと、
ネットワークとの通信を制御する通信制御部と、
前記マイクから入力された音声を、前記通信制御部を介して前記ネットワークに出力する出力処理部と、
前記マイクから入力された前記音声のうち、予め定められたワードである特定ワードを検出する検出部と、
前記通信制御部を介して前記ネットワークより入力された、前記検出部が前記特定ワードを検出するための感度の値である感度設定値を、前記検出部に設定する設定部と、
を具備する音声対話端末である。
【図面の簡単な説明】
【００１２】
【図１】図１は、本発明の実施形態が適用された音声対話端末１を含む音声対話システムの概要を示す図である。
【図２】図２は、図１に示した音声対話端末１の構成図である。
【図３】図３は、第１の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れを示す図である。
【図４】図４は、携帯端末２０あるいは携帯端末２１にインストールされている音声対話端末１を制御するための感度設定アプリの、感度設定値を設定する設定画面の例である。
【図５Ａ】図５Ａは、第１の実施形態における音声対話端末１のトリガワードの検出感度の感度設定値を、ユーザが携帯端末２０あるいは携帯端末２１にインストールされているアプリケーションを操作して変更する場合の、音声対話端末１の処理フローである。
【図５Ｂ】図５Ｂは、第１の実施形態における音声対話端末１のトリガワードの検出感度の感度設定値を、ユーザが携帯端末２０あるいは携帯端末２１にインストールされているアプリケーションを操作して変更する場合の、携帯端末２０あるいは携帯端末２１の処理フローである。
【図６】図６は、第２の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れの例を示す図である。
【図７】図７は、携帯端末２０あるいは携帯端末２１にインストールされている音声対話端末１を制御するためのアプリケーションにおける、周辺雑音の測定の開始および終了の設定画面の例である。
【図８Ａ】図８Ａは、第２の実施形態における音声対話端末が、携帯端末２１で設定した期間収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データにより算出した感度設定値を用いて、感度設定値を更新する場合の、音声対話端末１の処理フローである。
【図８Ｂ】図８Ｂは、第２の実施形態における音声対話端末が、携帯端末２１で設定した期間収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データにより算出した感度設定値を用いて、感度設定値を更新する場合の、携帯端末２１の処理フローである。
【図８Ｃ】図８Ｃは、第２の実施形態における音声対話端末が、携帯端末２１で設定した期間収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データにより算出した感度設定値を用いて、感度設定値を更新する場合の、検出感度算出部２－４の処理フローである。
【図９】図９は、第３の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れの例を示す図である。
【図１０Ａ】図１０Ａは、第３の実施形態における音声対話端末が、ユーザの発話の内容により設定した期間において収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データを用いて算出した感度設定値を用いて、感度設定値を更新する場合の、音声対話端末１の処理フローである。
【図１０Ｂ】図１０Ｂは、第３の実施形態における音声対話端末が、ユーザの発話の内容により設定した期間において収集した周辺の音声を音声対話サービス２に送信し、音声対

話サービス２において音声対話端末１が収集した音声データを用いて算出した感度設定値を用いて、感度設定値を更新する場合の、検出感度算出部２－４の処理フローである。
【図１１】図１１は、第４の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れの例を示す図である。
【図１２】図１２は、第４の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れの他の例を示す図である。
【発明を実施するための形態】
【００１３】
以下、本発明の実施の形態について図面を参照して説明する。
【００１４】
図１は、本発明の実施形態が適用された音声対話端末１を含む音声対話システムの概要を示す図である。本音声対話システムは、例えば家屋４に配置された音声対話端末１とクラウド上に存在するサーバ９に配置された音声対話サービス２とからなる、音声対話端末１と音声対話サービス２は、家屋４内に配置されたホームゲートウェイ５を経由してネットワーク３を介して互いに通信を行うことが可能である。
【００１５】
音声対話端末１は、またＢｌｕｅｔｏｏｔｈ（登録商標）、ＺｉｇＢｅｅ（登録商標）、Ｗｉ－Ｆｉ等の近距離無線通信システム７を介して、家屋４の中に設置されている照明１０やエア・コンディショナー（エアコン）１１、録画再生機器１２と通信を行うことが可能である。また音声対話端末１は、赤外線通信のようなペアリングを必要としない通信方式を介して周辺機器を制御することも可能である。また音声対話端末１は、ここに示した電子機器以外の電子機器とも通信を行うことも可能である。また、一例として近距離無線通信システムを用いるものとして説明するが必ずしも近距離無線通信システムである必要はなく、通常の無線システム或いは有線通信システムを用いてもよい。
【００１６】
音声対話端末１は、Ｂｌｕｅｔｏｏｔｈ、ＺｉｇＢｅｅ、Ｗｉ－Ｆｉ等の近距離無線通信システム７を介して携帯端末２０により、音声対話端末１の機能設定等種々の制御をすることが可能である。また音声対話端末１は、音声対話サービス２を含むサーバ９を経由してネットワーク３を介して携帯端末２１により、音声対話端末１の機能設定等種々の制御をすることが可能である。携帯端末２０および携帯端末２１には、音声対話端末１を制御するためのアプリケーション（感度設定アプリとも呼ぶ）がインストールされている。
【００１７】
音声対話サービス２は、例えば音声対話サービスＡ２－１（図示しない）と音声対話サービスＢ２－２（図示しない）の２つの音声対話サービスを含んでもよい。この場合音声対話サービスＡ２－１と音声対話サービスＢ２－２のいずれを用いるかは、ユーザが発する特定ワードによって決定される。音声対話サービス２は、３つ以上の音声対話サービスを含んでもよい。
【００１８】
音声対話端末１は、ユーザが発した言葉を備え付けのマイク１－１で収集し、特定ワードと認識すると、以降のユーザの発話内容を継続してマイク１－１を通して収集し、この収集したユーザの発話内容の音声データを、ネットワーク３を介して音声対話サービス２に送る。
【００１９】
音声対話サービス２は、音声認識部２－１、対話処理部２－２、音声合成部２－３、検出感度算出部２－４からなる。
【００２０】
音声対話端末１から送られてきた音声データを受信した音声対話サービス２は、受信した音声データの音声認識部２－１で解析を行い、この解析した内容に応じた応答を対話処理部２－２で生成する。
【００２１】

音声対話サービス２が生成する応答内容は、音声による応答とコマンドによる応答との２種類の応答を含む。いずれの種類の応答を生成するかは、音声対話端末１から送られてきた音声データを音声認識部２－１で解析した結果による。音声による応答とコマンドによる応答に関しては、以降で詳細に説明する。

【００２２】

応答を生成すると、この生成した応答の内容によりを音声合成部２－３で音声データに変換し、この変換した音声データを、ネットワーク３を介して音声対話端末１に送信する。

【００２３】

特定ワードとは、ユーザが発話する、音声対話端末１との対話処理を開始するための所定のキーワードのことであり、トリガワードあるいはウェイク表現あるいはボイストリガあるいは音声コマンドあるいは起動コマンドと呼んでもよい。以降の説明では、特定ワードをトリガワードと呼んで説明する。

【００２４】

音声対話端末１は、マイク１－１からユーザの発話内容を、図２に示すトリガワード検出処理部２０３によりトリガワードであると認識すると、トリガワード以降のユーザの発話内容を、音声対話端末１に対して話しかけられているものとして、ユーザの発話内容を継続して収集する。また音声対話端末１は、マイク１－１から一定時間ユーザの発話の入力がない場合、ユーザの発話が終了したと判断し、再びトリガワードが入力されるのを待つトリガワード入力待ち状態に遷移する。

【００２５】

トリガワードは、予め決めれたものが登録されてあってもよいし、音声対話端末１の初期設定時にユーザにより任意に登録されてもよい。

【００２６】

音声対話サービス２の対話処理部２－２が生成する応答は、音声による応答とコマンドによる応答の２種類の応答を含む。

【００２７】

音声による応答は、音声対話端末１から入力される音声データに応じて音声対話サービス２が、音声で生成する応答である。

【００２８】

コマンドによる応答は、音声対話端末１から入力される音声データに応じて、音声対話サービス２が、制御コマンドで生成する応答である。制御コマンドは、音声対話端末１がもつ電子機器（デバイス）あるいは音声対話端末１と近距離無線通信システム等を介して接続されている周辺機器（周辺デバイス）を制御するコマンドである。音声対話端末１がもつ電子機器（デバイス）は、例えば付属するカメラである。音声対話端末１と近距離無線通信システム等で接続されている周辺機器（周辺デバイス）は、例えば照明１０やエア・コンディショナー（エアコン）１１である。

【００２９】

音声データによる応答の応答内容は、ユーザが音声対話端末１に対して発話した例えば「おはよう」に対する「おはようございます。今日は元気ですか？」のようにユーザが発話した内容に対応した返事である。また例えば「今から新幹線に乗って大阪に行くと何時に着くかな？」という質問に対する「今から３０分後に出発すれば、夜の８時までに大阪駅に到着します」のようにユーザの問い掛けに対応した回答である。

【００３０】

音声対話サービス２から応答を受け取った音声対話端末１は、その応答が音声データによる応答の場合は、その応答の内容を、例えば備え付けのスピーカ１－２より音声として出力することができる。これによりユーザは、自らの発話に対する音声対話システムの応答を聞くことができる。

【００３１】

コマンドによる応答の応答内容は、ユーザが音声対話端末１に対して発話した例えば「

エアコンつけて」に対する「デバイス＝エアコン１０、操作＝ＯＮ、モード＝冷房、設定＝温度２６度、風量最大」の内容のコマンドである。また例えば「ちょっと電気つけて」に対する「デバイス＝照明１０、操作＝ＯＮ」の内容のコマンドである。

【００３２】

音声対話サービス２から応答を受け取った音声対話端末１は、その応答がコマンドによる応答の場合は、コマンドに含まれている制御対象のデバイスの制御を行う。例えばコマンドの内容が「デバイス＝エアコン１０、操作＝ＯＮ、モード＝冷房、設定＝温度２６度、風量最大」の場合、音声対話端末１は、エアコン１１を温度２６度、風量最大の設定で起動するように、内部に持つＷｉ-Ｆｉ、ＺｉｇＢｅｅ、Ｂｌｕｅｔｏｏｔｈ等の近距離無線通信システムを介して制御する。

【００３３】

コマンドによる応答の内容は、ユーザが音声対話端末１に対して発した例えば「ＡＡＡ動画サービスのＢＢＢコンテンツを再生して」に対する「ｐｌａｙ　ｆｒｏｍ　ｗｗｗ．ｘｘｘｘｘｘ．ｃｏ．ｊｐ／ｍｕｓｉｃＢＢＢ．ｗａｖ」のように、コマンドの部分である「ｐｌａｙ」とユーザの発話の内容をもとにテキストデータに変換した部分である「ｗｗｗ．ｘｘｘｘｘｘ．ｃｏ．ｊｐ／ｍｕｓｉｃＢＢＢ．ｗａｖ」から構成される場合もある。

【００３４】

音声対話サービス２から応答を受け取った音声対話端末１は、その応答がテキストデータを含むコマンドによる応答の場合は、コマンドの解釈に加えてテキストデータ部分の解釈も行い、制御対象のデバイスの制御を行う。例えばコマンドの内容が「ｐｌａｙ　ｆｒｏｍ　ｗｗｗ．ｘｘｘｘｘｘ．ｃｏ．ｊｐ／ｍｕｓｉｃＢＢＢ．ｗａｖ」の場合、音声対話端末１は、ｗｗｗ．ｘｘｘｘｘｘ．ｃｏ．ｊｐ／ｍｕｓｉｃＢＢＢ．ｗａｖのデータを取得して、この取得したデータを音声対話端末１内で再生してもよい。

【００３５】

このように音声対話サービス２は、ユーザとの対話に基づく情報の提供を行うことができる。

【００３６】

また音声対話サービス２は、音声対話端末１からの音声データの入力がない場合でも、自発的に音声対話端末１に情報を提供してもよい。

【００３７】

音声対話サービス２が自発的に提供する情報は、例えばユーザの近所のバス停へのバスの接近情報であったり、ユーザの居住地域への雨雲の接近情報であったりと、ユーザ個人のニーズに対応した情報であってもよいし、また例えば緊急地震速報や津波警報のように公共性の高い情報であってもよい。

【００３８】

図２は、図１に示した音声対話端末１の構成図である。

【００３９】

音声対話端末１は、ユーザが発した発話を収集するマイク１－１、収集した発話内容をノイズキャンセル等音響処理を行う音響処理部２０１、収集された発話内容からトリガワードを検出するトリガワード検出処理部２０３、トリガワード検出処理部２０３でトリガワードを検出したら、以降のユーザの発話内容を音声サービス２に送信するための処理を行う音声データ出力処理部２０２、出力する音声データを音声対話サービス２とやり取りする通信制御部２０４を含む。さらに音声対話端末１は、トリガワード検出処理部２０３でトリガワードを検出する際の検出感度を設定するトリガワード検出感度設定処理部２０５を含む。

【００４０】

検出感度は、マイク１－１から入力された音声データとトリガワードとの類似度を示すものである。トリガワード検出処理部２０３は、マイク１－１から入力された音声データとトリガワードとの類似度が、予め設定した閾値を超えた場合に、マイク１－１から入力

された音声データは、トリガワードであると判断する。
【００４１】
トリガワード検出処理部２０３に設定されている閾値は、トリガワード検出感度設定処理部２０５が設定する感度設定値により更新される。
【００４２】
音声対話端末１は、一定時間マイク１－１からユーザの発話の入力がない場合、トリガワードが入力されるのを待つトリガワード入力待ち状態に遷移する。音声対話端末１は、トリガワード検出処理部２０３において、マイク１－１から入力された音声が、予め登録されてトリガワードと一致するかの検出を常に行う。音声対話端末１は、トリガワード検出処理部２０３でトリガワードを検出すると、以降ユーザの発話内容をネットワーク３を介して音声サービス２に送信する。音声対話端末１は、トリガワード検出処理部２０３でトリガワードを検出した後も、一定時間マイク１－１からユーザの発話の入力がない場合は、再びトリガワード入力待ち状態に遷移する。
（第１の実施形態）
第１の実施形態の音声対話端末は、ユーザが携帯端末２０あるいは携帯端末２１にインストールされているアプリケーション（感度設定アプリ）を操作して、任意の検出感度の感度設定値を入力することで、トリガワード検出処理部２０３の検出感度を変更する音声対話端末である。
【００４３】
図３は、第１の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れを示す図である。
【００４４】
携帯端末２０を操作することでユーザが設定した感度設定値は、近距離無線通信システム７を介して音声対話端末１に送信することが出来る。また携帯端末２１を操作することでユーザが設定した感度設定値は、サーバ９を経由してネットワーク３を介して音声対話端末１に送信することが出来る。これにより音声対話端末１は、トリガワード検出処理部２０３の検出感度を、携帯端末２０あるいは携帯端末２１のユーザ操作により自由に変更することができる。
【００４５】
図４は、携帯端末２０あるいは携帯端末２１にインストールされている音声対話端末１を制御するための感度設定アプリの、感度設定値を設定する設定画面の例である。
【００４６】
携帯端末２０あるいは携帯端末２１のユーザは、スライドバー（スライダー）４０１を左右に移動させることで検出感度の感度設定値を上下させることができる。ユーザは、スライドバー４０１を任意の位置に設定し、設定ボタン４０２を押下することで、ユーザがスライドバー４０１に設定した位置に基づく感度設定値が、近距離無線通信システム７あるいはネットワーク３を介して音声対話端末１に送られ、トリガワード検出処理部２０３に設定される。
【００４７】
図４の例は、スライドバー４０を左右に移動させることで検出感度の設定値を上下させる場合の例であるが、検出感度の設定値（例えば０～１００）の数値を直接またはアップダウンさせて入力させる方法のユーザインターフェースであってもよい。
【００４８】
図５Ａは、第１の実施形態における音声対話端末１のトリガワードの検出感度の感度設定値を、ユーザが携帯端末２０あるいは携帯端末２１にインストールされているアプリケーションを操作して変更する場合の、音声対話端末１の処理フローである。
【００４９】
音声対話端末１の電源をＯＮすると、音声対話端末１は音声対話端末処理を開始する（Ｓ５００）。音声対話端末１は、起動後初期化処理を行い（Ｓ５０１）、マイク１－１から音声が入力されるのを待つ。

【００５０】
音声対話端末１は、マイク１－１から音声が入力される（Ｓ５０２）と、入力された音声データを音響処理部２０１でノイズの除去等音響処理を行う（Ｓ５０３）。次に音声対話端末１は、感度設定値を受信しているかを判定する（Ｓ５０４）。
【００５１】
Ｓ５０４の判定の結果、感度設定値を受信した場合（Ｓ５０４のＹｅｓ）、音声対話端末１は、トリガワード検出感度設定処理部２０５の検出感度を受信した感度設定値に設定する。音声対話端末１は、トリガワード検出感度設定処理部２０５の検出感度を設定すると、トリガワード検出処理を行う（Ｓ５０６）。
【００５２】
Ｓ５０４の判定の結果、感度設定値を受信しなかった場合（Ｓ５０４のＮｏ）、音声対話端末１は、トリガワード検出処理を行う（Ｓ５０６）。
【００５３】
音声対話端末１は、Ｓ５０２の処理でマイク１－１から入力された音声と、予めトリガワードとして登録されている内容との類似度を算出し、この算出した類似度が予め決められた閾値を超えたかどうかを判定する（Ｓ５０６）。
【００５４】
Ｓ５０６の判定の結果、閾値を超えてる場合（Ｓ５０７のＹｅｓ）、音声対話端末１は、Ｓ５０２でマイク１－１から入力された音声をトリガワードであると認識し、トリガワード以降に入力された音声データを、ネットワーク３を介して音声対話サービス２に送信する（Ｓ５０８）。
【００５５】
Ｓ５０６の判定の結果、閾値を超えていない場合（Ｓ５０７のＮｏ）、音声対話端末１は、Ｓ３０２でマイク１－１から入力された音声からトリガワードを認識できなかったと判断し、トリガワード以降に入力された音声データを、音声対話サービス２に送信しない。
【００５６】
音声対話端末１は、電源がＯＮであれば（Ｓ５０９のＹｅｓ）、Ｓ５０２の処理へ戻りＳ５０３以降の処理を、電源がＯＮである限り繰り返す。
【００５７】
図５Ｂは、第１の実施形態における音声対話端末１のトリガワードの検出感度の感度設定値を、ユーザが携帯端末２０あるいは携帯端末２１にインストールされているアプリケーションを操作して変更する場合の、携帯端末２０あるいは携帯端末２１の処理フローである。なお、携帯端末２０の処理フローと携帯端末２１の処理フローは同一であるので、携帯端末２１の処理フローを例に説明する。
【００５８】
携帯端末２０のユーザが、感度設定値を設定する感度設定アプリを起動するためのアイコンをタップすると、感度設定アプリは携帯端末処理を開始する（Ｓ５２０）。感度設定アプリは、アイコンがタップされたことにより起動する（Ｓ５２１）と、例えば図４に示す表示内容（ＧＵＩ画面）を携帯端末２１の表示画面に表示する。
【００５９】
ユーザは、携帯端末２０の表示画面に表示されている図４に示す表示内容により、感度設定値を調整して設定する（Ｓ５２２）。ユーザは、感度設定値を設定したあとに、図４の設定ボタン４０２を押下することで、設定された感度設定値がネットワーク３を介して音声対話端末１に送られ（Ｓ５２３）、トリガワード検出感度設定処理部２０５によりトリガワード検出処理部２０３の閾値が更新される。
【００６０】
なお第１の実施形態の音声対話端末は、ユーザが携帯端末２０あるいは携帯端末２１にインストールされているアプリケーション（感度設定アプリ）を操作して感度設定値を入力する場合を例に説明したが、例えば音声対話端末１に感度設定アプリがインストールさ

れていてもよい。この場合ユーザは、音声対話端末１にインストールされている感度設定アプリを直接操作することで、感度設定値を入力してもよい。

（第２の実施形態）

第２の実施形態の音声対話端末は、携帯端末２１で設定した期間において収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データを用いて算出した感度設定値を用いて、感度設定値を更新する音声対話端末である。音声対話端末１は、周辺の音声の収集を携帯端末２１からの指示で行う。

【００６１】

図６は、第２の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れを示す図である。

【００６２】

ユーザは、携帯端末２１にインストールされている感度設定アプリを操作することで、音声対話端末１が、周辺の音声をマイク１－１で収集する期間を指示することができる。ユーザは、携帯端末２１の感度設定アプリにより、期間の開始として感度設定開始、期間の終了として感度設定終了を設定する。

【００６３】

感度設定アプリにより設定された感度設定開始および感度設定終了は、感度設定開始イベントおよび感度設定終了イベントとして、ネットワーク３を介してサーバ９を経由して音声対話端末１に送信される（Ｓ６１）。

【００６４】

感度設定開始イベントおよび感度設定終了イベントを受信した音声対話端末１は、受信したイベントの内容に基づいて、設定された期間内にマイク１－１で収集した音声データを検出感度算出部２－４に送る（Ｓ６２）。設定された期間内の音声データの送信処理は、図８Ａを用いて説明する。なお、設定された期間中は、音声対話端末１に対して、ユーザはトリガワードを発しないことが望ましい。

【００６５】

検出感度算出部２－４は、受信した音声データを用いて、トリガワードを発話していないのに、トリガワードを発したと誤検出した回数が一定回数（一定割合）以下になるようにトリガワードの検出感度を算出する。検出感度算出部２－４は、検出感度の算出が完了すると、この算出した感度設定値を音声対話端末１に送信する（Ｓ６３）。感度設定の算出処理は、図８Ｃを用いて説明する。

【００６６】

音声対話端末１は、検出感度算出部２－４から送られてきた感度設定値を、トリガワード検出処理部２０３に設定する。

【００６７】

図７は、携帯端末２０あるいは携帯端末２１にインストールされている音声対話端末１を制御するためのアプリケーションにおける、周辺雑音の測定の開始および終了の設定画面の例である。携帯端末２０あるいは携帯端末２１のユーザは、任意のタイミンツで開始ボタン７０２と終了ボタン７０３を押下することができる。ユーザが開始ボタン７０２を押下すると、測定開始の通知（感度設定開始イベント）がネットワーク３を介して音声対話端末１に送られ、トリガワード検出処理部２０３が周辺雑音の測定を開始する。次にユーザが終了ボタン７０３を押下すると、測定終了の通知（感度設定終了イベント）がネットワーク３を介して音声対話端末１に送られ、トリガワード検出処理部２０３が周辺雑音の測定を終了する。

【００６８】

図７に示す設定アプリの設定画面の例は、感度設定開始および感度設定終了を設定する例であるが、例えば感度設定開始と期間の長さの２つを設定できる設定画面を表示してもよい。また、ユーザが開始ボタン７０２を押下した後に終了ボタン７０３を押下しなくても、所定数秒間（例えば、１０秒間）取得した周辺雑音の音声データを音声対話サービス２に送信し、この送信後に、感度設定終了のイベントを音声対話サービス２に送信する。

このようにすることで終了ボタン７０３を省くことも可能である。

【００６９】

図８Ａは、第２の実施形態における音声対話端末が、携帯端末２１で設定した期間収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データにより算出した感度設定値を用いて、感度設定値を設定する場合の、音声対話端末１の処理フローである。

【００７０】

音声対話端末１の電源をＯＮすると、音声対話端末１は音声対話端末処理を開始する（Ｓ８００）。音声対話端末１は、起動後初期化処理を行い（８０１）、マイク１－１から音声が入力されるのを待つ。

【００７１】

音声対話端末１は、マイク１－１から音声が入力される（Ｓ８０２）と、この入力された音声データを音響処理部２０１でノイズの除去等音響処理を行う（Ｓ８０３）。次に音声対話端末１は、感度設定開始イベントを受信しているかを判定する（Ｓ８０４）。

【００７２】

Ｓ８０４の判定の結果、感度設定開始イベントを受信している場合（Ｓ８０４のＹｅｓ）、音声対話端末１は、マイク１－１から入力された音声を、音声データとして音声対話サービス２に送信する（Ｓ８１２）。続いて音声対話端末１は、感度設定終了イベントを受信したかを判定する（Ｓ８１３）。

【００７３】

Ｓ８１３の判定の結果、感度設定終了イベントを受信していない場合（Ｓ８１３のＮｏ）、音声対話端末１は、Ｓ８０２の処理に戻り、Ｓ８０２以降の処理を続ける。

【００７４】

Ｓ８０４の判定の結果、感度設定開始イベントを受信していない場合（Ｓ８０４のＮｏ）、音声対話端末１は、図５Ａの処理フローのＳ５０４以降と同じ処理を行う。つまり音声対話端末１は、感度設定値を受信したかの判定（Ｓ８０５）と、トリガワードを検出したかの判定（８０８）を、電源がＯＮされている間繰り返し行う。

【００７５】

またＳ８１３の判定の結果、感度設定終了イベントを受信していた場合（Ｓ８１３のＹｅｓ）、音声対話端末１は、図５Ａの処理フローのＳ５０４以降と同じ処理を行う。つまり、音声対話端末１は、感度設定値を受信したかの判定（Ｓ８０５）と、トリガワードを検出したかの判定（８０８）を、電源がＯＮされている間繰り返し行う。

【００７６】

図８Ｂは、第２の実施形態における音声対話端末が、携帯端末２１で設定した期間収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データにより算出した感度設定値を用いて、感度設定値を更新する場合の、携帯端末２１の処理フローである。

【００７７】

携帯端末２１のユーザが感度設定アプリを起動するためのアイコンをタップすると、感度設定アプリは携帯端末処理を開始する（Ｓ８２０）。感度設定アプリは、アイコンがタップされたことにより起動する（Ｓ８２１）と、例えば図７に示す表示内容を携帯端末２１の表示画面に表示する。

【００７８】

ユーザは、携帯端末２１の表示画面に表示さている図７に示す表示内容の開始ボタン７０２を押下すると、感度設定開始イベントが、ネットワーク３を介して音声対話端末１に送信される。感度設定開始イベントを受信した音声対話端末１は、図８ＡのＳ８０４の処理においてＹｅｓと判定し、取得した周辺雑音の音声データを音声対話サービス２に送信を開始する。

【００７９】

つづいてユーザが、携帯端末２０の表示画面に表示されている図７に示す表示内容の終

了ボタン７０３を押下すると、感度設定終了イベントが、ネットワーク３を介して音声対話端末１に送られる。感度設定終了イベントを受信した音声対話端末１は、図８ＡのＳ８１３の処理においてをＹｅｓと判定し、周辺雑音の音声データの音声対話サービス２への送信を終了する。

【００８０】

図８Ｃは、第２の実施形態における音声対話端末が、携帯端末２１で設定した期間収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データにより算出した感度設定値を用いて、感度設定値を更新する場合の、検出感度算出部２－４の処理フローである。

【００８１】

音声対話サービス２の検出感度算出部２－４は、外部からイベントを受信するとクラウド処理を開始する（Ｓ８３０）。検出感度算出部２－４は、受信したイベントが感度設定開始かどうかを判定する（Ｓ８３１）。

【００８２】

Ｓ８３１の判定の結果、受信したイベントが感度設定開始イベントである場合（Ｓ８３１がＹｅｓ）、検出感度算出部２－４は、検出感度の調整を開始する（Ｓ８３２）。また検出感度算出部２－４は同時に、受信した感度設定開始イベントを音声対話端末１に送信する（Ｓ８３２）。

【００８３】

検出感度算出部２－４は、感度設定開始イベントを音声対話端末１に送ることで、音声対話端末１から音声データが送られてくる（図８Ａの８１２）のを待つ（Ｓ８３３）。検出感度算出部２－４は、音声対話端末１から音声データが送られてくる（Ｓ８３３）と、感度設定終了イベントを受信するまで、送られてくる音声データを受信し続ける（Ｓ８３４のＮｏ）。

【００８４】

検出感度算出部２－４は、感度設定終了イベントを受信すると（Ｓ８３４のＹｅｓ）検出感度調整を終了する（Ｓ８３５）。また検出感度算出部２－４は同時に、受信した感度設定終了イベントを音声対話端末１に送信する（Ｓ８３５）。

【００８５】

次に検出感度算出部２－４は、受信した音声データを用いて感度設定値を算出する（Ｓ８３６）。検出感度算出部２－４は、算出した感度設定値を音声対話端末１に送信し（Ｓ８３７）、処理を終了する。

【００８６】

以上のように第２の実施形態の音声対話端末は、音声対話端末が置かれている周辺の音の状況から算出された最適な感度設定値を用いて、トリガワードを検出する感度設定値を設定するので、音声対話端末１は、置かれている環境に適した感度でトリガワードを検出でき、ユーザが使い勝手が向上する。

（第３の実施形態）

第３の実施形態の音声対話端末は、ユーザの発話の内容により設定した期間において収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データを用いて算出した感度設定値を用いて、感度設定値を設定する音声対話端末である。音声対話端末１は、周辺の音声の収集をユーザの発話による指示で行う。

【００８７】

図９は、第３の実施形態の音声対話端末に感度設定を行う場合の、全体構成図および感度設定に関するデータの流れを示す図である。

【００８８】

ユーザは、音声対話端末１に発話することで、周辺の音声をマイク１－１で収集する期間を指示することができる。発話の内容は例えば、期間の開始を示す感度設定開始としての「トリガワード検出感度開始」、期間の終了を示す感度設定終了としての「トリガワー

ド検出感度終了」でもよい。ユーザが発話した「トリガワード検出感度開始」および「トリガワード検出感度終了」は、音声データとして検出感度算出部２－４に送られ（Ｓ９１）、検出感度算出部２－４において期間の開始と期間の終了として認識される。検出感度算出部２－４は、認識した期間の開始を感度設定開始イベントとして音声対話端末１に送信する（Ｓ９１）。また検出感度算出部２－４は、認識した期間の終了を感度設定終了イベントとして音声対話端末１に送信する（Ｓ９１）。

【００８９】

また音声対話端末１は、マイク１－１から入力された音声を、音声データとして検出感度算出部２－４に送信する（Ｓ９２）。

【００９０】

音声対話端末１から送られてきた期間の開始を示す音声データである「トリガワード検出感度開始」、期間の終了を示す音声データである「トリガワード検出感度終了」を認識した検出感度算出部２－４は、この期間に受信した音声データを用いて、トリガワードを発話していないのに、トリガワードを発したとご検出した回数が一定回数以下になるようにトリガワードの検出感度を算出する。検出感度算出部２－４は、検出感度の算出が完了すると、算出した感度設定値を音声対話端末１に送信する（Ｓ９３）。

【００９１】

図１０Ａは、第３の実施形態における音声対話端末が、ユーザの発話の内容により設定した期間において収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データを用いて算出した感度設定値を用いて、感度設定値を更新する場合の、音声対話端末１の処理フローである。図１０ＡのＳ１０００からＳ１０１３までの処理は、図８ＡのＳ８００からＳ８１３までの処理と同一である。

【００９２】

図１０Ｂは、第３の実施形態における音声対話端末が、ユーザの発話の内容により設定した期間において収集した周辺の音声を音声対話サービス２に送信し、音声対話サービス２において音声対話端末１が収集した音声データを用いて算出した感度設定値を用いて、感度設定値を更新する場合の、検出感度算出部２－４の処理フローである。

【００９３】

図８Ｃの処理フローとの違いは、期間の開始を示す感度設定開始イベントを受信する代わりに期間の開始を示す音声データである「トリガワード検出感度開始」を受信して認識する点および、期間の終了を示す感度設定終了イベントを受信する代わりに期間の終了を示す音声データである「トリガワード検出感度終了」を受信して認識する点である。

【００９４】

音声対話端末１は、トリガワードを認識し、マイク１－１から入力された音声を音声データとして音声対話サービス２に送信しているとする。音声対話サービス２は、音声対話端末１から送られてきた音声データを受信すると、クラウド処理（１０２０）を開始する。

【００９５】

検出感度算出部２－４は、受信した音声データの中から、期間の開始を示す音声データである「トリガワード検出感度開始」を認識すると（Ｓ１０２３のＹｅｓ）、検出感度算出部２－４は、検出感度の調整を開始する（Ｓ１０２４）。また検出感度算出部２－４は、同時に受信した感度設定開始のイベントを音声対話端末１に送信する（Ｓ１０２４）。

【００９６】

検出感度算出部２－４は、受信した音声データの中から、期間の終了を示す音声データである「トリガワード検出感度終了」を認識するまで音声データを受信し続ける（Ｓ１０２５）。

【００９７】

検出感度算出部２－４は、受信した音声データの中から、期間の終了を示す音声データである「トリガワード検出感度終了」を認識する（Ｓ１０２６のＹｅｓ）と検出感度調整

を終了する（Ｓ１０２７）。検出感度算出部２－４は、検出感度調整を終了する（Ｓ１０２７）と、受信した音声データを用いて感度設定値を算出する（Ｓ１０２８）。検出感度算出部２－４は、算出した感度設定値を音声対話端末１に送信し（Ｓ１０２９）、処理を終了する（１０３０）。

【００９８】

以上のように第３の実施形態の音声対話端末は、音声対話端末が置かれている周辺の音の状況から算出された最適な感度設定値を用いて、トリガワードを検出する感度設定値を更新することができ、かつ最適な感度設定値を算出を、ユーザの発話をトリガーにして実行することができるため、ユーザは携帯端末を操作する手間から解放され、さらにユーザの使い勝手が向上する。このよう音声対話端末１は、置かれている環境に適した感度でトリガワードを検出でき、ユーザが使い勝手が向上する。

【００９９】

以上のように、本実施形態の音声対話端末は、音声対話端末が設定されている状況の応じて、トリガワードを検出する感度の設定を更新することが可能である。例えば。比較的騒がしい環境においては、第２の実施形態の機能を使い周辺環境に基づいた検出感度の設定を行うのに加え、第１の実施形態のユーザ操作による検出感度を設定を行ってもよい。

【０１００】

また第１の実施形態の機能および第２の実施形態の機能の両方を同時に使用してもよい。たとえば、ある位置に置かれた音声対話端末１に対して他の位置にいるユーザがトリガワードを発する場合の、検出感度を調整するシーンを想定する。この場合、第１の実施形態の機能だけで検出感度を決定すると誤検出が増えることが予想される。そこで、第２の実施形態の機能を合わせて使用することで、音声対話端末１とユーザの立ち位置の相対関係において、周辺環境を考慮にいれた最適な検出感度を設定することが可能である。

（第４の実施形態）

第１の実施形態から第３の実施形態において音声対話サービス２の検出感度算出部２－４は、サーバ９に存在しているが、音声対話端末１の中に存在していてもよい。

【０１０１】

第４の実施形態の音声対話端末は、さらに検出感度算出部２－４を含む音声対話端末である。

【０１０２】

図１１および図１２に、検出感度算出部２－４が音声対話端末１の中に存在している場合の全体構成図および感度設定に関するデータの流れの例を示す図である。

【０１０３】

図１１は、第２の実施形態の音声対話端末１の場合の図６に対応する、検出感度算出部２－４が音声対話端末１の中に存在している場合の、全体構成図および感度設定に関するデータの流れを示す図である。

【０１０４】

感度設定アプリにより設定された感度設定開始および感度設定終了は、感度設定開始イベントおよび感度設定終了イベントとして、ネットワーク３を介して音声対話端末１に送信される（Ｓ１１１）。音声対話端末１は、この感度設定開始イベントと感度設定終了イベントとの間、図６に示す例の場合同様に、周辺の音声をマイク１－１で収集するが、図１１に示す例の場合は、図６に示すＳ６２およびＳ６３に相当する処理をすべて音声対話端末１の内部を行う点が、図６に示す例との相違点である。

【０１０５】

同様に図１２は、第３の実施形態の音声対話端末１の場合の図９に対応する、検出感度算出部２－４が音声対話端末１の中に存在している場合の全体構成図および感度設定に関するデータの流れを示す図である。

【０１０６】

ユーザは、図９に示す例の場合同様に、音声対話端末１に発話することで、周辺の音声をマイク１－１で収集する期間を指示することができる。音声対話端末１は、この指示さ

れた期間、図６に示す例の場合同様に、周辺の音声をマイク１－１で収集するが、図１２に示す例の場合は、図９に示すＳ９１、Ｓ９２およびＳ９３の処理に相当する処理をすべて音声対話端末１の内部を行う点が、図９に示す例との相違点である。

【０１０７】

本発明のいくつかの実施形態を説明したが、これらの実施形態は例として提示したものであり、発明の範囲を限定することは意図していない。これら新規な実施形態は、その他の様々な形態で実施されることが可能であり、発明の要旨を逸脱しない範囲で、種々の省略、置き換え、変更を行うことができる。これら実施形態やその変形は、発明の範囲や要旨に含まれるとともに、特許請求の範囲に記載された発明とその均等の範囲に含まれる。さらにまた、請求項の各構成要素において、構成要素を分割して表現した場合、或いは複数を合わせて表現した場合、或いはこれらを組み合わせて表現した場合であっても本発明の範疇である。また、複数の実施形態を組み合わせてもよく、この組み合わせで構成される実施例も発明の範疇である。

【０１０８】

また、本明細書と各図において、既出の図に関して前述したものと同一又は類似した機能を発揮する構成要素には同一の参照符号を付し、重複する詳細な説明を適宜省略することがある。また請求項を制御ロジックとして表現した場合、コンピュータを実行させるインストラクションを含むプログラムとして表現した場合、及び前記インストラクションを記載したコンピュータ読み取り可能な記録媒体として表現した場合でも本発明の装置を適用したものである。また、使用している名称や用語についても限定されるものではなく、他の表現であっても実質的に同一内容、同趣旨であれば、本発明に含まれるものである。

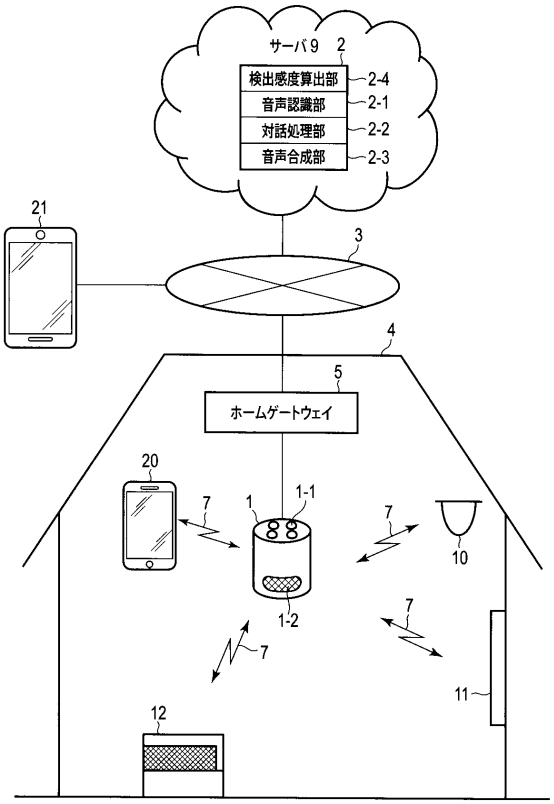
【符号の説明】

【０１０９】

１・・・音声対話端末、２・・・音声対話サービス、２－１・・・音声認識部、２－２・・・対話処理部、２－３・・・音声合成部、２－４検出感度算出部、３・・・ネットワーク、２０３・・・トリガワード検出処理部、２０５・・・トリガワード検出感度設定処理部。

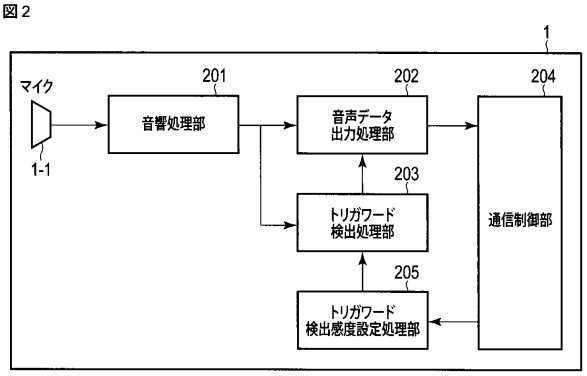
【図１】

図1



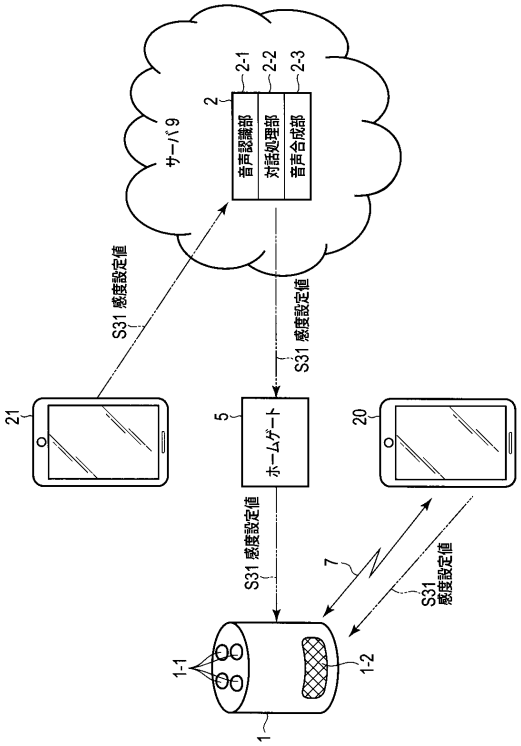
【図２】

図2



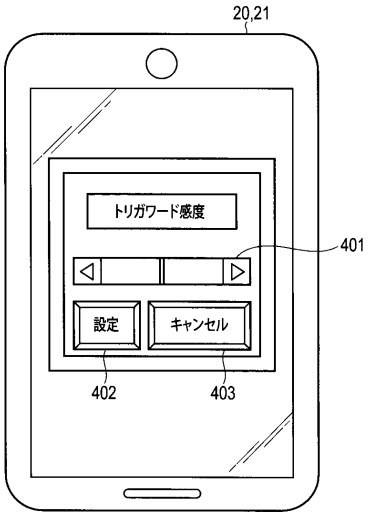
【図３】

図3



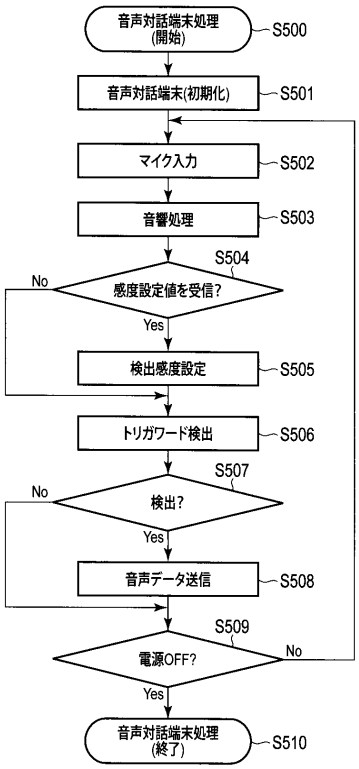
【図４】

図4



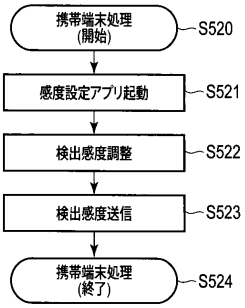
【図５Ａ】

図5A



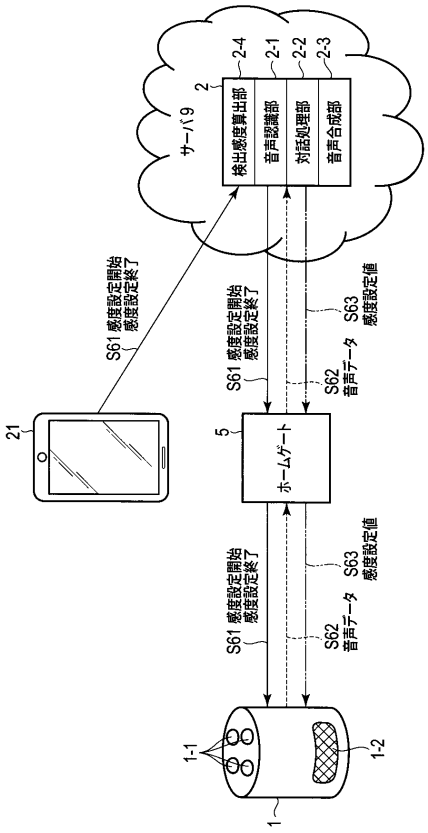
【図５Ｂ】

図5B



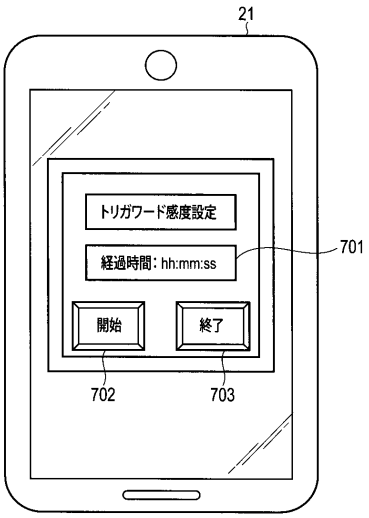
【図６】

図6



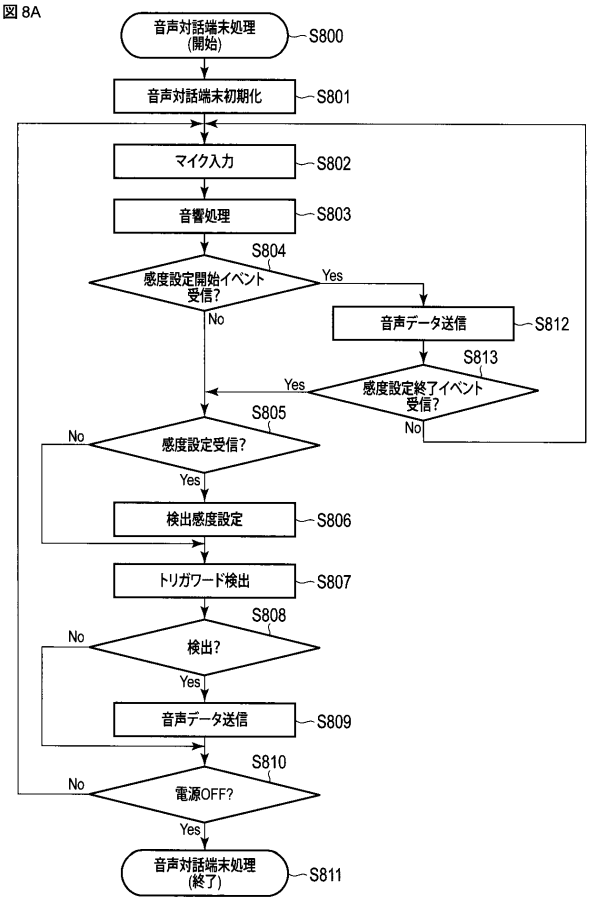
【図７】

図7



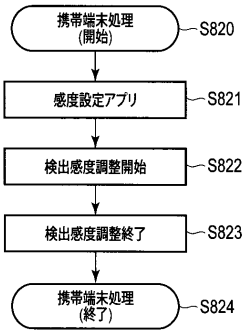
【図８Ａ】

図 8A



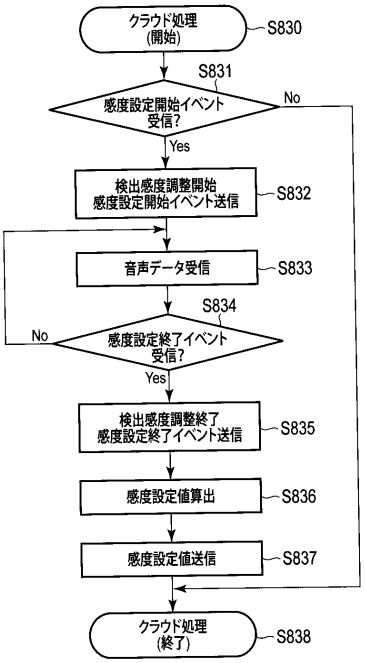
【図８Ｂ】

図 8B



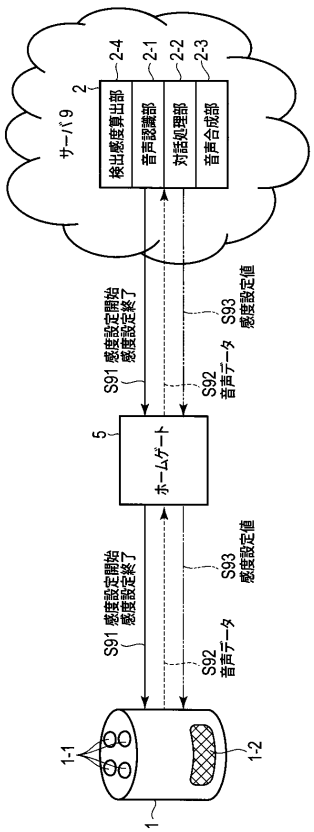
【図８Ｃ】

図 8C



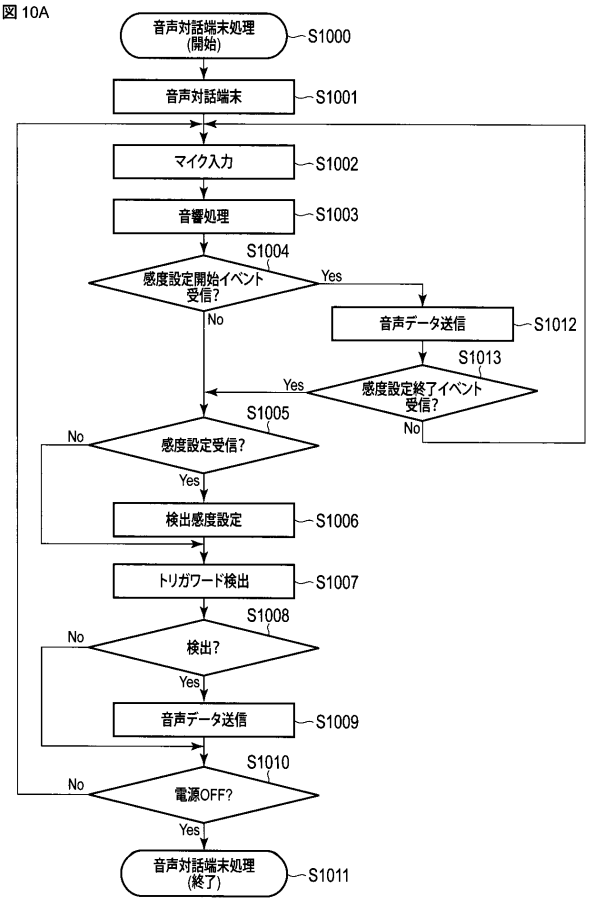
【図９】

図 9



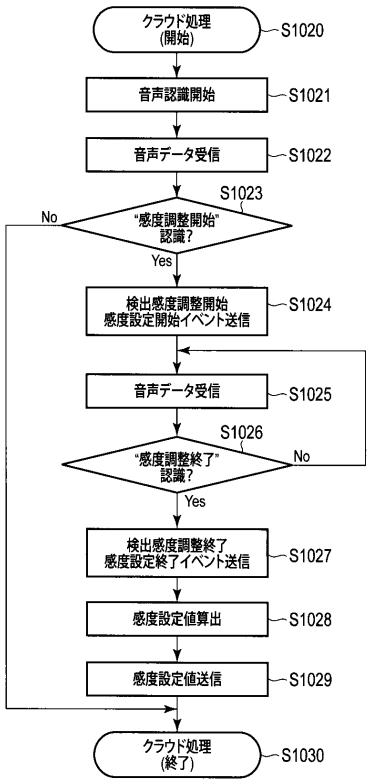
【図１０Ａ】

図 10A



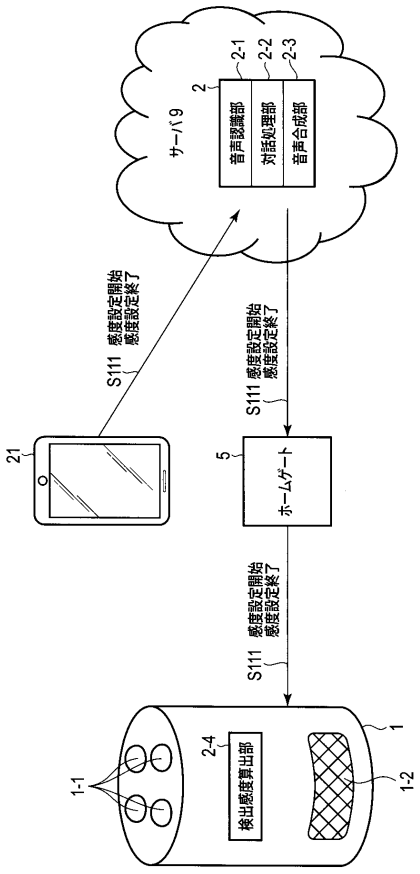
【図１０Ｂ】

図 10B



【図１１】

図 11



【図１２】

図 12

