

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号
特許第6318621号
(P6318621)

(45) 発行日 平成30年5月9日 (2018.5.9)

(24) 登録日 平成30年4月13日 (2018.4.13)

(51) Int.Cl.	F I
HO 4 M 1/00 (2006.01)	HO 4 M 1/00 V
HO 4 M 3/42 (2006.01)	HO 4 M 3/42 P
HO 4 M 11/00 (2006.01)	HO 4 M 11/00 3 O 2
HO 4 M 1/60 (2006.01)	HO 4 M 1/60 A
G 1 O L 15/00 (2013.01)	G 1 O L 15/00 2 O O A

請求項の数 10 (全 15 頁)

(21) 出願番号	特願2014-285 (P2014-285)	(73) 特許権者	000004260
(22) 出願日	平成26年1月6日 (2014.1.6)		株式会社デンソー
(65) 公開番号	特開2015-130554 (P2015-130554A)		愛知県刈谷市昭和町 1 丁目 1 番地
(43) 公開日	平成27年7月16日 (2015.7.16)	(74) 代理人	110000567
審査請求日	平成28年9月14日 (2016.9.14)		特許業務法人 サトー国際特許事務所
		(72) 発明者	伊藤 正也
			愛知県刈谷市昭和町 1 丁目 1 番地 株式会
			社デンソー内
		(72) 発明者	尾崎 義隆
			愛知県刈谷市昭和町 1 丁目 1 番地 株式会
			社デンソー内
		(72) 発明者	林 圭作
			愛知県刈谷市昭和町 1 丁目 1 番地 株式会
			社デンソー内

最終頁に続く

(54) 【発明の名称】 音声処理装置、音声処理システム、音声処理方法、音声処理プログラム

(57) 【特許請求の範囲】

【請求項 1】

音声データを取得する音声データ取得手段（31）と、
前記音声データ取得手段が取得した音声データを外部の携帯端末（12）に送信する音声データ送信手段（32）と、
前記音声データ送信手段が送信する音声データにノイズキャンセル処理を含む所定の音声処理を施す音声処理手段（33）と、を備え、
前記音声処理手段は、前記音声処理として、通話用の音声処理と通話以外用の音声処理を切り替えて実行可能に構成され、
前記音声データ送信手段は、通話用の前記音声データ及び通話以外用の前記音声データを同一の通信プロトコルによって送信することを特徴とする音声処理装置。

10

【請求項 2】

前記音声処理手段は、通話アプリケーションにおける自発的な操作または他発的な操作を検知した場合に、前記通話用の音声処理を実行する請求項 1 に記載の音声処理装置。

【請求項 3】

前記音声処理手段は、通話アプリケーション以外のアプリケーションが起動された場合に、前記通話以外用の音声処理を実行する請求項 1 または 2 に記載の音声処理装置。

【請求項 4】

前記音声処理手段は、通話アプリケーション以外のアプリケーションである音声認識アプリケーションが起動された場合に、前記通話以外用の音声処理である音声認識用の音声

20

処理を実行する請求項 1 から 3 の何れか 1 項に記載の音声処理装置。

【請求項 5】

前記音声処理手段は、通話用の音声処理よりも音声波形が残る通話以外用の音声処理を実行可能に構成され、通話アプリケーション以外のアプリケーションが起動された場合に、その通話以外用の音声処理を実行する請求項 1 から 4 の何れか 1 項に記載の音声処理装置。

【請求項 6】

前記音声処理手段は、通話アプリケーション以外のアプリケーションが起動された場合に、音声処理を実行しないように構成されている請求項 1 に記載の音声処理装置。

【請求項 7】

前記音声データ送信手段は、前記通信プロトコルとして、Bluetooth 通信規格 (Bluetooth:登録商標) のハンズフリー通話用のプロファイルを用いる請求項 1 から 6 の何れか 1 項に記載の音声処理装置。

【請求項 8】

音声処理装置 (11) と、

前記音声処理装置と通信可能に接続される携帯端末 (12) と、
で構築される音声処理システム (10) であって、

前記音声処理装置は、

音声データを取得する音声データ取得手段 (31) と、

前記音声データ取得手段が取得した音声データを外部の携帯端末に送信する音声データ送信手段 (32) と、

前記音声データ送信手段が送信する音声データにノイズキャンセル処理を含む所定の音声処理を施す音声処理手段 (33) と、を備え、

前記音声処理手段は、前記音声処理として、通話用の音声処理と通話以外用の音声処理を切り替えて実行可能に構成され、

前記音声データ送信手段は、通話用の前記音声データ及び通話以外用の前記音声データを同一の通信プロトコルによって送信することを特徴とする音声処理システム。

【請求項 9】

音声データを取得する音声データ取得ステップと、

前記音声データ取得ステップによって取得された音声データを外部の携帯端末に送信する音声データ送信ステップと、

前記音声データ送信ステップによって送信される音声データにノイズキャンセル処理を含む所定の音声処理を施す音声処理ステップと、を含み、

前記音声処理ステップでは、前記音声処理として、通話用の音声処理と通話以外用の音声処理を切り替えて実行し、

前記音声データ送信ステップは、通話用の前記音声データ及び通話以外用の前記音声データを同一の通信プロトコルによって送信することを特徴とする音声処理方法。

【請求項 10】

音声処理装置に組み込まれて実行される音声処理用のプログラムであって、

音声データを取得する音声データ取得ステップと、

前記音声データ取得ステップによって取得された音声データを外部の携帯端末に送信する音声データ送信ステップと、

前記音声データ送信ステップによって送信される音声データにノイズキャンセル処理を含む所定の音声処理を施す音声処理ステップと、を実行可能であり、

前記音声処理ステップでは、前記音声処理として、通話用の音声処理と通話以外用の音声処理を切り替えて実行し、

前記音声データ送信ステップは、通話用の前記音声データ及び通話以外用の前記音声データを同一の通信プロトコルによって送信することを特徴とする音声処理プログラム。

【発明の詳細な説明】

【技術分野】

10

20

30

40

50

【 0 0 0 1 】

本発明は、音声処理装置、音声処理システム、音声処理方法、音声処理プログラムに関する。

【 背景技術 】

【 0 0 0 2 】

近年、車両に搭載される車両用機器と携帯端末とを通信可能に接続し、携帯端末を手に持たなくとも通話を可能とするいわゆるハンズフリー通話を実現する技術が普及しつつある（例えば特許文献 1 参照）。この種のハンズフリー通話技術においては、多くの車両用機器で採用されている Bluetooth（登録商標）の HFP（HFP : Hands Free Profile）を通信プロトコルとして使用しており、車両用機器は、携帯端末に送信する音声データに、当該データを最適化するための音声処理を施している。

10

【 先行技術文献 】

【 特許文献 】

【 0 0 0 3 】

【 特許文献 1 】 特開 2 0 0 6 - 2 3 8 1 4 8 号公報

【 発明の概要 】

【 発明が解決しようとする課題 】

【 0 0 0 4 】

ところで、近年では、車両用機器と携帯端末とを相互に連携させながらアプリケーションを実行する技術の開発が進められており、この技術においては、ハンズフリー通話を可能とするいわゆる通話アプリケーションに限らず、例えば音声認識を利用した検索アプリケーションといった通話以外のアプリケーションも実行可能である。

20

【 0 0 0 5 】

この検索アプリケーションでは、車両用機器は、取得した音声データを、携帯端末を介して外部のセンターサーバに送信する。そして、センターサーバは、取得した音声データに基づき音声認識を実施し、その音声に対応する検索結果を車両用機器に返信する。ところが、従来では、車両用機器は、ハンズフリー通話の実行時において音声データを携帯端末に送信する場合、及び、音声認識を利用した検索の実行時において音声データを携帯端末に送信する場合の何れの場合においても、音声データに同一の音声処理、具体的には同一のノイズキャンセル処理、エコーキャンセル処理、ゲインコントロール処理等の音声処理を施している。一方で、通話に最適な音声処理と音声認識に最適な音声処理は、それぞれ異なる。即ち、例えば、ハンズフリー通話では、例えば人間の耳で聞こえる周波数の音に絞る音声処理を実施しているが、この音声処理と同様の処理を音声認識において実施すると、音声認識に必要な音声波形が歪んでしまい、認識率が下がってしまう。

30

【 0 0 0 6 】

本発明は上記した事情に鑑みてなされたものであり、その目的は、通話用の音声処理及び通話以外用の音声処理を何れも最適に実施することができる音声処理装置、この音声処理装置を含んで構築される音声処理システム、この音声処理装置において実行される音声処理方法、及び、この音声処理装置に組み込まれて実行される音声処理プログラムを提供することにある。

40

【 課題を解決するための手段 】

【 0 0 0 7 】

本発明によれば、取得した音声データを外部の携帯端末に送信する場合に、その送信する音声データに所定の音声処理が施される。そして、本発明によれば、その音声処理として、通話用の音声処理と通話以外用の音声処理とを切り替えて実行することが可能である。よって、実行されるアプリケーションに応じて通話用の音声処理及び通話以外用の音声処理を適宜切り替えて実行することができ、通話用の音声処理及び通話以外用の音声処理を何れも最適に実施することができる。

さらに、本発明によれば、通話用の音声データ及び通話以外用の音声データを同一の通信プロトコルによって送信する。

50

【図面の簡単な説明】

【 0 0 0 8 】

【図 1】一実施形態に係る音声処理システムの構成例を概略的に示す図

【図 2】音声処理装置の構成例を概略的に示す図

【図 3】携帯端末の構成例を概略的に示す図

【図 4】通話アプリケーションを実行する場合の制御内容の一例を示すフローチャート

【図 5】音声処理装置及び携帯端末が相互に連携してアプリケーションを実行する状態を概略的に示す図

10

【図 6】音声認識検索アプリケーションを実行する場合の制御内容の一例を示すフローチャート

【図 7】本実施形態の変形例を示す音声処理システムの概略構成図（その 1）

【図 8】本実施形態の変形例を示す音声処理システムの概略構成図（その 2）

【図 9】本実施形態の変形例を示す音声処理システムの概略構成図（その 3）

【図 10】本実施形態の変形例を示す音声処理システムの概略構成図（その 4）

【発明を実施するための形態】

【 0 0 0 9 】

以下、本発明の一実施形態について図面を参照しながら説明する。図 1 に示すように、音声処理システム 10 は、音声処理装置 11 と携帯端末 12 とで構築される。音声処理装置 11 は、例えば車両に搭載されるナビゲーション装置で構成される。この場合、音声処理装置 11 には、通話アプリケーション A が搭載されている。この通話アプリケーション A は、使用者が携帯端末 12 を手に持たなくとも通話を可能とするいわゆるハンズフリー通話機能を実現するためのアプリケーションである。また、携帯端末 12 は、例えば車両の搭乗者が所有する携帯通信端末であり、車室内に持ち込まれると、近距離無線通信規格の一例である Bluetooth 通信規格（Bluetooth：登録商標）により音声処理装置 11 に通信可能に接続される。

20

【 0 0 1 0 】

音声処理装置 11 及び携帯端末 12 は、通信網 100 を介して外部の配信センター 14 に接続することで、当該配信センター 14 から配信される各種のアプリケーションを取得するように構成されている。配信センター 14 は、上述の通話アプリケーション A のほか、例えば音声認識を利用した検索サービスを実現する音声認識検索アプリケーション B、さらには、インターネットラジオを実現するアプリケーション、音楽配信サービスを実現するアプリケーションなど各種のアプリケーションを格納しており、外部の端末や機器からアプリケーションの配信要求を受けると、該当するアプリケーションを通信網 100 を介して要求元に配信する。なお、配信センター 14 から配信されるアプリケーションには、そのアプリケーションを実行するために必要な各種のデータなどが含まれる。

30

【 0 0 1 1 】

また、音声処理装置 11 及び携帯端末 12 は、通信網 100 を介して音声認識検索サーバ 15（以下、音認検索サーバ 15 と称する）にも接続可能に構成されている。この音認検索サーバ 15 には、音声認識処理に必要な周知の辞書データ、検索処理に必要な検索処理用データなどが格納されている。検索処理用データには、地図データのほか、地図上に存在する店舗や施設などの名称や場所といったデータが含まれている。

40

【 0 0 1 2 】

次に、音声処理装置 11 の構成について図 2 を参照しながら説明する。即ち、音声処理装置 11 は、制御部 21、通信接続部 22、記憶部 23、音声入出力部 24、表示出力部 25、操作入力部 26 などを備えている。制御部 21 は、図示しない CPU、RAM、ROM 及び I/O バスなどを有する周知のマイクロコンピュータで構成されている。制御部 21 は、ROM あるいは記憶部 23 などに記憶されている各種のコンピュータプログラムに従って音声処理装置 11 の動作全般を制御する。また、制御部 21 は、コンピュータ

50

プログラムである音声処理プログラムを実行することにより、音声データ取得処理部 3 1、音声データ送信処理部 3 2、音声処理部 3 3 をソフトウェアによって仮想的に実現する。

【 0 0 1 3 】

通信接続部 2 2 は、例えば無線通信モジュールなどで構成され、携帯端末 1 2 が備える通信接続部 4 2 との間に無線通信回線を確立し、その無線通信回線を通じて携帯端末 1 2 との間で各種の通信を行う。この場合、通信接続部 2 2 は、ハンズフリー通話用のプロファイル (H F P : Hands Free Profile) やデータ通信用のプロファイルなど種々の通信プロトコルを備える。記憶部 2 3 は、例えばハードディスクドライブなどの不揮発性の記憶媒体で構成されており、各種のコンピュータプログラムやアプリケーションプログラム、外部の装置や端末と相互に連携してアプリケーションを実行する連携機能を実現する連携アプリケーションなどの各種のプログラム、及び、各プログラムで使用される各種のデータなどを記憶している。また、記憶部 2 3 は、取得した音声データを音声認識するための周知の辞書データなど音声認識処理に必要な各種のデータを格納している。よって、音声処理装置 1 1 は、音認検索サーバ 1 5 に依らずとも、単独でも音声認識処理が可能となっている。

10

【 0 0 1 4 】

音声入出力部 2 4 は、図示しないマイクロホン及びスピーカに接続しており、周知の音声入力機能及び音声出力機能を備える。この音声入出力部 2 4 は、音声処理装置 1 1 に携帯端末 1 2 が通信可能に接続された状態で通話アプリケーション A が起動されると、マイクロホンから入力された音声に対応する音声データを携帯端末 1 2 に送信し、また、携帯端末 1 2 から受信した音声データに基づき音声をスピーカから出力することが可能となる。これにより、音声処理装置 1 1 は、携帯端末 1 2 と協働して、いわゆるハンズフリー通話を実現可能となる。

20

【 0 0 1 5 】

表示出力部 2 5 は、例えば液晶表示器や有機 E L で構成されており、制御部 2 1 からの表示指令信号に基づいて各種の情報を表示する。この表示出力部 2 5 の画面には、周知の感圧方式、電磁誘導方式、静電容量方式あるいはそれらを組み合わせた方式で構成されるタッチパネルスイッチが設けられる。この表示出力部 2 5 には、アプリケーションに対する操作を入力するための操作入力画面などの入力インターフェース、アプリケーションの実行内容や実行結果を出力するための出力画面などの出力インターフェースなどの各種画面が表示される。

30

【 0 0 1 6 】

操作入力部 2 6 は、表示出力部 2 5 の画面上に設けられるタッチパネルスイッチ、及び、表示出力部 2 5 の周囲に設けられているメカニカルスイッチなど各種のスイッチ類を含む。操作入力部 2 6 は、使用者による各種のスイッチの操作に応じて、その操作検知信号を制御部 2 1 に出力する。制御部 2 1 は、操作入力部 2 6 から入力された操作検知信号を解析して使用者の操作内容を特定し、特定した操作内容に基づいて各種の処理を実行する。なお、図示はしないが、音声処理装置 1 1 は、図示しない測位用衛星から受信する衛星電波などに基づいて音声処理装置 1 1 の現在位置を特定するための周知の位置特定部を備えている。

40

【 0 0 1 7 】

音声データ取得処理部 3 1 は、音声データ取得手段の一例であり、音声入出力部 2 4 のマイクロホンから音声が入力されると、その取得した音声に対応する音声データを生成する。

音声データ送信処理部 3 2 は、音声データ送信手段の一例であり、音声データ取得処理部 3 1 によって取得された音声データを、通信接続部 2 2 が確立した通信回線を介して外部の携帯端末 1 2 に送信する。この場合、音声データ送信処理部 3 2 は、通話用の音声データ及び通話以外用の音声データを、何れも同一の通信プロトコルによって送信するように構成されている。なお、本実施形態では、その同一の通信プロトコルとして、B l u e t o o t h 通信規格のハンズフリー通話用のプロファイル (H F P) が採用されている。

50

しかし、採用可能な通信プロトコルは、これに限られるものではない。

【 0 0 1 8 】

音声処理部 3 3 は、音声処理手段の一例であり、音声データ送信処理部 3 2 によって送信される音声データに所定の音声処理を施す。詳しくは後述するが、この音声処理部 3 3 は、音声処理として、通話用の音声処理と通話以外用の音声処理の一例である音認検索用の音声処理を切り替えて実行可能に構成されている。なお、通話用の音声処理は、例えば人間の耳で聞こえる周波数の音のみに絞るための処理であり、通話用のノイズキャンセル処理、通話用のエコーキャンセル処理、通話用のゲインコントロール処理などを含む。この通話用の音声処理によれば、人間の耳で聞こえる周波数以外の音は完全にあるいは殆どキャンセルされる。一方、音認検索用の音声処理は、例えば人間の耳で聞こえる周波数の音を含んで音声認識が可能となる程度に音を絞るための処理であり、音認検索用のノイズキャンセル処理、音認検索用のエコーキャンセル処理、音認検索用のゲインコントロール処理などを含む。この音認検索用の音声処理によれば、人間の耳で聞こえる周波数以外の音もある程度はキャンセルされずに残る。

10

【 0 0 1 9 】

基本的には、音認検索用の音声処理よりも通話用の音声処理の方が、音声データに対し確実なノイズキャンセル、エコーキャンセル、ゲインコントロールがかけられる。一方、音認検索用の音声処理では、使用者が発した音声に極力近い生の音声を拾いたいことから、音声データに対し比較的緩いノイズキャンセル、エコーキャンセル、ゲインコントロールがかけられる。即ち、音認検索用の音声処理では、本来の音声情報（音声波形）が変化

20

【 0 0 2 0 】

例えば通話用の音声処理におけるゲインコントロールでは、音声データに含まれる各周波数帯域に対して、人の耳には聞こえにくい高周波数帯域と低周波数帯のゲインを落とし、人の耳に聞こえやすい中周波数帯域を増幅させる処理などが行われる。しかし、このような音声処理を音認検索用の音声データに施すと、本来の音声波形が歪んでしまうため、音声認識には向かない。因みに、音声波形は、母音や子音ごとに波形（周波数）が異なるため、本来の音声波形が崩れてしまうと、音声の認識が極めて困難となる。従って、音声認識用の音声処理におけるゲインコントロールとしては、例えば、ゲインを落とす高周波数帯域や低周波数帯域の設定値（パラメタ）を変更する、ゲインの落とし方を適宜調整する

30

【 0 0 2 1 】

次に、携帯端末 1 2 の構成について図 3 を参照しながら説明する。携帯端末 1 2 は、制御部 4 1、通信接続部 4 2、記憶部 4 3、音声入出力部 4 4、表示出力部 4 5、操作入力部 4 6、電話通信部 4 7などを備えている。制御部 4 1は、図示しないCPU、RAM、ROM及びI/Oバスなどを有する周知のマイクロコンピュータで構成されている。制御部 4 1は、ROMあるいは記憶部 4 3などに記憶されているコンピュータプログラムに従って携帯端末 1 2 の動作全般を制御する。

【 0 0 2 2 】

通信接続部 4 2 は、例えば無線通信モジュールなどで構成され、音声処理装置 1 1 が備える通信接続部 2 2 との間に無線通信回線を確立し、その無線通信回線を通じて音声処理装置 1 1 との間で各種の通信を行う。この場合、通信接続部 4 2 は、ハンズフリー通話用のプロファイル（HFP）やデータ通信用のプロファイルなど種々の通信プロトコルを備える。記憶部 4 3 は、例えばメモ리카ードなどの不揮発性の記憶媒体で構成されており、各種のコンピュータプログラムやアプリケーションプログラム、外部の装置や端末と相互に連携してアプリケーションを実行する連携機能を実現する連携アプリケーションなどの各種のプログラム、及び、各プログラムで使用される各種のデータなどを記憶している。

40

【 0 0 2 3 】

音声入出力部 4 4 は、図示しないマイクロホン及びスピーカに接続しており、周知の音

50

声入力機能及び音声出力機能を備える。この音声入出力部 44 は、携帯端末 12 に音声処理装置 11 が通信可能に接続された状態で音声処理装置 11 にて通話アプリケーション A が起動されている場合には、図示しない通話相手の携帯端末から入力された音声に対応する音声データを音声処理装置 11 に送信し、また、音声処理装置 11 から受信した音声データを通話相手の携帯端末に送信することが可能な状態となる。これにより、携帯端末 12 は、音声処理装置 11 と協働して、いわゆるハンズフリー通話を実現可能となる。なお、この音声入出力部 44 は、携帯端末 12 に音声処理装置 11 が通信可能に接続されていない状態では、マイクロホンから入力された発話音声を制御部 41 に出力する一方、制御部 41 から入力される受話音声をスピーカから出力する。これにより、携帯端末 12 は、単独でも通話機能を実現可能である。

10

【0024】

表示出力部 45 は、例えば液晶表示器や有機 EL 表示器で構成されており、制御部 41 からの表示指令信号に基づいて各種の情報を表示する。この表示出力部 45 の画面には、周知の感圧方式、電磁誘導方式、静電容量方式あるいはそれらを組み合わせた方式で構成されるタッチパネルスイッチが設けられる。この表示出力部 45 には、アプリケーションに対する操作を入力するための操作入力画面などの入力インターフェース、アプリケーションの実行内容や実行結果を出力するための出力画面などの出力インターフェースなどの各種画面が表示される。

【0025】

操作入力部 46 は、表示出力部 45 の画面上に設けられるタッチパネルスイッチ、及び、表示出力部 45 の周囲に設けられているメカニカルスイッチなど各種のスイッチ類を含む。操作入力部 46 は、使用者による各種のスイッチの操作に応じて操作検知信号を制御部 41 に出力する。制御部 41 は、操作入力部 46 から入力された操作検知信号を解析して使用者の操作内容を特定し、特定した操作内容に基づいて各種の処理を実行する。

20

【0026】

電話通信部 47 は、通信網 100 との間に無線の電話通信回線確立し、この電話通信回線を介して電話通信を実行する。この場合、通信網 100 は、図示しない携帯電話基地局や基地局制御装置などの周知の公衆回線網を使用する携帯電話通信サービスを提供する設備を含む。また、制御部 41 は、この電話通信部 47 を介して、通信網 100 に接続している配信センター 14 あるいは音認検索サーバ 15 に通信可能に接続される。

30

【0027】

次に、上記構成の音声処理システム 10 において、通話アプリケーション A（以下、通話アプリ A と称する）を実行する場合における制御内容の一例について説明する。即ち、例えば図 4 に示すように、音声処理装置 11 は、当該音声処理装置 11 にて通話アプリ A が起動されたか否か（A1）及び外部の携帯端末 12 から着信操作が入力されたか否か（A2）を監視している。音声処理装置 11 は、通話アプリ A が起動している場合（A1：YES）には、使用者が通話アプリ A を介して発信操作を入力したか否か（A3）を監視する。なお、発信操作は、通話アプリ A における自発的な操作の一例であり、外部の携帯端末に対し発信を行うことをいう。そして、音声処理装置 11 は、発信操作が入力されると（A3：YES）、通常モードからハンズフリー通話モードに移行する（A4）。また、音声処理装置 11 は、通話アプリ A が起動されていない状態で着信操作が入力されると（A2：YES）、通話アプリ A を起動する（A5）。そして、音声処理装置 11 は、通常モードからハンズフリー通話モードに移行する（A4）。なお、着信操作は、通話アプリ A における他発的な操作の一例であり、外部の携帯端末から着信を受けることをいう。携帯端末 12 は、外部の携帯端末から着信が有り、且つ、ハンズフリー通話モードに移行している場合には、音声処理装置 11 に対し着信操作を入力するように設定されている。

40

【0028】

ハンズフリー通話モードでは、音声処理装置 11 は、携帯端末 12 との間に HFP による無線通信回線確立して、マイクロホンから入力された音声に対応する音声データを携帯端末 12 に送信し、また、携帯端末 12 から受信した音声データに基づき音声をスピー

50

力から出力することが可能な状態となる。

【0029】

一方、携帯端末12は、図示しない外部の携帯端末から着信を受けると(B1:YES)、音声処理装置11との間にHFPによる無線通信回線が確立されているか否かを確認する(B2)。携帯端末12は、音声処理装置11との間にHFPによる無線通信回線が確立されてない場合には(B2:NO)、通常通話モードにて当該携帯端末12単独で通話を実行する(B3)。即ち、携帯端末12と通話相手の携帯端末との間で通常通話が行われる。

【0030】

一方、携帯端末12は、音声処理装置11との間にHFPによる無線通信回線が確立されている場合には(B2:YES)、通常通話モードからハンズフリー通話モードに移行する(B4)。このハンズフリー通話モードでは、携帯端末12は、音声処理装置11との間に確立されているHFPによる無線通信回線を介して、図示しない通話相手の携帯端末から入力された音声に対応する音声データを音声処理装置11に送信し、また、音声処理装置11から受信した音声データを通話相手の携帯端末に送信することが可能な状態となる。音声処理システム10は、このように音声処理装置11及び携帯端末12の双方がハンズフリー通話モードに移行することにより、いわゆるハンズフリー通話が可能な状態となる。

【0031】

音声処理装置11は、ハンズフリー通話モードに移行すると、音声データ取得処理部31によって音声データを取得し(A6)、その取得した音声データに、音声処理部33によって通話用の音声処理を施す(A7)。この場合、音声処理装置11は、通話アプリAの自発的な操作または他発的な操作を検知しており、これにより、実行中のアプリケーションが通話アプリAであることを確認している。よって、音声処理装置11は、音声データに施す音声処理を、通話用の音声処理に切り替えている。そして、音声処理装置11は、通話用の音声処理を施した音声データを携帯端末12に送信する(A8)。なお、ステップA6の処理は、音声データ取得ステップの一例であり、ステップA7の処理は、音声処理ステップの一例であり、ステップA8の処理は、音声データ送信ステップの一例である。

【0032】

携帯端末12は、音声処理装置11から受信した音声データを通話相手の携帯端末に送信する(B5)。また、携帯端末12は、通話相手の携帯端末から音声データを受信すると(B6)、その音声データを音声処理装置11に送信する(B7)。音声処理装置11は、携帯端末12から音声データを受信すると、その音声データに基づき音声をスピーカから出力する(A9)。これにより、通話相手の携帯端末からの受話音声は音声処理装置11から出力されるようになる。このように、携帯端末12を中継して音声処理装置11と通話相手の携帯端末との間で発話音声の音声データ及び受話音声の音声データが適宜送受信されることで、いわゆるハンズフリー通話が実現される。そして、この場合、音声処理装置11において通話アプリAの自発的な操作または他発的な操作が検知された場合には、音声処理装置11から携帯端末12に送信される音声データに、通話用の音声処理が施される。なお、このハンズフリー通話は、音声処理装置11または通話相手の携帯端末にて通話が終了されるまで継続される。

【0033】

次に、上記構成の音声処理システム10において、音声認識検索アプリケーションB(以下、音認検索アプリBと称する)を実行する場合における制御内容の一例について説明する。即ち、例えば図5に示すように、音声処理装置11に携帯端末12が通信可能に接続されて、これら音声処理装置11及び携帯端末12にてそれぞれ連携アプリケーションが起動されると、携帯端末12が有する音認検索アプリBの実行処理は当該携帯端末12にて実行され、この音認検索アプリBの入力インターフェース及び出力インターフェースは音声処理装置11にて提供される状態となる。なお、このような音認検索アプリBは、

10

20

30

40

50

例えば車両が走行していない状態など走行に影響を及ぼさない状態で実行することが好ましい。

【0034】

そして、例えば図6に示すように、音声処理装置11及び携帯端末12の双方にて連携アプリケーションが起動されると(C1, D1)、音声処理装置11には、携帯端末12が有するアプリケーションの起動ボタンが表示される(C2)。なお、この起動ボタンは、入力インターフェースの一例である。そして、音声処理装置11は、音認検索アプリBの起動ボタンが操作されると(C3: YES)、音認検索アプリBの起動指令信号を携帯端末12に送信する(C4)。このとき、音声処理装置11は、位置特定部によって得られる当該音声処理装置11の現在位置を示す現在位置情報も携帯端末12に送信する。

10

【0035】

携帯端末12は、音認検索アプリBの起動指令信号を受信すると、音認検索アプリBを起動する(D2)。そして、携帯端末12は、音認検索アプリBを起動したことを示す起動完了信号を音認検索サーバ15に送信する(D3)。このとき、携帯端末12は、音声処理装置11から受信した現在位置情報も音認検索サーバ15に送信する。

音認検索サーバ15は、音認検索アプリBの起動完了信号を受信すると、検索条件収集用の音声データを携帯端末12に送信する(E1)。この場合、検索条件収集用の音声データとして、例えば「ご用件を教えてください。」といったメッセージデータが設定される。携帯端末12は、音認検索サーバ15から受信した検索条件収集用の音声データを音声処理装置11に送信する(D4)。

20

【0036】

音声処理装置11は、検索条件収集用の音声データを受信すると、その音声データに基づき検索条件収集用の音声をスピーカから出力する(C5)。この場合、例えば「ご用件を教えてください。」といった案内音声出力される。この案内音声に応じて、使用者が例えば「イタリアン」などといった検索条件を発声すると、音声処理装置11は、その音声データを音声データ取得処理部31によって取得し(C6)、その取得した音声データに、音声処理部33によって音認検索用の音声処理を施す(C7)。この場合、音声処理装置11は、通話アプリAの自発的な操作または他発的な操作を検知しておらず、これにより、実行中のアプリケーションが通話アプリA以外のアプリケーションであることを確認している。よって、音声処理装置11は、音声データに施す音声処理を、通話以外用の音声処理の一例である音認検索用の音声処理に切り替えている。そして、音声処理装置11は、音認検索用の音声処理を施した音声データを携帯端末12に送信する(C8)。なお、ステップC6の処理は、音声データ取得ステップの一例であり、ステップC7の処理は、音声処理ステップの一例であり、ステップC8の処理は、音声データ送信ステップの一例である。

30

【0037】

また、本実施形態では、実行中のアプリケーションが通話アプリA以外のアプリケーションである場合には、一律、音認検索用のノイズキャンセル処理を施す例を述べた。しかし、例えば、実行中のアプリケーションを特定するためのアプリ特定データを携帯端末12から音声処理装置11に送信し、音声処理装置11は、そのアプリ特定データによって特定されるアプリケーションに適した音声処理を切り替えて実行するように構成してもよい。

40

【0038】

携帯端末12は、音声処理装置11から受信した音声データを音認検索サーバ15に送信する(D5)。一方、音認検索サーバ15は、携帯端末12から音声データを受信すると、その音声データに基づき周知の音声認識処理を実施する(E2)。そして、音認検索サーバ15は、認識した音声および音声処理装置11の位置情報に基づき周知の検索処理を実行し(E3)、その検索結果を示す検索結果データを携帯端末12に送信する(E4)。このとき、音認検索サーバ15は、検索結果出力用の音声データも携帯端末12に送信する。この場合、検索結果出力用の音声データとして、例えば「近くのイタリアンの店

50

を表示します。」といったメッセージデータが設定される。即ち、音認検索サーバ１５は、検索結果出力用の音声データに、例えば「イタリアン」といった検索条件も反映させる。

【００３９】

携帯端末１２は、音認検索サーバ１５から受信した検索結果データを音声処理装置１１に送信する（Ｄ６）。このとき、携帯端末１２は、音認検索サーバ１５から受信した検索結果出力用の音声データも音声処理装置１１に送信する。一方、音声処理装置１１は、検索結果出力用の音声データを受信すると、その音声データに基づき音声をスピーカから出力する（Ｃ９）。この場合、例えば「近くのイタリアンの店を表示します。」といった案内音声出力される。また、音声処理装置１１は、検索結果データを受信すると、その検索結果データに基づき検索結果を表示する（Ｃ１０）。なお、これら検索結果の出力音声及び検索結果の表示画面は、出力インターフェースの一例である。このように、携帯端末１２を中継して音声処理装置１１と音認検索サーバ１５との間で音声データ及び検索結果データが適宜送受信されることで、音声認識を利用した検索サービスが実現される。そして、この場合、音声処理装置１１において通話アプリＡの自発的な操作または他発的な操作が検知されず、従って、音声処理装置１１から携帯端末１２に送信される音声データに、音声認識用の音声処理が施される。

【００４０】

本実施形態によれば、音声処理装置１１は、取得した音声データを外部の携帯端末１２に送信する場合に、その送信する音声データに所定の音声処理を施す。そして、その音声処理として、通話用の音声処理の一例である通話用の音声処理と通話以外用の音声処理の一例である音認検索用の音声処理とを切り替えて実行することが可能である。よって、起動中のアプリケーションに応じて通話用の音声処理及び通話以外用の音声処理を適宜切り替えて実行することができ、通話用の音声処理及び通話以外用の音声処理を何れも最適に実施することができる。なお、音声データに施す音声処理としては、ノイズキャンセル処理、エコーキャンセル処理、ノイズキャンセル処理の絞りを徐々に大きくしていくオートゲインコントロール処理などといった処理を単発で実施するように構成してもよいし、あるいは各処理を適宜組み合わせるよう実施するように構成してもよい。

【００４１】

また、本実施形態によれば、音声処理装置１１は、通話アプリＡにおける自発的な操作または他発的な操作を検知した場合に、通話用の音声処理を実行する。即ち、通話アプリＡに特有の操作、換言すれば、通話アプリＡ以外のアプリケーションでは発生し得ない操作を検知したか否かに基づき、音声データに施す音声処理を通話用の音声処理に切り替える。従って、通話アプリＡの実行時に、確実に通話用の音声処理を実行することができる。また、通話アプリＡ以外のアプリケーションの実行時には、確実に通話以外用の音声処理を実行することができる。

【００４２】

また、本実施形態によれば、通話用の音声データ及び通話以外用の音声データである音声認識用の音声データを何れも同一の通信プロトコルによって送受信するように構成した。これにより、例えば通話以外用のアプリケーションを新たに追加する場合であっても、そのアプリケーションに係る音声データを同一のプロトコルで送受信することができる。また、アプリケーションを追加するたびに専用の通信プロトコルを開発する必要がなく、開発コストの低減を図ることができる。

【００４３】

なお、本発明は、上述した一実施形態のみに限定されるものではなく、その要旨を逸脱しない範囲で種々の実施形態に適用可能である。

例えば、通話アプリケーションは携帯端末で実行するように構成してもよい。また、音認検索アプリケーションは音声処理装置で実行するように構成してもよい。

【００４４】

また、音声処理装置１１、より具体的には音声処理部３３は、通話アプリケーション以

10

20

30

40

50

外のアプリケーションが起動された場合に音声処理を実行しないように構成し、代わりに、携帯端末 1 2 または音認検索サーバ 1 5 が音声処理を実行するように構成してもよい。この構成によれば、音声処理装置 1 1 の処理負荷を抑えることができる。また、携帯端末 1 2 あるいは音認検索サーバ 1 5 にて、特化した音声認識を実施することができる。

【 0 0 4 5 】

即ち、例えば図 7 に示すように、音声処理システム 1 0 は、音声処理装置 1 1 では音声認識用の音声処理、換言すれば音声データの信号処理を実行せず、携帯端末 1 2 にて音声認識用の信号処理を実行するように構成してもよい。また、例えば図 8 に示すように、音声処理システム 1 0 は、音声処理装置 1 1 及び携帯端末 1 2 では音声認識用の信号処理を実行せず、音認検索サーバ 1 5 にて音声認識用の信号処理を実行するように構成してもよい。

10

【 0 0 4 6 】

また、例えば図 9 に示すように、音声処理システム 1 0 は、音声処理装置 1 1 及び携帯端末 1 2 の双方に通話アプリを備え、音声処理装置 1 1 にて通話用の音声データに対し通話用の音声処理を施す構成とし、携帯端末 1 2 では通話用の音声データに対して通話用の音声処理を施さない、または、付加的な音声処理を施す構成としてもよい。なお、図示はしないが、音声処理システム 1 0 は、音声処理装置 1 1 では通話用の音声データに対して通話用の音声処理を施さない、または、付加的な音声処理を施す構成とし、携帯端末 1 2 にて通話用の音声データに対し通話用の音声処理を施す構成としてもよい。

【 0 0 4 7 】

20

また、例えば図 1 0 に示すように、音声処理システム 1 0 は、音認検索サーバ 1 5 に対応する音認検索アプリ 1 6 及び音認検索サーバ 1 5 に対応する音認検索アプリ 1 7 を携帯端末 1 2 に備える構成としてもよい。そして、音認検索アプリ 1 6 により音認検索サーバ 1 5 の検索サービスを利用する場合には、携帯端末 1 2 では音認用の音声データに対して音認用の音声処理を施さず、音認検索サーバ 1 5 にて音認用の音声データに対し音認用の音声処理を施す構成としてもよい。また、音認検索アプリ 1 7 により音認検索サーバ 1 5 の検索サービスを利用する場合には、携帯端末 1 2 にて音認用の音声データに対し音認用の音声処理を施し、音認検索サーバ 1 5 では音認用の音声データに対し音認用の音声処理を施さない構成としてもよい。つまり、音声処理システム 1 0 は、利用する音認検索アプリの種類に応じて、音声データに対し音認用の音声処理を施す実行主体を適宜変更する構成としてもよい。

30

【 0 0 4 8 】

通話アプリケーション以外のアプリケーションは、音声認識処理を要するサービスを実現するアプリケーションであればよく、音認検索アプリケーションに限られるものではない。

音声処理装置 1 1 は、例えばナビゲーション機能を有するアプリケーションプログラムがインストールされた機器で構成してもよい。また、音声処理装置 1 1 は、車両に組み込まれる車載装置で構成してもよいし、車両に着脱可能な携帯型の無線装置などで構成してもよい。

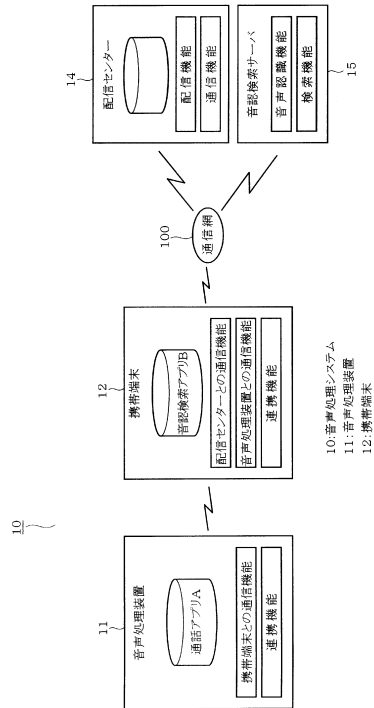
【 符号の説明 】

【 0 0 4 9 】

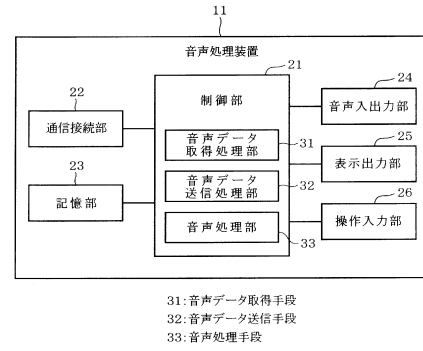
40

図面中、1 0 は音声処理システム、1 1 は音声処理装置、1 2 は携帯端末、3 1 は音声データ取得処理部（音声データ取得手段）、3 2 は音声データ送信処理部（音声データ送信手段）、3 3 は音声処理部（音声処理手段）を示す。

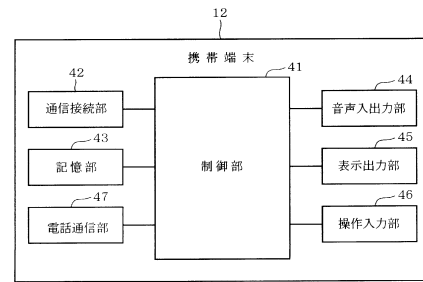
【図 1】



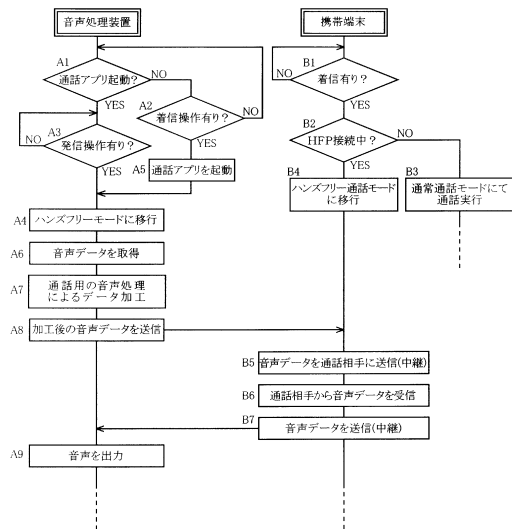
【図 2】



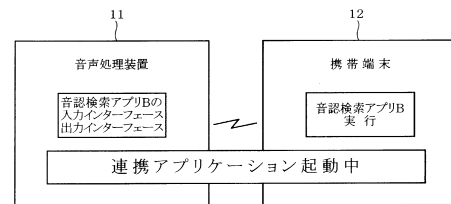
【図 3】



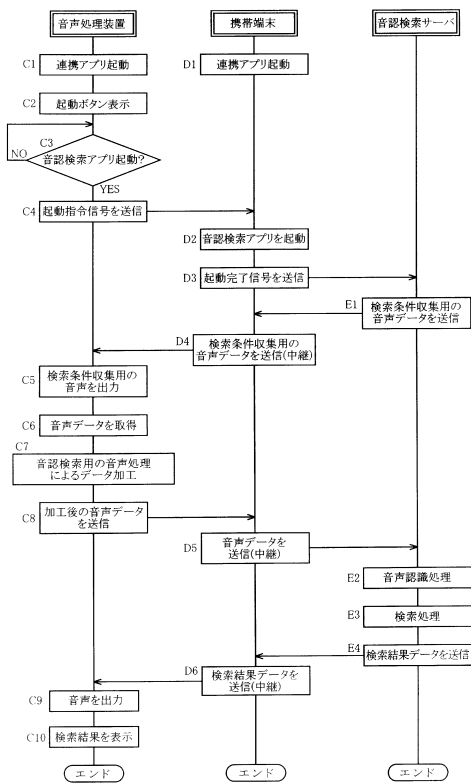
【図 4】



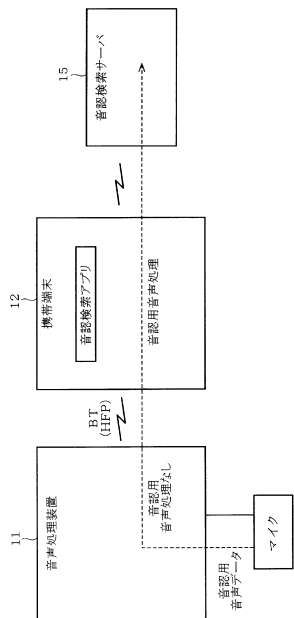
【図 5】



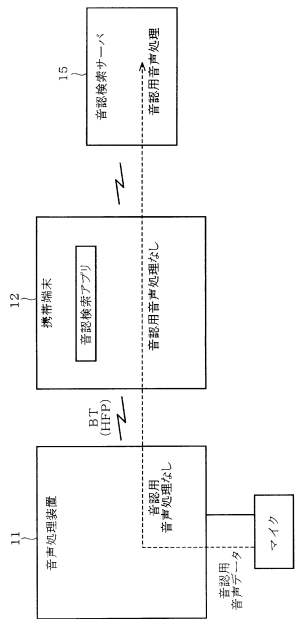
【図 6】



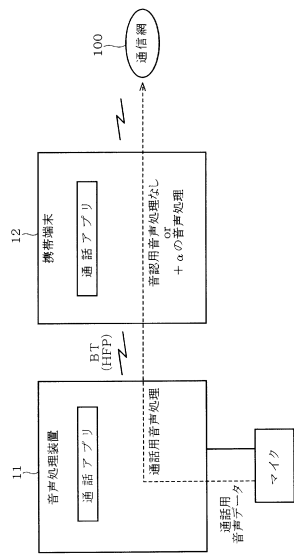
【図 7】

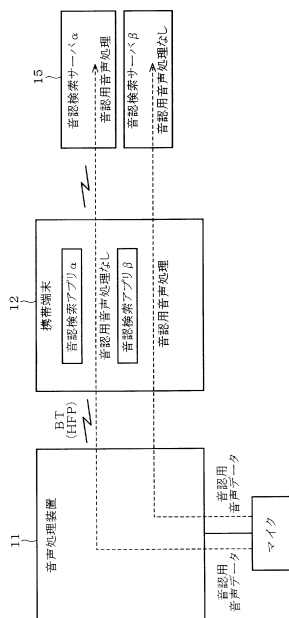


【図 8】



【図 9】





フロントページの続き

(72)発明者 鵜飼 拡基
愛知県刈谷市昭和町1丁目1番地 株式会社デンソー内

審査官 望月 章俊

(56)参考文献 特開2004-153380(JP,A)
特開2006-222969(JP,A)
特開2004-343335(JP,A)
米国特許出願公開第2013/0332886(US,A1)

(58)調査した分野(Int.Cl., DB名)
H04M1/00
G10L15/00
H04M1/60
H04M3/42
H04M11/00