

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2015年7月9日(09.07.2015)



(10) 国際公開番号

WO 2015/102040 A1

(51) 国際特許分類:
H04M 1/00 (2006.01) *H04M 3/42* (2006.01)
G10L 15/00 (2013.01) *H04M 11/00* (2006.01)
H04M 1/60 (2006.01)

(21) 国際出願番号: PCT/JP2014/006172

(22) 国際出願日: 2014年12月11日(11.12.2014)

(25) 国際出願の言語: 日本語

(26) 国際公開の言語: 日本語

(30) 優先権データ:
特願 2014-000285 2014年1月6日(06.01.2014) JP

(71) 出願人: 株式会社デンソー(DENSO CORPORATION) [JP/JP]; 〒4488661 愛知県刈谷市昭和町1丁目1番地 Aichi (JP).

(72) 発明者: 伊藤 正也(ITO, Masaya); 〒4488661 愛知県刈谷市昭和町1丁目1番地株式会社デンソー内 Aichi (JP). 尾崎 義隆(OZAKI, Yoshitaka); 〒4488661 愛知県刈谷市昭和町1丁目1番地株式会社デンソー内 Aichi (JP). 林 圭作(HAYASHI,

Keisaku); 〒4488661 愛知県刈谷市昭和町1丁目1番地株式会社デンソー内 Aichi (JP). 鶴飼 拡基(UKAI, Hiroki); 〒4488661 愛知県刈谷市昭和町1丁目1番地株式会社デンソー内 Aichi (JP).

(74) 代理人: 金 順姫(KIN, Junhi); 〒4600003 愛知県名古屋市中区錦2丁目13番19号 滝定ビル6階 Aichi (JP).

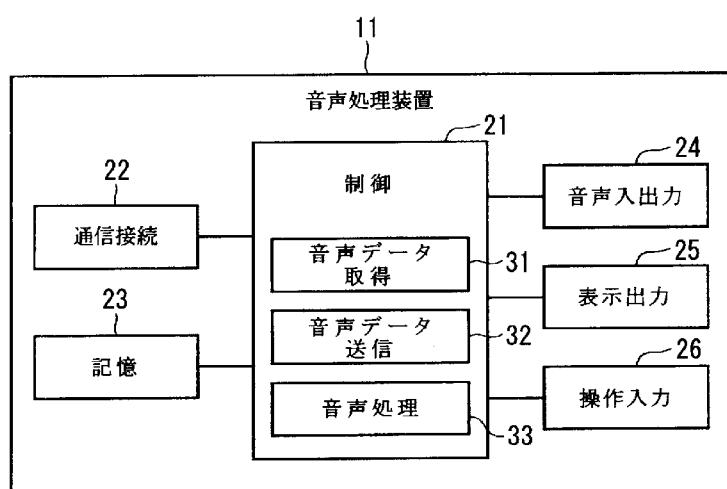
(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, KE, KG, KN, KP, KR, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, ZA, ZM, ZW.

(84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユー

[続葉有]

(54) Title: SPEECH PROCESSING APPARATUS, SPEECH PROCESSING SYSTEM, SPEECH PROCESSING METHOD, AND PROGRAM PRODUCT FOR SPEECH PROCESSING

(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、音声処理用のプログラム製品



(57) Abstract: A speech processing apparatus (11), when transmitting received speech data to an external mobile terminal (12), causes a speech processing unit (33) to apply predetermined speech processing to the speech data to be transmitted. The speech processing unit is capable of switching between performing, for speech processing, first speech processing for telephone conversations and second speech processing for purposes other than telephone conversations.

(57) 要約: 音声処理装置(Speech Processing Apparatus) (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、音声処理部(33)によって所定の音声処理を施す。音声処理部は、その音声処理として、通話用の第一の音声処理と通話以外用の第二の音声処理とを切り替えて、実行することが可能である。

- 11 Speech processing apparatus
21 Control
22 Communication connection
23 Memory
24 Speech input and output
25 Display output
26 Operation input
31 Speech data acquisition
32 Speech data transmission
33 Speech processing



ラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨー 添付公開書類:

ロツバ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE,
ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV,
MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK,
SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ,
GW, KM, ML, MR, NE, SN, TD, TG).

— 国際調査報告（条約第 21 条(3)）

明細書

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、音声処理用のプログラム製品

関連出願の相互参照

[0001] 本開示は、2014年1月6日に出願された日本出願番号2014-285号に基づくもので、ここにその記載内容を援用する。

技術分野

[0002] 本開示は、音声処理装置（Speech Processing Apparatus）、音声処理システム、音声処理方法、音声処理用のプログラム製品に関する。

背景技術

[0003] 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トーク技術においては、多くの車両用機器で採用されているBlueooth（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。

先行技術文献

特許文献

[0004] 特許文献1：JP 2006-238148 A

発明の概要

[0005] 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、例えばユーザの発声したスピーチを認識するところの音声認識（Speech R

ecognition) 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。

[0006] 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、エコーフィルタ処理、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。

[0007] 本開示の目的は、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用のプログラム製品を提供することにある。

[0008] 本開示の例によれば、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、その音声処理として、通話用の第一の音声処理と通話以外用の第二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第一の音声処理及び通話以外用の第二の音声処理を適宜切り替えて実行することができ、通話用の第一の音声処理及び通話以外用の第二の音声処理を何れも最適に実施することができる。

図面の簡単な説明

[0009] 本開示についての上記目的およびその他の目的、特徴や利点は、添付の図面を参照しながら下記の詳細な記述により、より明確になる。

[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

[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

[0010] 以下、本開示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（電話会話）を可能とするいわゆるハンズフリー通話機能（ハンズフリー電話会話機能とも言及される）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。

[0011] 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、例えばユーザの発声したスピーチを認識するところの音声認識（Speech Recognition）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。

[0012] 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、検索サーバ15とも称する）にも接続可能に構成されている。この音声認識検索サーバ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。

[0013] 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、制御回路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。制御回路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。制御回路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、本実施形態では、一例として、制御回路21は、コンピュータ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31、音声データ送信処理部32、音声処理部33をソフトウェアによって仮想的に実現する。但し、それぞれの処理部の各々において、機能の一部あるいは全てをハードウエ

アの構成要素として提供することも可能である。

- [0014] 通信接続部 22 は、例えば無線通信モジュールなどで構成され、携帯端末 12 が備える通信接続部 42 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 12 との間で各種の通信を行う。この場合、通信接続部 22 は、ハンズフリー通話用のプロファイル（HFP : 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。
- [0015] 記憶部 23 は、例えばハードディスクドライブなどのコンピュータ読取可能な非遷移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（インストラクションを含むプログラム製品とも言及される）、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 23 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 11 は、音声認識検索サーバ 15 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。
- [0016] 音声入出力部 24 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 24 は、音声処理装置 11 に携帯端末 12 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 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 11 は、携帯端末 12 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。
- [0017] 表示出力部 25 は、例えば液晶表示器や有機ELで構成されており、制御回路 21 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 25 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25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。

- [0018] 操作入力部26は、表示出力部25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25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26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回路21に出力する。制御回路21は、操作入力部26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11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11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。
- [0019] 音声データ取得処理部31は、音声データ取得部／デバイス／ミーンズとも言及され、音声入出力部24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。
- [0020] 音声データ送信処理部32は、音声データ送信部／デバイス／ミーンズとも言及され、音声データ取得処理部31によって取得された音声データを、通信接続部22が確立した通信回線を介して外部の携帯端末12に送信する。この場合、音声データ送信処理部32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（HFP）が採用されている。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。
- [0021] 音声処理部33は、音声処理デバイス／ミーンズとも言及され、音声データ送信処理部32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33は、音声処理として、通話用の音

声処理（第一の音声処理とも言及される）と通話以外用の音声処理（第二の音声処理とも言及される）の一例である音声認識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、通話用のエコーチャンセル処理、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、音声認識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、音声認識検索用のノイズキャンセル処理、音声認識検索用のエコーチャンセル処理、音声認識検索用のゲインコントロール処理などを含む。この音声認識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。

[0022] 基本的には、音声認識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、エコーチャンセル、ゲインコントロールがかけられる。一方、音声認識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、エコーチャンセル、ゲインコントロールがかけられる。即ち、音声認識検索用の音声処理では、本来の音声情報（音声波形）が変化してしまうことを極力防止することが求められる。

[0023] 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声認識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などして、本来の音声波形に極力近い音声波形が残るような処理、つまり、通話用の音声処理よりも、音声波形が原形に近い状態で残るような音声処理を行うことが好ましい。

[0024] 次に、携帯端末12の構成について図3を参照しながら説明する。携帯端末12は、制御回路41、通信接続部42、記憶部43、音声入出力部44、表示出力部45、操作入力部46、電話通信部47などを備えている。制御回路41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。本実施形態では、一例として、制御回路41は、ROMあるいは記憶部43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12の動作全般を制御する。一方、制御回路41の機能の一部あるいは全ては、ハードウェアの構成要素によっても、実行は可能である。

[0025] 通信接続部42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11が備える通信接続部22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11との間で各種の通信を行う。この場合、通信接続部42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43は、例えばメモリカードなどのコンピュータ読取可能な非遷移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（インストラクションを含むプログラム製品とも言及される）、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。

[0026] 音声入出力部44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回路 41 に出力する一方、制御回路 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。

[0027] 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、制御回路 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。

[0028] 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回路 41 に出力する。制御回路 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。

[0029] 電話通信部 47 は、通信網 100 との間に無線の電話通信回線を確立し、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、制御回路 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 あるいは音声認識検索サーバ 15 に通信可能に接続される。

- [0030] 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 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。
- [0031] ここで、この出願に記載されるフローチャート、あるいは、フローチャートの処理は、複数のセクション（あるいはステップと言及される）を含み、各セクションは、たとえば、A 1, B 1, C 1, D 1, E 1 と表現される。さらに、各セクションは、複数のサブセクションに分割することができる、一方、複数のセクションが合わさって一つのセクションにすることも可能である。さらに、各セクションは、デバイス、モジュール、ミーンズとして言及することができる。また、上記の複数のセクションの各々あるいは組合わさったものは、(i)ハードウェアユニット（例えば、コンピュータ）と組み合わさったソフトウェアのセクションのみならず、(ii)ハードウェア（例えば、集積回路、配線論理回路）のセクションとして、関連する装置の機能を含みあるいは含まずに実現できる。さらに、ハードウェアのセクションは、マイクロコンピュータの内部に含まれることもできる。
- [0032] 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 にて通話アプリケーション A が起動されたか否か (A 1) 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 (A 2) を監視している。音声処理装置 11 は、通話アプリケーション A が起動している場合 (A 1 : YES) には、使用者が通話アプリケーション A を介して発信操作を入力したか否か (A 3) を監視する。なお、発信操作は、通話アプリケーション 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 (A 3 : YES) 、通常モードからハンズフリー通話モードに移行する (A 4)。また、音声処理装置 11 は、通話アプリケーション A が起動されていない状態で着信操作が入力されると (A 2 : YES) 、通話アプリケーション A を起動する (A 5)。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 (A 4)。なお、着信操作は、通話アプリケーション 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。

[0033] 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を確立して、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カから出力することが可能な状態となる。

[0034] 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の通話が行われる。

[0035] 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。

[0036] 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、通話アプリケーション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ケーション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。

[0037] 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が音声処理装置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ケーション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。

[0038] 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、検索アプリケーションBとも称する）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が有する音声認識検索アプリケーションBの実行処理は当

該携帯端末12にて実行され、この音声認識検索アプリケーション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、このような音声認識検索アプリケーションBは、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。

- [0039] 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、音声認識検索アプリケーションBの起動ボタンが操作されると（C3：YES）、音声認識検索アプリケーション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。
- [0040] 携帯端末12は、音声認識検索アプリケーションBの起動指令信号を受信すると、音声認識検索アプリケーションBを起動する（D2）。そして、携帯端末12は、音声認識検索アプリケーションBを起動したことを示す起動完了信号を音声認識検索サーバ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声認識検索サーバ15に送信する。
- [0041] 音声認識検索サーバ15は、音声認識検索アプリケーション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「ご用件を言ってください。」といったメッセージデータが設定される。携帯端末12は、音声認識検索サーバ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）。
- [0042] 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「ご用件を言ってください。」といった案内音声が出力される。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によって音声認識検索用の音声処理を施す（C7）。この場合、音声処理装置11は、通話アプリケーション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ケーション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声認識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、音声認識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。

[0043] 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ケーションA以外のアプリケーションである場合には、一律、音声認識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。

[0044] 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声認識検索サーバ15に送信する（D5）。一方、音声認識検索サーバ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、音声認識検索サーバ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、音声認識検索サーバ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、音声認識検索サーバ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。

[0045] 携帯端末12は、音声認識検索サーバ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、音声認識検索サーバ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。」といった案内音声が出力される。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と音声認識検索サーバ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ケーション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。

[0046] 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声認識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て実施するように構成してもよい。

- [0047] また、本実施形態によれば、音声処理装置11は、通話アプリケーション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、通話アプリケーションAに特有の操作、換言すれば、通話アプリケーション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、通話アプリケーション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、通話アプリケーション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。
- [0048] 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。
- [0049] なお、本開示は、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。
- [0050] 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、音声認識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。
- [0051] 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12または音声認識検索サーバ15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、音声処理裝

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 あるいは音声認識検索サーバ 1 5 にて、特化した音声認識を実施することができる。

- [0052] 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、音声認識検索サーバ 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。
- [0053] 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。
- [0054] 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、音声認識検索サーバ α に対応する音声認識検索アプリ α 及び音声認識検索サーバ β に対応する音声認識検索アプリ β を携帯端末 1 2 に備える構成としてもよい。そして、音声認識検索アプリ α により音声認識検索サーバ α の検索サービスを利用する場合には、携帯端末 1 2 では音声認識用の音声データに対して音声認識用の音声処理を施さず、音声認識検索サーバ α にて音声認識用の音声データに対し音声認識用の音声処理を施す構成としてもよい。また、音声認識検索アプリ β により音声認識検索サーバ β の検索サービスを利用する場合には、携帯端末 1 2 にて音声認識用の音声データに対し音声認識用の音声処理を施し、音声認識検索サーバ β では音声認識用の音声データに対し音声認識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、利用する音声認識検索アプリの種類に応じて、音声データに対し音声認識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。

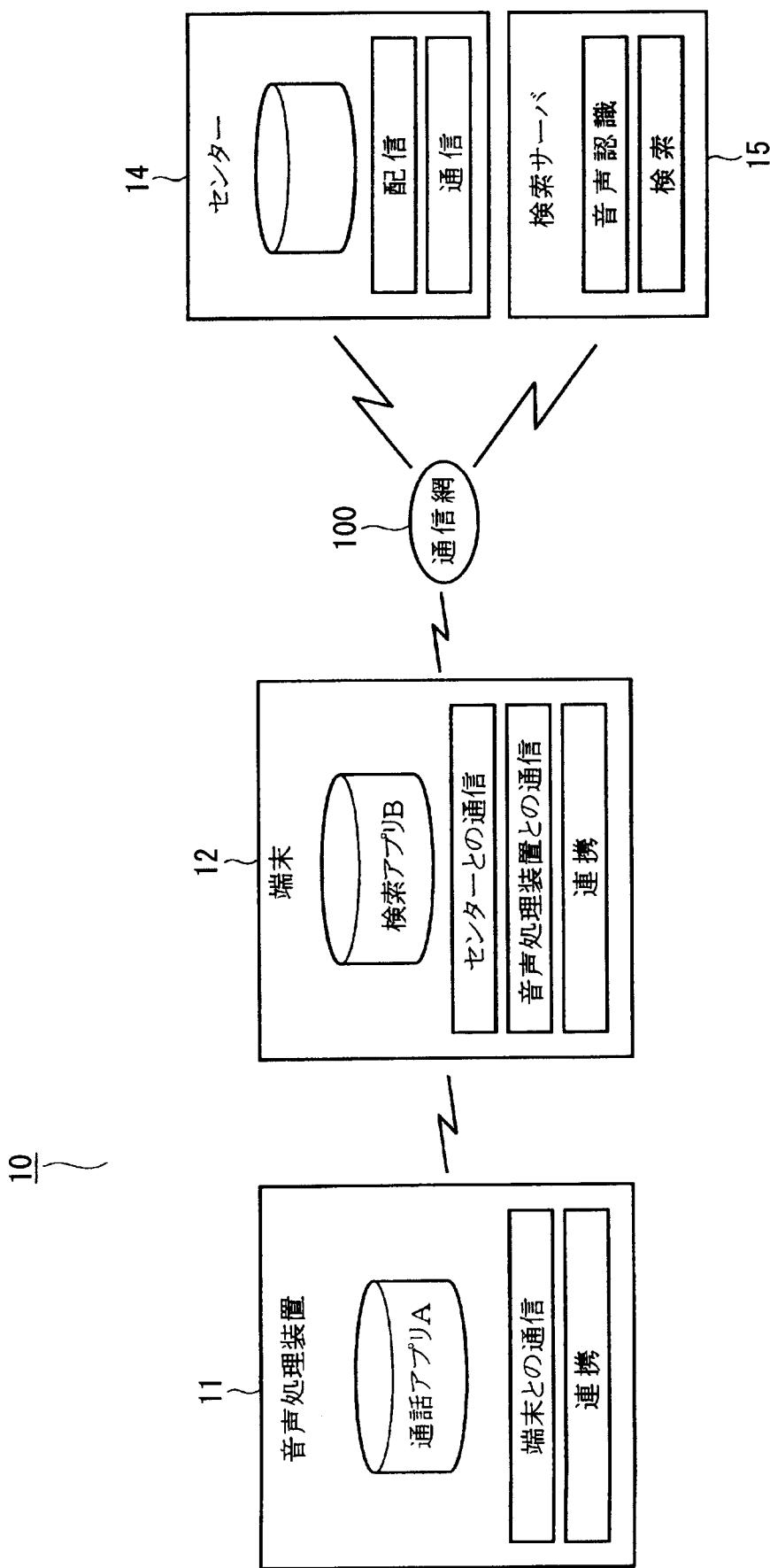
- [0055] 通話アプリケーション以外のアプリケーションは、音声認識処理を要するサービスを実現するアプリケーションであればよく、音声認識検索アプリケーションに限られるものではない。
- [0056] 音声処理装置11は、例えばナビゲーション機能を有するアプリケーションプログラムがインストールされた機器で構成してもよい。また、音声処理装置11は、車両に組み込まれる車載装置で構成してもよいし、車両に着脱可能な携帯型の無線装置などで構成してもよい。
- [0057] 本開示は、実施例に準拠して記述されたが、本開示は当該実施例や構造に限定されるものではないと理解される。本開示は、様々な変形例や均等範囲内の変形をも包含する。加えて、様々な組み合わせや形態、さらには、それらに一要素のみ、それ以上、あるいはそれ以下、を含む他の組み合わせや形態をも、本開示の範疇や思想範囲に入るものである。

請求の範囲

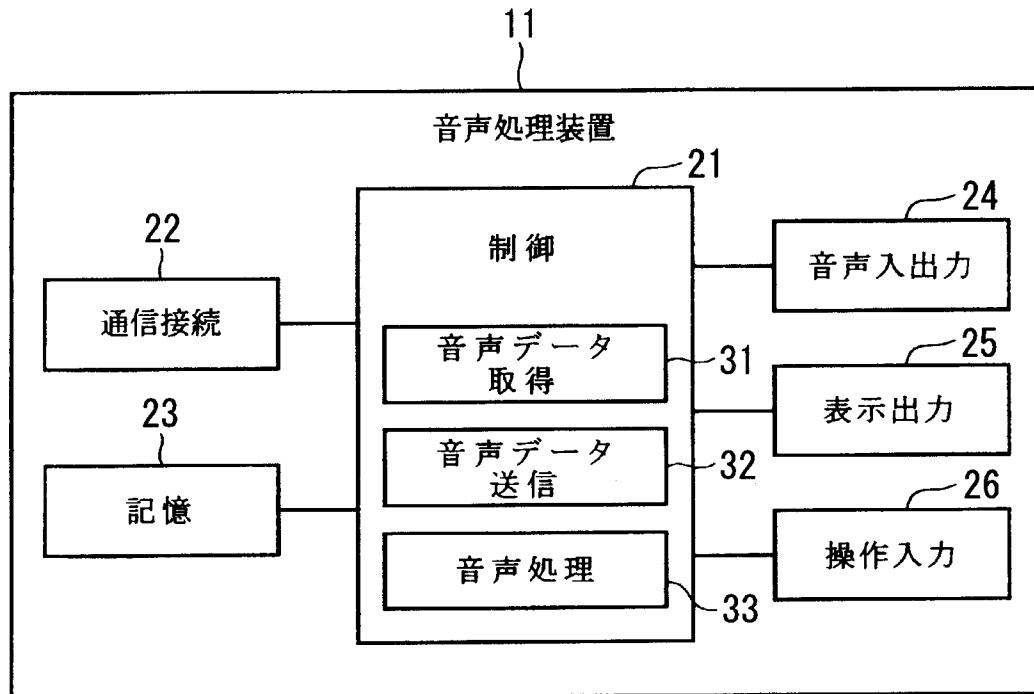
- [請求項1] 音声データを取得する音声データ取得部（31）と、
前記音声データ取得部が取得した音声データを外部の携帯端末（1
2）に送信する音声データ送信部（32）と、
前記音声データ送信部が送信する音声データに所定の音声処理を施
す音声処理部（33）と、を備え、
前記音声処理部は、前記所定の音声処理として、通話用の第一の音
声処理と通話以外用の第二の音声処理を切り替えて、実行可能に構成
されている
音声処理装置。
- [請求項2] 前記音声処理部は、通話アプリケーションにおける自発的な操作ま
たは他発的な操作を検知した場合に、前記通話用の第一の音声処理を
実行する請求項1に記載の音声処理装置。
- [請求項3] 前記音声処理部は、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、前記通話以外用の第二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] 前記音声処理部は、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
用の第二の音声処理である音声認識用の音声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] 前記音声処理部は、通話以外用の、通話用の音声処理よりも音声波
形が残る第二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、その通話以外用の第
二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] 前記音声処理部は、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部は、通話用の第一の音声データ及び通話以外用の第二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] 前記音声データ送信部は、前記通信プロトコルとして、Blueooth通信規格（Blueooth：登録商標）のハンズフリーブルートン用のプロファイルを用いる請求項7に記載の音声処理装置。
- [請求項9] 請求項1から8の何れかの1項に記載の音声処理装置（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。
- [請求項10] コンピュータによって、実行される音声処理方法であって、音声データを取得することと、前記取得された音声データを外部の携帯端末に送信することと、前記送信される音声データに所定の音声処理を施す音声処理を実行することと、を含み、前記音声処理を実行することでは、前記所定の音声処理として、通話用の第一の音声処理と通話以外用の第二の音声処理を切り替えて、実行する音声処理方法。
- [請求項11] 非遷移の記憶媒体に保管されるプログラム製品であり、コンピュータにより読み出されて実行されるインストラクションを含み、前記インストラクションは請求項10に記載の前記音声処理方法を含む、音声処理用のプログラム製品。

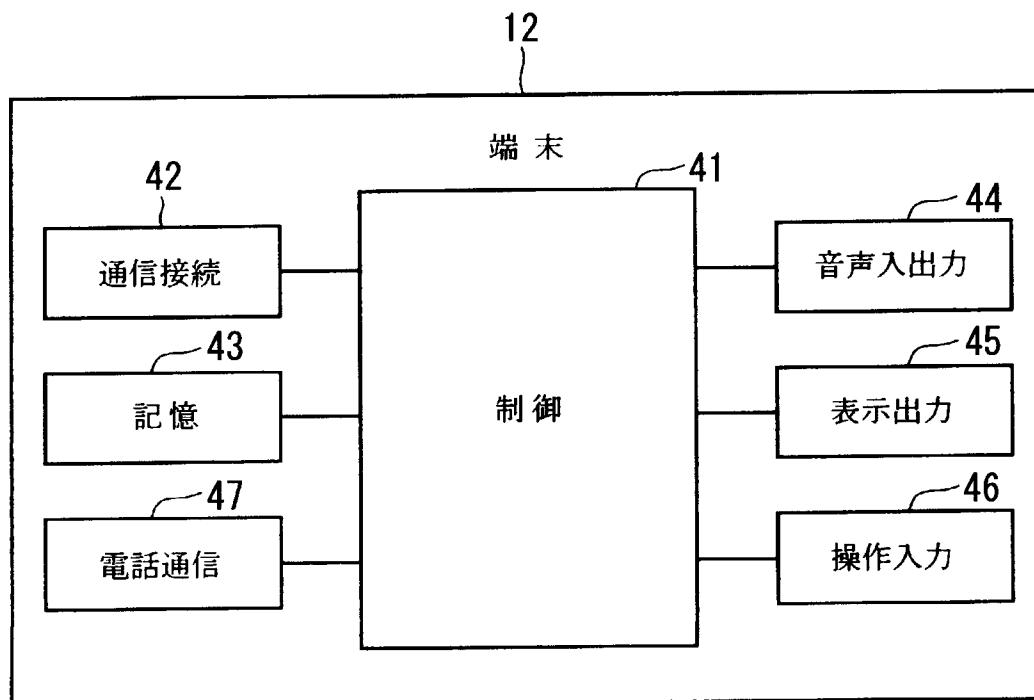
[図1]



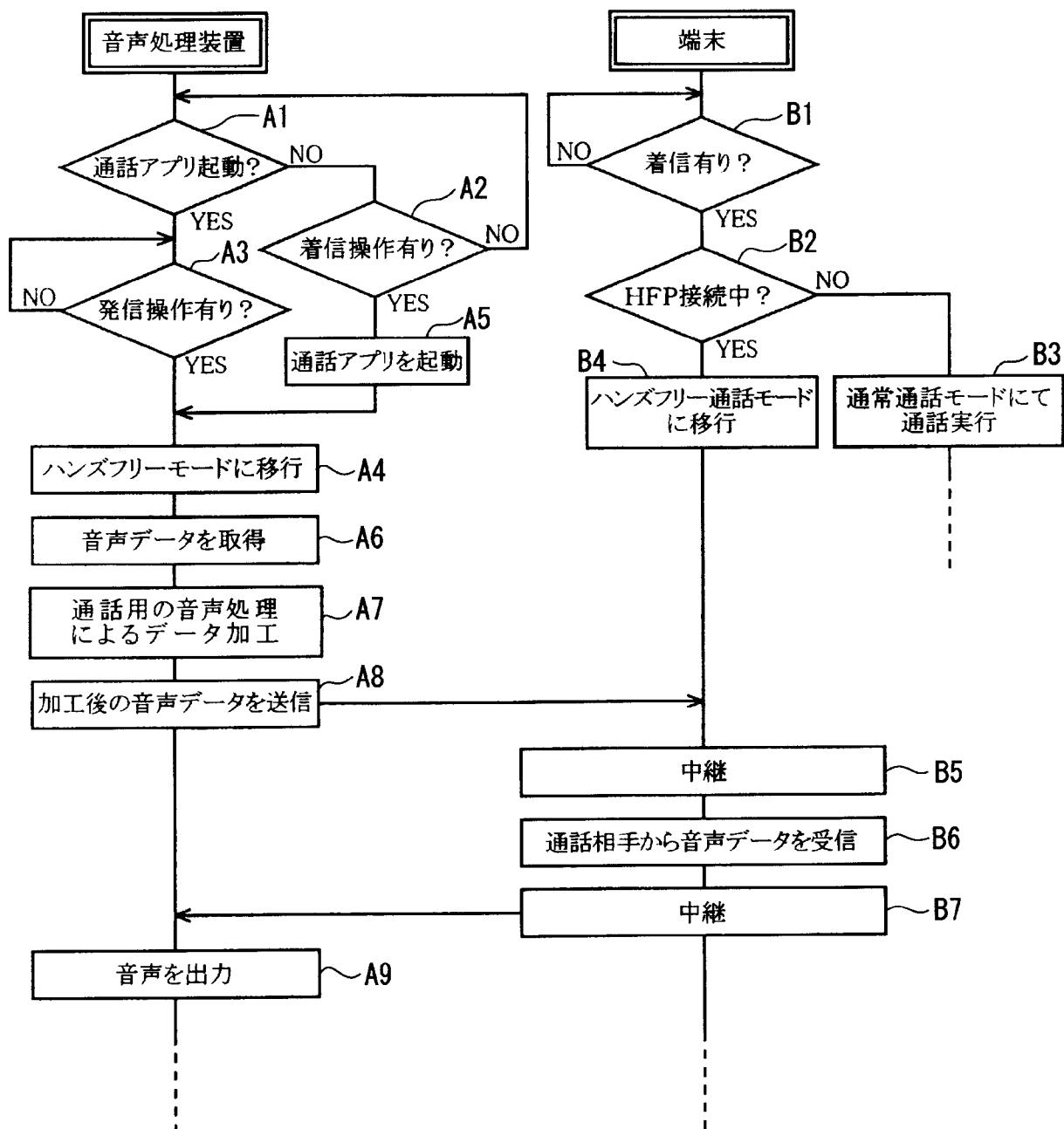
[図2]



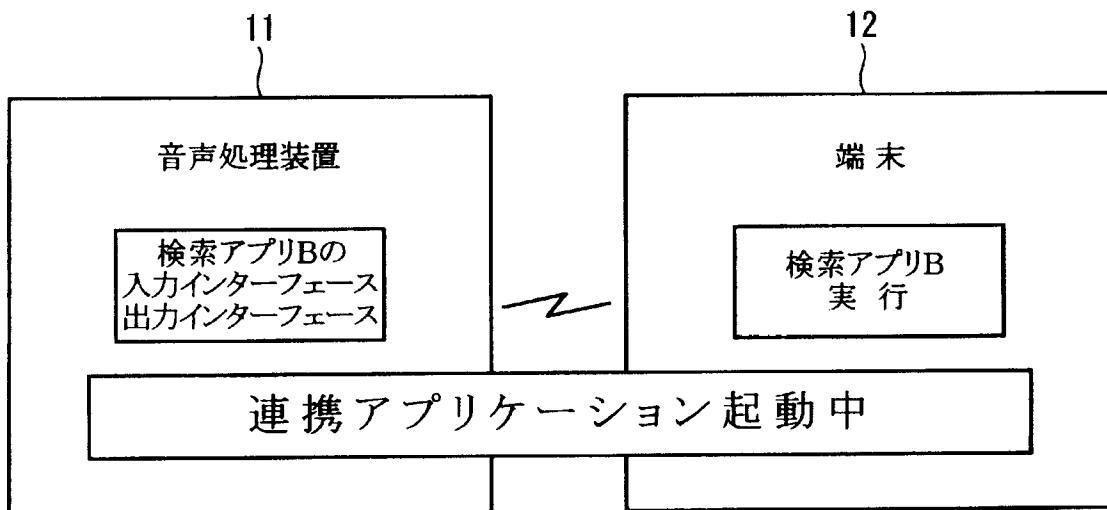
[図3]



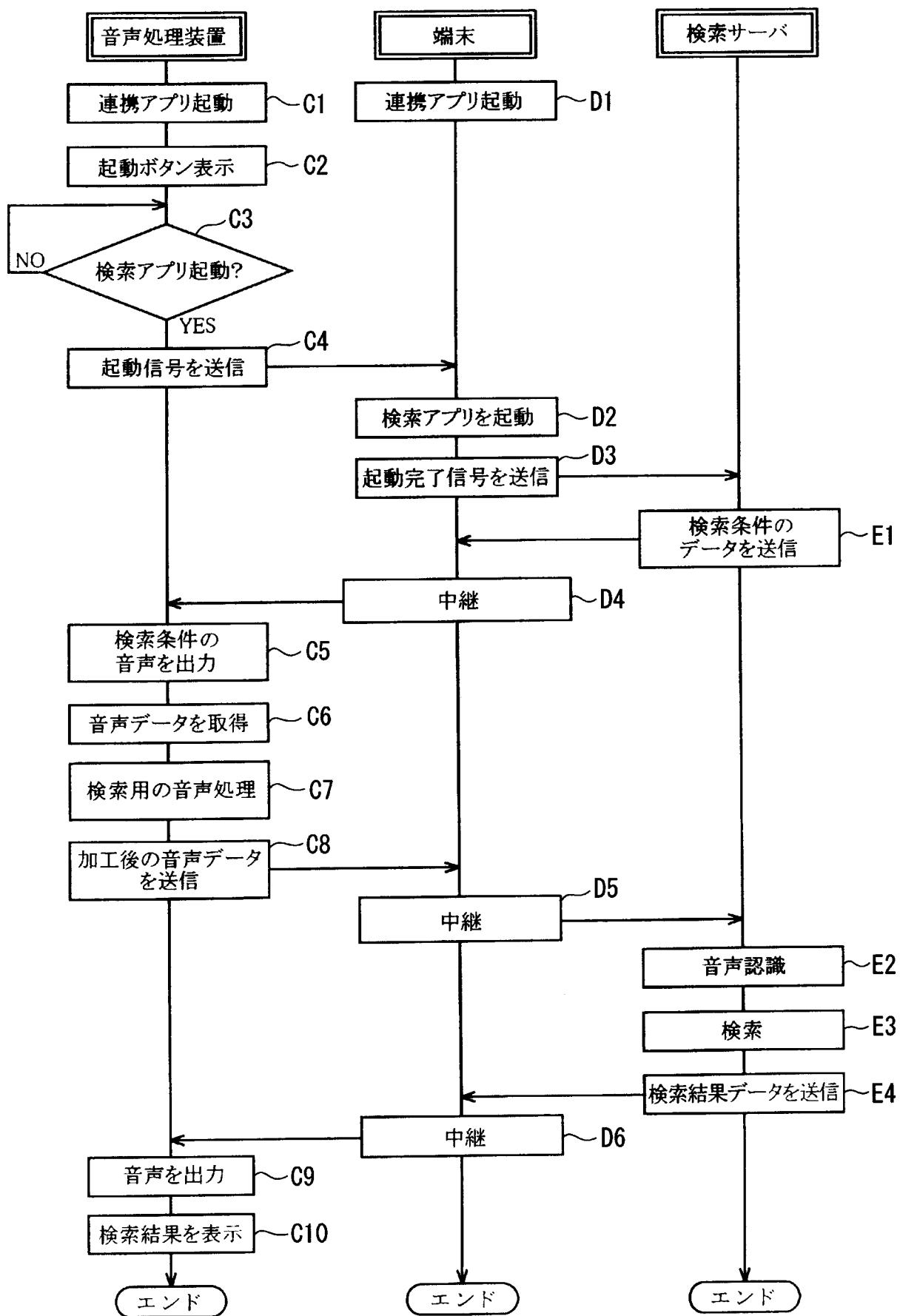
[図4]



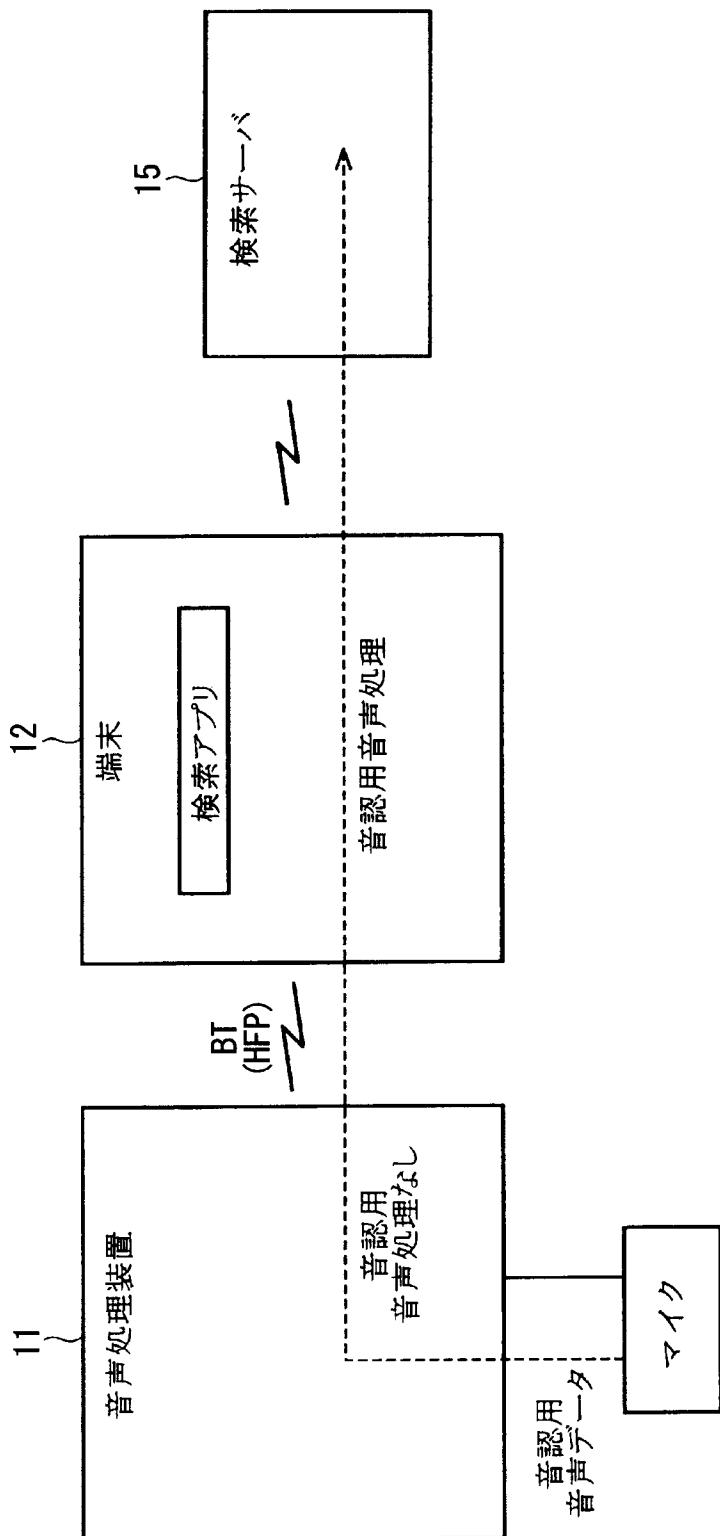
[図5]



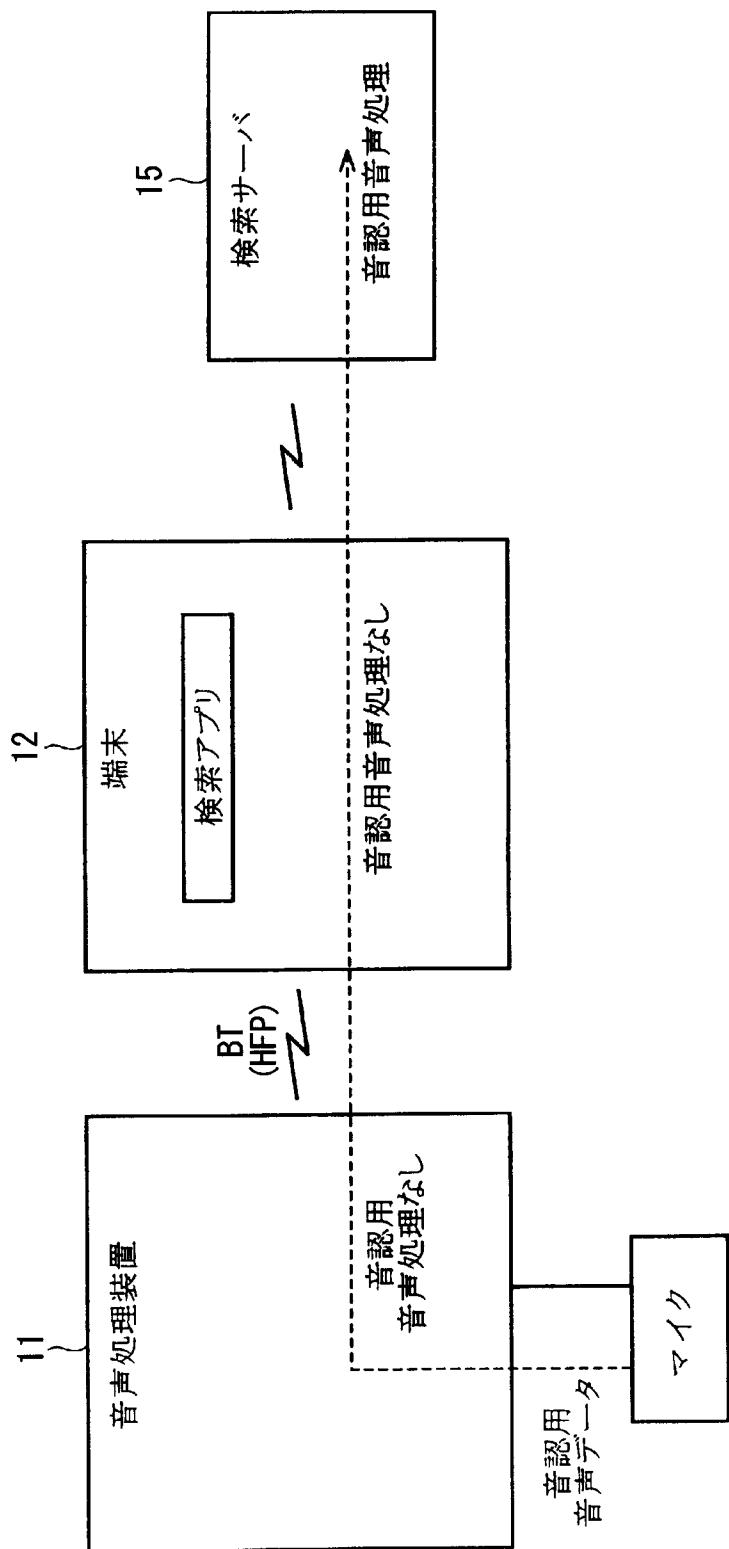
[図6]



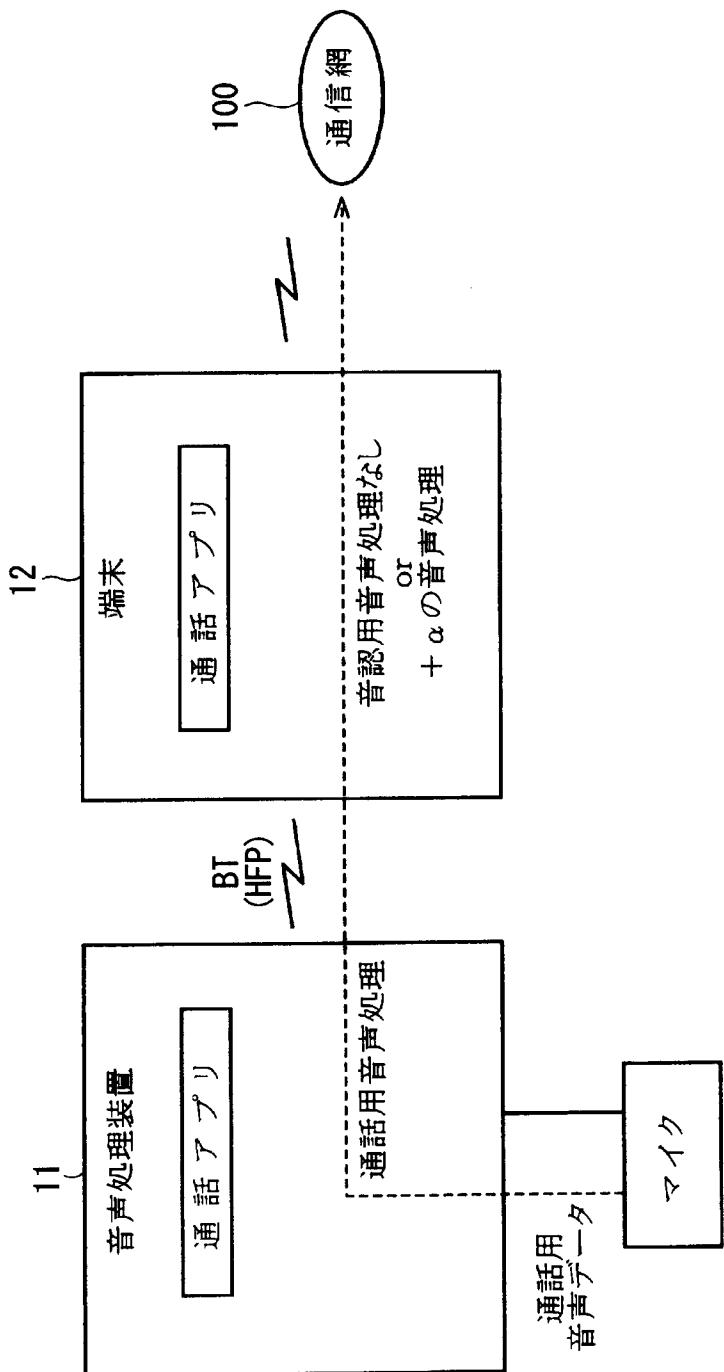
[図7]



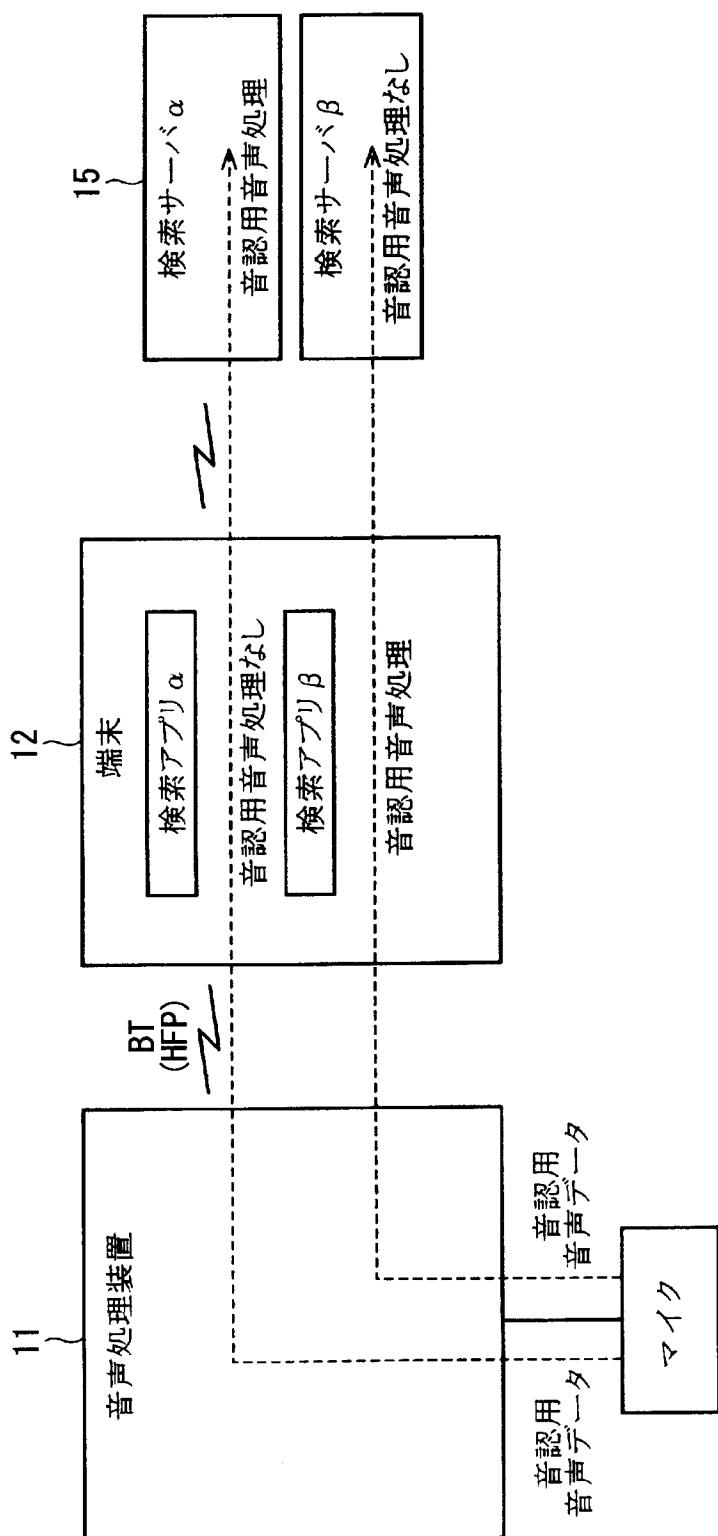
[図8]



[図9]



[図10]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/006172

A. CLASSIFICATION OF SUBJECT MATTER

H04M1/00(2006.01)i, G10L15/00(2013.01)i, H04M1/60(2006.01)i, H04M3/42 (2006.01)i, H04M11/00(2006.01)i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

H04M1/00, G10L15/00, H04M1/60, H04M3/42, H04M11/00

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

<i>Jitsuyo Shinan Koho</i>	<i>1922-1996</i>	<i>Jitsuyo Shinan Toroku Koho</i>	<i>1996-2015</i>
<i>Kokai Jitsuyo Shinan Koho</i>	<i>1971-2015</i>	<i>Toroku Jitsuyo Shinan Koho</i>	<i>1994-2015</i>

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	JP 2004-153380 A (Nissan Motor Co., Ltd.), 27 May 2004 (27.05.2004), paragraphs [0009] to [0028]; fig. 1 (Family: none)	1-11
A	JP 2006-222969 A (Bose Corp.), 24 August 2006 (24.08.2006), paragraphs [0077] to [0079]; fig. 1 & US 2006/0178169 A1 & EP 1691534 A1 & CN 1819477 A	1-11
A	JP 2004-343335 A (Denso Corp.), 02 December 2004 (02.12.2004), entire text; all drawings (Family: none)	1-11

Further documents are listed in the continuation of Box C.

See patent family annex.

* Special categories of cited documents:	
"A" document defining the general state of the art which is not considered to be of particular relevance	"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
"E" earlier application or patent but published on or after the international filing date	"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)	"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art
"O" document referring to an oral disclosure, use, exhibition or other means	"&" document member of the same patent family
"P" document published prior to the international filing date but later than the priority date claimed	

Date of the actual completion of the international search
06 March 2015 (06.03.15)

Date of mailing of the international search report
17 March 2015 (17.03.15)

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer
Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2014/006172

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 2013/0332886 A1 (APPLE INC.), 12 December 2013 (12.12.2013), paragraph [0028]; all drawings (Family: none)	1-11
P, A	WO 2014/141574 A1 (NEC Corp.), 18 September 2014 (18.09.2014), entire text; all drawings (Family: none)	1-11

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. H04M1/00(2006.01)i, G10L15/00(2013.01)i, H04M1/60(2006.01)i, H04M3/42(2006.01)i, H04M11/00(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. H04M1/00, G10L15/00, H04M1/60, H04M3/42, H04M11/00

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2015年
日本国実用新案登録公報	1996-2015年
日本国登録実用新案公報	1994-2015年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X	JP 2004-153380 A (日産自動車株式会社) 2004.05.27, 段落 [0009] – [0028], 第1図 (ファミリーなし)	1-11
A	JP 2006-222969 A (ボーズ・コーポレーション) 2006.08.24, 段落 [0077] – [0079], 第1図 & US 2006/0178169 A1 & EP 1691534 A1 & CN 1819477 A	1-11
A	JP 2004-343335 A (株式会社デンソー) 2004.12.02, 全文, 全図 (ファミリーなし)	1-11

C欄の続きにも文献が列挙されている。

パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの

「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの

「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）

「O」口頭による開示、使用、展示等に言及する文献

「P」国際出願目前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの

「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの

「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの

「&」同一パテントファミリー文献

国際調査を完了した日 06.03.2015	国際調査報告の発送日 17.03.2015
国際調査機関の名称及びあて先 日本国特許庁（ISA/JP） 郵便番号100-8915 東京都千代田区霞が関三丁目4番3号	特許庁審査官（権限のある職員） 間宮 嘉誉 電話番号 03-3581-1101 内線 3591 5Z 3454

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	US 2013/0332886 A1 (APPLE INC.) 2013. 12. 12, 段落 [0028], 全図 (ファミリーなし)	1-11
P, A	WO 2014/141574 A1 (日本電気株式会社) 2014. 09. 18, 全文, 全図 (フ アミリーなし)	1-11