

明 細 書

発明の名称：

情報処理装置、音声認識システム、及び、情報処理方法

技術分野

[0001] 本開示は、音声認識システムに用いられる情報処理装置に関する。

背景技術

[0002] VPA (Virtual Personal Assistance) と呼ばれる、音声で機器を操作するサービスが普及し始めている。特許文献1には、起動ワードを検出した場合にユーザの発話を予測する情報処理装置が開示されている。

先行技術文献

特許文献

[0003] 特許文献1：特開2017-138476号公報

発明の概要

発明が解決しようとする課題

[0004] 主な音声認識がクラウドサーバで行われる場合、情報処理装置は、例えば、起動ワードを認識することにより音声認識サーバへの音声信号の送信を開始する。

[0005] 本開示は、複数の音声認識サーバへ選択的に音声信号を送信することができる情報処理装置を提供する。

課題を解決するための手段

[0006] 本開示の一態様に係る情報処理装置は、ユーザの音声を取得する音声取得部と、前記音声取得部によって取得された音声が第一起動ワードであると認識した場合に、前記第一起動ワードに対応する音声信号を出力する第一制御部と、前記第一制御部によって出力された音声信号が前記第一起動ワードを示すと認識した場合に、前記音声取得部によって取得された音声の音声信号

を第一音声認識サーバへ送信する第一音声送信処理を開始する第二制御部とを備え、前記第一制御部は、前記第一音声送信処理中に前記音声取得部によって取得された音声信号が、前記第二制御部に第二音声送信処理を開始させるための第二起動ワードであると認識した場合に、所定の優先度に基づいて前記第二起動ワードに対応する音声信号を前記第二制御部に出力するか否かの判定を行い、前記第二音声送信処理は、前記音声取得部によって取得された音声信号を前記第一音声認識サーバと異なる第二音声認識サーバへ送信する処理である。

[0007] なお、これらの包括的または具体的な態様は、システム、方法、集積回路、コンピュータプログラムまたはコンピュータ読み取り可能なＣＤ－ＲＯＭなどの記録媒体で実現されてもよく、システム、方法、集積回路、コンピュータプログラムおよび記録媒体の任意な組み合わせで実現されてもよい。

発明の効果

[0008] 本開示の情報処理装置は、複数の音声認識サーバへ選択的に音声信号を送信することができる。

図面の簡単な説明

[0009] [図1]図1は、スマートスピーカの機能を説明するための図である。

[図2]図2は、スマートスピーカを用いた家電の制御を説明するための図である。

[図3]図3は、サービスと起動ワードとの関係を示す図である。

[図4]図4は、ユーザがスマートフォンを使ってサービスを呼び出す場合を説明するための図である。

[図5]図5は、実施の形態1に係る音声認識システムの構成を示すブロック図である。

[図6]図6は、実施の形態1に係る音声認識システムの動作のフローチャートである。

[図7]図7は、実施の形態1における、サービスと起動ワードとの関係を示す図である。

[図8]図8は、実施の形態2に係る音声認識システムの動作のフローチャートである。

[図9]図9は、実施の形態2における、サービスと起動ワードとの関係を示す図である。

[図10]図10は、実施の形態3に係る音声認識システムの構成を示すブロック図である。

[図11]図11は、実施の形態3に係る音声認識システムの動作のフローチャートである。

[図12]図12は、実施の形態4に係る音声認識システムの構成を示すブロック図である。

[図13]図13は、実施の形態4に係る音声認識システムの動作のフローチャートである。

[図14]図14は、実施の形態5に係る音声認識システムの構成を示す簡易ブロック図である。

[図15]図15は、実施の形態6に係る音声認識システムの構成を示す簡易ブロック図である。

発明を実施するための形態

[0010] (本開示の基礎となった知見)

米国アマゾン（登録商標）社が提供するアレクサ（登録商標）、米国Google（登録商標）社が提供するGoogle Assistant（登録商標）、米国マイクロソフト（登録商標）社が提供するCortana（登録商標）等のいわゆるVPA（Virtual Personal Assistance）と呼ばれる、音声で機器を操作するサービスが普及し始めている。

[0011] このようなサービスは、当初はスマートフォンまたはパーソナルコンピュータのアプリケーションとして提供されていたが、最近、スマートスピーカと呼ばれる単体の製品へのアプリケーションとしても提供されている。スマートスピーカは、ユーザが音声によって操作可能なスピーカであり、主とし

て音楽ストリーミングサービスの提供を受けるためのスピーカである。つまり、スマートスピーカの主たる機能は、ユーザの音声に応じて、オーディオストリーミングサービスから提供される音楽を再生させることである。図1は、スマートスピーカの機能を説明するための図である。

[0012] スマートスピーカ110の主たる機能は以下のように実現される。スマートスピーカ110が取得したユーザの音声の音声信号は、VPAクラウドサーバ120に送信される。VPAクラウドサーバ120は、受信した音声信号を音声認識処理してオーディオストリーミングサービスを提供するクラウドサーバ131に指令を送信し、クラウドサーバ131は、受信した指令に基づいてスマートスピーカ110に音楽を転送する。

[0013] このような仕組みは、オーディオストリーミングサービスを提供するクラウドサーバ131のみならず、サードパーティクラウドサーバ群130に含まれる他のクラウドサーバにも適用できる。サードパーティクラウドサーバ群130のサービスプロバイダによってVPAサービスのプロバイダにサードパーティクラウドI/Fが提供されれば、ユーザは、スマートフォンと同様に、スマートスピーカ110からサードパーティクラウドサーバ群130が提供するサービスを利用することが可能になる。

[0014] 例えば、ユーザは、スマートスピーカ110に「今日の天気は」と発話することにより、天気予報サービスを提供するクラウドサーバ132から天気予報情報の提供を受けることができる。具体的には、スマートスピーカ110から「晴れです」といった音声出力される。また、ユーザは、eコマースサイトを提供するクラウドサーバ133において、ユーザの商品の購入履歴が記憶されていれば、ユーザは、スマートスピーカ110に「同じ商品を買って」と発話することで同じ商品の購入を図ることができる。

[0015] このようなスマートスピーカ110を通じた多様な対話機能は、住宅などに置かれたスマートスピーカ110が備えるマイクロフォンが、ユーザの音声、テレビの音声、及び、ラジオの音声などを録音して、VPAクラウドサーバ120に転送することによって実現される。VPAクラウドサーバ12

0は、音声認識機能（ASR： automatic speech recognition）によって音声信号をテキストに変換し、自然言語処理機能（NLP： natural language processing）によってテキストを機械語化する。VPAクラウドサーバ120は、さらに、コンテキスト理解機能によって機械語を特定のコンテキストにあった意味に変換し、最後にパーソナライズド機能によって各ユーザの情報に応じた指示内容に変換する。VPAクラウドサーバ120は、このような指示内容を指令として送信することにより、サードパーティクラウドサーバ群130を呼び出すことができる。

- [0016] また、スマートスピーカ110を用いれば、VPAクラウドサーバ120のVPAによって住宅内の家電を制御することもできる。図2は、スマートスピーカ110を用いた家電の制御を説明するための図である。
- [0017] 図2に示されるように、VPAクラウドサーバ120が家電制御サーバ140に接続されれば、ユーザの住宅内のエアコン、照明、セキュリティシステム、または、テレビ等を音声によって制御することができる。
- [0018] 図2では、制御対象の機器としてエアコンa～エアコンdを含むエアコン群150が例示されている。ユーザがスマートスピーカ110に対して「外出するので設定お願い」と発話すると、この音声の音声信号は、VPAクラウドサーバ120に送信される。VPAクラウドサーバ120は、音声認識機能によって音声信号をテキスト化し、自然言語処理機能によって外出モードの設定を指示する機械語に変換する。また、VPAクラウドサーバ120は、コンテキスト理解機能によって、外出モードの設定を指示する機械語をユーザの住宅内の電気機器をオフするコマンドに変換し、パーソナライズド機能によって、ユーザ情報に基づき4台のエアコンをオフするコマンドに変換する。このように、VPAを利用した家電制御によれば、ユーザが「外出するので設定お願い」と発話することにより4台のエアコンa～dがオフされる。つまり、VPAを利用した家電制御によれば、個々のエアコンa～dをリモートコントローラの使用によってオフする以上の機能が実現される。

- [0019] ところで、ユーザは、スマートフォンによってこのようなV P Aを利用する場合は、スマートフォンが備えるボタンを操作するか、あるいは、スマートフォンに表示されたアイコンをタッチする等の操作を行うことでV P Aを起動することができる。
- [0020] 一方、スマートスピーカ1 1 0によってV P Aを利用する場合に、スマートスピーカ1 1 0が備えるボタンの操作によってV P Aを起動すると、音声によってハンズフリーで操作が可能なスマートスピーカ1 1 0の利点が得られない。このため、スマートスピーカ1 1 0によってV P Aを利用する場合、V P Aの起動は、音声によって行われることが望ましい。例えば、あらかじめ起動ワードが定められ、V P Aクラウドサーバ1 2 0によって起動ワードが認識されることでV P Aが起動される構成が考えられる。起動ワードは、ウェイクワード等と呼ばれる場合もある。
- [0021] しかしながら、このような構成では、スマートスピーカ1 1 0が備えるマイクロフォンを通じて音声信号が常にV P Aクラウドサーバ1 2 0に転送される必要がある。そうすると、ユーザは自分のプライバシーが守られないことを懸念する。したがって、ユーザのプライバシーを考慮するとこのような構成を採用することは難しい。
- [0022] そこで、起動ワードに関してはスマートスピーカ1 1 0において音声認識する構成が採用される。V P Aが起動された後にスマートスピーカ1 1 0が備えるマイクロフォンによって取得される音声の音声信号は、V P Aクラウドサーバ1 2 0に転送される。図3は、サービスと起動ワードとの関係を示す図である。
- [0023] 図3に示されるように、現状、ユーザは、スマートスピーカ1 1 0を経由して、オーディオストリーミングサービス、天気予報サービス、及び、e コマースサイトなどのV P A提供サービス（家電制御サービス以外のサービス）を呼び出すときも、家電制御サービスを呼び出すときも、V P Aサービスプロバイダが指定する起動ワードを使うことになる。V P AサービスプロバイダをB社とすると、起動ワードは、例えば、「B社」、「H e y B社」

などである。

[0024] V P Aが起動された後には、ユーザは、機能起動ワード、及び、コマンド内容をスマートスピーカ１１０に対して発話することでV P A提供サービス及び家電制御サービスを呼び出すことができる。

[0025] 例えば、上述のエアコン群１５０の製造・販売を行うA社が家電制御サービスを提供している場合には、ユーザは、機能起動ワードとして「A社」と発話し、制御対象及び制御内容を指定するためのコマンド内容として「エアコン ON」と発話する。これにより、住宅内のエアコン群１５０の制御が可能となる。

[0026] ところで、図４に示されるように、ユーザがスマートフォンを使う場合、V P A提供サービスを呼び出すために用いられるアプリケーションと、家電制御サービスを呼び出すために用いられるアプリケーションとは、通常異なる。図４は、ユーザがスマートフォンを使ってサービスを呼び出す場合を説明するための図である。このため、ユーザは、V P A提供サービスの世界と、家電制御サービスの世界とは別の世界であると認識し、２つの異なるサービス体系があると考ええる。

[0027] このように２つのサービス体系があることを認識したユーザが、A社が製造・販売するスマートスピーカ１１０のようなV P A機器を購入し使用する場面がある。V P A提供サービスを呼び出す場合に、V P Aサービスプロバイダが指定する起動ワードを使うことは、ユーザにとってスマートフォンを使ってアプリを実行することと同じような感覚である。このため、ユーザに違和感はなく自然である。

[0028] しかしながら、家電制御サービスを呼び出す場合に、V P Aサービスプロバイダが指定する起動ワードを使うと、ユーザに違和感を与えてしまう可能性がある。

[0029] 例えば、制御対象の家電機器、及び、V P A機器に製造・販売元のA社のロゴが付与されている場合、A社の家電機器を制御するために、V P Aサービスプロバイダが指定する「B社」「Hey B社」などの起動ワードを発

話することは不自然である。

[0030] また、ユーザは、家電制御サービスの提供元がA社ではなく、V P Aサービスプロバイダ（つまり、B社）であると勘違いしてしまう可能性もある。家電制御サーバ140において発生した障害により家電制御サービスが提供されなかった場合に、ユーザは、V P Aサービスプロバイダの問題であると考え、V P Aサービスプロバイダの電話相談窓口に電話する恐れがある。このように、V P Aサービスプロバイダが指定する起動ワードが使用される場合には、ユーザにとってサービスの提供者（つまり、責任者）がわかりづらくなることも課題となる。

[0031] そこで、発明者らは、V P Aサービスプロバイダが指定する起動ワード以外で音声認識システムを起動することができるスマートスピーカ110の構成を見出した。また、発明者らは、複数のV P Aクラウドサーバに接続することができるスマートスピーカ110の構成を見出した。

[0032] 以下、実施の形態について、図面を参照しながら具体的に説明する。なお、以下で説明する実施の形態は、いずれも包括的または具体的な例を示すものである。以下の実施の形態で示される数値、形状、材料、構成要素、構成要素の配置位置及び接続形態、ステップ、ステップの順序などは、一例であり、本開示を限定する主旨ではない。また、以下の実施の形態における構成要素のうち、最上位概念を示す独立請求項に記載されていない構成要素については、任意の構成要素として説明される。

[0033] なお、各図は模式図であり、必ずしも厳密に図示されたものではない。また、各図において、実質的に同一の構成に対しては同一の符号を付し、重複する説明は省略または簡略化される場合がある。

[0034] （実施の形態1）

〔構成〕

以下、実施の形態1に係る音声認識システムの構成について説明する。図5は、実施の形態1に係る音声認識システムの構成を示すブロック図である。

[0035] 図5に示されるように、実施の形態1に係る音声認識システム100は、スマートスピーカ110と、VPAクラウドサーバ120と、サードパーティクラウドサーバ群130と、家電制御サーバ140と、エアコン群150とを備える。スマートスピーカ110及びエアコン群150は、ユーザの住宅などに設置されている。

[0036] 音声認識システム100においては、スマートスピーカ110の音声取得部11によって取得された音声の音声信号はVPAクラウドサーバ120に送信され、音声認識処理される。ただし、上述した起動ワードについては、スマートスピーカ110内で音声認識処理される。以下、音声認識システム100を構成するスマートスピーカ110、VPAクラウドサーバ120、及び、家電制御サーバ140について詳細に説明する。

[0037] [スマートスピーカ]

スマートスピーカ110は、音声ユーザインターフェースとして、VPAクラウドサーバ120へ音声信号を送信する情報処理装置10を備える。情報処理装置10は、音声取得部11と、第一制御部12と、第二制御部13と、通信部14と、出音部15と、記憶部16とを備える。

[0038] 音声取得部11は、ユーザの音声を取得する。音声取得部11は、具体的には、マイクロフォンである。

[0039] 第一制御部12は、音声取得部11と第二制御部13との間に位置し、音声取得部11からユーザの音声の音声信号を取得し、取得したユーザの音声を常時モニタする。言い換えれば、第一制御部12は、音声取得部11から第二制御部13への音声の経路上に位置する。第一制御部12は、音声取得部11によって取得された音声が第一起動ワードであると認識した場合に、VPAサービスプロバイダ（つまり、B社）が指定する第二起動ワードに対応する音声信号を出力する。ここで、第一起動ワードは、例えば、ハードウェアプロバイダ（つまり、スマートスピーカ110の製造・販売元のA社）が指定するハード指定起動ワードであり、第二起動ワードとは異なる。

[0040] このように、第一制御部12は、第一起動ワードを第二起動ワードに変換

して第二制御部13に出力する。これにより、ユーザが第一起動ワードを話し、第一起動ワードが音声取得部11によって取得されても、第二制御部13は第二起動ワードが取得されたと認識する。第二起動ワードが取得されたと認識した第二制御部13は、従来通りVPAサービスプロバイダが提供するVPA SDK (Software Development Kit) に従ってVPA機能をONするための起動処理を行う。つまり、第一制御部12によれば、第二制御部13に第一起動ワードによって起動処理を行わせることができる。第一起動ワードとしてハード指定起動ワードが用いられる場合、このような第一制御部12の機能は、ハード指定起動ワード検索機能とも記載される。

[0041] なお、実施の形態1では、第一制御部12は、音声取得部11によって取得された音声の音声信号が第二起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を出力する。これにより、第二制御部13は、音声取得部11によって取得された音声が第一起動ワード及び第二起動ワードのどちらであっても起動処理を行うことができる。

[0042] 以上説明した第一制御部12は、例えば、マイクロコンピュータによって実現されるが、プロセッサによって実現されてもよい。

[0043] 第二制御部13は、第一制御部12によって出力された音声信号が第二起動ワードを示すと認識した場合に、音声取得部11によって取得された音声（より詳細には、音声取得部11によって取得された音声の音声信号）のVPAクラウドサーバ120への送信を開始するための起動処理を行う。第二制御部13は、具体的には、記憶部16に記憶されたVPA SDKを実行する。第二制御部13は、例えば、マイクロコンピュータによって実現されるが、プロセッサによって実現されてもよい。

[0044] 通信部14は、VPA機能がオンされた後、第二制御部13の制御に基づいて音声信号をVPAクラウドサーバ120の通信部121に送信する。通信部14は、具体的には、通信モジュールである。通信モジュールは、言い換えれば、通信回路である。通信部14は、有線通信を行ってもよいし、無

線通信を行ってもよい。なお、図示されないが、通信部14と通信部121との間には、ブロードバンドルータなどの中継装置、及び、インターネットなどの通信ネットワークが介在する。

[0045] 出音部15は、第二制御部13の制御に基づいて出音する。出音部15は、例えば、オーディオストリーミングサービスを提供するクラウドサーバ131から通信部14に転送される音楽を出音する。出音部15は、具体的には、スピーカである。

[0046] 記憶部16は、第一制御部12が第一起動ワード及び第二起動ワードを認識するために実行する音声認識用プログラム、第二制御部13が実行するVPA SDKなどが記憶される記憶装置である。また、記憶部16には、第一制御部12が第一起動ワードまたは第二起動ワードに対応する音声信号を出力するために読み出す音声データが記憶されてもよい。記憶部16は、音声取得部11によって取得された音声データとして一時的に記憶されるバッファメモリとして利用されてもよい。記憶部16は、具体的には、半導体メモリなどによって実現される。

[0047] [VPAクラウドサーバ]

VPAクラウドサーバ120は、スマートスピーカ110の起動後（VPA機能がオンした後）に音声取得部11によって取得された音声の音声信号を受信し、受信した音声信号を音声認識処理することにより、VPA提供サービスまたは家電制御サービスを実現する。VPAクラウドサーバ120は、音声認識サーバの一例である。VPAクラウドサーバ120は、通信部121と、VPA制御部122と、記憶部123とを備える。

[0048] 通信部121は、スマートスピーカ110の通信部14によって送信される音声信号を受信する。また、通信部121は、VPA制御部122の制御部に基づいて、サードパーティクラウドサーバ群130または家電制御サーバ140に指令を送信する。通信部121は、具体的には、通信モジュールである。通信モジュールは、言い換えれば、通信回路である。

[0049] VPA制御部122は、通信部121によって受信された音声信号を音声

認識処理し、音声認識処理の結果得られる指令を通信部 1 2 1 に送信させる。上述のように、音声認識処理においては、音声認識機能、自然言語処理機能、コンテキスト理解機能、及び、パーソナライズド機能などが用いられる。V P A 制御部 1 2 2 は、例えば、マイクロコンピュータによって実現されるが、プロセッサによって実現されてもよい。

[0050] 記憶部 1 2 3 は、V P A 制御部 1 2 2 が音声認識処理を行うために実行する音声認識用プログラムなどが記憶される記憶装置である。また、記憶部 1 2 3 には、サードパーティクラウドサーバ群 1 3 0 が提供する各種サービスに対応したプログラムであるサードパーティクラウド I / F も記憶される。記憶部 1 2 3 は、具体的には、半導体メモリなどによって実現される。

[0051] [家電制御サーバ]

家電制御サーバ 1 4 0 は、V P A クラウドサーバ 1 2 0 から指令を受信し、受信した指令に基づいてエアコン群 1 5 0 を制御することにより家電制御サービスをユーザに提供する。なお、エアコン群 1 5 0 は、制御対象の家電機器の一例であり、制御対象の家電機器は、エアコン以外の家電機器であってもよい。家電制御サーバ 1 4 0 は、通信部 1 4 1 と、家電制御部 1 4 2 と、記憶部 1 4 3 とを備える。

[0052] 通信部 1 4 1 は、V P A クラウドサーバ 1 2 0 の通信部 1 2 1 によって送信される指令を受信する。また、通信部 1 4 1 は、家電制御部 1 4 2 の制御部に基づいて、エアコン群 1 5 0 に制御信号を送信する。通信部 1 4 1 は、具体的には、通信モジュールである。通信モジュールは、言い換えれば、通信回路である。

[0053] 家電制御部 1 4 2 は、通信部 1 4 1 によって受信された指令に応じた制御信号を通信部 1 4 1 に送信させる。家電制御部 1 4 2 は、例えば、マイクロコンピュータによって実現されるが、プロセッサによって実現されてもよい。

[0054] 記憶部 1 4 3 は、家電制御部 1 4 2 がエアコン群 1 5 0 を制御するための制御プログラムなどが記憶される記憶装置である。記憶部 1 4 3 は、具体的

には、半導体メモリなどによって実現される。

[0055] [動作]

次に、音声認識システム１００の動作について説明する。図６は、音声認識システム１００の動作のフローチャートである。

[0056] まず、スマートスピーカ１１０の第二制御部１３は、初期化処理を行う（Ｓ１１）。初期化処理は、例えば、スマートスピーカ１１０への給電が開始された際に行われる。スマートスピーカ１１０は、初期化処理によって第一起動ワード及び第二起動ワードの認識が可能な待機状態となる。なお、待機状態においては、ＶＰＡクラウドサーバ１２０への音声信号の送信は停止されている。

[0057] 次に、音声取得部１１は、ユーザの音声を取得する（Ｓ１２）。第一制御部１２は、取得された音声の音声信号音声認識処理を行い、音声取得部１１によって取得された音声が第一起動ワードであるか否かを判定する（Ｓ１３）。上述のように、第一起動ワードは、例えば、ハードウェアプロバイダ（つまり、スマートスピーカ１１０の製造・販売元のＡ社）が指定するハード指定起動ワードである。第一起動ワードは、例えば、「Ａ社」である。

[0058] 第一制御部１２は、音声取得部１１によって取得された音声が第一起動ワードであると判定した場合（Ｓ１３でＹｅｓ）、記憶部１６にあらかじめ記憶された第二起動ワードに対応する音声データを読み出し（Ｓ１４）、第二起動ワードに対応する音声信号を第二制御部１３に出力する（Ｓ１６）。第二起動ワードは、ＶＰＡサービスプロバイダ（つまり、Ｂ社）が指定する起動ワードである。第二起動ワードは、例えば、「Ｂ社」、「Ｈｅｙ Ｂ社」などである。

[0059] 一方、第一制御部１２は、音声取得部１１によって取得された音声が第一起動ワードでないと判定した場合（Ｓ１３でＮｏ）、音声取得部１１によって取得された音声が第二起動ワードであるか否かを判定する（Ｓ１５）。

[0060] 第一制御部１２は、音声取得部１１によって取得された音声が第二起動ワードであると判定した場合（Ｓ１５でＹｅｓ）、第二起動ワードに対応する

音声信号を第二制御部 13 に出力する (S 16)。なお、この場合、第一制御部 12 は、例えば、ステップ S 12 において取得された音声の音声信号を一時的に記憶部 16 に記憶しておき、記憶された音声信号を第二制御部 13 に出力するが、ステップ S 14 と同様に記憶部 16 にあらかじめ記憶された第二起動ワードに対応する音声データを読み出してもよい。第一制御部 12 は、音声取得部 11 によって取得された音声の第二起動ワードでないと判定した場合 (S 15 で No)、待機状態が継続され、音声取得部 11 による音声の取得が行われる (S 11)。

[0061] なお、待機状態において、音声取得部 11 によって取得された第一起動ワード及び第二起動ワード以外の音声の音声信号は第一制御部 12 を通じて第二制御部 13 に出力されてもよいし、第一制御部 12 によって第二制御部 13 への出力が停止されてもよい。

[0062] 第二制御部 13 は、第一制御部 12 によって第二起動ワードに対応する音声信号が出力されると (S 16)、これを第二起動ワードであると認識し、起動処理を行う (S 17)。起動処理は、音声取得部 11 によって取得された音声の音声信号の VPA クラウドサーバ 120 への送信を開始するための処理であり、この結果、VPA 機能がオンされる。

[0063] 音声取得部 11 は、起動処理後も継続して音声を取得し (S 18)、第二制御部 13 は、起動処理後に音声取得部 11 によって取得された音声の音声信号を、通信部 14 を用いて VPA クラウドサーバ 120 にリアルタイム送信する (S 19)。

[0064] VPA クラウドサーバ 120 の通信部 121 は、通信部 14 から音声信号を受信し、取得した音声信号に対して音声認識処理を行う (S 20)。この結果、ステップ S 18 において取得された音声に応じて、ユーザに様々なサービスが提供される。

[0065] 以上説明したように、音声認識システム 100 においては、第一制御部 12 は、第一起動ワードを第二起動ワードに変換して第二制御部 13 に出力する。これにより、ユーザが第一起動ワードを発話し、第一起動ワードが音声

取得部 11 によって取得されても、第二制御部 13 は第二起動ワードが取得されたと認識する。第二起動ワードが取得されたと認識した第二制御部 13 は、これまで通り VPA サービスプロバイダが提供する VPA SDK に従って VPA 機能を ON するための起動処理を行い、VPA 機能をオンすることができる。言い換えれば、第一制御部 12 によれば、VPA SDK を変更することなく、第一起動ワードによって VPA 機能をオンすることができる。

[0066] また、第一制御部 12 は、音声取得部 11 によって取得された音声が第二起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を出力する。これにより、第二制御部 13 は、音声取得部 11 によって取得された音声が第一起動ワード及び第二起動ワードのどちらであっても起動処理を行い、VPA 機能をオンすることができる。図 7 は、実施の形態 1 における、サービスと起動ワードとの関係を示す図である。図 7 に示されるように、実施の形態 1 では、VPA サービスプロバイダが指定する単語（つまり、第二起動ワード）に加えて、A 社が指定する単語（つまり、第一起動ワード）が起動ワードとして追加される。

[0067] このような構成においては、ユーザは、VPA 提供サービスを受けたい場合には、第二起動ワードを発話し、家電制御サービスを受けたい場合には第一起動ワードを発話する、などの使い分けを行うこともできる。このような使い分けは、図 4 で説明したスマートフォン 160 におけるアプリケーションの使い分けに類似しているため、ユーザに違和感を与えにくいと考えられる。

[0068] なお、起動ワードは、ユーザによって指定されてもよく、例えば、図 7 に示されるように、ユーザのペットの名前等が用いられてもよい。この場合、記憶部 16 には、ユーザの音声に基づいて起動ワードをユーザ指定の単語にするための音声認識プログラムが記憶される。

[0069] （実施の形態 2）

ところで、図 7 に示されるように、ユーザは、例えば、エアコンに外出モ

ードの動作を行わせるために、第一起動ワードである「A社」と発話してスマートスピーカ110を起動した後、機能起動ワードとして「A社」と発話し、さらに、コマンド内容として「外出モード」と発話する。つまり、「A社」を2度発話する必要がある。

[0070] ここで、第一起動ワードが「A社」のような家電機器の製造・販売元の社名である場合、第一起動ワードの次に家電制御サービスを受けるための機能起動ワードが発話される可能性は高くなる。このような場合、第一制御部12は、音声取得部11によって取得された音声が第一起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を出力するだけでなく、さらに、機能起動ワードに対応する音声信号を出力してもよい。図8は、このような音声認識システム100の動作のフローチャートである。以下の実施の形態2では、主として図6のフローチャートとの差分について説明が行われ、既出事項については説明が省略される場合がある。

[0071] また、実施の形態2では、機能起動ワードは、指定ワードとも記載される。上述のようにVPAクラウドサーバ120は、起動処理の後にスマートスピーカ110（つまり、情報処理装置10）から受信した音声信号の音声認識結果に応じて他のサーバに指令を送信する。指定ワードは、この指令の送信先のサーバを指定するための単語である。

[0072] 実施の形態1と同様に、ステップS13において第一制御部12は、音声取得部11によって取得された音声が第一起動ワードであるか否かを判定する（S13）。第一制御部12は、音声取得部11によって取得された音声（S13でYes）、記憶部16にあらかじめ記憶された第二起動ワードに対応する音声データを読み出す（S14）。そして、第一制御部12は、指定ワード送信モードをオンにする（S21）。第一制御部12は、具体的には、例えば、記憶部16内で指定ワード送信モードに対応するフラグを立てる。

[0073] その後、第一制御部12は、第二起動ワードに対応する音声信号を第二制御部13に出力し（S16）、第二制御部13は、これを第二起動ワードで

あると認識し、起動処理を行う（S 1 7）。

[0074] 音声取得部 1 1 は、起動処理後も継続して音声を取得する（S 1 8）。ここで、第一制御部 1 2 は、指定ワード送信モードがオンであるか否かを判定する（S 2 2）。第一制御部 1 2 は、具体的には、記憶部 1 6 を参照することにより、指定ワード送信モードに対応するフラグが立っているか否かを判定する。

[0075] 第一制御部 1 2 は、指定ワード送信モードがオンであると判定すると（S 2 2 で Y e s）、記憶部 1 6 にあらかじめ記憶された指定ワードに対応する音声データを読み出し、指定ワードに対応する音声信号を第二制御部 1 3 に出力する。そうすると、第二制御部 1 3 は、指定ワードに対応する音声信号を通信部 1 4 に V P A クラウドサーバ 1 2 0 へ送信させる（S 2 3）。

[0076] その後、第二制御部 1 3 は、起動処理後に音声取得部 1 1 によって取得された音声の音声信号を、通信部 1 4 に送信させる（S 1 9）。第一制御部 1 2 によって指定ワード送信モードがオンでないと判定された場合には（S 2 2 で N o）、ステップ S 2 3 は、省略される。

[0077] 以上説明したように、図 8 のフローチャートにおいては、第一制御部 1 2 は、音声取得部 1 1 によって取得された音声が第一起動ワードであると認識した場合に（S 1 3 で Y e s）、第二起動ワードに対応する音声信号を出力し（S 1 6）、さらに、指定ワードに対応する音声信号を出力する（S 2 3）。指定ワードに対応する音声信号は、ユーザが指定ワードを発話しておらず、音声取得部 1 1 が指定ワードに対応する音声を取得していないにもかかわらず出力される。

[0078] これにより、ユーザは、指定ワードの発話を省略することができる。図 9 に示されるように、指定ワードとして「A 社」に対応する音声信号が送信されれば、ユーザは、「A 社」の発話を省略することができる。図 9 は、実施の形態 2 における、サービスと起動ワードとの関係を示す図である。

[0079] このような指定ワードの自動送信は、ユーザが、V P A 提供サービスを受けたい場合には第二起動ワードを発話し、家電制御サービスを受けたい場合

には第一起動ワードを発話する、などの使い分けを行う場合に有用である。

- [0080] なお、指定ワードは、例えば、第一起動ワードと同一の「A社」である。
つまり、第一制御部12は、指定ワードに対応する音声信号として、第一起動ワードに対応する音声信号を出力する。しかしながら、指定ワードと第一起動ワードは異なってもよい。例えば、図9に基づき、第一起動ワードが「A社」で、指定ワードが「テレビ」であってもよい。

- [0081] （実施の形態3）

情報処理装置10は、スマートスピーカ110以外の家電機器に実装されてもよい。例えば、情報処理装置10は、洗濯機に実装されてもよい。図10は、このような実施の形態3に係る音声認識システムの構成を示すブロック図である。

- [0082] 図10に示されるように、実施の形態3に係る音声認識システム100aは、洗濯機170と、VPAクラウドサーバ120と、家電制御サーバ140と、洗濯機群180とを備える。洗濯機170は、ユーザの住宅などに設置されており、洗濯機群180にも含まれる。

- [0083] 洗濯機170は、洗濯機能を実現するための洗濯制御部20に加えて、情報処理装置10を備える。つまり、洗濯機170は、VPAに対応した家電機器である。洗濯機170は、例えば、A社によって製造・販売される家電機器である。

- [0084] ユーザは、例えば、洗濯機170の洗濯開始ボタンを押した後に、「洗濯が終わったら教えて」と発話する。そうすると、洗濯の完了後、家電制御サーバ140からユーザのスマートフォン160に完了メッセージが送信される。

- [0085] 音声認識システム100aでは、主として洗濯機170に関連する家電制御サービスの提供を受けることが想定されている。このため、音声認識システム100aには、サードパーティクラウドサーバ群130は含まれていない。したがって、A社で製造・販売された洗濯機170の起動ワードとして、VPAサービスプロバイダが指定する第二起動ワード（例えば、「B社」

「Hey B社」) が用いられることは、非常に不自然である。

[0086] そこで、音声認識システム100aにおいては、情報処理装置10の第一制御部12は、音声取得部11によって取得された音声が第二起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を出力しない。つまり、第二起動ワードは、第一制御部12によってマスクされる。図11は、このような音声認識システム100aの動作のフローチャートである。

[0087] 図11のフローチャートにおいては、図6のフローチャートに示されるステップS15が省略されている。ステップS13において音声取得部11によって取得された音声が第一起動ワードでないと判定された場合(S13でNo)、待機状態が継続され、音声取得部11による音声の取得が行われる(S11)。

[0088] つまり、音声認識システム100aにおいては、第一起動ワードによってVPA機能をオンすることができるが、第二起動ワードによってVPA機能をオンすることはできない。言い換えれば、音声認識システム100aにおいては、VPAサービスプロバイダが指定する第二起動ワードからハードウェアプロバイダの指定する第一起動ワードへの起動ワードの変更が実現されている。

[0089] これにより、A社で製造・販売された洗濯機170の起動ワードとして、VPAサービスプロバイダが指定する第二起動ワードが用いられないため、ユーザに違和感を与えてしまうことが抑制される。

[0090] (実施の形態4)

上記実施の形態1～3では、情報処理装置10は、1つのVPAクラウドサーバ120にのみ接続可能であったが、情報処理装置10は、複数のVPAクラウドサーバに接続可能であってもよい。図12は、このような実施の形態4に係る音声認識システム100bの構成を示すブロック図である。

[0091] 図12に示されるように、実施の形態3に係る音声認識システム100bは、スマートスピーカ110bと、VPAクラウドサーバ120bと、VPAクラウドサーバ120cと、サードパーティクラウドサーバ群130と、

家電制御サーバ140と、エアコン群150とを備える。

[0092] スマートスピーカ110bは、音声ユーザインターフェースとして、VPAクラウドサーバ120b及びVPAクラウドサーバ120cへ音声信号を送信する情報処理装置10bを備える。情報処理装置10bは、音声取得部11と、第一制御部12bと、第二制御部13bと、通信部14bと、出音部15と、記憶部16bとを備える。以下、情報処理装置10bの情報処理装置10との相違点について説明する。

[0093] なお、実施の形態4では、第一起動ワード及び第二起動ワードの定義が実施の形態1～3と異なる。実施の形態4では、第一起動ワードは、スマートスピーカ110bがVPAクラウドサーバ120bに接続するための起動ワードであり、第二起動ワードは、スマートスピーカ110bがVPAクラウドサーバ120cに接続するための起動ワードである。

[0094] 第一制御部12bは、音声取得部11と第二制御部13bとの間に位置し、音声取得部11によって取得されたユーザの音声を常時モニタする。第一制御部12bは、例えば、音声取得部11によって取得された音声信号が第一起動ワードであると認識した場合に、第一起動ワードに対応する音声信号を第二制御部13bに出力する。また、第一制御部12bは、音声取得部11によって取得された音声信号が第二起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を第二制御部13bに出力する。第一制御部12bは、例えば、取得された音声信号を一時的に記憶部16bに記憶しておき、記憶された音声信号を第二制御部13bに出力する。第一制御部12bは、例えば、マイクロコンピュータによって実現されるが、プロセッサによって実現されてもよい。

[0095] 第二制御部13bは、第一制御部12bによって出力された音声信号が第一起動ワードであると認識した場合に、音声取得部11によって取得された音声信号をVPAクラウドサーバ120bへ送信する第一音声送信処理を開始する。第二制御部13bは、具体的には、記憶部16bに記憶されたVPA SDK-Bを実行する。VPA SDK-Bは、VPAクラウド

サーバ120bを用いて音声認識サービスを提供するB社によって提供される。

[0096] また、第二制御部13bは、第一制御部12bによって出力された音声信号が第二起動ワードであると認識した場合に、音声取得部11によって取得された音声信号をVPAクラウドサーバ120cへ送信する第二音声送信処理を開始する。第二制御部13bは、具体的には、記憶部16bに記憶されたVPA SDK-Cを実行する。VPA SDK-Cは、VPAクラウドサーバ120cを用いて音声認識サービスを提供するC社によって提供される。

[0097] 第二制御部13bは、例えば、マイクロコンピュータによって実現されるが、プロセッサによって実現されてもよい。

[0098] 通信部14bは、第一音声送信処理中には、第二制御部13bの制御に基づいて音声信号をVPAクラウドサーバ120bに送信し、第二音声送信処理中には、第二制御部13bの制御に基づいて音声信号をVPAクラウドサーバ120cに送信する。通信部14bは、具体的には、通信モジュールである。通信モジュールは、言い換えれば、通信回路である。

[0099] 記憶部16bは、第一制御部12bが第一起動ワード及び第二起動ワードを認識するために実行する音声認識用プログラム、第二制御部13bが実行するVPA SDK-B及びVPA SDK-Cなどが記憶される記憶装置である。また、記憶部16bには、後述する優先度情報が記憶される。記憶部16bは、音声取得部11によって取得された音声の音声信号が一時的に記憶されるバッファメモリとして利用されてもよい。記憶部16bは、具体的には、半導体メモリなどによって実現される。

[0100] VPAクラウドサーバ120bは、第一音声送信処理中に音声取得部11によって取得された音声の音声信号を受信し、受信した音声信号を音声認識処理することにより、VPA提供サービスを実現する。VPAクラウドサーバ120bは、第一音声認識サーバの一例である。VPAクラウドサーバ120bの具体的構成は、VPAクラウドサーバ120と同様である。

- [0101] V P Aクラウドサーバ120cは、第二音声送信処理中に音声取得部11によって取得された音声の音声信号を受信し、受信した音声信号を音声認識処理することにより、家電制御サービスを実現する。V P Aクラウドサーバ120cは、第二音声認識サーバの一例である。V P Aクラウドサーバ120cの具体的構成は、V P Aクラウドサーバ120と同様である。
- [0102] このように、音声認識システム100bにおいては、スマートスピーカ10b（言い換えれば、情報処理装置10b）は、2つのV P Aクラウドサーバに接続可能である。このような情報処理装置10bの機能は、起動ワード検索機能とも記載される。
- [0103] 音声認識システム100bにおいて、2つのV P Aクラウドサーバが互いに異なるサービスの提供に用いられれば、2つのV P Aクラウドサーバの役割分担が明確化される。ユーザは、V P A提供サービスを受けたい場合には第一起動ワードを発話し、家電制御サービスを受けたい場合には第二起動ワードを発話すればよい。このため、ユーザが起動ワードによって混乱してしまうことが抑制される。
- [0104] ところで、音声認識システム100bにおいては、ユーザが第一音声送信処理及び第二音声送信処理の一方が行われているときに他方に切り替えたい場合があると考えられる。このような場合、例えば、音声送信処理の切り替え用のワード、または、音声送信処理の切り替えボタン等が準備される。ここで、音声認識システム100bにおいては、第一制御部12bが所定の優先度によって第一音声送信処理及び第二音声送信処理の一方から他方への切り替えを制御してもよい。図13は、このような音声認識システム100bの動作のフローチャートである。なお、図13のフローチャートでは、第一音声送信処理中に第二音声送信処理に切り替えるか否かが判定されるが、第二音声送信処理中に第一音声送信処理に切り替えるか否かについても同様の動作となる。
- [0105] 第二制御部13bは、第一音声送信処理を行う（S31）。第一音声送信処理は、音声取得部11によって取得された音声の音声信号がV P Aクラウド

ドサーバ120bにリアルタイム送信される処理である。上述のように、第二制御部13bは、第一制御部12bによって出力された音声信号が第一起動ワードを示すと認識した場合に、第一音声送信処理を開始する。

[0106] 第一制御部12bは、第一音声送信処理中に音声取得部11によって取得された音声の音声信号が、第二制御部13bに第二音声送信処理を開始させるための第二起動ワードを示すか否かを判定する(S32)。ステップS32において音声信号が第二起動ワードを示すと判定された場合(S33でNo)、第一音声送信処理が継続される(S31)。

[0107] 一方、ステップS32において音声信号が第二起動ワードを示すと判定された場合(S33でYes)、つまり、第一制御部12bによって音声が第二起動ワードであると認識された場合、第一制御部12bは優先度に基づく判定を行う(S34)。

[0108] 優先度は、優先度情報としてあらかじめ記憶部16bに記憶され、ステップS34において、第一制御部12bは記憶部16bに記憶された優先度情報を参照する。優先度は、例えば、スマートスピーカ110bの接続先のVPAクラウドサーバごと(言い換えれば、起動ワードごと)に定められる。この場合、第一制御部12bは、VPAクラウドサーバ120bの第一優先度がVPAクラウドサーバ120cの第二優先度よりも低いかなかを判定する。

[0109] 第一制御部12bは、VPAクラウドサーバ120bの優先度がVPAクラウドサーバ120cの優先度よりも低いと判定した場合(S34でYes)、第二起動ワードに対応する音声信号を第二制御部13bに出力する(S35)。第二制御部13bは、これを第二起動ワードであると認識し、第一音声送信処理を停止して第二音声送信処理を開始する(S36)。第二音声送信処理は、音声取得部11によって取得された音声の音声信号がVPAクラウドサーバ120cにリアルタイム送信される処理である。

[0110] 一方、第一制御部12bは、VPAクラウドサーバ120bの優先度がVPAクラウドサーバ120cの優先度よりも高いと判定した場合(S34で

No)、第二起動ワードに対応する音声信号を第二制御部13bに出力しない。この結果、第一音声送信処理が継続される(S31)。

[0111] このように、第一制御部12bは、第一音声送信処理中に音声取得部11によって取得された音声が第二起動ワードであると認識した場合に、所定の優先度に基づいて第二起動ワードに対応する音声信号を第二制御部13bに出力するか否かの判定を行う。

[0112] これにより、音声認識システム100bは、所定の優先度を考慮して第一音声送信処理を第二音声送信処理に切り替えることができる。なお、音声認識システム100bにおいて図13のような優先度に基づく動作が不要である場合には、第一制御部12bは省略されてもよい。この場合、第二制御部13bは、音声取得部11から直接音声信号を取得する。

[0113] なお、優先度は、ユーザに提供されるサービスごとに定められてもよい。優先度は、例えば、オーディオストリーミングサービス、天気予報サービス、eコマースサービス、及び、家電制御サービスに対して定められてもよい。

[0114] この場合、第一制御部12bは、第一音声送信処理の結果として提供されるサービスの第一優先度が第二音声送信処理の結果として提供されるサービスの第二優先度よりも低い場合に(S34でYes)、第二起動ワードに対応する音声信号を第二制御部13bに出力する(S35)。また、第一制御部12bは、第一サービスの優先度が第二サービスの優先度よりも高い場合に、第二起動ワードに対応する音声信号を第二制御部13bに出力しない。この結果、第一音声送信処理が継続される(S31)。

[0115] 例えば、第一音声送信処理の結果としてオーディオストリーミングサービスが提供されているときに、ユーザが第二音声送信処理によって家電制御サービスを呼び出して住宅内の家電を外出モード(電源オフ)にしたい場合が考えられる。このような場合、オーディオストリーミングサービスの優先度が家電制御サービスの優先度よりも低ければ、ユーザは第二起動ワードを発話することにより、オーディオストリーミングサービスをキャンセルして家

電制御サービスの提供を受けることができる。つまり、ユーザは、音楽の再生をキャンセルして家電の電源をオフし、外出することができる。

[0116] なお、現在提供されているサービスがどのようなサービスであるかは、例えば、サービスの提供元であるサードパーティクラウドサーバ群 130 等から送信されるサービス内容を示す情報が通信部 14b によって受信されることで認識可能である。サービス内容と VPA クラウドサーバとが 1 対 1 の関係となる場合には、このような情報は不要となる。

[0117] (実施の形態 5)

上記実施の形態 1～3 で説明された起動ワードの変換機能を有する情報処理装置 10 によれば、図 14 に示されるような複数の VPA 機器が混在している音声認識システム 100c において、起動ワードを統一することができる。図 14 は、実施の形態 5 に係る音声認識システム 100c の構成を示す簡易ブロック図である。

[0118] 音声認識システム 100c は、スマートスピーカ 110 と、テレビ 190 と、洗濯機 170 と、VPA クラウドサーバ 120b と、VPA クラウドサーバ 120c と、サードパーティクラウドサーバ群 130 と、家電制御サーバ 140 と、エアコン群 150 と、家電制御サーバ 200 と、テレビ群 210 と、洗濯機群 180 とを備える。スマートスピーカ 110、テレビ 190、及び、洗濯機 170 は、A 社によって製造・販売される家電機器であり、ユーザの住宅などに設置されている。

[0119] スマートスピーカ 110 は、起動ワード変換機能を有する情報処理装置 10 を備える。この情報処理装置 10 が備える記憶部 16 には、VPA クラウドサーバ 120b を用いて音声認識サービスを提供する VPA サービスプロバイダである B 社から供給される VPA SDK-B が記憶されている。つまり、スマートスピーカ 110 は、VPA クラウドサーバ 120b に接続可能である。

[0120] テレビ 190 は、テレビ機能と、起動ワード変換機能を有する情報処理装置 10 とを備える。この情報処理装置 10 が備える記憶部 16 には、VPA

クラウドサーバ１２０ｃを用いて音声認識サービスを提供するＶＰＡサービスプロバイダであるＣ社から供給されるＶＰＡ SDK-Cが記憶されている。つまり、テレビ１９０は、ＶＰＡクラウドサーバ１２０ｃに接続可能である。

[0121] 洗濯機１７０は、洗濯機能と、起動ワード変換機能を有する情報処理装置１０とを備える。この情報処理装置１０が備える記憶部１６には、ＶＰＡクラウドサーバ１２０ｃを用いて音声認識サービスを提供するＶＰＡサービスプロバイダであるＣ社から供給されるＶＰＡ SDK-Cが記憶されている。つまり、洗濯機１７０は、ＶＰＡクラウドサーバ１２０ｃに接続可能である。

[0122] ＶＰＡクラウドサーバ１２０ｂは、Ｂ社によって管理され、サードパーティクラウドサーバ群１３０、及び、家電制御サーバ１４０に接続可能である。家電制御サーバ１４０は、エアコン群１５０を制御する機能を有し、Ａ社によって管理される。

[0123] ＶＰＡクラウドサーバ１２０ｃは、Ｃ社によって管理され、家電制御サーバ１４０、及び、家電制御サーバ２００に接続可能である。家電制御サーバ２００は、テレビ群２１０を制御する機能、及び、洗濯機群１８０を制御する機能を有し、Ａ社によって管理される。

[0124] このような音声認識システム１００ｃにおいて、仮に、スマートスピーカ１１０、テレビ１９０、及び、洗濯機１７０に起動ワード変換機能が実装されていないとすると、ユーザは、スマートスピーカ１１０のＶＰＡ機能をオンする場合には、Ｂ社によって指定される起動ワードを発話し、テレビ１９０及び洗濯機１７０のＶＰＡ機能をオンする場合には、Ｃ社によって指定される起動ワードを発話する必要がある。

[0125] これに対し、スマートスピーカ１１０、テレビ１９０、及び、洗濯機１７０に起動ワード変換機能が実装されれば、ユーザは、スマートスピーカ１１０、テレビ１９０、及び、洗濯機１７０に対する起動ワードを統一することができる。起動ワードは、Ｂ社によって指定される起動ワードに統一されて

もよいし、C社によって指定される起動ワードに統一されてもよいし、その他の起動ワードに統一されてもよい。

[0126] (実施の形態6)

上記実施の形態4で説明された複数のVPAクラウドサーバへの接続機能を有する情報処理装置10bによれば、図15に示されるような複数のVPA機器が混在している音声認識システム100dにおいて、起動ワードを整理することができる。図15は、実施の形態6に係る音声認識システム100dの構成を示す簡易ブロック図である。

[0127] 音声認識システム100dは、音声認識システム100cにおいて、スマートスピーカ110がスマートスピーカ110bに置き換えられた構成を有する。

[0128] スマートスピーカ110bは、複数のVPAクラウドサーバへの接続機能を有する情報処理装置10bを備える。この情報処理装置10bが備える記憶部16bには、VPA SDK-B、及び、VPA SDK-Cが記憶されている。つまり、スマートスピーカ110bは、VPAクラウドサーバ120b及びVPAクラウドサーバ120cに接続可能である。

[0129] このような音声認識システム100dにおいても、ユーザは、スマートスピーカ110、テレビ190、及び、洗濯機170に対する起動ワードを、B社によって指定される起動ワード、または、C社によって指定される起動ワードに統一することが可能である。

[0130] (実施の形態1～6のまとめ)

以上説明したように、情報処理装置10は、ユーザの音声を取得する音声取得部11と、音声取得部11によって取得された音声が第一起動ワードであると認識した場合に、第一起動ワードと異なる第二起動ワードに対応する音声信号を出力する第一制御部12と、第一制御部12によって出力された音声信号が第二起動ワードを示すと認識した場合に、音声取得部11によって取得された音声の音声信号のVPAクラウドサーバ120への送信を開始するための起動処理を行う第二制御部13とを備える。VPAクラウドサー

バ１２０は、音声認識サーバの一例である。

[0131] このような情報処理装置１０は、ＶＰＡサービスプロバイダが指定する第二起動ワード以外の第一起動ワードによってＶＰＡクラウドサーバ１２０への音声の送信を開始することができる。また、実施の形態５または実施の形態６のように複数のＶＰＡ機器が混在する音声認識システム１００ｃまたは音声認識システム１００ｄにおいて、起動ワードの統一を図ることができる。

[0132] また、上記実施の形態１においては、第一制御部１２は、音声取得部１１によって取得された音声が第二起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を第二制御部１３に出力する。

[0133] このような情報処理装置１０は、ＶＰＡサービスプロバイダが指定する第二起動ワード以外の第一起動ワードによってＶＰＡクラウドサーバ１２０への音声の送信を開始することができる。

[0134] また、上記実施の形態３においては、第一制御部１２は、音声取得部１１によって取得された音声が第二起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を第二制御部１３に出力しない。

[0135] このような情報処理装置１０は、第一起動ワード及び第二起動ワードのうち第一起動ワードのみによって音声認識システム１００ａを起動することができる。

[0136] また、上記実施の形態２においては、ＶＰＡクラウドサーバ１２０は、起動処理の後に情報処理装置１０から受信した音声信号の音声認識結果に応じて他のサーバに指令を送信する。第一制御部１２は、音声取得部１１によって取得された音声が第一起動ワードであると認識した場合に、第二起動ワードに対応する音声信号を出力し、さらに、指令の送信先を指定するための指定ワードに対応する音声信号を第二制御部１３に出力する。

[0137] これにより、ユーザは、指定ワードの発話を省略することができる。

[0138] また、例えば、第一制御部１２は、指定ワードに対応する音声信号として、第一起動ワードに対応する音声信号を第二制御部１３に出力する。

- [0139] これにより、ユーザは、通常は2度発話すべき第一起動ワードを1度発話することで、指令の送信先を指定することができる。
- [0140] また、音声認識システム100または音声認識システム100aは、情報処理装置10と、VPAクラウドサーバ120とを備える。
- [0141] このような音声認識システム100または音声認識システム100aは、VPAサービスプロバイダが指定する第二起動ワード以外の第一起動ワードによってVPAクラウドサーバ120への音声信号の送信を開始することができる。
- [0142] また、コンピュータによって実行される情報処理方法は、ユーザの音声を取得し、取得された音声が第一起動ワードであると認識した場合に、第一起動ワードと異なる第二起動ワードに対応する音声信号を出力し、出力された音声信号が第二起動ワードを示す認識した場合に、取得された音声の音声信号のVPAクラウドサーバへの送信を開始するための起動処理を行う。
- [0143] このような情報処理方法は、VPAサービスプロバイダが指定する第二起動ワード以外の第一起動ワードによってVPAクラウドサーバ120への音声の送信を開始することができる。
- [0144] また、上記実施の形態4においては、情報処理装置10bは、ユーザの音声を取得する音声取得部11と、音声取得部11によって取得された音声が第一起動ワードであると認識した場合に、第一起動ワードに対応する音声信号を出力する第一制御部12bと、第一制御部12bによって出力された音声信号が第一起動ワードを示すと認識した場合に、音声取得部11によって取得された音声の音声信号をVPAクラウドサーバ120bへ送信する第一音声送信処理を開始する第二制御部13bとを備える。第一制御部12bは、第一音声送信処理中に音声取得部11によって取得された音声は、第二制御部13bに第二音声送信処理を開始させるための第二起動ワードであると認識した場合に、所定の優先度に基づいて第二起動ワードに対応する音声信号を第二制御部13bに出力するか否かの判定を行い、第二音声送信処理は、音声取得部11によって取得された音声の音声信号をVPAクラウドサー

バ１２０ｂと異なるＶＰＡクラウドサーバ１２０ｃへ送信する処理である。

ＶＰＡクラウドサーバ１２０ｂは、第一音声認識サーバの一例であり、ＶＰＡクラウドサーバ１２０ｃは、第二音声認識サーバの一例である。

[0145] このような情報処理装置１０ｂは、第一起動ワード及び第二起動ワードの両方を認識し、ＶＰＡクラウドサーバ１２０ｂ、及び、ＶＰＡクラウドサーバ１２０ｃに選択的に音声を送信することができる。具体的には、情報処理装置１０ｂは、所定の優先度を考慮して第一音声送信処理を第二音声送信処理に切り替えることができる。

[0146] また、例えば、第一制御部１２ｂは、ＶＰＡクラウドサーバ１２０ｂの優先度がＶＰＡクラウドサーバ１２０ｃの優先度よりも低い場合に、上記判定に基づいて第二起動ワードに対応する音声信号を第二制御部１３ｂに出力する。

[0147] このような情報処理装置１０ｂは、ＶＰＡクラウドサーバの優先度に基づいて、第一音声送信処理を第二音声送信処理に切り替えることができる。

[0148] また、例えば、第一制御部１２ｂは、ＶＰＡクラウドサーバ１２０ｂの優先度がＶＰＡクラウドサーバ１２０ｃの優先度よりも高い場合に、上記判定に基づいて第二起動ワードに対応する音声信号を第二制御部１３ｂに出力しない。

[0149] このような情報処理装置１０ｂは、ＶＰＡクラウドサーバの優先度に基づいて、第一音声送信処理を継続することができる。

[0150] また、例えば、第一制御部１２ｂは、第一音声送信処理の結果として提供されるサービスの優先度が第二音声送信処理の結果として提供されるサービスの優先度よりも低い場合に、上記判定に基づいて第二起動ワードに対応する音声信号を第二制御部１３ｂに出力する。

[0151] このような情報処理装置１０ｂは、サービスの優先度に基づいて、第一音声送信処理を第二音声送信処理に切り替えることができる。

[0152] また、例えば、第一制御部１２ｂは、第一サービスの優先度が第二サービスの優先度よりも高い場合に、上記判定に基づいて第二起動ワードに対応する

る音声信号を第二制御部 13b に出力しない。

[0153] このような情報処理装置 10b は、サービスの優先度に基づいて、第一音声送信処理を継続することができる。

[0154] また、音声認識システム 100b は、情報処理装置 10b と、VPAクラウドサーバ 120b と、VPAクラウドサーバ 120c とを備える。

[0155] このような音声認識システム 100b は、第一起動ワード及び第二起動ワードの両方を認識し、VPAクラウドサーバ 120b、及び、VPAクラウドサーバ 120c に選択的に音声を送信することができる。具体的には、音声認識システム 100b は、所定の優先度を考慮して第一音声送信処理を第二音声送信処理に切り替えることができる。

[0156] また、コンピュータによって実行される情報処理方法は、ユーザの音声を取得し、取得された音声が第一起動ワードであると認識した場合に、第一起動ワードに対応する音声信号を出力し、出力された音声信号が第一起動ワードを示すと認識した場合に、取得された音声の音声信号をVPAクラウドサーバ 120b へ送信する第一音声送信処理を開始し、第一音声送信処理中に取得された音声、第二音声送信処理を開始するための第二起動ワードであると認識した場合に、所定の優先度に基づいて第二起動ワードに対応する音声信号を出力するか否かを決定し、第二音声送信処理は、取得された音声の音声信号をVPAクラウドサーバ 120b と異なるVPAクラウドサーバ 120c へ送信する処理である。

[0157] このような情報処理方法は、第一起動ワード及び第二起動ワードの両方を認識し、VPAクラウドサーバ 120b、及び、VPAクラウドサーバ 120c に選択的に音声信号を送信することができる。具体的には、情報処理方法は、所定の優先度を考慮して第一音声送信処理を第二音声送信処理に切り替えることができる。

[0158] (その他の実施の形態)

以上、実施の形態について説明したが、本開示は、上記実施の形態に限定されるものではない。

- [0159] 例えば、本開示の包括的または具体的な態様は、装置、システム、方法、集積回路、コンピュータプログラムまたはコンピュータ読み取り可能なＣＤ－ＲＯＭなどの記録媒体で実現されてもよく、装置、システム、方法、集積回路、コンピュータプログラムおよび記録媒体の任意な組み合わせで実現されてもよい。本開示は、上記実施の形態の情報処理方法をコンピュータに実行させるためのプログラムとして実現されてもよいし、このようなプログラムが記録されたコンピュータ読み取り可能な非一時的な記録媒体として実現されてもよい。
- [0160] また、上記実施の形態において、特定の処理部が実行する処理を別の処理部が実行してもよい。また、上記実施の形態において説明された音声認識システムの動作における複数の処理の順序は一例である。複数の処理の順序は、変更されてもよいし、複数の処理は、並行して実行されてもよい。
- [0161] また、上記実施の形態において、第一制御部及び第二制御部などの構成要素は、当該構成要素に適したソフトウェアプログラムを実行することによって実現されてもよい。各構成要素は、ＣＰＵまたはプロセッサなどのプログラム実行部が、ハードディスクまたは半導体メモリなどの記録媒体に記録されたソフトウェアプログラムを読み出して実行することによって実現されてもよい。第一制御部及び第二制御部は、単一のＣＰＵまたはプロセッサとして実現されてもよい。
- [0162] また、第一制御部及び第二制御部などの構成要素は、ハードウェアによって実現されてもよい。第一制御部及び第二制御部などの構成要素は、具体的には、回路または集積回路によって実現されてもよい。これらの回路は、全体として１つの回路を構成してもよいし、それぞれ別々の回路でもよい。また、これらの回路は、それぞれ、汎用的な回路でもよいし、専用の回路でもよい。
- [0163] その他、各実施の形態に対して当業者が思いつく各種変形を施して得られる形態、または、本開示の趣旨を逸脱しない範囲で各実施の形態における構成要素及び機能を任意に組み合わせることで実現される形態も本開示に含ま

れる。

産業上の利用可能性

[0164] 本開示の情報処理装置は、複数のV P Aクラウドサーバへ選択的に音声信号を送信することができる。本開示の情報処理装置は、接続先となるV P Aクラウドサーバの切り替えが容易であるため、V P A機器の普及、及び、V P Aクラウドサーバを用いたサービスの普及に寄与することができる。

符号の説明

- [0165] 1 0、1 0 b 情報処理装置
- 1 1 音声取得部
 - 1 2、1 2 b 第一制御部
 - 1 3、1 3 b 第二制御部
 - 1 4、1 4 b、1 2 1、1 4 1 通信部
 - 1 5 出音部
 - 1 6、1 6 b、1 2 3、1 4 3 記憶部
- 2 0 洗濯制御部
- 1 0 0、1 0 0 a、1 0 0 b、1 0 0 c、1 0 0 d 音声認識システム
 - 1 1 0、1 1 0 b スマートスピーカ
 - 1 2 0、1 2 0 b、1 2 0 c V P Aクラウドサーバ
 - 1 2 2 V P A制御部
 - 1 3 0 サードパーティクラウドサーバ群
 - 1 3 1、1 3 2、1 3 3 クラウドサーバ
 - 1 4 0、2 0 0 家電制御サーバ
 - 1 4 2 家電制御部
 - 1 5 0 エアコン群
 - 1 6 0 スマートフォン
 - 1 7 0 洗濯機
 - 1 8 0 洗濯機群
 - 1 9 0 テレビ

2 1 0 テレビ群