

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B1)

(11) 特許番号

特許第6571302号

(P6571302)

(45) 発行日 令和1年9月4日(2019.9.4)

(24) 登録日 令和1年8月16日(2019.8.16)

(51) Int.Cl. F I
G 1 O L 15/22 (2006.01) G 1 O L 15/22 4 5 3
G 1 O L 15/10 (2006.01) G 1 O L 15/10 2 O O W

請求項の数 17 (全 39 頁)

(21) 出願番号	特願2019-506379 (P2019-506379)	(73) 特許権者	506030756
(86) (22) 出願日	平成29年8月4日(2017.8.4)		ソノズ インコーポレイテッド
(86) 国際出願番号	PCT/US2017/045521		アメリカ合衆国93101カリフォルニア
(87) 国際公開番号	W02018/027142		州サンタ・バーバラ、チャバラ・ストリー
(87) 国際公開日	平成30年2月8日(2018.2.8)		ト614番
審査請求日	平成31年4月3日(2019.4.3)	(74) 代理人	100101454
(31) 優先権主張番号	15/229,868		弁理士 山田 卓二
(32) 優先日	平成28年8月5日(2016.8.5)	(74) 代理人	100132241
(33) 優先権主張国・地域又は機関	米国 (US)		弁理士 岡部 博史
		(74) 代理人	100183276
			弁理士 山田 裕三

早期審査対象出願

最終頁に続く

(54) 【発明の名称】 複数の音声サービス

(57) 【特許請求の範囲】

【請求項1】

ネットワークマイクロフォンデバイスのマイクロフォンを介して、音声入力を示す音声データを受信するステップ、

受信した音声データが、メディア再生システムに登録されている複数の音声サービスのうちの所定の音声サービスに対応する任意のウェイクワードを除外していることを決定するステップ、

上記決定に応じて、メディア再生システムに登録されている複数の音声サービスの中から、音声入力を処理するための1つ又は複数の音声サービスを識別するステップ、

ネットワークインタフェースを介して、識別された音声サービスに音声入力を処理させるステップ、

を含む、方法。

【請求項2】

音声入力は、特定の音声サービスには対応しない一般的なウェイクワードを含む、請求項1に記載の方法。

【請求項3】

一般的なウェイクワードは、ネットワーク化された特定のマイクロフォンデバイス、メディア再生システム、又はメディア再生システムの特定のゾーン、のうちの1つに対応するデフォルトのウェイクワードのうちの1つである、請求項2に記載の方法。

【請求項4】

10

20

音声入力を処理するための音声サービスを識別するステップは、複数の音声サービスのうちのデフォルトの音声サービスを、音声入力を処理するための音声サービスとして識別するステップを含む、請求項 1 に記載の方法。

【請求項 5】

音声入力を処理するための音声サービスを識別するステップは、特定の音声サービスによって処理された前の音声入力を受信されてから閾値期間内に次の音声入力を受信されたと決定された場合に、当該特定の音声サービスを次の音声入力を処理するための音声サービスとして識別するステップを含む、請求項 1 から 3 のいずれか 1 つに記載の方法。

【請求項 6】

音声入力を処理するための音声サービスを識別するステップは、特定の音声サービスによって処理された前の音声入力と同じ種類の動作を次の音声入力の対象としていると決定された場合に、当該特定の音声サービスを次の音声入力を処理するための音声サービスとして識別するステップを含む、請求項 1 から 3 のいずれか 1 つに記載の方法。

【請求項 7】

音声入力を処理するための音声サービスを識別するステップは、音声入力メディア再生コマンドを含むと決定された場合に、メディア再生コマンドを処理するように構成された特定の音声サービスを、当該音声入力を処理するための音声サービスとして識別するステップを含む、請求項 1 から 3 のいずれか 1 つに記載の方法。

【請求項 8】

音声入力を処理するための音声サービスを識別するステップは、音声入力無線照明デバイスを対象としていると決定された場合に、無線照明デバイスを対象とした音声入力を処理するように構成された特定の音声サービスを、音声入力を処理するための音声サービスとして識別するステップを含む、請求項 1 から 3 のいずれか 1 つに記載の方法。

【請求項 9】

メディア再生システムは複数のゾーンを含み、

音声入力を処理するための音声サービスを識別するステップは、

音声入力複数のゾーンのうちの特定のゾーンを対象としていることを決定するステップ、

当該決定に基づいて、当該特定のゾーンを対象とした音声入力を処理するように構成された特定の音声サービスを、音声入力を処理する音声サービスとして識別するステップ、を含む、請求項 1 から 3 のいずれか 1 つに記載の方法。

【請求項 10】

音声入力を処理するための音声サービスを識別するステップは、音声入力を提供した特定のユーザを識別するステップと、当該特定のユーザによって使用される特定の音声サービスを識別するステップとを含む、請求項 1 から 3 のいずれか 1 つに記載の方法。

【請求項 11】

第 1 の音声サービスが音声入力を処理するために現在利用不可能である場合に、第 1 の音声サービスとは異なる第 2 の音声サービスを、音声入力を処理するための音声サービスとして識別するステップをさらに含む、請求項 1 から 10 のいずれか 1 つに記載の方法。

【請求項 12】

音声入力は音声コマンドを含み、

識別された音声サービスに音声入力を処理させるステップは、識別された音声サービスの 1 つ又は複数のサーバに対して、ネットワークインタフェースを介して、音声コマンドを表すデータと、当該データの処理を命じるコマンドとを送信するステップを含む、請求項 1 から 11 のいずれか 1 つに記載の方法。

【請求項 13】

ネットワークマイクロフォンデバイスは、少なくとも第 2 のネットワークマイクロフォンデバイスを含むメディア再生システムの一部である第 1 のネットワークマイクロフォンデバイスであり、

さらに、

10

20

30

40

50

音声入力、音声コマンドと、第2のネットワークマイクロフォンデバイスが位置するゾーンの識別情報とを含むことが決定された場合に、第1のネットワークマイクロフォンデバイスにより音声入力を処理するステップと、第2のネットワークマイクロフォンデバイスに音声コマンドを実行させるステップをさらに含む、請求項1から12のいずれかに記載の方法。

【請求項14】

全ての登録済み音声サービスに問い合わせるステップと、これらの登録済み音声サービスから結果を受信するステップと、出力のために特定の1つ又は複数の結果を選択するステップとをさらに含む、請求項1から13のいずれかに記載の方法。

【請求項15】

請求項1から13のいずれかに記載の方法を実行するように構成された、ネットワークマイクロフォンデバイス。

【請求項16】

識別された音声サービスの結果を出力するように構成されたメディア再生デバイスである、請求項14に記載のネットワークマイクロフォンデバイス。

【請求項17】

請求項1から13のいずれかに記載の方法を実行するように構成されたメディア再生システム。

【発明の詳細な説明】

【関連出願の相互参照】

【0001】

本願は、2016年8月5日に出願された米国特許出願第15/229,868号に基づく優先権を主張するものであり、当該米国特許出願の内容は参照することによりその全体が本明細書に組み込まれる。

【技術分野】

【0002】

本願は、コンシューマ製品に関するものであり、特に、メディア再生に向けられた方法、システム、製品、機能、サービス、および他の要素に関するものや、それらのいくつかの態様に関する。

【背景技術】

【0003】

2003年に、ソノス・インコーポレイテッドが最初の特許出願のうちの1つである「複数のネットワークデバイス間のオーディオ再生を同期する方法」と題する特許出願をし、2005年にメディア再生システムの販売を開始するまで、アウトラウド設定におけるデジタルオーディオへのアクセスおよび試聴のオプションは制限されていた。ソノス無線HiFiシステムによって、人々は1つ又は複数のネットワーク再生デバイスを介して多くのソースから音楽を体験できるようになっている。スマートフォン、タブレット、又はコンピュータにインストールされたソフトウェアコントロールアプリケーションを通じて、ネットワーク再生デバイスを備えたいずれの部屋においても、人々は自分が望む音楽を再生することができる。また、例えばコントローラを用いて、再生デバイスを備えた各部屋に異なる歌をストリーミングすることもできるし、同期再生のために複数の部屋をグループ化することもできるし、全ての部屋において同期して同じ歌を聞くこともできる。

【0004】

これまでのデジタルメディアに対する関心の高まりを考えると、試聴体験を更に向上させることができるコンシューマアクセス可能な技術を更に開放出すことにニーズがある。

【0005】

本明細書で開示されている技術の特徴、態様、および利点は、以下の記載、添付の特許請求の範囲、および添付の図面を参照するとより理解しやすい。

【図面の簡単な説明】

【0006】

【図1】ある実施形態で実施可能な例示的なメディア再生システムの構成を示す図

【図2】例示的な再生デバイスの機能ブロック図

【図3】例示的な制御デバイスの機能ブロック図

【図4】例示的なコントローラインタフェースを示す図

【図5】例示的な複数のネットワークデバイスを示す図

【図6】例示的なネットワークマイクロフォンデバイスの機能ブロック図

【図7】例示的な実施形態における音声サービスに音声入力进行处理させる技術を示す図

【図8】例示的な実施形態における音声サービスに音声入力进行处理させる別の技術を示す図

10

【図9】例示的な実施形態における音声サービスを登録する技術を示す図

【発明を実施するための形態】

【0007】

図面は、いくつかの例示的な実施形態を説明することを目的としているが、本発明が、図面に示した配置および手段に限定されるものではないことは理解される。

【0008】

I. 概要

【0009】

ネットワークマイクロフォンデバイス (NMD、Networked microphone devices) を使用することで、音声コントロールを使用しながら家庭内をコントロールすることができる。NMDは、マイクロフォンを介して音声入力を受信することができる、例えばソノス (登録商標) の再生デバイス、サーバ、又はシステムであってもよい。またNMDは、マイクロフォンを介して音声入力を受信することができる、ソノス (登録商標) の再生デバイス、サーバ、又はシステム以外のデバイスであってもよい (例えば、アマゾン (登録商標) 社のECHO (登録商標)、アップル (登録商標) 社のIPHONE S (登録商標))。 「デフォルト再生デバイスの指定」と題する米国特許出願第15/098, 867号が本明細書に参照として援用されており、それは音声起動型家庭用アーキテクチャの例を提供している。音声コントロールは、再生デバイス、無線照明デバイス、サーモスタット、ドアロック、ホームオートメーション、および他の例を含む、「スマート」ホーム機能を有する様々なデバイスにとって有益であり得る。

20

30

【0010】

ある実施形態では、NMDによって検出される音声入力は、処理を目的として音声サービスへと送信されている。再生デバイスなどのNMDは、この音声サービスに対するマイクロフォンインタフェース又はスピーカインタフェースとして機能してもよい。音声入力はNMDのマイクロフォンによって検出され、次いで処理を目的として特定の音声サービスへと送信される。その後音声サービスは、コマンド又は音声入力の他の結果を返すことができる。

【0011】

特定の音声サービスが、場合によっては設定手順の間に、メディア再生システムに対して選択されてもよい。ユーザは、自身のスマートフォンやタブレットコンピュータ (又はその他のパーソナル電子装置) で使用中の音声サービスと同じサービスを選択し得、おそらくこれは、ユーザがその音声サービスに慣れ親しんでいたり、あるいは自身のスマートフォンで使用中の音声コントロールと同じ制御を再生デバイスでも使用して、同様の体験を得たいと望んでいる可能性があるからである。特定の音声サービスがユーザのスマートフォン上に設定されている場合、このスマートフォンは、その音声サービス用の設定情報 (例えば、ユーザ認証情報) をNMDに送信して、その音声サービスの自動設定がNMDで容易に行われるようにすることができる。

40

【0012】

場合によっては、複数の音声サービスをNMD、又はNMDのシステム (例えば、複数の再生デバイスを備えるメディア再生システム) に対して設定してもよい。設定手順の間

50

に、1つ又は複数のサービスが設定される場合がある。追加の音声サービスを、後で本システムに設定してもよい。したがって、本明細書に記載されるNMDは複数の音声サービスとのインタフェースとして機能してもよく、おそらくそれぞれの音声サービスと対話するために、それぞれの音声サービスからのNMDを備える必要性を軽減できる可能性がある。更にまた、NMDは、家庭内に存在するサービス固有のNMDと協調して動作して、所定の音声コマンドを処理することができる。

【0013】

NMDに対して2つ以上の音声サービスが設定されている場合、特定の音声サービスに対応する起動ワークを発することによって、この特定の音声サービスを起動することができる。例えば、アマゾン（登録商標）社のサービスに質問する場合、ユーザは「A l e x a」というウェイクワードを発し、これに続けて音声入力を行う可能性がある。その他の例としては、グーグル（登録商標）社のサービスに質問する場合の「オッケー、G o o g l e」や、アップル（登録商標）社のサービスに質問する場合の「hey、S i r i」が挙げられる。

10

【0014】

あるいは、所定の音声入力に対していずれのウェイクワードも使用されていない場合、NMDはその音声入力を処理するための音声サービスを識別することができる。場合によっては、NMDがデフォルトの音声サービスを識別することができる。あるいは、NMDは、コンテキストに基づいて特定の音声サービスを識別してもよい。例えば、NMDは、最近質問が行われた音声サービスを、ユーザが再び同じ音声サービスを使用することを望むという仮定に基づいて使用してもよい。他の例も可能である。

20

【0015】

上述したように、一般的ウェイクワードを使用してNMDへの音声入力を示すことができる場合がある。場合によっては、これは、任意の特定の音声サービスと関連付けられたウェイクワードではなく（例えば、NMDがソノス（登録商標）の再生デバイスである場合、「hey、S o n o s」）、製造元固有のウェイクワードであってもよい。そのようなウェイクワードを受信すると、NMDは当該要求を処理するために特定の音声サービスを識別することができる。例えば、ウェイクワードに続く音声入力が特定の種類のコマンド（例えば、音楽の再生）に関連している場合、その音声入力は、その種類のコマンドと関連付けられた特定の音声サービス（例えば、音声コマンド機能を有する音楽ストリーミングサービス）に送信されてもよい。

30

【0016】

NMDは、場合によっては音声入力を複数の音声サービスに送信してもよく、その結果として、質問が行われた音声サービスからそれぞれの結果を得てもよい。NMDはこれらの結果を評価し、「最良の」結果（例えば、目的の行動に最大限一致している結果）をもって応答することができる。例えば、音声入力が「hey、S o n o s、テイラー・スウィフトの曲を再生して」であった場合、第1音声サービスは「テイラー・スウィフト」に関する検索結果で応答する一方、第2音声サービスは、テイラー・スウィフトというアーティストによるオーディオトラックの識別子で応答する可能性がある。その場合、NMDは、第2音声サービスからのテイラー・スウィフトによるオーディオトラックの識別子を使用して、音声入力に従ってテイラー・スウィフトの曲を再生することができる。

40

【0017】

上述したように、例示的な技術は音声サービスに関連している。例示的な実施形態は、NMDがマイクロフォンを介して、音声入力を示す音声データを受信するステップを含んでもよい。NMDは、メディア再生システムに登録されている複数の音声サービスの中から、当該音声入力を処理するための音声サービスを識別し、識別したその音声サービスに当該音声入力を処理させてもよい。

【0018】

別の例示的な実施形態は、NMDが、メディア再生システムに1つ又は複数の音声サービスを登録するように命じるコマンドを示す入力データを受信するステップを含んでもよ

50

い。NMDは、NMDに登録されている音声サービスを検出することができる。NMDは、NMDに登録されている音声サービスがメディア再生システムに登録されるようにしてもよい。

【0019】

第3の例示的な実施形態は、NMDがマイクロフォンを介して、音声入力を示す音声データを受信するステップを含んでもよい。NMDは、受信した音声データの一部が、特定の音声サービスに対応していない一般的ウェイクワードを示していると決定してもよい。NMDは、複数の音声サービスに音声入力の処理を実行させてもよい。NMDは、複数の音声サービスのうちの所定の1つから得た結果を出力してもよい。

【0020】

これらの例示的な実施形態のそれぞれは、方法、本実施形態を実行するように構成されたデバイス、本実施形態を実行するように構成されたデバイスのシステム、又は1つ又は複数のプロセッサによって実行されて本実施形態を実行する命令を含む、非一時的なコンピュータ読取り可能媒体その他として具現化されてもよい。本開示が、本明細書に記載の例示的な特徴を組み合わせたものを含む、他の多数の実施形態を含むことは、当業者には理解されよう。また、ある技術を例示するために、所定のデバイスによって実行されるものとして記載されている任意の例示的な動作が、本明細書に記載のデバイスを含む、任意の適切なデバイスによって実行されてもよい。更にまた、任意のデバイスによって、別のデバイスに本明細書に記載の動作のうちのいずれかを実行させてもよい。

【0021】

本明細書に記載されるいくつかの例は、「ユーザ」および／又は他のエンティティなどの所定のアクターによって実行される機能を参照するが、これは説明の目的に過ぎない。特許請求の範囲自体の言葉によって明示的に要求されない限り、そのような例示的なアクターによるアクションが要求されるものと解釈されるべきではない。

【0022】

II. 動作環境の例

図1は、本明細書で開示されている1つ又は複数の実施形態で実施可能又は実装可能なメディア再生システム100の例示的な構成を示す。図示されるように、メディア再生システム100は、複数の部屋および空間、例えば、主寝室、オフィス、ダイニングルーム、およびリビングルームを有する例示的なホーム環境と関連付けられている。図1の例に示されるように、メディア再生システム100は、再生デバイス102、104、106、108、110、112、114、116、118、120、122、124、制御デバイス126および128、有線又は無線のネットワークルータ130を含む。

【0023】

更に、例示的なメディア再生システム100の異なる構成要素、および異なる構成要素がどのように作用してユーザにメディア体験を提供するかについての説明は、以下のセクションで述べられている。本明細書における説明は、概してメディア再生システム100を参照しているが、本明細書で述べられている技術は、図1に示されるホーム環境の用途に限定されるものではない。例えば、本明細書で述べられている技術は、マルチゾーンオーディオが望まれる環境、例えば、レストラン、モール、又は空港のような商業的環境、スポーツ用多目的車（SUV）、バス又は車のような車両、船、若しくはボード、飛行機などの環境において有益である。

【0024】

a. 例示的なゾンプレーヤー

図2は、図1のメディア再生システム100の再生デバイス102～124の1つ又は複数構成する例示的な再生デバイス200の機能ブロック図を示す。再生デバイス200は、プロセッサ202、ソフトウェアコンポーネント204、メモリ206、オーディオ処理コンポーネント208、オーディオアンプ210、スピーカ212、およびネットワークインタフェース214を含んでもよい。ネットワークインタフェース214は、無線インタフェース216、有線インタフェース218およびマイクロフォン220を含む

。ある場合では、再生デバイス200は、スピーカ212を含まないが、再生デバイス200を外部スピーカに接続するためのスピーカインタフェースを含んでいてもよい。別の場合では、再生デバイス200は、スピーカ212もオーディオアンプ210も含まないが、再生デバイス200を外部オーディオアンプ又はオーディオビジュアルレシーバーに接続するためのオーディオインタフェースを含んでもよい。

【0025】

ある例では、プロセッサ202は、メモリ206に記憶された命令に基づいて、入力データを処理するように構成されたクロック駆動コンピュータコンポーネントであってもよい。メモリ206は、プロセッサ202によって実行可能な命令を記憶するように構成された非一時的なコンピュータ読取り可能記録媒体であってもよい。例えば、メモリ206は、ある機能を実行するためにプロセッサ202によって実行可能なソフトウェアコンポーネント204の1つ又は複数をロードすることができるデータストレージであってもよい。ある例では、機能は、再生デバイス200がオーディオソース又は別の再生デバイスからオーディオデータを読み出すステップを含んでもよい。別の例では、機能は、再生デバイス200がネットワーク上の別のデバイス又は再生デバイスにオーディオデータを送信するステップを含んでもよい。更に別の例では、機能は、マルチチャネルオーディオ環境を作るために、再生デバイス200と1つ又は複数の再生デバイスとをペアリングするステップを含んでもよい。

【0026】

ある機能は、再生デバイス200が、1つ又は複数の他の再生デバイスと、オーディオコンテンツの再生を同期するステップを含む。再生を同期している間、再生デバイス200によるオーディオコンテンツの再生と1つ又は複数の他の再生デバイスによる再生との間の遅延を、リスナーが気づかないことが好ましい。「複数の独立クロックデジタルデータ処理デバイス間の動作を同期するシステムおよび方法」と題する米国特許第8,234,395号が本明細書に参照として援用されており、それは再生デバイス間のオーディオ再生を同期することが述べられたより詳細な例を提供している。

【0027】

更に、メモリ206は、データを記憶するように構成されていてもよい。データは、例えば、1つ又は複数のゾーンおよび／又はゾーングループに一部として含まれる再生デバイス200などの再生デバイス200、再生デバイス200によりアクセス可能なオーディオソース、又は再生デバイス200（又は他の再生デバイス）と関連付け可能な再生キュー、に関連付けられている。データは、定期的に更新され、再生デバイス200の状態を示す1つ又は複数の状態変数として記憶されてもよい。また、メモリ206は、メディアシステムの他のデバイスの状態と関連付けられたデータを含んでもよく、デバイス間で随時共有することによって、1つ又は複数のデバイスが、システムと関連するほぼ直近のデータを有することができる。他の実施形態も可能である。

【0028】

オーディオ処理コンポーネント208は、1つ又は複数のデジタルアナログ変換器（DAC）、オーディオ処理コンポーネント、オーディオ強化コンポーネント、又はデジタルシグナルプロセッサ（DSP）等を含んでいてもよい。ある実施形態では、1つ又は複数のオーディオ処理コンポーネント208は、プロセッサ202のサブコンポーネントであってもよい。ある実施形態では、オーディオコンテンツが、オーディオ処理コンポーネント208によって処理および／又は意図的に変更されることによって、オーディオ信号を生成してもよい。生成されたオーディオ信号は、オーディオアンプ210に送信され、増幅され、スピーカ212を通じて再生される。特に、オーディオアンプ210は、1つ又は複数のスピーカ212を駆動できるレベルまでオーディオ信号を増幅するように構成されたデバイスを含んでもよい。スピーカ212は、独立した変換器（例えば、「ドライバ」）又は1つ又は複数のドライバを内包する筐体を含む完全なスピーカシステムを備えてもよい。スピーカ212に備えられたあるドライバは、例えば、サブウーファ（例えば、低周波用）、ミドルレンジドライバ（例えば、中間周波用）、および／又はツイータ

10

20

30

40

50

ー（高周波用）を含んでもよい。ある場合では、1つ又は複数のスピーカ212のそれぞれの変換器は、オーディオアンプ210の対応する個々のオーディオアンプによって駆動されてもよい。再生デバイス200で再生するアナログ信号を生成することに加えて、オーディオ処理コンポーネント208は、オーディオコンテンツを処理し、そのオーディオコンテンツを1つ又は複数の他の再生デバイスに再生させるために送信する。

【0029】

再生デバイス200によって処理および／又は再生されるオーディオコンテンツは、外部ソース、例えば、オーディオライン—in入力接続（例えば、オートディテクティング3.5mmオーディオライン—in接続）又はネットワークインタフェース214を介して、受信されてもよい。

10

【0030】

ネットワークインタフェース214は、データネットワーク上で再生デバイス200と1つ又は複数の他のデバイスとの間のデータフローを可能にするように構成されてもよい。このように、再生デバイス200は、再生デバイスと通信する1つ又は複数の他の再生デバイス、ローカルエリアネットワーク内のネットワークデバイス、又は例えば、インターネット等のワイドエリアネットワーク上のオーディオコンテンツソースから、データネットワークを介してオーディオコンテンツを受信するように構成されていてもよい。ある例では、再生デバイス200によって送信および受信されたオーディオコンテンツおよび他の信号は、インターネットプロトコル（IP）に基づくソースアドレスおよびIPに基づく宛先アドレスを含むデジタルパケットの形で送信されてもよい。そのような場合、ネットワークインタフェース214は、デジタルパケットデータを解析することによって、再生デバイス200宛てのデータを、再生デバイス200によって適切に受信して処理することができる。

20

【0031】

図示されるように、ネットワークインタフェース214は、無線インタフェース216と有線インタフェース218とを含んでもよい。無線インタフェース216は、再生デバイス200用のネットワークインタフェース機能を提供し、通信プロトコル（例えば、無線規格IEEE802.11a、802.11b、802.11g、802.11n、802.11ac、802.15、4Gモバイル通信基準などを含む無線基準（規格）のいずれか）に基づいて、他のデバイス（例えば、再生デバイス200に関連付けられたデータネットワーク内の他の再生デバイス、スピーカ、レシーバー、ネットワークデバイス、制御デバイス）と無線通信してもよい。有線インタフェース218は、再生デバイス200用のネットワークインタフェース機能を提供し、通信プロトコル（例えば、IEEE802.3）に基づいて他のデバイスとの有線接続を介して通信してもよい。図2に示されるネットワークインタフェース214は、無線インタフェース216と有線インタフェース218との両方を含んでいるが、ネットワークインタフェース214は、ある実施形態において、無線インタフェースのみか、又は有線インタフェースのみを含んでもよい。

30

【0032】

マイクロフォン220は、再生デバイス200の環境内の音を検出するように構成してもよい。マイクロフォンは例えば、再生デバイスの筐体の外壁に取り付けてもよい。マイクロフォンは、コンデンサーマイクロフォン、エレクトレットコンデンサーマイクロフォン、またはダイナミックマイクロフォンなど、現在知られているか又は後で開発される任意の種類のマイクロフォンであってもよい。マイクロフォンは、スピーカ220の一部の周波数範囲に対して高感度であってもよい。スピーカ220のうちの1つ又は複数が、マイクロフォン220とは逆に動作してもよい。いくつかの態様では、再生デバイス200は、マイクロフォン220を含まなくてもよい。

40

【0033】

ある例では、再生デバイス200と他の再生デバイスとは、ペアにされて、オーディオコンテンツの2つの別々のオーディオコンポーネントを再生してもよい。例えば、再生デ

50

バイス200は、左チャンネルオーディオコンポーネントを再生するように構成される一方、他の再生デバイスは、右チャンネルオーディオコンポーネントを再生するように構成されてもよい。これにより、オーディオコンテンツのステレオ効果を生成するか、又は強化することができる。ペアにされた再生デバイス（「結合再生デバイス」とも言う）は、更に、他の再生デバイスと同期してオーディオコンテンツを再生してもよい。

【0034】

別の例では、再生デバイス200は、1つ又は複数の他の再生デバイスと音響的に統合され、単一の統合された再生デバイス（統合再生デバイス）を形成してもよい。統合再生デバイスは、統合されていない再生デバイス又はペアにされた再生デバイスと比べて、サウンドの処理や再現を異なるように構成することができる。なぜならば、統合再生デバイスは、オーディオコンテンツを再生するスピーカ追加することができるからである。例えば、再生デバイス200が、低周波レンジのオーディオコンテンツを再生するように設計されている場合（例えば、サブウーファー）、再生デバイス200は、全周波数レンジのオーディオコンテンツを再生するように設計された再生デバイスと統合されてもよい。この場合、全周波数レンジの再生デバイスは、低周波の再生デバイス200と統合されたとき、オーディオコンテンツの中高周波コンポーネントのみを再生するように構成されてもよい。一方で低周波レンジの再生デバイス200は、オーディオコンテンツの低周波コンポーネントを再生する。更に、統合再生デバイスは、単一の再生デバイス、又は更に他の統合再生デバイスとペアにされてもよい。

【0035】

例として、現在、ソノス・インコーポレイテッドは、「PLAY:1」、「PLAY:3」、「PLAY:5」、「PLAYBAR」、「CONNECT:AMP」、「CONNECT」、および「SUB」を含む再生デバイスを販売提供している。他の過去、現在、および／又は将来のいずれの再生デバイスにおいても、追加的に又は代替的に本明細書で開示された実施例の再生デバイスに実装して使用することができる。更に、再生デバイスは、図2に示された特定の例又は提供されるソノス製品に限定されないことは理解される。例えば、再生デバイスは、有線又は無線のヘッドホンを含んでもよい。別の例では、再生デバイスは、パーソナルモバイルメディア再生デバイス用のドッキングステーションを含むか、又は、それらと対話してもよい。更に別の例では、再生デバイスは、別のデバイス又はコンポーネント、例えば、テレビ、照明器具、又は屋内又は屋外で使用するためのいくつかの他のデバイスと一体化されてもよい。

【0036】

b. 例示的な再生ゾーン構成

図1のメディア再生システムに戻って、環境は、1つ又は複数の再生ゾーンを有しており、それぞれの再生ゾーンは1つ又は複数の再生デバイスを含んでいる。メディア再生システム100は、1つ又は複数の再生ゾーンで形成されており、後で1つ又は複数のゾーンが追加又は削除して、図1に示す例示的な構成としてもよい。それぞれのゾーンは、異なる部屋又は空間、例えば、オフィス、浴室、主寝室、寝室、キッチン、ダイニングルーム、リビングルーム、および／又はバルコニーに基づく名前が与えられてもよい。ある場合では、単一の再生ゾーンは複数の部屋又は空間を含んでいてもよい。別の場合では、単一の部屋又は空間は、複数の再生ゾーンを含んでいてもよい。

【0037】

図1に示されるように、バルコニー、ダイニングルーム、キッチン、浴室、オフィス、および寝室のゾーンのそれぞれは、1つの再生デバイスを有する一方、リビングルームおよび主寝室のゾーンのそれぞれは、複数の再生デバイスを有する。リビングルームゾーンは、再生デバイス104、106、108、および110が、別々の再生デバイスとしてか、1つ又は複数の結合再生デバイスとしてか、1つ又は複数の統合再生デバイスとしてか、又はこれらのいずれかの組み合わせで、オーディオコンテンツを同期して再生するように構成されてもよい。同様に、主寝室の場合では、再生デバイス122および124が、別々の再生デバイスとしてか、結合再生デバイスとしてか、又は統合再生デバイスとし

て、オーディオコンテンツを同期して再生するように構成されてもよい。

【0038】

ある例では、図1の環境における1つ又は複数の再生ゾーンは、それぞれ異なるオーディオコンテンツを再生している。例えば、ユーザは、バルコニーゾーンで Grillしながら、再生デバイス102によって再生されるヒップホップ音楽を聞くことができる。一方、別のユーザは、キッチンゾーンで食事を準備しながら、再生デバイス114によって再生されるクラシック音楽を聞くことができる。別の例では、再生ゾーンは、同じオーディオコンテンツを別の再生ゾーンと同期して再生してもよい。例えば、ユーザがオフィスゾーンにいる場合、オフィスゾーンの再生デバイス118が、バルコニーの再生デバイス102で再生されている音楽と同じ音楽を再生してもよい。そのような場合、再生デバイス102および118は、ロック音楽を同期して再生しているため、ユーザは、異なる再生ゾーン間を移動してもアウトラウドで再生されるオーディオコンテンツをシームレス（又は少なくともほぼシームレス）に楽しむことができる。再生ゾーン間の同期は、前述の米国特許第8,234,395号で述べられているような再生デバイス間の同期と同様の方法で行ってもよい。

10

【0039】

上述したように、メディア再生システム100のゾーン構成は、動的に変更してもよく、ある実施形態では、メディア再生システム100は、複数の構成をサポートする。例えば、ユーザが1つ又は複数の再生デバイスを、物理的にゾーンに移動させるか、又はゾーンから移動させる場合、メディア再生システム100は変更に対応するように再構成されてもよい。例えば、ユーザが再生デバイス102をバルコニーゾーンからオフィスゾーンに物理的に移動させる場合、オフィスゾーンは、再生デバイス118と再生デバイス102との両方を含んでもよい。必要に応じて、制御デバイス、例えば制御デバイス126と128とを介して、再生デバイス102が、ペアにされるか、又はオフィスゾーンにグループ化されるか、および／又はリネームされてもよい。一方、1つ又は複数の再生デバイスが、再生ゾーンを未だ設定していないホーム環境において、ある領域に移動させられた場合、新しい再生ゾーンがその領域に形成されてもよい。

20

【0040】

更に、メディア再生システム100の異なる再生ゾーンは、動的にゾーングループに組み合わされてもよいし、又は別々の再生ゾーンに分割されてもよい。例えば、ダイニングルームゾーンとキッチンゾーン114とがディナーパーティ用のゾーングループに組み合わされることによって、再生デバイス112と114とがオーディオコンテンツを同期して再生することができる。一方、あるユーザがテレビを見たい一方、他のユーザがリビングルーム空間の音楽を聞きたい場合、リビングルームゾーンが、再生デバイス104を含むテレビゾーンと、再生デバイス106、108および110を含むリスニングゾーンと、に分けられてもよい。

30

【0041】

c. 例示的な制御デバイス

図3は、メディア再生システム100の制御デバイス126と128とうちの一方又は両方を構成する例示的な制御デバイス300の機能ブロック図を示す。図示されるように、制御デバイス300は、プロセッサ302、メモリ304、ネットワークインタフェース306、ユーザインタフェース308、マイクロフォン310、およびソフトウェアコンポーネント312を含んでもよい。ある例では、制御デバイス300は、メディア再生システム100専用の制御デバイスであってもよい。別の例では、制御デバイス300は、メディア再生システムコントローラアプリケーションソフトウェアをインストールされたネットワークデバイス、例えば、iPhone（登録商標）、iPad（登録商標）、又は任意の他のスマートフォン、タブレットあるいはネットワークデバイス（例えば、PC又はMac（登録商標）などのネットワークコンピュータ）であってもよい。

40

【0042】

プロセッサ302は、メディア再生システム100のユーザアクセス、コントロール、

50

および構成を可能にすることに関する機能を実行するように構成されてもよい。メモリ 304 は、プロセッサ 302 によって実行されて機能を発揮する 1 つ又は複数のソフトウェアコンポーネントを搭載可能なデータストレージであってもよい。また、メモリ 304 は、メディア再生システムコントローラアプリケーションソフトウェアと、メディア再生システム 100 とユーザとに関連付けられた他のデータを記憶するように構成されていてもよい。

【0043】

ある例では、ネットワークインタフェース 306 は、工業規格（例えば、赤外線、無線、IEEE 802.3 などの有線規格、IEEE 802.11a、802.11b、802.11g、802.11n、802.11ac、802.15 などの無線規格、4G 通信規格など）に基づいてもよい。ネットワークインタフェース 306 においては、制御デバイス 300 がメディア再生システム 100 内の他のデバイスと通信するための手段を提供してもよい。ある例では、データおよび情報（例えば、状態変数）は、ネットワークインタフェース 306 を介して制御デバイス 300 と他のデバイスとの間で通信されてもよい。例えば、メディア再生システム 100 における再生ゾーンおよびゾーングループの構成は、制御デバイス 300 によって、再生デバイス又は別のネットワークデバイスから受信されてもよいし、あるいは制御デバイス 300 によって、ネットワークインタフェース 306 を介して別の再生デバイス又はネットワークデバイスに送信されてもよい。ある場合では、他のネットワークデバイスは、別の制御デバイスであってもよい。

【0044】

ボリュームコントロールおよびオーディオ再生コントロールなどの再生デバイス制御コマンドは、ネットワークインタフェース 306 を介して制御デバイス 300 から再生デバイスに通信されてもよい。上述したように、メディア再生システム 100 の構成の変更は、ユーザにより制御デバイス 300 を用いて行うことができる。構成の変更は、1 つ又は複数の再生デバイスをゾーンに追加すること、1 つ又は複数の再生デバイスをゾーンから取り除くこと、1 つ又は複数のゾーンをゾーングループに追加すること、1 つ又は複数のゾーンをゾーングループから取り除くこと、結合プレーヤー又は統合プレーヤーを形成すること、結合プレーヤー又は統合プレーヤーから 1 つ又は複数の再生デバイスに分けることなどを含んでもよい。このように、制御デバイス 300 は、コントローラと呼ばれてもよく、制御デバイス 300 は、メディア再生システムコントローラアプリケーションソフトウェアをインストールした専用のコントローラか、又はネットワークデバイスであってもよい。

【0045】

制御デバイス 300 は、マイクロフォン 310 を含んでもよい。マイクロフォン 310 は、制御デバイス 300 の環境内の音を検出するように構成してもよい。マイクロフォン 310 は、コンデンサマイク、エレクトレットコンデンサマイク、ダイナミックマイクなど、現在知られている又は今後開発される任意の種類 of マイクロフォンであってもよい。マイクロフォンは、一部の周波数範囲に対して高感度であってもよい。2 つ以上のマイクロフォン 310 が、音源（例えば、音声、可聴音）の位置情報を取得する、および／又は、背景雑音のフィルタリングを補助するために設けられてもよい。

【0046】

制御デバイス 300 のユーザインタフェース 308 は、図 4 に示されるコントローラインタフェース 400 などのようなコントローラインタフェースを提供することによって、メディア再生システム 100 のユーザアクセスおよび制御を可能にするように構成されていてもよい。コントローラインタフェース 400 は、再生制御領域 410、再生ゾーン領域 420、再生ステータス領域 430、再生キュー領域 440、およびオーディオコンテンツソース領域 450 を含む。図示されるユーザインタフェース 400 は、図 3 の制御デバイス 300 などのようなネットワークデバイス（および／又は図 1 の制御デバイス 126 および 128）を設けられたユーザインタフェースの単なる一例であって、ユーザによってメディア再生システム 100 などのようなメディア再生システムを制御するためにア

クセスされるものである。あるいは、様々なフォーマット、スタイル、および対話型シーケンスを他のユーザのインタフェースを1つ又は複数のネットワークデバイスに実装し、メディア再生システムへ類似の制御アクセスを提供してもよい。

【0047】

再生制御領域410は、（例えば、タッチ又はカーソルを用いることで）選択可能なアイコンを含んでもよい。このアイコンによって、選択された再生ゾーン又はゾーングループ内の再生デバイスが、再生又は停止、早送り、巻き戻し、次にスキップ、前にスキップ、シャッフルモードのオン／オフ、リピートモードのオン／オフ、クロスフェードモードのオン／オフを行う。再生制御領域410は、別の選択可能なアイコンを含んでもよい。別の選択可能なアイコンは、イコライゼーション設定、再生ボリュームなど他の設定などを変更してもよい。

10

【0048】

再生ゾーン領域420は、メディア再生システム100内の再生ゾーンの表示を含んでもよい。ある実施形態では、再生ゾーンのグラフィック表示が選択可能であってもよい。追加の選択可能なアイコンを移動させることによって、メディア再生システム内の再生ゾーンを管理又は構成することができる。例えば、結合ゾーンの作成、ゾーングループの作成、ゾーングループの分割、およびゾーングループのリネームなど他の管理又は構成を行うことができる。

【0049】

例えば、図示されるように、「グループ」アイコンは、再生ゾーンのグラフィック表示のそれぞれに設けられてもよい。あるゾーンのグラフィック表示内の「グループ」アイコンは、メディア再生システム内の1つ又は複数のゾーンを選択して、あるゾーンとグループ化するオプションを出せるように選択可能であってもよい。一度グループ化すると、あるゾーンとグループ化されたゾーン内の再生デバイスは、あるゾーン内の再生デバイスと同期してオーディオコンテンツを再生するように構成される。同様に、「グループ」アイコンは、ゾーングループのグラフィック表示内に設けられてもよい。この場合、「グループ」アイコンは、ゾーングループ内の1つ又は複数のゾーンをゾーングループから取り除くために、ゾーングループ内の1つ又は複数のゾーンを選択から外すというオプションを出すように選択可能であってもよい。ユーザインタフェース400等のユーザインタフェースを介してゾーンをグループ化およびグループ解除するための他の対話をすることも可能であるし、実施することも可能である。再生ゾーン領域420内の再生ゾーンの表示は、再生ゾーン又はゾーングループ構成が変更されると、動的に更新されてもよい。

20

30

【0050】

再生ステータス領域430は、現在再生されているオーディオコンテンツ、前に再生されたオーディオコンテンツ、又は選択された再生ゾーン又はゾーングループ内で次に再生するように予定されているオーディオコンテンツ、のグラフィック表示を含んでもよい。選択可能な再生ゾーン又は再生グループは、ユーザインタフェース上で、例えば、再生ゾーン領域420および／又は再生ステータス領域430内で視覚的に区別されてもよい。グラフィック表示は、トラックタイトル、アーティスト名、アルバム名、アルバム年、トラックの長さ、およびメディア再生システムを、ユーザインタフェース400を介して制御するときに、ユーザにとって有益な他の関連情報を含んでもよい。

40

【0051】

再生キュー領域440は、選択された再生ゾーン又はゾーングループと関連付けられた再生キュー内のオーディオコンテンツのグラフィック表示を含んでもよい。ある実施形態では、それぞれの再生ゾーン又はゾーングループは、再生ゾーン又は再生グループによって再生される0以上のオーディオアイテムに対応する情報を含む再生キューと関連付けられていてもよい。例えば、再生キュー内のそれぞれのオーディオアイテムは、ユー・アール・アイ（URI）、ユー・アール・エル（URL）、又は再生ゾーン又はゾーングループ内の再生デバイスによって使用可能な他の識別子を含んでもよい。これらによって、ローカルオーディオコンテンツソース又はネットワークオーディオコンテンツソース、

50

からオーディオアイテムを見つけ、および／又は取り出し、再生デバイスによって再生することができる。

【0052】

ある例では、プレイリストが再生キューに追加されてもよい。この場合、プレイリスト内のそれぞれのオーディオアイテムに対応する情報が再生キューに追加されてもよい。別の例では、再生キュー内のオーディオアイテムは、プレイリストとして保存されてもよい。更に別の例では、再生デバイスがストリーミングオーディオコンテンツ、例えば、再生時間を有することで連続して再生されないオーディオアイテムよりも、停止しない限り連続して再生されるインターネットラジオを再生し続けているとき、再生キューは、空であってもよいし、又は「未使用」であるが埋められていてもよい。別の実施形態では、再生キューは、インターネットラジオおよび／又は他のストリーミングオーディオコンテンツアイテムを含むことができ、且つ再生ゾーン又はゾーングループがそれらのアイテムを再生しているとき「未使用」とすることができる。他の例も可能である。

10

【0053】

再生ゾーン又はゾーングループが「グループ化される」か、又は「グループ解除」されるとき、影響を受ける再生ゾーン又はゾーングループに関連付けられた再生キューは、クリアされてもよいし、又は再び関連付けられてもよい。例えば、第1再生キューを含む第1再生ゾーンが、第2再生キューを含む第2再生ゾーンとグループ化された場合、形成されたゾーングループは、関連付けられた再生キューを有していてもよい。関連付けられた再生キューは、最初は空であるか、（例えば、第2再生ゾーンが第1再生ゾーンに追加された場合、）第1再生キューのオーディオアイテムを含むか、（例えば、第1再生ゾーンが第2再生ゾーンに追加された場合、）第2再生キューのオーディオアイテムを含むか、又は第1再生キューと第2再生キューとの両方のオーディオアイテムを組み合わせられる。その後、形成されたゾーングループがグループ解除された場合、グループ解除された第1再生ゾーンは、前の第1再生キューと再び関連付けられてもよいし、空の新しい再生キューと関連付けられてもよいし、あるいはゾーングループがグループ解除される前にゾーングループと関連付けられていた再生キューのオーディオアイテムを含む新しい再生キューと関連付けられてもよい。同様に、グループ解除された第2再生ゾーンは、前の第2再生キューと再び関連付けられてもよいし、空の新しい再生キューと関連付けられてもよいし、あるいはゾーングループがグループ解除される前にゾーングループと関連付けられていた再生キューのオーディオアイテムを含む新しい再生キューと関連付けられてもよい。その他の例も可能である。

20

30

【0054】

図4のユーザインタフェース400に戻って、再生キュー領域440内のオーディオコンテンツのグラフィック表示は、トラックタイトル、アーティスト名、トラックの長さ、および再生キュー内のオーディオコンテンツと関連付けられた他の関連情報を含んでもよい。ある例では、オーディオコンテンツのグラフィック表示は、追加の選択可能なアイコンを選択して移動させることができる。これにより、再生キューおよび／又は再生キューに表示されたオーディオコンテンツを管理および／又は操作することができる。例えば、表示されたオーディオコンテンツは、再生キューから取り除いてもよいし、再生キュー内の異なる位置に移動させてもよいし、すぐに再生させるか若しくは現在再生しているオーディオコンテンツの後に再生するように選択されてもよいし、あるいは他の動作を実行してもよい。再生ゾーン又はゾーングループに関連付けられた再生キューは、再生ゾーン又はゾーングループ内の1つ又は複数の再生デバイスのメモリ、再生ゾーン又はゾーングループに入っていない再生デバイスのメモリ、および／又は他の指定のデバイスのメモリに記憶されていてもよい。

40

【0055】

オーディオコンテンツソース領域450は、選択可能なオーディオコンテンツソースのグラフィック表示を含んでもよい。このオーディオコンテンツソースにおいては、オーディオコンテンツが選択された再生ゾーン又はゾーングループによって取り出され、再

50

生されてもよい。オーディオコンテンツソースに関する説明は、以降のセクションを参照することができる。

【0056】

d. 例示的なオーディオコンテンツソース

前回図示したように、ゾーン又はゾーングループ内の1つ又は複数の再生デバイスは、再生するオーディオコンテンツを、（例えば、オーディオコンテンツの対応するURI又はURLに基づいて、）複数の入手可能なオーディオコンテンツソースから取り出すように構成されていてもよい。ある例では、オーディオコンテンツは、再生デバイスによって、対応するオーディオコンテンツソース（例えば、ラインーイン接続）から直接取り出されてもよい。別の例では、オーディオコンテンツは、1つ又は複数の他の再生デバイス若しくはネットワークデバイスを介してネットワーク上の再生デバイスに提供されてもよい。

10

【0057】

例示的なオーディオコンテンツソースは、メディア再生システム内の1つ又は複数の再生デバイスのメモリを含んでもよい。メディア再生システムとしては、例えば、図1のメディア再生システム100、1つ又は複数のネットワークデバイス上のローカルミュージックライブラリ（例えば、制御デバイス、ネットワーク対応のパーソナルコンピュータ、又はネットワーク接続ストレージ（NAS）など）、インターネット（例えば、クラウド）を介してオーディオコンテンツを提供するストリーミングオーディオサービス、あるいは再生デバイス又はネットワークデバイスのラインーイン入力接続を介してメディア再生システムに接続されるオーディオソース、他の可能なシステムであってもよい。

20

【0058】

ある実施形態では、オーディオコンテンツソースは、図1のメディア再生システム100などのようなメディア再生システムに定期的に追加されてもよいし、定期的に取り除かれてもよい。ある例では、1つ又は複数のオーディオコンテンツソースが追加される、取り除かれる、又は更新される度に、オーディオアイテムのインデックス付けが行われてもよい。オーディオアイテムのインデックス付けは、ネットワーク上で共有される全てのフォルダ／ディレクトリ内の識別可能なオーディオアイテムをスキャンすることを含んでもよい。ここで、ネットワークは、メディア再生システム内の再生デバイスによってアクセス可能である。また、オーディオアイテムのインデックス付けは、メタデータ（例えば、タイトル、アーティスト、アルバム、トラックの長さなど）と他の関連情報とを含むオーディオコンテンツデータベースを作成すること、又は更新すること、を含んでもよい。他の関連情報とは、例えば、それぞれの識別可能なオーディオアイテムを見つけるためのURI又はURLを含んでもよい。オーディオコンテンツソースを管理し、且つ維持するための他の例も可能である。

30

【0059】

再生デバイス、制御デバイス、再生ゾーン構成、およびメディアコンテンツソースに関する上記の説明は、以下に記載される機能および方法を実装できる動作環境のうちの、一部の例を示しているに過ぎない。メディア再生システム、再生デバイス、およびネットワークデバイスに関して、本明細書に明示的に記載されていない他の動作環境および構成も同様に適用可能であり、本機能および本方法を実装するのに適している可能性がある。

40

【0060】

e. 複数の例示的なネットワークデバイス

図5は、音声コントロールに基づいてオーディオ再生体験を提供するように構成される、複数の例示的なデバイス500を示した図である。当業者であれば、図5に示されるデバイスは例示する目的のみのものであり、異なるおよび／又は追加のデバイスを含む変形例が実行可能であり得ることを理解するであろう。図示されるように、複数のデバイス500は、コンピューティングデバイス504、506、および508と、ネットワークマイクロフォンデバイス（NMD）512、514、および516と、再生デバイス（PBD）532、534、536、および538と、制御デバイス（CR）522とを含む。

50

【0061】

複数のデバイス500のそれぞれは、ワイドエリアネットワーク（WAN）、ローカルエリアネットワーク（LAN）、およびパーソナルエリアネットワーク（PAN）などの1つ又は複数の種類のネットワークを介して、NFC、Bluetooth（登録商標）、Ethernet、およびIEEE802.11その他の1つ又は複数のネットワークプロトコルに従って、複数のデバイスにおける1つ又は複数の他のデバイスとの通信を確立することができる、ネットワーク対応デバイスであってもよい。

【0062】

図示されるように、コンピューティングデバイス504、506、および508はクラウドネットワーク502の一部であってもよい。クラウドネットワーク502は追加のコンピューティングデバイスを含んでもよい。ある例では、コンピューティングデバイス504、506、および508は異なるサーバであってもよく、別の例では、コンピューティングデバイス504、506、および508のうちの2つ以上は単一サーバのモジュールであってもよい。同様に、コンピューティングデバイス504、506、および508のそれぞれは、1つ又は複数のモジュール若しくはサーバを含んでもよい。本明細書での例示を容易に行うために、コンピューティングデバイス504、506、および508のそれぞれは、クラウドネットワーク502内で特定の機能を実行するように構成されてもよい。例えば、コンピューティングデバイス508は、音楽ストリーミングサービス用のオーディオコンテンツのソースであってもよい。

【0063】

図示されるように、コンピューティングデバイス504は、通信経路542を介してNMD512、514、および516とインタフェースするように構成されてもよい。NMD512、514、および516は、1つ又は複数の「スマートホーム」システムのコンポーネントであってもよい。ある場合では、NMD512、514、および516は、図1に示されるデバイスの配置と同様に、家庭全体に物理的に配置されてもよい。別の場合では、NMD512、514、および516の2つ以上は、互いに比較的近接するように物理的に配置されてもよい。通信経路542は、インターネットを含むWAN、LAN、および／又はPANその他などの、1つ又は複数の種類のネットワークを備えてもよい。

【0064】

ある例では、NMD512、514、および516のうちの1つ又は複数のは、主に音声検出を行うように構成されたデバイスであってもよい。別の例では、NMD512、514、および516のうちの1つ又は複数のは、様々な主要ユーティリティを有するデバイスのコンポーネントであってもよい。例えば、図2および図3に関連して上述したように、NMD512、514、および516のうちの1つ又は複数のは、再生デバイス200のマイクロフォン（複数可）220又はネットワークデバイス300のマイクロフォン（複数可）310であってもよい。また場合によっては、NMD512、514、および516のうちの1つ又は複数のは、再生デバイス200又はネットワークデバイス300であってもよい。ある例では、NMD512、514、および／又は516のうちの1つ又は複数のは、マイクロフォンアレイに配置される複数のマイクロフォンを含んでもよい。

【0065】

図示されるように、コンピューティングデバイス506は、通信経路544を介してCR522並びにPBD532、534、536、および538とインタフェースするように構成されてもよい。ある例では、CR522は、図2のネットワークデバイス200などのネットワークデバイスであってもよい。したがって、CR522は、図4のコントローラインタフェース400を提供するように構成されてもよい。同様に、PBD532、534、536、および538は、図3の再生デバイス300などの再生デバイスであってもよい。このため、PBD532、534、536、および538は、図1に示されるように家庭全体に物理的に配置されてもよい。例示する目的上、PBD536および538は結合ゾーン530の一部であってもよく、その一方で、PBD532および534は自身が属するそれぞれのゾーンの一部であってもよい。上述したように、PBD532、

534、536、および538は動的に結合、グループ化、結合解除、かつグループ解除されてもよい。通信経路544は、インターネットを含むWAN、LAN、および／又はPANその他などの1つ又は複数の種類のネットワークを備えてもよい。

【0066】

ある例では、NMD512、514、および516と同様に、CR522並びにPBD532、534、536、および538も、1つ又は複数の「スマートホーム」システムのコンポーネントであってもよい。ある場合では、PBD532、534、536、および538は、NMD512、514、および516と同じ家庭全体に配置されてもよい。更に、上述したように、PBD532、534、536、および538のうちの1つ又は複数は、NMD512、514、および516のうちの1つ又は複数であってもよい。

10

【0067】

NMD512、514、および516はローカルエリアネットワークの一部であってもよく、通信経路542は、NMD512、514、および516が属するローカルエリアネットワークを、WANを介してコンピューティングデバイス504にリンクする（通信経路は図示せず）アクセスポイントを含んでもよい。同様に、NMD512、514、および516のそれぞれは、そのようなアクセスポイントを介して互いと通信してもよい。

【0068】

同様に、CR522並びにPBD532、534、536、および538は、前のセクションで説明したように、ローカルエリアネットワークおよび／又はローカル再生ネットワークの一部であってもよく、また通信経路544は、CR522並びにPBD532、534、536、および538が属するローカルエリアネットワークおよび／又はローカル再生ネットワークを、WANを介してコンピューティングデバイス506にリンクするアクセスポイントを含んでもよい。したがって、CR522並びにPBD532、534、536、および538のそれぞれもまた、そのようなアクセスポイントを介して互いと通信してもよい。

20

【0069】

ある例では、単一のアクセスポイントは通信経路542および544を含んでもよい。ある例では、NMD512、514、および516、CR522、並びにPBD532、534、536、および538のそれぞれは、家庭用の同じアクセスポイントを介してクラウドネットワーク502にアクセスしてもよい。

30

【0070】

図5に示されるように、NMD512、514、および516、CR522、並びにPBD532、534、536、および538のそれぞれはまた、通信手段546を介して他のデバイスのうちの1つ又は複数と直接通信してもよい。本明細書に記載の通信手段546は、1つ又は複数の種類のネットワークを介した、1つ又は複数のネットワークプロトコルによるデバイス間の1つ又は複数の形式の通信を含んでもよく、および／又は1つ又は複数の他のネットワークデバイスを介した通信を含んでもよい。例えば通信手段546は、一例として、Bluetooth（商標）（IEEE802.15）、NFC、Wireless Direct、および／又はプロプライエタリ無線その他のうちの1つ又は複数を含んでもよい。

40

【0071】

ある例では、CR522はBluetooth（商標）を介してNMD532と通信し、別のローカルエリアネットワークを介してPBD534と通信してもよい。別の例では、NMD514は別のローカルエリアネットワークを介してCR522と通信し、Bluetoothを介してPBD536と通信してもよい。更に別の例では、PBD532、534、536、および538のそれぞれは、ローカル再生ネットワークを介し、スパンニングツリープロトコルに従って互いと通信してもよく、その一方で、ローカル再生ネットワークとは異なるローカルエリアネットワークを介してCR522とそれぞれ通信してもよい。他の例も可能である。

50

【0072】

場合によっては、NMD 512、514、および516、CR 522、並びにPBD 532、534、536、および538間の通信手段は、デバイス間の通信の種類、ネットワーク状態、および／又は待ち時間要求に応じて変化してもよい。例えば、NMD 516をPBD 532、534、536、および538と共に最初に家庭に導入するときに、通信手段546を使用してもよい。ある場合では、NMD 516は、NFCを介してNMD 516に対応する識別情報をPBD 538に送信してもよく、またPBD 538はこれに応じて、NFC（又は他の何らかの通信形式）を介してローカルエリアネットワーク情報をNMD 516に送信してもよい。ただし、ひとたびNMD 516を家庭内に設置した後、NMD 516とPBD 538との通信手段は変化してもよい。例えばNMD 516は、通信経路542、クラウドネットワーク502、および通信経路544を連続的に経由してPBD 538と通信してもよい。別の例では、NMDとPBDとは、ローカル通信手段546を介して決して通信しないようにしてもよい。更に別の例では、NMDとPBDとは、主としてローカル通信手段546を介して通信してもよい。他の例も可能である。

10

【0073】

例示的な例では、NMD 512、514、および516は、PBD 532、534、536、および538を制御するための音声入力を受信するように構成されてもよい。利用可能な制御コマンドは、再生ボリュームコントロール、再生トランスポートコントロール、音楽ソース選択、およびグループ化その他など、前述した任意のメディア再生システムの制御を含んでいてもよい。例えばNMD 512は、PBD 532、534、536、および538のうちの1つ又は複数を制御するための音声入力を受信してもよい。音声入力を受信したことに応じて、NMD 512は通信経路542を介して、処理を目的としてその音声入力をコンピューティングデバイス504に送信してもよい。ある例では、コンピューティングデバイス504は、音声入力を同等のテキストコマンドに変換し、そのテキストコマンドを解析してコマンドを識別してもよい。次いでコンピューティングデバイス504は、続けてそのテキストコマンドをコンピューティングデバイス506に送信してもよい。別の例では、コンピューティングデバイス504は、音声入力を同等のテキストコマンドに変換し、次いで、続けてそのテキストコマンドをコンピューティングデバイス506に送信してもよい。その後コンピューティングデバイス506は、そのテキストコマンドを解析して、1つ又は複数の再生コマンドを識別してもよい。

20

30

【0074】

例えば、そのテキストコマンドが「『ストリーミングサービス1』から『アーティスト1』による『トラック1』を『ゾーン1』で再生して」である場合、コンピューティングデバイス506は（i）「ストリーミングサービス1」から入手可能な「アーティスト1」による「トラック1」のURLと、（ii）「ゾーン1」内の少なくとも1つの再生デバイスとを識別してもよい。本例では、「ストリーミングサービス1」からの「アーティスト1」による「トラック1」のURLは、コンピューティングデバイス508を指すURLであってもよく、「ゾーン1」は結合ゾーン530であってもよい。したがって、URLとPBD 536および538の一方又は両方とを識別すると、コンピューティングデバイス506は通信経路544を介して、PBD 536および538の一方又は両方に、識別された再生用のURLを送信してもよい。PBD 536および538の一方又は両方は、これに応じて、受信したURLに従ってコンピューティングデバイス508からオーディオコンテンツを取り出し、かつ「ストリーミングサービス1」から「アーティスト1」による「トラック1」の再生を開始してもよい。

40

【0075】

当業者であれば、上記は単なる例示的な例であり、他の実施形態も実行可能であることを理解する。ある場合では、上述したように、複数のデバイス500のうちの1つ又は複数によって実行される動作は、複数のデバイス500における1つ又は複数の他のデバイスによって実行されてもよい。例えば、音声入力からテキストコマンドへの変換は、NMD 512、コンピューティングデバイス506、PBD 536、および／又はPBD 53

50

8などの他のデバイスによって代替的に、部分的に若しくは完全に実行されてもよい。同様にURLの識別は、NMD 512、コンピューティングデバイス504、PBD 536、および／又はPBD 538などの別のデバイス若しくは複数のデバイスによって代替的に、部分的に若しくは完全に実行されてもよい。

【0076】

f. 例示的なネットワークマイクロフォンデバイス

図6は、図5のNMD 512、514、および516のうちの1つ又は複数を構成する例示的なネットワークマイクロフォンデバイス600の機能ブロック図を示す。図示されるように、ネットワークマイクロフォンデバイス600は、プロセッサ602、メモリ604、マイクロフォンアレイ606、ネットワークインタフェース608、ユーザインタフェース610、ソフトウェアコンポーネント612、およびスピーカ（複数可）614を含む当業者であれば、他のネットワークマイクロフォンデバイスの構成および配置も可能であることを理解する。例えば、ネットワークマイクロフォンデバイスは、代替案としてスピーカ（複数可）614を除外するか、又はマイクロフォンアレイ606の代わりに単一のマイクロフォンを有することができる。

【0077】

プロセッサ602は、汎用プロセッサ若しくはコントローラ又は専用プロセッサ若しくはコントローラの形態をとる、1つ又は複数のプロセッサおよび／又はコントローラを含んでいてもよい。例えば、処理ユニット602は、マイクロプロセッサ、マイクロコントローラ、特定用途向け集積回路、およびデジタル信号プロセッサなどを含んでいてもよい。メモリ604は、プロセッサ602によって実行されて機能を発揮する1つ又は複数のソフトウェアコンポーネントを搭載可能なデータストレージであってもよい。したがって、メモリ604は、ランダムアクセスメモリ、レジスタ、キャッシュなどが例として挙げられる1つ又は複数の非一時的なコンピュータ読取り可能記録媒体と、読取り専用メモリ、ハードディスクドライブ、ソリッドステートドライブ、フラッシュメモリ、および／又は光学記憶装置その他などの1つ又は複数の不揮発性記録媒体とを備えていてもよい。

【0078】

マイクロフォンアレイ606は、ネットワークマイクロフォンデバイス600の環境内の音を検出するように構成された、複数のマイクロフォンであってもよい。マイクロフォンアレイ606は、コンデンサマイク、エレクトレットコンデンサマイク、又はダイナミックマイクなど、現在知られている又は今後開発される任意の種類のマイクロフォンを含んでもよい。ある例では、マイクロフォンアレイは、ネットワークマイクロフォンデバイスに対する1つ又は複数の方向から音声を検出するように構成されてもよい。マイクロフォンアレイ606は、一部の周波数範囲に対して高感度であってもよく、ある例では、マイクロフォンアレイ606の第1サブセットは第1周波数範囲に対して高感度であってもよい一方、マイクロフォンアレイの第2サブセットは第2周波数範囲に対して高感度であってもよい。更に、マイクロフォンアレイ606は、オーディオソース（例えば、音声、可聴音）の位置情報を取得する、および／又は、背景雑音のフィルタリングを補助するために設けられてもよい。特にある実施形態では、マイクロフォンアレイは、複数のマイクロフォンではなく、単一のマイクロフォンのみから構成されてもよい。

【0079】

ネットワークインタフェース608は、図5に関連してCR 522、PBD 532～538、コンピューティングデバイス504～508などのクラウドネットワーク502内の様々なネットワークデバイス、および他のネットワークマイクロフォンデバイスなどの間における無線および／又は有線通信を容易にするように構成されてもよい。このため、ネットワークインタフェース608は、これらの機能を実行するのに適した任意の形態をとることができ、その例としては、Ethernetインタフェース、シリアルバスインタフェース（例えば、FireWire、USB 2.0など）、無線通信を容易にするように構成されたチップセットおよびアンテナ、並びに／又は有線および／又は無線通信を提供する他の任意のインタフェースが挙げられる。ある例では、ネットワークインタフェ

ース608は、工業規格（例えば、赤外線、無線、IEEE802.3などの有線規格、IEEE802.11a、802.11b、802.11g、802.11n、802.11ac、802.15などの無線規格、4G通信規格など）に基づいてもよい。

【0080】

ネットワークマイクロフォンデバイス600のユーザインタフェース610は、ネットワークマイクロフォンデバイスとのユーザ対話を容易にするように構成されていてもよい。ある例では、ユーザインタフェース608は、ユーザがネットワークマイクロフォンデバイス600に直接入力できるようにするために、物理的ボタン、タッチセンサ画面（複数可）および／又は面（複数可）に設けられるグラフィカルインタフェースその他のうちの1つ又は複数を含んでいてもよい。ユーザインタフェース610は、視覚的および／又は聴覚的フィードバックをユーザに提供するために、1つ又は複数の照明およびスピーカ（複数可）614を更に含んでいてもよい。ある例では、ネットワークマイクロフォンデバイス600は、スピーカ（複数可）614を介してオーディオコンテンツを再生するように更に構成されていてもよい。

【0081】

ここでいくつかの例示的な実施形態である、図7、図8、および図9に示される実施形態700、800、並びに900を参照すると、本明細書に記載される技術の例示的な実施形態がそれぞれ提示されている。例えば、図1のメディア再生システム100、図2の再生デバイス200の1つ又は複数、又は図3の制御デバイス300の1つ又は複数、並びに本明細書に記載の他のデバイスおよび／又は他の適切なデバイスを含む動作環境内に、これらの例示的な実施形態を実装することができる。更に、メディア再生システムによって実行されるものとして例として図示されている動作は、メディア再生システムの再生デバイス又は制御デバイスなどの任意の適切なデバイスによって実行されてもよい。実施形態700、800、および900は、図7、図8、および図9に示されるブロックのうちの1つ又は複数によって図示されるように、1つ又は複数の動作、機能、若しくはアクションを含んでもよい。ブロックを順番に図示しているが、これらのブロックは同時に、および／又は本明細書に記載の順番とは異なる順番で実行されてもよい。また、様々なブロックは、より少数のブロックに組み合わされるか、追加のブロックに分割されるか、および／又は所望の実施形態に基づいて取り除かれてもよい。

【0082】

更に、本明細書に開示される実施形態について、フローチャートは、本実施形態の実行可能な1つの実施態様の機能および動作を示す。これに関してそれぞれのブロックは、プロセッサによって実行されてプロセスにおける特定の論理機能又はステップを実装するための1つ又は複数の命令を含むプログラムコードのモジュール、セグメント、又は一部を表すことができる。このプログラムコードは、例えばディスク又はハードドライブを含むストレージデバイスなどの、任意の種類のコンピュータ読取り可能媒体に記憶されてもよい。コンピュータ読取り可能媒体としては、例えばレジスタメモリ、プロセッサキャッシュ、およびランダムアクセスメモリ（RAM）のように短期間データを記憶するコンピュータ読取り可能媒体などの、非一時的なコンピュータ読取り可能媒体が挙げられる。更に、コンピュータ読取り可能媒体として、例えば読取り専用メモリ（ROM）、光ディスク又は磁気ディスク、コンパクトディスク読取り専用メモリ（CD-ROM）のように二次的な、又は永続的な長期ストレージなどの非一時的な記録媒体も挙げることができる。また、コンピュータ読取り可能媒体は、他の任意の揮発性又は不揮発性ストレージシステムであってもよい。コンピュータ読取り可能媒体は、例えばコンピュータ読取り可能記録媒体、又は有形のストレージデバイスと見なすことができる。更に、本明細書に開示される実施形態について、それぞれのブロックは、プロセスにおける特定の論理機能を実行するように配線されている回路を表すことができる。

【0083】

III. 音声サービスを起動するための例示的なシステムおよび方法

上述したように、ある例では、コンピューティングデバイスは音声サービスを使用して

10

20

30

40

50

音声コマンドを処理することができる。実施形態 700 は、音声サービスに音声入力进行处理させるための例示的な技術である。

【0084】

a. 音声入力を示す音声データの受信

ブロック 702 で、実施形態 700 は、音声入力を示す音声データを受信するステップを含む。例えば、NMD 600 などの NMD はマイクロフォンを介して、音声入力を示す音声データを受信することができる。更に別の例として、図 1 の再生デバイス 102、104、106、108、110、112、114、116、118、120、122、および 124 又は制御デバイス 126 および 128 のいずれかは NMD であってもよく、また音声入力を示す音声データを受信してもよい。更に別の例では、NMD は、図 5 の NMD 512、514、および 516、PBD 532、534、536、および 538、並びに CR 522 を含む。

10

【0085】

NMD は、1 つ又は複数のマイクロフォンを介して、周辺雑音を継続的に録音（すなわち、音声入力を聞く）してもよい。NMD は、この継続的な録音をリングバッファ又はサーキュラバッファに記憶してもよい。そのようなバッファにおいては、その録音は音声入力を含まない限り上書き（すなわち、破棄される）されてもよい。このバッファは、本明細書に記載のデバイス又はサーバのいずれかを介して、ローカルおよび／又はリモートに記憶されてもよい。そのような場合、音声入力を示す音声データを受信するステップは、音声入力を含む音声データをバッファに録音するステップを含んでもよい。

20

【0086】

NMD は、音声データの一部がウェイクワード又はウェイクフレーズを含んでいることを検出することにより、音声入力を受信されたことを検出することができる。例えば音声入力は、ウェイクワードとそれに続く音声コマンドとを含んでいる場合がある。ウェイクワードは、NMD が音声入力を能動的に聞き取るための時間間隔又は時間枠を開始することができる。時間間隔又は時間枠は、一定の時間（例えば、NMD が最初の音声入力を受信してから 1 分後）が経過すると期限切れになるようにしてもよい。「ユーザ ID に基づくアクション」と題する米国特許出願第 15/131,776 号が本明細書に参照として援用されており、それには更に別の例が記載されている。商業的に使用されているいくつかの例示的なウェイクワードとしては、「ヘイ、Siri」（アップル（登録商標）社）、「オッケー、Google」（グーグル（登録商標）社）、および「Alexa」（アマゾン（登録商標）社）が挙げられる。あるいは、ウェイクワードは独自ののものであってもよい（例えば、ユーザ定義の）。

30

【0087】

例示のために図 1 に戻って、ユーザは、主寝室ゾーンにいる間に特定の音声入力を発することができる。NMD として機能している再生デバイス 122（および／又は再生デバイス 124）は、音声入力を聞き取り（すなわち、マイクロフォンを介した録音であり、おそらくはバッファに録音される）、かつユーザの声を音声入力として検出することができる。特定の音声入力は、NMD にユーザの声を音声入力として容易に認識させるようにするためのウェイクワードを含んでもよい。

40

【0088】

例示的な音声コマンドは、メディア再生システムの制御又は再生設定のいずれかを変更するように命じるコマンドを含んでもよい。再生設定には、例えば再生ボリューム、再生トランスポートコントロール、音楽ソース選択、およびグループ化その他が含まれていてもよい。他の音声コマンドは、他のデバイス操作の中でもとりわけ、テレビコントロール又は再生設定、携帯電話端末の設定、又は照明装置を調整するための操作を含んでもよい。より多くの家庭用装置が「スマート」になるにつれて（例えば、ネットワークインタフェースを搭載することによって）、音声コマンドを使用して様々な家庭用装置をコントロールすることができるようになる。

【0089】

50

一例として、NMDは音声入力を示す音声データを、ネットワークインタフェースを介して、おそらくは家庭内の別のNMDから受信することができる。NMDはマイクロフォンを介して、音声入力を示す音声データを受信することに加えて、当該録音を受信してもよい（例えば、2つのNMDが両方とも音声入力の検出範囲内にある場合）。

【0090】

そのような実施形態では、NMDは周辺雑音を継続的に録音しない可能性がある。むしろ、場合によってはNMDは、NMDに対して「起動」して音声入力又はコマンドの録音を開始するように指示する音声入力又は指示を受信してもよい。例えば、第1NMD（例えば、図1に示される再生デバイス104）は音声入力を受信し、本明細書に記載のある状況では、録音を開始するように指示する指示を1つ又は複数の第2NMD（例えば、再生デバイス106および／又は108その他など）に送信してもよい。

10

【0091】

いくつかの例においては、NMDデバイスが音声データを受信する前に、複数のNMDからの音声録音が精査され、処理され、および／又は単一の音声入力にまとめられてもよい。例として、NMD512は、514又は516などの1つ又は複数の他のNMDから音声録音を受信することができる。いくつかの実施形態では、PBD532、534、536および／又は538はNMDとして構成されてもよく、またNMD512は、PBD532、534、536および／又は538のうちの1つから音声録音を受信してもよい。NMD（又は複数のNMD）は、音声録音を精査し、処理し、および／又はまとめて単一の音声入力にし、さらなる処理を目的として、この単一の音声入力をコンピューティングデバイスに送信してもよい。

20

【0092】

b. 音声入力を処理するための音声サービス（複数可）の識別

ブロック704で、実施形態700は、音声入力を処理するための1つ又は複数の音声サービスを識別するステップを含む。例えば、NMDは、受信した音声データに示されている音声入力を処理するための特定の音声サービスを識別することができる。あるいは、NMDは、音声入力を処理するための複数の音声サービスを識別してもよい。

【0093】

NMDは、音声入力を処理するための特定の音声サービスを、利用可能な音声サービスの中から識別することができる。音声サービスは、様々な技術を使用してNMDで利用可能となってもよい。利用可能な音声サービスは、NMDに登録されている音声サービスを含んでいてもよい。NMDに所定の音声サービスを登録する作業は、音声サービスのユーザ認証情報（例えば、ユーザ名およびパスワード）をNMDに提供するステップ、および／又はNMDの識別子を音声サービスに提供するステップを含んでいてもよい。そのような登録作業において、音声サービスに代わって音声入力を受信するようにNMDを構成し、またおそらくは処理を目的として、NMDから音声入力を受信するように音声サービスを構成してもよい。登録作業は設定手順の間に行われてもよい。

30

【0094】

場合によっては、NMDはメディア再生システムと関連付けられてもよい。NMDは、メディア再生システム自体の一部として（例えば、制御デバイス又は再生デバイスとして）、又はメディア再生システムと相互接続された別のデバイスとして機能し、場合によってはメディア再生システムの特定の動作（例えば、再生デバイスの音声コントロール）を容易に行うことができる。1つ又は複数の音声サービスが所定のメディア再生システムに登録されてもよく、NMDは音声入力を処理するために、登録された音声サービスを識別することができる。

40

【0095】

メディア再生システムの登録作業において、音声サービスに代わって音声入力を受信するように、メディア再生システムのNMD（例えば、制御デバイス、再生デバイス、又はその他の関連デバイス）を構成してもよい。更に、こうした登録作業において、処理を目的としてこれらのデバイスから音声入力を受信するように、音声サービスを構成してもよ

50

い。メディア再生システムに音声サービスを登録する作業は、設定手順の間に行われてもよい。例示的な設定手順は、再生デバイス（又は複数の再生デバイス）および／又は制御デバイスを新たなメディア再生システムに設定する手順を含む。他の例示的な設定手順は、メディア再生システムを変更するための手順（例えば、デバイスを本システムに追加、又は本システムから取り除く手順、あるいは本システムに音声サービスを設定する手順）を含む。

【0096】

場合によっては、単一の音声サービスがNMDで利用可能であってもよく、これによって音声入力を処理するための音声サービスを簡便に識別することができるようになる。NMDによって受信された音声入力は音声サービスに直接送信されてもよく、またNMDによって応答が提供されてもよい。そのような実施形態では、NMDは単一の音声サービス用のマイクロフォンインタフェースおよびスピーカインタフェースとして機能することになる。

10

【0097】

他の場合では、音声入力を処理するために、複数の音声サービスがNMDで利用可能であってもよい。そのような場合、NMDは、音声入力を処理するための特定の音声サービスを、複数の音声サービスの中から識別することができる。例えば、NMDは、メディア再生システムに登録されている複数の音声サービスの中から、特定の音声サービスを識別することができる。上述したように、NMDはメディア再生システムの一部（例えば、再生デバイス又は制御デバイスとして）であってもよく、あるいは本システムと関連付けられていてもよい。

20

【0098】

音声入力を処理するための特定の音声サービスを識別するステップは、音声入力におけるウェイクワード又はウェイクフレーズに基づいてもよい。例えば、音声入力を示す音声データを受信した後、NMDは、音声データの一部分が特定のウェイクワードを表していると決定することができる。更に、NMDは、この特定のウェイクワードが特定の音声サービスに対応していると決定してもよい。換言すれば、NMDは、特定のウェイクワード又はウェイクフレーズが特定の音声サービスを起動するために使用されていると決定してもよい。例えば、特定のウェイクワードとしては、アップル（登録商標）社の音声サービスを起動するための「ヘイ、S i r i」、グーグル（登録商標）社の音声サービスを起動するための「オッケー、G o o g l e」、アマゾン（登録商標）社の音声サービスを起動するための「A l e x a」、又はマイクロソフト社の音声サービスを起動するための「ヘイ、C o r t a n a」が挙げられる。あるいは、特定の音声サービスを起動するために、独自のウェイクワード（例えば、ユーザ定義の）を定義することができる。NMDが、受信した音声データにおける特定のウェイクワードが特定の音声サービスに対応していると決定した場合、NMDはその特定の音声サービスを、音声データにおける音声入力を処理するための音声サービスとして識別することができる。

30

【0099】

特定のウェイクワードが特定の音声サービスに対応していると決定するステップは、音声データ（例えば、ウェイクワード又はウェイクフレーズに対応している音声データの一部）を用いて1つ又は複数の音声サービスにクエリを実行するステップを含んでもよい。例えば、音声サービスは、音声データがその音声サービスに対応するウェイクワード又はウェイクフレーズを含んでいるかどうかを決定するためにNMDが呼び出すことができる、アプリケーションプログラミングインタフェースを提供してもよい。NMDは、受信した音声データにおいてウェイクワード部分を表すデータと共に、その音声サービスに関する特定のクエリを音声サービスに送信することによって、A P Iを呼び出すことができる。あるいは、NMDはNMD自体のA P Iを呼び出すことができる。NMD又はメディア再生システムに音声サービスを登録する作業により、音声サービスのA P I又は他のアーキテクチャをNMDと一体化させることができる。

40

【0100】

50

複数の音声サービスがNMDで利用可能である場合、NMDは、複数の音声サービスにおけるそれぞれの音声サービスに対応するウェイクワード検出アルゴリズムでクエリを実行する可能性がある。上述したように、そのような検出アルゴリズムでクエリを実行するステップは、NMD上でローカルに、又はネットワークインタフェースを使用してリモートに、複数の音声サービスのそれぞれのAPIを呼び出すステップを含んでもよい。所定の音声サービスのウェイクワード検出アルゴリズムに関するクエリへの応答として、NMDは、クエリにおける音声データがその音声サービスに対応するウェイクワードを含んでいたかどうかを示す応答を受信することができる。受信した音声データが特定の音声サービスに対応する特定のウェイクワードを表していることを特定の音声サービスのウェイクワード検出アルゴリズムが検出した場合、NMDはその特定の音声サービスを、音声入力

10

【0101】

場合によっては、受信された音声データが、特定の音声サービスに対応する認識可能なウェイクワードを含んでいないにもかかわらず、音声入力を含んでいることがある。そのような状況は、周辺雑音又は他の要因のために、所定のウェイクワードが明確に検出されない場合に発生し、その結果、ウェイクワード検出アルゴリズム（複数可）が所定のウェイクワードを、任意の特定の音声サービスに対応するものとして認識しない可能性がある。あるいは、ユーザが特定の音声サービスに対応するウェイクワードを発していない可能性もある。例えば、特定の音声サービスに対応していない一般的ウェイクワードを使用して、音声入力の処理を呼び出している場合がある（例えば、「ヘイ、Sonos」）。

20

【0102】

そのような場合、NMDはコンテキストに基づいて、音声入力を処理するためのデフォルトの音声サービスを識別することができる。デフォルトの音声サービスはあらかじめ決定されていてもよい（例えば、上述の例示的な手順などの設定手順中に設定される）。その場合、NMDは、受信した音声データが特定の音声サービスに対応するウェイクワードを含んでいないと決定すると（例えば、NMDが音声データにおいて特定の音声サービスに対応するウェイクワードを検出しなかった場合）、音声入力を処理するためにデフォルトの音声サービスを選択することができる。

【0103】

上述したように、いくつかの例示的なシステムは、場合によっては複数のゾーンに設置された複数のNMD（例えば、リビングルーム、キッチン、ダイニングルーム、および寝室ゾーンを対象とし、各自がそれぞれの再生デバイスを有する図1のメディア再生システム100）を含んでもよい。このようなシステムでは、デフォルトの音声サービスはNMDごと、又はゾーンごとに設定されてもよい。その場合、所定のNMD又はゾーンによって検出された音声入力は、そのNMD又はゾーンのデフォルトの音声サービスによって処理されてもよい。場合によってはNMDは、所定のNMD又はゾーンによって検出された音声入力が、そのゾーンと関連付けられている音声サービスによって処理されることが意図されていると想定してもよい。ただし他の場合では、ウェイクワード又はウェイクフレーズにより、音声入力は特定のNMD又はゾーンに送信されてもよい（例えば、「ヘイ、キッチン」の場合は、音声入力はキッチンゾーンへと送信される）。

30

40

【0104】

例示のために図1を参照すると、再生デバイス122および／又は124は、主寝室ゾーンのNMDとして機能していてもよい。このゾーンによって検出された、および／又はこのゾーンに送信された音声入力（例えば、「ヘイ主寝室、今日の天気はどう？」）は、主寝室ゾーンのデフォルトの音声サービスによって処理されてもよい。例えば、主寝室ゾーンのデフォルトの音声サービスが「アマゾン（登録商標）社のAlexa（登録商標）」である場合、主寝室ゾーンにあるNMDの少なくとも1つが、Alexaに天気に関するクエリを実行することになる。音声入力が特定の音声サービスに対応するウェイクワード又はウェイクフレーズを含んでいる場合、そのウェイクワード又はウェイクフレーズによってデフォルトの音声サービスが無効にされ（特定の音声サービスがデフォルトの音声

50

サービスと異なる場合)、NMDがその特定の音声サービスを、音声入力进行处理するために識別できるようになる。

【0105】

いくつかの実施形態では、NMDは、その音声入力を提供しているユーザの識別情報に基づいて、音声サービスを識別してもよい。人間の声は高さ、声質、および他の特性によって異なる可能性があり、これらによってそのユーザの声によって特定のユーザを識別するための特性が、もたらされてもよい。場合によっては、家庭内のユーザがそれぞれの声を認識させるように、NMDを訓練してもよい。

【0106】

家庭内のユーザは各自、それぞれ独自の好みの音声サービスを利用してもよい。例えば、家庭内の第1ユーザおよび第2ユーザは、第1音声サービスと第2音声サービスとをそれぞれ使用するように、NMDを設定してもよい(例えば、SIRI(登録商標)およびCORTANA(登録商標))。NMDが音声入力において第1ユーザの声を認識した場合、NMDは音声コマンド进行处理するために、第1音声サービスを識別してもよい。ただし、NMDが音声入力において第2ユーザの声を認識した場合、NMDは音声コマンド进行处理するために、第2音声サービスを代わりに識別することができる。

【0107】

あるいは、NMDはコンテキストに基づいて、音声入力进行处理するための特定の音声サービスを識別してもよい。例えば、NMDは、コマンドの種類に基づいて特定の音声サービスを識別してもよい。NMD(例えば、メディア再生システムと関連付けられたNMD)は、特定の種類のコマンド(例えば、メディア再生コマンド)として、一定のコマンド(例えば、再生、停止、前方読みとばしなど)を認識することができる。そのような場合において、NMDは、その音声入力特定の種類のコマンド(例えば、メディア再生コマンド)を含んでいると決定したとき、その音声入力进行处理するための音声サービスとして、その種のコマンド进行处理するように構成されている特定の音声サービスを識別してもよい。更に例示すると、検索クエリは別の例示的な種類のコマンド(例えば、「今日の天気はどう?」や「デヴィッド・ボウイの出生地はどこ?」)であってもよい。NMDは、音声入力検索クエリを含んでいると決定した場合、検索クエリを含むその音声入力进行处理するために、特定の音声サービス(例えば、「GOOGLE」)を識別してもよい。

【0108】

NMDは場合によっては、その音声入力特定の種類のデバイスを対象とした音声コマンドを含んでいると決定してもよい。そのような場合、NMDは、その種のデバイスを対象とした音声入力进行处理するように構成されている特定の音声サービスを、当該音声入力进行处理するために識別してもよい。例えば、NMDは、所定の音声入力1つ又は複数の無線照明デバイスを対象としている(例えば、「ここの電気をつけて」は、NMDと同じ部屋にある「スマート」電球を対象としている)と決定し、その音声入力进行处理するための音声サービスとして、無線照明デバイスを対象とした音声入力进行处理するように構成されている特定の音声サービスを識別してもよい。別の例として、NMDは、所定の音声入力再生デバイスを対象としていると決定し、その音声入力进行处理するための音声サービスとして、再生デバイスを対象とした音声入力进行处理するように構成されている特定の音声サービスを識別してもよい。

【0109】

いくつかの例においては、NMDは以前の入力に基づいて、その音声入力进行处理するために特定の音声サービスを識別することができる。ユーザは、第1音声入力所定の音声サービスによって処理されていた場合において、後に続く第2音声入力、他の想定可能なコンテキスト上の要素の中でもとりわけ、同じ種類の同じデバイスを対象としているか、あるいは第1コマンドの直後にこれが提供されているとき、この第2音声入力も同様に、当該音声サービスによって処理されることを期待している可能性がある。例えば、NMDは、以前の音声入力所定の音声サービスによって処理されたことと、現在の音声入力以前の音声入力と同じ種類の動作を対象としていることを決定することができる(例

10

20

30

40

50

えば、両方ともメディア再生コマンドであると決定する)。そのような状況では、NMDは、現在の音声入力を処理するために当該音声サービスを識別してもよい。

【0110】

別の例として、NMDは、以前の音声入力が所定の音声サービスによって処理されたことと、現在の音声入力が、以前の音声入力受信後の閾値時間内(例えば、1~2分以内)に受信されたことを決定することができる。例示すると、再生デバイス114は、第1音声入力(「ヘイキッチン、ジャニス・ジョプリンの曲を再生して」)を受信し、その第1音声入力を処理するための音声サービスを識別し、その結果、再生デバイス114がジャニス・ジョプリンによるオーディオトラックを再生することができる。その後、再生デバイス114は、後に続く第2音声入力を受信し(「ボリュームを上げて」)、その第2音声入力を処理するための音声サービスを識別してもよい。メディア再生コマンドとしてのこの種のコマンド間の類似性および/又は2つの音声入力間における経過時間を所定として、再生デバイス114は、第2音声入力を処理するために、第1音声入力を処理するために識別したのと同じ音声サービスを識別してもよい。

10

【0111】

一例として、NMDは、音声入力を処理するために第1音声サービスを識別し、その後、第1音声サービスが音声入力を処理するために利用できないと決定してもよい(おそらく、一定時間内に結果を受信できなかったことによって)。音声サービスは、当該サービスの有効期限切れ、クラウドサービスに関する技術的問題、又は可用性を侵害する悪意のあるイベント(例えば、分散型サービス妨害の攻撃)などを含む、いくつかの理由で利用できなくなる可能性がある。

20

【0112】

そのような場合、NMDは音声入力を処理するために、代替の第2音声サービスを識別することができる。この代替音声サービスは、デフォルトの音声サービスである可能性がある。あるいは、システムに登録される複数の音声サービスが優先順位によってランク付けされてもよく、この代替音声サービスは次に優先順位の高い音声サービスであってもよい。他の例も可能である。

【0113】

場合によっては、NMDは、代替音声サービスを識別する際にユーザからの入力を要求してもよい。例えば、NMDは、ユーザが代替音声サービスを指定するように要求してもよい(例えば、「GOOGLE(登録商標)は現在応答していません。別のサービスを検索しますか?」となる)。更にまた、NMDは代替音声サービスを識別し、ユーザに対して、代わりにこの代替音声サービスを検索したいかどうかを確認してもよい(例えば、「SIRI(登録商標)は現在応答していません。代わりにALEXA(登録商標)を検索しますか?」となる)。あるいは、別の例として、NMDが代替音声サービスに対してクエリを実行し、結果を返すときにユーザに通知してもよい(例えば、「CORTANA(登録商標)は利用できませんでした。以下の結果はSIRI(登録商標)から得ました」となる)。元の音声サービスがひとたび利用可能になると、NMDはこの状況変化をユーザに通知し、おそらくは現在の音声サービスを変更してもよい(例えば、「SIRI(登録商標)は現在利用可能である。代わりにSIRI(登録商標)に問い合わせますか?」となる)。そのような応答は、NMDのデータストレージ上に記憶されたオーディオデータから、又はNMDがアクセスできるオーディオデータから生成されてもよい。

30

40

【0114】

代替の第2音声サービスにクエリを実行する場合、NMDは第1音声サービスの1つ又は複数の設定値を第2音声サービスに適用しようと試みることができる。例えば、そのクエリが特定のアーティストによるメディアコンテンツを再生することであり、デフォルトのオーディオサービスが第1音声サービス(例えば、特定のメディアストリーミングサービス)に対して設定されている場合、NMDは第2音声サービスに対して、デフォルトのオーディオサービスからの特定のアーティストによるオーディオトラックについてクエリの実行を試みてもよい。ただし、異なる設定値(例えば、異なるデフォルトサービス)が

50

第2音声サービスに対して設定されている場合、そのような設定値により、第2音声サービスにクエリを実行するときに第1音声サービスの設定値が上書きされてもよい。

【0115】

場合によっては、単一の音声サービスのみがNMDで利用可能となっている。例えば、メディア再生システムの設定中に、メディア再生システムに対して特定の音声サービスが選択されていてもよい。一例として、特定の音声サービスが選択されている場合、他の音声サービスに対応するウェイクワードが非アクティブになり、これらのウェイクワードを検出して処理が開始されないようにすることができる。音声サービスは、音声入力でクエリを実行されたときに音声サービスの動作を変更するための、様々な設定値を含んでもよい。例えば、優先されるメディアストリーミングサービス、又はデフォルトのメディアストリーミングサービスを設定することができる。メディア再生音声コマンド（例えば、「ケイティ・ペリーの曲を再生して」）は、その特定の音楽サービスからメディアコンテンツ（例えば、ケイティ・ペリーによるオーディオトラック）を参照することになる。

10

【0116】

c. 識別された音声サービス（複数可）による音声入力処理の実行

ブロック706で、実施形態700は、識別された音声サービス（複数可）に音声入力を処理させるステップを含む。例えば、NMDはネットワークインタフェースを介して、識別された音声サービス（複数可）の1つ又は複数のサーバに対して、音声入力を示すデータと、音声入力を示すデータを処理するように命じるコマンド又はクエリとを送信してもよい。このコマンド又はクエリは、識別された音声サービス（複数可）に音声コマンドを処理させてもよい。このコマンド又はクエリは、これらが識別された音声サービスに（例えば、当該音声サービスのAPIに）適合するように、識別された音声サービスに応じて異なってもよい。

20

【0117】

上述したように、音声データは音声入力を示していてもよく、この音声入力は、ウェイクワードを表す第1部分と音声コマンドを表す第2部分とを含んでもよい。NMDは場合によっては、音声入力において少なくとも第2部分（例えば、音声コマンドを表す部分）を示すデータのみを送信してもよい。第1部分を含まないことによって、NMDは、他の想定可能な利点の中でもとりわけ、コマンドを送信するのに必要な帯域幅を低減し、かつウェイクワードに起因して発生し得る音声入力に対する誤処理を回避することができる。あるいはNMDは、音声入力において両方の部分を示しているか、又は音声データの他の何らかの部分を示しているデータを送信してもよい。

30

【0118】

識別された音声サービスに音声入力を処理させた後、NMDはその処理の結果を受信することができる。例えば、その音声入力が検索クエリを示していた場合、NMDは検索結果を受信してもよい。別の例として、その音声入力がデバイスに対するコマンド（例えば、再生デバイスに対するメディア再生コマンド）を示していた場合、NMDはコマンドと、おそらくはそのコマンドと関連付けられた追加のデータ（例えば、コマンドと関連付けられたメディアのソース）とを受信してもよい。NMDはこれらの結果を、コマンドの種類と受信結果とに応じて適切に出力することができる。

40

【0119】

あるいは、音声コマンドが当該NMD以外の別のデバイスを対象としていた場合、結果は当該NMDではなくそのデバイスに送信される可能性がある。例えば、図1を参照すると、キッチンゾーンの再生デバイス114は、ダイニングルームゾーンの再生デバイス112を対象とした音声入力（例えば、再生デバイス112にメディア再生を調整させるための）を受信する場合がある。そのような実施形態では、再生デバイス114がその音声入力の処理を円滑に進めるが、この処理の結果（例えば、メディア再生の調整を命じるコマンドは再生デバイス112に送信されてもよい）。あるいは、音声サービスが結果を再生デバイス114に送信してもよく、再生デバイス114が当該コマンドを再生デバイス112に送信してもよく、又は再生デバイス112に当該コマンドを実行させてもよい。

50

【0120】

NMDは、識別された音声サービスにいくつかの音声入力を処理させることができるが、他の音声入力がNMD自体によって処理されてもよい。例えば、NMDがメディア再生システムの再生デバイス、制御デバイス、又は他のデバイスである場合、NMDはメディア再生コマンドの音声認識を含んでいてもよい。別の例として、NMDは音声入力のウェイクワード部分を処理してもよい。場合によっては、NMDが処理すると、音声サービスを使用して処理するよりも高速の応答時間が可能になり得る。ただし、場合によっては、音声サービスを使用して処理するとより有効な結果、および／又はNMDを介した処理では得られない結果が得られる可能性がある。いくつかの実施形態では、NMDと関連付けられた音声サービス（例えば、NMDの製造業者によって運営されている）は、そのような音声認識を容易に行うことができる。

10

【0121】

I V. 音声サービスを起動するための例示的システムおよび方法

上述したように、ある例では、コンピューティングデバイスは音声サービスを使用して音声コマンドを処理することができる。実施形態800は、音声サービスに音声入力を処理させるための例示的な技術である。

【0122】

a. 音声入力を示す音声データの受信

ブロック802で、実施形態800は、音声入力を示す音声データを受信するステップを含む。例えば、NMDは、他の実行可能な実施形態の中でもとりわけ、実施形態700のブロック702に関連して上述した例示的な技術のいずれかを使用して、マイクロフォンを介して音声入力を示す音声データを受信することができる。

20

【0123】

b. 受信した音声データに一般的ウェイクワードを表す部分が含まれている場合の決定

ブロック804で、実施形態800は、受信した音声データに一般的ウェイクワードを表す部分が含まれていると決定するステップを含む。一般的ウェイクワードは、特定の音声サービスに対応していない可能性がある。その代わりに、一般的ウェイクワードは、一般的にNMD又はメディア再生システムに対応していてもよい（例えば、ソノス（登録商標）のメディア再生システムの場合は「ヘイ、Sonos」、あるいはメディア再生システムのキッチンゾーンの場合は「ヘイ、キッチン」）。一般的であることで、一般的ウェイクワードによって特定の音声サービスが起動しないことが想定され得る。むしろ、複数の音声サービスが登録されている場合、一般的ウェイクワードによってこれらの音声サービス全てが起動して、最良の結果を得ようとするのが想定され得る。あるいは、単一の音声サービスが登録されている場合、一般的ウェイクワードによって、その音声サービスが起動することが想定され得る。

30

【0124】

c. 音声サービス（複数可）による音声入力処理の実行

ブロック806で、実施形態800は、1つ又は複数の音声サービス（複数可）に音声入力を処理させるステップを含む。例えば、NMDは、他の実行可能な実施形態の中でもとりわけ、実施形態700のブロック706に関連して上述した例示的な技術のいずれかを使用して、音声サービス（複数可）に音声入力を処理させることができる。

40

【0125】

場合によっては、複数の音声サービスがNMDで利用可能となっている。例えば、複数の音声サービスが、NMDと関連付けられたメディア再生システムに登録されている。そのような例では、NMDは利用可能な音声サービスのそれぞれに音声入力を処理させてもよい。例えば、NMDはネットワークインタフェースを介して、複数の音声サービス（複数可）のそれぞれのサーバに対して、音声入力を示すデータと、音声入力を示すデータを処理するように命じるコマンド又はクエリとを送信してもよい。このコマンド又はクエリは、識別された音声サービス（複数可）に音声コマンドを処理させてもよい。このコマンド又はクエリは、これらが当該音声サービスに（例えば、当該音声サービスのAPIに）

50

適合するように、それぞれの音声サービスに応じて異なってもよい。

【0126】

音声サービス（複数可）に音声入力を処理させた後、NMDはその処理の結果を受信することができる。例えば、その音声入力が検索クエリ又はメディア再生コマンドを示していた場合、NMDは検索結果又はコマンドをそれぞれ受信してもよい。NMDは、それぞれの音声サービス又は音声サービスのサブセットから結果を受信してもよい。一部の音声サービスでは、発生し得る全ての入力に対して結果を返すとは限らない。

【0127】

d. 音声サービス（複数可）のうちの特定の音声サービスからの出力結果

ブロック806で、実施形態800は、音声サービス（複数可）のうちの特定の音声サービスからの結果を出力するステップを含む。1つの音声サービスからのみ結果を受信した場合、NMDはその結果を出力してもよい。ただし、複数の音声サービスから結果を受信した場合、NMDは、複数の音声サービスからのそれぞれの結果の中から特定の結果を選択し、その結果を出力してもよい。

10

【0128】

例示として、ある例では、NMDは「ヘイキッチン、テイラー・スウィフトの曲を再生して」という音声入力を受信してもよい。NMDは、音声入力のウェイクワード部分（「ヘイ、キッチン」）が特定の音声サービスを指定していないことから、一般的であると決定してもよい。この種のウェイクワードを受信した場合、NMDは複数の音声サービスに音声入力を処理させてもよい。ただし、音声入力のウェイクワード部分が特定の音声サービスに対応するウェイクワード（例えば、「ヘイ、Siri」）を含んでいる場合、NMDは、代わりに対応する音声サービスのみに音声入力を処理させる可能性がある。

20

【0129】

複数の音声サービスに音声入力を処理させた後、NMDはこれらの複数の音声サービスからそれぞれの結果を受信することができる。例えば、「テイラー・スウィフトの曲を再生して」という音声コマンドに対して、NMDは、第1音声サービス（例えば、ALEXA（登録商標））からテイラー・スウィフトのオーディオトラックを受信し、第2音声サービス（例えば、GOOGLE（登録商標））からテイラー・スウィフトに関連する検索結果を受信してもよい。コマンドはテイラー・スウィフトの曲を「再生する」ことであったので、NMDは第2音声サービスからの検索結果よりも、第1音声サービスからのオーディオトラックを選択してもよい。NMDは、キッチンゾーンに当該オーディオトラックの再生を開始させることによって、この結果を出力してもよい。

30

【0130】

別の例では、処理作業に関係する音声サービスは、特定の種類のコマンドに固有のものであってもよい。例えば、メディアストリーミングサービス（例えば、SPOTIFY（登録商標））は、オーディオ再生関連のコマンド用の音声サービスコンポーネントを有している可能性がある。ある例では、NMDは「天気はどう？」という音声入力を受信してもよい。この入力に対して、メディアストリーミングサービスの音声サービスは有用な結果を返さない可能性がある（例えば、ヌル結果又はエラー結果）。NMDは、別の音声サービスからの結果を選択する可能性がある。

40

【0131】

V. 音声サービスを登録するための例示的なシステムおよび方法

上述したように、ある例では、コンピューティングデバイスは、音声コマンドを処理するために1つ又は複数の音声サービスを登録することができる。実施形態900は、NMDに少なくとも1つの音声サービスを登録させるための例示的な技術である。

【0132】

a. 音声サービス（複数可）を登録するように命じるコマンドを示す入力データの受信

ブロック902で、実施形態900は、1つ又は複数の第2デバイスに1つ又は複数の音声サービスを登録するように命じるコマンドを示す、入力データを受信するステップを含む。例えば、第1デバイス（例えば、NMD）はユーザインタフェース（例えば、タッ

50

チスクリーン)を介して、1つ又は複数の再生デバイスを含むメディア再生システムに対して、1つ又は複数の音声サービスを登録するように命じるコマンドを示す入力データを受信してもよい。ある例では、NMDは、他の実行可能な実施形態の中でもとりわけ、実施形態700のブロック702に関連して上述した例示的な技術のいずれかを使用して、メディア再生システムを設定する手順の一部として当該入力を受信する。

【0133】

b. NMDに登録されている音声サービスの検出

ブロック904で、実施形態900は、第1デバイス(例えば、NMD)に登録されている1つ又は複数の音声サービスを検出するステップを含む。そのような音声サービスは、NMDにインストールされているか、又はNMDに固有の(例えば、NMDのオペレーティングシステムの一部である)音声サービスを含んでいてもよい。

10

【0134】

例えば、NMDがスマートフォン又はタブレットである場合、音声サービスとインタフェースする1つ又は複数のアプリケーション(「アプリ」)をインストールしている可能性がある。NMDは、任意の適切な技術を用いてこれらのアプリケーションを検出することができる。こうした技術は、NMDの製造業者又はオペレーティングシステムによって異なってもよい。ある例ではNMDは、インストールされているアプリケーションのリスト又はデータベースを、サポートされている音声サービスのリストと比較して、NMDにインストール中の音声サービスのうちでどのサービスがサポートされているかを決定してもよい。

20

【0135】

他の例では、音声サービスはNMDに固有のものであってもよい。例えば、アップル(登録商標)社およびグーグル(登録商標)社の音声サービスは、それぞれiOSおよびAndroidのオペレーティングシステムを実行しているデバイスに組み込まれているか、又はプレインストールされている場合がある。更に、これらのオペレーティングシステムにおいてカスタマイズされたいくつかのディストリビューション(例えば、アマゾン(登録商標)社のFireOS(登録商標))は、固有の音声サービス(例えば、ALEXA(登録商標))を含んでいる場合がある。

【0136】

c. 検出された音声サービス(複数可)のデバイスへの登録の実行

30

ブロック906で、実施形態900は、検出された音声サービスのうちの少なくとも1つを、1つ又は複数の第2デバイスに登録させるステップを含む。例えば、NMDは、検出された音声サービスのうちの少なくとも1つを、1つ又は複数の再生デバイスを含むメディア再生システム(例えば、図1のメディア再生システム100)に登録させてもよい。この音声サービスを登録させるステップは、ネットワークインタフェースを介して、その音声サービスに関する認証情報を示すメッセージをメディア再生システム(すなわち、その少なくとも1つのデバイス)に送信するステップを含んでもよい。このメッセージは、NMDからの認証情報を使用して、メディア再生システムに音声サービスを登録するように命じるコマンド、要求、又は他のクエリを更に含んでいてもよい。このようにして、ユーザのメディア再生システムには、ユーザのNMD(例えば、スマートフォン)に登録されているのと同じ音声サービスのうちの1つ又は複数が、ユーザのNMDと同じ認証情報を利用して登録されていてもよく、これによって登録作業を早めることができる。他の利点も想定可能である。

40

【0137】

VI. 結論

本明細書は、様々な例示のシステム、方法、装置、および製品などを開示しており、それらは、他のコンポーネントの中で、ハードウェア上で実行されるファームウェアおよび/又はソフトウェアを含む。そのような例は、単なる例示であり、限定されるものとみなすべきではないと理解される。例えば、これらのファームウェア、ハードウェア、および/又はソフトウェアの態様又はコンポーネントのいくつか又はすべてが、専らハードウェ

50

アに、専らソフトウェアに、専らファームウェアに、又はハードウェア、ソフトウェア、および／又はファームウェアの任意の組み合わせを実施することができることが意図されている。したがって、提供されているそれらの例は、それらのシステム、方法、装置、および／又は生産物を実施する唯一の方法ではない。

【0138】

(特徴1) マイクロフォンを介して、音声入力を示す音声データを受信するステップと、メディア再生システムに登録されている複数の音声サービスの中から、前記音声入力を処理するための音声サービスを識別するステップと、ネットワークインタフェースを介して、前記識別された音声サービスに前記音声入力を処理させるステップとを含む、方法。

【0139】

(特徴2) 前記音声入力を処理するための音声サービスを識別するステップは、受信した前記音声データの一部が特定の音声サービスに対応する特定のウェイクワードを表していると決定するステップと、前記音声入力を処理するための音声サービスとして、前記特定のウェイクワードに対応する前記特定の音声サービスを識別するステップであって、前記メディア再生システムに登録されている複数の音声サービスのそれぞれのウェイクワードに対応している、ステップとを含む、特徴1に記載の方法。

【0140】

(特徴3) 前記受信した音声データの一部が特定の音声サービスに対応する特定のウェイクワードを表していると決定するステップは、前記複数の音声サービスのそれぞれの音声サービスに対応するウェイクワード検出アルゴリズムに対して、前記受信した音声データを用いてクエリを実行するステップと、前記受信した音声データの一部が前記特定の音声サービスに対応する特定のウェイクワードを表していることを、前記特定の音声サービスのウェイクワード検出アルゴリズムが検出したと決定するステップとを含む、特徴2に記載の方法。

【0141】

(特徴4) 前記音声入力を処理するための音声サービスを識別するステップは、前記受信した音声データが、前記メディア再生システムに登録されている前記複数の音声サービスのうちの、所定の音声サービスに対応する任意のウェイクワードを含んでいないと決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、前記複数の音声サービスの中からデフォルトの音声サービスを識別するステップとを含む、特徴1に記載の方法。

【0142】

(特徴5) 前記音声入力を処理するための音声サービスを識別するステップは、(i) 以前の音声入力が特定の音声サービスによって処理されたこと、および(ii) 前記音声入力が、前記以前の音声入力受信後の閾値時間内に受信されていることを決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、前記以前の音声入力を処理した前記特定の音声サービスを識別するステップとを含む、特徴1に記載の方法。

【0143】

(特徴6) 前記音声入力を処理するための音声サービスを識別するステップは、(i) 以前の音声入力が特定の音声サービスによって処理されたこと、および(ii) 前記音声入力が前記以前の音声入力と同じ種類の動作を対象としていることを決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、前記以前の音声入力を処理した前記特定の音声サービスを識別するステップとを含む、特徴1に記載の方法。

【0144】

(特徴7) 前記音声入力を処理するための音声サービスを識別するステップは、前記音声入力がメディア再生コマンドを含んでいると決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、メディア再生コマンドを処理するように構成されている特定の音声サービスを識別するステップとを含む、特徴1に記載の

10

20

30

40

50

方法。

【0145】

(特徴8) 前記音声入力を処理するための音声サービスを識別するステップは、前記音声入力が無線照明デバイスを対象としていると決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、無線照明デバイスを対象とした音声入力を処理するように構成されている特定の音声サービスを識別するステップとを含む、特徴1に記載の方法。

【0146】

(特徴9) 前記音声入力を処理するための音声サービスを識別するステップは、前記受信した音声データの一部が、任意の特定の音声サービスに対応していない一般的ウェイクワードを表していると決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、前記複数の音声サービスの中からデフォルトの音声サービスを識別するステップとを含む、特徴1に記載の方法。

10

【0147】

(特徴10) 前記メディア再生システムは複数のゾーンを含み、前記音声入力を処理するための音声サービスを識別するステップは、前記音声入力が、前記複数のゾーンのうちの特定のゾーンを対象としていると決定するステップと、前記決定に基づいて、前記音声入力を処理するための音声サービスとして、前記メディア再生システムの前記特定のゾーンを対象とした音声入力を処理するように構成されている、特定の音声サービスを識別するステップとを含む、特徴1に記載の方法。

20

【0148】

(特徴11) 前記音声入力を処理するための音声サービスを識別するステップは、前記受信した音声データの一部が第1音声サービスに対応する特定のウェイクワードを表していると決定するステップと、前記音声入力を処理するにあたり、前記第1音声サービスが現在利用できないと決定するステップと、前記音声入力を処理するための音声サービスとして、前記第1音声サービスとは異なる第2音声サービスを識別するステップとを含む、特徴1に記載の方法。

【0149】

(特徴12) 前記音声入力は、ウェイクワードを表す第1部分と音声コマンドを表す第2部分とを含み、前記識別された音声サービスに前記音声入力を処理させるステップは、ネットワークインタフェースを介して、前記識別された音声サービスの1つ又は複数のサーバに対して、(i) 前記音声入力において少なくとも前記第2部分を示すデータ、および(ii) 前記音声コマンドを示すデータの処理を命じるコマンドを送信するステップを含む、特徴1に記載の方法。

30

【0150】

(特徴13) 1つ又は複数のプロセッサによって実行されて特徴1から12のいずれか一項に記載の方法を実行させる命令を記憶している、有形の非一時的なコンピュータ読取り可能媒体。

【0151】

(特徴14) 特徴1から12のいずれか一項に記載の方法を実行するように構成された、デバイス。

40

【0152】

(特徴15) 特徴1から12のいずれか一項に記載の方法を実行するように構成された、メディア再生システム。

【0153】

(特徴16)

ネットワーク化されたマイクロフォンデバイスは、(i) マイクロフォンと、(ii) ネットワークインタフェースと、(iii) 1つ又は複数のプロセッサと、(iv) 1つ又は複数のプロセッサによって実行されると、ネットワーク化されたマイクロフォンデバイスにある方法を実行させるように動作可能な命令を記憶した有形非一時的コンピュータ

50

可読媒体とを備え、当該方法は、（a）マイクロフォンを介して、音声入力を示す音声データを受信するステップ、（b）メディア再生システムに登録されている複数の音声サービスの中から、音声入力进行处理するための音声サービスを識別するステップ、（c）ネットワークインタフェースを介して、識別された音声サービスに音声入力进行处理させるステップ、を含む。

【0154】

（特徴17）

音声入力进行处理するための音声サービスを識別するステップは、（i）受信した音声データの一部分が特定の音声サービスに対応する特定のウェイクワードを表すものであることを決定するステップ、（ii）当該特定のウェイクワードに対応する特定の音声サービスを、音声入力进行处理するための音声サービスとして識別するステップ、ここで、メディア再生システムに登録されている複数の音声サービスの各々は各ウェイクワードに対応する、を含む、特徴16に記載のマイクロフォンデバイス。

10

【0155】

（特徴18）

受信した音声データの一部分が特定の音声サービスに対応する特定のウェイクワードを表すものであることを決定するステップは、（i）受信した音声データを用いて、複数の音声サービスの各々に対応するウェイクワード検知アルゴリズムに問合せするステップ、（ii）特定の音声サービスのウェイクワード検知アルゴリズムが、受信した音声データの一部分が当該特定の音声サービスに対応する特定のウェイクワードを表すものであることを検知したことを決定するステップ、を含む、特徴17に記載のマイクロフォンデバイス。

20

【0156】

（特徴19）

音声入力进行处理するための音声サービスを識別するステップは、（i）受信した音声データが、メディア再生システムに登録されている複数の音声サービスのうちの所定の音声サービスに対応する任意のウェイクワードを除外していることを決定するステップ、（ii）当該決定に基づいて、複数の音声サービスのうちのデフォルトの音声サービスを、音声入力进行处理するための音声サービスとして識別するステップを含む、特徴16に記載のマイクロフォンデバイス。

【0157】

（特徴20）

音声入力进行处理するための音声サービスを識別するステップは、（i）（a）前の音声入力が特定の音声サービスによって処理されたこと、（b）前の音声入力を受信してから閾値期間内に次の音声入力を受信されたこと、を決定するステップ、（ii）当該決定に基づいて、当該特定の音声サービスを次の音声入力进行处理するための音声サービスとして識別するステップを含む、特徴16に記載のマイクロフォンデバイス。

30

【0158】

（特徴21）

音声入力进行处理するための音声サービスを識別するステップは、（i）（a）前の音声入力が特定の音声サービスによって処理されたこと、（b）次の音声入力が前の音声入力と同じ種類の動作を対象としていること、を決定するステップ、（ii）当該決定に基づいて、当該特定の音声サービスを次の音声入力进行处理するための音声サービスとして識別するステップを含む、特徴16に記載のマイクロフォンデバイス。

40

【0159】

（特徴22）

音声入力进行处理するための音声サービスを識別するステップは、（i）音声入力がメディア再生コマンドを含むことを決定するステップ、（ii）当該決定に基づいて、メディア再生コマンドを処理するように構成された特定の音声サービスを、当該音声入力进行处理するための音声サービスとして識別するステップ、を含む、特徴16に記載のマイクロフォンデバイス。

50

【0160】

(特徴23)

音声入力処理するための音声サービスを識別するステップは、(i) 音声入力が無線照明デバイスを対象としていると決定するステップ、(ii) 当該決定に基づいて、無線照明デバイスを対象とした音声入力処理するように構成された特定の音声サービスを、音声入力処理するための音声サービスとして識別するステップ、を含む、特徴16に記載のマイクロフォンデバイス。

【0161】

(特徴24)

音声入力処理するための音声サービスを識別するステップは、(i) 受信した音声データの一部分が任意の音声サービスに対応しない一般的なウェイクワードを表すものであることを決定するステップ、(ii) 当該決定に基づいて、複数の音声サービスのうちのデフォルトの音声サービスを、音声入力処理するための音声サービスとして識別するステップ、を含む、特徴16に記載のマイクロフォンデバイス。

10

【0162】

(特徴25)

メディア再生システムは複数のゾーンを含み、音声入力処理するための音声サービスを識別するステップは、(i) 音声入力複数のゾーンのうちの特定のゾーンを対象としていることを決定するステップ、(ii) 当該決定に基づいて、当該特定のゾーンを対象とした音声入力処理するように構成された特定の音声サービスを、音声入力処理する音声サービスとして識別するステップ、を含む、特徴16に記載のマイクロフォンデバイス。

20

【0163】

(特徴26)

音声入力処理するための音声サービスを識別するステップは、(i) 受信した音声データが第1の音声サービスに対応する特定のウェイクワードを表すことを決定するステップ、(ii) 第1の音声サービスが音声入力処理のために現在利用可能でないことを決定するステップ、(iii) 第1の音声サービスとは異なる第2の音声サービスを、音声入力処理するための音声サービスとして識別するステップ、を含む、特徴16に記載のマイクロフォンデバイス。

30

【0164】

(特徴27)

音声入力は、ウェイクワードを表す第1の部分と、音声コマンドを表す第2の部分とを含み、識別された音声サービスに音声入力処理させるステップは、識別された音声サービスの1つ又は複数のサーバに対して、ネットワークインタフェースを介して、(i) 音声入力の少なくとも第2の部分を表すデータと、(ii) 当該データの処理を命じるコマンドとを送信するステップを含む、特徴16に記載のマイクロフォンデバイス。

【0165】

(特徴28)

有形非一時的コンピュータ可読媒体は、1つ又は複数のプロセッサによって実行されると、ネットワーク化されたマイクロフォンデバイスにある方法を実行させるように動作可能な命令を記憶したており、当該方法は、(i) マイクロフォンを介して、音声入力を示す音声データを受信するステップ、(ii) メディア再生システムに登録されている複数の音声サービスの中から、音声入力処理するための音声サービスを識別するステップ、(iii) ネットワークインタフェースを介して、識別された音声サービスに音声入力処理させるステップ、を含む。

40

【0166】

(特徴29)

音声入力処理するための音声サービスを識別するステップは、(i) 受信した音声データの一部分が特定の音声サービスに対応する特定のウェイクワードを表すことを決定する

50

ステップ、(i i) 当該特定のウェイクワードに対応する特定の音声サービスを、音声入力を処理するための音声サービスとして識別するステップ、ここで、メディア再生システムに登録されている複数の音声サービスの各々は各ウェイクワードに対応する、を含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

【0167】

(特徴 30)

受信した音声データの一部が特定の音声サービスに対応する特定のウェイクワードを表すものであることを決定するステップは、(i) 受信した音声データを用いて、複数の音声サービスの各々に対応するウェイクワード検知アルゴリズムに問合せするステップ、(i i) 特定の音声サービスのウェイクワード検知アルゴリズムが、受信した音声データの一部が当該特定の音声サービスに対応する特定のウェイクワードを表すものであることを検知したことを決定するステップ、を含む、特徴 29 に記載の有形非一時的コンピュータ可読媒体。

10

【0168】

(特徴 31)

音声入力を処理するための音声サービスを識別するステップは、(i) 受信した音声データが、メディア再生システムに登録されている複数の音声サービスのうちの所定の音声サービスに対応する任意のウェイクワードを除外していることを決定するステップ、(i i) 当該決定に基づいて、複数の音声サービスのうちのデフォルトの音声サービスを、音声入力を処理するための音声サービスとして識別するステップを含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

20

【0169】

(特徴 32)

音声入力を処理するための音声サービスを識別するステップは、(i) (a) 前の音声入力が特定の音声サービスによって処理されたこと、(b) 前の音声入力を受信されてから閾値期間内に次の音声入力を受信されたこと、を決定するステップ、(i i) 当該決定に基づいて、当該特定の音声サービスを次の音声入力を処理するための音声サービスとして識別するステップを含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

【0170】

(特徴 33)

音声入力を処理するための音声サービスを識別するステップは、(i) (a) 前の音声入力が特定の音声サービスによって処理されたこと、(b) 次の音声入力が前の音声入力と同じ種類の動作を対象としていること、を決定するステップ、(i i) 当該決定に基づいて、当該特定の音声サービスを次の音声入力を処理するための音声サービスとして識別するステップを含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

30

【0171】

(特徴 34)

音声入力を処理するための音声サービスを識別するステップは、(i) 音声入力メディア再生コマンドを含むことを決定するステップ、(i i) 当該決定に基づいて、メディア再生コマンドを処理するように構成された特定の音声サービスを、当該音声入力を処理するための音声サービスとして識別するステップ、を含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

40

【0172】

(特徴 35)

音声入力を処理するための音声サービスを識別するステップは、(i) 受信した音声データの一部が任意の音声サービスに対応しない一般的なウェイクワードを表すものであることを決定するステップ、(i i) 当該決定に基づいて、複数の音声サービスのうちのデフォルトの音声サービスを、音声入力を処理するための音声サービスとして識別するステップ、を含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

【0173】

50

(特徴 36)

メディア再生システムは複数のゾーンを含み、音声入力処理するための音声サービスを識別するステップは、(i) 音声入力処理が複数のゾーンのうちの特定のゾーンを対象としていることを決定するステップ、(ii) 当該決定に基づいて、当該特定のゾーンを対象とした音声入力処理を行うように構成された特定の音声サービスを、音声入力処理する音声サービスとして識別するステップ、を含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

【0174】

(特徴 37)

音声入力処理するための音声サービスを識別するステップは、(i) 受信した音声データが第 1 の音声サービスに対応する特定のウェイクワードを表すことを決定するステップ、(ii) 第 1 の音声サービスが音声入力処理のために現在利用可能でないことを決定するステップ、(iii) 第 1 の音声サービスとは異なる第 2 の音声サービスを、音声入力処理するための音声サービスとして識別するステップ、を含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

【0175】

(特徴 38)

音声入力は、ウェイクワードを表す第 1 の部分と、音声コマンドを表す第 2 の部分を含み、識別された音声サービスに音声入力処理させるステップは、識別された音声サービスの 1 つ又は複数のサーバに対して、ネットワークインタフェースを介して、(i) 音声入力の少なくとも第 2 の部分を表すデータと、(ii) 当該データの処理を命じるコマンドとを送信するステップを含む、特徴 28 に記載の有形非一時的コンピュータ可読媒体。

【0176】

(特徴 39)

(i) ネットワーク化されたマイクロフォンデバイスのマイクロフォンを介して、音声入力を示す音声データを受信するステップ、(ii) 受信した音声データの一部が、メディア再生システムに登録されている複数の音声サービスのうちの特定の音声サービスに対応する特定のウェイクワードを表すことを決定するステップ、ここで、メディア再生システムに登録されている複数の音声サービスの各々は各ウェイクワードに対応するものである、(iii) ネットワーク化されたマイクロフォンデバイスのネットワークインタフェースを介して、特定の音声サービスに音声入力処理させるステップ、ここで、特定の音声サービスに音声入力処理させるステップは、マイクロフォンデバイスのネットワークインタフェースを介して、音声入力を示すデータを特定の音声サービスの 1 つ又は複数のサーバに送信するステップを含む、方法。

【0177】

さらに、本明細書での「実施形態」への言及は、その実施形態に関連して記載される特定の特徵、構造、又は特性が、本発明の少なくとも 1 つの例示的な実施形態に含まれ得ることを意味する。明細書における様々な部分でのこのフレーズの使用は、同じ実施形態を必ずしも言及しておらず、また、他の実施形態と互いに排他的な別個の又は代替的な実施形態でもない。このように、本明細書に記載の実施形態を他の実施形態と組み合わせてもよいことが、明示的および暗示的に当業者によって理解される。

【0178】

本明細書は、例示的な環境、システム、手順、ステップ、論理ブロック、処理、および他のシンボル表現に関して広く示されており、それらは直接又は間接的にネットワークに接続されるデータ処理デバイスの動作に類似するものである。これらの処理説明および表現は、一般的に当業者によって使用され、それらの仕事の内容を他の当業者に最も効率良く伝えることができる。多くの具体的な内容が、本開示を理解するために提供されている。しかしながら、当業者にとって、本開示の特定の実施形態が特定の、具体的な詳細なしに実施され得ることは理解される。他の例では、周知の方法、手順、コンポーネント、お

よび回路が、実施形態を不必要に曖昧にすることを避けるため、詳細に説明していない。したがって、本開示の範囲は、上記した実施形態よりむしろ添付された特許請求の範囲によって定義される。

【0179】

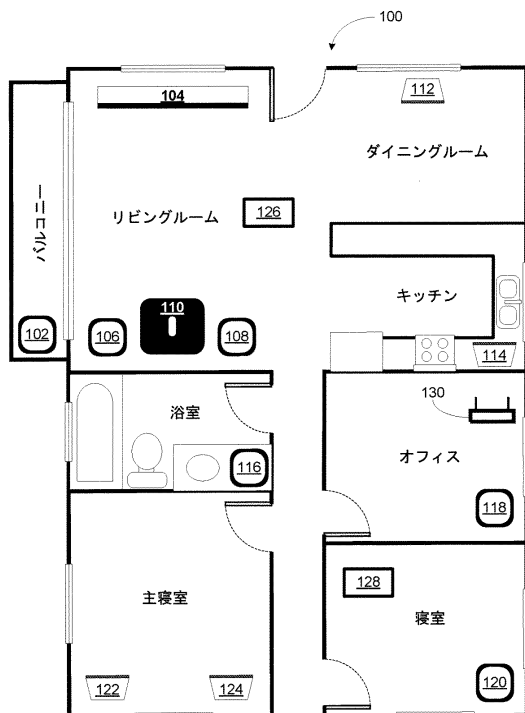
添付の特許請求の範囲のいずれかが単にソフトウェアおよび／又はファームウェアへの実装をカバーするように読み取ると、少なくとも1つの例における要素の1つ又は複数は、本明細書では、ソフトウェアおよび／又はファームウェアを記憶する有形の非一時的な記憶媒体、例えば、メモリ、DVD、CD、Blu-ray（登録商標）等を含むことが明確に定められている。

【要約】

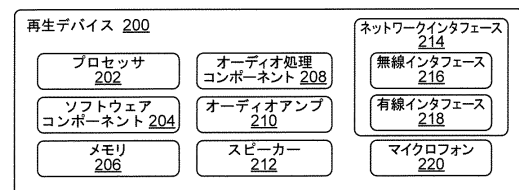
本明細書には、音声入力进行处理するための音声サービスを識別する例示的な技術が開示される。ある実施例は、NMDがマイクロフォンを介して音声入力を示す音声データを受信するステップを含む。NMDは、メディア再生システムに登録された複数の音声サービスの中から、音声入力进行处理するための音声サービスを識別し、ネットワークインタフェースを介して、識別した音声サービスに音声入力进行处理させる。

10

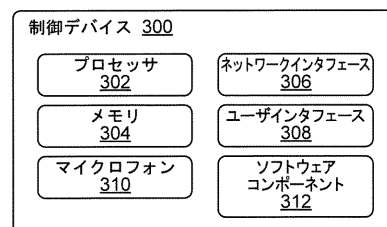
【図1】



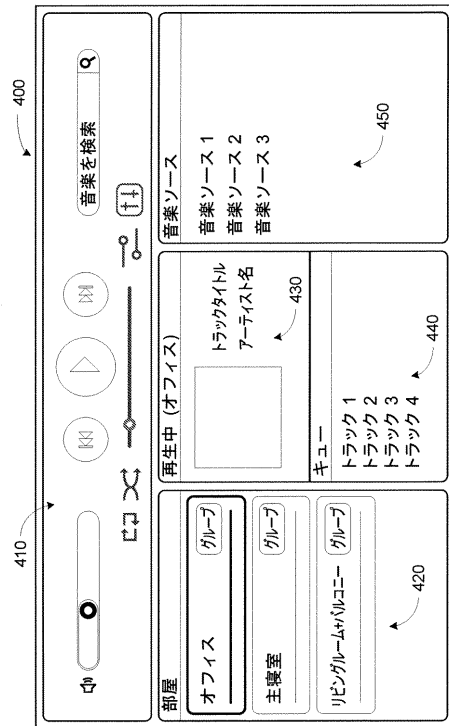
【図2】



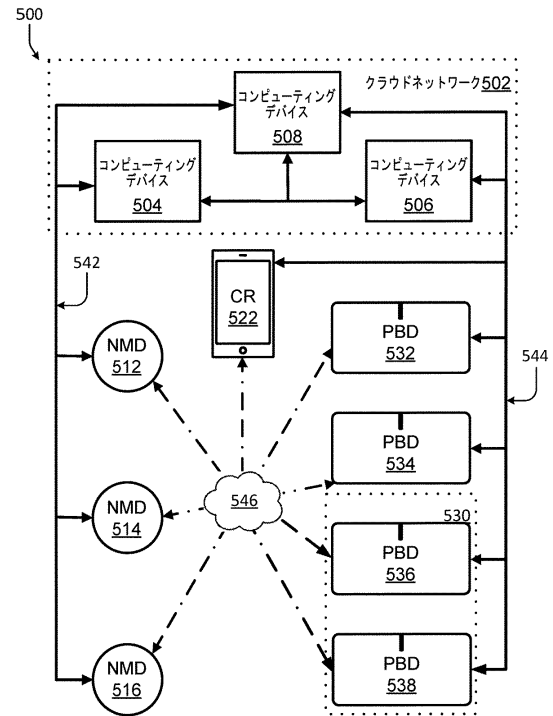
【図3】



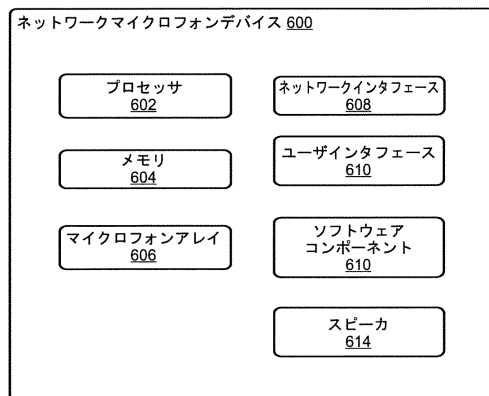
【図 4】



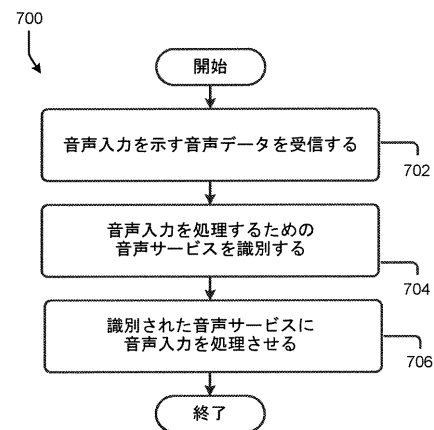
【図 5】



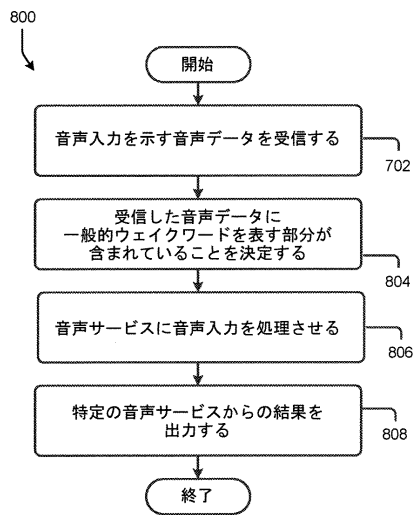
【図 6】



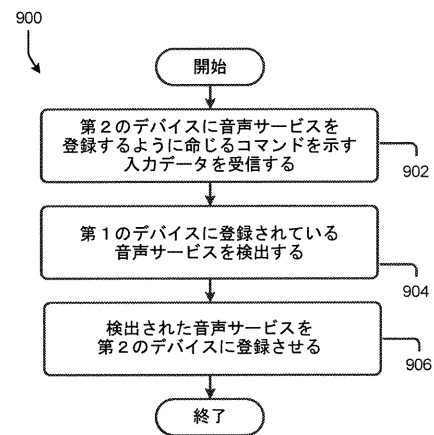
【図 7】



【図 8】



【図 9】



フロントページの続き

(72)発明者 デイン・ウィルバーディング
アメリカ合衆国93101カリフォルニア州サンタ・バーバラ、チャパラ・ストリート614番、
ソノズ・インコーポレイテッド内

審査官 千本 潤介

(56)参考文献 特開2016-151608 (JP, A)
特開2002-116797 (JP, A)
特開平11-146071 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G10L 15/00-15/34