

(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7436077号
(P7436077)

(45)発行日 令和6年2月21日(2024.2.21)

(24)登録日 令和6年2月13日(2024.2.13)

(51)国際特許分類 F I
G 1 0 L 15/10 5 0 0 T
G 1 0 L 15/10 2 0 0 W

請求項の数 9 (全13頁)

(21)出願番号	特願2022-540758(P2022-540758)	(73)特許権者	520235782
(86)(22)出願日	令和2年10月26日(2020.10.26)		エーアイスピーチカンパニーリミテッド
(65)公表番号	特表2023-506087(P2023-506087A)		AI SPEECH CO., LTD.
(43)公表日	令和5年2月14日(2023.2.14)		中華人民共和國江蘇省蘇州市蘇州工業園
(86)国際出願番号	PCT/CN2020/123643		區新平街388號騰飛科技園14幢
(87)国際公開番号	WO2021/135561		No. 14 Building, Ascendas IHub, No. 388 X
(87)国際公開日	令和3年7月8日(2021.7.8)		inping Street, Suzhou Industrial Park
審査請求日	令和5年1月19日(2023.1.19)		Suzhou, Jiangsu 215123 (CN)
(31)優先権主張番号	201911422397.2	(74)代理人	100137338
(32)優先日	令和1年12月31日(2019.12.31)		弁理士 辻田 朋子
(33)優先権主張国・地域又は機関	中国(CN)	(74)代理人	100224719
早期審査対象出願			

最終頁に続く

(54)【発明の名称】 スキルの音声ウェイクアップ方法および装置

(57)【特許請求の範囲】

【請求項1】

電子デバイスに適用されているスキルの音声ウェイクアップ方法であって、前記方法は、
処理する音声要求メッセージに対応するウェイクアップテキスト情報を識別するステップと、

業務スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定し、知識スキルセマンティックモデルを呼び出してウェイクアップテキスト情報に対応する知識応答回答および対応する第2信頼度を決定するステップと、

第1信頼度および第2信頼度に基づいて、ウェイクアップ知識スキルと目標業務分野に対応する目標業務スキルの1つを選択するステップとを含み、

前記業務スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定するステップは、

前記ウェイクアップテキスト情報中のウェイクアップ業務キーワードおよびウェイクアップ業務実體情報を抽出するステップと、

複数の業務キーワードおよび対応する業務実體情報を含む業務実體データベースに前記ウェイクアップ業務キーワードおよび対応する前記ウェイクアップ業務実體情報が存在するか否かを判定するステップと、

前記業務実體データベースに前記ウェイクアップ業務キーワードまたは前記対応する前記ウェイクアップ業務実體情報が存在しない場合、前記ウェイクアップ業務キーワードに対

10

20

応する業務関連度情報を取得し、前記ウェイクアップ業務キーワードおよび前記業務関連度情報を前記業務スキルセマンティックモデルに提供して、前記目標業務分野および対応する第1信頼度を決定するステップと、を含む、スキルの音声ウェイクアップ方法。

【請求項2】

前記業務実体データベースに前記ウェイクアップ業務キーワードおよび前記対応する前記ウェイクアップ業務実体情報が存在する場合、前記ウェイクアップ業務キーワードを前記業務スキルセマンティックモデルに提供して、前記目標業務分野および対応する第1信頼度を決定するステップ、をさらに含む、請求項1に記載の方法。

【請求項3】

前記ウェイクアップ業務キーワードに対応する前記業務関連度情報を取得するステップは、

10

検索エンジンに基づいて前記ウェイクアップ業務キーワードに対応する業務ホットネス情報および検索結果業務関連指標を決定するステップと、

前記業務ホットネス情報および前記検索結果業務関連指標に基づいて、前記業務関連度情報を決定するステップと、を含む、請求項2に記載の方法。

【請求項4】

前記検索エンジンに基づいて前記ウェイクアップ業務キーワードに対応する前記業務ホットネス情報および前記検索結果業務関連指標を決定するステップは、

前記検索エンジンに基づいて前記ウェイクアップ業務キーワードに対応する検索結果を決定するステップと、

20

予め設定された検索結果評価策略に基づいて、前記検索結果に対応する前記検索結果業務関連指標を決定するステップと、を含む、請求項3に記載の方法。

【請求項5】

前記検索エンジンに基づいて前記ウェイクアップ業務キーワードに対応する検索結果を決定するステップは、

前記ウェイクアップ業務キーワードを前記検索エンジンに提供して、対応する第1検索結果を決定するステップと、

前記ウェイクアップ業務キーワードおよび前記目標業務分野に対応する業務名称を前記検索エンジンに提供して、対応する第2検索結果を決定するステップと、を含む、請求項4に記載の方法。

30

【請求項6】

前記目標業務スキルは音楽スキルを含む、請求項1に記載の方法。

【請求項7】

電子デバイスに適用されているスキル音声ウェイクアップ装置であって、前記装置は、
処理する音声要求メッセージに対応するウェイクアップテキスト情報を識別するように構成された音声識別ユニットと、

業務スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定し、知識スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する知識応答回答および対応する第2信頼度を決定するように構成されたモデル呼び出しユニットと、

40

前記第1信頼度および前記第2信頼度に基づいて、ウェイクアップ知識スキルおよび前記目標業務分野に対応する目標業務スキルの1つを選択するように構成されたスキルウェイクアップユニットと、を含む、

前記業務スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定するステップは、

前記ウェイクアップテキスト情報中のウェイクアップ業務キーワードおよびウェイクアップ業務実体情報を抽出するステップと、

複数の業務キーワードおよび対応する業務実体情報を含む業務実体データベースに前記ウェイクアップ業務キーワードおよび対応する前記ウェイクアップ業務実体情報が存在するか否かを判定するステップと、

50

前記業務実体データベースに前記ウェイクアップ業務キーワードまたは前記対応する前記ウェイクアップ業務実体情報が存在しない場合、前記ウェイクアップ業務キーワードに対応する業務関連度情報を取得し、前記ウェイクアップ業務キーワードおよび前記業務関連度情報を前記業務スキルセマンティックモデルに提供して、前記目標業務分野および対応する第1信頼度を決定するステップと、を含む、スキル音声ウェイクアップ装置。

【請求項8】

少なくとも1つのプロセッサ、および前記少なくとも1つのプロセッサに通信可能に接続されたメモリを含み、前記メモリに前記少なくとも1つのプロセッサによって実行され得る指令が記憶され、前記指令は前記少なくとも1つのプロセッサによって実行されると、前記少なくとも1つのプロセッサは請求項1から6のいずれか1項に記載の方法のステップを実施する、電子デバイス。

10

【請求項9】

プロセッサによって実行されると請求項1から6のいずれか1項に記載の方法のステップが実施されるコンピュータプログラムが記憶された記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、インターネットの技術分野に属し、特にスキルの音声ウェイクアップ方法および装置に関する。

【背景技術】

20

【0002】

音声技術や人工知能技術の継続的な発展により、音声ウェイクアップ技術は、スマートホームなどのスマートデバイスの分野で大きな発展を遂げた。

【0003】

現在、スマートデバイスには、知識スキルや業務スキルが存在する。知識スキルは、スマートデバイスのユーザーに、個々のユーザーの質問に対する回答を提供する機能など、対応する知識ベースの質問と回答のサービスを提供するために使用される。業務スキルは、音楽サービス、タクシーサービス、天気予報サービスなど、スマートデバイスのユーザーに業務サービスを提供するために使用される。

【0004】

30

しかし、一般的なスマートデバイスでは、ユーザーの音声を識別する際、その音声が業務スキルを呼び出そうとしているのか、知識スキルを呼び出そうとしているのかを識別できず、誤ったスキルを呼び出してしまうことがある。例えば、スマートスピーカーがユーザーの音声メッセージ「Li Chenのお母さんは誰」を受信し、音楽スキルをウェイクアップさせると、歌手「Li Chen」の楽曲「Who is mother」が再生されてしまう。また、知識スキルをウェイクアップさせると、「Li Chenの母親は誰」という回答が再生されてしまう。

【0005】

上記の問題に対して、業界内には良い解決策がまだない。

【発明の概要】

40

【0006】

本発明の実施例は、少なくとも上記の技術的問題の1つを解決するためのスキルの音声ウェイクアップ方法および装置を提供する。

【0007】

第1側面によれば、本発明の実施例は電子デバイスに適用されているスキルの音声ウェイクアップ方法を提供し、前記方法は、処理する音声要求メッセージに対応するウェイクアップテキスト情報を識別するステップと、業務スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定し、知識スキルセマンティックモデルを呼び出してウェイクアップテキスト情報に対応する知識応答回答および対応する第2信頼度を決定するステップと、第1信頼度および

50

第2信頼度に基づいて、ウェイクアップ知識スキルと目標業務分野に対応する目標業務スキルの1つを選択するステップとを含む。

【0008】

第2側面によれば、本発明の実施例は、電子デバイスに適用されているスキル音声ウェイクアップ装置を提供し、この装置は、処理する音声要求メッセージに対応するウェイクアップテキスト情報を識別するように構成された音声識別ユニットと、業務スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定し、知識スキルセマンティックモデルを呼び出して前記ウェイクアップテキスト情報に対応する知識応答回答および対応する第2信頼度を決定するように構成されたモデル呼び出しユニットと、前記第1信頼度および前記第2信頼度に基づいて、ウェイクアップ知識スキルおよび前記目標業務分野に対応する目標業務スキルの1つを選択するように構成されたスキルウェイクアップユニットと、を含む。

10

【0009】

第3側面によれば、本発明の実施例は、電子デバイスを提供し、少なくとも1つのプロセッサ、および前記少なくとも1つのプロセッサに通信可能に接続されたメモリを含み、前記メモリに前記少なくとも1つのプロセッサによって実行され得る指令が記憶され、前記指令は前記少なくとも1つのプロセッサによって実行されると、前記少なくとも1つのプロセッサは上記の方法のステップを実施する。

【0010】

第4側面によれば、本発明の実施例は、プロセッサによって実行されると上記の方法のステップが実施されるコンピュータプログラムが記憶された記憶媒体を提供する。

20

【発明の効果】

【0011】

本発明の実施例は以下の有益な効果を有する。音声要求メッセージを受信すると、業務スキルセマンティックモデルと知識スキルセマンティックモデルを用いて、対応する業務分野と応答回答を並行して決定し、対応する信頼度を出力し、これにより、知識スキルまたは目標業務スキルへのウェイクアップを可能にすることである。その結果、業務スキルと知識スキルに対する音声メッセージのマッチングを比較し、音声メッセージに基づくスキルを誤ってウェイクアップさせる確率を低減させることができる。

【図面の簡単な説明】

30

【0012】

本発明の実施例の技術的解決策をより明確に説明するために、以下、実施例の説明において必要な図面を簡単に説明する。説明された図面は本発明のいくつかの実施形態であり、当業者は、創造的な努力を払うことない下で、他の図面を得ることができる。

【図1】本発明の実施例によるスキルの音声ウェイクアップ方法の一例を示すフローチャートである。

【図2】本発明の実施例による業務スキルセマンティックモデルを呼び出して実行される操作例を示すフローチャートである。

【図3】本発明の実施例による業務関連度情報を決定する操作例を示すフローチャートである。

40

【図4】本発明の実施例による音楽スキルの音声ウェイクアップ方法の一例を示す原理フローチャートである。

【図5】本発明の実施例によるスキル音声ウェイクアップ装置の一例を示す構造ブロック図である。

【発明を実施するための形態】

【0013】

従来技術の問題点を解決するために、本発明は、電子デバイスに適用されているスキルの音声ウェイクアップ方法および装置を提供し、この電子デバイスは、端末デバイスまたはサーバー側であり得る。その中で、端末デバイス例えば、スマートステレオ、自動車、スマートテレビ、スマートフォン、タブレット、スマートウォッチなど、人間とコンピュ

50

ータの音声対話機能を有する電子機器であり、本発明では制限されなく、サーバー側は、端末装置が人間とコンピュータの音声対話を実現するための技術サポートを提供するサーバー装置であり得る。もちろん、説明される実施形態は本発明の一部の実施形態にすぎず、本発明の全部の実施形態ではない。本発明の実施形態に基づいて、当業者が創造的な努力なしに得られる他の全ての実施形態も本発明の保護範囲に含まれるべきである。

【0014】

なお、本願の実施形態と実施形態の特徴とは、互いに矛盾することなければ、組み合わせることができる。

【0015】

本発明は、例えば、プログラムモジュールなどの、コンピューターによって実行されるコンピューター実行可能命令の一般的なコンテキストで説明することができる。プログラムモジュールには、通常、特定のタスクを実行したり、特定の抽象データ型を実現したりするルーチン、プログラム、オブジェクト、コンポーネント、データ構造などが含まれる。本発明は、分散コンピューティング環境で実施することもできる。これらの分散コンピューティング環境では、通信ネットワークを介して接続されたりリモート処理デバイスによって、タスクが実行される。分散コンピューティング環境では、プログラムモジュールは、ストレージデバイスを含むローカル及びリモートのコンピューターストレージメディアに配置できる。

10

【0016】

本発明では、「モジュール」、「システム」などは、例えば、ハードウェア、ハードウェアとソフトウェアの組み合わせ、ソフトウェア、又は実行中のソフトウェアなどのコンピューターに適用される関連エンティティを指す。具体的には、コンポーネントは、プロセッサで実行するプロセス、プロセッサ、オブジェクト、実行可能なコンポーネント、実行スレッド、プログラム、及び/又はコンピューターであるが、これらに限定されない。また、サーバーで実行するアプリケーションプログラムやスクリプトプログラム、サーバーがコンポーネントと呼ばれることもできる。実行のプロセス及び/又はスレッドには、一つ又は複数のコンポーネントを含んでもよく、且つ、コンポーネントは、一台のコンピューターにローカライズされ、及び/又は二台以上のコンピューターの間分布され、さまざまなコンピューター可読媒体で実行することができる。コンポーネントは、一つ以上のデータパケットを有する信号によって、ローカル及び/又はリモートプロセスを介して、通信することができる。信号は、例えば、ローカルシステム、分散システムにおけるもう一つのコンポーネントと相互作用するデータ、及び/又はインターネットのようなネットワークで信号を介して他のシステムと相互作用するデータからの信号である。

20

30

【0017】

最後に、用語「含む」、「有する」及びそれらのいずれの変更は、それらの要素だけでなく、明示的にリストされていない他の要素、又はそのようなプロセス、方法、オブジェクト、又は機器に固有の要素も含む。これ以上の制限が課されない場合、「含む」、「有する」及びそれらのいずれの変更によって定義された要素は、その要素を含むプロセス、方法、オブジェクト、又は機器に他の同一要素があることを除外しない。

【0018】

図1は、本発明の実施例によるスキルの音声ウェイクアップ方法の一例を示すフローチャートである。

40

【0019】

図1に示すように、ステップ110では、電子デバイスは、処理する音声要求メッセージに対応するウェイクアップテキスト情報を識別する。音声要求メッセージは、スマート音声デバイスのマイクキャプチャデバイスを介してキャプチャすることで取得することができる。さらに、様々な音声識別技術によって音声要求メッセージに対応するウェイクアップテキスト情報を決定することもでき、ここで制限されない。

【0020】

ステップ120では、電子デバイスは、業務スキルセマンティックモデルを呼び出して

50

ウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定し、知識スキルセマンティックモデルを呼び出してウェイクアップテキスト情報に対応する知識応答回答および対応する第2信頼度を決定する。例えば、業務スキルセマンティックモデルと知識スキルセマンティックモデルを並列に呼び出して、同期に予測を行い、対応する予測結果と信頼度を出力することができる。さらに、業務スキルセマンティックモデルは、業務分野タグセットでセマンティックを訓練し、知識スキルセマンティックモデルは知識問答タグセットでセマンティックを訓練し、様々な訓練方法を使用でき、ここで制限されない。

【0021】

ステップ130では、電子デバイスは、第1信頼度および第2信頼度に基づいて、ウェイクアップ知識スキルと目標業務分野に対応する目標業務スキルの1つを選択する。例えば、第1信頼度が第2信頼度よりも大きい場合、ウェイクアップ目標業務スキルを選択し、第1信頼度が第2信頼度以下である場合、ウェイクアップ知識スキルを選択することができる。信頼度は、予測確率を示し、0～1の数値を取る。したがって、本実施例によれば、音楽スキルと知識スキル間でのユーザ音声意図の予測確率を比較することができ、スキルが誤ってウェイクアップされる確率を低減することができる。

10

【0022】

図2は、本発明の実施例による業務スキルセマンティックモデルを呼び出して実行される操作の一例を示すフローチャートである。

【0023】

図2に示すように、ステップ210では、電子デバイスは、ウェイクアップテキスト情報中のウェイクアップ業務キーワードとウェイクアップ業務実體情報を抽出する。ウェイクアップ業務キーワードは、ウェイクアップテキスト情報において業務属性を有するキーワードであり、ウェイクアップ業務実體情報は、ウェイクアップテキスト情報において実體属性を有する単語であり得る。例えば、ウェイクアップ業務キーワードは、「追憶の夜」、ウェイクアップ業務実體情報は歌手「李谷一」とすることができる。さらに、様々なキーワード抽出モデルを用いて上記のウェイクアップ業務キーワードとウェイクアップ業務実體情報を抽出することができ、ここで制限されない。

20

【0024】

ステップ220では、電子デバイスは、業務実體データベースにウェイクアップ業務キーワードおよび対応するウェイクアップ業務実體情報が存在するか否かを判定する。業務実體データベースは、複数の業務キーワードおよび対応する業務実體情報を含み、1つの業務キーワードは複数の業務実體情報に対応することができる。上記の例には、業務実體データベースに「追憶の夜」に対応する例えば「李谷一」、「董文華」、「張也」などの数人の歌手が記憶されている。さらに、業務実體データベース中のデータ情報は、例えば、どの歌手が同じタイトルの歌を演奏したか、どの俳優が同じタイトルの映画を演奏したかといった業務キーワードと業務実體間の関係を反映するように予め収集され配置され得る。

30

【0025】

ステップ220中の業務実體データベースにウェイクアップ業務キーワードおよび対応するウェイクアップ業務実體情報が存在する場合、ステップ231に移行する。さらに、ステップ220中の業務実體データベースにウェイクアップ業務キーワードおよび対応するウェイクアップ業務実體情報が存在しない場合、ステップ232に移行する。

40

【0026】

ステップ231では、電子デバイスは、ウェイクアップ業務キーワードを業務スキルセマンティックモデルに提供して、目標業務分野および対応する第1信頼度を決定する。上記の例には、ウェイクアップ業務キーワードおよび対応するウェイクアップ業務実體情報はそれぞれ「追憶の夜」と「李谷一」である場合、「追憶の夜」を直接に業務スキルセマンティックモデルに提供して予測操作を行うことができる。

【0027】

ステップ232では、電子デバイスは、ウェイクアップ業務キーワードに対応する業務

50

関連度情報を取得する。例えば、ウェイクアップ業務キーワードを業務関連度分析ツールに提供し、業務関連度分析ツールから対応する業務関連度情報を取得する。さらに、以下に詳述するように、ウェイクアップ業務キーワードに対応するホットネス情報などを分析することで対応する業務関連度情報を決定することもできる。

【0028】

ステップ240では、電子デバイスはウェイクアップ業務キーワードと業務関連度情報を業務スキルセマンティックモデルに提供して、目標業務分野および対応する第1信頼度を決定する。このとき、業務スキルセマンティックモデルの入力は、ウェイクアップ業務キーワードに加え、業務関連度情報を有し、つまり、業務関連度情報は業務スキルセマンティックの予測過程に影響を与え、決定された目標業務分野に対する第1信頼度の結果が高い精度を有する。

10

【0029】

図3は、本発明の実施例による業務関連度情報の決定操作の一例を示すフローチャートである。

【0030】

ステップ310では、電子デバイスは、検索エンジンに基づいて、ウェイクアップ業務キーワードに対応する業務ホットネス情報と検索結果業務関連指標を決定する。例示的に、ウェイクアップ業務キーワードに基づいて、検索エンジンを呼び出し、この検索エンジンから上記の業務ホットネス情報と検索結果業務関連指標を取得し、または、この検索エンジンから検索結果を取得して分析することで対応する業務ホットネス情報と検索結果業務関連指標を決定する。検索結果業務関連指標は、検索エンジンで決定された検索結果と業務間の関連度合いを反映する。

20

【0031】

いくつかの実施形態では、検索エンジンに基づいて、ウェイクアップ業務キーワードに対応する検索結果を決定する。さらに、予め設定された検索結果評価策略に基づいて、この検索結果に対応する検索結果業務関連指標を決定する。例えば、上位ランキングの所定数（例えば10個）の検索結果を用いて目標業務分野の関連度を評価し、検索結果業務関連指標は、強い関連、一般的な関連または弱い関連性の多段指標を有し得る。

【0032】

業務関連指標の精度を確保するために、業務関連指標を特定するための検索を複数回行う検索バリエーションが実行されることもある。具体的に、ウェイクアップ業務キーワードを検索エンジンに提供して、対応する第1検索結果を決定し、ウェイクアップ業務キーワードと目標業務分野に対応する業務名称を検索エンジンに提供して、対応する第2検索結果を決定することができる。さらに、検索結果評価策略により、第1検索結果および第2検索結果を評価し、対応する検索結果業務関連指標を決定し、例えば、第1検索結果および第2検索結果と業務間の関連性を総合的に考慮する。

30

【0033】

ステップ320では、電子デバイスは、業務ホットネス情報と検索結果業務関連指標に基づいて、業務関連度情報を決定する。例えば、業務関連度情報は、業務ホットネス情報と検索結果業務関連指標を含み、または業務ホットネス情報と検索結果業務関連指標は業務関連度情報に対して重み付け構成がある。

40

【0034】

図4は、本発明の実施例による音楽スキルの音声ウェイクアップ方法の一例を示す原理フローチャートである。

【0035】

本実施例の業務スキルは、様々な業務に特化することが可能であり、以下の実施例では音楽スキルのみを例示して説明する。

【0036】

なお、音楽分野では、「曲名+歌手名」のような正規マッチングに基づくウェイクアップ方法が存在し、多くのスピーカー製品で採用されている。例えば、ユーザが直接に「劉

50

徳華の愛情忘却水」と言うと、「愛情忘却水」という曲を直接に再生することができる。拡張しやすいために、「劉徳華」に対応する曲名シソーラスが関連付けられ、「愛情忘却水」にも対応する歌手名シソーラスが関連付けられ、2つのシソーラスに多くの情報があり、業務スキルと知識スキルを誤って呼び出しやすい状況がある。

【0037】

現在の関連技術では、一般的に曲名または歌手名シソーラスの対応語句を直接削除するのが普通である。しかしながら、そうすると指定された曲名または歌手名を実際に言うと、セマンティック解析に失敗することがある。また、ある応用シナリオによって、ユーザが間違った歌手情報を言った場合でも、曲情報に基づいて正確な音楽を再生することも期待し、例えば、ユーザが「劉徳華の Kiss Goodbye」を言った場合、マッチングされないが、Kiss Goodbyeという曲を再生できる。

10

【0038】

本実施例では、クローラを用いて音楽曲名のホットネス情報や検索情報を取得し、このように、セマンティック解析において曲名にホットネス情報、検索情報などおよび信頼度情報を付与して解析するようにことができる。ここでは、検索情報の取得過程について、「曲名」を直接に検索エンジンに入れて第1語彙が音楽用語かどうかを判定し、第1語彙が音楽用語でない場合、検索エンジンは「曲名+歌曲」（例えば、「歌曲 Kiss Goodbye」）を検索し第1語彙が音楽用語かどうかを判定する。したがって、検索情報は様々な検索結果関連状態にある。

【0039】

また、本実施例では、音楽知識ベースを自己構築し、曲名から対応するすべての歌手リストを検査することができる。

20

【0040】

さらに、セマンティック解析によりセマンティックスロットに曲名+歌手名のみがあると解析した場合、音楽知識ベースを検査し、マッチングかどうかを確認する。マッチングの場合、タスク型スキル信頼度と知識型スキル信頼度を比較する。また、マッチングしない場合、歌曲ホットネス情報、検索情報および信頼度情報と併せて新しい信頼度を再び算出した後、知識型スキル信頼度とタスク型スキル信頼度を再度比較する。

【0041】

図4に示すフローでは、テキストが入力された後、タスク型スキルセマンティック解析と知識型スキルを並列にディスパッチし、タスク型スキルは複数分野のセマンティック解析結果（slotセマンティックスロット情報と信頼度情報を含み、セマンティックスロットが曲名である場合、ホットネスと検索情報を含む）を返答し、知識型スキルは回答結果および信頼度などの情報を返答する。

30

【0042】

さらに、両方から結果を得た後、返答されたタスク型が音楽分野を含んでいるかどうかを判定する（1文のタスク型が複数分野の解析結果を返答することもある）。

【0043】

次に、返答された業務分野に音楽分野が含まれる場合、解析されたセマンティックスロットは単純の「曲名+歌手名」であるかどうかを判定する。

40

【0044】

そして、「曲名+歌手名」である場合、音楽知識ベースを呼び出して、曲名と歌手名がマッチングするかどうかを判定する。

【0045】

その後、曲名+歌手名がマッチングする場合、タスク型セマンティック解析結果と知識型スキルスケジューリング融合モジュールを使用する。

【0046】

次に、曲名+歌手名がマッチングしない場合、タスク型スキル解析信頼度を再び算出し（検索情報、ホットネス情報、信頼度情報と併せて）、融合モジュールを呼び出して、タスク型または知識型を選択する。タスク型を選択すると、複数のタスク型スキルから融合

50

アルゴリズムによって1つを選択し、知識型スキルを選択すると、知識型ベースのデータを融合モジュールでプロトコルを従って直接組織して返す。

【0047】

表1は、本発明の実施例の音楽スキルの音声ウェイクアップ方法の使用前後の実験データテーブルを示す。

表1

方法	分野	合計	TP	FP	TN	FN	recall	precision	accuracy	F値
最適化前	音楽	1	8	3	9	5	0.93	0.69	0.69	0.80
		3	2	5	4	2	9570	7152	1021	0408
		7	2	7	9	9	4820	0596	3919	7193
		5 3	5	3			6534	7113	8554	4605
最適化後	音楽	1	8	2	2	7	0.91	0.79	0.78	0.85
		3	0	1	4	0	9236	2885	8415	1399
		7	4	0	2	7	9202	9986	1852	2487
		5 3	7	2	0		6502	2055	9678	9649

10

20

【0048】

表1に示すように、TP：正類、タスク型スキルに当たり、TN：負類、知識型スキルに当たり、FP：負類、正類として判定し、FN：正類、負類として判定し、recall：再現率($TP / (TP + FN)$)、precision：精度($TP / (TP + FP)$)、accuracy：正確さ($(TP + TN) / (TP + FP + TN + FN)$)、F値： $(2 * precision * recall / (precision + recall))$ 。

【0049】

最適化前後を比較すると、F値が5%向上したことが容易に分かる。さらに、融合アルゴリズムをさらに調整したり、case by caseを最適化したりすれば、より良質な最適化効果が得られる。

30

【0050】

図5は、本発明の実施例のスキル音声ウェイクアップ装置の一例を示す構造ブロック図である。

【0051】

図5に示すように、スキル音声ウェイクアップ装置500は、音声識別ユニット510、モデル呼び出しユニット520およびスキルウェイクアップユニット530を含む。

【0052】

音声識別ユニット510は、処理する音声要求メッセージに対応するウェイクアップテキスト情報を識別するように構成される。音声識別ユニット510の操作は、上記の図1のステップ110の説明を参照すればよい。

40

【0053】

モデル呼び出しユニット520は、業務スキルセマンティックモデルを呼び出してウェイクアップテキスト情報に対応する目標業務分野および対応する第1信頼度を決定し、知識スキルセマンティックモデルを呼び出してウェイクアップテキスト情報に対応する知識応答回答および対応する第2信頼度を決定するように構成される。モデル呼び出しユニット520の操作は、上記の図1のステップ120の説明を参照すればよい。

【0054】

スキルウェイクアップユニット530は、第1信頼度および第2信頼度に基づいて、ウェイクアップ知識スキルと目標業務分野に対応する目標業務スキルの1つを選択するよう

50

に構成される。スキルウェイクアップユニット 530 の操作は、上記の図 1 のステップ 130 の説明を参照すればよい。

【0055】

本発明の実施例の装置は、本発明の対応の方法の実施例を実施するために使用され、上記の本発明の方法の実施例の技術的効果を達成することができ、ここで繰り返さない。

【0056】

一方、本発明の実施例は、プロセッサによって実行されると上記のスキルの音声ウェイクアップ方法のステップを実施するコンピュータプログラムが記憶された記憶媒体を提供する。

【0057】

上記製品は、本発明の実施例に係る方法を実行可能であり、方法を実行する関連機能モジュール及び有利な作用効果を有する。本実施例において詳しく記述されていない技術的詳細は、本発明の実施例に係る方法を参照可能である。

【0058】

本発明の実施形態の電子デバイスは、以下のデバイスを含むが、これらに限定されない様々な形態で存在する。

【0059】

(1) モバイル通信デバイス：これらのデバイスは、その特徴がモバイル通信機能を備えることであり、音声及びデータ通信を提供することを主な目標として、スマートフォン（例えば、iPhone）、マルチメディア携帯電話、機能携帯電話、ローエンド携帯電話などを含む。

【0060】

(2) ウルトラモバイルパソコンデバイス：これらのデバイスは、パソコンのカテゴリに属し、計算及び処理の機能を持ち、一般にモバイルインターネットアクセスの特性を持って、例えば、iPad などの PDA、MID 及び UMPC デバイスなどを含む。

【0061】

(3) ポータブルエンターテイメントデバイス：これらのデバイスは、マルチメディアコンテンツを表示及び放送でき、オーディオ、ビデオプレーヤー（例えば、iPod）、ハンドヘルドゲームデバイス、電子書籍、インテリジェントおもちゃ及びポータブルカーナビゲーションデバイスを含む。

【0062】

(4) データー交換機能を備えたその他の電子デバイス。

【0063】

上記の装置の実施形態は、例示的であり、分離するコンポーネントとして記載されたユニットは、物理的に分離されるものであってもよく、分離されないものであってもよい。ユニットとして表示されるコンポーネントは、物理ユニットであってもよく、物理ユニットではなくてもよい。即ち、一つの場所に配置することもでき、複数のネットワークユニットに分散することもできる。実際のニーズに応じて、いくつかのモジュール又はすべてのモジュールを選択することができ、実施形態の目的を達成するようになる。

【0064】

上記の実施形態の説明を通じて、当業者は、各実施形態がソフトウェア及び共通のハードウェアプラットフォームによって実現することができ、もちろん、ハードウェアによって実現することもできることを明確に理解することができる。この理解に基づいて、上記の技術方案の本質又は関連技術に貢献する部分は、ソフトウェア製品の形式で具体化でき、コンピュータソフトウェア製品は、例えば、ROM/RAM、磁気ディスク、コンパクトディスクなどのコンピュータ可読記憶媒体に格納でき、コンピュータデバイス（パーソナルコンピュータ、サーバー又はネットワークデバイスなどである）に、各々の実施形態又は実施形態のある部分に記載された方法を実行させるように、いくつかの命令を含む。

【0065】

10

20

30

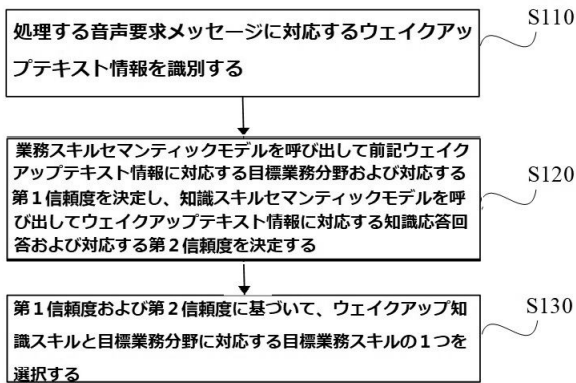
40

50

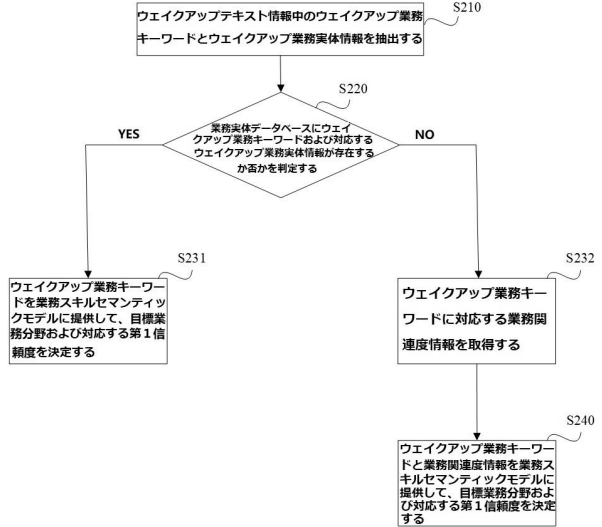
最後に、上記した実施形態は、本発明の技術構成を説明するためのものであり、本発明を限定するためのものではない。当業者であれば、本発明の技術構成の趣旨や範囲を逸脱しない前提下で、上述した実施形態に対して修正することもできるし、一部の技術的特徴を均等置換することもできる。これらの修正や置換は、いずれも本発明の保護範囲に含まれるべきである。

【図面】

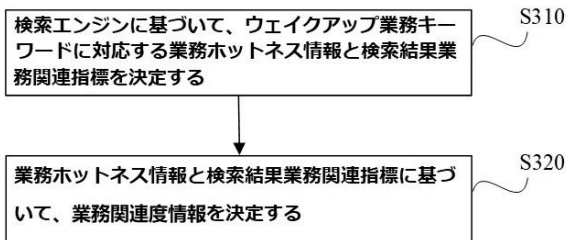
【図 1】



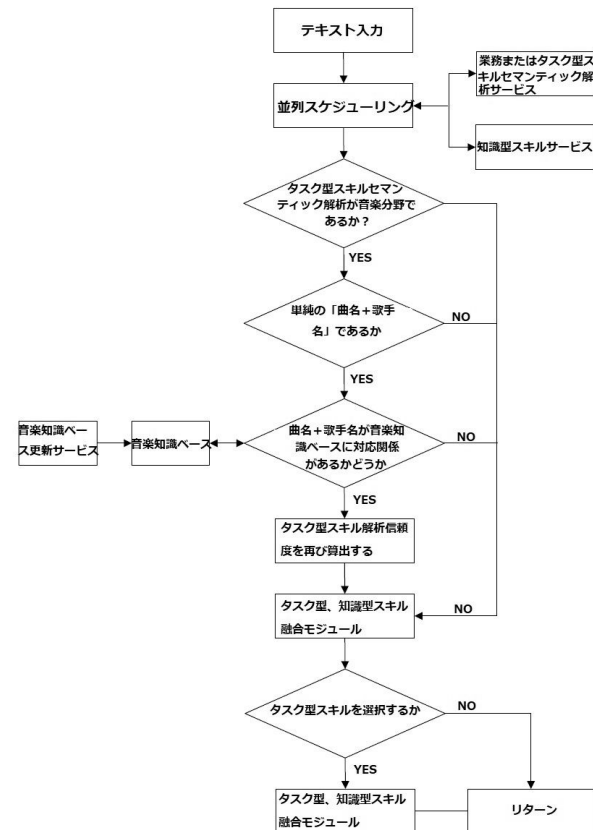
【図 2】



【図 3】



【図 4】



10

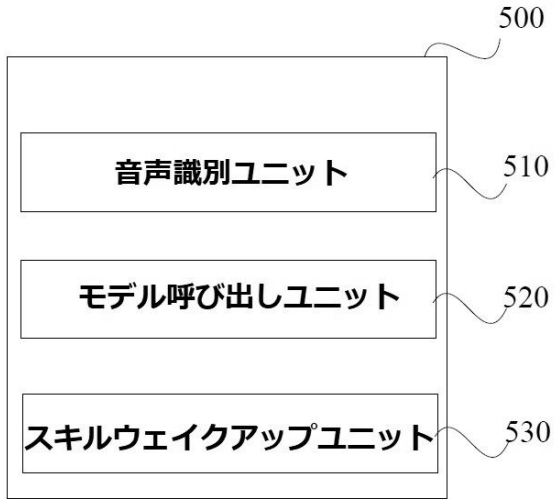
20

30

40

50

【図 5】



10

20

30

40

50

フロントページの続き

弁理士 長谷川 隆治

(72)発明者 朱 成亜

中華人民共和国江蘇省蘇州市蘇州工業園区新平街388号騰飛科技园14棟

審査官 中村 天真

(56)参考文献 中国特許出願公開第107316643(CN, A)

特開2013-190985(JP, A)

中国特許出願公開第109658271(CN, A)

特表2018-503857(JP, A)

特開2016-024652(JP, A)

(58)調査した分野 (Int.Cl., DB名)

G10L 15/00 - 15/34

G06F 40/00 - 40/58