

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6254583号
(P6254583)

(45) 発行日 平成29年12月27日(2017.12.27)

(24) 登録日 平成29年12月8日(2017.12.8)

(51) Int.Cl.	F 1
G 06 F 17/30	(2006.01) G 06 F 17/30 310Z
G 10 L 15/00	(2013.01) G 06 F 17/30 350C
G 10 L 15/065	(2013.01) G 10 L 15/00 200T
G 10 L 15/197	(2013.01) G 10 L 15/065 Z G 10 L 15/197

請求項の数 18 (全 23 頁)

(21) 出願番号	特願2015-518519 (P2015-518519)
(86) (22) 出願日	平成25年6月18日 (2013.6.18)
(65) 公表番号	特表2015-526797 (P2015-526797A)
(43) 公表日	平成27年9月10日 (2015.9.10)
(86) 国際出願番号	PCT/US2013/046381
(87) 国際公開番号	W02013/192218
(87) 国際公開日	平成25年12月27日 (2013.12.27)
審査請求日	平成28年4月25日 (2016.4.25)
(31) 優先権主張番号	61/662,889
(32) 優先日	平成24年6月21日 (2012.6.21)
(33) 優先権主張国	米国(US)
(31) 優先権主張番号	13/802,414
(32) 優先日	平成25年3月13日 (2013.3.13)
(33) 優先権主張国	米国(US)

(73) 特許権者	502208397 グーグル エルエルシー アメリカ合衆国 カリフォルニア州 94 043 マウンテン ビュー アンフィシ アター パークウェイ 1600
(74) 代理人	100108453 弁理士 村山 靖彦
(74) 代理人	100064908 弁理士 志賀 正武
(74) 代理人	100089037 弁理士 渡邊 隆
(74) 代理人	100110364 弁理士 実広 信哉

最終頁に続く

(54) 【発明の名称】動的言語モデル

(57) 【特許請求の範囲】

【請求項 1】

データ処理装置によって実施される方法であって、

音声認識のためのベース言語モデルを受信するステップであって、前記ベース言語モデルは、ベース確率値を有する第1の単語シーケンスを含む、ステップと、

音声検索クエリを受信するステップであって、前記音声検索クエリは、クエリコンテキストに関連付けられている、ステップと、

前記クエリコンテキストがカスタマイズされた言語モデルに関連付けられた1つまたは複数の基準を満たす場合、前記カスタマイズされた言語モデルが使用されるべきであると決定するステップと、

前記カスタマイズされた言語モデルを取得するステップであって、前記カスタマイズされた言語モデルは、調整済み確率値を有する前記第1の単語シーケンスを含み、前記調整済み確率値は、前記クエリコンテキストに従って1つまたは複数の調整係数に基づいて調整された前記ベース確率値である、ステップと、

1つまたは複数の確率に基づいて前記音声検索クエリをテキスト検索クエリに変換するステップであって、前記1つまたは複数の確率の各々は、1つまたは複数の単語シーケンスのグループにおける単語シーケンスに対応し、前記1つまたは複数の単語シーケンスのグループは、前記調整済み確率値を有する前記第1の単語シーケンスを含む、ステップとを含む、方法。

【請求項 2】

10

20

前記カスタマイズされた言語モデルを取得するステップは、
前記クエリコンテキストに基づいて1つまたは複数のカスタマイズされた言語モデルから前記カスタマイズされた言語モデルを選択するステップを含む、請求項1に記載の方法。
。

【請求項 3】

前記カスタマイズされた言語モデルを取得するステップは、
1つまたは複数の言語調整ルールを用いて、前記受信したクエリコンテキストに基づいて前記ベース言語モデルを調整するステップを含む、請求項1に記載の方法。

【請求項 4】

前記クエリコンテキストは、前記音声検索クエリを受信したユーザデバイスの地理的ロケーションを含む、請求項1に記載の方法。 10

【請求項 5】

前記単語シーケンスが前記地理的ロケーションに関連していると決定するステップであって、前記調整済み確率値は、前記単語シーケンスと前記地理的ロケーションとの関連度に基づいて調整された前記ベース確率値に相当する、ステップを含む、請求項4に記載の方法。

【請求項 6】

前記単語シーケンスは、前記地理的ロケーションにまたはその近傍に位置する地物の名称を含み、

前記関連度は、前記クエリコンテキスト内の前記地物と前記地理的ロケーションとの距離に基づいて決定される、請求項5に記載の方法。 20

【請求項 7】

前記単語シーケンスは、閾値を超える頻度で前記地理的ロケーションから受信した過去のクエリのタームを含み、

前記関連度は、前記頻度に基づいて決定される、請求項5に記載の方法。

【請求項 8】

前記地理的ロケーションは、経度および緯度からなる座標の1つまたは複数のセットによって定義される地理的区域である、請求項4に記載の方法。

【請求項 9】

前記クエリコンテキストは、前記音声検索クエリに関連付けるためのユーザ提供識別子を含む、請求項1に記載の方法。 30

【請求項 10】

前記ユーザ提供識別子に関連付けられたソーシャルグループを決定するステップであって、前記ソーシャルグループは、前記ユーザ提供識別子に関連付けられた1つまたは複数のユーザ識別子を含む、ステップと、

前記ソーシャルグループにおける前記ユーザ識別子に関連付けられたストアドクエリのログを用いて前記単語シーケンスの発生頻度を決定するステップと、

前記発生頻度に基づいて前記ベース確率値を調整するステップとを含む、請求項9に記載の方法。

【請求項 11】

前記ベース確率値を調整するステップは、前記ユーザ識別子の各々と前記ユーザ提供識別子との前記ソーシャルグループにおける関連度にさらに基づいている、請求項10に記載の方法。

【請求項 12】

前記クエリコンテキストは、前記音声検索クエリのタイムスタンプを含む、請求項1に記載の方法。

【請求項 13】

前記カスタマイズされた言語モデルを取得するステップは、
前記1つまたは複数の電子ドキュメントを特定するステップであって、1つまたは複数の電子ドキュメントは、クエリのログに従って検索クエリの数に関連付けられ、検索クエリ 50

の数は、閾値を満たし、前記検索クエリは、前記タイムスタンプの指定の時間ウィンドウ内に発生したものであり、前記1つまたは複数の電子ドキュメントの各々は、前記単語シーケンスを含む、ステップと、

前記検索クエリの数に基づいて前記ベース確率値を調整するステップとを含む、請求項12に記載の方法。

【請求項14】

前記1つまたは複数の電子ドキュメントの各々は、ニュース記事である、請求項13に記載の方法。

【請求項15】

前記音声検索クエリに対する応答として、前記テキスト検索クエリの1つまたは複数の検索結果をディスプレイデバイス上の表示を提供するステップを含む、請求項1に記載の方法。 10

【請求項16】

前記音声検索クエリをテキスト検索クエリに変換するステップは、前記カスタマイズされた言語モデルにおける対応する前記調整済み確率値に基づいて、前記音声検索クエリを前記第1の単語シーケンスを含むテキスト検索クエリに変換するステップを含む、請求項1に記載の方法。

【請求項17】

1つまたは複数のコンピュータと、1つまたは複数の記憶装置であって、前記1つまたは複数のコンピュータによって実行される際に、 20

音声認識のためのベース言語モデルを受信するステップであって、前記ベース言語モデルは、ベース確率値を有する第1の単語シーケンスを含む、ステップと、

音声検索クエリを受信するステップであって、前記音声検索クエリは、クエリコンテキストに関連付けられている、ステップと、

前記クエリコンテキストがカスタマイズされた言語モデルに関連付けられた1つまたは複数の基準を満たす場合、前記カスタマイズされた言語モデルが使用されるべきであると決定するステップと、

前記カスタマイズされた言語モデルを取得するステップであって、前記カスタマイズされた言語モデルは、調整済み確率値を有する前記第1の単語シーケンスを含み、前記調整済み確率値は、前記クエリコンテキストに従って1つまたは複数の調整係数に基づいて調整された前記ベース確率値である、ステップと、 30

1つまたは複数の確率に基づいて前記音声検索クエリをテキスト検索クエリに変換するステップであって、前記1つまたは複数の確率の各々は、1つまたは複数の単語シーケンスのグループにおける単語シーケンスに対応し、前記1つまたは複数の単語シーケンスのグループは、前記調整済み確率値を有する前記第1の単語シーケンスを含む、ステップとを含む処理を前記1つまたは複数のコンピュータに実施させるように処理可能な命令を記憶する、1つまたは複数の記憶装置とを備える、システム。

【請求項18】

コンピュータプログラムで符号化されたコンピュータ記憶媒体であって、前記プログラムは、データ処理装置によって実行される際に、 40

音声認識のためのベース言語モデルを受信するステップであって、前記ベース言語モデルは、ベース確率値を有する第1の単語シーケンスを含む、ステップと、

音声検索クエリを受信するステップであって、前記音声検索クエリは、クエリコンテキストに関連付けられている、ステップと、

前記クエリコンテキストがカスタマイズされた言語モデルに関連付けられた1つまたは複数の基準を満たす場合、前記カスタマイズされた言語モデルが使用されるべきであると決定するステップと、

前記カスタマイズされた言語モデルを取得するステップであって、前記カスタマイズされた言語モデルは、調整済み確率値を有する前記第1の単語シーケンスを含み、前記調整済み確率値は、前記クエリコンテキストに従って1つまたは複数の調整係数に基づいて調 50

整されたベース確率値である、ステップと、

1つまたは複数の確率に基づいて前記音声検索クエリをテキスト検索クエリに変換するステップであって、前記1つまたは複数の確率の各々は、1つまたは複数の単語シーケンスのグループにおける単語シーケンスに対応し、前記1つまたは複数の単語シーケンスのグループは、前記調整済み確率値を有する前記第1の単語シーケンスを含む、ステップとを含む処理を前記データ処理装置に実施せらるるように処理可能な命令からなる、コンピュータ記憶媒体。

【発明の詳細な説明】

【技術分野】

【0001】

10

本明細書は、音声認識に関する。

【背景技術】

【0002】

従来の音声検索システムは、ユーザからの音声入力をテキスト検索クエリに変換することを目的としている。一般的な音声検索システムでは、音声インターフェースは、ユーザの音声入力を受信し、音声入力を音声認識エンジンに提供する。音声認識エンジンは、言語モデルに基づいて音声入力の口述単語をテキスト検索クエリに変換し得る。そして、音声検索システムは、テキスト検索クエリを検索エンジンに送信する。したがって、音声検索システムは、ユーザがテキスト検索クエリをタイピングすることなく検索を実施することを可能にする。

20

【発明の概要】

【課題を解決するための手段】

【0003】

音声検索システムは、音声入力をテキスト検索クエリに変換するために、カスタマイズされた言語モデルを適用する。音声検索システムは、例えば、匿名検索ログデータのコレクションを用いて構築されたベース言語モデルを取得する。音声検索システムは、1つまたは複数のカスタマイズ基準、例えば、地理的ロケーション、時刻、または天候に基づいて、ベース言語モデルをカスタマイズする。カスタマイズされた言語モデルでは、カスタマイズ基準に関連した1つまたは複数の単語シーケンスに割り当てられた確率値が調整される。

30

【0004】

音声検索システムは音声検索クエリを受信した場合、音声検索システムは、音声検索クエリのコンテキストを決定する。音声検索システムは、コンテキストのコンポーネントがカスタマイズ基準を満たす場合、カスタマイズされた言語モデルを決定する。そして、音声検索システムは、カスタマイズされた言語モデルを用いて、音声検索クエリをテキスト検索クエリに変換する。

【0005】

一般的に、本明細書に記載の発明特定事項の革新的一態様は、音声認識のためのベース言語モデルを受信するステップであって、ベース言語モデルは、ベース確率値を有する第1の単語シーケンスを含む、ステップと、音声検索クエリを受信するステップであって、音声検索クエリは、クエリコンテキストに関連付けられている、ステップと、クエリコンテキストがカスタマイズされた言語モデルに関連付けられた1つまたは複数の基準を満たす場合、カスタマイズされた言語モデルが使用されるべきと決定するステップと、カスタマイズされた言語モデルを取得するステップであって、カスタマイズされた言語モデルは、調整済み確率値を有する第1の単語シーケンスを含み、調整済み確率値は、クエリコンテキストに従って調整されたベース確率値である、ステップと、1つまたは複数の確率に基づいて音声検索クエリをテキスト検索クエリに変換するステップであって、1つまたは複数の確率の各々は、1つまたは複数の単語シーケンスのグループにおける単語シーケンスに対応し、1つまたは複数の単語シーケンスのグループは、調整済み確率値を有する第1の単語シーケンスを含む、ステップとの処理を含む方法において具現化され得る。本態様

40

50

の他の実施形態は、その各々が方法の処理を実施するように構成される、対応するコンピュータシステム、装置、および1つまたは複数のコンピュータ記憶装置に記録されたコンピュータプログラムを含む。1つまたは複数のコンピュータのシステムは、処理においてシステムに動作を実施させる、システムにインストールされた、ソフトウェア、ファームウェア、ハードウェア、またはそれらの組合せを有しているので、特定の処理または動作を実施するように構成され得る。1つまたは複数のコンピュータプログラムは、データ処理装置によって実行される際に装置に処理を実施させる命令を備えることによって、特定の処理または動作を実施するように構成され得る。

【0006】

前述および他の実施形態の各々は、必要に応じて、1つまたは複数の以下の特徴そのものまたはその組合せを含む。カスタマイズされた言語モデルを取得するステップは、クエリコンテキストに基づいて1つまたは複数のカスタマイズされた言語モデルからカスタマイズされた言語モデルを選択するステップを含む。カスタマイズされた言語モデルを取得するステップは、1つまたは複数の言語調整ルールを用いて、受信したクエリコンテキストに基づいてベース言語モデルを調整するステップを含む。

10

【0007】

クエリコンテキストは、音声検索クエリを受信したユーザデバイスの地理的ロケーションを含む。方法は、単語シーケンスが地理的ロケーションに関連していると決定するステップであって、調整済み確率値は、単語シーケンスと地理的ロケーションとの関連度に基づいて調停されたベース確率値を含む、ステップを含む。単語シーケンスは、地理的ロケーションにまたはその近傍に位置する地物(feature)の名称を含み、関連度は、クエリコンテキスト内の地物と地理的ロケーションとの距離に基づいて決定される。単語シーケンスは、閾値を超える頻度で地理的ロケーションから受信した過去のクエリのタームを含み、関連度は、頻度に基づいて決定される。地理的ロケーションは、経度および緯度からなる座標の1つまたは複数のセットによって定義される地理的区域である。

20

【0008】

クエリコンテキストは、音声検索クエリに関連付けるためのユーザ提供識別子を含む。方法は、ユーザ提供識別子に関連付けられたソーシャルグループを決定するステップであって、ソーシャルグループは、ユーザ提供識別子に関連した1つまたは複数のユーザ識別子を含む、ステップと、ソーシャルグループにおけるユーザ識別子に関連付けられたストアドクエリのログを用いて単語シーケンスの発生頻度を決定するステップと、発生頻度に基づいてベース確率値を調整するステップとを含む。ベース確率値を調整するステップは、ユーザ識別子の各々とユーザ提供識別子とのソーシャルグループにおける関連度にさらに基づいている。クエリコンテキストは、音声検索クエリのタイムスタンプを含む。カスタマイズされた言語モデルを取得するステップは、1つまたは複数の電子ドキュメントを特定するステップであって、1つまたは複数の電子ドキュメントは、クエリのログに従って検索クエリの数に関連付けられ、検索クエリの数は、閾値を満たし、検索クエリは、タイムスタンプの指定の時間ウィンドウ内に発生したものであり、1つまたは複数の電子ドキュメントの各々は、単語シーケンスを含む、ステップと、検索クエリの数に基づいてベース確率値を調整するステップとを含む。1つまたは複数の電子ドキュメントの各々は、ニュース記事である。方法は、音声検索クエリに対する応答として、テキスト検索クエリの1つまたは複数の検索結果をディスプレイデバイス上の表示を提供するステップを含む。

30

【0009】

音声検索クエリをテキスト検索クエリに変換するステップは、カスタマイズされた言語モデルにおける対応する調整済み確率値に基づいて、音声検索クエリを第1の単語シーケンスを含むテキスト検索クエリに変換するステップを含む。

40

【0010】

本明細書に記載の発明特定事項の特定の実施形態は、以下の利点の1つまたは複数を実現するために実施され得る。

50

【0011】

従来の音声検索システムと比較して、動的言語モデルを用いる音声検索システムは、受信した音声入力に基づいて、より適切なテキスト検索クエリを提供し得る。システムが、例えば、音声入力を受信した時刻、音声入力をしたロケーション、および/または音声入力をしたユーザのユーザプロファイルを含む、音声-テキスト変換において、音声入力の特定の状況を考慮しているので、クエリの結果出力は、その特定の状況に応じてカスタマイズされ得る。したがって、システムは、向上したユーザエクスペリエンスを提供し得る。

【0012】

動的言語モデリング技法の1つまたは複数の実施形態の詳細を、添付の図面および以下の記載に記載する。技法の他の特徴、オブジェクト、および利点は、明細書および図面、並びに特許請求の範囲から明白であろう。

10

【図面の簡単な説明】**【0013】**

【図1】図1は、音声検索を提供するための例示的システムの概要を提供するブロック図である。

【図2】図2は、動的言語モデルを生成するための例示的ルールを表すブロック図である。

【図3】図3は、動的言語モデルの例示的地理学ベースのルールを表す概略図である。

【図4】図4は、動的言語モデルのルールに基づく例示的ソーシャルグループを表す概略図である。

20

【図5A】図5Aは、動的言語モデルの例示的実施形態を表すブロック図である。

【図5B】図5Bは、動的言語モデルの例示的実施形態を表すブロック図である。

【図6】図6は、カスタマイズされた言語モデルを用いるための例示的方法のフローチャートである。

【発明を実施するための形態】**【0014】**

様々な図面中の類似の参照符号は、類似の要素を示す。

【0015】

図1は、音声検索を提供するための例示的システムの概要を提供するブロック図である。音声検索システム102は、ユーザからユーザの音声入力を受信し、動的言語モデルを用いて音声入力をテキスト検索クエリに変換し、テキスト検索クエリに対する検索結果を受信し、音声入力に対する応答として検索結果をユーザに提供するようにプログラムされた1つまたは複数のコンピュータを備える。音声検索システム102は、例えば、サーバに含まれる1つまたは複数のコンピュータまたはユーザデバイスで実施され得る。

30

【0016】

音声検索システム102は、ユーザから音声入力106を受信するように構成される、音声インターフェース104を備える。音声インターフェース104は、例えば、ユーザデバイスに接続された音響電気変換器またはセンサ(例えば、マイクロフォン)を備え得る。音声入力106は、口述単語のシーケンスを含み得る。

40

【0017】

音声検索システム102は、音声認識サブシステム108を含む。音声認識サブシステム108は、音声入力106の口述単語をテキスト検索クエリ115に変換するように構成される、音声検索システム102のソフトウェアコンポーネントであり得る。音声認識サブシステム108は、ベース言語モデル110を含む。ベース言語モデル110は、単語のシーケンスに関連付けられた確率値を含み得る。ベース言語モデル110では、所与の単語シーケンス、例えば、 (W_a, W_b, \dots, W_m) は、確率値、例えば、 $P(W_a, W_b, \dots, W_m)$ に関連付けられており、 W_a, W_b, \dots, W_m の各々は、単語、熟語、または文を含む。ベース言語モデル110は、複数の単語シーケンスおよび単語シーケンスの各々に関連付けられた確率値を含み得る。

【0018】

50

ベース言語モデル110は、Nグラムモデル(例えば、ユニグラムモデルまたはマルチグラムモデル)であり得る。図示した例においては、ベース言語モデル110は、バイグラム言語モデルである。ベース言語モデルは、単語シーケンス(W_x, W_y)に関連付けられた確率値 $P(W_x, W_y)$ を含む。確率値 $P(W_x, W_y)$ は、以下の数式を用いて計算され得る。

【0019】

【数1】

$$P(W_x, W_y) \approx P(W_x|<s>)P(W_y|W_x), \quad (1)$$

【0020】

10

ここで、 $P(W_x, W_y)$ は、検索クエリに表れる単語シーケンス(W_x, W_y)の確率であり、 $<s>$ は、検索クエリの開始を示し、 $P(W_x|<s>)$ は、 $<s>$ に関して単語 W_x が $<s>$ の次にくる条件付き確率を示し、 $P(W_y|W_x)$ は、単語 W_x に関して単語 W_y が単語 W_x の次にくる条件付き確率を示す。

【0021】

加えて、ベース言語モデル110は、単語シーケンス(W_p, W_q, W_r)に関連付けられた確率値 $P(W_p, W_q, W_r)$ を含み得る。確率値 $P(W_p, W_q, W_r)$ は、例えば、バイグラムモデルにおいては、以下の数式を用いて計算され得る。

【0022】

【数2】

20

$$P(W_p, W_q, W_r) \approx P(W_p|<s>) P(W_q|W_p) P(W_r|W_q) \quad (2)$$

【0023】

確率値 $P(W_x|<s>)$ 、 $P(W_y|W_x)$ 、 $P(W_p|<s>)$ 、 $P(W_q|W_p)$ 、および $P(W_r|W_q)$ は、過去に受信したストアドクエリから決定され得る。

【0024】

30

図においては、単語シーケンス(W_p, W_q, W_r)は、世間で注目されている検索クエリ、例えば、他の検索クエリと比較してより高い頻度で表れる検索クエリを表し得る。単語シーケンス(W_x, W_y)は、特殊な状況下でのみ注目されている検索クエリを表し得る。例えば、(W_x, W_y)は、地理的区域、例えば、ABC町に位置する、名物レストランの名称であり得る。音声検索システム102は、地理的区域ABC以外にいるユーザから受信するよりは、地理的区域ABCにいるユーザからクエリターム(W_x, W_y)を有する検索クエリを受信する可能性が高い。

【0025】

本例では、地理的区域ABCは、クエリを受信する多くの地理的区域のうちで特に注目の地理的区域というわけではない。したがって、ベース言語モデルでは、 $P(W_p, W_q, W_r)$ は、 $P(W_x, W_y)$ より大きくなり得る。このことは、一般的に、音声認識サブシステム108が音声入力106を、テキスト検索クエリ(W_x, W_y)に変換するより、テキスト検索クエリ(W_p, W_q, W_r)に変換する可能性が高いことを示している。確率値の順位付けに従って、(W_p, W_q, W_r)は、1番目に順位付けされ得る一方で、(W_x, W_y)は、ベース言語モデル110における1番目より後位の第w番目のポジションに順位付けされ得る。

40

【0026】

音声認識サブシステム108は、モデルカスタマイズモジュール112を備える。モデルカスタマイズモジュール112は、コンピュータデバイスに、1つまたは複数の調整係数に基づいてベース言語モデル110を動的に調整させ、カスタマイズされた言語モデル114を生成させるように構成される、音声認識サブシステム108のソフトウェアコンポーネントを含み得る。調整係数は、音声入力106が特定のテキスト検索クエリに変換される可能性に影響を与えるベース言語モデル110とは独立した情報アイテムである。モデルカスタマイズモジュール112は、例えば、ロケーション、クエリタイムスタンプ、または他のパラメータ

50

のうちの1つまたは複数に基づいて、調整係数を決定し得る。

【0027】

例えば、所与の音声入力に対し、確率値 $P(W_x, W_y)$ は、ベース言語モデル110における確率値 $P(W_p, W_q, W_r)$ より低くなり得る。モデルカスタマイズモジュール112は、1つまたは複数の調整係数が考慮に入れられた場合、例えば、特定のロケーション条件を満たす場合、カスタマイズされた言語モデル114における確率値 $P(W_x, W_y)$ がベース言語モデル110における確率値 $P(W_x, W_y)$ より高くなるように、確率値 $P(W_x, W_y)$ を増大させ得る。モデルカスタマイズモジュール112は、カスタマイズされた言語モデル114をロケーション(例えば、地理的領域ABC)に関連付け得る。モデルカスタマイズモジュール112は、カスタマイズされた言語モデル114における1つまたは複数の特定単語シーケンスの確率がベース言語モデル110におけるその同一単語シーケンスの確率とは異なるように、ベース言語モデル110を動的に調整しカスタマイズされた言語モデル114をクエリ時に生成し得る。
10

【0028】

音声認識サブシステム108は、音声入力106に関連付けられたクエリコンテキスト117を受信するように構成される。クエリコンテキスト117は、音声入力106が生成された状況を記載しているデータを含む。クエリコンテキスト117は、音声入力106された場所を示すロケーション情報を含み得る。例えば、ロケーション情報は、それに基づいてロケーションが推定され得る、緯度および経度パラメータ、全地球測位網の座標、またはインターネットプロトコル(IP)アドレスを含み得る。ロケーション情報は、ユーザデバイスから取得され得る。
20

【0029】

音声認識サブシステム108は、音声入力106をテキスト検索クエリ115に変換するためには、ベース言語モデル110またはカスタマイズされた言語モデル114を使用するかどうかを決定するように構成される。例えば、音声認識サブシステム108は、クエリコンテキスト117に基づいて、音声入力106を提供するユーザが地理的領域ABCに位置していると決定すると、音声認識サブシステム108は、カスタマイズされた言語モデル114を使用し得る。しかしながら、音声認識サブシステム108は、クエリコンテキスト115に基づいて、音声入力106を提供するユーザが地理的領域ABCに位置していないと決定すると、音声認識サブシステム108は、音声入力106をテキスト検索クエリ115に変換するためにベース言語モデル110を使用し得る。
30

【0030】

したがって、音声認識サブシステム108は、音声入力106を提供するユーザが地理的領域ABCに位置していると決定すると、音声認識サブシステム108は、カスタマイズされた言語モデル114を音声認識に適用し、音声入力106から文字列を生成する。カスタマイズされた言語モデル114における確率値の順位付けと、他の全てが同一であることとに従って、音声入力106は、カスタマイズされた言語モデル114において増大した確率値 $P(W_x, W_y)$ が確率値 $P(W_p, W_q, W_r)$ より高いので、 (W_p, W_q, W_r) より (W_x, W_y) に関連している可能性が高くなる。それゆえ、音声認識サブシステム108は、音声入力106をターム (W_x, W_y) を含む文字列に変換する、またはユーザがテキスト検索クエリを選択するために、 (W_x, W_y) が (W_p, W_q, W_r) より高位に、例えば、クエリ提案リストのトップまたはその近くに配置されるようにするオプションのリストの表示を提供する。
40

【0031】

音声認識サブシステム108は、音声入力106を提供するユーザがカスタマイズされた言語モデルに関連付けられた特定の地理的領域に位置していないと決定すると、音声認識サブシステム108は、ベース言語モデル110を音声認識に適用し、例えば、ターム (W_p, W_q, W_r) を含む音声入力106から文字列を生成する。

【0032】

音声検索システム102は、テキスト検索クエリ115として音声認識サブシステム108によって生成された文字列を検索エンジン116に送信する。テキスト検索クエリ115は、カスタマイズされた言語モデル114に基づいて検索ターム「 $W_x W_y$ 」を、またはベース言語モデル
50

110に基づいて「W_p W_q W_r」を含み得る。検索エンジン116は、音声検索システム102と接続されたウェブ検索エンジンを備え得る。音声検索システム102は、検索エンジン116から1つまたは複数のクエリ結果118を受信し、ユーザデバイス上のコンポーネント、例えば、ウェブブラウザであり得る、またはサーバのコンポーネント、例えば、動的ウェブページであり得る、提示インターフェース120上に表示するためにクエリ結果118を提供する。提示インターフェース120は、ディスプレイインターフェースまたは音声出力インターフェースを備え得る。提示インターフェース120は、ユーザ1つまたは複数のクエリ結果118、例えば、表示されたクエリ結果ページ122を表示し得るし、それは、例えば結果「W_x W_y レストラン」を含み得る。

【0033】

10

図2は、カスタマイズされた言語モデル、例えば、カスタマイズされた言語モデル114を生成するための例示的ルールを表すブロック図200である。動的言語モデルを実装するシステムは、ベース言語モデル202を取得する。システムは、別のシステムからベース言語モデル202を取得し得る。あるいは、システムは、周知の利用可能な言語モデリング技法、例えば、RandLMまたはCMU-Cambridge Statistical Language Modeling Toolkitを用いて、検索ログ204からベース言語モデル202を構築し得る。システムは、システムと通信する、またはシステムの一部である検索エンジンから検索ログを取得し得る。いくつかの実施形態では、システムは、システムによってサポートされている各言語に関する個別のベース言語モデル202を取得する。

【0034】

20

システムは、地理的言語モデルルール206を用いてベース言語モデル202をカスタマイズし得る。地理的言語モデルルール206は、単語シーケンスが地理的地物に関連している場合、または単語シーケンスを指定の地理的区域から受信した場合、ベース言語モデル202における単語シーケンスの確率値を調整するための調整係数を含む。単語シーケンスは、例えば、地理的地物の名称、アドレス、または特徴などにより地理的地物に関連し得る。

【0035】

地理的地物は、関連する任意の地物、または地理的区域に位置するエンティティを含み得る。例えば、地理的地物は、企業、ランドマーク、地方官庁、通り、公園、学校、または特定のロケーションで一般的に発生するイベントを含み得る。地理的区域は、国、州、郡、市、商業地区、または経度および緯度によって決定されるグリッドを含み得る。

30

【0036】

地理的言語モデルルール206は、地理的区域ごとにグループ化され得る。各地理的区域について、システムは、地理的区域内の地理的地物に関連する単語シーケンスを調整係数に関連付ける。調整係数は、他の地理的区域と比較して、指定の地理的区域からの検索クエリが指定の地理的区域内の地理的地物に関連する単語シーケンスを含む可能性がより高くなるようにする、可能性を含み得る。調整係数は、ベース言語モデル202における単語シーケンスに関連付けられた確率値に適用される乗数を含み得る。

【0037】

例えば、地理的言語モデルルールは、本例では地理的領域ABCに位置する企業の名称である単語シーケンス(W_x, W_y)のための調整係数を含む。調整係数は、単語シーケンス(W_x, W_y)の確率値P(W_x, W_y)を増大させる乗数を含み得る。システムは、カスタマイズされた言語モデルを生成するために、乗数をベース言語モデル202における確率値P(W_x, W_y)に適用する。カスタマイズされた言語モデルは、地理的領域ABCに関連付けられ得る。システムは、音声入力(例えば、音声クエリ)を提供するユーザが地理的領域ABCに位置していると決定すると、システムは、音声認識のためのカスタマイズされた言語モデルを使用する。

40

【0038】

システムは、地理的データストア208、検索ログ204、またはその双方からの情報に基づいて、地理的言語モデルルール206を生成し得る。地理的データストア208は、各地理的区域に関連付けられた地理的地物の名称およびロケーションを記憶する。システムは、登録

50

企業名、公的記録、またはユーザ入力のリストから名称およびロケーションを取得し得る。例えば、地理的領域ABC内の名称(Wx, Wy)を有するレストランの経営者は、地理的領域ABCに関連付けて名称(Wx, Wy)ウェブに登録し得る

【 0 0 3 9 】

「職業別広告欄」または公企業の他のデータベースは、地理的データストア208にリスト化される。データベースに基づいて、システムは、名称(Wx, Wy)を地理的領域ABCに関連付け得る。地理的データストア208に記憶された名称は、ベース言語モデル202における単語シーケンスにマッピングするために、正規化(例えば、大文字解除、スペル訂正、または標準化)され得る。

【 0 0 4 0 】

システムは、システムが調整係数の大きさを決定するために使用するクエリ頻度情報を生成するために、検索ログ204を解析し得る。クエリ頻度情報は、システムまたは検索エンジンによってこれまで受信された特定のクエリの回数についてのカウントを含み得る。システムは、閾値を用いてクエリ頻度情報をフィルタリングし得る。システムは、確率値を増大させるかどうかを決定するために閾値を使用し得る。検索ログ204に記憶された情報を用いて、システムは、様々なロケーションでユーザから受信したクエリターム(Wx, Wy)を含む音声またはテキスト検索クエリの頻度を決定し得る。システムは、ユーザが地理的領域ABCにまたは近くに位置している場合に頻度が閾値を超えていると決定すると、システムは、ベース言語モデル202における値P(Wx, Wy)を増大させる、調整係数を有する地理的言語モデルルールを指定し得る。

10

【 0 0 4 1 】

加えてまたはあるいは、システムは、ソーシャルグループ言語モデルルール210を用いてベース言語モデル202をカスタマイズし得る。ソーシャルグループ言語モデルルール210は、ユーザのソーシャルグループに基づいてベース言語モデル202における確率値を調整するための調整係数を含む。ベース言語モデル202における単語シーケンスの確率値は、ユーザAのソーシャルグループにおけるユーザによって送信された情報に基づいて、ユーザAに関して調整される。情報は、例えば、ソーシャルグループにおけるユーザの1つまたは複数のストアドクエリを含み得る。

20

【 0 0 4 2 】

システムがは、閾値を超える頻度でユーザAのソーシャルグループのユーザから特定のクエリを受信した場合、システムは、クエリに対応する単語シーケンスの確率値を調整し得る。例えば、システムは、ユーザAのソーシャルグループ内のユーザからのクエリに基づいて、システムが単語シーケンス(Wa, Wb, Wc, Wd)を含む検索クエリの閾値を受信していると決定し得る。いくつかの実施形態では、送信ユーザが(例えば、ストアドクエリから情報を特定し得る出所を除去することによって)わからなくなるように、クエリは、クエリのログにおいて匿名化されている。そのような決定に基づいて、システムは、ユーザAに関する単語シーケンス(Wa, Wb, Wc, Wd)の確率値を増大させるように、またユーザAのソーシャルグループ内のユーザに関して同様に増大させるように、ベース言語モデル202を調整し得る。したがって、ユーザAのソーシャルグループ内のユーザに関して、システムは、ベース言語モデル202においてより高い確率を有するいくつかの他の単語シーケンス、例えば(We, Wf, Wg, Wh)よりむしろ単語シーケンス(Wa, Wb, Wc, Wd)として、音声入力を認識する可能性を増大させ得る。

30

40

【 0 0 4 3 】

システムは、ソーシャルグループデータストア212、検索ログ204、または双方から取得した情報に基づいて、ソーシャルグループ言語モデルルール210を生成し得る。ソーシャルグループデータストア212は、ユーザに関連する情報、および関連ユーザの各々が言語モデルをカスタマイズするためにシステムに送信することを選択した情報を記憶し得る。情報は、ユーザとの社会的に交流があると特定される1つまたは複数のユーザを含む、交流ユーザのサークルを含み得る。情報は、ソーシャルサイト、仮想グループ、仮想チャットルーム、連絡先リスト、または前述した任意の組合せから取り出し得る。検索ログ

50

204は、システムによってこれまで受信された特定のクエリの回数についてのカウントを含み、ユーザのソーシャルグループにおけるユーザ識別子に対応づけられ得る、クエリ頻度情報を記憶している。

【 0 0 4 4 】

加えてまたはあるいは、システムは、イベント言語モデルルール214を用いてベース言語モデル202をカスタマイズし得る。イベント言語モデルルール214は、ベース言語モデル202における確率値を調整するための調整係数を含む。単語シーケンスが注目のイベントに関連している場合、単語シーケンスに関連付けられた確率値は、増大され得る。システムは、イベントデータストア216、または検索ログ204、またはその双方に記憶された情報に基づいて、イベント言語モデルルール214を生成する。指定のイベントは、音声入力を特定の単語シーケンスに変換する確率に影響を与え得る、任意のイベントを含み得る。10

【 0 0 4 5 】

いくつかの実施形態では、システムは、時間イベントに基づいて、確率値を調整する。システムは、一日における時刻または一週間における日数に基づいて、イベント言語モデルルール214を生成し得る。例えば、システムは、指定の日時(例えば、正午から午後1時の間、または午後6時から午後8時の間)において、システムがレストランに関連するクエリの数値またはパーセンテージであって閾値を超えている数値またはパーセンテージを受信していると決定し得る。システムは、数値またはパーセンテージに基づいてレストランに関連する、ベース言語モデル202における単語シーケンスの確率値を増大させ得る。20

【 0 0 4 6 】

いくつかの実施形態では、システムは、天候イベントに基づいて、確率値を調整する。システムは、天候情報に基づいて、イベント言語モデルルール214を生成し得る。システムは、所与の天候状況下において、単語シーケンスの指定のセットに調整済み確率値が付与されるように構成され得る。イベントデータストア216は、1つまたは複数の天候サービスから受信した情報を含み得る。天候サービスの各々が雨の確率が指定の閾値を超えて天気予報を提供している場合、道路状況レポート、洪水警報、または屋根の修理に関連するサービスの名称により高い重み付けがなされる。20

【 0 0 4 7 】

いくつかの実施形態では、システムは、注目のニュースイベントを使用する。システムは、1つまたは複数のニュースレポートの話題の注目度に基づいて、特定のイベント言語モデルルール214を生成し得る。例えば、イベントデータストア216は、様々なニュースソースからの情報を記憶し得る。単語シーケンスが、閾値を超える頻度でニュースソースに表れる場合、システムは、単語シーケンスに関連付けられた確率値を増大させるイベント言語モデルルール214を生成し得る。30

【 0 0 4 8 】

図3は、地理的情報に基づく例示的動的言語モデルを表す概略図である。地理的情報に基づいて動的言語モデルを実施するシステムにおいて、システムは、地理的情報を用いて単語シーケンスにアノテーションを付加する。クエリ時において、システムは、音声入力を受信したユーザデバイスの地理的ロケーションに基づいて、動的言語モデルを使用する。40

【 0 0 4 9 】

システムは、ユーザデバイスからロケーション情報を受信するように構成されるロケーションインターフェースを備える。いくつかの実施形態では、ロケーションインターフェースは、経度、緯度、または高度情報をロケーション認識ユーザデバイスから受信するよう構成される、様々なモジュールを含み得る。いくつかの実施形態では、ロケーションインターフェースは、ユーザデバイスのインターネットプロトコル(IP)アドレス、ユーザデバイスに接続される無線アクセスポイント、または他の情報に基づいて、ユーザデバイスのロケーションを決定するように構成される、様々なモジュールを含み得る。システムは、同一の音声入力を、ユーザデバイスが異なる地理的ロケーションに位置する際は異なるテキスト検索クエリに変換し得る。50

【 0 0 5 0 】

システムは、地理的区域を仮想ブロック310の集合に分割する。仮想ブロック310の集合における各ブロックは、ブロックに関してカスタマイズされた言語モデルに対応し得る。ブロックは、各ブロックの中心の、または各ブロックの1つまたは複数の頂点の、経度および緯度からなる座標に基づいて定義され得る。いくつかの実施形態では、ブロックは、各ブロックのコーナー、およびブロックサイズに基づいて定義され得る。ブロックサイズは、キロメータ、マイル、または経度若しくは緯度で定義され得る。

【 0 0 5 1 】

システムは、仮想ブロック310の集合のブロック7に位置するエンティティ312を特定する。エンティティ312は、ブロック7に恒久的または半恒久的に関連付けられた地物であり得る。例えば、エンティティ312は、ブロック7に位置する企業であり得る。この企業は、システムに登録された、またはシステムに接続されたレジストリに登録された、企業名(例えば、単語シーケンス(W_x, W_y))を有する。システムは、ブロック7に関連付けられた動的言語モデルを生成し得る。この動的言語モデルにおいて、エンティティに関連付けられた単語シーケンスの確率値 $P(W_x, W_y)$ は、確率値を増大させるために、乗数によって調整される。

10

【 0 0 5 2 】

システムは、仮想ブロック310の集合の各ブロックに関するカスタマイズされた言語モデルを生成し得る。カスタマイズされた言語モデルの各々において、エンティティ312に関連付けられた単語シーケンスの確率値は、ブロック7と動的言語モデルに関連付けられたブロックとの距離に基づいて決定される数値によって調整される。具体的には、数値は、距離に基づいて小さくなり得る。距離は、例えば、ユークリッド距離、マンハッタン距離、または推定移動時間によって求まる距離であり得る。ブロック7の直接隣接しているブロック(例えば、ブロック4、5、および8)に関して、対応する動的言語モデルは、確率値 $P(W_x, W_y)$ の第1の調整数値を有し得る。ブロック7の遠隔隣接しているブロック(例えば、ブロック1乃至3、6、および9)に関して、対応する動的言語モデルは、確率値 $P(W_x, W_y)$ の第2の調整数値を有し得る。ここで、第2の調整数値は、第1の調整数値より小さい。ブロック7からさらに離れているブロック(図示せず)に関して、確率値 $P(W_x, W_y)$ は、調整されないままであり得る。

20

【 0 0 5 3 】

30

システムはユーザデバイスから音声入力を受信する際、システムは、仮想ブロック310の集合のどこにユーザデバイスが位置しているかを決定する。そして、システムは、音声入力をテキスト検索クエリに変換するために、ブロックに関連付けられた言語モデルを適用する。

【 0 0 5 4 】

図4は、ユーザのソーシャルグループに基づいた例示的動的言語モデル400を表す概略図である。所与のユーザに関して、システムは、ある閾値を満たす頻度で単語シーケンスがその所与のユーザのソーシャルグループにおけるユーザによって使用される場合の言語モデルにおいて、単語シーケンスの確率値を調整する。頻度は、絶対頻度または相対頻度であり得る。

40

【 0 0 5 5 】

システムは、特定のソーシャルサイトの登録ユーザ(ユーザA)のユーザデータ402を取り出し得る。ユーザデータ402は、ユーザAに関するユーザ識別子と、ユーザAが音声認識改善を要求した後にユーザAによって送信されたユーザ情報とを含み得る。ユーザ情報は、例えば、ユーザのホームロケーション、ユーザの好み、およびユーザに関連付けられた他の情報を含み得る。

【 0 0 5 6 】

システムは、ユーザデータ402に基づいて、ユーザAに関する言語モデル調整ルール404を生成する。システムは、ロケーション、好み、および他の情報に基づいて、言語モデル調整ルール404を生成し得る。例えば、ユーザAが特定のクエリタームを含むクエリを送信

50

していた場合、およびユーザAがそのクエリを指定の回数(例えば、一度、または二度以上)送信していた場合、システムは、クエリタームを含む単語シーケンスの確率を増大させる、言語モデル調整ルール404におけるルールを生成し得る。

【0057】

システムは、ユーザによって提供または公認されたユーザデータ402に基づいて、ユーザAのソーシャルネットワークを特定し得る。ソーシャルネットワークは、ユーザ指定のソーシャル交流サークル(例えば、友人サークル)、ユーザAの連絡先リストに基づいたユーザのリスト、またはユーザとユーザAとの共通のアクティビティ(例えば、チャットルームへの参加)に基づいたユーザのリストを含み得る。システムは、ユーザとユーザAとの関連度を決定し得る。例えば、システムは、ユーザBおよびユーザCがユーザAに直接関連していると、第1の度合いでユーザAと関連付けられることを決定し得る。システムは、ユーザD、E、F、およびGがユーザBおよびCを介してユーザAに関連していると、第2の度合いでユーザAと関連付けられることを決定し得る。

10

【0058】

システムは、ユーザAのソーシャルネットワークにおけるユーザ(ユーザBからユーザG)からユーザデータ412、414、416、418、420、および422を、これら自身およびソーシャル交流サークルのユーザに関する音声認識を改善するとのユーザ要求の後に、取り出す。システムは、ユーザBからユーザGに関するモデル調整ルール432、434、436、438、440、および442をそれぞれ生成し得る。システムは、ユーザデータ412乃至422、またはモデル調整ルール432乃至442、または双方に基づいて、言語モデル調整ルール404の少なくとも一部を生成し得る。

20

【0059】

いくつかの実施形態では、システムは、ユーザのソーシャル交流のクエリを使用してカスタマイズされた言語モデルを構築する。例えば、ユーザデータ412に従って、ユーザBは、1回または複数回あるクエリタームを含むクエリを用いて、検索を実施していたとする。システムは、ユーザBの検索の頻度と、ユーザBとユーザAとの関連度とに基づいて、クエリタームを含む単語シーケンスの確率値を増大させる。

【0060】

いくつかの実施形態では、システムは、カスタマイズされた言語モデルを構築するために、ユーザの友人のクエリ以外の情報を使用する。例えば、ユーザAが参加していようがいまいが、ユーザBおよびユーザCが、名称「Wx Wy」を有するレストランを話題にしていた場合、またはクエリ「Wx Wy」を用いて検索していた場合、システムは、ユーザAについてカスタマイズされた言語モデルにおける単語シーケンス(Wx, Wy)の確率を増大させ得る。

30

【0061】

ユーザAのソーシャルネットワークのアプリケーションは、各ユーザのプライバシー設定に基づき得る。例えば、システムは、ユーザBおよびユーザCの各々が友人のサークル間の検索情報を共有することに許可を与えていた場合、またはユーザBおよびユーザCの各々がユーザの情報を使用してユーザの言語モデルを調整することに許可を与えていた場合、ユーザBおよびユーザCの検索を使用してユーザAの言語モデルを調整し得る。

40

【0062】

システムは、関連度に基づいて調整の数値を決定し得る。より小さい関連度は、2人のユーザ間のより密接な関係を表し、2人のユーザが関心および検索傾向を共有していたより高い可能性を示し得る。それゆえ、システムは、関連度が増大すると調整の数値を低減させる。例えば、システムは、ユーザBによって実施されたクエリに基づいて増大している場合、第1の数値にユーザAに関する単語シーケンスの確率値を増大させ得る。システムは、ユーザDによって実施されたクエリに基づいて増大している場合、第2の数値にユーザAに関する単語シーケンスの確率値を増大させ得る。第1の数値は、第2の数値より大きくなり得る。

【0063】

50

システムは、特定された言語モデル調整ルール404を用いてベース言語モデルをクエリ時にカスタマイズし得る、またはカスタマイズされた言語モデルをクエリ時以前に記憶し得る。システムはユーザAから音声入力を受信する際、システムは、ユーザAの識別子に基づいてユーザAに関する言語モデル調整ルール404を特定し得るし、言語モデル調整ルール404を用いてベース言語モデルをカスタマイズし得る。加えてまたはあるいは、システムは、ユーザA固有の事前に記憶されたカスタム言語モデルを用いて、音声認識を実施し得る。

【 0 0 6 4 】

図5Aおよび5Bは、動的言語モデルの例示的実施形態を表すブロック図である。動的言語モデル技法を実施するシステムは、図5A、図5B、または2つの組合せを参照して記載される実施形態に採用され得る。10

【 0 0 6 5 】

図5Aは、システムが、例えば受信したクエリに応じる以外のいくつかの時に、音声クエリを用いて非同期にカスタマイズされた言語モデル502、504、506、および508を生成する、例示的実施形態を表すブロック図500である。カスタマイズされた言語モデル502、504、506、および508の各々は、ベース言語モデルにおける単語シーケンスのいくつかまたは全てを含み、カスタマイズされた言語モデルを適用するための1つまたは複数の基準に関連付けられる。カスタマイズされた言語モデルにおけるそれぞれの単語シーケンスに関連付けられた確率値は、適用基準に基づいて決定される。適用基準の各々は、ロケーション、時刻、天候、または他の因子(例えば、母集団の年齢、性別、所得水準、または購買傾向によって定義されるユーザセグメント)を指定し得る。例えば、動的言語モデル506は、動的言語モデル506が指定の天候状況下の指定の時間ウィンドウ内の指定のロケーションで適用されることを指定している適用基準に関連付けられている。システムは、カスタマイズされた言語モデル502、504、506、および508を、周期的またはイベントの発生時に構築又は更新し得る。イベントの発生は、例えば、ユーザのセグメント化の新規方法が実装された場合を含み得る。20

【 0 0 6 6 】

クエリ時において、システムは、音声入力を受信し、音声入力のコンテキストを取得する。例えば、システムは、ユーザのロケーション、音声入力の時刻を決定し、天候サービスから天気情報をロケーションおよび時刻に基づいて取り出し得る。システムは、ユーザシステムにログインする際にユーザによって提供されるユーザ識別子に関連付けられたプロファイルを用いて、ユーザが属するユーザセグメントを特定し得る。システムは、コンテキストを言語モデル選択部510に送信する。言語モデル選択部510は、音声入力のコンテキストとカスタマイズされた言語モデル502、504、506、または508に関連付けられた適用基準とをマッチングし、そのマッチングに基づいて動的言語モデルを特定するように構成される、システムのコンポーネントである。30

【 0 0 6 7 】

例えば、システムは、ユーザロケーション、クエリ時刻、およびユーザロケーションにおけるクエリ時の天候を受信し得る。言語モデル選択部510は、ユーザロケーションが動的言語モデル506に関連付けられた適用基準の適用可能なロケーションと重なり、クエリ時刻が適用基準のタイムスタンプの適用可能な時間ウィンドウ内であり、およびユーザロケーションにおけるクエリ時の天候が適用基準の適用可能な天候に適合すると決定し得る。40

【 0 0 6 8 】

それに応じて、言語モデル選択部510は、カスタマイズされた言語モデル506を選択する。その後、システムは、音声認識を実施し、カスタマイズされた言語モデル506を用いてユーザの音声入力をテキスト検索クエリに変換する。例えば、システムは異なるロケーションまたは異なる時間に同一の音声入力を受信した場合、システムは、異なる言語モデル(例えば、言語モデル504)を適用し得る。

【 0 0 6 9 】

50

図5Bは、システムが動的に言語モデル522を調整する、例示的実施形態を表すプロック図520である。言語モデル522は、ベース言語モデルまたはカスタマイズされた言語モデルであり得る。

【0070】

システムは、言語モデル変更部524を備える。言語モデル変更部は、言語モデル522を調整するために1つまたは複数のルール526を適用するように構成されるユーザデバイスまたはサーバデバイスのコンポーネントである。言語モデル変更部524は、個人別情報、ロケーション情報、日付/時間情報、天候情報、最新ニュース/最新事件情報、または1つまたは複数のユーザからのクエリストリームを含む、入力を受信する。言語モデル変更部524は、入力に基づいてルール526を言語モデル522に適用する。ルール526の適用後、言語モデル522は、カスタマイズされた言語モデルとなり得る。10

【0071】

例えば、ルール526のうちの1つは、タームがある閾値を満たす頻度で指定の期間(例えば、ルール526が言語モデル522に適用される前の直近の3日間)内にニュースレポートに現れる場合、システムがそのタームを含む単語シーケンスの確率値を増大させることを指定し得る。別のルールは、システムが情報を収集することをユーザが許可する選択をした場合、システムが過去ユーザによって送信されたクエリを分析し、検索パターンを決定してユーザに関連付けることをユーザが許可し得ることを指定し得る。

【0072】

いくつかの実施形態では、言語モデル変更部524は、クエリ時に言語モデル522を調整し得る。いくつかの実施形態では、言語モデル変更部524は、特定のユーザに関してまたは特定の現在イベントに応じてカスタマイズされた言語モデルを生成し、ユーザデバイスまたはサーバデバイスにカスタマイズされた言語モデルを記憶し得る。20

【0073】

図6は、カスタマイズされた言語モデルを用いるための例示的方法600のフローチャートである。便宜上、方法は、例えば上述したような方法600を実施するシステムについて記載している。システムは、音声認識のためのベース言語モデルを受信する(602)。ベース言語モデルは、複数の単語シーケンスを含み、各単語シーケンスは確率値に関連付けられている。単語シーケンスは、文字列、圧縮テキスト、オブジェクト、またはインデックスとして表され得る。ベース言語モデルは、ユーザのコレクションの匿名クエリのログから生成され得る。30

【0074】

システムは、音声検索クエリを受信する(604)。音声検索クエリは、クエリコンテキストに関連付けられる。クエリコンテキストは、音声検索クエリを受信したユーザデバイスの地理的ロケーションを含み得る。地理的ロケーションは、ある地点またはある地理的区域であり得る。地点または地理的区域は、例えば、経度、緯度または高度からなる座標の1つまたは複数のセットによって定義され得る。クエリコンテキストは、例えば、クエリのタイムスタンプ、音声検索クエリを提供するユーザのまたはユーザ識別子を含み得る。

【0075】

システムは、クエリコンテキストがカスタマイズされた言語モデルに関連付けられた1つまたは複数の基準を満たす場合、カスタマイズされた言語モデルが使用されるべきと決定する(605)。クエリコンテキストが任意の基準を満たしていない場合、システムは、ベース言語モデルを使用して音声検索クエリをテキスト検索クエリに変換し得る。40

【0076】

システムは、カスタマイズされた言語モデルを取得する(606)。カスタマイズされた言語モデルは、調整済み確率値に対応する1つまたは複数の単語シーケンスを含む。調整済み確率値は、クエリコンテキストが1つまたは複数の調整基準に適合する場合、クエリコンテキストに従って調整されたベース確率値である。カスタマイズされた言語モデルを取得するステップは、(1)クエリコンテキストに基づいてカスタマイズされた言語モデルを1つまたは複数の事前に記憶されたカスタマイズされた言語モデルから選択するステップ、50

または(2)受信したクエリコンテキストに基づいて、カスタマイズされた言語モデルを生成するための1つまたは複数の言語調整ルールを用いて、クエリ時にベース言語モデルを調整するステップのうちの少なくとも1つを含み得る。カスタマイズされた言語モデルは、図1乃至5に関して上述したように取得され得る。

【0077】

いくつかの実施形態では、クエリコンテキストは、音声検索クエリを受信したユーザデバイスの地理的ロケーションを含む。システムは、地理的ロケーションに関連している1つまたは複数の単語シーケンスを特定し得る。ベース確率値を含む単語シーケンスの調整済み確率値は、単語シーケンスと地理的ロケーションとの関連度に基づいて調査される。例えば、単語シーケンスは、ユーザデバイスの地理的ロケーションにまたはその近傍に位置する地物(例えば、企業、公園、または通り)の名称を含み得る。関連度は、クエリコンテキスト内の地物と地理的ロケーションとの距離に基づいて決定され得る。地理的ロケーションは、経度、緯度、または高度からなる座標の1つまたは複数のセットによって定義される地理的区域であり得る。10

【0078】

いくつかの実施形態では、単語シーケンスは、閾値を超える頻度で地理的ロケーションから受信した過去のクエリのタームを含む。関連度は、頻度に基づいて決定される。例えば、システムが、特定のロケーションにおいて、ターム「climate best by government test」を含むクエリの十分な数値またはパーセンテージを受信した場合、システムは、その特定のロケーションに関連付けられたカスタマイズされた言語モデルにおける単語シーケンス「climate best by government test」の確率値を増大させ得る。いくつかの実施形態では、数値またはパーセンテージが高くなればなるほど確率値は増大する。20

【0079】

いくつかの実施形態では、クエリコンテキストは、音声検索クエリに関連付けるためのユーザ提供識別子(例えば、ユーザ名称またはユーザ識別子)を含む。システムは、ユーザ提供識別子に関連付けられたソーシャルグループを決定する。ソーシャルグループは、ユーザの記憶済みソーシャルネットワークであり得る。ソーシャルグループは、ソーシャルネットワークにおけるユーザ提供識別子に関連付けられた1つまたは複数のユーザ識別子を含み得る。システムは、ソーシャルグループにおけるユーザ識別子に関連付けられたストアドクエリのログを用いて単語シーケンスの発生頻度を決定し、発生頻度に基づいてベース確率値を調整する。例えば、システムは、検索ターム「swim meet」を含む、音声またはテキストのいずれかの検索クエリが、育児関連グループでは高頻度で発生していると決定し得る。システムは、単語シーケンス「swim meet」の確率値がベース言語モデルにおける確率値より高い、カスタマイズされた言語モデルを生成し、カスタマイズされた言語モデルを育児グループのユーザによる音声入力に適用し得る。加えてまたはあるいは、システムは、ソーシャルグループにおける関連度に基づいて、ベース確率値を調整する。30

【0080】

いくつかの実施形態では、クエリコンテキストは、音声検索クエリのタイムスタンプを含む。カスタマイズされた言語モデルを取得するステップは、1つまたは複数の電子ドキュメント(例えば、ニュース記事)を特定するステップを含む。電子ドキュメントは、各々個別にまたはコレクションとして、クエリのログに従って検索クエリの数に関連付けられる。電子ドキュメントは、検索エンジンによってクエリの検索結果として指定されていた場合、または格納先としてユーザによって検索結果のリストから選択されていた場合、クエリに関連付けられ得る。システムは、各々が1つまたは複数の単語シーケンスを含む、電子ドキュメントを特定する。関連付けられた検索クエリがタイムスタンプから指定の時間ウィンドウ(例えば、一日または一週間)内に発生する場合、システムは、カスタマイズされた言語モデルにおける1つまたは複数の単語シーケンスに関連付けられた確率値を増大させ得る。例えば、期間中に多くの教育関連のニュース記事が単語シーケンス「tiger mom」を含んでいた場合、ベース言語モデルにおいて低い確率値に関連付けられ得る単語シーケンス「tiger mom」は、カスタマイズされた言語モデルにおいてより高い確率値に4050

関連付けられ得る。

【0081】

システムは、カスタマイズされた言語モデルにおける対応する調整済み確率値に基づいて、音声検索クエリを単語シーケンスを含むテキスト検索クエリに変換する(608)。システムは、ディスプレイデバイスに表示するために音声検索クエリに対する応答としてテキスト検索クエリの1つまたは複数の検索結果を提供する(610)。

【0082】

本明細書に記載の発明特定事項および処理の実施形態は、本明細書に記載の構造物およびその構造的等価物を含む、デジタル電子回路、またはコンピュータソフトウェア、ファームウェア、若しくはハードウェア、またはそれらのうちの1つまたは複数の組合せにおいて実施され得る。本明細書に記載の発明特定事項の実施形態は、1つまたは複数のコンピュータプログラム、すなわち、データ処理装置によってまたはデータ処理装置の処理を制御するためにコンピュータ記憶媒体に符号化されたコンピュータプログラム命令の1つまたは複数のモジュールとして実施され得る。加えてまたはあるいは、プログラム命令は、人工的に生成された伝搬シグナル、例えば、データ処理装置による実行のための適切な受信機装置との送受信用情報を符号化するために生成される、機械生成された電気的、光学的、または電磁的シグナルに符号化され得る。コンピュータ記憶媒体は、コンピュータ可読記憶装置、コンピュータ可読ストレージ回路基板、ランダム若しくはシリアルアクセスメモリアレイ若しくはデバイス、またはそれらのうちの1つまたは複数の組合せであり得る、またはそれらに備えられ得る。さらに、コンピュータ記憶媒体は伝搬シグナルではないが、コンピュータ記憶媒体は、人工的に生成された伝搬シグナルに符号化されたコンピュータプログラム命令の送信元または送信先であり得る。コンピュータ記憶媒体は、1つまたは複数の個別の物理コンポーネントまたはメディア、例えば、複数のCD、ディスク、または他の記憶装置でもあり得る、またはそれらに備えられ得る。

10

【0083】

本明細書に記載の処理は、1つまたは複数のコンピュータ可読記憶装置に記憶されたデータ、または他のソースから受信したデータについて、データ処理装置によって実行される処理として実施され得る。

【0084】

「データ処理装置」という用語は、一例として、プログラマブル処理ユニット、コンピュータ、システム・オン・チップ、パーソナルコンピュータシステム、デスクトップコンピュータ、ラップトップ、ノートブック、ネットブックコンピュータ、メインフレームコンピュータシステム、携帯コンピュータ、ワークステーション、ネットワークコンピュータ、アプリケーションサーバ、記憶装置、カメラ、ビデオカメラ、セットトップボックス、モバイルデバイス、ビデオゲームコンソール、携帯ビデオゲームデバイスなどの家庭用電子デバイス、スイッチ、モ뎀、ルータ、または一般的な任意のタイプのコンピュータデバイスまたは電子デバイスなどの周辺デバイス、または複数個からなるそれら、またはそれらの組合せを含む、あらゆる種類の装置、デバイス、およびデータ処理機器を包含する。装置は、専用論理回路、例えば、FPGA(フィールド・プログラマブル・ゲート・アレイ)またはASIC(特定用途向け集積回路)を備え得る。装置は、ハードウェアに加えて、当該コンピュータプログラム用実行環境を生成するコード、例えば、プロセッサファームウェア、プロトコルスタック、データベース管理システム、オペレーティングシステム、クロスプラットフォームランタイム環境、仮想マシン、またはそれらのうちの1つまたは複数の組合せの構成要素であるコードをも含み得る。装置および実行環境は、様々な異なる、ウェブサービス、分散コンピューティングなどのコンピューティングモデルインフラ、およびグリッドコンピューティングインフラを実現し得る。

20

【0085】

(プログラム、ソフトウェア、ソフトウェアアプリケーション、スクリプト、またはコードとしても知られる)コンピュータプログラムは、コンパイルまたはインタプリタ言語、宣言型または手続き型言語を含むプログラミング言語の任意の形式で記述され得るし、

40

50

スタンドアロンプログラムとして、またはモジュールとして、コンポーネント、サブルーチン、オブジェクト、またはコンピュータ環境においての使用に適した他のユニットを含む、任意の形式で展開され得る。コンピュータプログラムは、ファイルシステム内のファイルに対応し得るが、必ずしも対応させる必要はない。プログラムは、他のプログラムまたはデータ(例えば、マークアップ言語ドキュメントに格納された1つまたは複数のスクリプト)を保持しているファイルの一部、当該プログラム専用の単一のファイル、または複数の協調型ファイル(例えば、1つまたは複数のモジュール、サブプログラム、またはコードの一部を格納しているファイル)に格納され得る。コンピュータプログラムは、1つのサイトまたは分散された複数のサイトに位置し通信ネットワークによって相互接続されている、1つのコンピュータまたは複数のコンピュータにて実行されるように展開され得る。

10

【 0 0 8 6 】

本明細書に記載のプロセスおよび論理フローは、入力データを処理して出力を生成することによって処理を実施するために1つまたは複数のコンピュータプログラムを実行する、1つまたは複数のプログラマブル処理ユニットによって実施され得る。プロセスおよび論理フローは、専用論理回路、例えば、FPGA(フィールド・プログラマブル・ゲート・アレイ)またはASIC(特定用途向け集積回路)によって実施され得るし、装置は、専用論理回路、例えば、FPGA(フィールド・プログラマブル・ゲート・アレイ)またはASIC(特定用途向け集積回路)として実施され得る。

【 0 0 8 7 】

コンピュータプログラムの実行に適した処理ユニットは、一例として、汎用および専用マイクロプロセッサの双方、および任意の種類のデジタルコンピュータの任意の1つまたは複数の処理ユニットを含む。一般的に、処理ユニットは、リードオンリーメモリまたはランダムアクセスメモリまたはその双方から命令およびデータを受信する。コンピュータ必須要素は、命令に従って処理を実施する処理ユニットと、命令およびデータを記憶するための1つまたは複数のメモリデバイスとである。一般的に、コンピュータは、データを記憶するための1つまたは複数のマストレージデバイス、例えば、磁気、光磁気ディスク、または光ディスクを備え、またはそれらからデータを受信、またはそれらへと送信、またはその双方を行うために動作可能なよう接続される。しかしながら、コンピュータは、必ずしもこのようなデバイスである必要はない。さらに、コンピュータは、僅かばかりではあるが名前をあげると、例えば、携帯電話、携帯情報端末(PDA)、モバイルオーディオ若しくはビデオプレイヤ、ゲームコンソール、全地球測位網(GPS)受信機、ネットワークルーティングデバイス、または携帯記憶装置(例えば、ユニバーサルシリアルバス(USB)フラッシュドライブ)といった、別のデバイスに組み込まれ得る。コンピュータプログラム命令およびデータを記憶するのに適したデバイスは、一例として、例えば、EPROM、EEPROM、およびフラッシュメモリデバイスといった半導体メモリデバイス、例えば、内蔵ハードディスクまたはリムーバルディスクといった磁気ディスク、光磁気ディスク、並びにCD-ROMおよびDVD-ROMディスクを含む、不揮発性メモリ、メディア、およびメモリデバイスの全ての形式を含む。処理ユニットおよびメモリは、専用論理回路によって保管され得る、または専用論理回路に内蔵され得る。

20

【 0 0 8 8 】

ユーザとのインタラクションを提供するために、本明細書に記載の発明特定事項の実施形態は、例えばCRT(陰極線管)またはLCD(液晶ディスプレイ)モニタといった、情報をユーザに表示するためのディスプレイデバイスと、例えばマウスまたはトラックボールとを有するコンピュータといった、ユーザがコンピュータに入力することを可能とするキーボードおよびポインティングデバイスとで実施され得る。他の種類のデバイスも、同様に、ユーザとのインタラクション用に提供されるために使用され得る。例として、ユーザに提供されるフィードバックは、例えば、視覚フィードバック、聴覚フィードバック、または触覚フィードバックといった感覚フィードバックの任意の形式であり得るし、ユーザからの入力は、聴覚、音声、または触覚入力を含む、任意の形式で受信され得る。加えて、コンピュータは、ユーザによって使用されるデバイスにドキュメントを送信しデバイスからド

40

50

キュメントを受信することによって、例えば、ウェブブラウザから受信した要求に応じてウェブページをユーザのクライアントデバイス上のウェブブラウザに送信することによって、ユーザとやりとりし得る。

【0089】

本明細書に記載の発明特定事項の実施形態は、例えばデータサーバのような、バックエンドコンポーネントを備える、または、例えばアプリケーションサーバのような、ミドルウェアコンポーネントを備える、または、例えば、ユーザがそれを介して本明細書に記載の発明特定事項の実施形態とやりとりすることができる、グラフィカルユーザインターフェースまたはウェブブラウザを有するクライアントコンピュータのような、フロントエンドコンポーネントを備えるコンピュータシステムにおいて、または、例えばネットワークルータのような、ルーティングデバイス、またはそのようなバックエンド、ミドルウェア、またはフロントエンドコンポーネントのうちの1つまたは複数の任意の組合せにおいて実施され得る。システムのコンポーネントは、デジタルデータ通信の任意の形式または媒体、例えば、通信ネットワークによって、相互接続され得る。通信ネットワークの例には、ローカルエリアネットワーク(「LAN」)および広域エリアネットワーク(「WAN」)、相互ネットワーク(例えば、インターネット)、並びにピア・ツー・ピアネットワーク(例えば、アドホックピア・ツー・ピアネットワーク)が含まれる。10

【0090】

コンピュータシステムは、クライアントおよびサーバを含み得る。クライアントおよびサーバは、通常は互いに隔離されており、一般的に通信ネットワークを介してやりとりを行う。それぞれのコンピュータで実行され、互いにクライアント-サーバ関係を有するコンピュータプログラムのおかげで、クライアントとサーバとの関係が生まれる。いくつかの実施形態では、サーバは、データ(例えば、HTMLページ)を(例えば、データを表示し、クライアントデバイスとやり取りをするユーザからユーザ入力を受信するため)クライアントデバイスに送信する。クライアントデバイスで生成されたデータ(例えば、ユーザインタラクションの結果)は、サーバにおいて、クライアントデバイスから受信され得る。20

【0091】

1つまたは複数のコンピュータのシステムは、処理においてシステムに動作を実施させる、システムにインストールされた、ソフトウェア、ファームウェア、ハードウェア、またはそれらの組合せを有することによって、特定の処理を実施するように構成され得る。1つまたは複数のコンピュータプログラムは、データ処理装置によって実行される際に装置に処理を実施させる命令を備えることによって、特定の動作を実施するように構成され得る。30

【0092】

電子ドキュメント(略して、単にドキュメントとも称される)は、必ずしも対応する必要があるわけではないが、ファイルに対応している。ドキュメントは、他のドキュメントを保持しているファイルの一部、当該ドキュメント専用の单一のファイル、または複数の協調型ファイルに格納され得る。

【0093】

本明細書は、具体的な実施形態の詳細を内包しているが、これらは、如何なる発明についてもその範囲またはその主張され得ることの限定として解釈されるべきではなく、むしろ特定の発明の特定の実施形態固有の特徴を記載しているものとして解釈されるべきである。個々の実施形態に関連して本明細書において記載されるある特徴は、单一の実施形態の組合せとしても実施され得る。反対に、单一の実施形態に関連して記載される様々な特徴は、それぞれ複数の実施形態として、または任意の適切なサブコンビネーションとしても実施され得る。さらに、特徴のある組合せにおいて実施されるものとして記載し、当初そのように主張していたとしても、主張した組合せにおける1つまたは複数の特徴は、いくつかの場合では、組合せから削除され得るし、主張した組合せは、サブコンビネーションまたはサブコンビネーションの変形を対象にし得る。4050

【 0 0 9 4 】

同様に、処理が特定の順序で図面に記載されているが、そのような処理が示した特定の順序または一連の順序で実施される必要があると、または所望の結果を達成するために記載した処理全てが実施される必要があると、理解するべきではない。ある状況では、マルチタスク処理および並列処理が有利となり得る。さらに、上述した実施形態における様々なシステムコンポーネントの区分は、全ての実施形態においてそのように区分される必要があるものとして理解されるべきではなく、記載したプログラムコンポーネントおよびシステムが、一般的に、単一のソフトウェア製品に一纏めにして統合され、または複数のソフトウェア製品にパッケージ化され得る、と理解されるべきである。

【 0 0 9 5 】

10

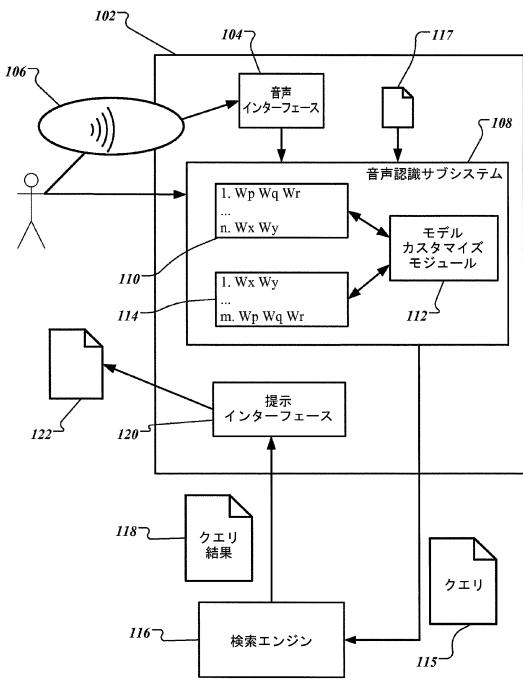
上に述べたように、発明特定事項の特定の実施形態をこれまで記載してきた。他の実施形態も、以下の特許請求の範囲内に存在する。いくつかの場合において、特許請求の範囲に記載の処理は、異なる順序で実施され得るし、それでも所望の結果を達成し得る。加えて、添付の図面に記載のプロセスは、所望の結果を達成するために、必ずしも示した特定の順序または一連の順序である必要はない。ある実施形態では、マルチタスク処理および並列処理が有利となり得る。それゆえ、他の実施形態も、以下の特許請求の範囲内に存在する。

【 符号の説明 】**【 0 0 9 6 】**

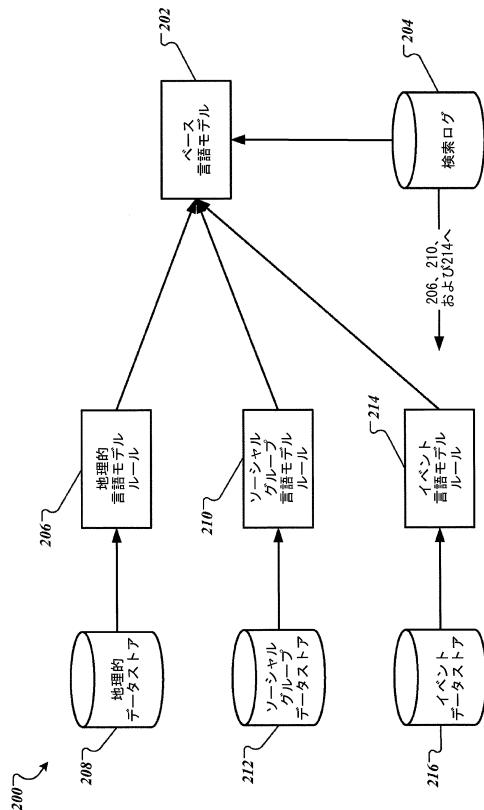
20

- 104 音声インターフェース
- 108 音声認識サブシステム
- 112 モデルカスタマイズモジュール
- 115 クエリ
- 116 検索エンジン
- 118 クエリ結果
- 120 提示インターフェース

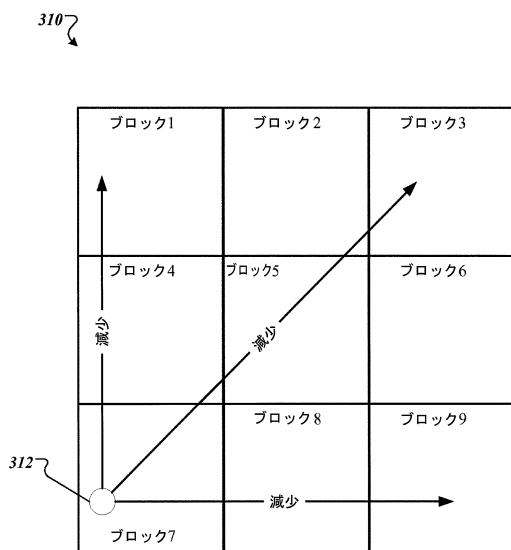
【図1】



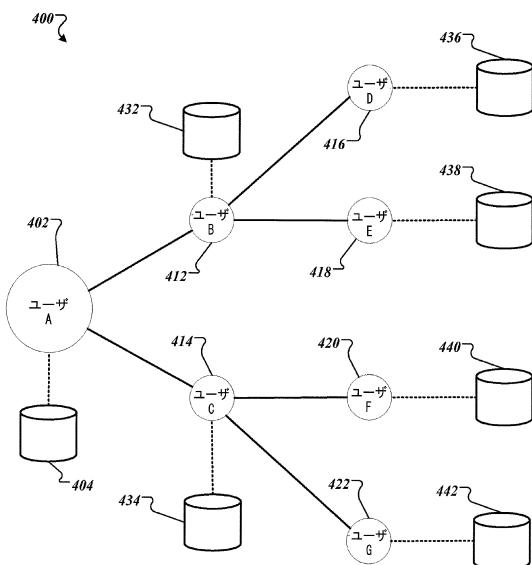
【図2】



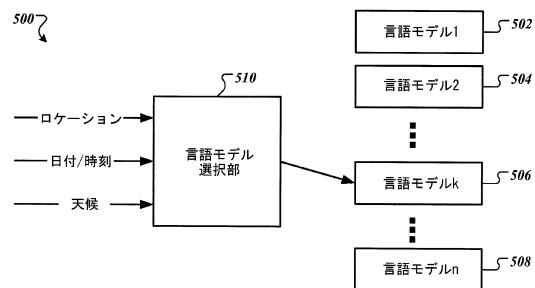
【図3】



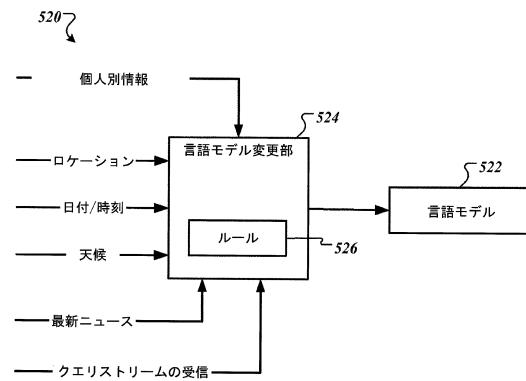
【図4】



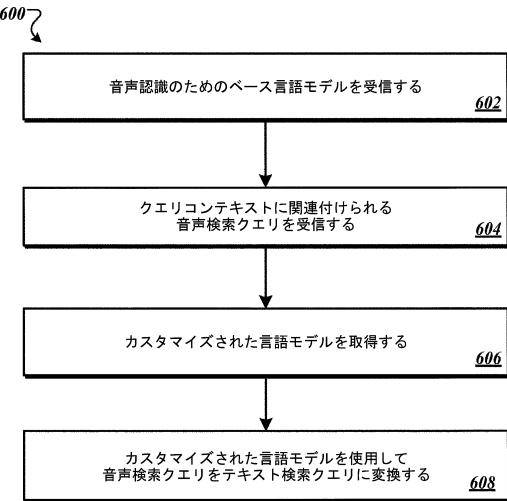
【図5A】



【図5B】



【図6】



フロントページの続き

(72)発明者 ペドロ・ジェー・モレーノ・メンギバー

アメリカ合衆国・ニュージャージー・07302・ジャージー・シティ・フィフス・ストリート・
310

(72)発明者 マイケル・エイチ・コーエン

アメリカ合衆国・カリフォルニア・94028・ポートラ・ヴァレー・ヒルブルック・ドライブ・
21

審査官 吉田 誠

(56)参考文献 特開2002-229584(JP,A)

米国特許出願公開第2011/0144973(US,A1)

特表2002-538534(JP,A)

米国特許出願公開第2011/0060587(US,A1)

米国特許出願公開第2011/0093265(US,A1)

特許第3955880(JP,B2)

特表2002-510076(JP,A)

国際公開第2006/059451(WO,A1)

特開2005-141049(JP,A)

(58)調査した分野(Int.Cl., DB名)

G06F 17/30

G10L 15/00

G10L 15/065

G10L 15/197