

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2013-178384

(P2013-178384A)

(43) 公開日 平成25年9月9日(2013.9.9)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 O L 15/183 (2013.01)	G 1 O L 15/18 2 0 0 Z	
G 1 O L 15/18 (2013.01)	G 1 O L 15/18 3 0 0 G	
G 1 O L 15/06 (2013.01)	G 1 O L 15/06 4 0 0 Z	

審査請求 未請求 請求項の数 9 O L (全 14 頁)

(21) 出願番号	特願2012-42144 (P2012-42144)	(71) 出願人	000004237
(22) 出願日	平成24年2月28日 (2012. 2. 28)		日本電気株式会社
			東京都港区芝五丁目7番1号
		(74) 代理人	100103090
			弁理士 岩壁 冬樹
		(74) 代理人	100124501
			弁理士 塩川 誠人
		(72) 発明者	稲垣 敬子
			東京都港区芝五丁目7番1号 日本電気株式会社社内

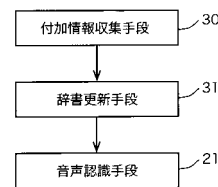
(54) 【発明の名称】 音声認識用辞書登録システム、音声認識システム、音声認識サービスシステム、方法およびプログラム

(57) 【要約】

【課題】複数のユーザで共通の言語モデルを使用する構成でありながらも、個々のユーザに適した音声認識を行うことができ、かつ音声入力時の状況を考慮して音声認識の精度を向上させることができるシステムを提供する。

【解決手段】入力された音声に関連する付加情報を収集する付加情報収集手段30と、音声認識による音声の認識結果と付加情報収集手段30が収集した付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新手段31と、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、音声の音声認識を行う音声認識手段21とを備えている。

【選択図】 図8



【特許請求の範囲】**【請求項 1】**

入力された音声に関連する付加情報を収集する付加情報収集手段と、
音声認識による前記音声の認識結果と前記付加情報収集手段が収集した前記付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新手段とを備えたことを特徴とする音声認識用辞書登録システム。

【請求項 2】

音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識手段を備えたことを特徴とする音声認識システム。 10

【請求項 3】

入力された音声に関連する付加情報を収集する付加情報収集手段と、
音声認識による前記音声の認識結果と前記付加情報収集手段が収集した前記付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新手段と、
音声に対応付けられた付加情報と前記音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識手段とを備えたことを特徴とする音声認識サービスシステム。

【請求項 4】

辞書更新手段は、音声認識手段による音声の認識結果が修正された修正後の認識結果と付加情報収集手段が収集した該音声に関連する付加情報とに基づいて、音声認識用辞書を更新する 20

請求項 3 記載の音声認識サービスシステム。

【請求項 5】

付加情報収集手段は、入力された音声に関連する複数種類の付加情報を収集し、
音声認識手段は、前記付加情報収集手段が収集した複数種類の付加情報のうち、あらかじめ定められた優先順位にしたがって選択した付加情報と、音声認識用辞書とに基づいて、前記音声の音声認識を行う

請求項 3 又は請求項 4 記載の音声認識サービスシステム。

【請求項 6】

音声認識用辞書は、単語に対応付けて付加情報を登録するための付加情報登録辞書と、言語モデルとを含み、 30

辞書更新手段は、音声の認識結果と付加情報収集手段が収集した該音声に関連する付加情報とに基づいて、前記付加情報登録辞書を更新し、

音声認識手段は、入力された音声に関連する付加情報と前記付加情報登録辞書および前記言語モデルとに基づいて、該入力された音声の音声認識を行う

請求項 3 から請求項 5 のうちのいずれか 1 項に記載の音声認識サービスシステム。

【請求項 7】

音声認識手段は、入力された音声に関連する付加情報と付加情報登録辞書とに基づいて、言語モデルに含まれる単語の出現確率を更新し、該入力された音声の音声認識を行う

請求項 3 から請求項 6 のうちのいずれか 1 項に記載の音声認識サービスシステム。 40

【請求項 8】

入力された音声に関連する付加情報を収集し、

音声認識による前記音声の認識結果と収集した前記付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新し、

音声に対応付けられた付加情報と前記音声認識用辞書とに基づいて、該音声の音声認識を行う

ことを特徴とする音声認識サービス方法。

【請求項 9】

コンピュータに、

入力された音声に関連する付加情報を収集する付加情報収集処理と、 50

音声認識による前記音声の認識結果と前記付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新処理と、

音声に対応付けられた付加情報と前記音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識処理とを

実行させるための音声認識サービスプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声認識用の辞書登録を行う音声認識用辞書登録システム、音声認識を行う音声認識システム、音声認識サービスシステム、方法およびプログラムに関する。

10

【背景技術】

【0002】

モバイル端末に入力された音声を認識して日報作成を行うシステムにおいて、音声認識の精度を向上させるためには、ユーザに適した単語を選択可能な辞書データ（以下、単に辞書という）を使用することが重要である。

【0003】

音声認識に用いられる辞書に登録されている単語が増えすぎると、認識処理の遅延や類似単語への誤認識による精度劣化が起こりやすくなる。また、辞書に登録されている単語が少ない場合には、当該単語を認識することができず、認識精度が劣化する。音声入力の内容は、ユーザ毎に異なる。そのため、一般的には、全てのユーザに適用する共通辞書（以下言語モデル）とは別に、個人別のユーザ辞書が用いられることが多い。

20

【0004】

一般的な音声認識システムでは、言語モデルは、認識エンジンと共に提供されている。また、ユーザが単語を登録する場合には、別途ユーザ辞書が作成されている。このユーザ辞書に登録されるのは、単語の表記、読みおよび品詞情報のみである。そのため、言語モデルに単語が登録される時よりも認識精度が落ちるという問題がある。この問題を解決するため、最近では、ユーザが登録したい単語を言語モデルに登録できる仕組みが提供されている。

【0005】

関連する技術として、例えば特許文献1に記載されたシステムは、携帯端末でメール文を作成する場合に、入力された音声に基づいて文字列の作成を行う。特許文献1に記載されたシステムは、携帯端末とネットワーク上にある音声認識サーバとから構成されている。音声認識サーバは、携帯端末から送信された音声データを認識して文字データを生成し、生成した文字データを携帯端末に送信する。

30

【0006】

また、特許文献1には音声認識に用いる辞書を更新する技術が記載されている。特許文献1に記載されたシステムでは、音声認識サーバから携帯端末に送信された認識結果がユーザによって修正された後、修正結果が音声認識サーバに送信されると、音声認識サーバは、修正結果に基づいて音声認識用の辞書を修正する。

【先行技術文献】

40

【特許文献】

【0007】

【特許文献1】特開2009-075582号公報

【発明の概要】

【発明が解決しようとする課題】

【0008】

言語モデルは、学習対象として用意された大量の文章が解析され、所定の統計情報処理が行われることにより作成される。言語モデルは、辞書に記述された各単語について、出現確率や接続確率をデータ化したものである。

【0009】

50

通常話し方はユーザ毎に異なるため、各単語の出現確率や接続確率もユーザ毎に異なる。したがって、音声認識の精度の向上を図るためには、このような違いを考慮して音声認識を行うことが望ましい。しかし、言語モデルをユーザ毎に作成することは、システムリソースの問題や運用の問題などから困難である。

【0010】

そのため、ある特定のグループで共通の言語モデルを使用し、各ユーザが個別にその言語モデルに単語や文章を登録するという運用が考えられている。しかし、この場合には、ユーザが言語モデルに単語を登録し、その単語を認識しやすくするためには、該当する単語が入った大量の文章を登録しないと効果が出にくいという問題がある。また、人により登録したい単語が異なる場合には、その調整ができず、結果的に、大量に登録された単語が出やすくなるという問題がある。

10

【0011】

また、日報作成システム等に入力する音声入力の内容は、同じユーザであっても状況（例えば顧客や商談のフェーズなど）によって異なる。したがって、音声認識の精度を向上させるためには、各単語の出現確率や接続確率はそれらを考慮して決められることが必要になる。

【0012】

特許文献1に記載されたシステムは、辞書を更新する手段を備えているが、文字列を修正した部分からのみ辞書を修正するため、修正していない情報を追加することが出来ない。また、特許文献1に記載されたシステムは、音声入力を行う状況（例えば日報作成時の顧客や商談のフェーズの違い）を考慮して言語モデルを作成することはできない。すなわち、特許文献1に記載されたシステムは、顧客毎に異なる用語や商談フェーズによる用語の使い分けを行うことによって音声認識の精度を向上させることができない。

20

【0013】

そこで、本発明は、複数のユーザで共通の言語モデルを使用する構成でありながらも、個々のユーザに適した音声認識を行うことができ、かつ音声入力時の状況を考慮して音声認識の精度を向上させることができる音声認識用辞書登録システム、音声認識システム、音声認識サービスシステム、音声認識サービス方法および音声認識サービスプログラムを提供することを目的とする。

【課題を解決するための手段】

30

【0014】

本発明による音声認識用辞書登録システムは、入力された音声に関連する付加情報を収集する付加情報収集手段と、音声認識による音声の認識結果と付加情報収集手段が収集した付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新手段とを備えたことを特徴とする。

【0015】

本発明による音声認識システムは、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識手段を備えたことを特徴とする。

【0016】

本発明による音声認識サービスシステムは、入力された音声に関連する付加情報を収集する付加情報収集手段と、音声認識による音声の認識結果と付加情報収集手段が収集した付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新手段と、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識手段とを備えたことを特徴とする。

40

【0017】

本発明による音声認識サービス方法は、入力された音声に関連する付加情報を収集し、音声認識による音声の認識結果と収集した付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新し、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行うことを特徴とする。

【0018】

50

本発明による音声認識サービスプログラムは、コンピュータに、入力された音声に関連する付加情報を収集する付加情報収集処理と、音声認識による音声の認識結果と付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する辞書更新処理と、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識処理とを実行させることを特徴とする。

【発明の効果】

【0019】

本発明によれば、複数のユーザで共通の言語モデルを使用する構成でありながらも、個々のユーザに適した音声認識を行うことができ、かつ音声入力時の状況を考慮して音声認識の精度を向上させることができる。

【図面の簡単な説明】

【0020】

【図1】本発明による音声認識システムと音声認識用辞書登録システムとが適用された音声認識サービスシステムの構成の一例を示すブロック図である。

【図2】音声認識システムの音声認識手段および音声認識用辞書の構成例を示すブロック図である。

【図3】言語モデル内の単語に付加情報に基づいて出現確率の重み付けをするために用いられる追加単語辞書の具体例を示す説明図である。

【図4】音声認識用辞書登録システムの動作手順を示すフローチャートである。

【図5】一般的な音声認識システムの構成を示すブロック図である。

【図6】音声認識システムと音声認識用辞書登録システムとが適用された営業日報システムの構成の一例を示すブロック図である。

【図7】日報データベースに含まれる日報データの具体例を示す説明図である。

【図8】音声認識サービスシステムの最小の構成例を示すブロック図である。

【発明を実施するための形態】

【0021】

以下、本発明の実施形態を図面を参照して説明する。図1は、本発明による音声認識システムと音声認識用辞書登録システムとが適用されたシステムの構成の一例を示すブロック図である。図1に示されるように、本発明による音声認識システムと音声認識用辞書登録システムとが適用されたシステムは、モバイル端末1と音声認識システム2と音声認識用辞書登録システム3とを含む。また、モバイル端末1と音声認識システム2と音声認識用辞書登録システム3とは、LANやインターネット等の通信ネットワークを介して相互に接続されている。

【0022】

なお、本実施形態では、音声認識システム2と音声認識用辞書登録システム3とが異なるシステムとして構成されている例を説明するが、本発明の適用形態はこれに限られない。例えば、辞書登録可能な音声認識サービスシステムとして、図1の音声認識システム2および音声認識用辞書登録システム3の各手段を全て備えるように構成されていてもよい。

【0023】

図1に示されるように、モバイル端末1は、音声入力手段11、付加情報取得手段12、データ送信手段13、認識結果表示手段14および認識結果修正手段15を備えている。モバイル端末1は、具体的には、携帯電話機やスマートフォン、ノート型パーソナルコンピュータ等の端末装置によって実現される。

【0024】

音声入力手段11は、マイクロフォン等によって実現され、モバイル端末1に音声を入力する機能を備えている。

【0025】

付加情報取得手段12は、音声入力手段11が入力する音声に関連する付加情報として、モバイル端末1の位置情報や端末ログイン情報、仕事のフェーズ情報などを取得する機

10

20

30

40

50

能を備えている。付加情報取得手段12は、例えば、GPS衛星からGPS信号を受信することによって位置情報を取得する。また、付加情報取得手段12は、例えば、ユーザがモバイル端末の入力装置を操作することによって入力された端末ログイン情報や仕事のフェーズ情報などの付加情報を取得する。

【0026】

データ送信手段13は、音声入力手段11が入力した音声データや、付加情報取得手段12が取得した付加情報、認識結果修正手段15が修正した認識結果情報を、ネットワークを介して、音声認識システム2又は音声認識用辞書登録システム3に送信する機能を備えている。データ送信手段13は、例えば、プログラムに従って動作するモバイル端末1のネットワークインタフェース部などによって実現される。

10

【0027】

認識結果表示手段14は、音声認識システム2から受信した認識結果情報を表示する機能を備えている。認識結果表示手段14は、例えば、ディスプレイ装置などの表示装置によって実現される。

【0028】

認識結果修正手段15は、音声認識システム2から受信した認識結果情報の誤りを修正する機能を備えている。認識結果修正手段15は、例えば、ユーザの入力操作に従って、認識結果情報の誤りを修正する。認識結果修正手段15は、具体的には、プログラムに従って動作するモバイル端末1のCPUによって実現される。

20

【0029】

また、図1に示されるように、音声認識システム2は、音声認識手段21および音声認識用辞書記憶手段22を含む。音声認識システム2は、具体的には、プログラムに従って動作するサーバ装置などの情報処理装置によって実現される。

【0030】

音声認識手段21は、モバイル端末1から受信した音声データを分析し、文字データに変換する機能を備えている。具体的には、音声認識手段21は、音声データを分析し、分析結果に基づいて文字データを生成する。音声認識手段21は、プログラムに従って動作する情報処理装置のCPUによって実現される。

【0031】

音声認識用辞書記憶手段22は、音声認識処理で音声データを分析して文字データを生成する際に用いられる音声認識用辞書を記憶している。音声認識用辞書記憶手段22は、具体的には、光ディスク装置や磁気ディスク装置などの記憶装置によって実現される。

30

【0032】

また、図1に示されるように、音声認識用辞書登録システム3は、付加情報収集手段30、辞書更新手段31および認識結果データベース記憶手段32を含む。音声認識用辞書登録システム3は、具体的には、プログラムに従って動作するサーバ装置などの情報処理装置によって実現される。

【0033】

付加情報収集手段30は、音声の認識結果に対応付けられた付加情報を収集する機能を備えている。付加情報収集手段30は、例えば、モバイル端末1の付加情報取得手段12によって取得され、音声認識用辞書登録システム3に送信された付加情報を受信する。付加情報収集手段30は、例えば、プログラムに従って動作する情報処理装置のCPUやネットワークインタフェース部などによって実現される。

40

【0034】

辞書更新手段31は、モバイル端末1で修正された認識結果情報に基づいて、音声認識用辞書記憶手段22が記憶する音声認識用辞書を更新する機能を備えている。辞書更新手段31は、例えば、プログラムに従って動作する情報処理装置のCPUやネットワークインタフェース部などによって実現される。

【0035】

認識結果データベース記憶手段32は、音声認識用辞書の更新に用いられる認識結果情

50

報が蓄積された認識結果データベースを記憶する。認識結果データベース記憶手段32は、具体的には、光ディスク装置や磁気ディスク装置などの記憶装置によって実現される。

【0036】

本実施形態では、音声認識手段21は、モバイル端末1から受信した音声データを分析し、音声認識用辞書記憶手段21が記憶する音声認識用辞書を参照して、音声と最も近い語彙を推定し文字に変換する。具体的には、音声認識手段21は、音声データを分析し、音声認識用辞書を参照して、文字データ(すなわち、認識結果情報)を生成する。

【0037】

音声認識手段21が変換した文字データは、モバイル端末1に送信され、認識結果表示手段14に表示される。そして、ユーザは表示された文字データを確認し、修正がある場合には入力装置を用いて修正操作を行う。すると、認識結果修正手段15は、ユーザの修正操作に従って、認識結果情報の修正を行う。その後、認識結果修正手段15によって修正された認識結果情報は、データ送信手段13によって音声認識用辞書登録システム3に送信され、認識結果データベース記憶手段32に蓄積される。

10

【0038】

また、本実施形態では、音声認識用辞書登録システム3は、自動的に(例えば所定期間ごとに)又はユーザの手動操作によって、音声認識用辞書22を更新する。

【0039】

図2は、図1の音声認識システム2の音声認識手段21および音声認識用辞書22の構成例を示すブロック図である。図2に示すように、音声認識手段21は、音声検出部211、音声分析部212および音声照合部213を含む。また、音声認識用辞書記憶手段22は、音響モデル221、言語モデル222、ユーザ単語辞書223および追加登録辞書224を記憶している。

20

【0040】

音声検出部211は、入力された音声データから音声と雑音とを切り分け、音声を含む区間を検出して、音声分析部212に出力する。音声検出部211は、例えば音声データのパワーを使う方法で音声の検出を行う。具体的には、音声検出部211は、音声データのパワーを逐次計算して、音声データのパワーがあらかじめ定められた閾値を一定時間連続して上回る時点を音声の開始時点と判定する。また、音声検出部211は、音声データのパワーがあらかじめ定められた閾値を一定時間連続して下回る時点を音声の終了時点と判定する。音声検出部211は、音声開始時点から音声終了時点までを音声として切り出し、音声分析部212に逐次出力する。

30

【0041】

音声分析部212は、音声検出部211により切り出された音声の音響分析を行い、音声の特徴を表す音響的特徴を音声照合部213に出力する。音声分析部212は、例えばスペクトル分析などによって音響分析を行う。本実施形態で音声分析部212によって出力される音響的特徴は、具体的には、後述する音声照合部213において候補としての単語列を検出できる特徴を示す情報である。音声分析部212は、音響的特徴として、例えば、ケプストラムやスペクトルピッチ、パワー等、音声データの音声波形に基づいて抽出される情報を音声照合部213に出力する。

40

【0042】

音響モデル221は、日本語の音素毎の標準的なパターンを含む。音響モデル211は、例えば隠れマルコフモデルが適用される。

【0043】

言語モデル222は、単語の出現確率がデータ化されたものであって、日本語の単語間や音素間の接続関係の規定や単語間の接続関係を規定する文法規則などを含む。本実施形態では、言語モデル222に含まれる単語の出現確率は、追加登録辞書224に基づいて、動的に変化する。

【0044】

ユーザ辞書223は、ユーザが辞書に単語を登録する時に用いられる辞書である。ユー

50

辞書 2 2 3 には、ユーザ自身が単語の表記、読みおよび品詞情報を登録することができる。なお、品詞情報は、あらかじめ定められた品詞の中から選択されて登録される。

【 0 0 4 5 】

追加登録辞書 2 2 4 は、言語モデル 2 2 2 に含まれる単語の出現確率を動的に変化させる際に用いられる。追加登録辞書 2 2 4 は、単語と付加情報とが対応付けられたデータを含む。本実施形態では、追加登録辞書 2 2 4 には、辞書更新手段 3 1 によって、認識結果データベース記憶手段 3 2 に蓄積されているモバイル端末 1 で修正された認識結果情報に基づいて、単語の読み、品詞情報および付加情報が登録されている。

【 0 0 4 6 】

図 3 は、言語モデル 2 2 2 内の単語に付加情報に基づいて出現確率の重み付けをするために用いられる追加単語辞書 2 2 4 の具体例を示す説明図である。図 3 に示すように、追加単語辞書には、単語の表記、読み、品詞および付加情報（担当情報、位置情報、仕事フェーズなど）が登録されている。図 3 に示す追加単語辞書 2 2 4 では、表記「検収」と「研修」とは、読みおよび品詞が同じであるが、付加情報に含まれる仕事ステータスが「5」と「6」とで異なっている。

10

【 0 0 4 7 】

本実施形態では、モバイル端末 1 は、入力された音声データにも付加情報を付与して音声認識システム 2 に送信する。そのため、音声認識システム 2 は、音声認識を行う際に、入力された音声データに付加された付加情報と追加単語辞書 2 2 4 の付加情報とが一致している単語について、言語モデル 2 2 2 中の該当する単語の出現確率に重み付けをして音声認識処理を行う。

20

【 0 0 4 8 】

例えば、音声データが「ケンシュウが必要です。」である場合を想定する。この場合、一般的な音声認識システムでは、「検収」と「研修」との使い分けは、言語モデル作成時の文例にどちらが多く入っているか（文例中の単語の出現確率）によって決定される。これに対して、本実施形態では、モバイル端末 1 は、音声データ「ケンシュウが必要です。」とともに付加情報（例えば仕事フェーズ：5）を音声認識システム 2 に送信する。すると、追加単語辞書 2 2 4 のうち「検収」と「研修」の読みおよび品詞が一致しているが、受信した付加情報と「検収」に対応付けられている付加情報とが一致しているため、言語モデル 2 2 2 中の「検収」の出現確率が上げられ、「研修」ではなく「検収」が選択されるようになる。

30

【 0 0 4 9 】

上述のように、本実施形態では、言語モデル 2 2 2 に含まれる単語の出現確率は動的に変化する。この出現確率の動的な変化は、例えば次のように実現される。

【 0 0 5 0 】

音声認識手段 2 1 は、例えば、言語モデル 2 2 2 における単語の出現確率に重み付けを行う重み付け手段（図示せず）を備える。重み付け手段は、音声認識を行う際に、追加単語辞書 2 2 4 を参照し、受信した音声データと付加情報とに基づいて、読み、品詞および付加情報が一致する単語を選択する。そして、重み付け手段は、選択した単語について、言語モデル 2 2 2 における単語の出現確率に重み付けし、出現確率を上げる。

40

【 0 0 5 1 】

なお、上記の例に限らず、音声認識手段 2 1 は、追加単語辞書 2 2 4 を参照し、受信した音声データと付加情報とに基づいて単語を選択する際に、複数の付加情報のうちのいずれかを優先するようにしてもよい。例えば、あらかじめ付加情報の優先順位を定めておき、音声認識手段 2 1 は、読みおよび品詞が一致する場合には、複数の付加情報（例えば、位置情報や仕事フェーズなど）のうち、優先順位が高い付加情報（仕事フェーズ）が一致するものを優先する。

【 0 0 5 2 】

また、例えば、音声認識手段 2 1 は、複数の付加情報が一致する場合には、一致したものの合計をスコア等として求めて、最もスコアが高いものを優先して音声認識結果を求め

50

る際に用いるようにしてもよい。

【0053】

次に、音声認識システムと音声認識用辞書登録システムとが適用された音声認識サービスシステムの動作を説明する。図4は、音声認識サービスシステムの動作手順を示すフローチャートである。

【0054】

図4に示す例では、動作を開始するにあたって、ユーザがモバイル端末1の音声入力手段11に音声を入力する。このとき、付加情報取得手段12は、モバイル端末1の位置情報や端末ログイン情報、仕事のフェーズ情報などの付加情報を取得する。付加情報取得手段12は、例えば、ユーザがモバイル端末の入力装置を操作することによって入力された端末ログイン情報や仕事のフェーズ情報などの付加情報を取得する。

10

【0055】

音声入力手段11が音声を入力し、付加情報取得手段12が付加情報を取得すると、データ送信手段13は、音声データおよび付加情報を、音声認識手段21に送信する(ステップA1)。

【0056】

音声認識手段21は、音声データおよび付加情報を受信すると、音声認識用辞書22を参照して、音声と音声認識用辞書のデータとを照合し、音声認識を行う(ステップA2)。上述のように、本実施形態では、音声認識手段21は、音声認識を行う際に、入力された音声データに付加された付加情報と追加単語辞書224の付加情報とが一致している単語について、言語モデル222中の該当する単語の出現確率に重み付けをして音声認識処理を行う。

20

【0057】

次いで、音声認識システム2は、ステップA2の音声認識による認識結果として、文字データをモバイル端末1に送信する。モバイル端末1は、音声認識システム2から受信した認識結果を認識結果表示手段14に表示する。そして、モバイル端末1は、認識結果に修正が必要か否かを判断する(ステップA3)。例えば、ユーザが認識結果表示手段14に表示された認識結果を確認し、修正が必要か否かをを入力する操作を行うと、モバイル端末1は、ユーザの操作に従って、修正が必要か否かを判断する。

【0058】

修正が必要であると判断された場合(ステップA3のYES)、認識結果修正手段15は、例えばユーザの修正操作に従って、認識結果の誤りを修正する(ステップA4)。

30

【0059】

その後、データ送信手段13は、認識結果修正手段15が修正した認識結果と、ステップA1で入力した付加情報とを、音声認識用辞書登録システム3に送信する(ステップA5)。なお、ステップA3で修正が必要でないと判断された場合には(ステップA3のNO)、データ送信手段13は、音声認識システム2から受信した認識結果と、ステップA1で入力した付加情報とを、音声認識用辞書登録システム3に送信する(ステップA5)。そして、データ送信手段13が送信した認識結果と付加情報とは、音声認識用辞書登録システム3の認識結果データベース記憶手段32に蓄積される。例えば、音声認識用辞書登録システム3の付加情報収集手段30は、データ送信手段13が送信した認識結果と付加情報とを受信し、認識結果データベース記憶手段32に蓄積する。

40

【0060】

次いで、音声認識用辞書登録システム3の辞書更新手段31は、認識結果データベース記憶手段32に蓄積された認識結果と認識結果に付加された付加情報とに基づいて、追加登録辞書を作成する。そして、辞書更新手段31は、音声認識システム2の音声認識用辞書22を更新する(ステップA6)。具体的には、辞書更新手段31が作成した追加登録辞書を音声認識システム2に送信すると、音声認識システム2は、受信した追加登録辞書で追加登録辞書224を更新する。

【0061】

50

なお、例えば、音声認識用辞書登録システム 3 は、無制限に付加情報が付加されていくことを防止するため、使用頻度が低い付加情報は音声認識用辞書 2 2 から随時削除するようにしてもよい。また、例えば、音声認識用辞書登録システム 3 は、逆に使用頻度が高いもののみを付加情報として登録していくようにしてもよい。

【 0 0 6 2 】

図 5 は、一般的な音声認識システムの構成を示すブロック図である。一般的な音声認識システムは、本発明のように、認識結果を用いて音声認識用辞書を更新する際に認識結果に付加情報を収集する手段を備えていない。したがって、図 5 に示されるような一般的な音声認識システムは、付加情報に基づいて単語の重み付けを行うことができず、同じような言い回しや音が似ている単語を含む文が登録されると、それぞれ区別されることなく出現確率が上がり、該当する単語が出やすくなってしまふ。

10

【 0 0 6 3 】

次に、本発明の他の適用例を図面を参照して説明する。図 6 は、音声認識システムと音声認識用辞書登録システムとが適用された営業日報システムの構成の一例を示すブロック図である。図 6 に示す適用例では、図 1 に示される構成に加えて、営業日報システム 4 が含まれている。また、営業日報システム 4 は、日報作成手段 4 1 および日報データベース記憶手段 4 2 を含む。

【 0 0 6 4 】

図 6 に示す適用例では、ユーザがモバイル端末 1 に日報データを入力すると（例えば、図 4 に示すように音声を入力し、音声認識を行うことによって日報データを入力する）、入力された日報データは、営業日報システム 4 に送信され、日報データベース 4 2 に蓄積される。そして、日報データベース 4 2 に蓄積された日報データは、認識結果データベース 3 2 に登録され、音声認識用辞書 2 2 を作成・更新する際に用いられる。

20

【 0 0 6 5 】

音声認識用辞書 2 2 を作成・更新する具体例を、図 7 に示される日報データベースに含まれる日報データの一例を参照して説明する。営業日報システム 4 は、図 7 に示す日報データベースから、備考欄に含まれるデータを音声認識結果情報として抽出する。また、営業日報システム 4 は、訪問先やフェーズ、報告者などの情報を付加情報として抽出する。そして、営業日報システム 4 は、抽出した音声認識結果情報と付加情報とを音声認識用辞書登録システム 3 の付加情報収集手段 3 0 に送信する。すると、付加情報収集手段 3 0 は、受信した音声認識結果情報と付加情報とを認識結果データベース 3 2 に登録する。これらの処理は、自動的に又はユーザの手動操作に従って行われる。その後、図 4 のステップ A 6 と同様に、音声認識用辞書登録システム 3 の辞書更新手段 3 1 は、認識結果データベース記憶手段 3 2 に蓄積された認識結果と認識結果に付加された付加情報とに基づいて、追加登録辞書を作成する。そして、辞書更新手段 3 1 は、音声認識システム 2 の音声認識用辞書 2 2 を更新する。

30

【 0 0 6 6 】

以上に説明したように、本発明による音声認識用辞書登録システムは、モバイル端末を活用して営業日報等を作成するシステム等において、単語や文章を音声認識用の辞書に登録する際に、その単語を使ったユーザの情報や使われた日時情報、位置情報、商談情報などを付加情報として登録する。また、本発明による音声認識システムは、音声認識を行う際に、それら付加情報を考慮した単語の選択をするため、音声認識システムは、適切な音声認識結果を提示することができる。

40

【 0 0 6 7 】

したがって、音声認識用辞書登録システムは、複数のユーザが音声認識用の辞書を共同でメンテナンスする場合でも、個々のユーザにあった単語が選択されるような音声認識用の辞書を作成することができる。また、音声認識システムは、音声入力時の状況（顧客や商談フェーズなど）を考慮して異なる単語を選択し、音声認識の精度を向上させることができる。

【 0 0 6 8 】

50

次に、本発明による音声認識サービスシステムの最小構成を説明する。図8は、音声認識サービスシステムの最小の構成例を示すブロック図である。図8に示されるように、音声認識サービスシステムは、最小の構成要素として、付加情報収集手段30と、辞書更新手段31と、音声認識手段21とを含む。

【0069】

図10に示す最小構成の音声認識サービスシステムでは、付加情報収集手段30は、入力された音声に関連する付加情報を収集する。また、辞書更新手段31は、音声認識による音声の認識結果と付加情報収集手段30が収集した付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書を更新する。そして、音声認識手段21は、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行う。

10

【0070】

したがって、最小構成の音声認識サービスシステムによれば、複数のユーザで共通の言語モデルを使用する構成であっても、個々のユーザに適した音声認識を行うことができ、かつ音声入力時の状況を考慮して音声認識の精度を向上させることができる。

【0071】

なお、本実施形態では、以下の(1)~(5)に示すような音声認識サービスシステムの特徴的構成が示されている。

【0072】

(1) 音声認識サービスシステム(例えば、図1に示される音声認識システム2および音声認識用辞書登録システム3によって実現される)は、入力された音声に関連する付加情報(例えば、位置情報や時間情報、ユーザ情報、顧客情報、仕事フェーズなど)を収集する付加情報収集手段(例えば、付加情報収集手段30によって実現される)と、音声認識による音声の認識結果と付加情報収集手段が収集した付加情報とに基づいて、音声認識の際に用いられる音声認識用辞書(例えば、音声認識辞書記憶手段22が記憶する音声認識辞書)を更新する辞書更新手段(例えば、辞書更新手段31によって実現される)と、音声に対応付けられた付加情報と音声認識用辞書とに基づいて、該音声の音声認識を行う音声認識手段(例えば、音声認識手段21によって実現される)とを備えたことを特徴とする。

20

【0073】

(2) 音声認識サービスシステムにおいて、辞書更新手段は、音声認識手段による音声の認識結果が修正された修正後の認識結果(例えば、認識結果修正手段15によって修正された修正後の認識結果)と付加情報収集手段が収集した音声に関連する付加情報とに基づいて、音声認識用辞書を更新するように構成されていてもよい。

30

【0074】

(3) 音声認識サービスシステムにおいて、付加情報収集手段は、入力された音声に関連する複数種類の付加情報を収集し、音声認識手段は、付加情報収集手段が収集した複数種類の付加情報のうち、あらかじめ定められた優先順位にしたがって選択した付加情報と、音声認識用辞書とに基づいて、音声の音声認識を行うように構成されていてもよい。

【0075】

(4) 音声認識サービスシステムにおいて、音声認識用辞書は、単語に対応付けて付加情報を登録するための付加情報登録辞書(例えば、追加登録辞書224)と、言語モデル(例えば、言語モデル222)とを含み、辞書更新手段は、音声の認識結果と付加情報収集手段が収集した該音声に関連する付加情報とに基づいて、付加情報登録辞書を更新し、音声認識手段は、入力された音声に関連する付加情報と付加情報登録辞書および前記言語モデルとに基づいて、該入力された音声の音声認識を行うように構成されていてもよい。

40

【0076】

(5) 音声認識サービスシステムにおいて、音声認識手段は、入力された音声に関連する付加情報と付加情報登録辞書とに基づいて、言語モデルに含まれる単語の出現確率を更新し(例えば、重み付け手段が処理を実行することによって実現される)、該入力された音声の音声認識を行うように構成されていてもよい。

50

【産業上の利用可能性】

【0077】

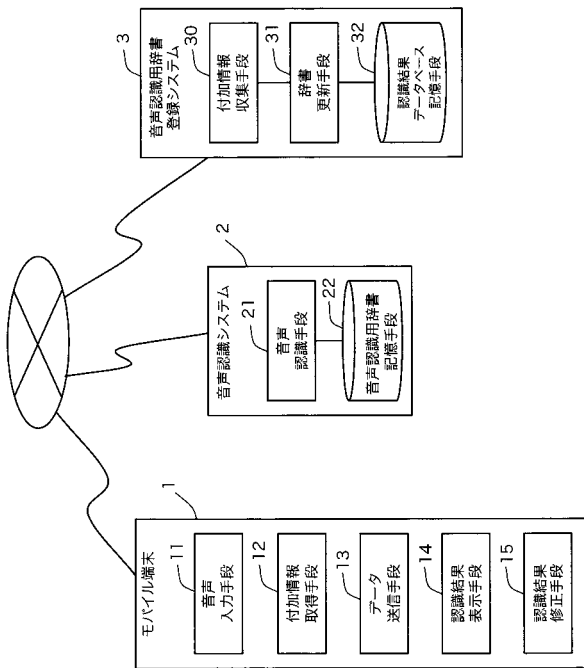
本発明は、メール文などの文字入力支援、コールセンターでの顧客とオペレータとの会話認識、音声認識による議事録作成など、話題が人や時間、場所によりある程度特定できる領域において、音声認識やテキスト入力支援、コンテンツ配信などを行うシステムに適用可能である。

【符号の説明】

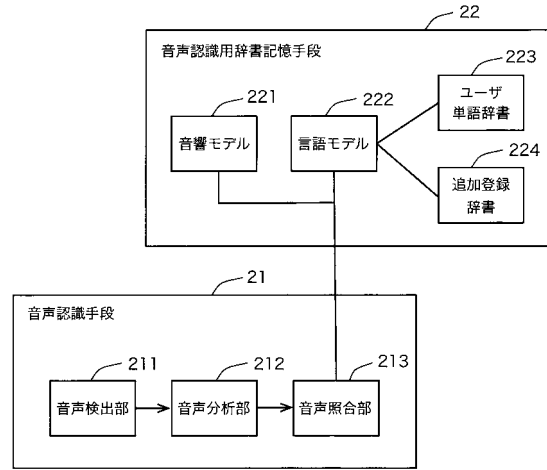
【0078】

- | | | |
|-------|----------------|----|
| 1 | モバイル端末 | |
| 1 1 | 音声入力手段 | 10 |
| 1 2 | 付加情報取得手段 | |
| 1 3 | データ送信手段 | |
| 1 4 | 認識結果表示手段 | |
| 1 5 | 認識結果修正手段 | |
| 2 | 音声認識システム | |
| 2 1 | 音声認識手段 | |
| 2 1 1 | 音声検出部 | |
| 2 1 2 | 音声分析部 | |
| 2 1 3 | 音声照合部 | |
| 2 2 | 音声認識用辞書記憶手段 | 20 |
| 2 2 1 | 音響モデル | |
| 2 2 2 | 言語モデル | |
| 2 2 3 | ユーザ単語辞書 | |
| 2 2 4 | 追加登録辞書 | |
| 3 | 音声認識用辞書登録システム | |
| 3 0 | 付加情報収集手段 | |
| 3 1 | 辞書更新手段 | |
| 3 2 | 認識結果データベース記憶手段 | |

【 図 1 】



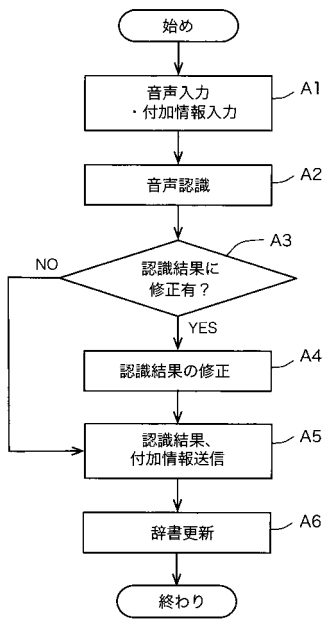
【 図 2 】



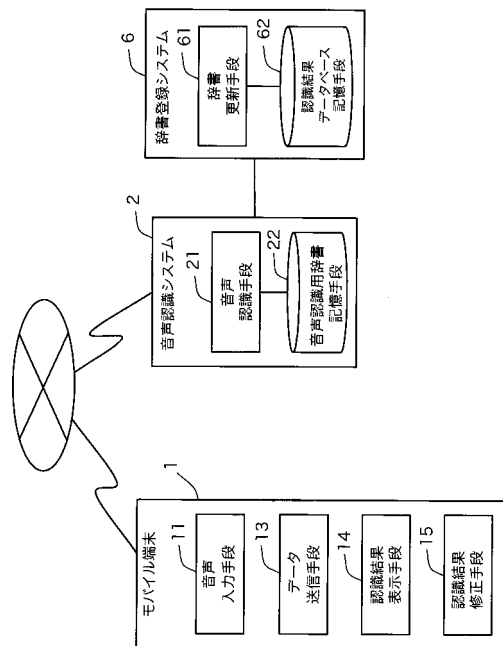
【 図 3 】

表記	読み	品詞	付加情報
AB社	えいびーしゃ	固有名詞(企業)	位置情報: 赤坂
AP社	えいびーしゃ	固有名詞(企業)	位置情報: 田町
エータ	えーた	固有名詞(製品名)	担当: 第一営業
ベータ	べーた	固有名詞(製品名)	担当: 第二営業
:	:	:	:
検収	けんしゅー	名詞	仕事フェーズ: 5
研修	けんしゅー	名詞	仕事フェーズ: 6

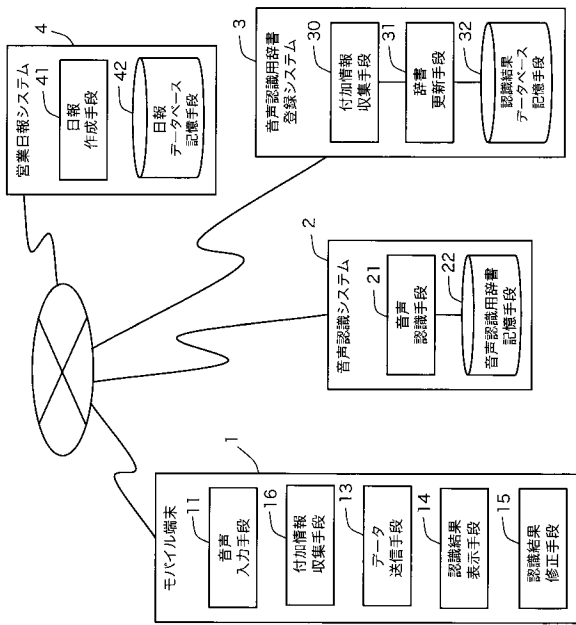
【 図 4 】



【 図 5 】



【 図 6 】



【 図 7 】

管理番号	訪問日	訪問先	担当者	報告者	フェーズ	次回訪問	備考
0001	10/9	AA社	鈴木様	山田	提案	..	サーバの入れ替え 商談
0002	10/9	BB社	佐藤様	武田	導入済	..	次回までに研修 日時を調整
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
0088	10/11	ZZ社	田中様	池田	納品	..	検収は月末までに

【 図 8 】

