

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4538954号
(P4538954)

(45) 発行日 平成22年9月8日(2010.9.8)

(24) 登録日 平成22年7月2日(2010.7.2)

(51) Int.Cl. F I
G 1 0 L 15/00 (2006.01) G 1 0 L 15/00 2 0 0 C
G 1 0 L 15/22 (2006.01) G 1 0 L 15/22 2 0 0 Z

請求項の数 4 (全 11 頁)

<p>(21) 出願番号 特願2000-600260 (P2000-600260) (86) (22) 出願日 平成12年2月17日(2000.2.17) (86) 国際出願番号 PCT/JP2000/000905 (87) 国際公開番号 W02000/049599 (87) 国際公開日 平成12年8月24日(2000.8.24) 審査請求日 平成19年2月9日(2007.2.9) (31) 優先権主張番号 特願平11-41452 (32) 優先日 平成11年2月19日(1999.2.19) (33) 優先権主張国 日本国(JP)</p>	<p>(73) 特許権者 000002185 ソニー株式会社 東京都港区港南1丁目7番1号 (74) 代理人 100067736 弁理士 小池 晃 (74) 代理人 100086335 弁理士 田村 榮一 (74) 代理人 100096677 弁理士 伊賀 誠司 (72) 発明者 田島 和彦 日本国東京都品川区北品川6丁目7番35号 ソニー株式会社内 審査官 井上 健一</p>
--	---

最終頁に続く

(54) 【発明の名称】 音声翻訳装置、音声翻訳方法及び音声翻訳制御プログラムを記録した記録媒体

(57) 【特許請求の範囲】

【請求項1】

入力された音声の特徴を抽出して特徴ベクトルを出力する特徴抽出手段と、
 上記特徴抽出手段からの特徴ベクトルに対して、音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声の音声認識を行い、認識した音声を示す文章を出力する音声認識照合手段と、
 上記音声認識照合手段が認識した音声を示す文章を、入力された音声の言語と異なる言語の文章に翻訳する翻訳処理手段と、
 上記翻訳処理手段により翻訳された言語の文章を、複数の修正情報から最適な修正情報を用いて修正する文修正手段と、
 上記文修正手段の修正結果を出力する出力手段と、
 上記音声認識照合手段が認識した文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積する履歴情報蓄積手段と、
 上記音声認識照合手段が音声認識結果を出力する毎に、当該音声認識結果と上記履歴情報蓄積手段に蓄積されている履歴情報とを対比して、累積学習機能により上記履歴情報蓄積手段に蓄積されている履歴情報を更新する累積学習手段とを備え、
 上記履歴情報蓄積手段は、蓄積している履歴情報に基づいて、上記音声認識情報手段が用いる照合情報の絞り込み及び/又は上記文修正手段が用いる修正情報の絞り込みを制御し、

上記累積学習手段は、上記履歴情報蓄積手段に蓄積される上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正手段が用いる修正情報の絞り込みを行うことを特徴とする音声翻訳装置。

【請求項 2】

上記履歴情報蓄積手段は、リセットにより、蓄積していた履歴情報を消去して、上記音声認識照合手段の認識結果を新たな履歴情報として蓄積することを特徴とする請求項 1 記載の音声翻訳装置。

【請求項 3】

入力された音声の特徴を抽出して特徴ベクトルを生成し、生成された特徴ベクトルに対して、音声認識用の複数の照合情報をそれぞれ照合して、上記入力された音声を認識する音声認識処理を行い、上記音声認識処理により認識した音声が表示文章を上記入力された音声の言語と異なる言語の文章に翻訳する翻訳処理を行い、上記翻訳処理により翻訳した文章を複数の修正情報から最適な修正情報を用いて修正する修正処理を行い、上記修正処理により修正した文章を出力するに当たり、

10

上記音声認識処理により、入力された音声の特徴を抽出して生成された特徴ベクトルに対して、上記音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声を認識し、認識した音声が表示文章を生成し、

上記音声認識処理により認識した音声が表示文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積し、

20

上記音声認識処理を行う毎に、当該音声認識結果と音声認識済みの文章の傾向を示す上記履歴情報とを対比して、累積学習機能により上記履歴情報を更新し、

更新された履歴情報に基づいて、上記音声認識処理に用いる照合情報の絞り込み及び/又は上記文修正処理に用いる修正情報の絞り込みを制御し、

上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正処理に用いる修正情報の絞り込みを行い、

上記翻訳処理により翻訳された文章を上記文修正処理により修正して出力することを特徴とする音声翻訳方法。

30

【請求項 4】

音声翻訳装置に搭載されるコンピュータにより、読み取り実行可能に音声翻訳制御プログラムを記録した記録媒体であって、

入力された音声の特徴を抽出して特徴ベクトルを生成し、生成された特徴ベクトルに対して、音声認識用の複数の照合情報をそれぞれ照合して、上記入力された音声を認識する音声認識処理を行い、上記音声認識処理により認識した音声が表示文章を上記入力された音声の言語と異なる言語の文章に翻訳する翻訳処理を行い、上記翻訳処理により翻訳した文章を複数の修正情報から最適な修正情報を用いて修正する修正処理を行い、上記修正処理により修正した文章を出力するに当たり、上記音声認識処理により、入力された音声の特徴を抽出して生成された特徴ベクトルに対して、上記音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声を認識し、認識した音声が表示文章を生成し、上記音声認識処理により認識した音声が表示文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積し、上記音声認識処理を行う毎に、当該音声認識結果と音声認識済みの文章の傾向を示す上記履歴情報とを対比して、累積学習機能により上記履歴情報を更新し、更新された履歴情報に基づいて、上記音声認識処理に用いる照合情報の絞り込み及び/又は上記文修正処理に用いる修正情報の絞り込みを制御し、上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正処理

40

50

に用いる修正情報の絞り込みを行い、上記翻訳処理により翻訳された文章を上記文修正処理により修正して出力する処理を上記コンピュータに実行させる音声翻訳制御プログラムを記録してなることを特徴とする記録媒体。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、入力された音声を他の言語に翻訳する音声翻訳装置、音声翻訳方法及び音声翻訳制御プログラムを記録した記録媒体に関する。

【0002】

【従来の技術】

今日、音声認識技術の進歩に伴い、認識した音声を他の言語に翻訳する音声翻訳装置が開発されている。音声翻訳装置は、一般に、様々なユーザに対応することができるように、不特定話者対応の認識エンジンを使用している。

【0003】

このような音声翻訳装置100は、図1に示すように、例えばマイクからなる入力部102と、音声から特徴を抽出する特徴抽出部103と、音声認識照合部104と、音韻などの関係を表す音響モデルを記憶する音響モデル記憶部105と、単語と音響モデルの対応を表す単語辞書を記憶する単語辞書記憶部106と、単語の接続を表す文法（一般的に統計言語モデルが使用される。）を記憶する文法記憶部107とを備える。

【0004】

入力部102は、入力される音声を音声信号に変換し、さらにデジタル化して特徴抽出部103に供給する。特徴抽出部103は、音声信号から特徴ベクトル列を算出し、これを音声認識照合部104に供給する。

【0005】

音声認識照合部104は、上記特徴ベクトル、音響モデル記憶部105に記憶されている音響モデル、単語辞書記憶部106に記憶されている単語辞書、文法記憶部107に記憶されている文法を用いて、どの単語列が発声されたかを照合する。

【0006】

また、上記音声翻訳装置100は、音声認識照合部104からの照合結果が供給される翻訳処理部108と、変換情報記憶部109と、文修正部110と、誤りのある文章を修正するための修正ルールに関する情報が記憶されている修正情報記憶部111と、例文の置き換えに関する情報が記憶されている文法情報記憶部112と、例えばCRT(Cathode Ray Tube)やLCD(Liquid Crystal Display)からなる結果提示部113とを備える。

【0007】

翻訳処理部108は、上記照合結果について、多数の例文情報を記憶する変換情報記憶部109を用いて粗い翻訳を実行し、この翻訳結果を文修正部110に供給する。文修正部110は、修正情報記憶部111や文法情報記憶部112に記憶されている情報を用いて、翻訳処理部108の翻訳結果の細かい修正を実行する。

【0008】

そして、結果提示部113は、文修正部110の修正結果を表示することによって、入力音声の翻訳結果を表示している。

【0009】

このような音声翻訳装置100は、不特定話者対応の認識エンジンを使用しているため、ユーザを特定しないで済む利点がある一方で、特定話者の発話形態、使用語彙の範囲、発話の癖等が一定である事実を判別することが困難である。このため、上記音声翻訳装置100は、各個人用に合わせて音声認識処理を行っていないので、必要でないものまで処理しなければならない。

【0010】

また、上記音声翻訳装置100は、ユーザ適応機能がなく、利用者の利用頻度、傾向な

10

20

30

40

50

どの結果保存手段を備えていないので、発話者としてのユーザと翻訳機のやり取りを最適化することができない。このため、音声翻訳装置100は起こらなくていいエラーなどを引き起こす可能性がある。さらに、ユーザの個人的な好み、利用形態が特定できず、ユーザ個人に全く合致していない問題もある。

【0011】

【発明が解決しようとする課題】

本発明は、このような実情に鑑みて提案されたものであり、処理速度を向上させるとともに、ユーザと翻訳機との最適化を図ることができる音声翻訳装置、音声翻訳方法及び音声翻訳制御プログラムを記録した記録媒体を提供することを目的とする。

【0012】

【課題を解決するための手段】

本発明は、音声翻訳装置であって、入力された音声の特徴を抽出して特徴ベクトルを出力する特徴抽出手段と、上記特徴抽出手段からの特徴ベクトルに対して、音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声の音声認識を行い、認識した音声が表示する文章を出力する音声認識照合手段と、上記音声認識照合手段が認識した音声が表示する文章を、入力された音声の言語と異なる言語の文章に翻訳する翻訳処理手段と、上記翻訳処理手段により翻訳された言語の文章を、複数の修正情報から最適な修正情報を用いて修正する文修正手段と、上記文修正手段の修正結果を出力する出力手段と、上記音声認識照合手段が認識した文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者
による利用頻度により重み付けして示した履歴情報を蓄積する履歴情報蓄積手段と、上記音声認識照合手段が音声認識結果を出力する毎に、当該音声認識結果と上記履歴情報蓄積手段に蓄積されている履歴情報とを対比して、累積学習機能により上記履歴情報蓄積手段に蓄積されている履歴情報を更新する累積学習手段とを備え、上記履歴情報蓄積手段は、蓄積している履歴情報に基づいて、上記音声認識情報手段が用いる照合情報の絞り込み及び/又は上記文修正手段が用いる修正情報の絞り込みを制御し、上記累積学習手段は、上記履歴情報蓄積手段に蓄積される上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正手段が用いる修正情報の絞り込みを行うことを特徴とする。

【0013】

上記音声翻訳装置では、音声認識照合手段が認識した文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を履歴情報蓄積手段に蓄積し、音声認識照合手段が音声認識結果を出力する毎に、累積学習機能により当該音声認識結果と上記履歴情報蓄積手段に蓄積されている履歴情報とを対比し、上記履歴情報蓄積手段に蓄積されている履歴情報を更新する。そして、上記履歴情報蓄積手段は、蓄積している履歴情報に基づいて、上記音声認識情報手段が用いる照合情報の絞り込み及び/又は上記文修正手段が用いる修正情報の絞り込みを制御し、上記累積学習手段は、上記履歴情報蓄積手段に蓄積される上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正手段が用いる修正情報の絞り込みを行う。
すなわち、人工知能による累積学習により、音声翻訳時における照合情報の絞り込みや翻訳された文章を修正するときの修正情報の絞り込みを行い、しかも履歴情報蓄積手段に蓄積される意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正手段が用いる修正情報の絞り込みを行うので、音声翻訳の処理速度及び性能を向上させ、個々のユーザに対する最適化を行うことができる。

【0014】

また、本発明は、音声翻訳方法であって、入力された音声の特徴を抽出して特徴ベクトルを生成し、生成された特徴ベクトルに対して、音声認識用の複数の照合情報をそれぞれ照合して、上記入力された音声を認識する音声認識処理を行い、上記音声認識処理により

10

20

30

40

50

認識した音声¹⁰が示す文章を上記入力された音声の言語と異なる言語の文章に翻訳する翻訳処理を行い、上記翻訳処理により翻訳した文章を複数の修正情報から最適な修正情報を用いて修正する修正処理を行い、上記修正処理により修正した文章を出力するに当たり、上記音声認識処理により、入力された音声の特徴を抽出して生成された特徴ベクトルに対して、上記音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声を認識し、認識した音声¹⁰が示す文章を生成し、上記音声認識処理により認識した音声¹⁰が示す文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積し、上記音声認識処理を行う毎に、当該音声認識結果と音声認識済みの文章の傾向を示す上記履歴情報とを対比して、累積学習機能により上記履歴情報を更新し、更新された履歴情報に基づいて、上記音声認識処理に用いる照合情報の絞り込み及び/又は上記文修正処理に用いる修正情報の絞り込みを制御し、上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正処理に用いる修正情報の絞り込みを行い、上記翻訳処理により翻訳された文章を上記文修正処理により修正して出力することを特徴とする。

【0015】

上記音声翻訳方法では、音声認識処理により、入力された音声の特徴を抽出して生成された特徴ベクトルに対して、音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声を認識し、認識した音声²⁰が示す文章を生成し、上記音声認識処理により認識した音声²⁰が示す文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積し、上記音声認識処理を行う毎に、当該音声認識結果と音声認識済みの文章の傾向を示す上記履歴情報とを対比して、累積学習機能により上記履歴情報を更新し、更新された履歴情報に基づいて、上記音声認識処理に用いる照合情報の絞り込み及び/又は上記文修正処理に用いる修正情報の絞り込みを制御し、上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正処理に用いる修正情報の絞り込みを行い、上記翻訳処理により翻訳された文章を上記文修正処理により修正して出力する。²⁰
すなわち、人工知能による累積学習により、音声翻訳時における照合情報の絞り込みや翻訳された文章を修正するときの修正情報の絞り込みを行い、しかも履歴情報蓄積手段に蓄積される意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正手段が用いる修正情報の絞り込みを行うので、音声翻訳の処理速度及び性能を向上させ、個々のユーザに対する最適化を行うことができる。³⁰

【0016】

さらに、本発明は、音声翻訳装置に搭載されるコンピュータにより、読み取り実行可能に音声翻訳制御プログラムを記録した記録媒体であって、入力された音声の特徴を抽出して特徴ベクトルを生成し、生成された特徴ベクトルに対して、音声認識用の複数の照合情報をそれぞれ照合して、上記入力された音声を認識する音声認識処理を行い、上記音声認識処理により認識した音声⁴⁰が示す文章を上記入力された音声の言語と異なる言語の文章に翻訳する翻訳処理を行い、上記翻訳処理により翻訳した文章を複数の修正情報から最適な修正情報を用いて修正する修正処理を行い、上記修正処理により修正した文章を出力するに当たり、上記音声認識処理により、入力された音声の特徴を抽出して生成された特徴ベクトルに対して、上記音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声を認識し、認識した音声⁴⁰が示す文章を生成し、上記音声認識処理により認識した音声⁴⁰が示す文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積し、上記音声認識処理を行う毎に、当該音声認識結果と音声認識済みの文章の傾向を示す上記履歴情報とを対比して、累積学習機能により上記履歴情報を更新し、更新された履歴情報に基づいて、上記音声認識処理に用い⁴⁰
⁵⁰

る照合情報の絞り込み及び／又は上記文修正処理に用いる修正情報の絞り込みを制御し、上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正処理に用いる修正情報の絞り込みを行い、上記翻訳処理により翻訳された文章を上記文修正処理により修正して出力する処理を上記コンピュータに実行させる音声翻訳制御プログラムを記録してなることを特徴とする

【 0 0 1 7 】

上記音声翻訳制御プログラムを記録した記録媒体では、音声翻訳装置に搭載されるコンピュータにより上記音声翻訳制御プログラムを読み取り実行させることができる。

【 0 0 1 8 】

【発明の実施の形態】

以下、本発明を実施するための最良の形態について図面を参照しながら詳細に説明する。

【 0 0 1 9 】

本発明は、例えば図 2 に示す構成の音声翻訳装置 1 に適用することができる。

【 0 0 2 0 】

上記音声翻訳装置 1 は、音声が入力される入力部 2 と、入力された音声の特徴を抽出する特徴抽出部 3 と、音声認識照合部 4 と、HMM (Hidden Markov Model) モデルが記憶されている音響モデル記憶部 5 と、単語と音響モデルの対応を表す単語辞書を記憶する単語辞書記憶部 6 と、単語の接続を表す文法等を記憶する意味構文解析文法記憶部 7 と、認識した音声結果を累積学習する人工知能学習部 8 と、履歴情報を記憶する学習履歴部 9 とを備える。

【 0 0 2 1 】

入力部 2 は、例えばマイクからなり、入力された音声を音声信号に変換し、さらにこれをデジタル化して、特徴抽出部 3 に供給する。

【 0 0 2 2 】

特徴抽出部 3 は、上記音声信号の周波数分析を行って音声信号のパワースペクトルを得る。そして、音声区間ではない信号区間をノイズ区間としてこのノイズ区間のスペクトルを求めておいて、音声区間でこのノイズスペクトルを引く処理を行う。さらに、周波数分析結果のパワースペクトルをそのまま持っていたのではパラメータの次元が多いことから、適当な次元に圧縮処理して特徴量たる特徴ベクトルを時系列的に抽出する。

【 0 0 2 3 】

音声認識照合部 4 は、上記特徴ベクトルについて、音響モデル記憶部 5、単語辞書記憶部 6、意味構文解析文法記憶部 7 とを用いて、どの単語列が発声されたかを照合する。ここで、単語辞書記憶部 6 は、名詞、動詞、接続詞等の様々の単語に関する情報を記憶しているものである。意味構文解析文法記憶部 7 は、統計的 (ストカスティック: stochastic) 言語モデル 7 a と、ルールベース (シンボリック: symbolic) 文法 7 b とから構成される。これにより、例えば、Head-Driven Phrase Structure Grammar 等の汎用の言語解析理論を利用することができ、人工知能学習機能を用いて、容易に音声認識を行うことができる。

【 0 0 2 4 】

すなわち、音声認識照合部 4 は、特徴ベクトルと HMM モデルとを照合して 1 語 1 語を認識し、次に、単語辞書記憶部 6 との照合を行うことで各単語を認識し、そして、意味構文解析文法記憶部 7 との照合を行うことで構文を認識して、入力された音声を文章として認識し、この音声認識結果を出力する。このとき、音声認識照合部 4 は、音響モデル記憶部 5、単語辞書記憶部 6、意味構文解析文法記憶部 7 に記憶されている情報のうち複数の認識候補から絞り込みを行って、認識結果を人工知能学習部 8 に供給する。

人工知能学習部 8 は、文献 (S. Russell & Norvig (1995) Artificial Intelligence; A Modern Approach, Prentice Hall (日本語訳; 古川康一 (監訳) 「人工知能 エージェント・アプロ

10

20

30

40

50

ーチ」共立出版、1998年)にあるような累積学習(cumulative learning)を用いた学習機能を有する。人工知能学習部8は、学習エージェントが蓄積される履歴情報から学習された情報(又はそのウェイト)をその都度改善しながら、ユーザの好みの情報等を帰納論理的(inductive logical)に的確に推論する。

【0025】

文献P. Langley「Machine Learning for Adaptive User Interfaces」(Proceedings of the 21st German Annual Conference on Artificial Intelligence, pp. 53-62. Freiburg, Germany: Springer, 1997年)は、既存のシステム・ユーザ間の適応(adaptive user interface)が、ほとんど一般ユーザ全体に対してのマクロ的な適応処理であり、各個人の好み、知識、経験、目的等に対しての適応といった処理が、ほとんど存在しなかったことを指摘している。しかし、ここでいう人工知能部8とは、各ユーザがシステムを調整することなしに、システム自体がユーザからの反応・使用方法に応じてその個人の使用目的に最も適合したシステムを実現するために学習し、適応する機能を有する。

10

【0026】

ここにいうユーザの好みとは、例えばユーザたる発話者の構文的・意味的発話形態、使用する語彙の範囲、頻度、さらに発話の癖等である。人工知能学習部8は、これらのユーザの好みを、音響モデル、単語辞書、及び意味構文解析文法に分類して、統計的優先度の指標を付して、学習履歴部9に蓄積する。具体的には、人工知能学習部8は、学習履歴部9に蓄積されている履歴情報に何回同じアクセスがあったかをカウントし、学習履歴情報に重み付けをして、重み付けの小さい項目を除外していく。このように処理された知識は、ユーザの使用のたびに更新され、「使いやすさ」も同時に向上する。

20

【0027】

人工知能学習部8の学習機能は、文献Langley(1997年)のいうところの「Informative User Interface」及び「Generative User Interface」の両機能を有する。「Informative User Interface」とは、システム側からの反応においてユーザにとって興味のあるもの、有益なもののみを提供するように、システムが応答としての情報を選択又はフィルタする機能をいう。一方、「Generative User Interface」とは、ユーザにとって興味のある有益な情報をフィルタする機能とは違い、累積学習により新規にユーザの使用目的を満たすような知識構造(knowledge structure)を生成することを指す。これらの両機能があいまって発話者と翻訳機のやり取りを最適化する。

30

【0028】

具体的には、「Generative User Interface」の一例として、ここでは発話者の利用形態に順応する機能から、翻訳機の不特定話者対応の音声認識部の処理を特定話者対応に変換し、音声認識処理を軽減させる機能がそれに当たる。システムは累積学習の結果、不特定対応言語モデルを新規に生成するが、これは上記の知識構造(knowledge structure)である。また、「Informative User Interface」としては、ここでの音声翻訳機としてのユーザ適応機能(例えば、ユーザが知っている情報などは詳しく翻訳する必要がない等)は、システム側からの応答においてユーザにとって興味のあるもの、有益なものを提供するために、システムが応答としての情報を選択又はフィルタする機能の例である。

40

【0029】

学習履歴部9は、蓄積された学習履歴に基づいて、単語辞書記憶部6、意味構文解析文法記憶部7、後述する修正情報13に制約をかけて、発話者の使用する頻度の少ない構文的・意味的言語構造や語彙の範囲を絞り込む。

50

【 0 0 3 0 】

そして、音声認識照合部 4 は、絞り込まれた範囲内で優先度の高い単語辞書記憶部 6 や意味構文解析文法記憶部 7 の情報を用いて、特徴ベクトルとの照合を行って、音声認識処理を行うことができる。

【 0 0 3 1 】

これにより、各ユーザに合致した「ユーザモデル」を構築し、音声認識処理の対象ドメインを縮小して、性能の向上及び処理の高速化を図ることができる。

【 0 0 3 2 】

なお、学習履歴部 9 は、リセット機能を有し、リセットされたときは蓄積していた履歴情報を消去して新たに音声認識された結果を履歴情報として蓄積する。

10

【 0 0 3 3 】

また、上記音声翻訳装置 1 は、音声認識照合部 4 の音声認識結果を所定の言語に翻訳する翻訳処理部 1 0 と、例文情報が記憶されている変換情報記憶部 1 1 と、翻訳済みの文章に修正を加える文修正部 1 2 と、詳しい文章に修正するための情報が記憶されている修正情報記憶部 1 3 と、文法に関する情報が記憶されている文法情報記憶部 1 4 と、正確に翻訳された文章をユーザに提示する結果提示部 1 5 とを備える。

【 0 0 3 4 】

翻訳処理部 1 0 は、変換情報記憶部 1 1 に記憶されている多数の例文情報を用いて、上記音声認識結果について粗い翻訳を実行し、この翻訳結果を文修正部 1 2 に供給する。文修正部 1 2 は、修正情報記憶部 1 3 に記憶されている修正情報を用いて、粗く翻訳された文章を詳しく正確な文章に修正したり、文法情報記憶部 1 4 に記憶されている文法上の誤りを訂正するための情報を用いて、翻訳結果の細かな修正を行う。

20

【 0 0 3 5 】

ここで、修正情報記憶部 1 3 に記憶されている修正情報は、学習履歴部 9 にアクセスして蓄積された意味構文解析文法記憶部 7 の情報に作用される。すなわち、学習履歴部 9 は、意味構文解析文法記憶部 7 に関する履歴情報のうち、繰り返し使用される意味構文については詳しく翻訳する必要がないものと判定し、それほど詳しく翻訳しないように修正情報記憶部 1 3 を制御することで最適化を図っている。

【 0 0 3 6 】

そして、文修正部 1 2 は、何度も繰り返される意味構文については、あまり詳しくしないで文章を修正する。これにより、発話者が同じ内容の文章を繰り返しはなした場合には、繰り返しの文章部分についてはあまり詳しく翻訳しないので、翻訳された文章がくどくなることを回避することができる。

30

【 0 0 3 7 】

結果提示部 1 5 は、例えば CRT (Cathode Ray Tube) や LCD (Liquid Crystal Display) からなり、文修正部 1 2 からの翻訳結果を文章等により提示する。これにより、ユーザは、翻訳された文章を視覚により理解することができる。

【 0 0 3 8 】

以上のように、上記音声翻訳装置 1 は、ユーザの好み等の個人情報で対象ドメインを縮小し、性能の向上化、処理の高速化を図り、発話者と翻訳機とのやり取りを最適化することができる。

40

【 0 0 3 9 】

また、人工知能学習部 8 を備えることによって、ユーザの特有の構文的・意味的発話形態、使用する語彙の範囲や頻度を学習することができるので、より高速に音声認識処理を行うことができる。

【 0 0 4 0 】

また、累積学習により学習履歴部 9 に蓄積された履歴情報は、リセット機能によりクリアすることができる。したがって、使用するユーザが変わっても、当該ユーザに対応して音声翻訳を行うことができる。すなわち、不特定話者対応型と特定話者対応型の切替えが

50

可能であり、それぞれの利点を考慮しながら、状況に合わせて使用することができる。

【0041】

なお、本実施の形態では、結果提示部15はCRTやLCDからなるディスプレイとして説明したが、本発明はこれに限定されるものではなく、結果提示部15は例えば翻訳された言語を音声で出力するスピーカであってもよいし、翻訳された結果のデータを外部に送信する送信回路等であってもよい。

【0042】

さらに、上述した処理を行う音声翻訳制御プログラムを記録した記録媒体から当該制御プログラムをコンピュータ装置にインストールすることによって、上述した処理を実行することもできる。

【0043】

【発明の効果】

以上詳細に説明したように、本発明に係る音声翻訳装置、音声翻訳方法及び音声翻訳制御プログラムを記録した記録媒体によれば、音声認識処理により、入力された音声の特徴を抽出して生成された特徴ベクトルに対して、音声認識用の複数の照合情報として、音響モデル、単語情報、意味構文解析文法情報をそれぞれ照合して、上記入力された音声を認識し、認識した音声を示す文章を生成し、上記音声認識処理により認識した音声を示す文章の傾向を上記音響モデル、単語情報及び意味構文解析文法情報に分類して、入力された音声の発話者による利用頻度により重み付けして示した履歴情報を蓄積し、上記音声認識処理を行う毎に、当該音声認識結果と音声認識済みの文章の傾向を示す上記履歴情報とを対比して、累積学習機能により上記履歴情報を更新し、更新された履歴情報に基づいて、上記音声認識処理に用いる照合情報の絞り込み及び/又は上記文修正処理に用いる修正情報の絞り込みを制御し、上記意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正処理に用いる修正情報の絞り込みを行い、上記翻訳処理により翻訳された文章を上記文修正処理により修正して出力する。すなわち、人工知能による累積学習により、音声翻訳時における照合情報の絞り込みや翻訳された文章を修正するときの修正情報の絞り込みを行い、しかも履歴情報蓄積手段に蓄積される意味構文解析文法情報に関する履歴情報のうち、使用頻度の高い意味構文解析文法情報については詳しく修正しないように、上記文修正手段が用いる修正情報の絞り込みを行い、音声翻訳の処理速度及び性能を向上させるとともに、個々のユーザに対する最適化を行うことができる。

【図面の簡単な説明】

【図1】 図1は、従来の音声翻訳装置の構成を示すブロック図である。

【図2】 図2は、本発明を適用した音声翻訳装置の構成を示すブロック図である。

10

20

30

フロントページの続き

(56)参考文献 特開平 1 0 - 1 3 6 3 2 7 (J P , A)

(58)調査した分野(Int.Cl. , D B 名)

G10L 15/00-15/28