

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6723907号
(P6723907)

(45) 発行日 令和2年7月15日(2020.7.15)

(24) 登録日 令和2年6月26日(2020.6.26)

(51) Int.Cl.

F I

G 1 O L 15/10 (2006.01)

G 1 O L 15/10 5 O O Z

G 1 O L 15/06 (2013.01)

G 1 O L 15/06 5 O O Z

G 1 O L 15/06 4 O O V

請求項の数 15 (全 20 頁)

(21) 出願番号 特願2016-232115 (P2016-232115)
 (22) 出願日 平成28年11月30日(2016.11.30)
 (65) 公開番号 特開2018-87945 (P2018-87945A)
 (43) 公開日 平成30年6月7日(2018.6.7)
 審査請求日 平成31年2月26日(2019.2.26)

(73) 特許権者 000005108
 株式会社日立製作所
 東京都千代田区丸の内一丁目6番6号
 (74) 代理人 110000176
 一色国際特許業務法人
 (72) 発明者 藤田 雄介
 東京都千代田区丸の内一丁目6番6号 株
 式会社日立製作所内

審査官 上田 雄

最終頁に続く

(54) 【発明の名称】 言語認識システム、言語認識方法、及び言語認識プログラム

(57) 【特許請求の範囲】

【請求項1】

プロセッサ及びメモリを備え、発せられた音声の言語を認識する言語認識システムであって、

音声を取得する音声取得部と、

第1の言語における第1の語彙に関する情報である第1言語一般語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語一般翻訳句を生成する第1言語一般翻訳部と、

前記第1の言語における第2の語彙に関する情報である第1言語専門語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語専門翻訳句を生成する第1言語専門翻訳部と、

前記生成した第1言語一般翻訳句、及び前記生成した第1言語専門翻訳句の間で相互に対応する部分である第1言語対応部分を抽出する第1対応区間算出部と、

第2の言語における第3の語彙に関する情報である第2言語一般語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語一般翻訳句を生成する第2言語一般翻訳部と、

前記第2の言語における第4の語彙に関する情報である第2言語専門語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語専門翻訳句を生成する第2言語専門翻訳部と、

前記生成した第2言語一般翻訳句、及び前記生成した第2言語専門翻訳句の間で相互に

10

20

対応する部分である第 2 言語対応部分を抽出する第 2 対応区間算出部と、

前記抽出した前記第 1 言語対応部分、及び前記抽出した前記第 2 言語対応部分に基づき、前記取得した音声の言語を特定する言語識別部と、
を備える言語認識システム。

【請求項 2】

前記第 1 対応区間算出部は、前記第 1 言語対応部分を、前記第 1 言語一般翻訳句を表す文字列と前記第 1 言語専門翻訳句を表す文字列との共通性に基づき抽出し、

前記第 2 対応区間算出部は、前記第 2 言語対応部分を、前記第 2 言語一般翻訳句を表す文字列と前記第 2 言語専門翻訳句を表す文字列との共通性に基づき抽出する、

請求項 1 に記載の言語認識システム。

10

【請求項 3】

前記第 1 対応区間算出部は、前記第 1 言語一般翻訳句における各単語の発声のタイミング、及び前記第 1 言語専門翻訳句における各単語の発声のタイミングを取得し、取得したこれらの単語の発声のタイミングの共通性に基づき、前記第 1 言語対応部分を抽出し、

前記第 2 対応区間算出部は、前記第 2 言語一般翻訳句における各単語の発声のタイミング、及び前記第 2 言語専門翻訳句における各単語の発声のタイミングを取得し、取得したこれらの単語の発声のタイミングの共通性に基づき、前記第 2 言語対応部分を抽出する、

請求項 1 に記載の言語認識システム。

【請求項 4】

前記言語識別部は、前記第 1 言語対応部分の、前記第 1 の言語としての確からしさを示す指標である第 1 信頼度と、前記第 2 言語対応部分の、前記第 2 の言語としての確からしさを示す指標である第 2 信頼度とに基づき、前記取得した音声の言語を特定する、

請求項 1 に記載の言語認識システム。

20

【請求項 5】

前記取得した音声に対応する、前記特定した言語による音声を入力する応答部を備える、請求項 1 に記載の言語認識システム。

【請求項 6】

前記第 1 言語一般語彙情報、前記第 1 言語専門語彙情報、前記第 2 言語一般語彙情報、又は前記第 2 言語専門語彙情報の少なくともいずれかの入力を受け付け、入力を受け付けた情報を出力するデータベース更新部を備える、請求項 1 に記載の言語認識システム。

30

【請求項 7】

前記第 1 の言語の言葉を取得し、取得した前記第 1 の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第 1 の言語の言葉の一部を前記第 1 言語専門語彙情報として記憶し、

前記第 2 の言語の言葉を取得し、取得した前記第 2 の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第 2 の言語の言葉の一部を前記第 2 言語専門語彙情報として記憶する、

専門言語更新部を備える、請求項 1 に記載の言語認識システム。

【請求項 8】

前記抽出した前記第 1 の言語の言葉の一部を前記第 2 の言語の言葉に変換し、変換した前記第 2 の言語の言葉を前記第 2 言語専門語彙情報として記憶し、

前記抽出した前記第 2 の言語の言葉の一部を前記第 1 の言語の言葉に変換し、変換した前記第 1 の言語の言葉を前記第 1 言語専門語彙情報として記憶する、

専門言語変換部を備える、請求項 7 に記載の言語認識システム。

40

【請求項 9】

前記取得した音声に対応する、前記特定した言語による音声を入力する応答部と、

前記第 1 言語一般語彙情報、前記第 1 言語専門語彙情報、前記第 2 言語一般語彙情報、又は前記第 2 言語専門語彙情報の少なくともいずれかの入力を受け付け、入力を受け付けた情報を出力するデータベース更新部と、

前記第 1 の言語の言葉を取得し、取得した前記第 1 の言語の言葉の一部を、言葉の使用

50

頻度に関する情報に基づき抽出し、抽出した前記第 1 の言語の言葉の一部を前記第 1 言語専門語彙情報として記憶し、

前記第 2 の言語の言葉を取得し、取得した前記第 2 の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第 2 の言語の言葉の一部を前記第 2 言語専門語彙情報として記憶する専門言語更新部と、

前記抽出した前記第 1 の言語の言葉の一部を前記第 2 の言語の言葉に変換し、変換した前記第 2 の言語の言葉を前記第 2 言語専門語彙情報として記憶し、

前記抽出した前記第 2 の言語の言葉の一部を前記第 1 の言語の言葉に変換し、変換した前記第 1 の言語の言葉を前記第 1 言語専門語彙情報として記憶する専門言語変換部とを備え、

10

前記第 1 対応区間算出部は、前記第 1 言語対応部分を、前記第 1 言語一般翻訳句を表す文字列と前記第 1 言語専門翻訳句を表す文字列との共通性に基づき抽出し、及び / 又は、前記第 1 言語一般翻訳句における各単語の発声のタイミング、及び前記第 1 言語専門翻訳句における各単語の発声のタイミングを取得し、取得したこれらの単語の発声のタイミングの共通性に基づき、前記第 1 言語対応部分を抽出し、

前記第 2 対応区間算出部は、前記第 2 言語対応部分を、前記第 2 言語一般翻訳句を表す文字列と前記第 2 言語専門翻訳句を表す文字列との共通性に基づき抽出し、及び / 又は、前記第 2 言語一般翻訳句における各単語の発声のタイミング、及び前記第 2 言語専門翻訳句における各単語の発声のタイミングを取得し、取得したこれらの単語の発声のタイミングの共通性に基づき、前記第 2 言語対応部分を抽出し、

20

前記言語識別部は、前記第 1 言語対応部分の、前記第 1 の言語としての確からしさを示す指標である第 1 信頼度と、前記第 2 言語対応部分の、前記第 2 の言語としての確からしさを示す指標である第 2 信頼度とに基づき、前記取得した音声の言語を特定する、

請求項 1 に記載の言語認識システム。

【請求項 10】

発せられた音声の言語を認識する言語認識方法であって、

プロセッサ及びメモリを備える情報処理装置が、

音声を取得する音声取得処理と、

第 1 の言語における第 1 の語彙に関する情報である第 1 言語一般語彙情報に基づき、前記取得した音声に対応する前記第 1 の言語の言葉である第 1 言語一般翻訳句を生成する第 1 言語一般翻訳処理と、

30

前記第 1 の言語における第 2 の語彙に関する情報である第 1 言語専門語彙情報に基づき、前記取得した音声に対応する前記第 1 の言語の言葉である第 1 言語専門翻訳句を生成する第 1 言語専門翻訳処理と、

前記生成した第 1 言語一般翻訳句、及び前記生成した第 1 言語専門翻訳句の間で相互に対応する部分である第 1 言語対応部分を抽出する第 1 対応区間算出処理と、

第 2 の言語における第 3 の語彙に関する情報である第 2 言語一般語彙情報に基づき、前記取得した音声に対応する前記第 2 の言語の言葉である第 2 言語一般翻訳句を生成する第 2 言語一般翻訳処理と、

前記第 2 の言語における第 4 の語彙に関する情報である第 2 言語専門語彙情報に基づき、前記取得した音声に対応する前記第 2 の言語の言葉である第 2 言語専門翻訳句を生成する第 2 言語専門翻訳処理と、

40

前記生成した第 2 言語一般翻訳句、及び前記生成した第 2 言語専門翻訳句の間で相互に対応する部分である第 2 言語対応部分を抽出する第 2 対応区間算出処理と、

前記抽出した前記第 1 言語対応部分、及び前記抽出した前記第 2 言語対応部分に基づき、前記取得した音声の言語を特定する言語識別処理と、

を実行する、言語認識方法。

【請求項 11】

前記第 1 対応区間算出処理は、前記第 1 言語対応部分を、前記第 1 言語一般翻訳句を表す文字列と前記第 1 言語専門翻訳句を表す文字列との共通性に基づき抽出し、

50

前記第2対応区間算出処理は、前記第2言語対応部分を、前記第2言語一般翻訳句を表す文字列と前記第2言語専門翻訳句を表す文字列との共通性に基づき抽出する処理を含む、

請求項10に記載の言語認識方法。

【請求項12】

前記第1の言語の言葉を取得し、取得した前記第1の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第1の言語の言葉の一部を前記第1言語専門語彙情報として記憶し、

前記第2の言語の言葉を取得し、取得した前記第2の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第2の言語の言葉の一部を前記第2言語専門語彙情報として記憶する、

専門言語更新処理を実行する、請求項10に記載の言語認識方法。

【請求項13】

発せられた音声の言語を認識する言語認識プログラムであって、

プロセッサ及びメモリを備える情報処理装置に、

音声を取得する音声取得処理と、

第1の言語における第1の語彙に関する情報である第1言語一般語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語一般翻訳句を生成する第1言語一般翻訳処理と、

前記第1の言語における第2の語彙に関する情報である第1言語専門語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語専門翻訳句を生成する第1言語専門翻訳処理と、

前記生成した第1言語一般翻訳句、及び前記生成した第1言語専門翻訳句の間で相互に対応する部分である第1言語対応部分を抽出する第1対応区間算出処理と、

第2の言語における第3の語彙に関する情報である第2言語一般語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語一般翻訳句を生成する第2言語一般翻訳処理と、

前記第2の言語における第4の語彙に関する情報である第2言語専門語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語専門翻訳句を生成する第2言語専門翻訳処理と、

前記生成した第2言語一般翻訳句、及び前記生成した第2言語専門翻訳句の間で相互に対応する部分である第2言語対応部分を抽出する第2対応区間算出処理と、

前記抽出した前記第1言語対応部分、及び前記抽出した前記第2言語対応部分に基づき、前記取得した音声の言語を特定する言語識別処理と、

を実行させる、言語認識プログラム。

【請求項14】

前記第1対応区間算出処理は、前記第1言語対応部分を、前記第1言語一般翻訳句を表す文字列と前記第1言語専門翻訳句を表す文字列との共通性に基づき抽出し、

前記第2対応区間算出処理は、前記第2言語対応部分を、前記第2言語一般翻訳句を表す文字列と前記第2言語専門翻訳句を表す文字列との共通性に基づき抽出する処理を含む、

請求項13に記載の言語認識プログラム。

【請求項15】

前記第1の言語の言葉を取得し、取得した前記第1の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第1の言語の言葉の一部を前記第1言語専門語彙情報として記憶し、

前記第2の言語の言葉を取得し、取得した前記第2の言語の言葉の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第2の言語の言葉の一部を前記第2言語専門語彙情報として記憶する、

専門言語更新処理を実行させる、請求項13に記載の言語認識プログラム。

10

20

30

40

50

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、言語認識システム、言語認識方法、及び言語認識プログラムに関する。

【背景技術】

【0002】

人が発した言葉の言語（例えば、日本語、英語、中国語、韓国語等）を識別する技術に関する種々の開発が現在行われている。例えば、特許文献1には、利用者に対して所定の呼びかけ動作を行い、この呼びかけ動作に対して返されるであろう限定されたパターンの言葉を言語識別における認識候補とする言語識別用音声認識エンジンが記載されている。

10

【先行技術文献】

【特許文献】

【0003】

【特許文献1】特開2004-53825号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかし、特許文献1では、銀行のATMなどの特定の業務における言語識別においては有効であるものの、言語を認識するためになされる発話が限定されており、応用範囲が狭い。

20

【0005】

そこで、各言語の音声認識装置に入力された発話から出力された認識結果に基づき、各言語についての信頼度を算出し、算出した信頼度のそれぞれを比較することにより言語を識別することが考えられる。ただ、このような方法によっても、発話が短い場合には精度良く識別することが難しい。

【0006】

特に、特定の業務において特殊な言葉が発せされた場合、例えば、いずれの言語の音声認識装置でも認識が困難な言葉が発言者の発言に含まれていた場合には、算出される各言語の信頼度がいずれも低いものとなり、その結果、正しい言語認識ができなくなるおそれがある。例えば、観光業務において、外国からの観光客が発した言葉にその観光客の母国以外の国の地名が含まれている場合、その観光客が発している言語の識別は非常に難しいものとなる。このように、発する言葉の種類によっては、言語識別が難しくなるという問題があった。

30

【0007】

本発明はこのような背景に基づきなされたものであり、言語認識を正確に行うための言語認識システム、言語認識方法、及び言語認識プログラムを提供することにある。

【課題を解決するための手段】

【0008】

上記課題を解決するための本発明の一つは、プロセッサ及びメモリを備え、発せられた音声の言語を認識する言語認識システムであって、音声を取得する音声取得部と、第1の言語における第1の語彙に関する情報である第1言語一般語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語一般翻訳句を生成する第1言語一般翻訳部と、前記第1の言語における第2の語彙に関する情報である第1言語専門語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語専門翻訳句を生成する第1言語専門翻訳部と、前記生成した第1言語一般翻訳句、及び前記生成した第1言語専門翻訳句の間で相互に対応する部分である第1言語対応部分を抽出する第1対応区間算出部と、第2の言語における第3の語彙に関する情報である第2言語一般語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語一般翻訳句を生成する第2言語一般翻訳部と、前記第2の言語における第4の語彙に関する情報である第2言語専門語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言

40

50

葉である第2言語専門翻訳句を生成する第2言語専門翻訳部と、前記生成した第2言語一般翻訳句、及び前記生成した第2言語専門翻訳句の間で相互に対応する部分である第2言語対応部分を抽出する第2対応区間算出部と、前記抽出した前記第1言語対応部分、及び前記抽出した前記第2言語対応部分に基づき、前記取得した音声の言語を特定する言語識別部と、を備える。

【発明の効果】

【0009】

本発明によれば、言語認識を正確に行うことができる。

【0010】

上記した以外の課題、構成及び効果は、以下の実施形態の説明により明らかにされる。

10

【図面の簡単な説明】

【0011】

【図1】図1は、本実施形態に係る言語認識システム1の構成の一例を示す図である。

【図2】図2は、サービスロボット200のハードウェア構成の一例を説明する図である。

。

【図3】図3は、言語識別サーバ100が備えるハードウェア構成の一例を説明する図である。

【図4】図4は、サービスロボット200が備える機能の一例を説明する図である。

【図5】図5は、言語識別サーバ100が備える機能の一例を説明する図である。

【図6】図6は、言語識別処理の一例を説明するフローチャートである。

20

【図7】図7は、言語識別処理の一例を説明するフローチャートである。

【図8】図8は、言語識別処理の具体例を説明する図である。

【図9】図9は、データベース更新処理の一例を説明するフローチャートである。

【図10】図10は、データベース更新画面の一例を示す図である。

【図11】図11は、サービスロボット200による外国人観光客への観光案内の一例を説明する図である。

【発明を実施するための形態】

【0012】

本発明を実施するための形態につき、以下図面を用いて説明する。

【0013】

30

図1は、本実施形態に係る言語認識システム1の構成の一例を示す図である。同図に示すように、言語認識システム1は、言語の識別を行う情報処理装置である言語識別サーバ100と、サービスロボット200と、無線アクセスポイント300とを含んで構成されている。言語認識システム1は、主に国外からの観光客（以下、単に観光客という）に対応する業務（以下、本件業務という）において、その観光客が発した言葉の言語を識別する言語認識システムである。

【0014】

なお、観光客が発する言葉の言語には、日本語、英語、中国語、韓国語などの様々な言語（以下、対象言語という）がありうるが、本実施形態の言語認識システム1は、日本語（以下、第1の言語という）、及び英語（以下、第2の言語という）を識別するものとする。

40

【0015】

サービスロボット200は、観光客が発した音声を認識し、認識した音声を言語識別サーバ100に送信する、自律的な移動が可能な情報処理装置である。また、サービスロボット200は、言語識別サーバ100から送信されてきた各種の情報に基づき音声を生成し、生成した音声を発することで、観光客との対話を行う。

【0016】

無線アクセスポイント300は、サービスロボット200と言語識別サーバ100との間の通信を制御する情報処理装置である。無線アクセスポイント300は、サービスロボット200と無線通信網5等を介して通信可能に接続し、また、言語識別サーバ100と

50

有線 L A N 4 等の有線の通信網、又は無線の通信網を介して通信可能に接続している。

【 0 0 1 7 】

(ハードウェア構成)

次に、言語認識システム 1 を構成する各情報処理装置のハードウェア構成について説明する。

【 0 0 1 8 】

図 2 は、サービスロボット 2 0 0 のハードウェア構成の一例を説明する図である。同図に示すように、サービスロボット 2 0 0 は、C P U (Central Processing Unit) 等からなるプロセッサ 2 1 と、R A M (Random Access Memory)、R O M (Read Only Memory)、又は N V R A M (Non-Volatile RAM) 等の主記憶装置や、ハードディスク (Hard Disk Drive) や S S D (Solid State Drive) 等の補助記憶装置からなる記憶装置 2 2 と、集音装置 2 3 と、スピーカ 2 4 と、カメラ 2 5 と、移動装置 2 6 と、無線アクセスポイント 3 0 0 を介して言語識別サーバ 1 0 0 と通信する通信装置 2 7 とを備える。

10

【 0 0 1 9 】

記憶装置 2 2 は、サービスロボット 2 0 0 が行う対話の内容に関する情報 (シナリオデータ) や、シナリオデータに基づき行う対話の順序や応答方法を制御するプログラム (シナリオ動作プログラム) を記憶している。

【 0 0 2 0 】

集音装置 2 3 は、観光客の音声をデジタルデータに変換する。集音装置 2 3 は、例えば、音声を電気信号に変換するマイクロフォンと、マイクロフォンから出力された電気信号をデジタルデータに変換する A D コンバータとを含んで構成される。本実施形態では、集音装置 2 3 は指向性の異なる複数のマイクロフォンを備え、これらのマイクロフォンにより複数の方向からの音声が入力されるものとする。

20

【 0 0 2 1 】

スピーカ 2 4 は、音声を出力する。スピーカ 2 4 は、例えば、言語識別サーバ 1 0 0 から受信した、音声に関する情報 (デジタルデータ) をアナログ電気信号に変換する D A コンバータと、変換されたアナログ電気信号を増幅するアンプ (amplifier) と、増幅されたアナログ電気信号を物理振動に変換して音声を生成するスピーカーとを含んで構成される。

【 0 0 2 2 】

カメラ 2 5 は、所定の範囲の被写体を撮影する撮影装置であり、例えば、サービスロボット 2 0 0 に相対する人の位置や動きなどを認識する。

30

【 0 0 2 3 】

移動装置 2 6 は、サービスロボット 2 0 0 を走行させるための装置であり、例えば、モータ等の動力機構や、車輪等の操舵機構から構成される。

【 0 0 2 4 】

図 3 は、言語識別サーバ 1 0 0 が備えるハードウェア構成の一例を説明する図である。同図に示すように、言語識別サーバ 1 0 0 は、C P U (Central Processing Unit) 等のプロセッサ 1 1 と、R A M (Random Access Memory)、R O M (Read Only Memory)、又は N V R A M (Non-Volatile RAM) 等からなる主記憶装置 1 2、ハードディスクや S S D (Solid State Drive) 等からなる補助記憶装置 1 3 と、タッチパネルや操作ボタン等の入力装置 1 4 と、液晶ディスプレイ、プリンタ等の出力装置 1 5 と、無線アクセスポイント 3 0 0 を介してサービスロボット 2 0 0 と通信するための通信装置 1 6 とを備える。

40

【 0 0 2 5 】

(機能)

次に、各情報処理装置が備える機能について説明する。

【 0 0 2 6 】

図 4 は、サービスロボット 2 0 0 が備える機能の一例を説明する図である。同図に示すように、サービスロボット 2 0 0 は、音声取得部 2 1 0、音声送信部 2 2 0、情報受信部 2 3 0、及び応答部 2 4 0 を備える。

50

【0027】

音声取得部210は、音声を取得する。具体的には、音声取得部210は、集音装置23に入力された音から雑音等を除去することにより、人の音声のみを抽出する。

【0028】

音声送信部220は、音声取得部210が抽出した音声の情報を言語識別サーバ100に送信する。

【0029】

情報受信部230は、言語識別サーバ100から送信されてくる、言語識別サーバ100が特定した言語（詳細は後述）に関する情報等を受信する。

【0030】

応答部240は、音声取得部210が前記取得した音声に対応する音声を、言語識別サーバ100が前記特定した言語に基づき出力する。

【0031】

以上のサービスロボット200の機能は、サービスロボット200のハードウェアによって、もしくは、サービスロボット200のプロセッサ21が、記憶装置22に記憶されているプログラムを読み出して実行することにより実現される。

【0032】

図5は、言語識別サーバ100が備える機能の一例を説明する図である。同図に示すように、言語識別サーバ100は、音声取得部110、第1言語処理部120、第2言語処理部130、言語識別部140、応答生成部150、及びデータベース更新部160を備える。

【0033】

音声取得部110は、サービスロボット200から音声の情報を取得する。

【0034】

第1言語処理部120は、第1言語音声認識部121、第1対応区間算出部122、第1信頼度算出部123を備える。また、第1言語処理部120は、第1言語一般語彙情報125、第1言語専門語彙情報126、及び第1言語音節情報127の各データベースを管理している。

【0035】

第1言語音声認識部121は、第1言語一般翻訳部1211、及び第1言語専門翻訳部1212を備える。

【0036】

第1言語一般翻訳部1211は、第1の言語における第1の語彙に関する情報である第1言語一般語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語一般翻訳句を生成する。なお、第1の語彙は、本実施形態では、第1言語で一般に発せられる、本件業務において使用頻度が高い（本件業務で日常的に発せられる、又は本件業務で聞くことが多い）語彙であるものとする。

【0037】

第1言語専門翻訳部1212は、前記第1の言語における第2の語彙に関する情報である第1言語専門語彙情報に基づき、前記取得した音声に対応する前記第1の言語の言葉である第1言語専門翻訳句を生成する。なお、第2の語彙は、本実施形態では、本件業務において使用頻度が低い（本件業務で発せられることが少ない、又は本件業務では聞き慣れない）語彙であるものとする。例えば、第2の語彙は、日本に来訪する外国人が発する日本語の固有名詞（地名等）など、一般に発音が人によって不明確となるような語彙を含む。

【0038】

第1対応区間算出部122は、前記生成した第1言語一般翻訳句、及び前記生成した第1言語専門翻訳句の間で相互に対応する部分である第1言語対応部分を抽出する。

【0039】

例えば、前記第1対応区間算出部は、前記第1言語対応部分を、前記第1言語一般翻訳句を表す文字列と前記第1言語専門翻訳句を表す文字列との共通性に基づき抽出する。ま

10

20

30

40

50

た、例えば、前記第1対応区間算出部は、前記第1言語一般翻訳句における各単語の発声のタイミング、及び前記第1言語専門翻訳句における各単語の発声のタイミングを取得し、取得したこれらの単語の発声のタイミングの共通性に基づき、前記第1言語対応部分を抽出する。

【0040】

第1言語一般語彙情報125は、第1の言語における第1の語彙を記憶したデータベース（言語モデル）である。具体的には、第1言語一般語彙情報125は、第1の言語の第1の語彙における単語、定型句、文等の情報を記憶している。

【0041】

第1言語専門語彙情報126は、前記第1の言語における第2の語彙を記憶したデータベース（言語モデル）である。具体的には、第1言語専門語彙情報126は、第1の言語の第2の語彙における単語、定型句、文等の情報を記憶している。

10

【0042】

第1言語音節情報127は、第1の言語において使用される音声の情報（音響モデル）を記憶しているデータベースである。第1言語音節情報127には、例えば、第1の言語における音節（発声のタイミング等を含む）の情報が含まれる。

【0043】

第1信頼度算出部123は、後述する第1信頼度を算出する。

【0044】

第2言語処理部130は、第2言語音声認識部131、第2対応区間算出部132、及び第2信頼度算出部133を備える。また、第2言語処理部130は、第2言語一般語彙情報135、第2言語専門語彙情報136、及び第2言語音節情報137を管理している。

20

【0045】

第2言語音声認識部131は、第2言語一般翻訳部1311、及び第2言語専門翻訳部1312を備える。

【0046】

第2言語一般翻訳部1311は、第2の言語における第3の語彙に関する情報である第2言語一般語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語一般翻訳句を生成する。なお、第3の語彙は、本実施形態では、第2言語で一般に発せられる、本件業務において使用頻度が高い（本件業務で日常的に発せられる、又は本件業務で聞くことが多い）語彙であるとする。

30

【0047】

第2言語専門翻訳部1312は、前記第2の言語における第4の語彙に関する情報である第2言語専門語彙情報に基づき、前記取得した音声に対応する前記第2の言語の言葉である第2言語専門翻訳句を生成する。なお、第4の語彙は、本実施形態では、本件業務において使用頻度が低い（本件業務で発声されることが少ない、又は本件業務では聞き慣れない）語彙であるとする。例えば、第4の語彙は、日本に來訪する外国人が発する日本語の固有名詞（地名等）など、一般に発音が人によって不明確となるような語彙を含む。

【0048】

40

第2対応区間算出部132は、前記生成した第2言語一般翻訳句、及び前記生成した第2言語専門翻訳句の間で相互に対応する部分である第2言語対応部分を抽出する。

【0049】

例えば、前記第2対応区間算出部132は、前記第2言語対応部分を、前記第2言語一般翻訳句を表す文字列と前記第2言語専門翻訳句を表す文字列との共通性に基づき抽出する。また、例えば、前記第2対応区間算出部132は、前記第2言語一般翻訳句における各単語の発声のタイミング、及び前記第2言語専門翻訳句における各単語の発声のタイミングを取得し、取得したこれらの単語の発声のタイミングの共通性に基づき、前記第2言語対応部分を抽出する。

【0050】

50

第2言語一般語彙情報135は、第2の言語における第3の語彙を記憶したデータベース（言語モデル）である。具体的には、第2言語一般語彙情報135は、第2の言語の第3の語彙における単語、定型句、文等の情報を、文字列として記憶している。

【0051】

第2言語専門語彙情報136は、前記第2の言語における第4の語彙を記憶したデータベース（言語モデル）である。具体的には、第2言語専門語彙情報136は、第2の言語の第4の語彙における単語、定型句、文等の情報を、文字列として記憶している。

【0052】

第2言語音節情報137は、第2の言語において使用される音声の情報（音響モデル）を記憶しているデータベースである。第2言語音節情報137には、例えば、第2の言語における音節（発声のタイミング等を含む）の情報が含まれる。

10

【0053】

第2信頼度算出部133は、後述する第2信頼度を算出する。

【0054】

言語識別部140は、前記抽出した前記第1言語対応部分、及び前記抽出した前記第2言語対応部分に基づき、前記入力された音声の言語を特定する。

【0055】

具体的には、前記言語識別部140は、前記第1言語対応部分の、前記第1の言語としての確からしさを示す指標である第1信頼度と、前記第2言語対応部分の、前記第2の言語としての確からしさを示す指標である第2信頼度とに基づき、前記取得した音声の言語を特定する。

20

【0056】

なお、信頼度（第1信頼度、第2信頼度）は各単語ごとに算出され、各信頼度は所定の範囲の値（又はスコア。例えば、0.0～1.0）を有する。この信頼度の値が下限に近い（例えば、0.0に近い）ほど、当該算出対象となった単語と同じ程度のスコアを有する他の単語（単語仮説）が多く存在する。

【0057】

また、信頼度は、例えば、認識した音声における各単語の単語仮説についての事後確率を算出することによって求められる。この算出方法の詳細は、例えば、「Confidence Measures for Large Vocabulary Continuous Speech Recognition, IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 9, NO. 3, MARCH 2001」に開示されている。

30

【0058】

応答生成部150は、前記取得した音声に対応する、前記特定した言語による音声を出力する。

【0059】

データベース更新部160は、前記第1言語一般語彙情報、前記第1言語専門語彙情報、前記第2言語一般語彙情報、又は前記第2言語専門語彙情報の入力を受け付け、入力を受け付けた情報を出力する。

【0060】

また、データベース更新部160は、専門言語更新部161、及び専門言語変換部162を備える。

40

【0061】

専門言語更新部161は、前記第1の言語の言葉を取得し、取得した前記第1の言語の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第1の言語の一部を前記第1言語専門語彙情報として記憶し、前記第2の言語の言葉を取得し、取得した前記第2の言語の一部を、言葉の使用頻度に関する情報に基づき抽出し、抽出した前記第2の言語の一部を前記第2言語専門語彙情報として記憶する。

【0062】

専門言語変換部162は、前記抽出した前記第1の言語の言葉の一部を前記第2の言語の言葉に変換し、変換した前記第2の言語の言葉を前記第2言語専門語彙情報として記憶

50

し、前記抽出した前記第2の言語の言葉の一部を前記第1の言語の言葉に変換し、変換した前記第1の言語の言葉を前記第1言語専門語彙情報として記憶する。

【0063】

以上に説明した言語識別サーバ100の機能は、言語識別サーバ100のハードウェアによって、もしくは、言語識別サーバ100のプロセッサ11が、主記憶装置12や補助記憶装置13に記憶されているプログラムを読み出して実行することにより実現される。

【0064】

また、以上に説明した言語識別サーバ100の機能の一部は、サービスロボット200に設けてもよい。また、サービスロボット200の機能の一部も、言語識別サーバ100に設けてもよい。

10

【0065】

(処理)

次に、言語認識システム1において行われる処理について説明する。

【0066】

<言語識別処理>

図6、7は、言語認識システム1において行われる処理のうち、観光客が発した言葉の言語を識別する処理(以下、言語識別処理という)の一例を説明するフローチャートである(紙面の都合上、2つの図に分けている)。言語識別処理は、例えば、言語識別サーバ100及びサービスロボット200に電源が投入された場合に開始される。

【0067】

20

同図に示すように、まずサービスロボット200は、観光客からの音声の受信を待機する(S1)。

【0068】

音声を受信されると、サービスロボット200は、受信した音声から、観光客の音声以外の音(物音や雑音等)を除去することにより、観光客の音声のみを取得する(S3)。具体的には、サービスロボット200は、取得した複数のチャンネルの音声の位相差に基づき、雑音を除去する。雑音等が除去された音声は、例えば、サービスロボット200が正面方向から取得した音声(1チャンネルの音声)のみとなる。

【0069】

サービスロボット200は、S3で取得した音声を、言語識別サーバ100に送信する(S5)。

30

【0070】

言語識別サーバ100は、サービスロボット200から音声を受信すると(S7)、第1言語一般語彙情報125に基づき、S7で受信した音声に対応する、第1の言語の言葉の文字列(第1言語一般翻訳句)を生成する(S9)。

【0071】

また、言語識別サーバ100は、第1言語専門語彙情報126に基づき、S7で受信した音声に対応する、第1の言語の言葉の文字列(第1言語専門翻訳句)を生成する(S11)。

【0072】

40

そして、言語識別サーバ100は、S9で生成した第1言語一般翻訳句と、S11で生成した第1言語専門翻訳句との間で相互に対応する部分(第1言語対応部分)を抽出する(S13)。具体的には、言語識別サーバ100は、第1言語対応部分として、第1言語一般翻訳句に対応する文字列と第1言語専門翻訳句に対応する文字列との共通する部分を抽出する。

【0073】

そして、言語識別サーバ100は、S13で抽出した第1言語抽出分の第1信頼度を算出する(S15)。

【0074】

また、言語識別サーバ100は、第2言語一般語彙情報に基づき、S7で受信した音声

50

に対応する、第2の言語の言葉の文字列（第2言語一般翻訳句）を生成する（S17）。また、言語識別サーバ100は、第2言語専門語彙情報に基づき、S7で受信した音声に対応する、第2の言語の言葉の文字列（第2言語専門翻訳句）を生成する（S19）。

【0075】

また、言語識別サーバ100は、S17で生成した第2の言語の文字列と、S19で生成した第2の言語の文字列との間で相互に対応する部分（第2言語対応部分）を抽出する（S21）。具体的には、第2言語対応部分として、第2言語一般翻訳句に対応する文字列と第2言語専門翻訳句に対応する文字列との共通部分を抽出する。

【0076】

そして、言語識別サーバ100は、S21で抽出した第2言語対応部分の第2信頼度を算出する（S23）。

【0077】

言語識別サーバ100は、S15で算出した第1信頼度と、S23で算出した第2信頼度とに基づき、S7で受信した音声の言語を特定する（S25）。具体的には、言語識別サーバ100は、第1信頼度及び第2信頼度を比較し、信頼度が高かった方の言語を、S7で受信した音声の言語として記憶する。

【0078】

続いて、言語識別サーバ100は、S7で受信した音声の意味内容を解析する（S27）。

【0079】

具体的には、言語識別サーバ100は、S7で受信した音声に対応する、S25で特定した言語の文字列を取得する。例えば、S25で特定した言語が第1の言語である場合は、言語識別サーバ100は、S9又はS11で生成した文字列を取得し、S25で特定した言語が第2の言語である場合は、言語識別サーバ100は、S17又はS19で生成した文字列を取得する。そして、言語識別サーバ100は、取得した文字列の意味内容を解析する。

【0080】

言語識別サーバ100は、S27で解析した内容を、サービスロボット200に送信する（S29）

【0081】

サービスロボット200は、言語識別サーバ100から受信した情報に基づき、音声を生成する（S31、S33）。具体的には、例えば、サービスロボット200は、受信した情報、及びシナリオ動作プログラムに従って、シナリオデータが示す対話の情報を音声に変換する。

【0082】

そして、サービスロボット200は、S33で生成した音声を出力する（S35）。以上で言語識別処理は終了する（S37）。

【0083】

< 言語識別処理の具体例 >

ここで、言語識別処理の具体例について説明する。

図8は、言語識別処理の具体例を説明する図である。同図に示すように、まず、言語識別サーバ100が、"How to get to KASAI RINKAI KOEN?"という音声の情報（以下、本音声という）をサービスロボット200から受信したとする（符号1001）。

【0084】

すると、言語識別サーバ100は、第1言語専門語彙情報126に基づき、本音声の言語が第1の言語（日本語）であると仮定して、本音声を、「初月葛西臨海公園」という文字列に変換する（符号1002）。なお、言語識別サーバ100は、この文字列を、「初月」なる文字列と、「葛西臨海公園」なる文字列との組み合わせとして認識する。

【0085】

また、言語識別サーバ100は、第1言語一般語彙情報125に基づき、本音声の言語

10

20

30

40

50

が第1の言語（日本語）であると仮定して、本音声を、「初月火災臨海公園」という文字列に変換する（符号1003）。なお、言語識別サーバ100は、この文字列を、「初月」なる文字列と、「火災臨海公園」なる文字列の組み合わせとして認識する。

【0086】

言語識別サーバ100は、符号1002に示した文字列と、符号1003に示した文字列とに基づき、第1言語対応部分を抽出する（符号1004）。すなわち、言語識別サーバ100は、符号1002に示した文字列と、符号1003に示した文字列との間の、連続する共通の文字列である「初月」を、第1言語対応部分として抽出する。

【0087】

なお、共通の文字列が複数存在する場合は、共通の文字列のうち、その文字列の長さが最大である共通の文字列が抽出される。後記する第2言語対応部分についても同様である。

【0088】

次に、言語識別サーバ100は、第2言語専門語彙情報136に基づき、本音声の言語が第2の言語（英語）であると仮定して、本音声を、「How to get to KASAI RINKAI KOEN?」なる文字列に変換する（符号1005）。なお、言語識別サーバ100は、この文字列を、「How to get to」なる文字列と、「KASAI RINKAI KOEN」なる文字列との組み合わせとして認識する。

【0089】

また、言語識別サーバ100は、第2言語一般語彙情報135に基づき、本音声の言語が第2の言語（英語）であると仮定して、本音声を、「How to get to as I bring eye hole?」という文字列に変換する（符号1006）。なお、言語識別サーバ100は、この文字列を、「How to get to」なる文字列と、「as I bring eye hole」なる文字列の組み合わせとして認識する。

【0090】

言語識別サーバ100は、符号1005に示した文字列と、符号1006に示した文字列とに基づき、第2言語対応部分を抽出する（符号1007）。すなわち、言語識別サーバ100は、符号1005に示した文字列と、符号1006に示した文字列との間の、連続する共通の文字列である「How to get to」を、第2言語対応部分として抽出する。

【0091】

なお、言語識別サーバ100は、上記のように文字列に基づき第2言語対応部分を抽出する代わりに、又はそれに加えて、第1言語一般翻訳句における各言葉の発声のタイミング、及び、第1言語専門翻訳句における各言葉の発声のタイミングに基づき（音響モデルに基づき）第2言語対応部分を抽出してもよい。すなわち、言語識別サーバ100は、符号1005の文字列の発声と、符号1006の文字列の発声とでは、「How to get to」の部分において各単語が発声されるタイミングや無音のタイミングが共通するので、「How to get to」を第2言語対応部分として抽出する。なお、このような方法は、第1言語対応部分についても同様に採用してもよい。

【0092】

そして、言語識別サーバ100は、符号1004に示した第1言語対応部分の信頼度（第1信頼度）と、符号1007に示した第2言語対応部分の信頼度（第2信頼度）とを算出する（符号1008）。

【0093】

例えば、言語識別サーバ100は、第1信頼度の方が第2信頼度より高いと判断した場合、第1信頼度を算出した第1言語対応部分の言語である第1の言語（日本語）が、符号1001に示した音声に対応する言語であると認識する。

【0094】

<データベース更新処理>

次に、言語認識システム1において行われる各データベースの更新処理について説明する。

10

20

30

40

50

【 0 0 9 5 】

図 9 は、データベースを更新する処理（以下、データベース更新処理という）の一例を説明するフローチャートである。データベース更新処理は、例えば、言語識別サーバ 1 0 0 に所定の操作入力が行われた際に開始される。

【 0 0 9 6 】

同図に示すように、まず言語識別サーバ 1 0 0 は、本件業務で使用される、第 1 の言語の言葉（例えば、文章、単語。以下、業務データという。）の入力をユーザから受け付ける（S 4 1）。なお、業務データは、言語識別サーバ 1 0 0 が取得した音声（サービスロボット 2 0 0 が受信した音声）に基づき生成された第 1 の言語であってもよい。

【 0 0 9 7 】

業務データの入力が受け付けられると、言語識別サーバ 1 0 0 は、業務データから、第 1 の言語の第 2 の語彙（専門用語や、外国人が発する固有名詞等）を抽出し、抽出した第 2 の語彙を第 1 言語専門語彙情報 1 2 6 に記憶する（S 4 2）。なお、第 2 の語彙の抽出は、例えば、言語識別サーバ 1 0 0 が記憶している、第 1 の言語における各単語の使用頻度に関する情報に基づき抽出される。例えば、言語識別サーバ 1 0 0 は、業務データが示す単語の使用頻度が所定値以下、又は所定の範囲にある場合に、その業務データを第 1 言語専門語彙情報 1 2 6 に記憶する。

【 0 0 9 8 】

また、言語識別サーバ 1 0 0 は、S 4 2 で抽出した第 2 の語彙を、第 4 の語彙に変換し、変換した第 4 の語彙を、第 2 言語専門語彙情報 1 3 6 に記憶する（S 4 3）。以上で処理は終了する（S 4 4）。

【 0 0 9 9 】

なお、以上のデータベース更新処理は、第 1 の言語だけでなく第 2 の言語についても同様に行われる。すなわち、言語識別サーバ 1 0 0 は、業務データから、第 2 の言語の第 4 の語彙（専門用語や、外国人が発する固有名詞等）を抽出し、抽出した第 4 の語彙を第 2 言語専門語彙情報 1 3 6 に記憶する。また、言語識別サーバ 1 0 0 は、抽出した第 4 の語彙を第 2 の語彙に変換し、変換した第 2 の語彙を、第 1 言語専門語彙情報 1 2 6 に記憶する。

【 0 1 0 0 】

< データベース更新画面 >

ここで、データベース更新処理において表示される画面について説明する。

【 0 1 0 1 】

図 1 0 は、データベース更新処理において表示される、各データベースを編集するための画面（以下、データベース更新画面という）の一例を示す図である。同図に示すように、データベース更新画面は、一般言語モデル編集画面 5 0 0、及び専門言語モデル編集画面 6 0 0 を備える。

【 0 1 0 2 】

一般言語モデル編集画面 5 0 0 は、第 1 の語彙、又は第 3 の語彙の入力を受け付ける。一般言語モデル編集画面 5 0 0 は、第 1 言語入力欄 5 0 1、第 2 言語入力欄 5 0 2、及び登録確定欄 5 0 3 を備える。第 1 言語入力欄 5 0 1 には、第 1 言語の文字列（第 1 の語彙）の入力が受け付けられる。第 2 言語入力欄 5 0 2 には、第 2 言語の文字列（第 3 の語彙）の入力が受け付けられる。登録確定欄 5 0 3 には、第 1 言語入力欄 5 0 1 に入力された文字列を第 1 言語一般語彙情報 1 2 5 に記憶させ、第 2 言語入力欄 5 0 2 に入力された文字列を第 2 言語一般語彙情報 1 3 5 に記憶させる旨の入力が受け付けられる。

【 0 1 0 3 】

専門言語モデル編集画面 6 0 0 は、第 2 の語彙、又は第 4 の語彙の入力を受け付ける。専門言語モデル編集画面 6 0 0 は、第 1 言語入力欄 6 0 1、第 2 言語入力欄 6 0 2、及び登録確定欄 6 0 3 を備える。第 1 言語入力欄 6 0 1 には、第 1 言語の文字列（第 2 の語彙）の入力が受け付けられる。第 2 言語入力欄 6 0 2 には、第 2 言語の文字列の入力（第 4 の語彙）が受け付けられる。登録確定欄 6 0 3 には、第 1 言語入力欄 6 0 1 に入力された

10

20

30

40

50

文字列を第1言語専門語彙情報126に記憶させ、第2言語入力欄602に入力された文字列を第2言語専門語彙情報136に記憶させる旨の入力が受け付けられる。

【0104】

<サービスロボットによる観光案内の例>

ここで、サービスロボット200による観光案内の例を説明する。

図11は、サービスロボット200による外国人観光客への観光案内の一例を説明する図である。同図に示すように、14チャンネルのマイクを備えるサービスロボット200に、外国人観光客が発した、複数の種類の言語（日本語、英語、中国語、韓国語）のうちある一つの言語（以下、入力言語という）による言葉が入力される（符号1101）。この言葉の音声と同時に入力された雑音等の他の不要な音声は、所定の雑音抑制処理により除去され、1チャンネルの音声（以下、入力音声という）となる。

10

【0105】

サービスロボット200から入力音声を受信した言語識別サーバ100は、言語識別処理を行う。

【0106】

具体的には、言語識別サーバ100は、入力音声の、日本語としての信頼度（第1信頼度）を、第1言語一般語彙情報125、第1言語専門語彙情報126、及び第1言語音節情報127に基づき算出する（符号1102）。また、言語識別サーバ100は、入力音声の、英語としての信頼度（第2信頼度）を、第2言語一般語彙情報135、第2言語専門語彙情報136、及び第2言語音節情報137に基づき算出する（符号1103）。なお、不図示であるが、言語識別サーバ100は、その他の言語（中国語、韓国語等）についても同様に、各言語の信頼度を算出する。

20

【0107】

そして言語識別サーバ100は、算出した各信頼度を比較し、最も信頼度が高い言語（本例では英語とする）が、入力音声の入力言語であると特定する（符号1104）。さらに、言語識別サーバ100は、特定した言語の意味内容（入力音声の意味内容）を解析することにより、入力音声の、所定の行き先への行き方の質問であることを認識し、認識した結果をサービスロボット200に送信する（符号1105）。

【0108】

サービスロボット200は、受信した情報に基づき、上記質問に対応する発声、すなわち、所定の行き先への行き方を説明する文章の発声を行う（符号1106）。

30

【0109】

以上のように、本実施形態の言語認識システム1によれば、第1言語一般語彙情報に基づき生成した第1言語一般翻訳句と、第1言語専門語彙情報に基づき生成した第1言語専門翻訳句との間で相互に対応する第1言語対応部分を抽出し、第2言語一般語彙情報に基づき生成した第2言語一般翻訳句と、第2言語専門語彙情報に基づき生成した第2言語専門翻訳句との間で相互に対応する第2言語対応部分を抽出し、第1言語対応部分、及び第2言語対応部分に基づき音声の言語を特定する。すなわち、第1の言語及び第2の言語について、異なる語彙の情報に基づき抽出した言葉の間の対応部分に基づいて言語を認識するので、言語認識を行うのに適切な部分のみを取り出して言語を識別することができる。

40

【0110】

例えば、専門語彙の言語モデルによる音声の認識結果と、一般語彙の言語モデルによる音声の認識結果とが対応する部分に基づき言語を識別することができる。したがって、発言者が発する言葉に外来語が含まれていることによって音声の認識結果に誤りが生じるおそれがある場合であっても、発言者が発した言語を正確に認識することができる。

【0111】

このように、本実施形態の言語認識システム1によれば、言語認識を正確に行うことができる。

【0112】

また、本実施形態の言語認識システム1によれば、第1言語対応部分を、第1言語一般

50

翻訳句を表す文字列と第1言語専門翻訳句を表す文字列との共通性に基づき抽出し、第2言語対応部分を、第2言語一般翻訳句を表す文字列と第2言語専門翻訳句を表す文字列との共通性に基づき抽出するので、第1言語対応部分及び第2言語対応部分において意味内容が共通する部分を確実に抽出することができる。これにより、言語認識を正確に行うことができる。

【0113】

また、本実施形態の言語認識システム1によれば、第1言語一般翻訳句における各単語の発声のタイミング、及び第1言語専門翻訳句における各単語の発声のタイミングの共通性に基づき第1言語対応部分を抽出し、第2言語一般翻訳句における各単語の発声のタイミング、及び第2言語専門翻訳句における各単語の発声のタイミングの共通性に基づき第2言語対応部分を抽出するので、各言語に特有に存在する発声のタイミングやリズムに基づき言語認識を行うことができる。これにより、言語認識を正確に行うことができる。

10

【0114】

また、本実施形態の言語認識システム1によれば、第1言語対応部分の、第1の言語としての確からしさを示す第1信頼度と、第2言語対応部分の、第2の言語としての確からしさを示す第2信頼度とに基づき、音声の言語を特定するので、言語認識を安定した精度で行うことができる。

【0115】

また、本実施形態の言語認識システム1によれば、第1言語一般語彙情報、第1言語専門語彙情報、第2言語一般語彙情報、又は第2言語専門語彙情報の入力を受け付け、入力を受け付けた情報を出力するので、言語認識システム1のユーザは、自身による入力内容を確認しながら各情報を適宜に更新し、言語認識の精度を向上させることができる。

20

【0116】

また、本実施形態の言語認識システム1によれば、発せられた音声に対応する、上記特定した言語の音声出力するので、人との的確な対話が可能となる。例えば、ロボットによる外国人観光客との自動音声対話を的確に行うことができる。

【0117】

また、本実施形態の言語認識システム1によれば、取得した第1及び第2の言語の一部を言葉の使用頻度に関する情報に基づき抽出し、抽出した第1及び第2の言語の一部を第1及び第2言語専門語彙情報として記憶するので、第1及び第2言語専門語彙情報を適切な内容に更新することができる。これにより、言語認識の精度を向上させることができる。

30

【0118】

また、本実施形態の言語認識システム1によれば、抽出した第1及び第2の言語の言葉の一部を第2及び第1の言語の言葉に変換し、変換した第2及び第1の言語の言葉を第2及び第1言語専門語彙情報として記憶するので、第1の言語及び第2の言語の語彙を同時に蓄積することができる。これにより、言語認識の精度を向上させることができる。

【0119】

以上の実施形態の説明は、本発明の理解を容易にするためのものであり、本発明を限定するものではない。本発明はその趣旨を逸脱することなく、変更、改良され得ると共に本発明にはその等価物が含まれる。

40

【0120】

例えば、本実施形態では、言語認識システム1が、観光客の発した言葉を認識するシステムである場合を説明したが、発言者は観光客に限られず、人一般でよい。例えば、サービスロボット200が看護用ロボットや介護用ロボットである等、発言者が病院施設や介護施設における患者や医師であるという場合でもよい。

【0121】

また、言語識別サーバ100は、第1言語対応部分及び第2言語対応部分を、予め記憶しておいた、第1言語対応部分及び第2言語対応部分に関する語彙の情報から抽出してもよい。

50

【 0 1 2 2 】

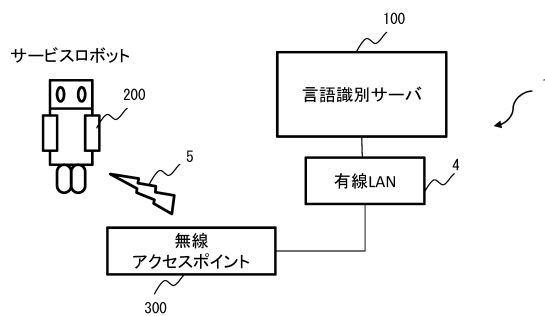
また、本実施形態では、対象言語として、第 1 の言語（日本語）、及び第 2 の言語（英語）のみを例として挙げたが、言語の種類や数はこれに限られず、中国語、韓国語等の他の言語も、言語認識の対象の言語とすることが可能である。

【符号の説明】

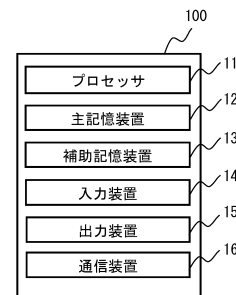
【 0 1 2 3 】

1 言語認識システム、100 言語識別サーバ、200 サービスロボット、210 音声取得部、1211 第 1 言語一般翻訳部、1212 第 1 言語専門翻訳部、122 第 1 対応区間算出部、1311 第 2 言語一般翻訳部、1312 第 2 言語専門翻訳部、132 第 2 対応区間算出部、133 第 2 信頼度算出部、140 言語識別部、150 応答生成部

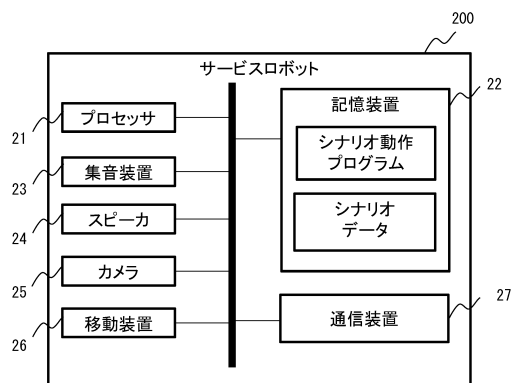
【図 1】



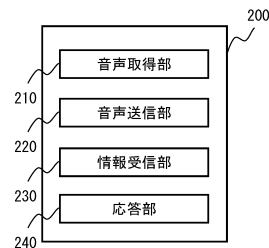
【図 3】



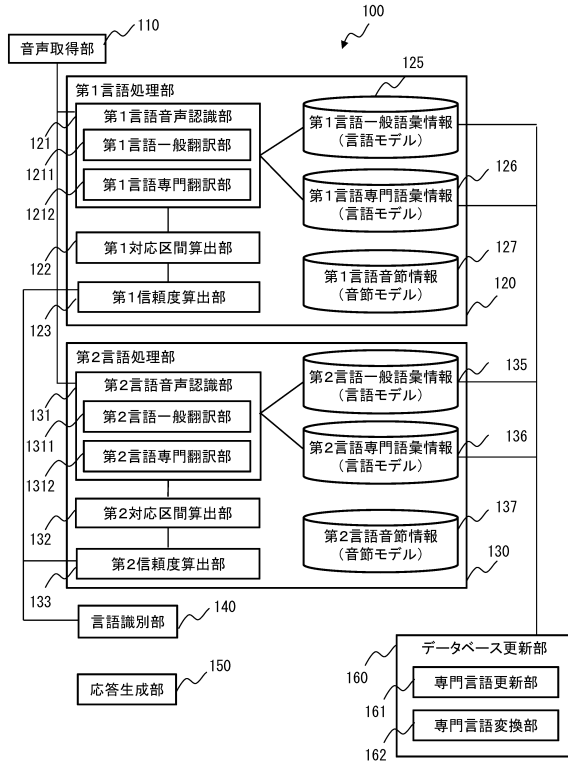
【図 2】



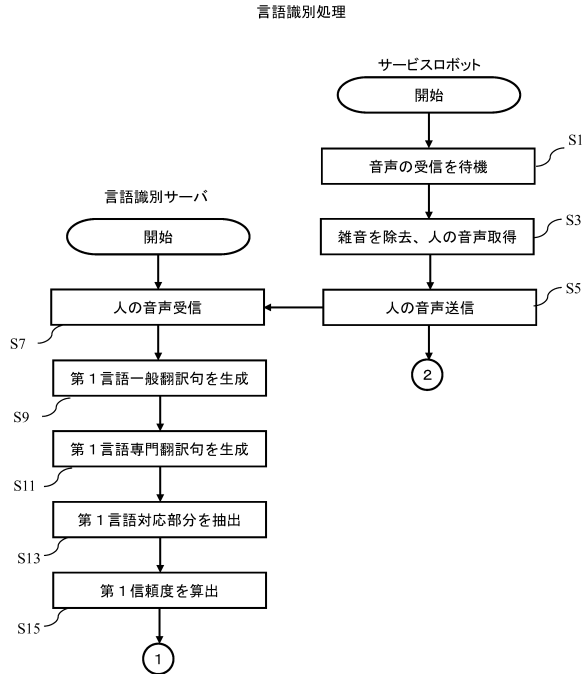
【図 4】



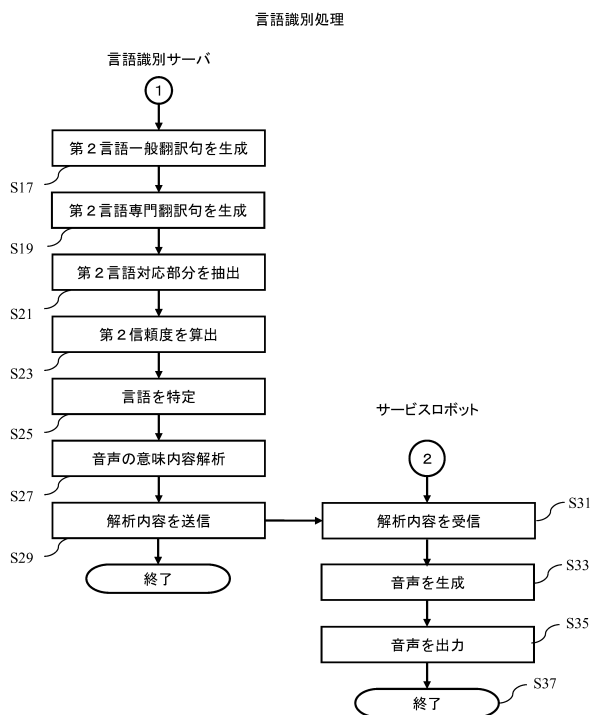
【図 5】



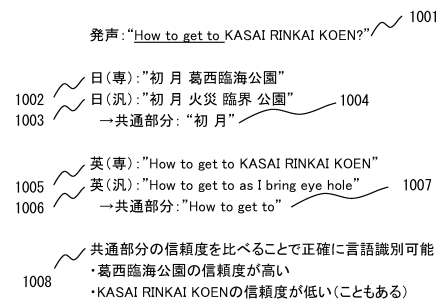
【図 6】



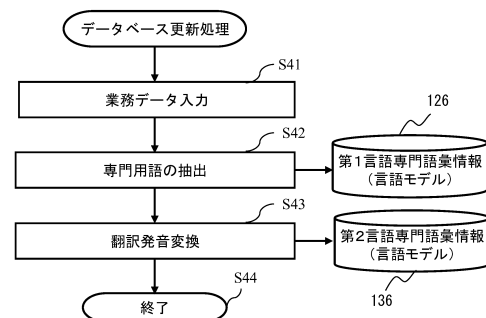
【図 7】



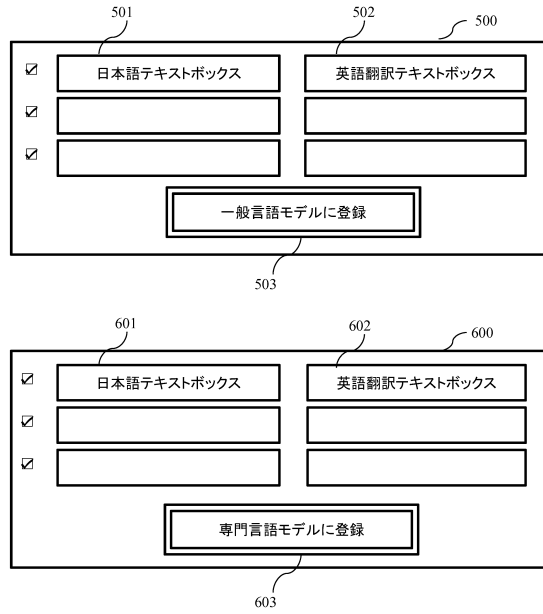
【図 8】



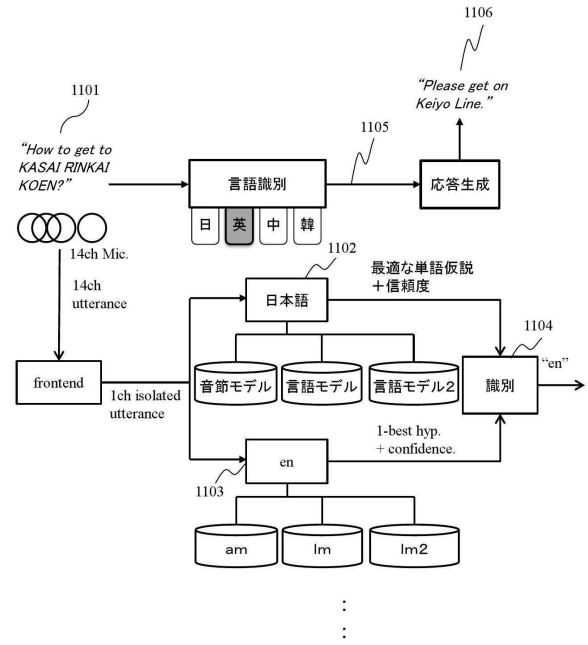
【図 9】



【図 10】



【図 11】



フロントページの続き

(56)参考文献 特開2004-347732(JP,A)
特開2003-316383(JP,A)
特開平07-230294(JP,A)
米国特許出願公開第2013/0238336(US,A1)

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