

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第6766991号

(P6766991)

(45) 発行日 令和2年10月14日(2020.10.14)

(24) 登録日 令和2年9月23日(2020.9.23)

(51) Int.Cl.

F I

G 1 0 L 15/00 (2013.01)

G 1 0 L 15/00 2 0 0 C

G 1 0 L 15/18 (2013.01)

G 1 0 L 15/18 3 0 0 H

G 1 0 L 15/22 (2006.01)

G 1 0 L 15/22 4 6 0 Z

G 0 6 F 3/16 (2006.01)

G 0 6 F 3/16 6 5 0

G 0 6 F 40/58 (2020.01)

G 0 6 F 40/58

請求項の数 6 (全 22 頁)

(21) 出願番号 特願2016-256486 (P2016-256486)

(22) 出願日 平成28年12月28日(2016.12.28)

(65) 公開番号 特開2018-18052 (P2018-18052A)

(43) 公開日 平成30年2月1日(2018.2.1)

審査請求日 令和1年8月27日(2019.8.27)

(31) 優先権主張番号 特願2016-138919 (P2016-138919)

(32) 優先日 平成28年7月13日(2016.7.13)

(33) 優先権主張国・地域又は機関

日本国(JP)

(73) 特許権者 591128763

株式会社富士通ソーシャルサイエンスラ
ボトリ神奈川県川崎市中原区小杉町一丁目403
番地

(74) 代理人 100113608

弁理士 平川 明

(74) 代理人 100105407

弁理士 高田 大輔

(72) 発明者 福岡 寿和

神奈川県川崎市中原区小杉町1丁目403
番地 株式会社富士通ソーシャルサイエ
ンスラボトリ内

審査官 岩田 淳

最終頁に続く

(54) 【発明の名称】 端末装置、翻訳方法、及び、翻訳プログラム

(57) 【特許請求の範囲】

【請求項1】

音声の入力を受け付ける音声入力部と、

所定時間毎に、前記音声入力部が受け付けた音声の音声データを前記音声入力部が以前に受け付けた音声の音声データと結合した結合音声データの音声で、第1言語の文字情報に変換することを、音声認識手段に指令する音声認識指令部と、

前記音声認識手段によって変換された前記第1言語の第1文字情報を受信する毎に、他の端末装置に向けて、前記第1言語の第1文字情報を送信する文字送信部と、

他の端末装置から第2言語の第2文字情報を受信する毎に、前記第2言語の第2文字情報を前記第1言語に翻訳することを、翻訳手段に指令する翻訳指令部と、

前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、最後に表示した前記第1言語の第2文字情報に改行を示す制御文字が含まれない場合、前記最後に表示した前記第1言語の第2文字情報を消去してから表示し、前記最後に表示した前記第1言語の第2文字情報に改行を示す制御文字が含まれる場合、改行してから表示する表示部とを備える端末装置。

【請求項2】

コンピュータが、

音声の入力を受け付け、

所定時間毎に、受け付けた音声の音声データを以前に受け付けた音声の音声データと結合した結合音声データの音声で、第1言語の文字情報に変換することを、音声認識手段に

指令し、

前記音声認識手段によって変換された前記第1言語の第1文字情報を受信し、他の端末装置に向けて、前記第1言語の第1文字情報を送信し、

他の端末装置から第2言語の第2文字情報を受信する毎に、前記第2言語の第2文字情報を前記第1言語に翻訳することを、翻訳手段に指令し、

前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、最後に表示した前記第1言語の第2文字情報に改行を示す制御文字が含まれない場合、前記最後に表示した前記第1言語の第2文字情報を消去してから表示し、前記最後に表示した前記第1言語の第2文字情報に改行を示す制御文字が含まれる場合、改行してから表示する

ことを実行する翻訳方法。

10

【請求項3】

コンピュータが、

音声の入力を受け付け、

所定時間毎に、受け付けた音声の音声データを以前に受け付けた音声の音声データと結合した結合音声データの音声を、第1言語の文字情報に変換することを、音声認識手段に指令し、

前記音声認識手段によって変換された前記第1言語の第1文字情報を受信する毎に、他の端末装置に向けて、前記第1言語の第1文字情報を送信し、

他の端末装置から第2言語の第2文字情報を受信する毎に、前記第2言語の第2文字情報を前記第1言語に翻訳することを、翻訳手段に指令し、

20

前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、最後に表示した前記第1言語の第2文字情報に改行を示す制御文字が含まれない場合、前記最後に表示した前記第1言語の第2文字情報を消去してから表示し、前記最後に表示した前記第1言語の第2文字情報に改行を示す制御文字が含まれる場合、改行してから表示する

ことを実行するための翻訳プログラム。

【請求項4】

前記表示部は、前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、前記第2言語の第2文字情報とともに表示する、

請求項1に記載の端末装置。

【請求項5】

30

前記コンピュータが、

前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、前記第2言語の第2文字情報とともに表示する、

ことを実行する請求項2に記載の翻訳方法。

【請求項6】

前記コンピュータが、

前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、前記第2言語の第2文字情報とともに表示する、

ことを実行するための請求項3に記載の翻訳プログラム。

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、端末装置、翻訳方法、及び、翻訳プログラムに関する。

【背景技術】

【0002】

利用者が発声するある言語の言葉を入力して、他の言語の言葉に翻訳する翻訳装置がある。

【先行技術文献】

【特許文献】

【0003】

50

【特許文献1】特開2011-182125号公報

【特許文献2】特開2014-149644号公報

【発明の概要】

【発明が解決しようとする課題】

【0004】

しかしながら、従来の翻訳装置では、利用者が言葉を発声してから翻訳を開始するまでのタイムラグが大きいなどのため、翻訳装置を通じて、他の利用者と会話をするには支障があった。

【0005】

本発明は、異なる言語を使用する利用者間の会話をよりスムーズにする技術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】

【0006】

上記課題を解決するために、以下の手段を採用する。

即ち、第1の態様は、

音声の入力を受け付ける音声入力部と、

所定時間毎に、前記音声入力部が受け付けた音声の音声データを前記音声入力部が以前に受け付けた音声の音声データと結合した結合音声データの音声を、第1言語の文字情報に変換することを、音声認識手段に指令する音声認識指令部と、

前記音声認識手段によって変換された前記第1言語の第1文字情報を受信する毎に、他の端末装置に向けて、前記第1言語の第1文字情報を送信する文字送信部と、

他の端末装置から第2言語の第2文字情報を受信する毎に、前記第2言語の第2文字情報を前記第1言語に翻訳することを、翻訳手段に指令する翻訳指令部と、

前記翻訳手段によって翻訳された前記第1言語の第2文字情報を、前記第2言語の第2文字情報とともに表示する表示部とを備える端末装置とする。

【0007】

開示の態様は、プログラムが情報処理装置によって実行されることによって実現されてもよい。即ち、開示の構成は、上記した態様における各手段が実行する処理を、情報処理装置に対して実行させるためのプログラム、或いは当該プログラムを記録したコンピュータ読み取り可能な記録媒体として特定することができる。また、開示の構成は、上記した各手段が実行する処理を情報処理装置が実行する方法をもって特定されてもよい。開示の構成は、上記した各手段が実行する処理を行う情報処理装置を含むシステムとして特定されてもよい。

【発明の効果】

【0008】

本発明によれば、異なる言語を使用する利用者間の会話をよりスムーズにする技術を提供することを目的とする。

【図面の簡単な説明】

【0009】

【図1】図1は、本実施形態の翻訳システムの構成例を示す図である。

【図2】図2は、端末装置の機能ブロックの例を示す図である。

【図3】図3は、情報処理装置のハードウェア構成例を示す図である。

【図4】図4は、端末装置における音声入力の動作フローの例を示す図である。

【図5】図5は、端末装置における翻訳文表示の動作フローの例を示す図である。

【図6】図6は、実施形態の翻訳システムの動作の具体例を示す図である。

【図7】図7は、変形例2の端末装置における音声入力の動作フローの例を示す図である。

。

【図8】図8は、変形例3の端末装置の機能ブロックの例を示す図である。

【図9】図9は、変形例3の端末装置における音声入力の動作フローの例を示す図である

10

20

30

40

50

。

【発明を実施するための形態】

【0010】

以下、図面を参照して実施形態について説明する。実施形態の構成は例示であり、発明の構成は、開示の実施形態の具体的構成に限定されない。発明の実施にあたって、実施形態に応じた具体的構成が適宜採用されてもよい。

【0011】

〔実施形態〕

本実施形態の翻訳システムは、利用者が端末装置に向けて発声した言葉を、他の利用者の端末装置に表示するとともに、当該言葉を他の利用者が使用する言語（例えば、日本語、英語、スペイン語など）の言葉に翻訳して表示する。本実施形態の翻訳システムでは、端末装置が利用者による音声の入力を受け付け、音声認識装置に、入力された音声の音声データを音声認識装置に送信する。音声認識装置は、音声データを文字情報に変換して、端末装置に出力する。端末装置は、所定の端末装置に向けて、文字情報を当該文字情報の言語を示す言語情報とともに、送信する。端末装置は、言語情報及び文字情報を受信すると、当該文字情報を表示し、翻訳装置に、言語情報及び文字情報と、翻訳後の言語を示す言語情報とを送信する。翻訳装置は、文字情報と当該文字情報の言語（第1言語）を示す言語情報と翻訳後の言語（第2言語）を示す言語情報を受信すると、文字情報を第1言語から第2言語に翻訳し、翻訳後の文字情報を、端末装置に送信する。端末装置は、翻訳後の文字情報を受信すると、当該文字情報を表示する。

【0012】

（構成例）

図1は、本実施形態の翻訳システムの構成例を示す図である。図1の翻訳システム10は、端末装置100A、端末装置100B、音声認識装置200、翻訳装置300、ネットワーク400、LAN（Local Area Network）500を含む。端末装置100Aは、ネットワーク400及びLAN500に接続される。端末装置100Bは、ネットワーク400及びLAN500に接続される。端末装置100A及び端末装置100Bは、同様の構成を有する。端末装置100A及び端末装置100Bを区別しない場合は、単に、端末装置100と呼ぶ。音声認識装置200は、入力される音声データを、文字情報に変換して、出力する。翻訳装置300は、入力される第1言語の文字情報を、第1言語と異なる第2言語の文字情報に翻訳して、出力する。ネットワーク400は、端末装置100、音声認識装置200、翻訳装置300等に接続されるインターネット等のネットワークである。LAN500は、端末装置100A、端末装置100Bが接続されるローカルエリアネットワークである。LAN500は、ルータを含み得る。端末装置100A、端末装置100Bは、LAN500を介して、ネットワーク500に接続されてもよい。図1のシステム10では、端末装置100の2台は、2台であるが、端末装置100の数は、2台に限定されるものではない。

【0013】

端末装置100は、利用者による音声の入力を受け付け、音声を音声データに変換して、音声認識装置200に送信する。端末装置100には、あらかじめ、端末装置100の利用者が使用する言語（使用言語）が登録されている。使用言語は、端末装置100で動作するオペレーティングシステムにおいて設定されるパラメータに含まれるものであってもよい。端末装置100は、音声データとともに使用言語の情報も、音声認識装置200に送信する。端末装置100は、音声認識装置200から変換後の文字情報を受信する。端末装置100は、受信した文字情報を、使用言語の情報とともに他の端末装置100に向けて送信する。端末装置100は、他の端末装置100などから使用言語の情報と文字情報を受信すると、当該文字情報を表示する。また、端末装置100は、当該使用言語の情報と当該文字情報と自装置の使用言語の情報とを、翻訳装置300に送信する。端末装置100は、翻訳装置300から、翻訳後の文字情報を受信すると、当該文字情報を表示する。端末装置100の構成については、後に説明する。

【0014】

音声認識装置200は、端末装置100から、当該端末装置100における使用言語の情報と、音声データとを受け付け、音声データの音声を当該使用言語の文字情報に変換して、端末装置100に送信する。音声認識装置200は、音声に所定時間継続する無音が含まれている場合、当該音声を改行を示す制御文字に変換してもよい。音声認識装置200は、端末装置100から送信される音声データをバッファし、バッファされている一連の音声データの音声を、一連の文字情報に変換してもよい。音声認識装置200は、音声を改行を示す制御文字を含む文字情報に変換した場合、当該制御文字に対応する音声データ以前の音声データをバッファから削除してもよい。これにより、音声認識装置200は、常に、改行以降の音声データについて、文字情報に変換することになる。音声認識装置200として、周知の音声認識手段が使用され得る。

10

【0015】

翻訳装置300は、端末装置100から、第1言語の情報と、文字情報と、第2言語の情報とを受け付ける。翻訳装置300は、文字情報を第1言語から第2言語に翻訳する。翻訳装置300は、翻訳した文字情報を、端末装置100に送信する。翻訳装置300として、周知の翻訳手段が使用され得る。

【0016】

〈端末装置の構成〉

図2は、端末装置の機能ブロックの例を示す図である。図2の端末装置100は、表示部102、記憶部104、音声入力部106、音声認識コマンド発行部108、音声認識結果受付部110、文字配布部112、文字受付部114、翻訳コマンド発行部116、翻訳結果受付部118、通信部120を含む。

20

【0017】

表示部102は、端末装置100の利用者に示す情報を表示する。表示部102は、文字配布部112で他の端末装置100等に向けて送信される文字情報や、翻訳結果受付部118で受信される翻訳結果である文字情報を表示する。

【0018】

記憶部104は、端末装置100で使用されるデータや、端末装置100の利用者が使用する言語（使用言語）の情報、他の装置から受信した文字情報、音声データ等を格納する。記憶部104は、自装置や他の装置等の識別情報、ネットワーク上のアドレスを格納しうる。

30

【0019】

音声入力部106は、マイクロフォンなどの入力手段により、端末装置100の利用者が発声した音声の入力を受け付ける。音声入力部106は、入力された音声を電気的な信号に変換して、音声データとして、音声認識コマンド発行部108に送信する。音声入力部106は、所定時間毎に音声データを送信する。

【0020】

音声認識コマンド発行部108は、音声入力部106より音声データを受信する。音声認識コマンド発行部108は、受信した音声データを記憶部104に格納する。音声認識コマンド発行部108は、記憶部104から、端末装置100の利用者が使用する言語として登録されている使用言語の情報を取得する。音声認識コマンド発行部108は、音声認識装置200で音声認識をさせるための音声認識コマンドを生成して、通信部120を介して、音声認識装置200に送信する。音声認識コマンドは、通信部120からネットワークを介して、音声認識装置200に送信される。音声認識コマンドとともに、音声データ、使用言語の情報も送信される。音声認識コマンドは、音声データを、使用言語の文字情報に変換することを指令するコマンドである。音声認識コマンド発行部108は、音声認識指令部の一例である。

40

【0021】

音声認識結果受付部110は、音声認識装置200による音声認識結果を、通信部120を介して、受信する。音声認識結果受付部110は、受信した音声認識結果を、文字配

50

布部 1 1 2 に送信する。

【0 0 2 2】

文字配布部 1 1 2 は、音声認識結果受付部 1 1 0 から、音声認識結果を受信する。文字配布部 1 1 2 は、受信した音声認識結果を記憶部 1 0 4 に格納する。文字配布部 1 1 2 は、受信した音声認識結果を表示部 1 0 2 に表示させる。文字配布部 1 1 2 は、記憶部 1 0 4 より、端末装置 1 0 0 の使用言語の情報を取得する。文字配布部 1 1 2 は、マルチキャスト通信により、LAN 5 0 0 上の他の端末装置に向けて、受信した音声認識結果及び取得した使用言語の情報を、通信部 1 2 0 を介して、送信する。文字配布部 1 1 2 は、文字送信部の一例である。

【0 0 2 3】

文字受付部 1 1 4 は、他の端末装置 1 0 0 からの音声認識結果及び使用言語の情報を、通信部 1 2 0 を介して、受信する。文字受付部 1 1 4 は、受信した音声認識結果及び使用言語の情報を、記憶部 1 0 4 に格納する。文字受付部 1 1 4 は、記憶部 1 0 4 より、自身の端末装置 1 0 0 の使用言語の情報を取得する。文字受付部 1 1 4 は、自身の端末の使用言語と受信した使用言語とが異なる場合、翻訳コマンド発行部 1 1 6 に、自身の端末の使用言語の情報と受信した使用言語の情報及び音声認識結果とを、翻訳コマンド発行部 1 1 6 に送信する。

【0 0 2 4】

翻訳コマンド発行部 1 1 6 は、音声認識結果及び他の端末装置 1 0 0 の使用言語（第 1 言語）の情報と、自身の端末装置 1 0 0 の使用言語（第 2 言語）の情報とを、文字受付部 1 1 4 から受信する。翻訳コマンド発行部 1 1 6 は、翻訳装置 3 0 0 で翻訳をさせるための翻訳コマンドを生成して、通信部 1 2 0 を介して、翻訳装置 3 0 0 に送信する。翻訳コマンドは、音声認識結果、第 1 言語の情報、第 2 言語の情報を含む。翻訳コマンドは、第 1 言語の文字情報である音声認識結果を、第 2 言語に翻訳することを指令するコマンドである。翻訳コマンド発行部 1 1 6 は、翻訳指令部の一例である。

【0 0 2 5】

翻訳結果受付部 1 1 8 は、翻訳装置 3 0 0 による翻訳結果を、通信部 1 2 0 を介して、受信する。翻訳結果受付部 1 1 8 は、受信した翻訳結果を、記憶部 1 0 4 に格納する。翻訳結果受付部 1 1 8 は、受信した翻訳結果を表示部 1 0 2 に表示させる。

【0 0 2 6】

通信部 1 2 0 は、他の装置等と、データの送受信を行う。通信部 1 2 0 は、所定のコマンドを、ネットワーク 4 0 0 を介して、音声認識装置 2 0 0 や翻訳装置 3 0 0 に送信する。通信部 1 2 0 は、音声認識装置 2 0 0 や翻訳装置 3 0 0 から、所定の結果を受信する。通信部 1 2 0 は、マルチキャスト通信により、LAN 5 0 0 上の他の端末装置 1 0 0 に向けて、文字情報等を送信する。通信部 1 2 0 は、マルチキャスト通信によって送信される他の端末装置 1 0 0 からの文字情報等を受信する。マルチキャスト通信は、パケット送信方法の 1 つであり、あらかじめ指定した 1 以上の端末装置に一斉にパケット（データ）を送信することができる。他の端末装置 1 0 0 へのデータの通信方法は、マルチキャスト通信に限定されるものではなく、他の通信方法が使用されてもよい。マルチキャスト通信における端末装置の指定は、例えば、端末装置 1 0 0 が、LAN 5 0 0 上のルータに、所定のアドレスで配信されるマルチキャスト通信によるデータの送信を要求することで行われる。端末装置 1 0 0 は、当該指定により、ルータから所定のアドレスで配信されるデータを受信することができる。また、端末装置 1 0 0 は、所定のアドレス宛にデータを送信することで、マルチキャスト通信による配信を行うことができる。

【0 0 2 7】

〈装置のハードウェア構成〉

端末装置 1 0 0 は、スマートフォン、携帯電話、タブレット型端末、カーナビゲーション装置、PDA (Personal Digital Assistant)、PC (Personal Computer) のような専用または汎用のコンピュータ、あるいは、コンピュータを搭載した電子機器を使用して実現可能である。音声認識装置 2 0 0、翻訳装置 3 0 0 は、PC、ワークステーション (

10

20

30

40

50

WS、Work Station) のような専用または汎用のコンピュータ、あるいは、コンピュータを搭載した電子機器を使用して実現可能である。音声認識装置 200、翻訳装置 300 は、端末装置 100 に含まれてもよい。

【0028】

図 3 は、情報処理装置のハードウェア構成例を示す図である。図 3 に示す情報処理装置 90 は、一般的なコンピュータの構成を有している。制御装置 20、端末装置 100、音声認識装置 200、翻訳装置 300 は、図 3 に示すような情報処理装置 90 によって実現される。情報処理装置 90 は、プロセッサ 91、メモリ 92、記憶部 93、入力部 94、出力部 95、通信制御部 96 を有する。これらは、互いにバスによって接続される。メモリ 92 及び記憶部 93 は、コンピュータ読み取り可能な記録媒体である。情報処理装置のハードウェア構成は、図 3 に示される例に限らず、適宜構成要素の省略、置換、追加が行われてもよい。

10

【0029】

情報処理装置 90 は、プロセッサ 91 が記録媒体に記憶されたプログラムをメモリ 92 の作業領域にロードして実行し、プログラムの実行を通じて各構成部等が制御されることによって、所定の目的に合致した機能を実現することができる。

【0030】

プロセッサ 91 は、例えば、CPU (Central Processing Unit) や DSP (Digital Signal Processor) である。

【0031】

メモリ 92 は、例えば、RAM (Random Access Memory) や ROM (Read Only Memory) を含む。メモリ 92 は、主記憶装置とも呼ばれる。

20

【0032】

記憶部 93 は、例えば、EPROM (Erasable Programmable ROM)、ハードディスクドライブ (HDD、Hard Disk Drive)、ソリッドステートドライブ (SSD、Solid State Drive) である。また、記憶部 93 は、リムーバブルメディア、即ち可搬記録媒体を含むことができる。リムーバブルメディアは、例えば、USB (Universal Serial Bus) メモリ、あるいは、CD (Compact Disc) や DVD (Digital Versatile Disc) のようなディスク記録媒体である。記憶部 93 は、二次記憶装置とも呼ばれる。

【0033】

記憶部 93 は、各種のプログラム、各種のデータ及び各種のテーブルを読み書き自在に記録媒体に格納する。記憶部 93 には、オペレーティングシステム (Operating System : OS)、各種プログラム、各種テーブル等が格納される。記憶部 93 に格納される情報は、メモリ 92 に格納されてもよい。また、メモリ 92 に格納される情報は、記憶部 93 に格納されてもよい。

30

【0034】

オペレーティングシステムは、ソフトウェアとハードウェアとの仲介、メモリ空間の管理、ファイル管理、プロセスやタスクの管理等を行うソフトウェアである。オペレーティングシステムは、通信インタフェースを含む。通信インタフェースは、通信制御部 96 を介して接続される他の外部装置等とデータのやり取りを行うプログラムである。外部装置等には、例えば、他の情報処理装置、外部記憶装置等が含まれる。

40

【0035】

入力部 94 は、キーボード、ポインティングデバイス、ワイヤレスリモコン、タッチパネル等を含む。また、入力部 94 は、カメラのような映像や画像の入力装置や、マイクロフォンのような音声の入力装置を含むことができる。

【0036】

出力部 95 は、CRT (Cathode Ray Tube) ディスプレイ、LCD (Liquid Crystal Display)、PDP (Plasma Display Panel)、EL (Electroluminescence) パネル等の表示装置、プリンタ等の出力装置を含む。また、出力部 95 は、スピーカのような音声の出力装置を含むことができる。

50

【0037】

通信制御部96は、他の装置と接続し、情報処理装置90と他の装置との間の通信を制御する。通信制御部96は、例えば、LAN (Local Area Network) インタフェースボード、Bluetooth (登録商標) などの無線通信のための無線通信回路、電話通信のための通信回路である。LANインタフェースボードや無線通信回路は、インターネット等のネットワークに接続される。

【0038】

端末装置100、音声認識装置200、翻訳装置300を実現するコンピュータは、プロセッサが二次記憶装置に記憶されているプログラムを主記憶装置にロードして実行することによって、各機能を実現する。また、各装置の記憶部は、主記憶装置または二次記憶装置の記憶領域に設けられる。

10

【0039】

プログラムを記述するステップは、記載された順序に沿って時系列的に行われる処理はもちろん、必ずしも時系列的に処理されなくとも、並列的または個別に実行される処理を含む。プログラムを記述するステップの一部が省略されてもよい。

【0040】

(動作例)

端末装置100の動作例について、説明する。ここでは、端末装置100の動作を、音声入力の動作フローと、翻訳文表示の動作フローとに分けて説明するが、これらの動作は、端末装置100において並行して実行されている。端末装置100では、これらの動作を実現するアプリケーションが実行されている。

20

【0041】

端末装置100は、あらかじめ、所定のアドレスを用いて、マルチキャスト通信によるデータの送受信をできるように設定されている。LAN500には、マルチキャスト通信によるデータの送受信をできる1以上の端末装置100が接続されている。上記の動作を実現するアプリケーションを端末装置100で実行することによって、所定のアドレスを用いてマルチキャスト通信によるデータの送受信をできるように設定されてもよい。

【0042】

〈音声入力〉

図4は、端末装置における音声入力の動作フローの例を示す図である。端末装置100は、利用者によって使用されている。利用者は、端末装置100に対する発声により音声入力を行う。図4の動作フローは、例えば、音声入力の1区間の長さである所定時間 (例えば、200ms) 毎に実行される。

30

【0043】

S101では、端末装置100の音声入力部106は、端末装置100の利用者が発声した音声の入力を受け付ける。音声入力部106は、入力された音声を電気的な信号に変換して、音声データとして、音声認識コマンド発行部108に送信する。音声入力部106は、所定時間毎 (例えば、200ms 毎) に音声データを送信する。当該所定時間が短くなるほど、発声と翻訳文表示とのタイムラグを短くすることができるが、端末装置100や音声認識装置200における処理量が増大する。当該所定時間は、発声における一音や一語程度の長さにすることが望ましい。一音や一語よりも短い音声を音声認識により文字情報に変換することが困難であるからである。

40

【0044】

S102では、音声認識コマンド発行部108は、音声入力部106より音声データを、所定時間毎に受信する。音声認識コマンド発行部108は、受信した音声データを記憶部104に格納する。音声認識コマンド発行部108は、記憶部104から、端末装置100の利用者が使用する言語として登録されている使用言語の情報を取得する。音声認識コマンド発行部108は、音声認識装置200で音声認識をさせるための音声認識コマンドを生成 (発行) して、通信部120を介して、音声認識装置200に送信する。音声認識処理は、音声データを、あらかじめ音声と所定の言語の文字とを対応付けたデータに基

50

づいて、所定の言語の文字情報に変換する処理である。音声認識コマンドは、通信部120からネットワーク400を介して、音声認識装置200に送信される。音声認識コマンドは、音声データを、使用言語の文字情報に変換することを指令するコマンドである。音声認識コマンドとともに、音声データ、使用言語の情報も送信される。音声データには一連の通し番号が含まれてもよい。また、音声認識コマンドとともに、端末装置100のアドレスや端末装置100を識別する識別子が送信されてもよい。

【0045】

音声認識装置200は、端末装置100から音声認識コマンドを受信すると、音声認識コマンドに従って、音声認識コマンドとともに送信される使用言語の文字に、音声データの音声を変換して、音声認識結果として、端末装置100に送信する。音声認識装置200は、音声に所定時間継続する無音が含まれている場合、当該音声を改行を示す制御文字に変換する。音声認識装置200は、端末装置100から送信される音声データをバッファし、バッファされている一連の音声データを結合して1つの結合音声データを作成し、当該音声データの音声を一連の文字情報に変換する。音声認識装置200は、1つの音声データを作成する際、音声データに含まれる通し番号を用いて、順番が入れ替わらないように音声データを作成してもよい。音声認識装置200は、音声を改行を示す制御文字を含む文字情報に変換した場合、当該制御文字に対応する音声データ以前の音声データをバッファから削除する。これにより、音声認識装置200は、常に、無音（改行）以降の音声データについて、文字情報に変換することになる。音声認識コマンドには、すでに送信済みの一連の音声データを1つの音声データにして音声認識をする指令、音声データに所定時間以上の無音が含まれている場合に音声認識後にバッファを削除する指令が含まれてもよい。

【0046】

S103では、音声認識結果受付部110は、S102で送信した音声認識コマンドに対する、音声認識装置200による音声認識結果を、受信したか否かを判定する。音声認識結果を受信した場合（S103；YES）、処理がS104に進む。音声認識結果を受信していない場合（S103；YES）、S103の処理を繰り返す。

【0047】

S104では、音声認識結果受付部110は、今回受信した音声認識結果と、記憶部104に格納される前回受信した音声認識結果とを比較する。今回受信した音声結果と前回受信した音声認識結果とが同じである場合（S104；YES）、この動作フローの処理が終了する。この場合、前回の音声認識結果に対してS105の処理が行われているため、同じ音声認識結果に対して同じ処理を行う必要が無いからである。また、音声認識結果に、文字情報が含まれない場合も、同様にする。文字情報が含まれない場合には、表示などの以後の処理を行う必要が無いからである。また、今回受信した音声結果と前回受信した音声認識結果とが異なる場合（S104；NO）、音声認識結果受付部110は音声認識結果を文字配布部112に送信し、処理がS105に進む。

【0048】

S105では、文字配布部112は、音声認識結果受付部110から音声認識結果を受信する。文字配布部112は、受信した音声認識結果を、今回の音声認識結果として、記憶部104に格納する。この後、新たに音声認識結果が格納される際には、今回の音声認識結果は、前回の音声認識結果となる。文字配布部112は、受信した音声認識結果を表示部102に表示させる。このとき、文字配布部112は、最後に表示した音声認識結果を消去してから、今回受信した音声認識結果の文字情報を表示する。即ち、表示される音声認識結果が更新される。ただし、文字配布部112は、最後に表示した音声認識結果に改行を示す制御文字が含まれる場合、最後に表示した音声認識結果を消去せずに、改行してから、今回受信した音声認識結果の文字情報を表示する。音声認識結果に改行を示す制御文字が含まれることで、一連の音声データに対する音声認識結果が確定する。文字配布部112は、記憶部104より、端末装置100の使用言語の情報を取得する。文字配布部112は、マルチキャスト通信により、所定のアドレスを指定して、LAN500上の

他の端末装置に向けて、受信した音声認識結果及び取得した使用言語の情報を、通信部120を介して、送信する。当該使用言語は、音声認識結果の文字情報の言語である。

【0049】

ここでは、文字配布部112は、受信した音声認識結果を表示部102に表示させるとしたが、文字配布部112は、受信した音声認識結果を表示部102に表示させなくてもよい。端末装置100がマルチキャスト通信で送信する音声認識結果は、送信した端末装置100自身にも配信される。このため、文字配布部112が表示させなくても、マルチキャスト通信で配信された音声認識結果を、後に説明する文字受付部114が表示部102に表示させることができるからである。

【0050】

〈翻訳文表示〉

図5は、端末装置における翻訳文表示の動作フローの例を示す図である。端末装置100は、利用者によって使用されている。

【0051】

S201では、端末装置100の文字受付部114は、他の端末装置100からの音声認識結果（文字情報）及び使用言語の情報を、マルチキャスト通信により、通信部120を介して、受信したか否かを判定する。受信した場合（S201；YES）、処理がS202に進む。受信していない場合（S201；NO）、S201の処理を繰り返す。

【0052】

S202では、文字受付部114は、受信した文字情報及び使用言語の情報を、記憶部104に格納する。文字受付部114は、受信した文字情報を、表示部102に表示させる。このとき、文字受付部114は、S202で最後に表示した文字情報を消去してから、今回受信した文字情報を表示する。即ち、表示される文字情報が更新される。ただし、文字受付部114は、S202で最後に表示した文字情報に改行を示す制御文字が含まれる場合、最後に表示した文字情報を消去せずに、改行してから、今回受信した文字情報を表示する。文字情報に改行を示す制御文字が含まれる場合、表示される文字情報は更新されない。

【0053】

S203では、文字受付部114は、記憶部104より、自身の端末装置100の使用言語の情報を取得する。文字受付部114は、自身の端末装置100の使用言語とS201で受信した使用言語とが一致するか否かを判定する。自身の端末装置100の使用言語とS201で受信した使用言語とが一致する場合（S203；YES）、処理がS201に戻る。この場合、受信した文字情報は翻訳されない。自身の端末装置100の使用言語とS201で受信した使用言語とが異なる場合（S203；NO）、処理がS204に進む。

【0054】

S204では、文字受付部114は、翻訳コマンド発行部116に、自身の端末装置100の使用言語の情報と受信した使用言語の情報及び文字情報（音声認識結果）とを、翻訳コマンド発行部116に送信する。翻訳コマンド発行部116は、受信した音声認識結果及び受信した使用言語（第1言語）の情報と、自身の端末装置100の使用言語（第2言語）の情報とを、文字受付部114から受信する。翻訳コマンド発行部116は、翻訳装置300で翻訳をさせるための翻訳コマンドを生成（発行）して、通信部120を介して、翻訳装置300に送信する。翻訳処理は、所定の辞書等に基づいて、第1言語の文字情報を、第2言語の文字情報に変換（翻訳）する処理である。翻訳コマンドは、通信部120からネットワーク400を介して、翻訳装置300に送信される。翻訳コマンドは、第1言語の文字情報を、第2言語の文字情報に翻訳することを指令するコマンドである。翻訳コマンドとともに、第1言語の情報、文字情報、第2言語の情報が送信される。翻訳コマンドとともに、端末装置100のアドレスや端末装置100を識別する識別子が送信されてもよい。

【0055】

10

20

30

40

50

翻訳装置 300 は、端末装置 100 から、翻訳コマンドと、第 1 言語の情報と、文字情報と、第 2 言語の情報とを受け付ける。翻訳装置 300 は、翻訳コマンドを受信すると、翻訳コマンドに従って、第 1 言語の文字情報を、第 2 言語に翻訳する。翻訳装置 300 は、翻訳した文字情報を、端末装置 100 に送信する。翻訳装置 300 は、文字情報に改行を示す制御文字が含まれる場合、翻訳後の文字情報にも当該制御文字を含むようにする。

【0056】

S205では、翻訳結果受付部 118 は、翻訳装置 300 からの翻訳結果を、通信部 120 を介して、受信したか否かを判定する。受信した場合（S205；YES）、処理が S206 に進む。受信していない場合（S205；NO）、処理が S201 に戻る。

【0057】

S206では、翻訳結果受付部 118 は、受信した翻訳結果を、記憶部 104 に格納する。翻訳結果受付部 118 は、受信した翻訳結果を、表示部 102 に表示させる。翻訳結果は、S202 で表示した文字情報の近傍に表示させる。近傍に表示させることで、端末装置 100 の利用者が翻訳前の文字情報と翻訳結果とを容易に比較することができる。このとき、翻訳結果受付部 118 は、S206 で最後に表示した翻訳結果を消去してから、今回受信した翻訳結果を表示する。即ち、表示される翻訳結果が更新される。ただし、翻訳結果受付部 118 は、S206 で最後に表示した翻訳結果に改行を示す制御文字が含まれる場合、最後に表示した翻訳結果を消去せずに、改行してから、今回受信した翻訳結果を表示する。文字情報に改行を示す制御文字が含まれる場合、表示される文字情報は更新されない。表示される文字情報の近傍には、当該文字情報の基となる音声を発声した利用者（絵など）を表示してもよい。これにより、どの文字情報がどの利用者によって発声されたものなのかを認識しやすくなる。

【0058】

S207では、端末装置 100 は、本動作フローの動作を実現するアプリケーションにおいて、アプリケーションの終了が選択されたか否かを判定する。アプリケーションの終了には、端末装置 100 自体の電源 OFF が含まれる。終了が選択された場合（S207；YES）、動作フローの処理が終了する。終了が選択されていない場合（S207；NO）、処理が S201 に戻る。

【0059】

（具体例）

図 6 は、本実施形態の翻訳システムの動作の具体例を示す図である。ここでは、利用者によって、端末装置 100 A に約 2.5 秒間（時刻 0 s から時刻 2.5 s まで）で「おはようございます」と発声された場合について説明する。端末装置 100 A の使用言語は、日本語とする。端末装置 100 B の使用言語は、英語とする。

【0060】

端末装置 100 A の音声入力部 106 は、図 6 のように、第 1 区間（時刻 0 s から時刻 0.2 s まで）である最初の 200 ms に入力される「おはようございます」の「お」の一部の音声を、電気信号に変換して、音声認識コマンド発行部 108 に送信する（図 4 の S101）。音声認識コマンド発行部 108 は、受信した音声データ、使用言語の情報とともに、音声認識コマンドを、音声認識装置 200 に送信する（図 4 の S102）。音声認識装置 200 は、音声認識コマンドにしたがって、音声データの音声の音声認識を実行するが、音声データの音声は「お」の一部であるため、ここでは音声を使用言語である日本語の文字に変換することができない。音声認識装置 200 は、音声認識結果として、文字情報の含まれない「」を端末装置 100 A に送信する。端末装置 100 A の音声認識結果受付部 110 は、文字情報の含まれない音声認識結果を受信すると、表示する文字情報が存在しないため（図 4 の S104；NO）、第 1 区間の処理が終了する。

【0061】

次に、端末装置 100 A の音声入力部 106 は、第 2 区間である 200 ms に入力される「おはようございます」の「お」の一部の音声及び「は」の一部の音声を、電気信号に変換して、音声認識コマンド発行部 108 に送信する（図 4 の S101）。音声認識コマ

ンド発行部108は、受信した音声データ、使用言語の情報とともに、音声認識コマンドを、音声認識装置200に送信する(図4のS102)。音声認識装置200は、音声認識コマンドにしたがって、第1区間の音声データと第2区間の音声データとを結合して1つの音声データとして、当該音声データの音声の音声認識を実行する。ここでは、音声データの音声が「お」及び「は」の一部であるため、ここでは音声を使用言語である日本語の文字「お」に変換する。音声認識装置200は、音声認識結果として、文字情報「お」を端末装置100Aに送信する。端末装置100Aの音声認識結果受付部110は、文字情報「お」を受信すると、文字配布部112に送信する(図4のS104;YES)。文字配布部112は、受信した音声認識結果「お」を表示部102に表示させる(上述のように表示させなくてもよい)。また、文字配布部112は、マルチキャスト通信により、文字情報「お」を使用言語の情報とともに送信する(図4のS105)。

10

【0062】

一方、マルチキャスト通信により文字情報「お」を使用言語の情報とともに受信した端末装置100Bの文字受付部114は、受信した文字情報「お」を、表示部102に表示させる(図5のS202)。翻訳コマンド発行部116は、翻訳装置300で翻訳をさせるための翻訳コマンドを生成する。翻訳コマンド発行部116は、翻訳コマンドとともに、第1言語(日本語)の情報、文字情報「お」、第2言語(英語)の情報を、翻訳装置300に送信する。翻訳装置300は、日本語の文字情報「お」を、英語に翻訳する。ここでは、英語の「your」に翻訳されたとする。翻訳装置300は、翻訳結果として、文字情報「your」を、端末装置100Bに送信する。端末装置100Bの翻訳結果受付部118は、翻訳装置300から受信した翻訳結果「your」を、文字情報「お」の近傍に、表示させる(図5のS206)。これによって、発声される音声を所定時間で区切って、所定時間毎に音声認識をし、翻訳を行うことができる。

20

【0063】

同様にして、第3区間、第4区間の処理では、それぞれ、音声認識の結果として、「おは」「おはよ」が出力され、それぞれに対して、翻訳処理が行われる。音声認識の結果は、端末装置100A、100Bの表示部102において、逐次、更新される。また、翻訳処理の結果も、端末装置100Bの表示部102において、逐次、更新される。第6区間、第7区間、第9区間、第10区間、第12区間、第13区間の処理においても同様である。

30

【0064】

第5区間の処理では、音声認識の結果が第4区間の処理と同様になるため、翻訳処理が行われない(図4のS104;YES)。よって、ここでは、音声認識の結果も、翻訳処理の結果も更新されない。第8区間、第11区間の処理においても同様である。

【0065】

次に、端末装置100Aの音声入力部106は、第14区間である200msに入力される「おはようございます」の後の無音部分を、電気信号に変換して、音声認識コマンド発行部108に送信する(図4のS101)。音声認識コマンド発行部108は、受信した音声データ、使用言語の情報とともに、音声認識コマンドを、音声認識装置200に送信する(図4のS102)。音声認識装置200は、音声認識コマンドにしたがって、第1区間から第14区間までの音声データを結合して1つの音声データとして、当該音声データの音声の音声認識を実行する。ここでは、音声データの音声が「おはようございます」であるため、ここでは音声を使用言語である日本語の文字「おはようございます」に変換する。さらに、音声認識装置200は、音声データに所定期間以上の無音部分が含まれているため、無音部分に改行を示す制御文字「[EOL]」を追加する。なお、EOLは、End of Lineの頭文字である。音声認識装置200は、音声認識結果として、文字情報「おはようございます[EOL]」を端末装置100Aに送信する。端末装置100Aの音声認識結果受付部110は、文字情報「おはようございます[EOL]」を受信すると、文字配布部112に送信する(図4のS104;YES)。文字配布部112は、受信した音声認識結果「おはようございます[EOL]」を表示部102に表示させる(上述

40

50

のように表示させなくてもよい)。改行の制御文字が含まれるため、音声認識結果は確定される。また、文字配布部112は、マルチキャスト通信により、文字情報「おはようございます[EOL]」を使用言語の情報とともに送信する(図4のS105)。

【0066】

一方、マルチキャスト通信により文字情報「おはようございます[EOL]」を使用言語の情報とともに受信した端末装置100Bの文字受付部114は、受信した文字情報「おはようございます[EOL]」を、表示部102に表示させる(図5のS202)。翻訳コマンド発行部116は、翻訳装置300で翻訳をさせるための翻訳コマンドを生成する。翻訳コマンド発行部116は、翻訳コマンドとともに、第1言語(日本語)の情報、文字情報「お」、第2言語(英語)の情報を、翻訳装置300に送信する。翻訳装置300は、日本語の文字情報「おはようございます[EOL]」を、英語に翻訳する。ここでは、英語の「Good Morning[EOL]」に翻訳されたとする。制御文字は、翻訳後もそのままにされる。翻訳装置300は、翻訳結果として、文字情報「Good Morning[EOL]」を、端末装置100Bに送信する。端末装置100Bの翻訳結果受付部118は、翻訳装置300から受信した翻訳結果「Good Morning[EOL]」を、文字情報「おはようございます[EOL]」の近傍に、表示させる(図5のS206)。表示部102に制御文字は表示されない。翻訳結果に、改行の制御文字が含まれるため、音声認識結果は確定される。

【0067】

(実施形態の作用、効果)

音声が入力される端末装置100Aでは、音声が入力されるのにしたがって、表示部102に音声認識された文字が逐次更新されて表示される。また、端末装置100Bでは、端末装置100Aで音声が入力されるのにしたがって、表示部102に、音声認識された文字と、当該文字を翻訳した翻訳結果とが、逐次、更新されて表示される。これにより、音声を入力する端末装置100Aの利用者が話している最中であっても、音声認識がされ、翻訳処理がされる。このため、端末装置100Bの利用者は、逐次更新される翻訳文を見ることで、端末装置100Bの利用者が話し終わる前に、端末装置100Bの利用者が話しているおおよその内容を理解することができる。話している最中にも翻訳を行うため翻訳結果が不正確なることもあるが、話している間に徐々に正確になるため、おおよその内容を理解するには、支障が少ない。これによって、端末装置100Bの利用者は、端末装置100Aの利用者が話し終わってすぐに話しを始めることができ、異なる言語を話す話者同士の会話をスムーズにすることができる。また、同様にして、端末装置100Bの利用者が端末装置100Bに向かって話した音声は、音声認識され、端末装置100Aで音声認識された文字と、当該文字を翻訳した翻訳結果とが、逐次、更新されて表示される。また、マルチキャスト通信を利用して、複数の端末装置100に音声認識結果を送信することで、異なる言語を使用する多くの人と一度に話をするすることができる。音声認識された文字情報が翻訳される前に、他の端末装置100に送信されることで、それぞれの端末装置100で、端末装置100に応じた翻訳コマンドが発行されることになり、それぞれの端末装置100において、それぞれの端末装置100の使用言語に翻訳され得る。

【0068】

ここでは、端末装置100が2台の例を示したが、3台以上の端末装置100が存在してもよい。このとき、例えば、日本語を使用する第1利用者と、英語を利用する第2利用者と、スペイン語を使用する第3利用者が存在する場合には、第1利用者の端末装置100には、第1利用者が話した日本語と、第2利用者が話した英語と当該英語の日本語訳と、第3利用者が話したスペイン語と当該スペイン語の日本語訳とが、逐次、表示される。

【0069】

(変形例1)

ここでは、端末装置100の音声認識コマンド発行部108は、音声認識コマンドとともに、1区間分ずつ音声データを送るとしたが、音声認識装置200で行う音声データの結合を端末装置100で行い、音声認識を行う結合した一連の音声データを音声認識コマ

10

20

30

40

50

ンドとともに音声認識装置 200 に送信してもよい。このとき、音声認識装置 200 における負荷を軽減することができる。

【0070】

(変形例 2)

ここでは、標準言語を利用する変形例について説明する。

上記の図 3 の音声入力では、端末装置 100 は、音声を、音声を受け付けた言語の文字情報に変換し、他の端末に向けて、当該文字情報を他の端末に向けて、送信していた。ここでは、音声を、音声を受け付けた言語の文字情報に変換し、さらに、標準言語の文字情報に変換して、他の端末に向けて送信する。

【0071】

図 7 は、変形例 2 の端末装置における音声入力の動作フローの例を示す図である。端末装置 100 は、利用者によって使用されている。利用者は、端末装置 100 に対する発声により音声入力を行う。図 7 の動作フローは、例えば、音声入力の 1 区間の長さである所定時間（例えば、200ms）毎に実行される。

【0072】

S301 から S303 までの動作は、図 4 の S101 から S103 までの動作と同様である。ここでは、説明を省略する。

【0073】

S304 では、音声認識結果受付部 110 は、今回受信した音声認識結果と、記憶部 104 に格納される前回受信した音声認識結果とを比較する。今回受信した音声結果と前回受信した音声認識結果とが同じである場合（S304；YES）、処理が S301 に戻る。この場合、前回の音声認識結果に対して S305 以降の処理が行われているため、同じ音声認識結果に対して同じ処理を行う必要が無いからである。また、音声認識結果に、文字情報が含まれない場合も、同様にする。文字情報が含まれない場合には、表示などの以後の処理を行う必要が無いからである。また、今回受信した音声結果と前回受信した音声認識結果とが異なる場合（S304；NO）、音声認識結果受付部 110 は音声認識結果を文字配布部 112 に送信し、処理が S305 に進む。

【0074】

S305 では、文字配布部 112 は、音声認識結果受付部 110 から音声認識結果を受信する。文字配布部 112 は、受信した音声認識結果を、今回の音声認識結果として、記憶部 104 に格納する。この後、新たに音声認識結果が格納される際には、今回の音声認識結果は、前回の音声認識結果となる。文字配布部 112 は、受信した音声認識結果を表示部 102 に表示させる。このとき、文字配布部 112 は、最後に表示した音声認識結果を消去してから、今回受信した音声認識結果の文字情報を表示する。即ち、表示される音声認識結果が更新される。ただし、文字配布部 112 は、最後に表示した音声認識結果に改行を示す制御文字が含まれる場合、最後に表示した音声認識結果を消去せずに、改行してから、今回受信した音声認識結果の文字情報を表示する。音声認識結果に改行を示す制御文字が含まれることで、一連の音声データに対する音声認識結果が確定する。

【0075】

S306 では、文字配布部 112 は、記憶部 104 より、端末装置 100 の使用言語の情報を取得する。さらに、文字配布部 112 は、記憶部 104 より、LAN500 上の他の端末と共通に使用する標準言語の情報を取得する。標準言語の情報は、あらかじめ、記憶部 104 に格納されている。文字配布部 112 は、自身の端末装置 100 の使用言語の情報と、標準言語との情報を比較して、使用言語と標準言語とが一致するか否かを判定する。自身の端末装置 100 の使用言語と標準言語とが一致する場合（S306；YES）、処理が S309 に進む。この場合、音声認識結果は翻訳されない。自身の端末装置 100 の使用言語と標準言語とが異なる場合（S306；NO）、処理が S307 に進む。標準言語は、例えば、英語である。

【0076】

S307 では、文字配布部 112 は、翻訳コマンド発行部 116 に、標準言語の情報と

10

20

30

40

50

自身の端末装置100の使用言語の情報及び今回受信した音声認識結果とを、翻訳コマンド発行部116に送信する。翻訳コマンド発行部116は、自身の端末装置100の使用言語（第1言語）の情報及び今回受信した音声認識結果と標準言語（第2言語）の情報とを、文字配布部112から受信する。翻訳コマンド発行部116は、翻訳装置300で翻訳をさせるための翻訳コマンドを生成（発行）して、通信部120を介して、翻訳装置300に送信する。翻訳処理は、所定の辞書等に基づいて、第1言語の文字情報を、第2言語の文字情報に変換（翻訳）する処理である。翻訳コマンドは、通信部120からネットワーク400を介して、翻訳装置300に送信される。翻訳コマンドは、第1言語の文字情報を、第2言語の文字情報に翻訳することを指令するコマンドである。翻訳コマンドとともに、第1言語の情報、文字情報、第2言語の情報が送信される。翻訳コマンドとともに、端末装置100のアドレスや端末装置100を識別する識別子が送信されてもよい。

10

【0077】

翻訳装置300は、端末装置100から、翻訳コマンドと、第1言語の情報と、文字情報と、第2言語の情報とを受け付ける。翻訳装置300は、翻訳コマンドを受信すると、翻訳コマンドに従って、第1言語の文字情報を、第2言語に翻訳する。翻訳装置300は、翻訳した文字情報を、端末装置100に送信する。翻訳装置300は、文字情報に改行を示す制御文字が含まれる場合、翻訳後の文字情報にも当該制御文字を含むようにする。

【0078】

S308では、翻訳結果受付部118は、翻訳装置300からの翻訳結果を、通信部120を介して、受信したか否かを判定する。受信した場合（S308；YES）、処理がS308に進む。受信していない場合（S308；NO）、処理がS301に戻る。

20

【0079】

S309では、翻訳結果受付部118は、受信した翻訳結果を、記憶部104に格納する。翻訳結果受付部118は、受信した翻訳結果を、表示部102に表示させる。さらに、文字配布部112は、マルチキャスト通信により、所定のアドレスを指定して、LAN500上の他の端末装置に向けて、受信した翻訳結果及び標準言語の情報を、通信部120を介して、送信する。当該標準言語は、翻訳結果の文字情報の言語である。また、S306で標準言語と一致した場合（S306；YES）には、文字配布部112は、マルチキャスト通信により、所定のアドレスを指定して、LAN500上の他の端末装置に向けて、受信した音声認識結果及び標準言語の情報を、通信部120を介して、送信する。

30

【0080】

S310では、端末装置100は、本動作フローの動作を実現するアプリケーションにおいて、アプリケーションの終了が選択されたか否かを判定する。アプリケーションの終了には、端末装置100自体の電源OFFが含まれる。終了が選択された場合（S310；YES）、動作フローの処理が終了する。終了が選択されていない場合（S310；NO）、処理がS301に戻る。

【0081】

変形例2によると、端末装置100は、第1言語で話された音声を、一旦、英語などの標準言語に翻訳して、LAN500上の他の端末に向けて送信する。また、標準言語の文字情報を受信した他の端末では、図5の動作フローに従って、標準言語の文字情報から他の端末の使用言語に翻訳されて表示される。これにより、翻訳装置300において、第1言語から第2言語への翻訳精度が低い場合であっても、第1言語から標準言語、標準言語から第2言語への翻訳精度が高ければ、結果的に、第1言語から第2言語への翻訳精度を高くすることができる。英語やフランス語のような汎用の言語を標準言語とすることで、どの言語からどの言語への翻訳でも翻訳精度を高めることができる。一般に、翻訳装置300において、どの言語であっても汎用の言語との間の翻訳精度は高いと考えられるからである。また、第1言語から第2言語に翻訳する翻訳装置が存在しない場合であっても、第1言語から標準言語に、標準言語から第2言語に翻訳する翻訳装置が存在すれば、第1言語から第2言語への翻訳を実現することができる。これにより、すべての言語間で翻訳装置を用意しなくても、標準言語との間の翻訳装置さえ用意すれば、すべての言語間の翻

40

50

訳を実現することができる。

【0082】

(変形例3)

ここでは、音声入力した端末で、翻訳結果も表示する変形例について説明する。上記の例と共通部分については、説明を省略する。

【0083】

〈端末装置の構成〉

図8は、変形例3の端末装置の機能ブロックの例を示す図である。図8の端末装置1100は、表示部1102、記憶部1104、音声入力部1106、音声認識コマンド発行部1108、音声認識結果受付部1110、文字配布受付部1113、翻訳コマンド発行部1116、翻訳結果受付部1118、通信部1120を含む。

10

【0084】

表示部1102は、端末装置100の表示部102と同様である。記憶部1104は、端末装置100の記憶部104と同様である。音声入力部1106は、端末装置100の音声入力部1106と同様である。音声認識コマンド発行部1108は、端末装置100の音声認識コマンド発行部108と同様である。音声認識結果受付部1110は、端末装置100の音声認識結果受付部110と同様である。

【0085】

文字配布受付部1113は、音声認識結果受付部1110から、音声認識結果を受信する。文字配布受付部1113は、受信した音声認識結果を記憶部1104に格納する。文字配布受付部1113は、受信した音声認識結果を表示部1102に表示させる。文字配布受付部1113は、記憶部1104より、端末装置1100の使用言語の情報を取得する。文字配布受付部1113は、受信した音声認識結果及び使用言語の情報を、記憶部1104に格納する。文字配布受付部1113は、翻訳コマンド発行部1116に、自身の端末の使用言語の情報と翻訳先の言語の情報及び音声認識結果とを、翻訳コマンド発行部1116に送信する。

20

【0086】

翻訳コマンド発行部1116は、端末装置100の翻訳コマンド発行部116と同様である。翻訳結果受付部1118は、端末装置100の翻訳結果受付部118と同様である。通信部1120は、端末装置100の通信部120と同様である。

30

(変形例3の動作例)

図9は、変形例3の端末装置における音声入力の動作フローの例を示す図である。端末装置100は、利用者によって使用されている。利用者は、端末装置100に対する発声により音声入力を行う。図9の動作フローは、例えば、音声入力の1区間の長さである所定時間（例えば、200ms）毎に実行される。

【0087】

S401からS403までの動作は、図4のS101からS103までの動作と同様である。ここでは、説明を省略する。

【0088】

S404では、音声認識結果受付部1110は、今回受信した音声認識結果と、記憶部1104に格納される前回受信した音声認識結果とを比較する。今回受信した音声結果と前回受信した音声認識結果とが同じである場合（S404；YES）、処理がS401に戻る。この場合、前回の音声認識結果に対してS405以降の処理が行われているため、同じ音声認識結果に対して同じ処理を行う必要が無いからである。また、音声認識結果に、文字情報が含まれない場合も、同様にする。文字情報が含まれない場合には、表示などの以後の処理を行う必要が無いからである。また、今回受信した音声結果と前回受信した音声認識結果とが異なる場合（S404；NO）、音声認識結果受付部1110は音声認識結果を文字配布受付部1113に送信し、処理がS405に進む。

40

【0089】

S405では、文字配布受付部1113は、音声認識結果受付部1110から音声認識

50

結果を受信する。文字配布受付部1113は、受信した音声認識結果を、今回の音声認識結果として、記憶部1104に格納する。この後、新たに音声認識結果が格納される際には、今回の音声認識結果は、前回の音声認識結果となる。文字配布受付部1113は、受信した音声認識結果を表示部1102に表示させる。このとき、文字配布受付部1113は、最後に表示した音声認識結果を消去してから、今回受信した音声認識結果の文字情報を表示する。即ち、表示される音声認識結果が更新される。ただし、文字配布受付部1113は、最後に表示した音声認識結果に改行を示す制御文字が含まれる場合、最後に表示した音声認識結果を消去せずに、改行してから、今回受信した音声認識結果の文字情報を表示する。音声認識結果に改行を示す制御文字が含まれることで、一連の音声データに対する音声認識結果が確定する。

10

【0090】

S406では、文字配布受付部1113は、翻訳コマンド発行部1116に、標準言語の情報と自身の端末装置1100の使用言語の情報及び今回受信した音声認識結果とを、翻訳コマンド発行部1116に送信する。翻訳コマンド発行部1116は、自身の端末装置1100の使用言語（第1言語）の情報及び今回受信した音声認識結果と翻訳先の言語（第2言語）の情報とを、文字配布受付部1113から受信する。翻訳コマンド発行部1116は、翻訳装置300で翻訳をさせるための翻訳コマンドを生成（発行）して、通信部1120を介して、翻訳装置300に送信する。翻訳処理は、所定の辞書等に基づいて、第1言語の文字情報を、第2言語の文字情報に変換（翻訳）する処理である。翻訳コマンドは、通信部1120からネットワーク400を介して、翻訳装置300に送信される。翻訳コマンドは、第1言語の文字情報を、第2言語の文字情報に翻訳することを指令するコマンドである。翻訳コマンドとともに、第1言語の情報、文字情報、第2言語の情報が送信される。翻訳コマンドとともに、端末装置1100のアドレスや端末装置1100を識別する識別子が送信されてもよい。

20

【0091】

翻訳装置300は、端末装置100から、翻訳コマンドと、第1言語の情報と、文字情報と、第2言語の情報とを受け付ける。翻訳装置300は、翻訳コマンドを受信すると、翻訳コマンドに従って、第1言語の文字情報を、第2言語に翻訳する。翻訳装置300は、翻訳した文字情報を、端末装置100に送信する。翻訳装置300は、文字情報に改行を示す制御文字が含まれる場合、翻訳後の文字情報にも当該制御文字を含むようにする。

30

【0092】

S407では、翻訳結果受付部1118は、翻訳装置300からの翻訳結果を、通信部120を介して、受信したか否かを判定する。受信した場合（S407；YES）、処理がS408に進む。受信していない場合（S407；NO）、処理がS401に戻る。

【0093】

S408では、翻訳結果受付部1118は、受信した翻訳結果を、記憶部1104に格納する。翻訳結果受付部1118は、受信した翻訳結果を、表示部1102に表示させる。翻訳結果は、S405で表示した文字情報の近傍に表示させる。近傍に表示させることで、端末装置100の利用者が翻訳前の文字情報と翻訳結果とを容易に比較することができる。このとき、翻訳結果受付部1118は、S408で最後に表示した翻訳結果を消去してから、今回受信した翻訳結果を表示する。即ち、表示される翻訳結果が更新される。ただし、翻訳結果受付部1118は、S408で最後に表示した翻訳結果に改行を示す制御文字が含まれる場合、最後に表示した翻訳結果を消去せずに、改行してから、今回受信した翻訳結果を表示する。文字情報に改行を示す制御文字が含まれる場合、表示される文字情報は更新されない。

40

【0094】

S409では、端末装置1100は、本動作フローの動作を実現するアプリケーションにおいて、アプリケーションの終了が選択されたか否かを判定する。アプリケーションの終了には、端末装置1100自体の電源OFFが含まれる。終了が選択された場合（S409；YES）、動作フローの処理が終了する。終了が選択されていない場合（S409

50

; NO)、処理がS 4 0 1に戻る。

【0095】

ここでは、翻訳先の言語を第2言語の1言語としたが、翻訳先の言語が2言語以上存在してもよい。このとき、翻訳コマンド発行部1116は、複数の言語に翻訳するため、複数の翻訳コマンドを発行する。また、翻訳結果受付部1118は、複数言語の翻訳結果を、表示部1102に表示させる。変形例3によれば、端末装置1100において、1つの言語で音声を入力することで、入力した言語の文字情報と、翻訳された言語の文字情報とが、表示される。例えば、駅などに設置されるデジタルサイネージ等を端末装置1100とすれば、駅の放送などを入力音声とすることで、駅の放送が文字情報にされ、更に、リアルタイムで、他の言語に翻訳されて表示部に表示される。

10

【0096】

本発明は、上述した実施の形態に限定されるものではなく、本発明の趣旨を逸脱しない範囲内において変更したり組み合わせたりすることができる。

【0097】

〈コンピュータ読み取り可能な記録媒体〉

コンピュータその他の機械、装置（以下、コンピュータ等）に上記いずれかの機能を実現させるプログラムをコンピュータ等が読み取り可能な記録媒体に記録することができる。そして、コンピュータ等に、この記録媒体のプログラムを読み込ませて実行させることにより、その機能を提供させることができる。

20

【0098】

ここで、コンピュータ等が読み取り可能な記録媒体とは、データやプログラム等の情報を電氣的、磁氣的、光学的、機械的、または化学的作用によって蓄積し、コンピュータ等から読み取ることができる記録媒体をいう。このような記録媒体内には、CPU、メモリ等のコンピュータを構成する要素を設け、そのCPUにプログラムを実行させてもよい。

【0099】

また、このような記録媒体のうちコンピュータ等から取り外し可能なものとしては、例えばフレキシブルディスク、光磁気ディスク、CD-ROM、CD-R/W、DVD、DAT、8mmテープ、メモ리카ード等がある。

【0100】

また、コンピュータ等に固定された記録媒体としてハードディスクやROM等がある。

30

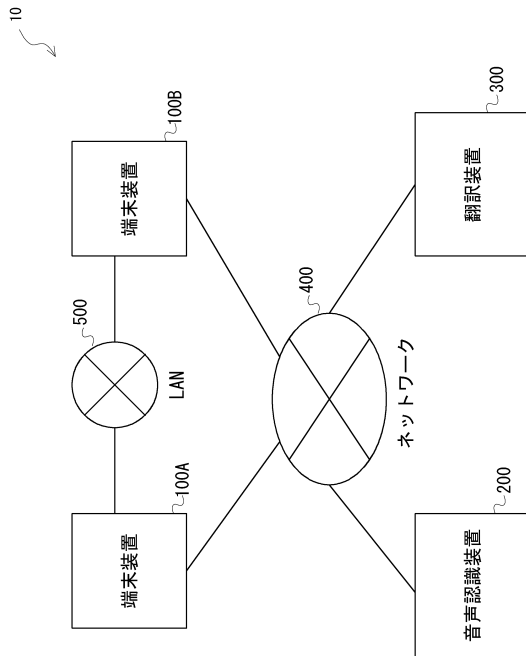
【符号の説明】

【0101】

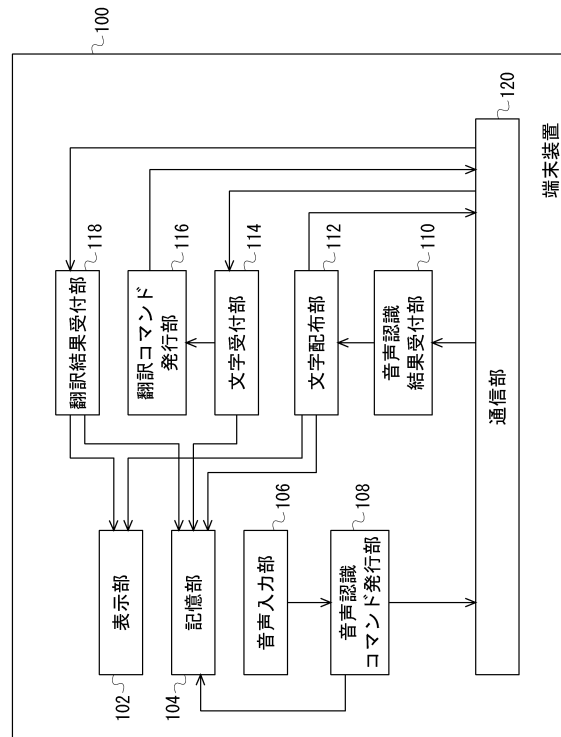
10	翻訳システム
100 (A、B)	端末装置
102	表示部
104	記憶部
106	音声入力部
108	音声認識コマンド発行部
110	音声認識結果受付部
112	文字配布部
114	文字受付部
116	翻訳コマンド発行部
118	翻訳結果受付部
120	通信部
200	音声認識装置
300	翻訳装置
400	ネットワーク
500	L A N

40

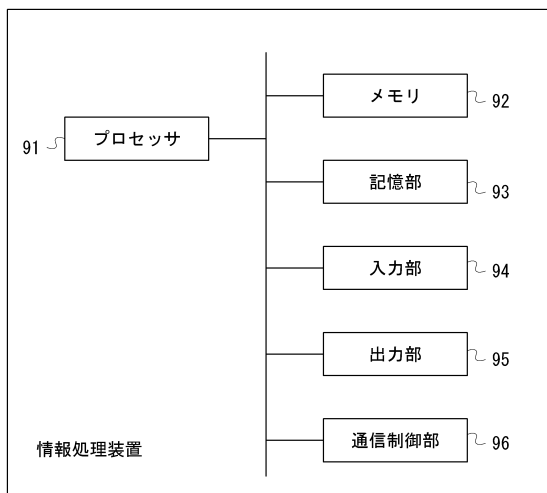
【図 1】



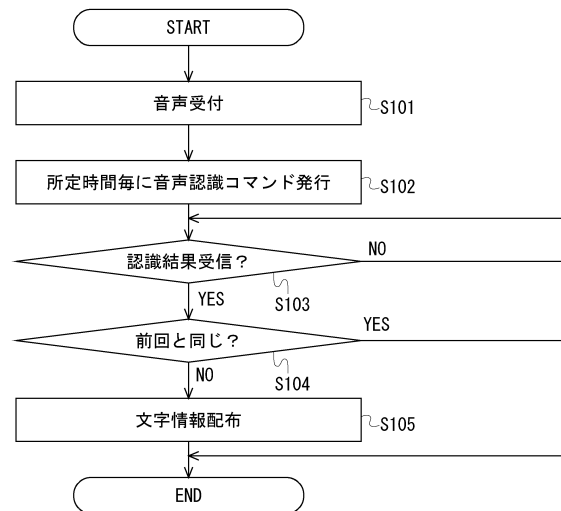
【図 2】



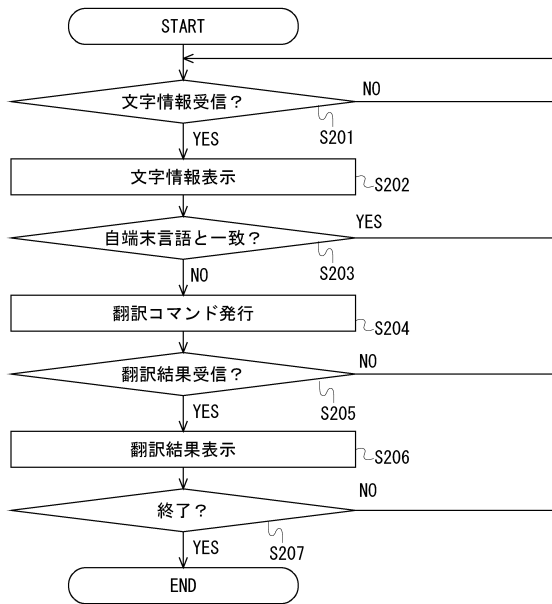
【図 3】



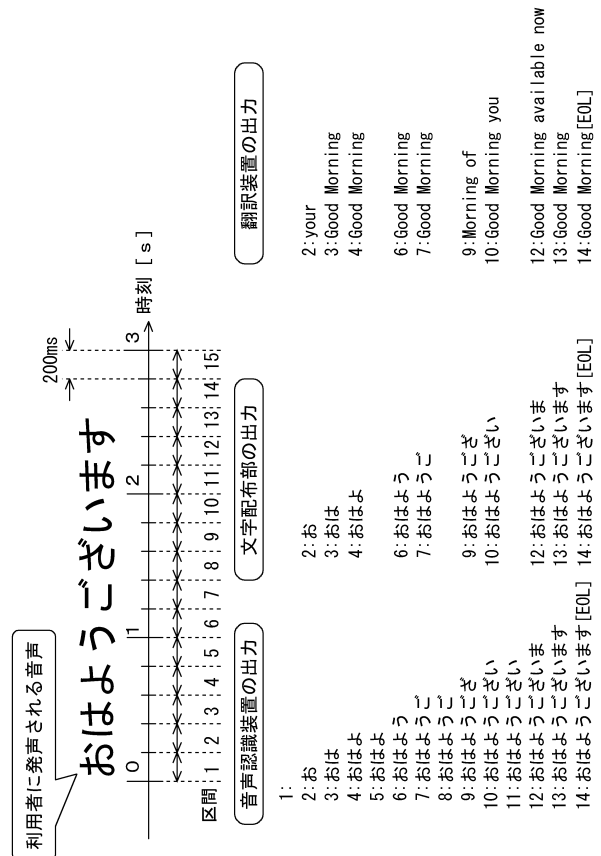
【図 4】



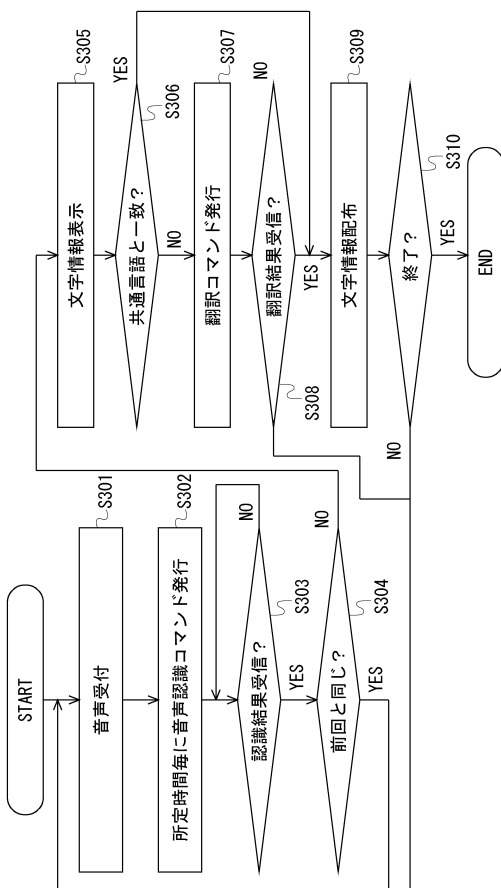
【図 5】



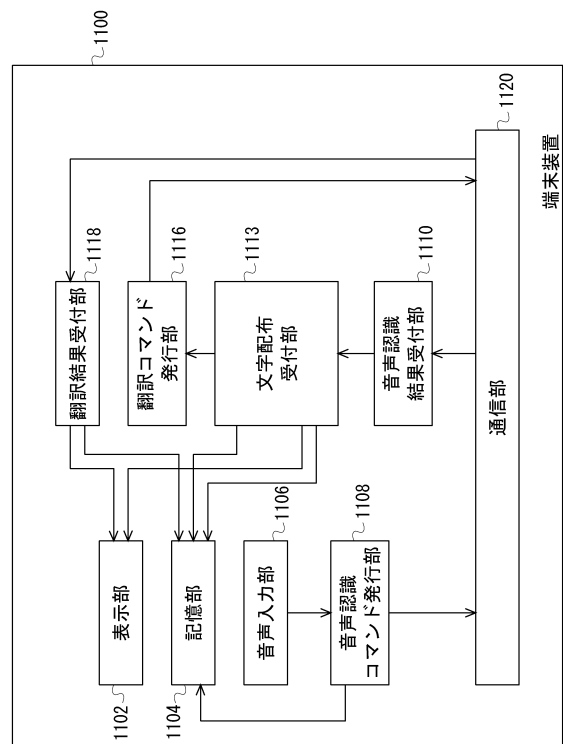
【図 6】



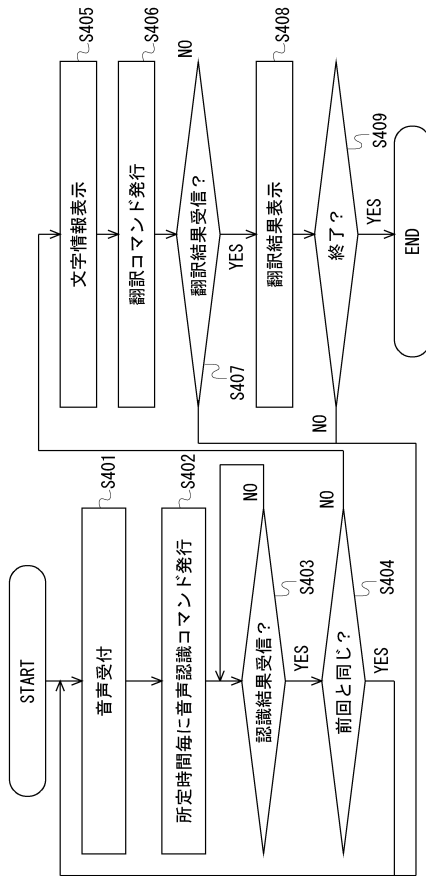
【図 7】



【図 8】



【図 9】



フロントページの続き

特許法第30条第2項適用 (1) 平成28年5月17日～5月20日に富士通フォーラム東京(東京国際フォーラム(東京都千代田区丸の内3丁目5番1号))で公開 (2) 平成28年6月25日に一般社団法人全国高等教育障害学生支援協議会第2回大会(東京大学先端科学技術研究センター(東京都目黒区駒場4丁目6-1))で公開 (3) 平成28年6月27日に富士通株式会社株主総会(新横浜プリンスホテル(神奈川県横浜市港北区新横浜3丁目4))で公開 (4) 平成28年7月5日に多言語対応・ICT化推進フォーラム(東京自治会館(東京都府中市新町2-77-1))で公開

(56)参考文献 特開2001-117920(JP, A)
特開2015-011170(JP, A)
国際公開第2016/055864(WO, A1)
特開平09-292971(JP, A)
特表2006-512647(JP, A)
特開2008-158328(JP, A)
特開2011-182125(JP, A)
米国特許出願公開第2014/0242955(US, A1)

(58)調査した分野(Int.Cl., DB名)
G10L 15/00-17/26
G06F 3/16
G06F 40/00-40/58