

(12) 特許協力条約に基づいて公開された国際出願

(19) 世界知的所有権機関
国際事務局

(43) 国際公開日
2020年4月2日(02.04.2020)



(10) 国際公開番号

WO 2020/065840 A1

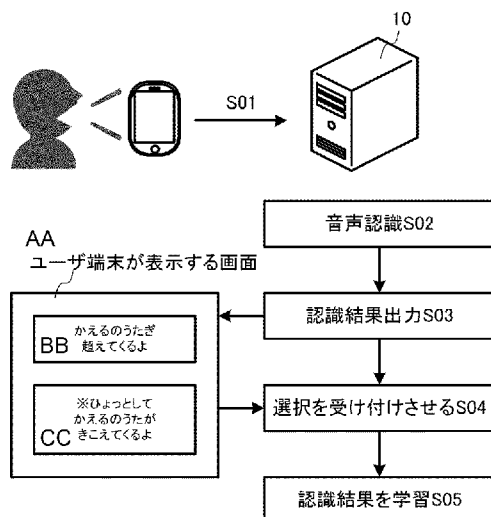
(51) 国際特許分類:
G10L 15/32 (2013.01) *G10L 15/22* (2006.01)
(21) 国際出願番号: PCT/JP2018/036001
(22) 国際出願日: 2018年9月27日(27.09.2018)
(25) 国際出願の言語: 日本語
(26) 国際公開の言語: 日本語
(71) 出願人: 株式会社 オプティム (OPTIM CORPORATION) [JP/JP]; 〒8408502 佐賀県佐賀市本庄町 1 Saga (JP).
(72) 発明者: 菅谷 俊二(SUGAYA Shunji); 〒1050022 東京都港区海岸 1 丁目 2 番 20 号 汐留ビルディング 2 1 階 株式会社オプティム内 Tokyo (JP).

(74) 代理人: 小木 智彦(KOGI Tomohiko); 〒8800804 宮城県宮崎市宮田町 1 1 - 2 4 黒木ビル 1 F Miyazaki (JP).

(81) 指定国(表示のない限り、全ての種類の国内保護が可能): AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, ME, MG, MK, MN, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL,

(54) Title: COMPUTER SYSTEM, SPEECH RECOGNITION METHOD, AND PROGRAM

(54) 発明の名称: コンピュータシステム、音声認識方法及びプログラム



S02 Perform speech recognition
S03 Output recognition result
S04 Cause selection to be received
S05 Learn recognition result
AA Screen image displayed by user terminal
BB I gear the song of the frogs
CC *Did you mean, "I hear the song of the frogs"?

(57) Abstract: [Problem] The purpose of the present invention is to provide a computer system, a speech recognition method, and a program, whereby the accuracy of speech recognition with respect to recognition results is easily enhanced. [Solution] This computer system acquires speech data, performs speech recognition of the acquired speech data, performs speech recognition of the acquired speech data using a different algorithm or database than the first recognition means, and outputs both recognition results when the recognition results of each respective speech recognition are different. The computer system also acquires speech data, performs speech recognition of the acquired speech data, performs N speech recognitions using mutually different algorithms or databases, and outputs only the recognition results that differ from

SM, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA,
UG, US, UZ, VC, VN, ZA, ZM, ZW.

- (84) 指定国(表示のない限り、全ての種類の広域保護が可能): ARIPO (BW, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SD, SL, ST, SZ, TZ, UG, ZM, ZW), ユーラシア (AM, AZ, BY, KG, KZ, RU, TJ, TM), ヨーロッパ (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

添付公開書類:

- 一 国際調査報告 (条約第21条(3))

among the N speech recognitions performed.

(57) 要約: 【課題】本発明は、音声認識の認識結果に対する正確性を向上させることが容易なコンピュータシステム、音声認識方法及びプログラムを提供することを目的とする。【解決手段】コンピュータシステムは、音声データを取得し、取得した前記音声データの音声認識を行い、取得した前記音声データの音声認識を、前記第一認識手段とは異なるアルゴリズム又はデータベースで行い、其々の音声認識の認識結果が異なる場合、双方の認識結果を出力させる。また、コンピュータシステムは、音声データを取得し、取得した前記音声データの音声認識を行い、互いに異なるアルゴリズム又はデータベースでN通りの音声認識を行い、前記N通りで行った音声認識のうち、認識結果が異なるもののみを出力させる。

明 細 書

発明の名称：

コンピュータシステム、音声認識方法及びプログラム

技術分野

[0001] 本発明は、音声認識を実行するコンピュータシステム、音声認識方法及びプログラムに関する。

背景技術

[0002] 近年、様々な分野において、音声入力が増進に行われている。このような音声入力の例としては、スマートフォンやタブレット端末等の携帯端末や、スマートスピーカ等に音声入力を行い、これらの端末類の操作、情報の検索又は連携家電の操作等を行うものがある。そのため、より正確な音声認識技術の需要が高まっている。

[0003] このような音声認識技術として、音響モデルと言語モデルとの異なるモデルにおける其々の音声認識の認識結果を結合することにより、最終的な認識結果を出力する構成が開示されている（特許文献1参照）。

先行技術文献

特許文献

[0004] 特許文献1：特開2017-40919号公報

発明の概要

発明が解決しようとする課題

[0005] しかしながら、特許文献1の構成では、複数の音声認識エンジンではなく、単一の音声認識エンジンが複数のモデルで音声認識するものに過ぎないことから、音声認識の正確性が十分なものではなかった。

[0006] 本発明は、音声認識の認識結果に対する正確性を向上させることが容易なコンピュータシステム、音声認識方法及び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

- [0007] 本発明では、以下のような解決手段を提供する。
- [0008] 本発明は、音声データを取得する取得手段と、
取得した前記音声データの音声認識を行う第一認識手段と、
取得した前記音声データの音声認識を、前記第一認識手段とは異なるアルゴリズム又はデータベースで行う第二認識手段と、
其々の音声認識の認識結果が異なる場合、双方の認識結果を出力させる出力手段と、
を備えることを特徴とするコンピュータシステムを提供する。
- [0009] 本発明によれば、コンピュータシステムは、音声データを取得し、取得した前記音声データの音声認識を行い、取得した前記音声データの音声認識を、前記第一認識手段とは異なるアルゴリズム又はデータベースで行い、其々の音声認識の認識結果が異なる場合、双方の認識結果を出力させる。
- [0010] 本発明は、コンピュータシステムのカテゴリであるが、方法及びプログラム等の他のカテゴリにおいても、そのカテゴリに応じた同様の作用・効果を発揮する。
- [0011] また、本発明は、音声データを取得する取得手段と、
取得した前記音声データの音声認識を行い、互いに異なるアルゴリズム又はデータベースでN通りの音声認識を行うN通りの認識手段と、
前記N通りで行った音声認識のうち、認識結果が異なるもののみを出力させる出力手段と、
を備えることを特徴とするコンピュータシステムを提供する。
- [0012] 本発明によれば、コンピュータシステムは、音声データを取得し、取得した前記音声データの音声認識を行い、互いに異なるアルゴリズム又はデータベースでN通りの音声認識を行い、前記N通りで行った音声認識のうち、認識結果が異なるもののみを出力させる。
- [0013] 本発明は、コンピュータシステムのカテゴリであるが、方法及びプログラム等の他のカテゴリにおいても、同様の作用・効果を発揮する。

発明の効果

[0014] 本発明によれば、音声認識の認識結果に対する正確性を向上させることが容易なコンピュータシステム、音声認識方法及びプログラムを提供することが容易となる。

図面の簡単な説明

[0015] [図1]図1は、音声認識システム1の概要を示す図である。

[図2]図2は、音声認識システム1の全体構成図である。

[図3]図3は、コンピュータ10が実行する第一の音声認識処理を示すフローチャートである。

[図4]図4は、コンピュータ10が実行する第二の音声認識処理を示すフローチャートである。

[図5]図5は、コンピュータ10が認識結果データをユーザ端末の表示部に出力させた状態を示す図である。

[図6]図6は、コンピュータ10が認識結果データをユーザ端末の表示部に出力させた状態を示す図である。

[図7]図7は、コンピュータ10が認識結果データをユーザ端末の表示部に出力させた状態を示す図である。

発明を実施するための形態

[0016] 以下、本発明を実施するための最良の形態について図を参照しながら説明する。なお、これはあくまでも一例であって、本発明の技術的範囲はこれに限られるものではない。

[0017] [音声認識システム1の概要]

本発明の好適な実施形態の概要について、図1に基づいて説明する。図1は、本発明の好適な実施形態である音声認識システム1の概要を説明するための図である。音声認識システム1は、コンピュータ10から構成され、音声認識を実行するコンピュータシステムである。

[0018] なお、音声認識システム1は、ユーザが所持するユーザ端末（携帯端末やスマートスピーカ等）等の他の端末類が含まれていてもよい。

- [0019] コンピュータ 10 は、ユーザが発した音声を、音声データとして取得する。この音声データは、ユーザ端末に内蔵されたマイク等の集音装置によりユーザが発した音声を集音し、ユーザ端末がこの集音した音声を、音声データとしてコンピュータ 10 に送信する。コンピュータ 10 は、この音声データを受信することにより、音声データを取得する。
- [0020] コンピュータ 10 は、この取得した音声データを、第一の音声解析エンジンにより音声認識を行う。また、コンピュータ 10 は、同時に、この取得した音声データを、第二の音声解析エンジンにより音声認識を行う。この第一の音声解析エンジンと第二の音声解析エンジンとは、其々、異なるアルゴリズム又はデータベースによるものである。
- [0021] コンピュータ 10 は、第一の音声解析エンジンの認識結果と、第二の音声解析エンジンの認識結果とが異なる場合、双方の認識結果をユーザ端末に出力させる。ユーザ端末はこの双方の認識結果を、自身の表示部等に表示又はスピーカ等から放音することにより、ユーザに双方の認識結果を通知する。その結果、コンピュータ 10 は、双方の認識結果を、ユーザに通知させることになる。
- [0022] コンピュータ 10 は、出力させた双方の認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる。ユーザ端末は、表示した認識結果へのタップ操作等の入力を受け付け、正しい認識結果の選択を受け付ける。また、ユーザ端末は、放音した認識結果への音声入力を受け付け、正しい認識結果の選択を受け付ける。ユーザ端末は、この選択された認識結果を、コンピュータ 10 に送信する。コンピュータ 10 は、この認識結果を取得することにより、ユーザが選択した正しい認識結果を取得する。その結果、コンピュータ 10 は、正しい認識結果の選択を受け付けさせることになる。
- [0023] コンピュータ 10 は、第一の音声解析エンジンと第二の音声解析エンジンのうち、正しい認識結果として選択されなかった音声解析エンジンに対して、選択された正しい認識結果に基づいて学習させる。例えば、第一の音声解析エンジンの認識結果が正しい認識結果として選択を受け付けさせていた場

合、第二の音声解析エンジンに、この第一の音声解析エンジンの認識結果を学習させる。

[0024] また、コンピュータ 10 は、この取得した音声データを、N 通りの音声解析エンジンにより音声認識を行う。このとき、N 通りの音声解析エンジンは、其々、互いに異なるアルゴリズム又はデータベースによるものである。

[0025] コンピュータ 10 は、N 通りの音声解析エンジンによる認識結果のうち、認識結果が異なるものをユーザ端末に出力させる。ユーザ端末この認識結果が異なるものを自身の表示部等に表示又はスピーカ等から放音することにより、ユーザに認識結果が異なるものを通知する。その結果、コンピュータ 10 は、N 通りの認識結果のうち、認識結果が異なるものをユーザに通知させることになる。

[0026] コンピュータ 10 は、出力させた認識結果が異なるもののうち、ユーザから正しい認識結果の選択を受け付けさせる。ユーザ端末は、表示した認識結果へのタップ操作等の入力を受け付け、正しい認識結果の選択を受け付ける。また、ユーザ端末は、放音した認識結果への音声入力を受け付け、正しい認識結果の選択を受け付ける。ユーザ端末は、この選択された認識結果を、コンピュータ 10 に送信する。コンピュータ 10 は、この認識結果を取得することにより、ユーザが選択した正しい認識結果を取得する。その結果、コンピュータ 10 は、正しい認識結果の選択を受け付けさせることになる。

[0027] コンピュータ 10 は、認識結果が異なるもののうち、正しい認識結果として選択されなかった音声解析エンジンに対して、選択された正しい認識結果に基づいて学習させる。例えば、第一の音声解析エンジンの認識結果が正しい認識結果として選択を受け付けさせていた場合、それ以外の認識結果の音声解析エンジンに、この第一の音声解析エンジンの認識結果を学習させる。

[0028] 音声認識システム 1 が実行する処理の概要について説明する。

[0029] はじめに、コンピュータ 10 は、音声データを取得する（ステップ S 0 1）。コンピュータ 10 は、ユーザ端末が入力を受け付けた音声を、音声データとして取得する。ユーザ端末は、自身に内蔵された集音装置によりユーザ

が発した音声を集音し、この集音した音声を音声データとしてコンピュータ 10 に送信する。コンピュータ 10 は、この音声データを受信することにより、音声データを取得する。

[0030] コンピュータ 10 は、この音声データを、第一の音声解析エンジン及び第二の音声解析エンジンにより音声認識する（ステップ S 0 2）。第一の音声解析エンジンと第二の音声解析エンジンとは、其々が、異なるアルゴリズム又はデータベースによるものであり、コンピュータ 10 は、一の音声データに対して、2つの音声認識を実行するものである。コンピュータ 10 は、例えば、スペクトラムアナライザ等により音声認識し、音声波形に基づいて、音声を認識する。コンピュータ 10 は、提供者が異なる音声解析エンジンや、異なるソフトウェアによる音声解析エンジンを用いて音声認識を実行する。コンピュータ 10 は、其々の音声認識の結果として、音声を其々の認識結果のテキストに変換する。

[0031] コンピュータ 10 は、第一の音声解析エンジンの認識結果と、第二の音声解析エンジンの認識結果とが異なる場合、双方の認識結果を、ユーザ端末に出力させる（ステップ S 0 3）。コンピュータ 10 は、双方の認識結果のテキストをユーザ端末に出力させる。ユーザ端末は、この双方の認識結果のテキストを、自身の表示部に表示又は音声により放音する。このとき、認識結果のテキストの一方には、認識結果が異なることをユーザに類推させるテキストが含まれる。

[0032] コンピュータ 10 は、ユーザ端末に出力させた双方の認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる（ステップ S 0 4）。コンピュータ 10 は、ユーザからのタップ操作や音声入力により、認識結果に対する正解の選択を受け付けさせる。例えば、コンピュータ 10 は、ユーザ端末に表示させたテキストの何れかに対する選択操作を受け付けさせることにより、認識結果に対する正解の選択を受け付けさせる。

[0033] コンピュータ 10 は、出力させた認識結果のうち、ユーザから正しい認識結果の選択を受け付けなかった音声解析エンジンに、この選択された正しい

認識結果を正解データとして、誤った音声認識を実行した音声解析エンジンに学習させる（ステップS05）。コンピュータ10は、第一の音声解析エンジンによる認識結果が正解データであった場合、第二の音声解析エンジンにこの正解データに基づいて学習させる。また、コンピュータ10は、第二の音声解析エンジンによる認識結果が正解データであった場合、第一の音声解析エンジンにこの正解データに基づいて学習させる。

[0034] なお、コンピュータ10は、2つの音声解析エンジンに限らず、三つ以上のN通りの音声解析エンジンにより音声認識を実行してもよい。このN通りの音声解析エンジンは、其々が異なるアルゴリズム又はデータベースによるものである。この場合、コンピュータ10は、取得した音声データを、N通りの音声解析エンジンにより音声認識する。コンピュータ10は、一の音声データに対してN通りの音声認識を実行するものである。コンピュータ10は、N通りの音声認識の結果として、音声将其の認識結果のテキストに変換する。

[0035] コンピュータ10は、N通りの音声解析エンジンの認識結果において、認識結果が異なるものを、ユーザ端末に出力させる。コンピュータ10は、認識結果が異なるテキストをユーザ端末に出力させる。ユーザ端末は、この異なる認識結果のテキストを、自身の表示部に表示又は音声により放音する。このとき、認識結果のテキストのうち、認識結果が異なることをユーザに類推するテキストが含まれる。

[0036] コンピュータ10は、ユーザ端末に出力した認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる。コンピュータ10は、ユーザからのタップ操作や音声入力により、認識結果に対する正解の選択を受け付けさせる。例えば、コンピュータ10は、ユーザ端末に表示させたテキストの何れかに対する選択操作を受け付けさせることにより、認識結果に対する正解の選択を受け付けさせる。

[0037] コンピュータ10は、出力させた認識結果のうち、ユーザから正しい認識結果の選択を受け付けなかった音声解析エンジンに、この選択された正しい

認識結果を正解データとして、誤った音声認識を実行した音声解析エンジンに学習させる。

[0038] 以上が、音声認識システム１の概要である。

[0039] [音声認識システム１のシステム構成]

図２に基づいて、本発明の好適な実施形態である音声認識システム１のシステム構成について説明する。図２は、本発明の好適な実施形態である音声認識システム１のシステム構成を示す図である。図２において、音声認識システム１は、コンピュータ１０から構成され、音声認識を実行するコンピュータシステムである。

[0040] なお、音声認識システム１は、図示していないユーザ端末等の他の端末類が含まれていてもよい。

[0041] コンピュータ１０は、上述した通り、図示していないユーザ端末等と公衆回線網等を介してデータ通信可能に接続されており、必要なデータの送受信を実行するとともに、音声認識を実行する。

[0042] コンピュータ１０は、CPU (Central Processing Unit)、RAM (Random Access Memory)、ROM (Read Only Memory)等を備え、通信部として、ユーザ端末や他のコンピュータ１０と通信可能にするためのデバイス、例えば、IEEE 802.11に準拠したWi-Fi (Wireless Fidelity) 対応デバイス等を備える。また、コンピュータ１０は、記録部として、ハードディスクや半導体メモリ、記録媒体、メモリカード等によるデータのストレージ部を備える。また、コンピュータ１０は、処理部として、各種処理を実行する各種デバイス等を備える。

[0043] コンピュータ１０において、制御部が所定のプログラムを読み込むことにより、通信部と協働して、音声取得モジュール２０、出力モジュール２１、選択受付モジュール２２、正解取得モジュール２３を実現する。また、コンピュータ１０において、制御部が所定のプログラムを読み込むことにより、処理部と協働して、音声認識モジュール４０、認識結果判定モジュール４１

を実現する。

[0044] [第一の音声認識処理]

図3に基づいて、音声認識システム1が実行する第一の音声認識処理について説明する。図3は、コンピュータ10が実行する第一の音声認識処理のフローチャートを示す図である。上述した各モジュールが実行する処理について、本処理に併せて説明する。

[0045] 音声取得モジュール20は、音声データを取得する（ステップS10）。ステップS10において、音声取得モジュール20は、ユーザ端末が入力を受け付けた音声を音声データとして取得する。ユーザ端末は、自身に内蔵された集音装置により、ユーザが発した音声を集音する。ユーザ端末は、この集音した音声を、音声データとしてコンピュータ10に送信する。音声取得モジュール20は、この音声データを受信することにより、音声データを取得する。

[0046] 音声認識モジュール40は、この音声データを、第一の音声解析エンジンにより、音声認識する（ステップS11）。ステップS11において、音声認識モジュール40は、スペクトラムアナライザ等による音波波形に基づいて、音声を認識する。音声認識モジュール40は、認識した音声を、テキスト変換する。このテキストを第一の認識テキストと称す。すなわち、第一の音声解析エンジンによる認識結果が、第一の認識テキストである。

[0047] 音声認識モジュール40は、この音声データを、第二の音声解析エンジンにより、音声認識する（ステップS12）。ステップS12において、音声認識モジュール40は、スペクトラムアナライザ等による音波波形に基づいて、音声を認識する。音声認識モジュール40は、認識した音声を、テキスト変換する。このテキストを、第二の認識テキストと称す。すなわち、第二の音声解析エンジンによる認識結果が、第二の認識テキストである。

[0048] 上述した第一の音声解析エンジンと第二の音声解析エンジンとは、其々が、異なるアルゴリズム又はデータベースによるものである。その結果、音声認識モジュール40は、一の音声データに基づいて、2つの音声認識を実行

することになる。この第一の音声解析エンジンと第二の音声解析エンジンとは、其々が、提供者が異なる音声解析エンジンや、異なるソフトウェアによる音声解析エンジンを用いて音声認識を実行する。

[0049] 認識結果判定モジュール41は、其々の認識結果が、一致するか否かを判定する（ステップS13）。ステップS13において、認識結果判定モジュール41は、第一の認識テキストと、第二の認識テキストとが一致するか否かを判定する。

[0050] ステップS13において、認識結果判定モジュール41は、一致すると判定した場合（ステップS13 YES）、出力モジュール21は、第一の認識テキストと第二の認識テキストとの何れか一方を、認識結果データとしてユーザ端末に出力させる（ステップS14）。ステップS14において、出力モジュール21は、其々の音声解析エンジンによる認識結果のうち、何れか一方のみの認識結果を、認識結果データとして出力させる。本例では、出力モジュール21は、第一の認識テキストを、認識結果データとして出力させたものとして説明する。

[0051] ユーザ端末は、この認識結果データを受信し、この認識結果データに基づいて、第一の認識テキストを、自身の表示部に表示する。あるいは、ユーザ端末は、この認識結果データに基づいて、第一の認識テキストに基づいた音声を自身のスピーカから出力する。

[0052] 選択受付モジュール22は、この第一の認識テキストが正しい認識結果であった場合又は誤った認識結果であった場合の選択を受け付けさせる（ステップS15）。ステップS15において、選択受付モジュール22は、ユーザ端末にユーザからのタップ操作や音声入力等の操作を受け付けさせることにより、正誤の認識結果の選択を受け付けさせる。正しい認識結果であった場合、正の認識結果の選択を受け付けさせる。また、誤った認識結果であった場合、誤の認識結果の選択を受け付けさせるとともに、タップ操作や音声入力等の操作を受け付けさせることにより、正の認識結果（正しいテキスト）の入力を受け付けさせる。

[0053] 図5は、ユーザ端末が認識結果データを自身の表示部に表示した状態を示す図である。図5において、ユーザ端末は、認識テキスト表示欄100、正解アイコン110、誤りアイコン120を表示する。認識テキスト表示欄100は、認識結果であるテキストを表示する。すなわち、認識テキスト表示欄100は、第一の認識テキスト「かえるのうたが きこえてくるよ」を表示する。

[0054] 選択受付モジュール22は、正解アイコン110又は誤りアイコン120への入力を受け付けさせることにより、この第一の認識テキストが正しい認識結果であるか又は誤った認識結果であるかの選択を受け付けさせる。選択受付モジュール22は、正しい認識結果であった場合、正の認識結果の操作として、ユーザに正解アイコン110への選択を受け付けさせ、誤った認識結果であった場合、誤の認識結果の操作として、ユーザに誤りアイコン120への選択を受け付けさせる。選択受付モジュール22は、誤りアイコン120への入力を受け付けさせた場合、さらに、正の認識結果として、正しいテキストの入力を受け付けさせる。

[0055] 正解取得モジュール23は、選択を受け付けさせた正誤の認識結果を、正解データとして取得する（ステップS16）。ステップS16において、正解取得モジュール23は、ユーザ端末が送信した正解データを受信することにより、正解データを取得する。

[0056] 音声認識モジュール40は、この正解データに基づいて、音声解析エンジンに、正誤の認識結果を学習させる（ステップS17）。ステップS17において、音声認識モジュール40は、正の認識結果を、正解データとして取得した場合、第一の音声解析エンジン及び第二の音声解析エンジンの其々に、今回の認識結果が正しいものであったことを学習させる。一方、音声認識モジュール40は、誤の認識結果を、正解データとして取得した場合、正の認識結果として受け付けさせた正しいテキストを、第一の音声解析エンジン及び第二の音声解析エンジンの其々に学習させる。

[0057] 一方、ステップS13において、認識結果判定モジュール41は、一致し

ないと判定した場合（ステップS 13 NO）、出力モジュール21は、第一の認識テキストと、第二の認識テキストとの双方を、認識結果データとしてユーザ端末に出力させる（ステップS 18）。ステップS 18において、出力モジュール21は、其々の音声解析エンジンによる認識結果の双方を、認識結果データとして出力させる。この認識結果データには、一方の認識テキストに、認識結果が異なっていることをユーザに類推させるテキスト（ひょっとして、もしかして等の可能性を認める表現）が含まれる。本例では、出力モジュール21は、第二の認識テキストにこの認識結果が異なっていることをユーザに類推させるテキストが含まれるものとして説明する。

[0058] ユーザ端末は、この認識結果データを受信し、この認識結果データに基づいて、第一の認識テキストと、第二の認識テキストとの双方を、自身の表示部に表示する。あるいは、ユーザ端末、この認識結果データに基づいて、第一の認識テキストと、第二の認識テキストとに基づいた音声を自身のスピーカから出力する。

[0059] 選択受付モジュール22は、ユーザ端末に出力させた認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる（ステップS 19）。ステップS 19において、選択受付モジュール22は、ユーザ端末にタップ操作や音声入力等の操作を受け付けさせることにより、何れの認識テキストが正しい認識結果であるかの選択を受け付けさせる。認識テキストのうち、正しい認識結果のものに、正の認識結果の選択（例えば、この認識テキストをタップ入力、この認識テキストを音声入力）を受け付けさせる。

[0060] なお、選択受付モジュール22は、何れの認識テキストも正しい認識結果ではない場合、誤の認識結果の選択を受け付けさせるとともに、タップ操作や音声入力等の選択を受け付けさせることにより、正の認識結果（正しいテキスト）の入力を受け付けさせてもよい。

[0061] 図6は、ユーザ端末が認識結果データを自身の表示部に表示した状態を示す図である。図6において、ユーザ端末は、第一の認識テキスト表示欄200、第二の認識テキスト表示欄210、誤りアイコン220を表示する。第

一の認識テキスト表示欄 200 は、第一の認識テキストを表示する。第二の認識テキスト表示欄 210 は、第二の認識テキストを表示する。この第二の認識テキストには、上述した第一の認識テキストと認識結果が異なっていることをユーザに類推させるテキストが含まれる。すなわち、第一の認識テキスト表示欄 200 は、第一の認識テキスト「かえるのうたぎ 超えてくるよ」を表示する。また、第二の認識テキスト表示欄 210 は、「※ひょっとして かえるのうたが きこえてくるよ」を表示する。

[0062] 選択受付モジュール 22 は、第一の認識テキスト表示欄 200 又は第二の認識テキスト表示欄 210 の何れかへの入力を受け付けさせることにより、この第一の認識テキスト又は第二の認識テキストの何れが正しい認識結果あるかの選択を受け付けさせる。選択受付モジュール 22 は、第一の認識テキストが正しい認識結果であった場合、正の認識結果の操作として、第一の認識テキスト表示欄 200 へのタップ操作や音声による選択を受け付けさせる。また、選択受付モジュール 22 は、第二の認識テキストが正しい認識結果であった場合、正の認識結果の操作として、第二の認識テキスト表示欄 210 へのタップ操作や音声による選択を受け付けさせる。また、選択受付モジュール 22 は、第一の認識テキスト及び第二の認識テキストの何れの認識テキストも正しい認識結果でなかった場合、誤の認識結果の選択として、誤りアイコン 220 への選択を受け付けさせる。選択受付モジュール 22 は、誤りアイコン 220 への選択を受け付けさせた場合、さらに、正の認識結果として、正しいテキストの入力を受け付けさせる。

[0063] 正解取得モジュール 23 は、選択を受け付けさせた正しい認識結果を、正解データとして取得する（ステップ S20）。ステップ S20 において、正解取得モジュール 23 は、ユーザ端末が送信した正解データを、受信することにより、正解データを取得する。

[0064] 音声認識モジュール 40 は、この正解データに基づいて、正しい認識結果の選択を受け付けなかった音声解析エンジンに、この選択された正しい認識結果を学習させる（ステップ S21）。ステップ S21 において、音声認識

モジュール40は、正解データが、第一の認識テキストであった場合、正しい認識結果である第一の認識テキストを、第二の音声解析エンジンに学習させるとともに、第一の音声解析エンジンに、今回の認識結果が正しいものであったことを学習させる。また、音声認識モジュール40は、正解データが、第二の認識テキストであった場合、正しい認識結果である第二の認識テキストを、正解データとして、第一の音声解析エンジンに学習させるとともに、第二の音声解析エンジンに、今回の認識結果が正しいものであったことを学習させる。また、音声認識モジュール40は、正解データが、第一の認識テキスト及び第二の認識テキストの何れでもない場合、正の認識結果として受け付けさせた正しいテキストを、第一の音声解析エンジン及び第二の音声解析エンジンに学習させる。

[0065] 音声認識モジュール23は、次回以降の音声認識に際して、学習させた結果を加味した第一の音声解析エンジン及び第二の音声解析エンジンを用いる。

[0066] 以上が、第一の音声認識処理である。

[0067] [第二の音声認識処理]

図4に基づいて、音声認識システム1が実行する第二の音声認識処理について説明する。図4は、コンピュータ10が実行する第二の音声認識処理のフローチャートを示す図である。上述した各モジュールが実行する処理について、本処理に併せて説明する。

[0068] なお、上述した第一の音声認識処理と同様の処理については、その詳細な説明を省略する。また、第一の音声認識処理と、第二の音声処理とは、音声認識モジュール40が用いる音声解析エンジンの総数が異なっている。

[0069] 音声取得モジュール20は、音声データを取得する（ステップS30）。ステップS30の処理は、上述したステップS10の処理と同様である。

[0070] 音声認識モジュール40は、この音声データを、第一の音声解析エンジンにより、音声認識する（ステップS31）。ステップS31の処理は、上述したステップS11の処理と同様である。

- [0071] 音声認識モジュール40は、この音声データを、第二の音声解析エンジンにより、音声認識する（ステップS32）。ステップS32の処理は、上述したステップS12の処理と同様である。
- [0072] 音声認識モジュール40は、この音声データを、第三の音声解析エンジンにより、音声認識する（ステップS33）。ステップS33において、音声認識モジュール40は、スペクトラムアナライザ等による音波波形に基づいて、音声进行認識する。音声認識モジュール40は、認識した音声を、テキスト変換する。このテキストを、第三の認識テキストと称す。すなわち、第三の音声解析エンジンによる認識結果が、第三の認識テキストである。
- [0073] 上述した第一の音声解析エンジンと、第二の音声解析エンジンと、第三の音声解析エンジンとは、其々が、異なるアルゴリズム又はデータベースによるものである。その結果、音声認識モジュール40は、一の音声データに基づいて、三通りの音声認識を実行することになる。この第一の音声解析エンジンと、第二の音声解析エンジンと、第三の音声解析エンジンとは、其々が、提供者が異なる音声解析エンジンや、異なるソフトウェアによる音声解析エンジンを用いて音声認識を実行する。
- [0074] なお、上述した処理は、三通りの音声解析エンジンにおり音声認識を実行するものであるが、音声解析エンジンの数は、三通り以上のN通りのものであってもよい。この場合、N通りの音声解析の其々は、異なるアルゴリズム又はデータベースで音声認識を行うものである。N通りの音声解析エンジンを用いる場合、後述する処理において、N通りの認識テキストにおいて、後述する処理を実行することになる。
- [0075] 認識結果判定モジュール41は、其々の認識結果が、一致するか否かを判定する（ステップS34）。ステップS34において、認識結果判定モジュール41は、第一の認識テキストと、第二の認識テキストと、第三の認識テキストとが一致するか否かを判定する。
- [0076] ステップS34において、認識結果判定モジュール41は、一致すると判定した場合（ステップS34 YES）、出力モジュール21は、第一の認

識テキスト、第二の認識テキスト又は第三の認識テキストの何れかを、認識結果データとしてユーザ端末に出力させる（ステップS 3 5）。ステップS 3 5の処理は、上述したステップS 1 4の処理と略同様であり、相違点は、第三の認識テキストが含まれる点である。本例では、出力モジュール2 1は、第一の認識テキストを、認識結果データとして出力させたものとして説明する。

[0077] ユーザ端末は、この認識結果データを受信し、この認識結果データに基づいて、第一の認識テキストを、自身の表示部に表示する。あるいは、ユーザ端末は、この認識結果データに基づいて、第一の認識テキストに基づいた音声を自身のスピーカから出力する。

[0078] 選択受付モジュール2 2は、この第一の認識テキストが正しい認識結果であった場合又は誤った認識結果であった場合の選択を受け付けさせる（ステップS 3 6）。ステップS 3 6の処理は、上述したステップS 1 5の処理と同様である。

[0079] 正解取得モジュール2 3は、選択を受け付けさせた正誤の認識結果を、正解データとして取得する（ステップS 3 7）。ステップS 3 7の処理は、上述したステップS 1 6の処理と同様である。

[0080] 音声認識モジュール4 0は、この正解データに基づいて、音声解析エンジンに、正誤の認識結果を学習させる（ステップS 3 8）。ステップS 3 8において、音声認識モジュール4 0は、正の認識結果を、正解データとして取得した場合、第一の音声解析エンジン、第二の音声解析エンジン及び第三の音声解析エンジンの其々に、今回の認識結果が正しいものであったことを学習させる。一方、音声認識モジュール4 0は、誤の認識結果を、正解データとして取得した場合、正しい認識結果として受け付けさせた正しいテキストを、第一の音声解析エンジン、第二の音声解析エンジン及び第三の音声解析エンジンの其々に学習させる。

[0081] 一方、ステップS 3 4において、認識結果判定モジュール4 1は、一致しないと判定した場合（ステップS 3 4 NO）、出力モジュール2 1は、第

一の認識テキスト、第二の認識テキスト又は第三の認識テキストのうち、認識結果が異なるもののみを、認識結果データとしてユーザ端末に出力させる（ステップS39）。ステップS39において、出力モジュール21は、其々の音声解析エンジンによる認識結果のうち、認識結果が異なるものを、認識結果データとして出力させる。また、この認識結果データには、認識結果が異なっていることをユーザに類推させるテキストが含まれる。

[0082] 例えば、出力モジュール21は、第一の認識テキストと、第二の認識テキストと、第三の認識テキストとが其々異なる場合、これら三つの認識テキストを認識結果データとしてユーザ端末に出力させる。このとき、第二の認識テキスト及び第三の認識テキストには、認識結果が異なっていることをユーザに類推させるテキストが含まれる。

[0083] また、例えば、出力モジュール21は、第一の認識テキストと、第二の認識テキストとが同一で、第三の認識テキストが異なる場合、第一の認識テキストと、第三の認識テキストとを認識結果データとしてユーザ端末に出力させる。このとき、第三の認識テキストには、認識結果が異なっていることをユーザに類推させるテキストが含まれる。また、出力モジュール21は、第一の認識テキストと、第三の認識テキストとが同一で、第二の認識テキストが異なる場合、第一の認識テキストと、第二の認識テキストとを認識結果データとしてユーザ端末に出力させる。このとき、第二の認識テキストには、認識結果が異なっていることをユーザに類推させるテキストが含まれる。また、出力モジュール21は、第二の認識テキストと、第三の認識テキストとが同一で、第一の認識テキストが異なる場合、第一の認識テキストと、第二の認識テキストとを認識結果データとしてユーザ端末に出力させる。このとき、第二の認識テキストには、認識結果が異なっていることをユーザに類推させるテキストが含まれる。このように、認識結果データにおいて、認識テキストの一致率（複数の音声解析エンジンによる認識結果のうち、一致する認識結果の割合）が最も高いものをそのままの認識テキストとして出力させ、それ以外のものに認識結果が異なっていることをユーザに類推させるテキ

ストを含めて出力させる。これは、音声解析エンジンの数が、4つ以上であっても同様である。

[0084] 本例では、出力モジュール21は、全ての認識テキストが異なっている場合と、第一の認識テキストと、第二の認識テキストとが同一で、第三の認識テキストが異なる場合とを例として説明する。

[0085] ユーザ端末は、この認識結果データを受信し、この認識結果データに基づいて、第一の認識テキストと、第二の認識テキストと、第三の認識テキストとの其々を、自身の表示部に表示する。あるいは、ユーザ端末は、この認識結果データに基づいて、第一の認識テキストと、第二の認識テキストと、第三の認識テキストとの其々に基づいた音声を自身のスピーカから出力する。

[0086] また、ユーザ端末は、この認識結果データを受信し、この認識結果データに基づいて、第一の認識テキストと、第三の認識テキストとを、自身の表示部に表示する。あるいは、ユーザ端末は、この認識結果データに基づいて、第一の認識テキストと、第三の認識テキストとの其々に基づいた音声を自身のスピーカから出力する。

[0087] 選択受付モジュール22は、ユーザ端末に出力させた認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる（ステップS40）。ステップS40の処理は、上述したステップS19の処理と同様である。

[0088] ユーザ端末が第一の認識テキストと、第二の認識テキストと、第三の認識テキストとの其々を、自身の表示部に表示する例について説明する。

[0089] 図7は、ユーザ端末が認識結果データを自身の表示部に表示した状態を示す図である。図7において、ユーザ端末は、第一の認識テキスト表示欄300、第二の認識テキスト表示欄310、第三の認識テキスト表示欄312、誤りアイコン330を表示する。第一の認識テキスト表示欄300は、第一の認識テキストを表示する。第二の認識テキスト表示欄310は、第二の認識テキストを表示する。この第二の認識テキストには、上述した第一の認識テキスト及び第三の認識テキストと認識結果が異なっていることをユーザに類推させるテキストが含まれる。第三の認識テキスト表示欄320は、第三

の認識テキストを表示する。この第三の認識テキストには、上述した第一の認識テキスト及び第二の認識テキストと認識結果が異なっていることをユーザに類推させるテキストが含まれる。すなわち、第一の認識テキスト表示欄 300 は、第一の認識テキスト「かえるのうたぎ 超えてくるよ」を表示する。また、第二の認識テキスト表示欄 310 は、「※ひょっとして かえるのうたが きこえてくるよ」を表示する。また、第三の認識テキスト 320 は、「※ひょっとして かえるのぶたが こえてくるよ」を表示する。

[0090] 選択受付モジュール 22 は、第一の認識テキスト表示欄 300、第二の認識テキスト表示欄 310 又は第三の認識テキスト表示欄 320 の何れかの選択を受け付けさせることにより、この第一の認識テキスト、第二の認識テキスト又は第三の認識テキストの何れが正しい認識結果あるかの選択を受け付けさせる。選択受付モジュール 22 は、第一の認識テキストが正しい認識結果であった場合、正の認識結果の操作として、第一の認識テキスト表示欄 300 へのタップ操作や音声による選択を受け付けさせる。また、選択受付モジュール 22 は、第二の認識テキストが正しい認識結果であった場合、正の認識結果の操作として、第二の認識テキスト表示欄 310 へのタップ操作や音声による選択を受け付けさせる。また、選択受付モジュール 22 は、第三の認識テキストが正しい認識結果であった場合、正の認識結果の操作として、第三の認識テキスト表示欄 320 へのタップ操作や音声による選択を受け付けさせる。また、選択受付モジュール 22 は、第一の認識テキスト、第二の認識テキスト及び第三の認識テキストの何れの認識テキストも正しい認識結果でなかった場合、誤の認識結果の操作として、誤りアイコン 330 への選択を受け付けさせる。選択受付モジュール 22 は、誤りアイコン 330 への選択を受け付けさせた場合、さらに、正の認識結果として、正しいテキストの入力を受け付けさせる。

[0091] ユーザ端末が第一の認識テキストと、第三の認識テキストとの其々を、自身の表示部に表示する例については、上述した図 6 のものと同様であるため、説明は省略するが、相違点としては、第二の認識テキスト表示欄 210 に

、第三の認識テキストを表示することになる。

[0092] 正解取得モジュール23は、選択を受け付けさせた正しい認識結果を、正解データとして取得する（ステップS41）。ステップS41の処理は、上述したステップS20の処理と同様である。

[0093] 音声認識モジュール40は、この正解データに基づいて、正しい認識結果の選択を受け付けなかった音声解析エンジンに、この選択された正しい認識結果を学習させる（ステップS42）。ステップS42において、音声認識モジュール40は、正解データが、第一の認識テキストであった場合、正しい認識結果である第一の認識テキストを、第二の音声解析エンジン及び第三の音声解析エンジンに学習させるとともに、第一の音声解析エンジンに、今回の認識結果が正しいものであったことを学習させる。また、音声認識モジュール40は、正解データが、第二の認識テキストであった場合、正しい認識結果である第二の認識テキストを、正解データとして、第一の音声解析エンジン及び第三の音声解析エンジンに学習させるとともに、第二の音声解析エンジンに、今回の認識結果が正しいものであったことを学習させる。また、音声認識モジュール40は、正解データが、第三の認識テキストであった場合、正しい認識結果である第三の認識テキストを、正解データとして、第一の音声解析エンジン及び第二の音声解析エンジンに学習させるとともに、第三の音声解析エンジンに、今回の認識結果が正しいものであったことを学習させる。また、音声認識モジュール40は、正解データが、第一の認識テキスト、第二の認識テキスト及び第三の認識テキストの何れでもない場合、正の認識結果として受け付けさせた正しいテキストを、第一の音声解析エンジン、第二の音声解析エンジン及び第三の音声解析エンジンに学習させる。

[0094] 以上が、第二の音声認識処理である。

[0095] なお、音声認識システム1は、三通りの音声解析エンジンで行った処理と同様の処理を、N通りの音声解析エンジンで行ってもよい。すなわち、音声認識システム1は、N通りで行った音声認識のうち、音声認識結果が異なるもののみを出力させ、この出力させた認識結果のうち、ユーザから正しい音

声認識の選択を受け付けさせる。音声認識システム 1 は、正しい音声認識として選択されなかった場合に、選択された正しい音声認識結果に基づいて学習する。

[0096] 上述した手段、機能は、コンピュータ（CPU、情報処理装置、各種端末を含む）が、所定のプログラムを読み込んで、実行することによって実現される。プログラムは、例えば、コンピュータからネットワーク経由で提供される（SaaS：ソフトウェア・アズ・ア・サービス）形態で提供される。また、プログラムは、例えば、フレキシブルディスク、CD（CD-ROM など）、DVD（DVD-ROM、DVD-RAM など）等のコンピュータ読取可能な記録媒体に記録された形態で提供される。この場合、コンピュータはその記録媒体からプログラムを読み取って内部記録装置又は外部記録装置に転送し記録して実行する。また、そのプログラムを、例えば、磁気ディスク、光ディスク、光磁気ディスク等の記録装置（記録媒体）に予め記録しておき、その記録装置から通信回線を介してコンピュータに提供するようにしてもよい。

[0097] 以上、本発明の実施形態について説明したが、本発明は上述したこれらの実施形態に限るものではない。また、本発明の実施形態に記載された効果は、本発明から生じる最も好適な効果を列挙したに過ぎず、本発明による効果は、本発明の実施形態に記載されたものに限定されるものではない。

符号の説明

[0098] 1 音声認識システム、10 コンピュータ

請求の範囲

- [請求項1] 音声データを取得する取得手段と、
取得した前記音声データの音声認識を行う第一認識手段と、
取得した前記音声データの音声認識を、前記第一認識手段とは異なるアルゴリズム又はデータベースで行う第二認識手段と、
其々の音声認識の認識結果が異なる場合、双方の認識結果を出力させる出力手段と、
を備えることを特徴とするコンピュータシステム。
- [請求項2] 出力させた前記双方の認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる選択手段と、
をさらに備え、
前記第一認識手段又は前記第二認識手段は、前記正しい認識結果として選択されなかった場合、選択された正しい認識結果に基づいて学習する、
ことを特徴とする請求項1に記載のコンピュータシステム。
- [請求項3] 音声データを取得する取得手段と、
取得した前記音声データの音声認識を行い、互いに異なるアルゴリズム又はデータベースでN通りの音声認識を行うN通りの認識手段と、
、
前記N通りで行った音声認識のうち、認識結果が異なるもののみを出力させる出力手段と、
を備えることを特徴とするコンピュータシステム。
- [請求項4] 出力させた前記認識結果のうち、ユーザから正しい認識結果の選択を受け付けさせる選択手段と、
をさらに備え、
前記N通りの認識手段は、前記正しい認識結果として選択されなかった場合、選択された正しい認識結果に基づいて学習する、
ことを特徴とする請求項3に記載のコンピュータシステム。

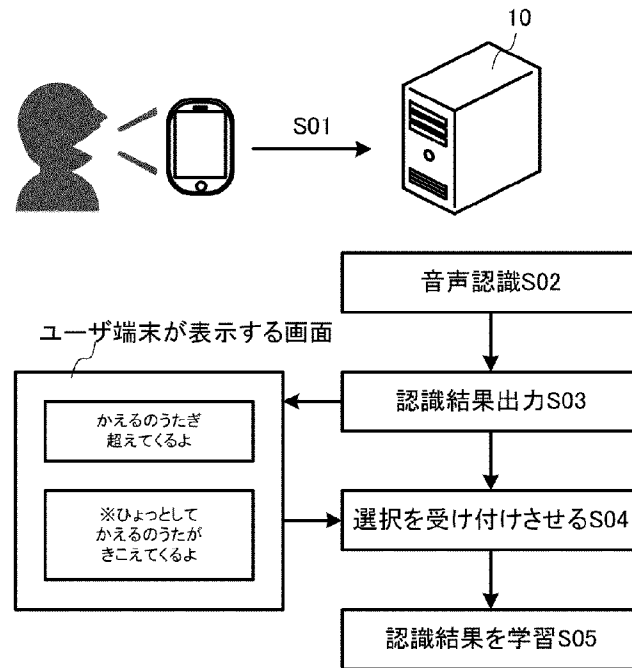
- [請求項5] コンピュータシステムが実行する音声認識方法であって、
 音声データを取得する取得ステップと、
 取得した前記音声データの音声認識を行う第一認識ステップと、
 取得した前記音声データの音声認識を、前記第一認識ステップとは異なるアルゴリズム又はデータベースで行う第二認識ステップと、
 其々の音声認識の認識結果が異なる場合、双方の認識結果を出力させる出力ステップと、
 を備えることを特徴とする音声認識方法。
- [請求項6] コンピュータシステムが実行する音声認識方法であって、
 音声データを取得する取得ステップと、
 取得した前記音声データの音声認識を行い、互いに異なるアルゴリズム又はデータベースでN通りの音声認識を行うN通りの認識ステップと、
 前記N通りで行った音声認識のうち、認識結果が異なるもののみを出力させる出力ステップと、
 を備えることを特徴とする音声認識方法。
- [請求項7] コンピュータシステムに、
 音声データを取得する取得ステップ、
 取得した前記音声データの音声認識を行う第一認識ステップ、
 取得した前記音声データの音声認識を、前記第一認識ステップとは異なるアルゴリズム又はデータベースで行う第二認識ステップ、
 其々の音声認識の認識結果が異なる場合、双方の認識結果を出力させる出力ステップ、
 を実行させるためのコンピュータ読み取り可能なプログラム。
- [請求項8] コンピュータシステムに、
 音声データを取得する取得ステップ、
 取得した前記音声データの音声認識を行い、互いに異なるアルゴリズム又はデータベースでN通りの音声認識を行うN通りの認識ステップ

プ、

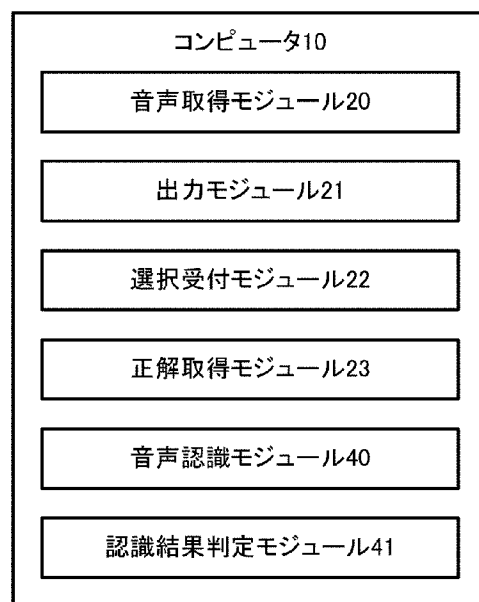
前記N通りで行った音声認識のうち、認識結果が異なるもののみを
出力させる出力ステップ、

実行させるためのコンピュータ読み取り可能なプログラム。

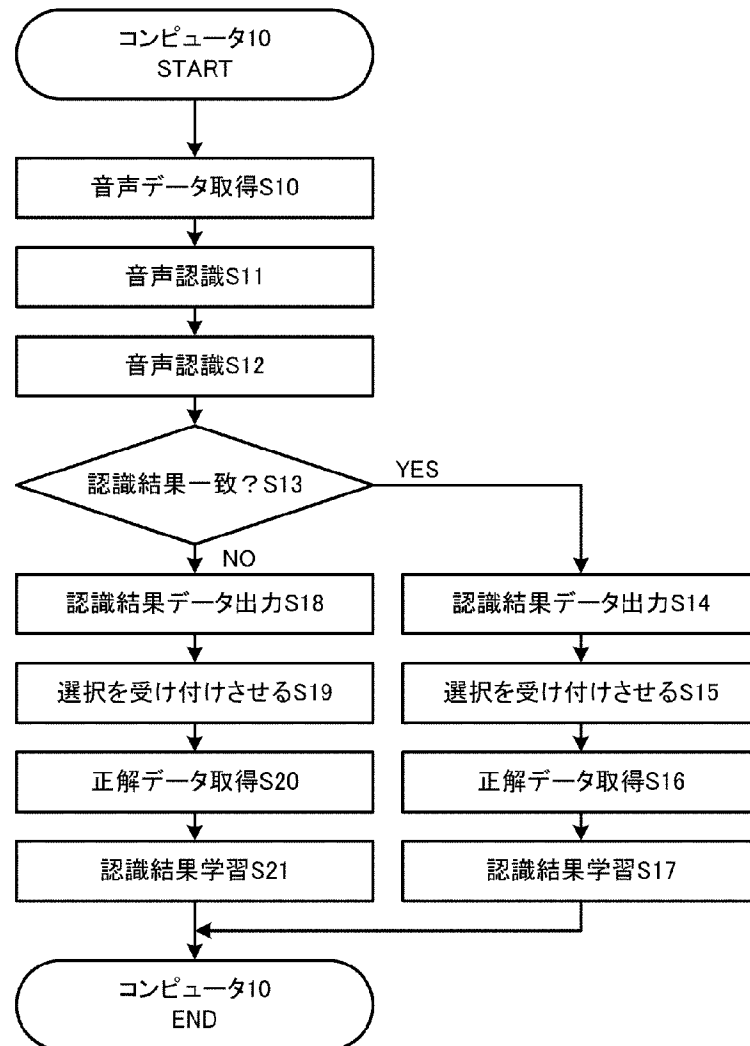
[図1]



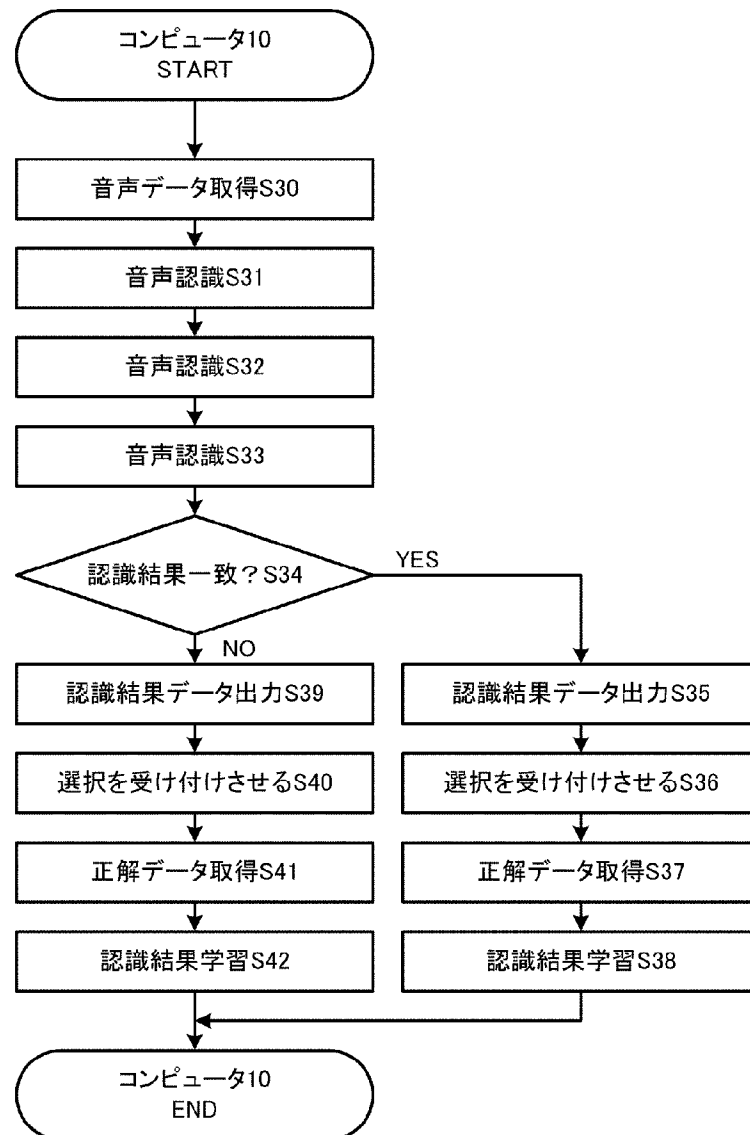
[図2]



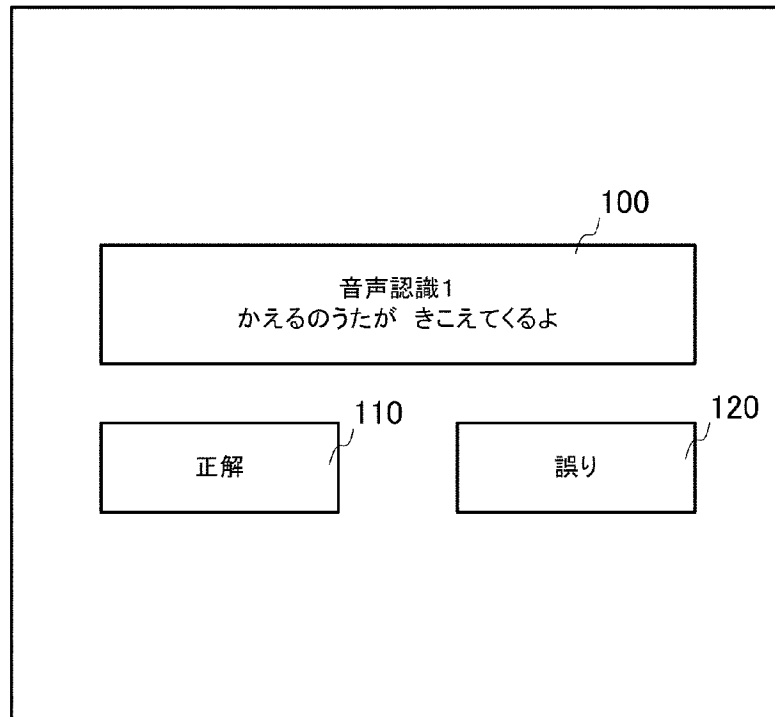
[図3]



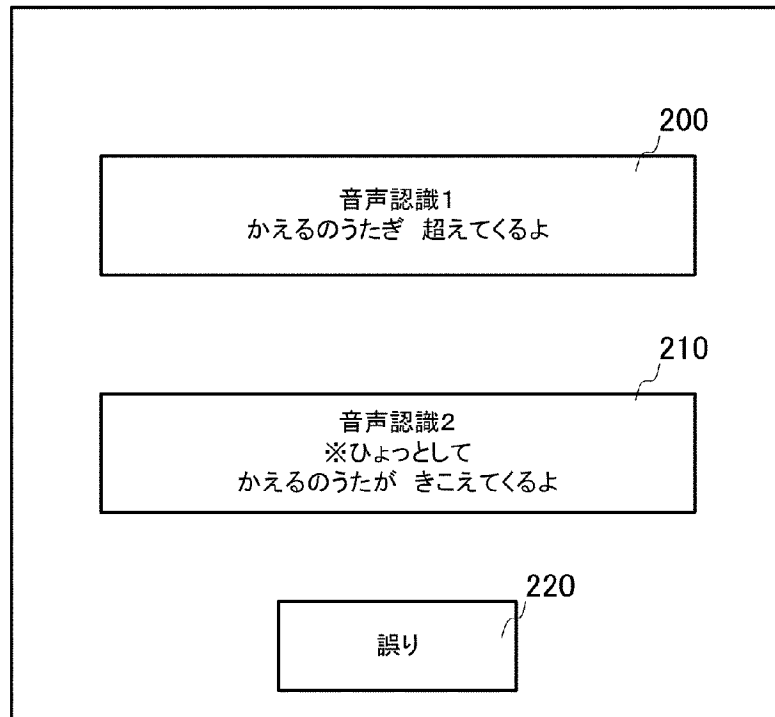
[図4]



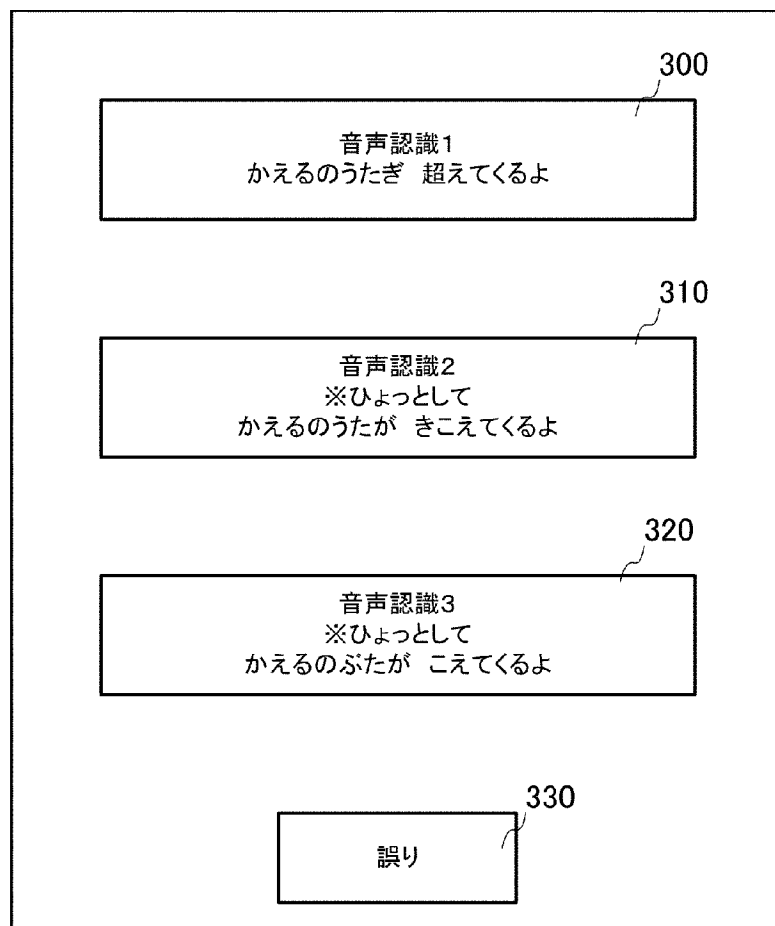
[図5]



[図6]



[図7]



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/036001

A. CLASSIFICATION OF SUBJECT MATTER

Int. Cl. G10L15/32 (2013.01) i, G10L15/22 (2006.01) i

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

Int. Cl. G10L15/00-15/34

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Published examined utility model applications of Japan 1922-1996

Published unexamined utility model applications of Japan 1971-2018

Registered utility model specifications of Japan 1996-2018

Published registered utility model applications of Japan 1994-2018

Electronic data base consulted during the international search (name of data base and, where practicable, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 2010-85536 A (FUETREK CO., LTD.) 15 April 2010, paragraphs [0034]-[0073] (Family: none)	1-3, 5-8 4
A	JP 11-154231 A (TOSHIBA CORP.) 08 June 1999, paragraphs [0001]-[0013], [0036]-[0069] (Family: none)	1-8
A	JP 2009-265307 A (TOYOTA MOTOR CORP.) 12 November 2009, paragraphs [0013]-[0030] (Family: none)	1-8
A	WO 2013/005248 A1 (MITSUBISHI ELECTRIC CORP.) 10 January 2013, paragraphs [0010]-[0057] & US 2014/0100847 A1, paragraphs [0020]-[0068] & CN 103650034 A	1-8



Further documents are listed in the continuation of Box C.



See patent family annex.

* Special categories of cited documents:

“A” document defining the general state of the art which is not considered to be of particular relevance

“E” earlier application or patent but published on or after the international filing date

“L” document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

“O” document referring to an oral disclosure, use, exhibition or other means

“P” document published prior to the international filing date but later than the priority date claimed

“T” later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

“X” document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

“Y” document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art

“&” document member of the same patent family

Date of the actual completion of the international search
07.12.2018

Date of mailing of the international search report
18.12.2018

Name and mailing address of the ISA/
Japan Patent Office
3-4-3, Kasumigaseki, Chiyoda-ku,
Tokyo 100-8915, Japan

Authorized officer

Telephone No.

INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2018/036001

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	JP 2002-116796 A (CANON INC.) 19 April 2002, paragraphs [0008]-[0053] & US 2002/0055845 A1, paragraphs [0011]-[0056]	1-8

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G10L15/32(2013.01)i, G10L15/22(2006.01)i

B. 調査を行った分野

調査を行った最小限資料（国際特許分類（IPC））

Int.Cl. G10L15/00-15/34

最小限資料以外の資料で調査を行った分野に含まれるもの

日本国実用新案公報	1922-1996年
日本国公開実用新案公報	1971-2018年
日本国実用新案登録公報	1996-2018年
日本国登録実用新案公報	1994-2018年

国際調査で使用した電子データベース（データベースの名称、調査に使用した用語）

C. 関連すると認められる文献

引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
X A	JP 2010-85536 A（株式会社フュートレック）2010.04.15, [0034]-[0073]（ファミリーなし）	1-3, 5-8 4
A	JP 11-154231 A（株式会社東芝）1999.06.08, [0001]-[0013], [0036]-[0069]（ファミリーなし）	1-8
A	JP 2009-265307 A（トヨタ自動車株式会社）2009.11.12, [0013]-[0030]（ファミリーなし）	1-8

☒ C欄の続きにも文献が列挙されている。

☐ パテントファミリーに関する別紙を参照。

* 引用文献のカテゴリー

「A」特に関連のある文献ではなく、一般的技術水準を示すもの
「E」国際出願日前の出願または特許であるが、国際出願日以後に公表されたもの
「L」優先権主張に疑義を提起する文献又は他の文献の発行日若しくは他の特別な理由を確立するために引用する文献（理由を付す）
「O」口頭による開示、使用、展示等に言及する文献
「P」国際出願日前で、かつ優先権の主張の基礎となる出願

の日の後に公表された文献

「T」国際出願日又は優先日後に公表された文献であって出願と矛盾するものではなく、発明の原理又は理論の理解のために引用するもの
「X」特に関連のある文献であって、当該文献のみで発明の新規性又は進歩性がないと考えられるもの
「Y」特に関連のある文献であって、当該文献と他の1以上の文献との、当業者にとって自明である組合せによって進歩性がないと考えられるもの
「&」同一パテントファミリー文献

国際調査を完了した日

07.12.2018

国際調査報告の発送日

18.12.2018

国際調査機関の名称及びあて先

日本国特許庁（ISA/J P）

郵便番号100-8915

東京都千代田区霞が関三丁目4番3号

特許庁審査官（権限のある職員）

菊池 智紀

5Z

3352

電話番号 03-3581-1101 内線 3591

C (続き) . 関連すると認められる文献		
引用文献の カテゴリー*	引用文献名 及び一部の箇所が関連するときは、その関連する箇所の表示	関連する 請求項の番号
A	WO 2013/005248 A1 (三菱電機株式会社) 2013. 01. 10, [0010]-[0057] & US 2014/0100847 A1, [0020]-[0068] & CN 103650034 A	1-8
A	JP 2002-116796 A (キヤノン株式会社) 2002. 04. 19, [0008]-[0053] & US 2002/0055845 A1, [0011]-[0056]	1-8