

(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
2016 年 12 月 1 日(01.12.2016)



(10) 国際公開番号
WO 2016/190060 A1

- (51) 国際特許分類:
G06F 3/16 (2006.01) *G10L 15/22* (2006.01)
G06F 3/01 (2006.01) *G10L 15/24* (2013.01)
- (21) 国際出願番号: PCT/JP2016/063631
- (22) 国際出願日: 2016 年 5 月 6 日(06.05.2016)
- (25) 国際出願の言語: 日本語
- (26) 国際公開の言語: 日本語
- (30) 優先権データ:
特願 2015-104706 2015 年 5 月 22 日(22.05.2015) JP
- (71) 出願人: ソニー株式会社(SONY CORPORATION)
[JP/JP]; 〒1080075 東京都港区港南 1 丁目 7 番 1 号 Tokyo (JP).
- (72) 発明者: 小川 浩明(OGAWA Hiroaki); 〒1080075 東京都港区港南 1 丁目 7 番 1 号 ソニー株式会社内 Tokyo (JP).
- (74) 代理人: 西川 孝, 外(NISHIKAWA Takashi et al.); 〒1600023 東京都新宿区西新宿 7 丁目 5 番 2 5 号 西新宿木村屋ビルディング 9 階 Tokyo (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, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IR, IS, KE, KG, KN, KP, KR, 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, 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).

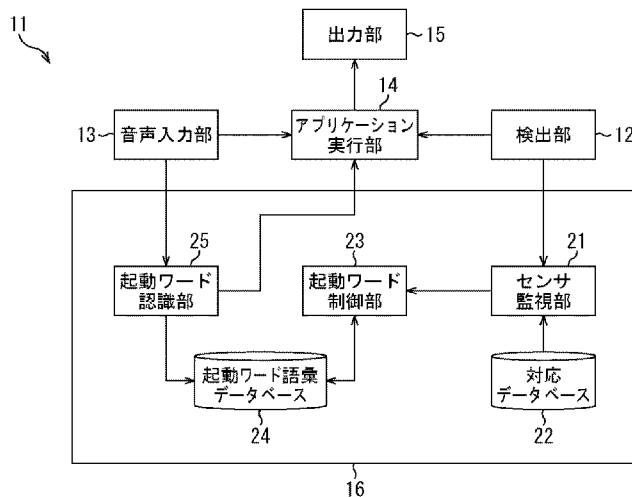
添付公開書類:

— 国際調査報告 (条約第 21 条(3))

(54) Title: INFORMATION PROCESSING DEVICE, INFORMATION PROCESSING METHOD, AND PROGRAM

(54) 発明の名称: 情報処理装置および情報処理方法、並びにプログラム

FIG.1



- 12 Detecting unit
13 Speech input unit
14 Application executing unit
15 Output unit
21 Sensor monitoring unit
22 Correspondence database
23 Activation word control unit
24 Activation word vocabulary database
25 Activation word recognizing unit

(57) Abstract: The present disclosure pertains to an information processing device, an information processing method, and a program which enable provision of an excellent user experience. This information processing device is provided with an activation word setting unit that sets a vocabulary utilized as activation words so as to, on the basis of a detection result of the detection of an action of the user, increase or decrease the number of activation words vocalized by a user in order to activate a prescribed function, and an activation word recognizing unit which performs speech recognition on the speech vocalized by the user, and which recognizes vocalization of the vocabulary set to be utilized as an activation word by the activation word setting unit. The present technology is applicable to, for example, a wearable terminal provided with a speech recognition function.

(57) 要約: 本開示は、より良いユーザ体験を提供することができるようにする情報処理装置および情報処理方法、並びにプログラムに関する。情報処理装置は、ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、起動ワードとして用いる語彙を設定する起動ワード設定部と、ユーザが発話する音声に対する音声認識を行って、起動ワード設定部により起動ワードとして用いるように設定されている語彙が発話されたことを認識する起動ワード認識部とを備える。本技術は、例えば、音声認識機能を備えたウェアラブル端末に適用できる。

明 細 書

発明の名称：

情報処理装置および情報処理方法、並びにプログラム

技術分野

[0001] 本開示は、情報処理装置および情報処理方法、並びにプログラムに関し、特に、より良いユーザ体験を提供することができるようにした情報処理装置および情報処理方法、並びにプログラムに関する。

背景技術

[0002] 近年、ユーザが身に着けて持ち歩くことができる小型デバイスとして、例えば、眼鏡型や時計型などのウェアラブル端末が開発されており、多くのウェアラブル端末には、ユーザが発話する音声を認識する音声認識機能が搭載されている。このようなウェアラブル端末における音声認識機能を、常時、利用可能とすることによって、ユーザの利便性を向上させることができる。しかしながら、ウェアラブル端末が、ユーザの独り言や周囲のノイズなどに反応して音声認識することも想定され、音声認識機能を常に利用可能とした場合には、誤作動が発生することが懸念される。

[0003] そのため、通常時には音声認識機能を停止させておき、音声認識機能を起動させるための起動ボタンをウェアラブル端末に設けることが検討される。しかしながら、小型のウェアラブル端末では、起動ボタンを設けることが困難であったり、起動ボタンに対する良好な操作性を提供することが困難であったりすることが想定される。これに対し、ユーザの利便性を考慮して、所定の起動ワードをウェアラブル端末に設定しておき、その起動ワードが認識されたときに一定期間だけ、音声認識機能を起動する技術が開発されている。

[0004] 一方、ユーザが発話する音声だけでなく、ユーザのジェスチャも組み合わせた解析を行うことによって、より複雑で具体的な入力情報に対応することができる技術が開発されている。

[0005] 例えば、特許文献1には、発話音声を音声認識した音声認識結果と、画面に対するジェスチャの軌道とを、発話音声およびジェスチャが入力された時間関係に基づいて対応付けることにより、ジェスチャによって指示された指示物の情報を出力することができる情報処理装置が開示されている。

先行技術文献

特許文献

[0006] 特許文献1：特開平9－114634号公報

発明の概要

発明が解決しようとする課題

[0007] ところで、従来のウェアラブル端末において、音声認識を利用したアプリケーションを実行しているときに、音声認識機能を起動させるための起動ワードを発話することによって、アプリケーションにより提供されるユーザ体験を損ねてしまうことがある。例えば、アプリケーションが特定の世界観を提供するゲームである場合に、その世界観と無関係の語彙が起動ワードとして設定されていると、その語彙を発話することによって、アプリケーションが提供する世界観からユーザを引き離すような感覚を与えてしまう。このようなユーザ体験を損ねるようなことを回避して、より良いユーザ体験を提供することが求められている。

[0008] 本開示は、このような状況に鑑みてなされたものであり、より良いユーザ体験を提供することができるようにするものである。

課題を解決するための手段

[0009] 本開示の一側面の情報処理装置は、ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定する起動ワード設定部と、前記ユーザの発話する音声に対する音声認識を行って、前記起動ワード設定部により前記起動ワードとして用いるように設定されている語彙が発話されたことを認識する起動ワード認識部とを備える。

[0010] 本開示の一側面の情報処理方法またはプログラムは、ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定し、前記ユーザが発話する音声に対する音声認識を行って、前記起動ワードとして用いるように設定されている語彙が発話されたことを認識するステップを含む。

[0011] 本開示の一側面においては、ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、起動ワードとして用いる語彙が設定され、ユーザが発話する音声に対する音声認識を行って、起動ワードとして用いるように設定されている語彙が発話されたことが認識される。

発明の効果

[0012] 本開示の一側面によれば、より良いユーザ体験を提供することができる。

図面の簡単な説明

[0013] [図1]本技術を適用した情報処理装置の第1の実施の形態の構成例を示すブロック図である。

[図2]対応データベースに登録される語彙およびセンサパタンを説明する図である。

[図3]起動ワード語彙データベースに登録される語彙を説明する図である。

[図4]起動ワード認識制御装置において実行される起動ワード認識制御処理を説明するフローチャートである。

[図5]本技術を適用した情報処理装置の第2の実施の形態の構成例を示すブロック図である。

[図6]アプリケーション実行装置の構成例を示すブロック図である。

[図7]コマンドデータベースの一例を示す図である。

[図8]アプリケーション実行部において行われる処理について説明する図である。

[図9]アプリケーション実行部において行われる処理を説明するフローチャー

トである。

[図10]本技術を適用したコンピュータの一実施の形態の構成例を示すブロック図である。

発明を実施するための形態

[0014] 以下、本技術を適用した具体的な実施の形態について、図面を参照しながら詳細に説明する。

[0015] 図1は、本技術を適用した情報処理装置の第1の実施の形態の構成例を示すブロック図である。

[0016] 図1に示すように、情報処理装置11は、検出部12、音声入力部13、アプリケーション実行部14、出力部15、および起動ワード認識制御部16を備えて構成される。例えば、情報処理装置11は、ユーザが身に着けて持ち歩くことが可能なウェアラブル端末であり、ユーザの音声および動作に応じて音声認識機能を起動し、音声認識を利用したアプリケーションによって様々なユーザ体験を提供することができる。

[0017] 検出部12は、例えば、位置センサや、地磁気センサ、ジャイロセンサなどの各種のセンサを有して構成され、それらのセンサにより検出された検出結果を、アプリケーション実行部14および起動ワード認識制御部16に供給する。例えば、検出部12は、位置センサにより検出される情報処理装置11の現在位置情報を、検出結果としてアプリケーション実行部14および起動ワード認識制御部16に供給する。また、検出部12は、地磁気センサおよびジャイロセンサにより検出される情報処理装置11の動作情報（どの方向に、どのような動きをしたかを示す情報）を、検出結果としてアプリケーション実行部14および起動ワード認識制御部16に供給する。

[0018] 音声入力部13は、例えば、マイクロホンにより構成され、ユーザが発話する音声を電気信号に変換して、アプリケーション実行部14および起動ワード認識制御部16に供給する。

[0019] アプリケーション実行部14は、音声認識を利用した各種のアプリケーションを実行する。例えば、アプリケーション実行部14は、音声認識機能の

起動に用いるように設定された起動ワードが認識されたタイミングから一定期間だけ、音声認識機能を起動する。そして、アプリケーション実行部 14 は、音声認識機能が起動している期間におけるユーザの音声を認識し、その認識結果に基づいて、実行するアプリケーションに応じた出力を出力部 15 に供給する。なお、アプリケーション実行部 14 の構成については、図 6 を参照して後述する。

[0020] 出力部 15 は、例えば、ディスプレイやスピーカなどにより構成され、アプリケーション実行部 14 において実行されるアプリケーションによる認識結果に応じて、スピーカから音声を出し、ディスプレイに画像を表示する。

[0021] 起動ワード認識制御部 16 は、センサ監視部 21、対応データベース 22、起動ワード制御部 23、起動ワード語彙データベース 24、および起動ワード認識部 25 を備えて構成される。そして、起動ワード認識制御部 16 は、検出部 12 から供給される検出結果に基づいて、アプリケーション実行部 14 の音声認識機能を起動するための起動ワードとして用いる語彙を増減する制御を行う。

[0022] センサ監視部 21 は、検出部 12 が備える各種のセンサの状態を監視し、対応データベース 22 に登録されているセンサパタンおよび語彙を参照して、起動ワード制御部 23 に対する指示を行う。例えば、センサ監視部 21 は、検出部 12 から供給される検出結果に基づくユーザの動作が、所定の語彙を起動ワードとして用いることを開始する開始条件を示すセンサパタンに該当する状況となったとき、その語彙を起動ワードとして用いるように、起動ワード制御部 23 に対する指示を行う。また、このとき、センサ監視部 21 は、その語彙を起動ワードとして用いることを終了する終了条件を示すセンサパタンを対応データベース 22 から読み出して保持しておく。そして、センサ監視部 21 は、起動ワードとして用いている語彙の終了条件を示すセンサパタンに該当する状況となったとき、その語彙を起動ワードとして用いることを終了するように、起動ワード制御部 23 に対する指示を行う。

- [0023] 対応データベース 22 は、図 2 を参照して後述するように、起動ワードとして用いられる様々な語彙と、それらの語彙の開始条件および終了条件を示すセンサパタンとが対応付けられて登録されている。
- [0024] 起動ワード制御部 23 は、センサ監視部 21 からの指示に従って、起動ワード語彙データベース 24 に語彙を登録し、また、起動ワード語彙データベース 24 から語彙を削除することにより、起動ワードとして用いる語彙の増減を制御する。
- [0025] 起動ワード語彙データベース 24 には、起動ワードとして用いられる語彙が登録される。
- [0026] 起動ワード認識部 25 は、音声入力部 13 から入力される音声に対する音声認識処理を行う。そして、起動ワード認識部 25 は、音声認識処理を行った結果、起動ワード語彙データベース 24 に登録されている起動ワードをユーザが発話したことを認識した場合、その旨をアプリケーション実行部 14 に通知する。
- [0027] 以上のように構成される情報処理装置 11 では、センサ監視部 21 および起動ワード制御部 23 により、検出部 12 による検出結果に基づいて、アプリケーション実行部 14 の音声認識機能を起動させるためにユーザが発話する起動ワードが増減するように、起動ワードとして用いる語彙を制御することができる。これにより、例えば、アプリケーション実行部 14 の音声認識機能を常に起動させる構成と比較して、雑音などによって誤作動が発生することを抑制することができる。
- [0028] また、情報処理装置 11 では、アプリケーション実行部 14 が実行するアプリケーションに応じて、そのアプリケーションにより提供される世界観に沿った語彙を起動ワードとすることができる。これにより、アプリケーションが提供する世界観からユーザを引き離すような感覚を与えることを回避することができる、より良いユーザ体験を提供することができる。
- [0029] 次に、図 2 には、対応データベース 22 に登録される語彙およびセンサパタンが示されている。

[0030] 対応データベース 22 には、情報処理装置 11 において起動ワードとして用いられる全ての語彙が登録されている。そして、対応データベース 22 では、それらの語彙に対応付けて、起動ワードとして用いることを開始する開始条件を示すセンサパタン、および、起動ワードとして用いることを終了する終了条件を示すセンサパタンが登録されている。センサパタンとしては、例えば、検出部 12 により検出されるユーザの動作や、語彙を起動ワードとして用い始めてからの経過時間などが登録される。

[0031] 例えば、対応データベース 22 には、語彙「アブラカダブラ」に対応付けて、開始条件を示すセンサパタン「北緯35.6197、東経139.728553、10mの範囲内である」と、終了条件を示すセンサパタン「北緯35.6197、東経139.728553、10mの範囲外である」とが登録されている。

[0032] 従って、例えば、センサパタンで規定される範囲内にユーザが移動すると、検出部 12 から検出結果としてセンサ監視部 21 に供給される位置情報が、開始条件を示すセンサパタンに該当する状況を示すことになる。これに従い、センサ監視部 21 は、語彙「アブラカダブラ」を起動ワードとして用いるように、起動ワード制御部 23 に対する指示を行う。その後、例えば、センサパタンで規定される範囲外にユーザが移動すると、検出部 12 から検出結果としてセンサ監視部 21 に供給される位置情報が、終了条件を示すセンサパタンに該当する状況を示すことになる。これに従い、センサ監視部 21 は、語彙「アブラカダブラ」を起動ワードとして用いることを終了するように、起動ワード制御部 23 に対する指示を行う。

[0033] また、対応データベース 22 には、語彙「ジャンプ」に対応付けて、開始条件を示すセンサパタン「ユーザが3回ジャンプした」と、終了条件を示すセンサパタン「開始から10秒経過した」とが登録されている。

[0034] 従って、例えば、検出部 12 の加速度センサの出力に従って、ユーザが3回ジャンプしたことを示す検出結果が供給されると、センサ監視部 21 は、語彙「ジャンプ」の開始条件を示すセンサパタンに該当する状況になったと判断する。そして、センサ監視部 21 は、語彙「ジャンプ」を起動ワードと

して用いるように、起動ワード制御部 23 に対する指示を行う。このとき、センサ監視部 21 は、語彙「ジャンプ」を起動ワードとして用い始めてからの経過時刻を計時する。その後、センサ監視部 21 は、経過時刻が 10 秒を経過したタイミングで、終了条件を示すセンサパタンに該当する状況になったと判断し、語彙「ジャンプ」を起動ワードとして用いることを終了するように、起動ワード制御部 23 に対する指示を行う。

[0035] 同様に、対応データベース 22 には、語彙「加速装置」に対応付けて、開始条件を示すセンサパタン「ユーザの歩行速度が 25km/h 以上である」と、終了条件を示すセンサパタン「ユーザの歩行速度が 25km/h 未満である」とが登録されている。従って、例えば、センサ監視部 21 は、検出部 12 の加速度センサ、ジャイロセンサ、および位置センサの出力から求められるユーザの歩行速度に従って、語彙「加速装置」を起動ワードとして用いることを制御する。

[0036] また、対応データベース 22 には、語彙「召喚」に対応付けて、開始条件を示すセンサパタン「ユーザが顔を上に向けた」と、終了条件を示すセンサパタン「ユーザが顔を前に向けた」とが登録されている。例えば、情報処理装置 11 が眼鏡型のウェアラブル端末である場合、検出部 12 のジャイロセンサによりユーザの顔の向きを検出することができ、センサ監視部 21 は、ユーザの顔の向きに従って、語彙「加速装置」を起動ワードとして用いることを制御する。

[0037] このように、対応データベース 22 に登録されているセンサパタンを参照し、センサ監視部 21 は、それぞれのセンサパタンに対応付けられている語彙を、起動ワードとして用いることを制御することができる。

[0038] なお、図 2 に示したようなセンサパタンの他、例えば、加速度センサなどにより検知されるユーザの動作が所定の条件を満たした場合、一定時間だけ、所定の語彙を起動ワードとして用いることができる。また、ユーザの動作が第 1 の動作を行ってから、第 2 の動作を行うまでの間、所定の語彙を起動ワードとして用いるようにしてもよい。例えば、ユーザが、顔を右側に傾け

てから左側に傾けるまでの間や、腕を振り上げてから振り下ろすまでの間などをセンサパタンとすることができる。

[0039] 次に、図3には、起動ワード語彙データベース24に登録される語彙が示されている。

[0040] 例えば、起動ワード語彙データベース24には、常に起動ワードとして用いるように設定されている語彙「ハロースマホ」が登録されている。

[0041] そして、図3の上段に示すように、語彙「ハロースマホ」だけが登録されているときに、例えば、図2に示したセンサパタン「北緯35.6197、東経139.728553、10mの範囲内」にユーザが移動したとする。この場合、センサ監視部21は、語彙「アブラカダブラ」を起動ワードとして用いるように、起動ワード制御部23に対する指示を行う。この指示に応じて、起動ワード制御部23は、図3の中段に示すように、起動ワード語彙データベース24に語彙「アブラカダブラ」を追加する。

[0042] その後、ユーザが移動し、センサパタン「北緯35.6197、東経139.728553、10mの範囲外」となると、センサ監視部21は、語彙「アブラカダブラ」を起動ワードとして用いることを終了するように、起動ワード制御部23に対する指示を行う。これに従い、起動ワード制御部23は、図3の下段に示すように、起動ワード語彙データベース24から語彙「アブラカダブラ」を削除する。

[0043] 従って、起動ワード認識部25は、図3に示した語彙「アブラカダブラ」に対応するセンサパタン「北緯35.6197、東経139.728553、10mの範囲内」にユーザが居るときだけ、語彙「アブラカダブラ」を起動ワードとして認識することができる。

[0044] これにより、例えば、アプリケーション実行部14が実行するアプリケーションが、呪文を使って冒険を行うようなゲームである場合、情報処理装置11は、ゲーム内で所定のドアが表示されている場所にユーザが居る時だけ、語彙「アブラカダブラ」を起動ワードとすることができる。そして、そのドアを開けるための呪文をユーザが発話すると、アプリケーション実行部1

4の音声認識機能により呪文を認識して、ドアが開くような画像を表示させることができる。一方、例えば、音声認識機能を起動するために、情報処理装置11に対する通常の語彙「ハロースマホ」をユーザが発話した場合には、そのゲームの世界観からユーザを引き離すような感覚を与えてしまうことが想定される。これに対し、情報処理装置11は、ゲームが提供する世界観に合った語彙「アブラカダブラ」をユーザが発話することで、アプリケーション実行部14の音声認識機能を起動することができるため、ゲームの世界観に沿ったままのユーザ体験を提供することができる。

[0045] 次に、図4は、起動ワード認識制御部16において実行される起動ワード認識制御処理を説明するフローチャートである。

[0046] 例えば、情報処理装置11に対してアプリケーション実行部14による音声認識を利用したアプリケーションを実行するように操作が行われると、起動ワード認識制御部16において起動ワード認識制御処理が開始される。そして、ステップS11において、センサ監視部21は、検出部12から供給される検出結果に基づいて、対応データベース22に登録されているセンサパタンの開始条件に該当する状況になったか否かを判定する。

[0047] ステップS11において、センサ監視部21が、センサパタンの開始条件に該当する状況になったと判定した場合、処理はステップS12に進む。

[0048] ステップS12において、センサ監視部21は、開始条件に該当する状況になったと判定されたセンサパタンに対応付けられている語彙を起動ワードとして用いるように、起動ワード制御部23に対する指示を行う。これに従い、起動ワード制御部23は、センサ監視部21から指示された語彙が起動ワードとして用いられるように、起動ワード語彙データベース24に登録する。

[0049] ステップS13において、センサ監視部21は、ステップS12で起動ワードとして用いるように指示した語彙と、その語彙に対応付けられているセンサパタンの終了条件を、内部に保持する。

[0050] ステップS13の処理後、または、ステップS11でセンサパタンの開始

条件に該当する状況になっていないと判定された場合、処理はステップ S 1 4 に進む。

[0051] ステップ S 1 4 において、センサ監視部 2 1 は、検出部 1 2 から供給される検出結果に基づいて、対応データベース 2 2 に登録されているセンサパタンの終了条件に該当する状況になったか否かを判定する。

[0052] ステップ S 1 4 において、センサ監視部 2 1 が、センサパタンの開始条件に該当する状況になったと判定した場合、処理はステップ S 1 5 に進む。

[0053] ステップ S 1 5 において、センサ監視部 2 1 は、開始条件に該当する状況になったと判定されたセンサパタンに対応付けられている語彙を起動ワードから除外するように、起動ワード制御部 2 3 に対する指示を行う。これに従い、起動ワード制御部 2 3 は、センサ監視部 2 1 から指示された語彙を、起動ワード語彙データベース 2 4 から削除する。

[0054] ステップ S 1 6 において、センサ監視部 2 1 は、ステップ S 1 5 で起動ワードから除外するように指示した語彙と、その語彙に対応付けられているセンサパタンの終了条件を、内部から消去する。

[0055] ステップ S 1 6 の処理後、または、ステップ S 1 4 でセンサパタンの終了条件に該当する状況になっていないと判定された場合、処理はステップ S 1 7 に進む。

[0056] ステップ S 1 7 において、起動ワード認識部 2 5 は、音声入力部 1 3 から入力される音声に対する音声認識処理を行った結果、起動ワード語彙データベース 2 4 に登録されている起動ワードが発話されたか否かを判定する。

[0057] ステップ S 1 7 において、起動ワードが発話されたと判定された場合、処理はステップ S 1 8 に進み、起動ワードが発話された旨をアプリケーション実行部 1 4 に通知し、音声認識機能を起動するように指示を行う。これにより、アプリケーション実行部 1 4 において音声認識処理が開始される。

[0058] ステップ S 1 8 の処理後、または、ステップ S 1 7 で起動ワードが発話されていないと判定された場合、処理はステップ S 1 1 に戻り、以下、アプリケーション実行部 1 4 がアプリケーションの実行を終了するまで、同様の処

理が繰り返される。

[0059] 以上のように、情報処理装置 11 は、ユーザの行動がセンサパタンの開始条件に該当すると、その開始条件に対応付けられている語彙を起動ワードとして用いるように登録し、ユーザの行動がセンサパタンの終了条件に該当すると、その終了条件に対応付けられている語彙を起動ワードから除外することで、ユーザの行動に従って、起動ワードの増減を確実に制御することができる。

[0060] なお、情報処理装置 11 では、例えば、起動ワードデータベース 24 に登録され、起動ワードとして認識可能な語彙を、出力部 15 のディスプレイに表示することができる。これにより、使用可能な起動ワードをユーザに認識させることができる。

[0061] 次に、図 5 は、本技術を適用した情報処理装置の第 2 の実施の形態の構成例を示すブロック図である。

[0062] 図 5 に示すように、情報処理装置 11 A は、検出部 12、音声入力部 13、アプリケーション実行部 14、出力部 15、および起動ワード認識制御部 16 A を備えて構成される。なお、検出部 12、音声入力部 13、アプリケーション実行部 14、および出力部 15 の構成は、図 1 の情報処理装置 11 と同様であり、その詳細な説明は省略する。

[0063] 起動ワード認識制御部 16 A は、センサ監視部 21、対応データベース 22、起動ワード語彙データベース 24、起動ワード認識部 25、および起動ワードフィルタ部 26 を備えて構成される。

[0064] センサ監視部 21 は、検出部 12 が備える各種のセンサの状態を監視し、対応データベース 22 に登録されているセンサパタンを参照し、起動ワードフィルタ部 26 に対して、起動ワードとして用いる語彙を増減する制御を行う。対応データベース 22 には、図 2 を参照して上述したように、センサパタンおよび語彙が登録されている。

[0065] 起動ワード語彙データベース 24 には、情報処理装置 11 A において利用する可能性のある全ての起動ワードとなる語彙が、予め登録されている。起

動ワード認識部 25 は、音声入力部 13 から入力される音声に対する音声認識処理を行い、起動ワード語彙データベース 24 に登録されている語彙が入力されたことを認識した場合、その語彙を起動ワードフィルタ部 26 に供給する。

[0066] 起動ワードフィルタ部 26 は、起動ワード認識部 25 から供給される語彙が、センサ監視部 21 から起動ワードとして用いるように指示された語彙であるか否かを判定することにより、起動ワードのフィルタリングを行う。そして、起動ワードフィルタ部 26 は、起動ワード認識部 25 から供給される語彙が、センサ監視部 21 から起動ワードとして用いるように指示された語彙である場合には、アプリケーション実行部 14 に対して、ユーザが起動ワードを発話した旨を通知する。一方、起動ワードフィルタ部 26 は、起動ワード認識部 25 から供給される語彙が、センサ監視部 21 から起動ワードとして用いるように指示された語彙でない場合には、ユーザが起動ワードを発話していないと判断する。

[0067] このように構成される情報処理装置 11A は、起動ワードを確実に認識することができる。

[0068] つまり、図 1 の情報処理装置 11 では、起動ワード認識部 25 が起動ワードとして認識する語彙を増減させることで、語彙数を最小限に保つことができることより、高い認識性を維持することができる。しかしながら、図 1 の情報処理装置 11 では、検出部 12 による検出結果が変化した後起動ワードを起動ワード語彙データベース 24 に登録するため、起動ワードを登録する処理に時間がかかった場合、起動ワードの増減に対応する反応が遅くなってしまう。そのため、例えば、起動ワードの登録中に、その起動ワードが発話されると、認識することができない恐れがある。

[0069] これに対し、情報処理装置 11A では、起動ワードを起動ワード語彙データベース 24 に登録する処理が行われないので、情報処理装置 11 と比較して、起動ワードの増減に対して確実に対応して、起動ワードを認識することができないような状況を回避することができる。これにより、情報処理装置

1 1 Aは、より確実に、起動ワードを認識することができる。

[0070] 次に、図6は、アプリケーション実行部14の構成例を示すブロック図である。

[0071] 図6に示すように、アプリケーション実行部14は、音声認識部31、動作認識部32、センサ情報記憶部33、コマンドデータベース34、および制御部35を備えて構成される。

[0072] 音声認識部31は、制御部35の制御に従って音声認識を開始し、音声入力部13から入力される音声に対する音声認識処理を行う。例えば、ユーザが、アプリケーション実行部14で実行されるアプリケーションで用いられるコマンドを発話すると、音声認識部31は、そのコマンドの音声を認識し、認識結果と、その発話に関連する時刻情報を制御部35に供給する。

[0073] 動作認識部32は、検出部12が備える各種のセンサにより検出された検出結果に基づいた動作認識処理を行い、ユーザの動作を認識する。例えば、動作認識部32は、検出部12が備えるセンサの種類に応じて、ユーザが手を振り下ろした方向や、ユーザの顔および体の向いている方向などを認識して、それらの動作認識結果を、センサ情報としてセンサ情報記憶部33に記憶させる。また、動作認識部32は、制御部35からセンサ情報について問い合わせがあると、検出部12から供給される現在のセンサ情報、および、センサ情報記憶部33に記憶されているセンサ情報の中から、問い合わせられたセンサ情報に対応する動作認識結果を、制御部35に供給する。

[0074] センサ情報記憶部33は、動作認識部32による動作認識処理により求められるセンサ情報を記憶する。

[0075] コマンドデータベース34には、後述する図7に示すように、アプリケーション実行部14において実行されるアプリケーションで用いられるコマンドと、時刻範囲およびセンサ情報とが対応付けられて登録されている。

[0076] 制御部35は、起動ワード認識制御部16の起動ワード認識部25により起動ワードが認識され、音声認識を起動するように起動ワード認識部25から指示が入力されると、音声認識処理を開始するように音声認識部31に対

する起動制御を行う。そして、制御部 35 は、音声認識部 31 から供給されるコマンドの音声の認識結果および時刻情報に基づいて、コマンドデータベース 34 を参照し、認識されたコマンドに対応付けられているセンサ情報を動作認識部 32 に対して、動作認識結果の問い合わせを行う。その後、制御部 35 は、動作認識部 32 からの動作認識結果に従った出力を、出力部 15 に供給する。なお、例えば、図 5 に示した情報処理装置 11A の構成においては、起動ワードフィルタ部 26（図 5）から制御部 35 に対して、音声認識を起動するように指示が入力され、同様の処理が行われる。

[0077] 例えば、図 7 には、コマンドデータベース 34 の一例が示されている。

[0078] 図 7 に示すように、コマンドデータベース 34 では、コマンド、時刻範囲、およびセンサ情報が対応付けられて登録されている。

[0079] 例えば、コマンド「ビーム発射」には、時刻範囲「音声発話後 3 秒以内」およびセンサ情報「右手が振り下ろされた方向」が対応付けられている。即ち、ユーザにより発話されるコマンド「ビーム発射」に対して、音声認識部 31 によりコマンド「ビーム発射」が発話されたと認識され、その発話が終了した時刻から 3 秒以内にユーザの右腕が振り下ろされた方向を、動作認識部 32 は、動作認識結果として制御部 35 に供給する。この動作認識結果に応じ、制御部 35 は、ユーザの右腕が振り下ろされた方向に、コマンド「ビーム発射」に応じた出力（例えば、ビームのエフェクトの表示）を行う。

[0080] また、コマンド「ロケット発射」には、時刻範囲「発話直後」およびセンサ情報「体の向いている方向」が対応付けられている。即ち、ユーザにより発話されるコマンド「ロケット発射」に対して、音声認識部 31 によりコマンド「ロケット発射」が発話されたと認識され、その発話が終了した直後にユーザの体が向いている方向を、動作認識部 32 は、動作認識結果として制御部 35 に供給する。この動作認識結果に応じ、制御部 35 は、ユーザの体が向いている方向に、コマンド「ロケット発射」に応じた出力（例えば、発射されたロケットが進行するエフェクトの表示）を行う。

[0081] また、コマンド「なぎ払え」には、時刻範囲「音声区間開始 0.1 秒前から音

声終了まで」およびセンサ情報「右手が向いていた方向の範囲」が対応付けられている。即ち、ユーザにより発話されるコマンド「なぎ払え」に対して、音声認識部 31 によりコマンド「なぎ払え」が発話されたと認識される直前の時刻（0.1秒前）から、その発話が終了する時刻までに、ユーザの右手が向いていた範囲を、動作認識部 32 は、動作認識結果として制御部 35 に供給する。

[0082] また、コマンド「ココをマーク」には、時刻範囲「音声認識結果中の単語「ココ」の単語の終端時刻」およびセンサ情報「ユーザの顔の向いていた方向、ユーザの位置」が対応付けられている。即ち、ユーザにより発話されるコマンド「ココをマーク」に対して、音声認識部 31 によりコマンド「ココをマーク」が発話されたと認識されたうちの、単語「ココ」の発話が終了した時刻に、ユーザの顔の向いていた方向とユーザの位置とを、動作認識部 32 は、動作認識結果として制御部 35 に供給する。この動作認識結果に応じ、制御部 35 は、ユーザの体が向いている方向の前方箇所（ユーザの目の前の位置）に、コマンド「ココをマーク」に応じた出力（例えば、マークの固定的な表示）を行う。

[0083] このように、コマンドデータベース 34 には、音声認識部 31 により認識されるコマンドに対応付けて、そのコマンドが発話された時刻を基準とした時刻範囲と、動作認識部 32 により認識されるセンサ情報（方位や、方位の範囲、向き、高度など）とが登録される。

[0084] ここで、図 8 を参照して、ユーザがコマンド「ビーム発射」を発話したときに、アプリケーション実行部 14 において行われる処理の例について説明する。

[0085] 例えば、情報処理装置 11 は、腕時計型のウェアラブル端末であって、検出部 12 が備える地磁気センサおよびジャイロセンサにより、ユーザの腕の動きや方向を検出することができるものとする。この場合、例えば、上述したように、起動ワードの発話を検出して音声認識部 31 の音声認識処理を開始する他、図 8 の最上段に示すように、ユーザが右手を上げたことを検出し

たときに、音声認識部 31 の音声認識処理を開始することができる。

[0086] 制御部 35 は、ユーザにより起動ワードが発話されたこと、または、ユーザにより所定の動作（この例では、右手を上げる）が行われたことが認識され、起動ワード認識制御部 16 から音声認識処理を起動するように指示されると、音声認識部 31 に対して音声認識処理を開始するように指示する処理を行う。そして、制御部 35 は、ユーザにより発話された音声コマンドが音声認識部 31 により認識されるまで処理を待機する。

[0087] そして、図 8 の上から 2 段目に示すように、ユーザが、コマンド「ビーム発射」を発話すると、音声認識部 31 は、そのコマンド「ビーム発射」を認識したことを示す認識結果と、コマンド「ビーム発射」が発話された時刻を示す時刻情報を制御部 35 に供給する。

[0088] これに応じて、制御部 35 は、音声認識部 31 により認識されたコマンド「ビーム発射」に対応付けられている時刻範囲およびセンサ情報をコマンドデータベース 34 から読み出して、その時刻範囲におけるセンサ情報を動作認識部 32 に対して要求する。

[0089] その後、図 8 の上から 3 段目に示すように、ユーザが、上げていた右手を振り下ろす動作を行う。このような動作が行われたことを示す検出結果を検出部 12 が動作認識部 32 に供給すると、動作認識部 32 は、ユーザの右手が振り下ろされた方向を動作認識結果として求め、制御部 35 に供給する。

[0090] そして、制御部 35 は、図 8 の上から 4 段目に示すように、ユーザの右手が振り下ろされた方向に、ビームが発射されたようなエフェクトを表現する画像を、拡張現実的に出力部 15 のディスプレイに表示させる。

[0091] このように、アプリケーション実行部 14 は、ユーザがコマンドを発話した時刻に基づいて、ユーザの動作を認識し、そのユーザの動作に合った出力を行うことができる。

[0092] 次に、図 9 のフローチャートを参照して、アプリケーション実行部 14 において実行される音声認識を利用したアプリケーションの処理について説明する。

- [0093] 例えば、起動ワード認識制御部 16 の起動ワード認識部 25 により起動ワードが認識されると処理が開始され、ステップ S 21 において、制御部 35 は、音声認識部 31 に対して音声認識処理を開始するように制御を行う。
- [0094] ステップ S 22 において、制御部 35 は、音声認識部 31 によりコマンドが認識されたか否かを判定し、音声認識部 31 によりコマンドが認識されたと判定するまで処理を待機する。そして、制御部 35 は、音声認識部 31 からコマンドの音声の認識結果および時刻情報が供給されると、コマンドが認識されたと判定して、処理はステップ S 23 に進む。
- [0095] ステップ S 23 において、制御部 35 は、コマンドデータベース 34 を参照して、認識されたコマンドに対応付けられているセンサ情報を動作認識部 32 に対して問い合わせを行う。
- [0096] ステップ S 24 において、動作認識部 32 は、問い合わせられたセンサ情報に対応する動作認識結果を制御部 35 に供給し、制御部 35 は、その動作認識結果に従った出力を、出力部 15 に供給する。
- [0097] ステップ S 24 の処理後、処理はステップ S 22 に戻り、以下、同様の処理が繰り返される。
- [0098] 以上のように、アプリケーション実行部 14 は、ユーザが発話するコマンドと、そのコマンドが発話された時刻を基準としたユーザの動作とを統合した処理を実行することができる。これにより、アプリケーション実行部 14 は、アプリケーションが提供する世界観に合わせたコマンドおよび動作をユーザに行わせることで、そのコマンドおよび動作に従った出力を行うことができる。
- [0099] 例えば、情報処理装置 11 が、眼鏡型のウェアラブル端末であって、出力部 15 が、ユーザの視界に重畳させて画像を表示することができる場合、ユーザが走っている途中にコマンド「加速装置」を発話すると、走っている方向に関連するビジュアルイフェクトを出力部 15 に表示させることができる。なお、コマンド「加速装置」を発話した後、ユーザが走り出したのを検出したときに、走っている方向に関連するビジュアルイフェクトを表示しても

よい。

[0100] また、例えば、情報処理装置 11 が、腕時計型のウェアラブル端末である場合、右手をある方向に上から振り出し、それをトリガとして音声コマンドが入力可能になり（例えば「ビーム発射」）、その後に認識される音声コマンドに応じて情報を出力する方向（例えば「ビームの方向」）を、振り下ろした右手の方向とすることもできる。また、情報処理装置 11 では、所定の位置で、ユーザがコマンド「罨を仕掛ける」を発話すると、その位置情報を取得して、アプリケーション実行部 14 が実行するゲーム上に、そのコマンドに対応する仕掛けを設定することができる。

[0101] このように、情報処理装置 11 では、ユーザがコマンドを発話した時刻を基準としたユーザの動作（ジェスチャ）を検出することで、アプリケーション実行部 14 が実行するアプリケーションに応じた様々な出力を行うことができる。また、情報処理装置 11 は、ユーザが連続して発話する単語「コレと、コレと、コレ」に応じて、それぞれの「コレ」が認識された時刻と、その時刻におけるユーザの手の方向とを連動して、ユーザが指示した物体を認識することができる。即ち、制御部 35 は、ユーザが物体を指示する単語を発話し、その単語を音声認識部 31 が認識した時刻において、動作認識部 32 により認識されたユーザの手の方向の先にある物体を、ユーザの指示対象物として認識する。このように、情報処理装置 11 は、画面上に表示される表示物を指示対象とするのではなく、実際の物体を指示対象とすることができる。

[0102] また、情報処理装置 11 は、検出部 12 により検出されたユーザの動作を、音声認識部 31 が音声認識する際の単語の切れ目を判断するのに利用することができる。

[0103] なお、情報処理装置 11 の検出部 12 が備えるセンサとしては、高度センサ、地磁気センサ、光センサ、血糖センサ、体温センサ、時計、および歩数計を用いることができる。例えば、高度センサを利用して、ある高さ以上または以下の場所にユーザがいる時のみ、所定の語彙が起動ワードとして用い

られるようにすることができる。また、地磁気センサを利用して、ある方向に向かっている時のみ、所定の語彙が起動ワードとして用いられるようにすることができる。また、光センサを利用して、明るい場所でのみ、または、暗い場所でのみ、所定の語彙が起動ワードとして用いられるようにすることができる。また、血糖センサを利用したときにはユーザが空腹である時のみ、時計を利用したときには所定の時刻のみ、歩数計を利用したときには所定歩数を歩いた後に、所定の語彙が起動ワードとして用いられるようにすることができる。

[0104] また、これらを組み合わせても利用してもよく、例えば、高度センサと位置センサを組み合わせることで、所定の位置にある建物の屋上にユーザが移動したときに、所定の語彙を起動ワードとして用いることができるようにしてもよい。また、情報処理装置 11 が撮像装置を備えている場合には、特定の対象物を撮像したときに、所定の語彙を起動ワードとして用いることができるようにしてもよい。

[0105] なお、情報処理装置 11 では、アプリケーション実行部 14 において特定のアプリケーションが実行されているとき、例えば、呪文を使用したゲームが実行されているとき、通常時に起動ワードとして使用される語彙（図 3 に示した語彙「ハロースマホ」）を、音声認識を起動させる起動ワードとして用いることを禁止することができる。また、情報処理装置 11 では、図 2 を参照して上述したようなセンサパタンに該当する状況となったときに、自動的に、音声認識機能を起動するようにしてもよい。

[0106] なお、上述のフローチャートを参照して説明した各処理は、必ずしもフローチャートとして記載された順序に沿って時系列に処理する必要はなく、並列的あるいは個別に実行される処理（例えば、並列処理あるいはオブジェクトによる処理）も含むものである。また、プログラムは、1 の CPU により処理されるものであっても良いし、複数の CPU によって分散処理されるものであっても良い。

[0107] また、上述した一連の処理（情報処理方法）は、ハードウェアにより実行

することもできるし、ソフトウェアにより実行することもできる。一連の処理をソフトウェアにより実行する場合には、そのソフトウェアを構成するプログラムが、専用のハードウェアに組み込まれているコンピュータ、または、各種のプログラムをインストールすることで、各種の機能を実行することが可能な、例えば汎用のパーソナルコンピュータなどに、プログラムが記録されたプログラム記録媒体からインストールされる。

[0108] 図10は、上述した一連の処理をプログラムにより実行するコンピュータのハードウェアの構成例を示すブロック図である。

[0109] コンピュータにおいて、CPU (Central Processing Unit) 101, ROM (Read Only Memory) 102, RAM (Random Access Memory) 103は、バス104により相互に接続されている。

[0110] バス104には、さらに、入出力インタフェース105が接続されている。入出力インタフェース105には、キーボード、マウス、マイクロホンなどよりなる入力部106、ディスプレイ、スピーカなどよりなる出力部107、ハードディスクや不揮発性のメモリなどよりなる記憶部108、ネットワークインタフェースなどよりなる通信部109、磁気ディスク、光ディスク、光磁気ディスク、或いは半導体メモリなどのリムーバブルメディア111を駆動するドライブ110が接続されている。

[0111] 以上のように構成されるコンピュータでは、CPU101が、例えば、記憶部108に記憶されているプログラムを、入出力インタフェース105及びバス104を介して、RAM103にロードして実行することにより、上述した一連の処理が行われる。

[0112] コンピュータ (CPU101) が実行するプログラムは、例えば、磁気ディスク (フレキシブルディスクを含む)、光ディスク (CD-ROM (Compact Disc-Read Only Memory), DVD (Digital Versatile Disc) 等)、光磁気ディスク、もしくは半導体メモリなどよりなるパッケージメディアであるリムーバブルメディア111に記録して、あるいは、ローカルエリアネットワーク、インターネット、デジタル衛星放送といった、有線または無線の伝送媒体を介して提

供される。

[0113] そして、プログラムは、リムーバブルメディア 111 をドライブ 110 に装着することにより、入出力インタフェース 105 を介して、記憶部 108 にインストールすることができる。また、プログラムは、有線または無線の伝送媒体を介して、通信部 109 で受信し、記憶部 108 にインストールすることができる。その他、プログラムは、ROM 102 や記憶部 108 に、あらかじめインストールしておくことができる。

[0114] なお、本技術は以下のような構成も取ることができる。

(1)

ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定する起動ワード設定部と、

前記ユーザの発話する音声に対する音声認識を行って、前記起動ワード設定部により前記起動ワードとして用いるように設定されている語彙が発話されたことを認識する起動ワード認識部と

を備える情報処理装置。

(2)

前記起動ワード設定部は、前記ユーザの動作が、所定の語彙を前記起動ワードとして用いることを開始する開始条件に該当すると、その開始条件に対応付けられている語彙を、前記起動ワードとして用いるように登録する

上記(1)に記載の情報処理装置。

(3)

前記起動ワード設定部は、前記所定の語彙が前記起動ワードとして用いるように登録されている場合に、前記ユーザの動作が、前記所定の語彙を前記起動ワードとして用いることを終了する終了条件に該当すると、前記所定の語彙を前記起動ワードから除外する

上記(2)に記載の情報処理装置。

(4)

前記起動ワード設定部は、前記所定の語彙を前記起動ワードとして用いることを終了する終了条件として、前記開始条件となった時刻からの経過時間が設定されている場合、その所定の語彙を前記起動ワードとして用いるように登録してからの時刻を計時し、前記経過時間が経過すると、その所定の語彙を前記起動ワードから除外する

上記（２）に記載の情報処理装置。

（５）

前記起動ワード設定部は、前記検出結果として供給される位置情報に基づいて、前記ユーザの位置に従い、前記起動ワードとして用いる語彙の増減を制御する

上記（１）から（４）までのいずれかに記載の情報処理装置。

（６）

前記起動ワード設定部は、前記検出結果として供給される少なくとも加速度情報に基づいて、前記ユーザの動作に従い、前記起動ワードとして用いる語彙の増減を制御する

上記（１）から（５）までのいずれかに記載の情報処理装置。

（７）

前記起動ワード認識部により前記ユーザが前記起動ワードを発話したことが認識された場合に、音声認識を利用したアプリケーションにおける音声認識処理を開始するアプリケーション実行部をさらに備える

上記（１）から（６）までのいずれかに記載の情報処理装置。

（８）

前記起動ワード認識部は、前記起動ワードとして利用する可能性のある全ての語彙を認識する音声認識を行い、

前記起動ワード設定部は、前記起動ワード認識部により認識された語彙に対して、前記ユーザの動作が検出された検出結果に従って前記起動ワードとして用いることとした前記語彙であるか否かによるフィルタリングを行うことにより、前記起動ワードの増減に対応する

上記（１）に記載の情報処理装置。

（９）

前記アプリケーション実行部は、

ユーザが所定のコマンドを発話したことを認識する音声認識部と、

前記音声認識部により認識されたコマンドが発話された時刻を基準とした時刻範囲における、前記コマンドに対応するユーザの動作を認識する動作認識部と、

前記動作認識部による動作認識結果に従って、前記コマンドに応じた出力を行う制御部と

を有する

上記（７）に記載の情報処理装置。

（１０）

前記動作認識部は、前記音声認識部により認識されたコマンドの発話が終了したときに、ユーザの体が向いている方向を認識し、

前記制御部は、前記動作認識部により認識された方向に、前記コマンドに応じた出力を行う

上記（９）に記載の情報処理装置。

（１１）

前記動作認識部は、前記音声認識部により認識されたコマンドが発話された時刻から所定の時刻範囲内に、ユーザの腕が振り下ろされた方向を認識し、

前記制御部は、前記動作認識部により認識された方向に、前記コマンドに応じた出力を行う

上記（９）または（１０）に記載の情報処理装置。

（１２）

前記制御部は、前記音声認識部により任意の物体を指示する単語が認識された時刻において、前記動作認識部により認識されたユーザの手の方向の先にある物体を、ユーザの指示対象物として認識する

上記（９）乃至（１１）までのいずれかに記載の情報処理装置。

（１３）

ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定し、

前記ユーザが発話する音声に対する音声認識を行って、前記起動ワードとして用いるように設定されている語彙が発話されたことを認識するステップを含む情報処理方法。

（１４）

ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定し、

前記ユーザが発話する音声に対する音声認識を行って、前記起動ワードとして用いるように設定されている語彙が発話されたことを認識するステップを含む情報処理をコンピュータに実行させるプログラム。

[0115] なお、本実施の形態は、上述した実施の形態に限定されるものではなく、本開示の要旨を逸脱しない範囲において種々の変更が可能である。

符号の説明

[0116] １１ 情報処理装置、 １２ 検出部、 １３ 音声入力部、 １４ アプリケーション実行部、 １５ 出力部、 １６ 起動ワード認識制御部、 ２１ センサ監視部、 ２２ 対応データベース、 ２３ 起動ワード制御部、 ２４ 起動ワード語彙データベース、 ２５ 起動ワード認識部、 ２６ 起動ワードフィルタ部、 ３１ 音声認識部、 ３２ 動作認識部、 ３３ センサ情報記憶部、 ３４ コマンドデータベース、 ３５ 制御部

請求の範囲

- [請求項1] ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定する起動ワード設定部と、
前記ユーザの発話する音声に対する音声認識を行って、前記起動ワード設定部により前記起動ワードとして用いるように設定されている語彙が発話されたことを認識する起動ワード認識部と
を備える情報処理装置。
- [請求項2] 前記起動ワード設定部は、前記ユーザの動作が、所定の語彙を前記起動ワードとして用いることを開始する開始条件に該当すると、その開始条件に対応付けられている語彙を、前記起動ワードとして用いるように登録する
請求項1に記載の情報処理装置。
- [請求項3] 前記起動ワード設定部は、前記所定の語彙が前記起動ワードとして用いるように登録されている場合に、前記ユーザの動作が、前記所定の語彙を前記起動ワードとして用いることを終了する終了条件に該当すると、前記所定の語彙を前記起動ワードから除外する
請求項2に記載の情報処理装置。
- [請求項4] 前記起動ワード設定部は、前記所定の語彙を前記起動ワードとして用いることを終了する終了条件として、前記開始条件となった時刻からの経過時間が設定されている場合、その所定の語彙を前記起動ワードとして用いるように登録してからの時刻を計時し、前記経過時間が経過すると、その所定の語彙を前記起動ワードから除外する
請求項2に記載の情報処理装置。
- [請求項5] 前記起動ワード設定部は、前記検出結果として供給される位置情報に基づいて、前記ユーザの位置に従い、前記起動ワードとして用いる語彙の増減を制御する
請求項1に記載の情報処理装置。

- [請求項6] 前記起動ワード設定部は、前記検出結果として供給される少なくとも加速度情報に基づいて、前記ユーザの動作に従い、前記起動ワードとして用いる語彙の増減を制御する
- 請求項1に記載の情報処理装置。
- [請求項7] 前記起動ワード認識部により前記ユーザが前記起動ワードを発話したことが認識された場合に、音声認識を利用したアプリケーションにおける音声認識処理を開始するアプリケーション実行部をさらに備える
- 請求項1に記載の情報処理装置。
- [請求項8] 前記起動ワード認識部は、前記起動ワードとして利用する可能性のある全ての語彙を認識する音声認識を行い、
- 前記起動ワード設定部は、前記起動ワード認識部により認識された語彙に対して、前記検出結果に従って前記起動ワードとして用いることとした前記語彙であるか否かによるフィルタリングを行うことにより、前記起動ワードの増減に対応する
- 請求項1に記載の情報処理装置。
- [請求項9] 前記アプリケーション実行部は、
- ユーザが所定のコマンドを発話したことを認識する音声認識部と、
- 前記音声認識部により認識されたコマンドが発話された時刻を基準とした時刻範囲における、前記コマンドに対応するユーザの動作を認識する動作認識部と、
- 前記動作認識部による動作認識結果に従って、前記コマンドに応じた出力を行う制御部と
- を有する
- 請求項7に記載の情報処理装置。
- [請求項10] 前記動作認識部は、前記音声認識部により認識されたコマンドの発話が終了したときに、ユーザの体が向いている方向を認識し、

前記制御部は、前記動作認識部により認識された方向に、前記コマンドに応じた出力を行う

請求項 9 に記載の情報処理装置。

[請求項11] 前記動作認識部は、前記音声認識部により認識されたコマンドが発話された時刻から所定の時刻範囲内に、ユーザの腕が振り下ろされた方向を認識し、

前記制御部は、前記動作認識部により認識された方向に、前記コマンドに応じた出力を行う

請求項 9 に記載の情報処理装置。

[請求項12] 前記制御部は、前記音声認識部により任意の物体を指示する単語が認識された時刻において、前記動作認識部により認識されたユーザの手の方向の先にある物体を、ユーザの指示対象物として認識する

請求項 9 に記載の情報処理装置。

[請求項13] ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定し、

前記ユーザが発話する音声に対する音声認識を行って、前記起動ワードとして用いるように設定されている語彙が発話されたことを認識する

ステップを含む情報処理方法。

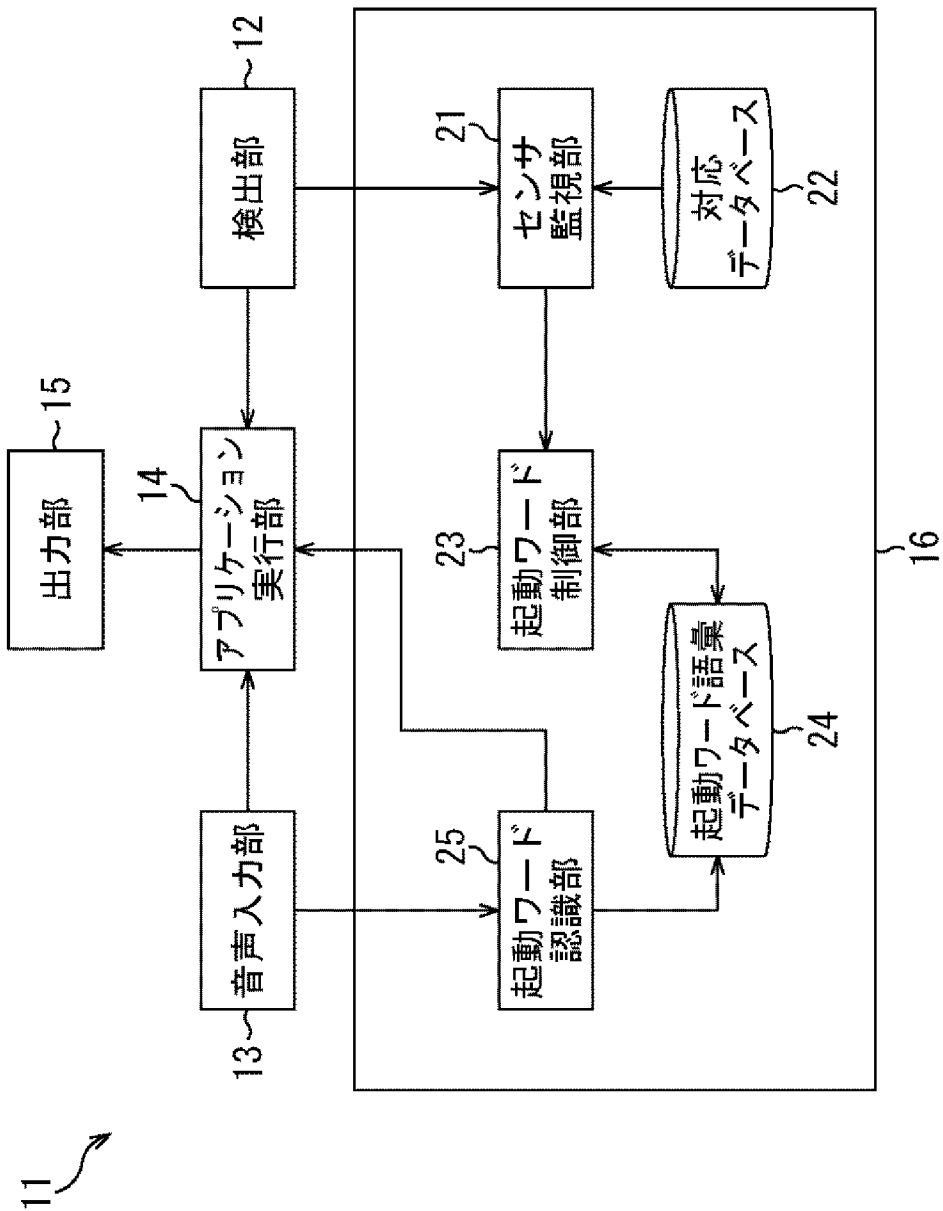
[請求項14] ユーザの動作が検出された検出結果に基づいて、所定の機能を起動させるためにユーザが発話する起動ワードが増減するように、前記起動ワードとして用いる語彙を設定し、

前記ユーザが発話する音声に対する音声認識を行って、前記起動ワードとして用いるように設定されている語彙が発話されたことを認識する

ステップを含む情報処理をコンピュータに実行させるプログラム。

[図1]

FIG.1



[図2]
FIG.2

語彙	センサパターン(開始)	センサパターン(終了)
「アブラカダブラ」	北緯35.6197, 東経139.728553, 10mの範囲内である	北緯35.6197, 東経139.728553, 10mの範囲外である
「ジャンプ」	ユーザが3回ジャンプした	開始から10秒経過した
「加速装置」	ユーザの歩行速度が 25km/h以上である	ユーザの歩行速度が 25km/h未満である
「召喚」	ユーザが顔を上に向けた	ユーザが顔を前に向けた

[図3]
FIG.3

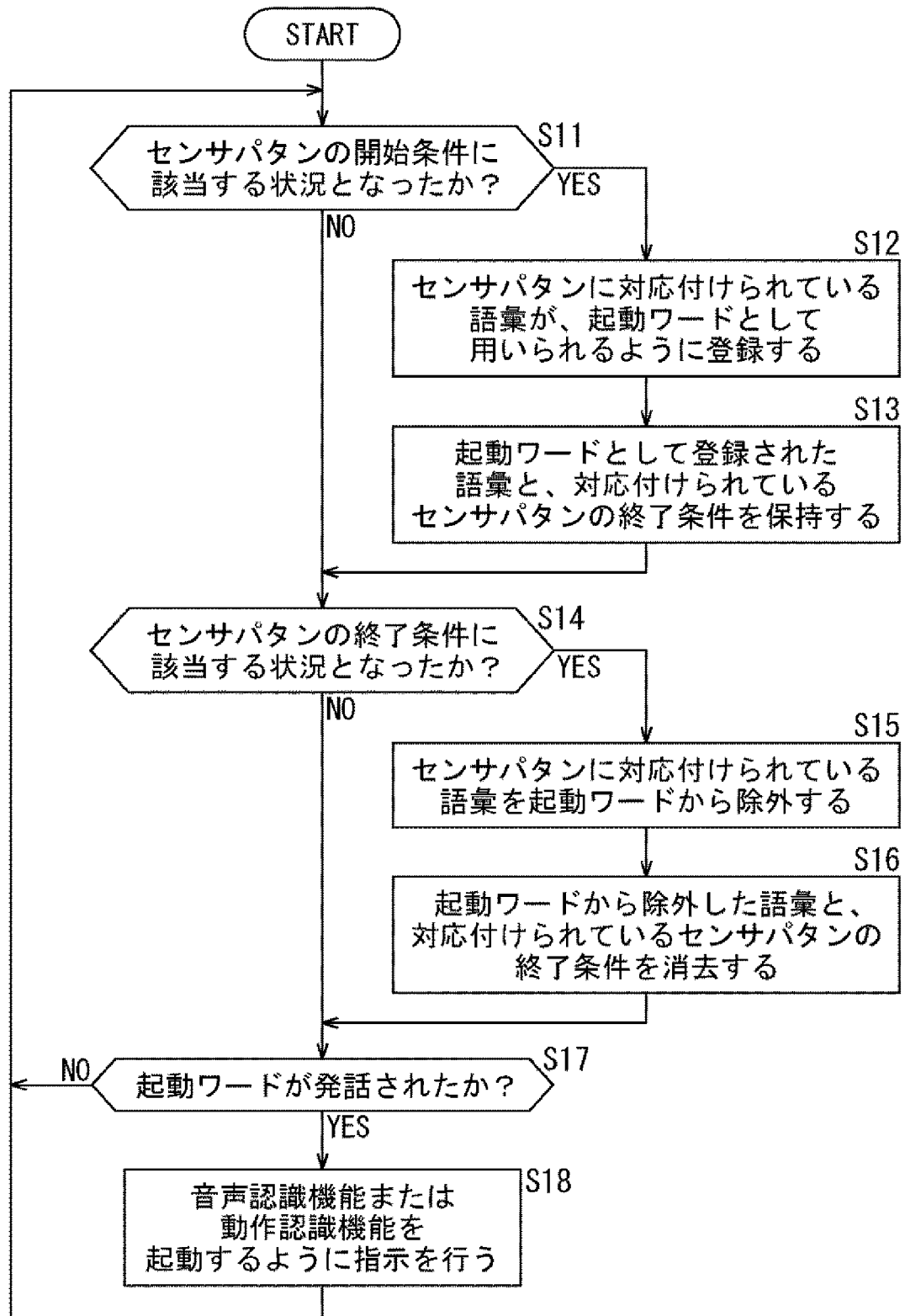
語彙
ハロースマホ



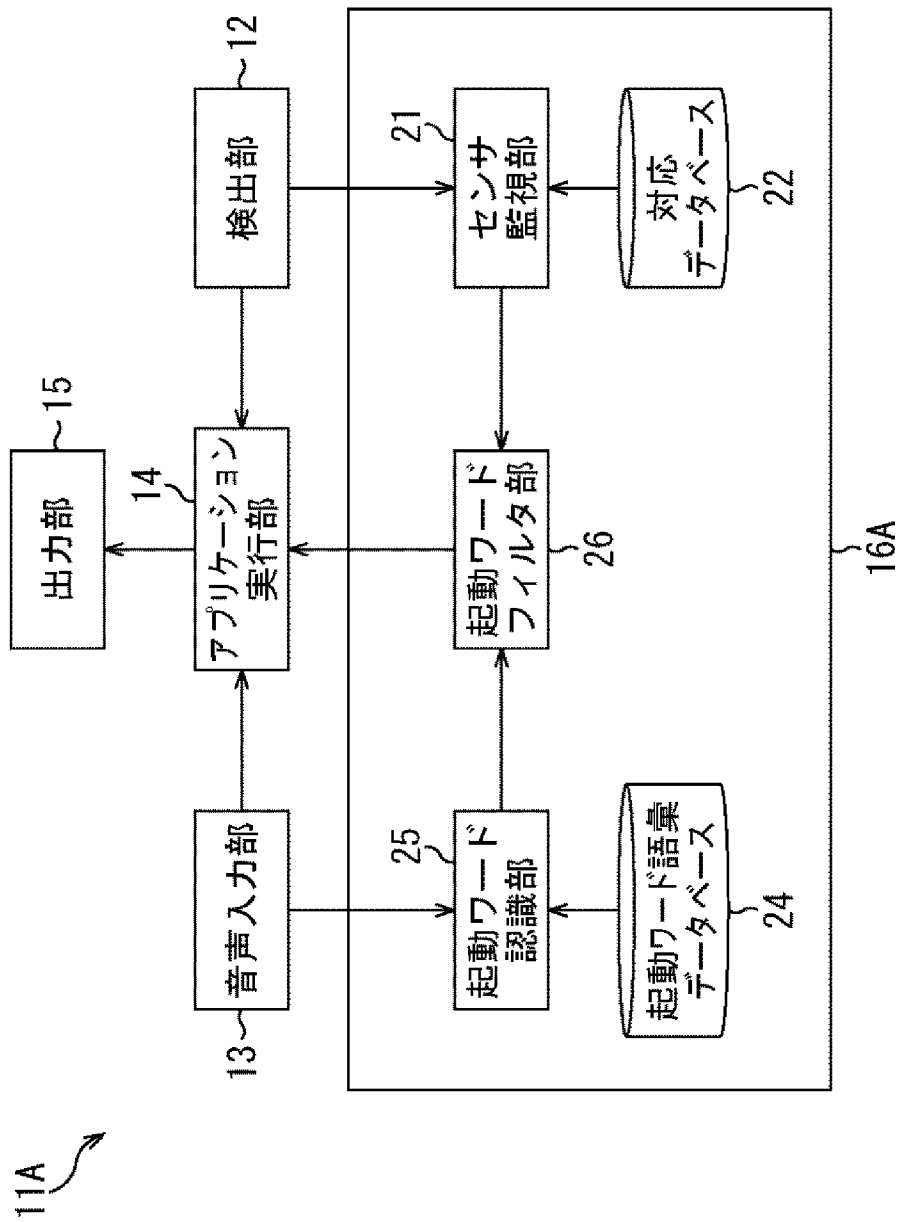
語彙
ハロースマホ
アブラカダブラ

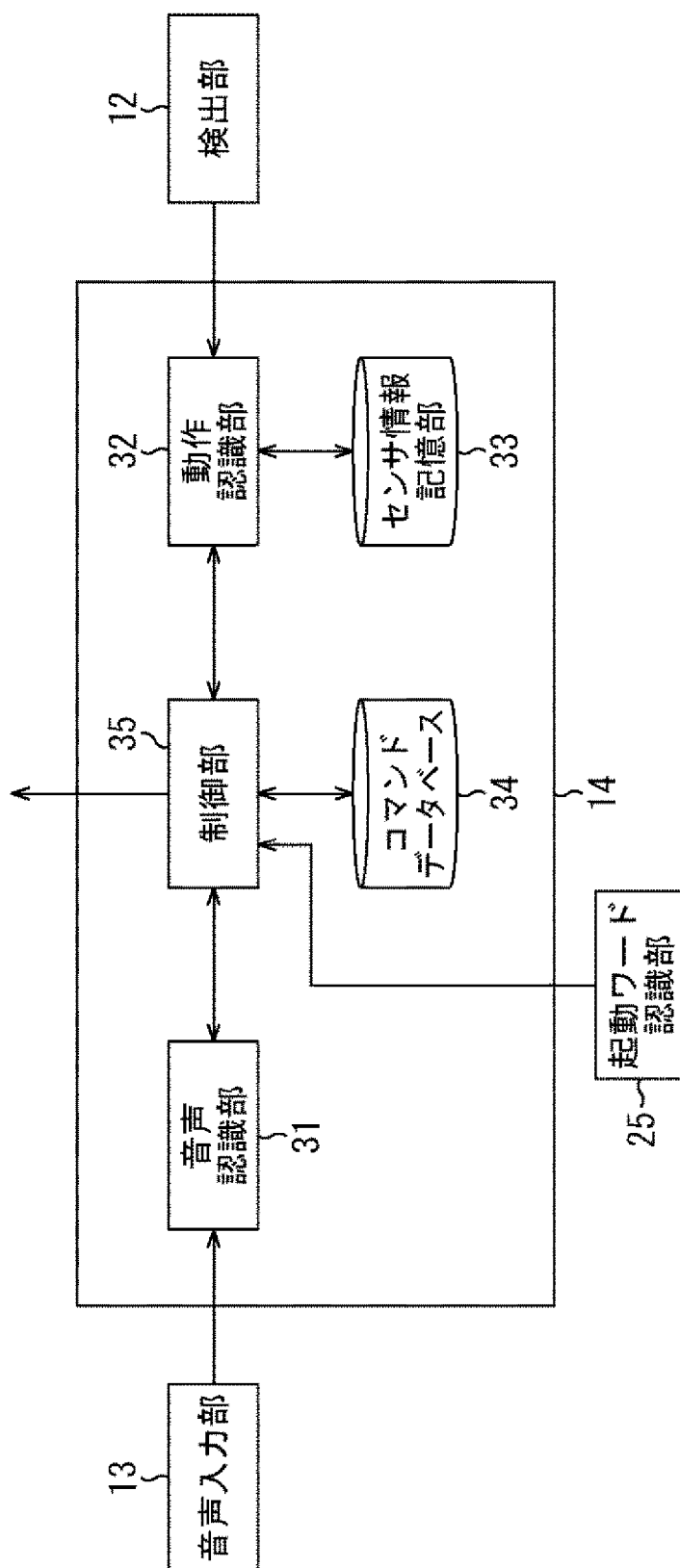


語彙
ハロースマホ

[図4]
FIG.4

[図5]
FIG. 5

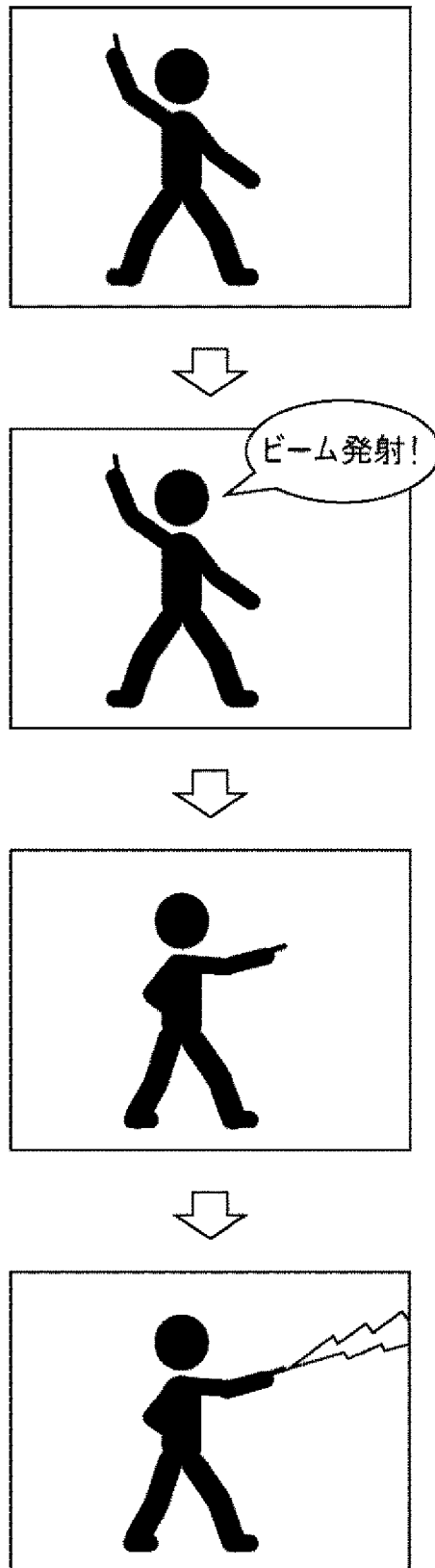


[図6]
FIG. 6

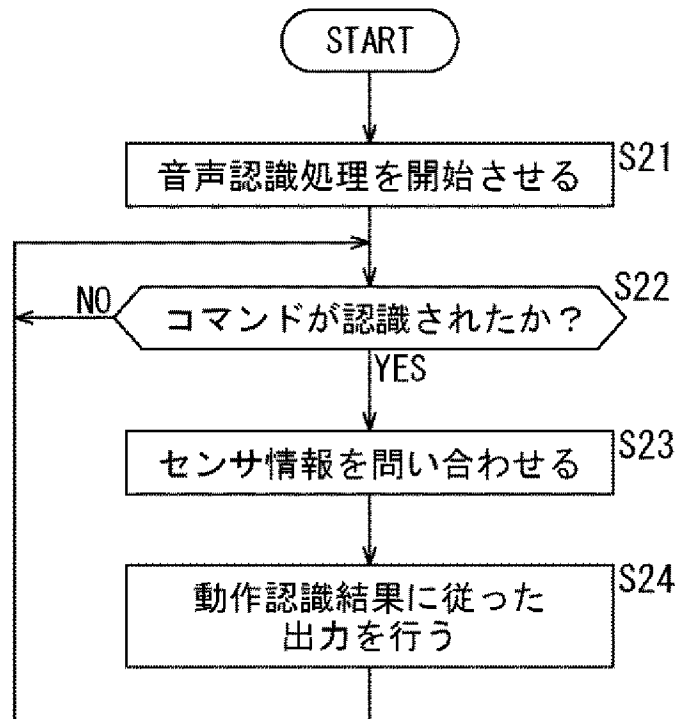
[図7]
FIG.7

コマンド	時刻範囲	センサ情報
「ビーム発射」	音声発話後3秒以内	右手が振り下ろされた方向
「ロケット発射」	発話直後	体の向いている方向
「なぎ払え」	音声区間開始0.1秒前から 音声終了まで	右手が向いていた方向の 範囲
「ココをマーク」	音声認識結果中の単語 「ココ」の単語の終端時刻	ユーザの顔の向いていた 方向、ユーザの位置

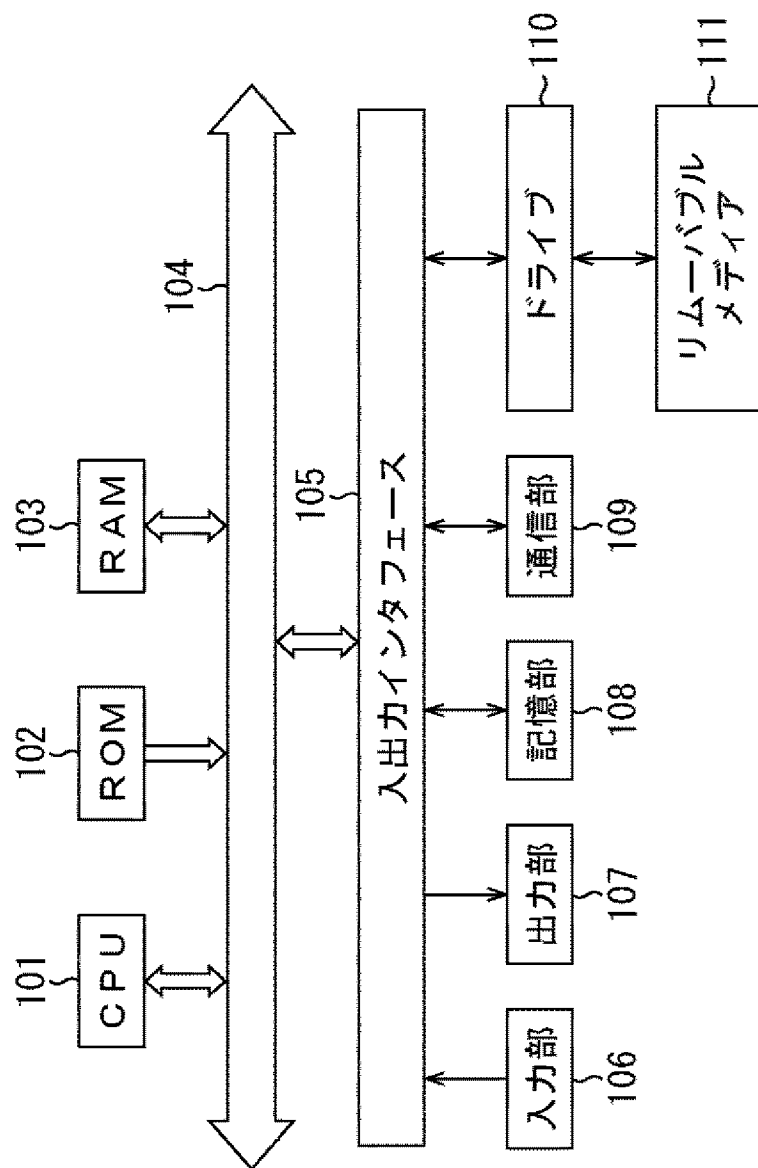
[図8]
FIG. 8



[図9]
FIG. 9



[図10]
FIG.10



INTERNATIONAL SEARCH REPORT

International application No.

PCT/JP2016/063631

A. CLASSIFICATION OF SUBJECT MATTER

G06F3/16(2006.01)i, G06F3/01(2006.01)i, G10L15/22(2006.01)i, G10L15/24(2013.01)n

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)

G06F3/16, G06F3/01, G10L15/22, G10L15/24

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

Jitsuyo Shinan Koho	1922-1996	Jitsuyo Shinan Toroku Koho	1996-2016
Kokai Jitsuyo Shinan Koho	1971-2016	Toroku Jitsuyo Shinan Koho	1994-2016

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 2015-076877 A (Panasonic Intellectual Property Corporation of America), 20 April 2015 (20.04.2015), paragraphs [0026] to [0068], [0113] to [0122], [0229] to [0246]; fig. 1 to 4, 9, 22 to 25 & US 2015/0100323 A1 paragraphs [0061] to [0103], [0148] to [0157], [0273] to [0281]; fig. 1 to 4, 9, 22 to 25	1-6, 8, 13-14 9-12
X A	WO 2013/188002 A1 (Honda Motor Co., Ltd.), 19 December 2013 (19.12.2013), paragraphs [0021] to [0073]; fig. 1 to 5 & JP 2015-526753 A & KR 10-2015-0044874 A & CN 104620257 A	1-3, 8, 13-14 4-7, 9-12

☒ 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
08 June 2016 (08.06.16)

Date of mailing of the international search report
21 June 2016 (21.06.16)

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/JP2016/063631

C (Continuation). DOCUMENTS CONSIDERED TO BE RELEVANT

Category*	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X A	JP 07-064480 A (Honda Motor Co., Ltd.), 10 March 1995 (10.03.1995), paragraphs [0008] to [0038]; fig. 1 to 3 (Family: none)	1-3, 5, 8, 13-14 4, 6-7, 9-12
X A	JP 2002-132290 A (Kenwood Corp.), 09 May 2002 (09.05.2002), paragraphs [0009] to [0029]; fig. 1 to 4 & WO 2002/035518 A1	1-3, 7-8, 13-14 4-6, 9-12
A	JP 09-114634 A (ATR Interpreting Telecommunications Research Laboratories), 02 May 1997 (02.05.1997), entire text; all drawings (Family: none)	9-12
A	JP 2005-178473 A (Denso Corp.), 07 July 2005 (07.07.2005), entire text; all drawings & US 2005/0134117 A1 entire text; all drawings	9-12

A. 発明の属する分野の分類（国際特許分類（IPC））

Int.Cl. G06F3/16(2006.01)i, G06F3/01(2006.01)i, G10L15/22(2006.01)i, G10L15/24(2013.01)n

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. G06F3/16, G06F3/01, G10L15/22, G10L15/24

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	1971-2016年
日本国実用新案登録公報	1996-2016年
日本国登録実用新案公報	1994-2016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 2015-076877 A（パナソニック インテレクチュアル プロパティ コーポレーション オブ アメリカ）2015.04.20, 段落[0026]-[0068], [0113]-[0122], [0229]-[0246], [図1]-[図4], [図9], [図22]-[図25] & US 2015/0100323 A1, 段落[0061]-[0103], [0148]-[0157], [0273]-[0281], 図1-図4, 図9, 図22-図25	1-6, 8, 13-14 9-12
X A	WO 2013/188002 A1（HONDA MOTOR CO., LTD.）2013.12.19, 段落[0021]-[0073], FIG.1-FIG.5 & JP 2015-526753 A & KR 10-2015-0044874 A & CN 104620257 A	1-3, 8, 13-14 4-7, 9-12

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

08.06.2016

国際調査報告の発送日

21.06.2016

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）

岩橋 龍太郎

5E

3790

電話番号 03-3581-1101 内線 3521

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
X A	JP 07-064480 A (本田技研工業株式会社) 1995. 03. 10, 段落[0008]-[0038], [図 1]-[図 3] (ファミリーなし)	1-3, 5, 8, 13-14 4, 6-7, 9-12
X A	JP 2002-132290 A (株式会社ケンウッド) 2002. 05. 09, 段落[0009]-[0029], [図 1]-[図 4] & WO 2002/035518 A1	1-3, 7-8, 13-14 4-6, 9-12
A	JP 09-114634 A (株式会社エイ・ティ・アール音声翻訳通信研究所) 1997. 05. 02, 全文, 全図 (ファミリーなし)	9-12
A	JP 2005-178473 A (株式会社デンソー) 2005. 07. 07, 全文, 全図 & US 2005/0134117 A1, 全文, 全図	9-12