(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号

特許第4574390号
(P4574390)

(45)発行日　平成22年11月4日(2010.11.4)　(24)登録日　平成22年8月27日(2010.8.27)

(51)Int.Cl.　F I
G10L 15/22 (2006.01)　G10L 15/22　460Z
G10L 15/28 (2006.01)　G10L 15/28　353

請求項の数 3　(全 18 頁)

(21)出願番号　特願2005-45618(P2005-45618)
(22)出願日　平成17年2月22日(2005.2.22)
(65)公開番号　特開2006-234907(P2006-234907A)
(43)公開日　平成18年9月7日(2006.9.7)
審査請求日　平成20年2月20日(2008.2.20)

(73)特許権者　000001007
キヤノン株式会社
東京都大田区下丸子３丁目３０番２号
(74)代理人　100126240
弁理士　阿部　琢磨
(74)代理人　100124442
弁理士　黒岩　創吾
(72)発明者　深田　俊明
東京都大田区下丸子３丁目３０番２号キヤノン株式会社内

審査官　前田　祐希

最終頁に続く

(54)【発明の名称】音声認識方法

(57)【特許請求の範囲】
【請求項１】
第１、第２のボタンを備え、受信した音声を認識してコマンドを得る音声認識装置の音声認識方法であって、
第１の音声を受信する第１の受信工程と、
前記第１の音声を認識し、第１の認識結果として１つまたは２つのコマンドを得る第１の音声認識工程と、
前記第１の認識結果として得られたコマンドを出力する認識結果出力工程と、
前記第１の認識結果が出力された後、第２の音声を受信する第２の受信工程と、
前記認識結果出力工程で１つまたは２つのコマンドが出力され、前記第１、第２のボタンのいずれも押下されなかった場合、前記第１の認識結果として得られたコマンドを正解と確定し、
前記認識結果出力工程で１つまたは２つのコマンドが出力され、前記第１のボタンが２回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを不正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、
前記認識結果出力工程で１つまたは２つのコマンドが出力され、前記第２のボタンが２回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを不正解と確定し、前記第２の音声から第２の認識結果として２つのコマンドを認識し、

前記認識結果出力工程で１つのコマンドが出力され、前記第２のボタンが押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、
前記認識結果出力工程で２つのコマンドが出力され、前記第１のボタンが１回押下され、前記第２の音声が受信されなかった場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを不正解と確定し、２番目のコマンドを正解と確定し、
前記認識結果出力工程で２つのコマンドが出力され、前記第２のボタンが１回押下され、前記第２の音声が受信されなかった場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを正解と確定し、２番目のコマンドを不正解と確定し、
前記認識結果出力工程で２つのコマンドが出力され、前記第１のボタンが１回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを不正解と確定し、２番目のコマンドを正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、
前記認識結果出力工程で２つのコマンドが出力され、前記第２のボタンが１回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを正解と確定し、２番目のコマンドを不正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識する第２の音声認識工程とを備えた音声認識方法。

【請求項２】

請求項１に記載の音声認識方法をコンピュータに実行させるための制御プログラム。

【請求項３】

第１、第２のボタンと、
第１の音声を受信する第１の受信手段と、
前記第１の音声を認識し、第１の認識結果として１つまたは２つのコマンドを得る第１の音声認識手段と、
前記第１の認識結果として得られたコマンドを出力する認識結果出力手段と、
前記第１の認識結果が出力された後、第２の音声を受信する第２の受信手段と、
前記認識結果出力手段によって１つまたは２つのコマンドが出力され、前記第１、第２のボタンのいずれも押下されなかった場合、前記第１の認識結果として得られたコマンドを正解と確定し、
前記認識結果出力手段によって１つまたは２つのコマンドが出力され、前記第１のボタンが２回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを不正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、
前記認識結果出力手段によって１つまたは２つのコマンドが出力され、前記第２のボタンが２回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを不正解と確定し、前記第２の音声から第２の認識結果として２つのコマンドを認識し、
前記認識結果出力手段によって１つのコマンドが出力され、前記第２のボタンが押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、
前記認識結果出力手段によって２つのコマンドが出力され、前記第１のボタンが１回押下され、前記第２の音声が受信されなかった場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを不正解と確定し、２番目のコマンドを正解と確定し、
前記認識結果出力手段によって２つのコマンドが出力され、前記第２のボタンが１回押下され、前記第２の音声が受信されなかった場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを正解と確定し、２番目のコマンドを不正解と確定し、
前記認識結果出力手段によって２つのコマンドが出力され、前記第１のボタンが１回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを不正解と確定し、２番目のコマンドを正解と確定し、前記第２

の音声から第２の認識結果として１つのコマンドを認識し、
前記認識結果出力手段によって２つのコマンドが出力され、前記第２のボタンが１回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを正解と確定し、２番目のコマンドを不正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識する第２の音声認識手段とを備えた音声認識装置。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、音声認識の結果の修正を簡便な操作で実現しうる方法に関するものである。

【背景技術】

【０００２】

連続音声認識を実用化する際の重要な課題の一つとして、簡便な操作による誤認識の訂正がある。例えば、機器操作における複数コマンドの設定は、複数コマンドを連続音声入力することにより可能となるが、ここで、「Ａ、Ｂ」という２つのコマンドを発声した際に、「Ｃ、Ｂ」や、「Ａ、Ｂ、Ｃ」という誤った認識結果が得られた際に、どのようにしてＣの部分を指示し、言い直すもしくは削除するかという課題である。こういった訂正は、視覚障害者や視覚が利用できない利用者に対して、あるいは画面表示が行えない機器を利用する場合には、とりわけ大きな困難を伴う。

【０００３】

この課題に対して、簡便な方法によって音声認識の結果を修正する方法がいくつか開示されている。特許文献１では、入力ボタンとは別の修正ボタンを用意することにより、発声が過去の発声の修正か、新たな発話の認識かが判断できるようにしている。この方法では、修正位置の指定は利用者ではなく装置側が行っているため、修正すべき部分の同定を誤ってしまうという問題がある。また、修正ボタンを用いずに、音声によって修正コマンドを入力する方法（例えば、「違う、会議」の「違う」の部分が修正コマンド）も開示されているが、修正コマンド自体が誤認識してしまうという問題がある。

【０００４】

また、特許文献２では、認識結果を認識単位区切りで表示し、例えば「Ｆ５」を押下すると５番目の単位の修正候補（Ｎベスト）が表示されるという方法が開示されている。しかしながら、この方法は、認識誤りとして置換誤りのみを扱っており、挿入誤りや脱落誤りの修正ができないといった問題がある。また、認識結果は修正候補を表示し、この中から選択する、もしくはこれらの修正候補を音声で読み上げ、正解がある場合にこれを指示するという方法を用いているため、視覚障害者に対しては必ずしも使い勝手のよい方法であるとは言えない。

【０００５】

また、特許文献３では、認識結果の文字列（ひらがな）の各文字に異なる符号（数字）を付けて表示し、利用者が符号を指定して修正用音声を発声して置換する方法が開示されている。しかしながら、この方法も認識誤りとして置換誤りのみを扱っており、挿入誤りや脱落誤りの修正ができないといった問題がある。また、訂正単位が文字単位であるため、単語を訂正する場合には訂正に時間を要するため操作性が悪いといった問題がある。さらに、認識結果は表示装置によって利用者に提示されるため、視覚障害者は操作できないといった問題もある。

【特許文献１】特開平１１－３３８４９３号公報
【特許文献２】特開２０００－２５９１７８号公報
【特許文献３】特開２００４－９３６９８号公報

【発明の開示】

【発明が解決しようとする課題】

【０００６】

本発明は上述の問題を鑑みてなされたもので、連続音声認識の認識結果の出力に対して

、利用者は誤認識の位置を物理ボタンを用いて指示することによって、視覚障害者や視覚が利用できないユーザに対して、あるいは画面表示が行えない機器を利用する場合においても、簡便な操作で認識結果の修正が行える手段を提供することを目的としている。ここで、連続音声認識結果としては、置換誤りの他、脱落、挿入誤りも生じ得るため、これら全ての誤りに対して、統一的な操作感で修正を行える手段を提供することも目的としている。

【課題を解決するための手段】

【０００７】

上記目的を達成するために、本発明は以下のような構成を備える。すなわち、第１、第２のボタンを備え、受信した音声を認識してコマンドを得る音声認識装置の音声認識方法であって、第１の音声を受信する第１の受信工程と、前記第１の音声を認識し、第１の認識結果として１つまたは２つのコマンドを得る第１の音声認識工程と、前記第１の認識結果として得られたコマンドを出力する認識結果出力工程と、前記第１の認識結果が出力された後、第２の音声を受信する第２の受信工程と、前記認識結果出力工程で１つまたは２つのコマンドが出力され、前記第１、第２のボタンのいずれも押下されなかった場合、前記第１の認識結果として得られたコマンドを正解と確定し、前記認識結果出力工程で１つまたは２つのコマンドが出力され、前記第１のボタンが２回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを不正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、前記認識結果出力工程で１つまたは２つのコマンドが出力され、前記第２のボタンが２回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを不正解と確定し、前記第２の音声から第２の認識結果として２つのコマンドを認識し、前記認識結果出力工程で１つのコマンドが出力され、前記第２のボタンが押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドを正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、前記認識結果出力工程で２つのコマンドが出力され、前記第１のボタンが１回押下され、前記第２の音声が受信されなかった場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを不正解と確定し、２番目のコマンドを正解と確定し、前記認識結果出力工程で２つのコマンドが出力され、前記第２のボタンが１回押下され、前記第２の音声が受信されなかった場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを正解と確定し、２番目のコマンドを不正解と確定し、前記認識結果出力工程で２つのコマンドが出力され、前記第１のボタンが１回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを不正解と確定し、２番目のコマンドを正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識し、前記認識結果出力工程で２つのコマンドが出力され、前記第２のボタンが１回押下され、前記第２の音声が受信された場合、前記第１の認識結果として得られたコマンドのうち１番目のコマンドを正解と確定し、２番目のコマンドを不正解と確定し、前記第２の音声から第２の認識結果として１つのコマンドを認識する第２の音声認識工程とを備える。

【発明の効果】

【００１１】

本発明によれば、簡便な操作によって、連続音声認識の誤認識を訂正する手段が提供できる。

【発明を実施するための最良の形態】

【００１２】

以下、図面を参照しながら本発明の好適な実施例について説明していく。

【実施例１】

【００１３】

図１は、本発明の第１の実施形態に係る音声認識装置の構成を示すブロック図である。１０１はＣＰＵで、ＲＯＭ１０２に記憶された制御プログラム或いは外部記憶装置１０４からＲＡＭ１０３にロードされた制御プログラムに従って、本実施形態の音声認識装置に

おける各種制御を行う。ＲＯＭ１０２は各種パラメータやＣＰＵ１０１が実行する制御プログラムなどを格納している。ＲＡＭ１０３は、ＣＰＵ１０１による各種制御の実行時に作業領域を提供するとともに、ＣＰＵ１０１により実行される制御プログラムを記憶する。１０４はハードディスク、フロッピー（登録商標）ディスク、ＣＤ－ＲＯＭ、ＤＶＤ－ＲＯＭ、メモリカード等の外部記憶装置で、この外部記憶装置がハードディスクの場合には、ＣＤ－ＲＯＭやフロッピー（登録商標）ディスク等からインストールされた各種プログラムが記憶されている。１０５はマイクロフォンなどによる音声入力装置であり、取り込まれた音声に対して音声認識が実行される。１０６はＣＲＴ、液晶ディスプレイなどの表示装置であり、処理内容の設定・入力に関する表示・出力を行う。１０７はボタン、テンキー、キーボード、マウス、ペンなどの補助入力装置であり、これらの入力装置を用いて利用者が発声する音声の取り込みを開始するための指示を与える。１０８はスピーカなどの補助出力装置であり、音声認識結果を音声で確認する場合などに用いる。１０９は上記各部を接続するバスである。

【００１４】

図２は、音声認識結果修正方法のモジュール構成を示したブロック図である。２０１は音声入力部であり、１０５から音声信号を受信する。２０２は２０１で入力された音声を認識する音声認識部であり、入力音声の分析、参照パターンとの距離計算、探索処理などを行う。２０３は認識結果出力部であり、２０２で認識された結果を１０６もしくは１０８に出力し利用者に対して出力する。２０４は認識結果修正部であり、２０３で出力された認識結果の中に含まれる正解部分を１０７で指定した後、認識誤りに対する言い直しを音声によって行い１０５から入力する。

【００１５】

図３は、１発声で２つまでのコマンドを同時に認識可能な場合の入力されるコマンド（音声入力コマンド）と出力されるコマンド（認識コマンド）の正誤の全組み合わせを示す図である。図中のＣは正解（Ｃｏｒｒｅｃｔ）、Ｓは置換誤り（Ｓｕｂｓｔｉｔｕｔｉｏｎ）、Ｄは脱落誤り（Ｄｅｌｅｔｉｏｎ）、Ｉは挿入誤り（Ｉｎｓｅｒｔｉｏｎ）を表わし、例えば、（Ｃ，Ｓ）は２０３で認識結果を２つ出力し、そのうちの一つが正解であり、もう一方は置換誤りであったことを表わしている。ここで、最初のコマンドが正解であるか、２番目のコマンドが正解であるかは、この図の表記上は区別していない。いま、コピー操作のコマンド入力を音声によって行うタスクを考える。認識対象語彙は、出力用紙サイズに関するコマンド（Ａ４、Ａ３、Ｂ４、Ｂ５の４単語）と部数に関するコマンド（１部から１００部）であるとする。また、同時に２コマンドまで（１コマンドもしくは２コマンド）の認識が可能であるとする。また、コマンドの発声順序は自由であるとする。この場合の発声例としては、「Ａ４、５部」、「８０部、Ｂ５」、「４部」、「Ａ３」などとなる。なお、出力用紙サイズもしくは部数に関する入力が行われなかった場合には、デフォルト値（例えば、出力用紙サイズは「自動」、部数は「１部」）が設定される。この場合、「Ａ４、５部」という音声入力（音声入力コマンド数２）に対して、「Ａ４、１５部」という認識結果（認識コマンド数２）が得られた場合、「５部」が「１５部」に誤っている（置換誤り）ため、図３の正誤パターンの（Ｃ，Ｓ）に当たる。同様に、「Ａ４、１５部」（音声入力コマンド数２）という音声入力に対して、「Ａ４」という認識結果（認識コマンド数１）が得られた場合、「１５部」が認識されない（脱落誤り）であるため、図３の正誤パターンの（Ｃ，Ｄ）に当たる。また、「Ａ４」（音声入力コマンド数１）という音声入力に対して、「Ａ４、４部」という認識結果（認識コマンド数２）が得られた場合、「４部」が余計に認識された（挿入誤り）であるため、図３の正誤パターンの（Ｃ，Ｉ）に当たる。本実施例では、図３で示される全ての組み合わせに対して、物理キーを用いて利用者が正解部分を指定することによって正解部分を確定する。図４は、この指定に用いる物理キーの一例であり、一般的な数字キーである。

【００１６】

図５は、図３の組み合わせに対して認識結果の正解部分を指定する際の物理キーの押下例を示した図である。「（Ｃ）：１」は、音声入力コマンド数と認識コマンド数がともに

１で、かつ正解（Ｃ）であった場合に、数字キーの「１」を押下することを表わしている。この「１」は、認識結果として出力される「１番目」の認識コマンドが正解であるという意味である。同様に、「（Ｃ，Ｃ）：１，２」は、音声入力コマンド数と認識コマンド数がともに２で、かつともに正解であった場合に、「１番目」および「２番目」の認識コマンドが正解であるため、数字キーの「１」と「２」を押下する。

【００１７】

また、「（Ｃ，Ｉ）：ｍ」は、前述の「Ａ４」（音声入力コマンド数１）という音声入力に対して、「Ａ４、４部」という認識結果（認識コマンド数２）が得られた例に当たり、この例では、「１番目」の認識コマンドが正解であるため、「１」を押下する（ｍ＝１）。なお、仮に「４部、Ａ４」という認識結果が得られたとすると、「２番目」の認識コマンドが正解であるため、「２」を押下する（ｍ＝２）。このように、ｍは１または２のいずれかの値を取る。

【００１８】

また、「（Ｓ）：Ｒ」は、音声入力コマンド数と認識コマンド数がともに１で、かつ誤り（Ｓ）であった場合である。この場合は、正解がないため、正解部分の指定は行わず、認識誤りを音声で言い直すための再発声Ｒ（Ｒｅｓｐｅａｋ）を行う。ここで、再発声する場合には、何らかのボタンを押下した後に発声を行ってもよいし、ボタンの押下なく発声を開始してもよい。同様に、「（Ｓ，Ｄ）：Ｒ」、「（Ｓ，Ｉ）：Ｒ」、「（Ｓ，Ｓ）：Ｒ」の場合も正解がないため、正解部分の指定は行わず、認識誤りを音声で言い直すための再発声Ｒを行う。

【００１９】

また、「（Ｃ，Ｓ）：ｍ，Ｒ」は、前述の「Ａ４、５部」という音声入力（音声入力コマンド数２）に対して、「Ａ４、１５部」という認識結果（認識コマンド数２）が得られた例に当たり、この例では、「１番目」の認識コマンドが正解であるため、「１」を押下し（ｍ＝１）、その後、再発声Ｒを行う。なお、仮に「Ｂ４、５部」という認識結果が得られたとすると、「２番目」の認識コマンドが正解であるため、「２」を押下し（ｍ＝２）、その後、再発声Ｒを行う。このように、ｍは１または２のいずれかの値を取る。

【００２０】

また、「（Ｃ，Ｄ）：１，Ｒ」は、前述の「Ａ４、１５部」という音声入力（音声入力コマンド数２）に対して、「Ａ４」という認識結果（認識コマンド数１）が得られた例に当たり、この例では、「１番目」の認識コマンドが正解であるため、「１」を押下し、その後、再発声Ｒを行う。

【００２１】

図６は、認識結果の正解部分を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートであり、この図を用いて全体の処理を更に詳細に説明する。まず、Ｓ３０１で音声入力を行う。次に、Ｓ３０２では、Ｓ３０１で入力された音声を音声分析し、音声の特徴パラメータを求めた後、３１０の認識文法もしくは言語モデルに基づいて探索処理を行う（その他、音響モデルや発音辞書なども用いるが図には示していない）。Ｓ３０３では、Ｓ３０２で認識された結果を利用者に対して提示する。提示の方法の例としては、１０６の表示装置を用いた画面表示や１０８の補助出力装置としてスピーカを用いた音声出力などがある。音声出力は、認識結果の文字情報（表記や読みなど）を音声合成することによって実現できる。ここで、正解部分が何番目であるかを利用者が正確に指定するためには、認識対象単位を正確に利用者へ伝える必要がある。具体的には、「Ａ４、４部」という結果に対して、「Ａ４」が１番目、「４部」が２番目であるという提示である。画面表示を行う場合には、認識対象単位の区切りが分かるように、「、」などの区切り記号を挿入して表示する、１つのボックス（矩形窓）の中に１つの認識対象単位を入れるなどの方法を用いればよい。また、音声出力を行う場合には、区切りが分かるような聴覚信号を挿入すればよい。聴覚信号の例としては、無音（認識対象後間に無音区間を挿入する）、「ピッ」といった報知音、「１．（イチ）Ａ４、２．（ニ）４部」といった数字の読み上げ音声などがある。これらの手段で認識対象単位を利用者へ伝えることによって、

例えば、拡大・縮小のコマンドとして、「Ａ４からＢ５」があった場合に、「Ａ４」と「Ｂ５」が別々なのか、「Ａ４からＢ５」で一つのいずれであるかを利用者へ正確に伝えることが可能となる。

【００２２】

次に、Ｓ３０４において、正解部分を指定するキー入力が行われるか否かを判定する。キー入力がある場合、すなわち、（Ｃ）、（Ｃ，Ｉ）、（Ｃ，Ｄ）、（Ｃ，Ｃ）、（Ｃ，Ｓ）の場合、Ｓ３０５で再発声が行われるか否かを判定する。ここで再発声がある場合、すなわち、（Ｃ，Ｄ）、（Ｃ，Ｓ）の場合、Ｓ３０６において、正解部分の認識結果を確定する。ここで、（Ｃ，Ｄ）の場合は、利用者は２つのコマンドを入力し、そのうち１つは正解で、もう１つは認識結果として出力されていなかったことが分かる。同様に、（Ｃ，Ｓ）の場合は、利用者は２つのコマンドを入力し、そのうち１つは正解で、もう１つは誤っていたことが分かる。つまり、これらの場合の再発声は、１つのコマンドが発声されると期待できる。また、例えば、コピー部数が正解であれば、再発声は出力用紙サイズに関するものであることも期待できる。すなわち、これらの場合には、再発声の認識を行う際に、２コマンドまでの連続発声を認識する音声認識を行うのではなく、出力用紙サイズのみに関する１コマンドの音声認識を行えばよい。つまり、再発声の音声認識を行う際に、制約を追加することが可能となる。Ｓ３０７はこのような認識制約の追加を行う処理であり、具体的には、再発声の音声を認識する際に３１０の認識文法や言語モデルに制約をかけてＳ３０１に戻る（もしくは、Ｓ３０３で再発声の音声認識結果から制約を満たす結果のみを出力するといった処理を行うことも可能である）。なお、キー入力の有無、もしくは再発声の有無は、タイマを用いて所定の時間内にこれらのイベント入力があるか否かで判定することができる。Ｓ３０５で再発声がないと判定された場合、すなわち、（Ｃ）、（Ｃ，Ｉ）、（Ｃ，Ｃ）の場合（あるいは（Ｃ，Ｄ）もしくは（Ｃ，Ｓ）でタイムアウトとなった場合）、正解部分は確定されているため、Ｓ３０９で正解部分を確定し、処理を終える。また、Ｓ３０４でキー入力が無い場合には、次にＳ３０８で再発声が行われるか否かを判定する。ここで再発声が無いと判定された場合（これは図５のいずれにも当たらない）、何も確定せずに処理を終了する。また、Ｓ３０８で再発声が行われた場合、すなわち、（Ｓ）、（Ｓ，Ｉ）、（Ｓ，Ｄ）、（Ｓ，Ｓ）の場合、正解部分が何も指定されていないため、Ｓ３０７で行ったような認識制約追加は行えないため、そのままＳ３０１へ戻る。

【００２３】

前述の実施例では、１発声で２つまでのコマンドを同時に認識可能な場合の正誤の全組み合わせについて述べたが、本発明はこれに限らず、任意のコマンド数に対して適用することができる。図１４は、１発声で３つまでのコマンドを同時に認識可能な場合の入力されるコマンド（音声入力コマンド）と出力されるコマンド（認識コマンド）の正誤の全組み合わせを示す図である。図中のＣ、Ｓ、Ｄ、Ｉは図５と同じ意味である。この図において、例えば、（Ｃ，Ｓ，Ｉ）は、２つの音声入力コマンドに対して認識結果が３つ出力され、そのうちの１つが正解であり、残りの２つは誤り（１つは置換誤り、もう１つは挿入誤り）であったことを表わしている。図５の場合と同様に、これらの表記は、組み合わせを示しており、順序は区別していない。

【００２４】

図１５は、図１４の組み合わせに対して認識結果の正解部分を指定する際の物理キーの押下例を示した図である。（音声入力コマンド数，認識コマンド数）のペアが、（１，１）、（１，２）、（２，１）、（２，２）の部分は、前述の図５と全く同じであるため説明は省略する。また、残りのペアについても図５の場合と同様であるが、図中のｊ、ｋは１から３の値を取り、また、ｊとｋは異なる値を取る（ｊ！＝ｋ）。例えば、（Ｃ，Ｉ，Ｉ）は、音声入力コマンド数が１で認識コマンド数が３である場合で、かつ音声入力コマンドは正解であった場合である。この場合、出力される１番目から３番目の中のいずれかが正解であるため、「１番目」の場合は「１」（ｊ＝１）を、「２番目」の場合は「２」（ｊ＝２）を、「３番目」の場合は「３」（ｊ＝３）を押下する。このように、ｊは１か

ら３のいずれかの値を取る。また、（Ｃ，Ｃ，Ｓ）は、音声入力コマンド数と認識コマンド数がともに３であったときに、２つは正解であり、１つは置換誤りであった場合である。この場合、出力される１番目から３番目の中の２箇所が正解であるため、その２箇所ｊ、ｋ（ｊ，ｋ＝｛１，２，３｝，ｊ！＝ｋ）を押下する。

【００２５】

以上のような構成をとることにより、簡便かつ統一的な操作によって、連続音声認識の誤認識を訂正する手段が提供でき、視覚障害者や視覚が利用できないユーザに対して、あるいは画面表示が行えない機器に対して、実用に耐え得る音声認識装置を提供することが可能となる。

【実施例２】

【００２６】

前記実施例では、図３もしくは図１４の組み合わせに対して認識結果の正解部分を指定していたが、誤り部分を指定してもよい。図７は、図３の組み合わせに対して認識結果の誤り部分を指定する際の物理キーの押下例を示した図である。ここで、Ｎ／Ａは、全てが正解で誤りがないため、誤り部分を指定する必要がないことを示している。その他については、図５と同様であり、違いは正解部分の代わりに誤り部分を指定する。

【００２７】

図８は、認識結果の誤り部分を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートであり、この図を用いて全体の処理を更に詳細に説明する。ここで、Ｓ４０１～Ｓ４０３はＳ３０１～Ｓ３０３と、４１３は３１０と同じであるため説明は省略する。Ｓ４０４において、誤り部分を指定するキー入力が行われるか否かを判定する。キー入力がある場合、すなわち、（Ｓ）、（Ｃ，Ｉ）、（Ｓ，Ｉ）、（Ｓ，Ｄ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）の場合、Ｓ４０５で再発声が行われるか否かを判定する。ここで再発声がある場合、すなわち、（Ｓ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）の場合、Ｓ４０６において、正解部分が確定できる場合に関して、すなわち、（Ｃ，Ｓ）のＣに対して認識結果を確定する（その他の場合は確定処理を行わない）。ここで、（Ｃ，Ｓ）の場合は、利用者は２つのコマンドを入力し、そのうち１つは正解で、もう１つは置換誤りであることが分かる。つまり、これらの場合の再発声は、１つのコマンドが発声されると期待できる。よって、前記実施例におけるＳ３０７と同様、再発声の音声認識を行う際に、制約を追加することが可能となる。Ｓ４０７はこのような認識制約の追加を行う処理であり、具体的には、再発声の音声を認識する際に４１３の認識文法や言語モデルに制約をかけてＳ４０１に戻る（もしくは、Ｓ４０３で再発声の音声認識結果から制約を満たす結果のみを出力するといった処理を行うことも可能である）。ここで、制約がかけられない場合は認識制約追加処理を行わない。なお、キー入力の有無、もしくは再発声の有無の判定は、前記実施例と同様にすればよい。Ｓ４０５で再発声がないと判定された場合、すなわち、（Ｃ，Ｉ）の場合（あるいは、（Ｓ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）でタイムアウトとなった場合）、正解部分が確定できるものについては、Ｓ４０９で正解部分を確定し、処理を終える。また、Ｓ４０４でキー入力が無い場合には、次にＳ４０８で再発声が行われるか否かを判定する。ここで再発声が無いと判定された場合、すなわち、（Ｃ）、（Ｃ，Ｃ）の場合、Ｓ４１２で認識結果を正解と確定して処理を終了する。また、Ｓ４０８で再発声が行われた場合、すなわち、（Ｃ，Ｄ）の場合、Ｓ４１０で認識結果を正解と確定し、Ｓ４１１で認識制約追加を行い、Ｓ４０１へ戻る。

【００２８】

前述の実施例では、１発声で２つまでのコマンドを同時に認識可能な場合の正誤の全組み合わせについて述べたが、前記実施例と同様に、任意のコマンド数に対して適用することができる。

【００２９】

図１６は、図１４の組み合わせに対して認識結果の誤り部分を指定する際の物理キーの押下例を示した図である。（音声入力コマンド数，認識コマンド数）のペアが、（１，１）、（１，２）、（２，１）、（２，２）の部分は、前述の図７と全く同じであるため説

明は省略する。また、残りのペアについても図７の場合と同様であるが、図中のｊ、ｋは、図１５と同じであり、１から３の値を取り、また、ｊとｋは異なる値を取る（ｊ！＝ｋ）。

【実施例３】

【００３０】

前記実施例では、図３もしくは図１４の組み合わせに対して認識結果の正解部分もしくは誤り部分を指定していたが、全ての認識結果に対してそれぞれ正誤を指定してもよい。正誤の指定は様々な方法が考えられるが、以下の例では、正解の場合には「１」を、誤りの場合には「２」を押下する場合について説明する。図９は、図３の組み合わせに対して認識結果の正誤を指定する際の物理キーの押下例を示す図である。

【００３１】

「（Ｃ）：１」は、音声入力コマンド数と認識コマンド数がともに１で、かつ正解（Ｃ）であった場合に、数字キーの「１」を押下することを表わしている。この「１」は、認識結果として出力される認識コマンドが「正解」であるという意味である。同様に、「（Ｃ，Ｃ）：１，１」は、音声入力コマンド数と認識コマンド数がともに２で、かつともに正解であった場合に、１番目および２番目の認識コマンドが「ともに正解」であるため「１」と「１」を押下する。

【００３２】

また、「（Ｓ）：２，Ｒ」は、音声入力コマンド数と認識コマンド数がともに１で、かつ誤り（Ｓ）であった場合である。この場合は、誤りであるため「２」を押下した後、認識誤りを音声で言い直すための再発声Ｒを行う。同様に、「（Ｓ，Ｄ）：２，Ｒ」、「（Ｓ，Ｉ）：２，２，Ｒ」、「（Ｓ，Ｓ）：２，２，Ｒ」の場合も正解がないため、認識結果に対する認識誤りの回数だけ「２」を押下した後、再発声Ｒを行う。

【００３３】

また、「（Ｃ，Ｄ）：１，Ｒ」は、音声入力コマンド数が２で、認識コマンド数が１で、１つは正解で、もう１つは脱落誤り（Ｄ）であった場合である。この場合は、認識コマンドとして出力される結果は正解であるため「１」を押下した後、脱落誤りとなったコマンドを入力するために再発声Ｒを行う。

【００３４】

また、「（Ｃ，Ｉ）：１，２」は、音声入力コマンド数が１で、認識コマンド数が２の場合で、１つは正解で、もう１つは挿入誤り（Ｉ）であった場合である。この場合は、Ｃに対応する部分は正解であるため「１」を押下し、挿入誤りに対応する部分は誤りであるため「２」を押下する。なお、「１」と「２」の押下順序は、結果出力の順序に従うとする。つまり、１番目が正解（Ｃ）、２番目が挿入誤り（Ｉ）の場合は、「１」、「２」の順で押下し、１番目が挿入誤り（Ｉ）、２番目が正解（Ｃ）の場合は、「２」、「１」の順で押下する。同様に、「（Ｃ，Ｓ）：１，２，Ｒ」は、正解部分に対して「１」を、置換誤り部分に対して「２」を押下した後、置換誤りとなったコマンドを入力するために再発声Ｒを行う。

【００３５】

図１０は、認識結果の正誤を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートであり、この図を用いて全体の処理を更に詳細に説明する。ここで、Ｓ５０１～Ｓ５０３はＳ３０１～Ｓ３０３と、５０９は３１０と同じであるため説明は省略する。Ｓ５０４において、正誤を指定するキー入力の取り込みを行う。次に、Ｓ５０５で再発声が行われるか否かを判定する。ここで再発声がある場合、すなわち、（Ｓ）、（Ｃ，Ｄ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）の場合、Ｓ５０６において、正解部分の認識結果を確定する。ここで、例えば、（Ｃ，Ｄ）の場合は、利用者は２つのコマンドを入力し、そのうち１つは正解で、もう１つは脱落誤りであることが分かる。つまり、これらの場合の再発声は、１つのコマンドが発声されると期待できる。よって、前記実施例におけるＳ３０７と同様、再発声の音声認識を行う際に、制約を追加することが可能となる。Ｓ５０７はこのような認識制約の追加を行う処理であり、具体的には、再発声

の音声を認識する際に５０９の認識文法や言語モデルに制約をかけてＳ５０１に戻る（もしくは、Ｓ５０３で再発声の音声認識結果から制約を満たす結果のみを出力するといった処理を行うことも可能である）。ここで、制約がかけられない場合は認識制約追加処理を行わない。なお、再発声の有無の判定は、前記実施例と同様にすればよい。Ｓ５０５で再発声がないと判定された場合、すなわち、（Ｃ）、（Ｃ，Ｉ）、（Ｃ，Ｃ）の場合（あるいは、（Ｓ）、（Ｃ，Ｄ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）でタイムアウトとなった場合）、正解部分が確定できるものについては、Ｓ５０８で正解部分を確定し、処理を終える。

【００３６】

前述の実施例では、認識結果を全て出力した後、正誤の指定を行う方法について述べたが、認識対象単位ごとに１つずつ結果を出力し、逐次正誤を指定していくこともできる。図１１は、認識結果の正誤を認識単位ごとに逐次的に指定する際の音声認識結果修正方法の全体の処理を示したフローチャートである。ここで、Ｓ６０１、Ｓ６０２、Ｓ６１２、Ｓ６０８～Ｓ６１１は、それぞれＳ５０１、Ｓ５０２、Ｓ５０９、Ｓ５０５～Ｓ５０８と同じであるため説明は省略する。Ｓ６０３では、Ｓ６０２で得られる認識結果から認識単位の結果数をＮに、カウンタｉを１にセットする。次に、Ｓ６０４で、ｉ番目の認識結果を出力する。次にＳ６０５でキー入力の取り込み（前記実施例では、正解の場合は「１」、誤りの場合は「２」のいずれか１つ）を行う。次に、Ｓ６０６でカウンタｉに１を加える。Ｓ６０７でｉがＮ以下であるかを判定し、Ｎ以下の場合にはＳ６０４へ戻り、Ｎより大きい場合にはＳ６０８へ進む。

【００３７】

前述の実施例では、１発声で２つまでのコマンドを同時に認識可能な場合の正誤の全組み合わせについて述べたが、前記実施例と同様に、任意のコマンド数に対して適用することができる。

【００３８】

図１７は、図１４の組み合わせに対して認識結果の正誤を指定する際の物理キーの押下例を示した図である。（音声入力コマンド数，認識コマンド数）のペアが、（１，１）、（１，２）、（２，１）、（２，２）の部分は、前述の図９と全く同じであり、残りのペアについても図９の場合と同様である。

【実施例４】

【００３９】

前記実施例２では、図３もしくは図１４の組み合わせに対して認識結果の誤り部分を指定していたが、例えば、図７において、「１，Ｒ」は、出力された１つの認識結果が誤りであることは判定できるが、入力された音声コマンド数は１つであるか、２つであるかは分からない。すなわち、認識誤りの組み合わせが、（Ｓ）であるか（Ｓ，Ｄ）であるかの区別をすることができない。同様に、「１，２，Ｒ」の場合も（Ｓ，Ｉ）か（Ｓ，Ｓ）かの区別ができない。よって、これらの場合には、再発声を認識する際に、何の制約もかけることができないため、同様の誤りを生じる可能性があり、なかなか正解が得られない場合がある。

【００４０】

本実施例は、このような問題を鑑みてなされたもので、認識結果の誤り部分に加え、誤りの種類を直接的あるいは間接的な方法で指定することによって、全ての組み合わせに対して、再発声を認識する際に、制約をかけることを可能とするものである。

【００４１】

いま、以下に示すような物理キーの押下規則を適用することを考える。すなわち、音声入力コマンドに対する認識コマンドが、全て誤りの場合は発声単語数を２回押下し（規則１）、誤りはないが正解が不足している場合は追加対象となる位置を押下し（規則２）、音声入力コマンドは全てもしくは一部認識されたが誤りも含まれている場合は誤り部分の認識コマンド位置を押下する（規則３）。これらの規則を図３の組み合わせに対して適用すると、図１２のようになる（Ｎ／Ａは、全てが正解で誤りがないため、誤り部分を指定

する必要がないことを示している）。このとき、（Ｓ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｓ，Ｓ）の押下例は規則１が、（Ｃ，Ｄ）は規則２が、（Ｃ，Ｉ）、（Ｃ，Ｓ）は規則３がそれぞれ適用される。ここで、（Ｃ，Ｉ）：ｍは、認識コマンドの１番目が挿入誤りである場合は「１」を押下し（ｍ＝１）、２番目が挿入誤りである場合は「２」を押下する（ｍ＝２）ことを表している。同様に、（Ｃ，Ｓ）：ｍ，Ｒは、認識コマンドの１番目が置換誤りである場合は「１」を押下し（ｍ＝１）、２番目が置換誤りである場合は「２」を押下した後（ｍ＝２）、再発声を行うことを表している。このようなキー押下を適用すれば、誤りの部分が特定できることに加え、同じ認識コマンド数の組み合わせにおけるボタン押下のパタンが全て異なるため、図１２のいずれの誤りパタンであるかが一意に同定できる。すなわち、図１２に示したボタン押下を用いれば、誤り部分と誤りの種類（置換誤り、挿入誤り、脱落誤り）が直接的もしくは間接的に指定されることになる。このような指定方法を用いれば、再発声時に常に認識に制約をかけられるため、再発声が正しく認識される可能性を高めることができる。

【００４２】

図１３は、認識結果の誤り部分と誤りの種類を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートであり、この図を用いて全体の処理を更に詳細に説明する。ここで、Ｓ７０１～Ｓ７０３はＳ３０１～Ｓ３０３と、７１０は３１０と同じであるため説明は省略する。Ｓ７０４において、誤り部分と誤りの種類を指定するキー入力が行われるか否かを判定する。キー入力がある場合、すなわち、（Ｃ）、（Ｃ，Ｃ）以外の場合、Ｓ７０５で再発声が行われるか否かを判定する。ここで再発声がある場合、すなわち、（Ｓ）、（Ｃ，Ｄ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）の場合、Ｓ７０６において、正解部分が確定できる場合に関して、すなわち、（Ｃ，Ｄ）、（Ｃ，Ｓ）のＣに対して認識結果を確定する（その他の場合は確定処理を行わない）。ここで、再発声の音声入力コマンド数は、（Ｓ）、（Ｃ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）の場合は１、（Ｓ，Ｄ）、（Ｓ，Ｓ）の場合は２であると確定することが可能である。よって、再発声の音声認識を行う際に、これらの制約を追加することが可能となる。Ｓ７０７はこのような認識制約の追加を行う処理であり、具体的には、再発声の音声を認識する際に７１０の認識文法や言語モデルに制約をかけてＳ７０１に戻る（もしくは、Ｓ７０３で再発声の音声認識結果から制約を満たす結果のみを出力するといった処理を行うことも可能である）。なお、キー入力の有無、もしくは再発声の有無の判定は、前記実施例と同様にすればよい。Ｓ７０５で再発声がないと判定された場合、すなわち、（Ｃ，Ｉ）の場合（あるいは、（Ｓ）、（Ｃ，Ｄ）、（Ｓ，Ｄ）、（Ｓ，Ｉ）、（Ｃ，Ｓ）、（Ｓ，Ｓ）でタイムアウトとなった場合）、正解部分が確定できるものについては、Ｓ７０８で正解部分を確定し、処理を終える。また、Ｓ７０４でキー入力が無い場合、すなわち、（Ｃ）、（Ｃ，Ｃ）の場合、Ｓ７０９で認識結果を正解と確定して処理を終了する。

【００４３】

前述の実施例では、１発声で２つまでのコマンドを同時に認識可能な場合の正誤の全組み合わせについて述べたが、前記実施例と同様に、任意のコマンド数に対して適用することができる。図１８は、図１４の組み合わせに対して認識結果の誤り部分と誤りの種類を指定する際の物理キーの押下例を示す図である。（音声入力コマンド数，認識コマンド数）のペアが、（１，１）、（１，２）、（２，１）、（２，２）の部分は、前述の図１２と全く同じであるため説明は省略する。また、残りのペアについても前述の規則１～規則３を適用したキー押下のパタンとなっているが、正解と２種類の誤りが混在する場合、すなわち（Ｃ，Ｓ，Ｄ）および（Ｃ，Ｓ，Ｉ）の場合は（他に（Ｃ，Ｄ，Ｉ）も考えられるが、これは（Ｃ，Ｓ）と見なす）、規則３を適用することも可能であるが、図１８のいずれの誤りパタンであるかを一意に同定するために、以下の規則３の変形規則を用いる。すなわち、音声入力コマンドは正解と誤りが混在して、音声入力コマンド数よりも認識コマンド数が少ない場合は誤り部分の認識コマンド位置に続いて３を押下する（規則３－１）。また、音声入力コマンドは正解と誤りが混在して、音声入力コマンド数よりも認識コマンド数が多い場合は誤り部分の認識コマンド位置に続いて３を押下する（規則３－２）。

図中のｊ、ｋは、図１５と同じであり、１から３の値を取り、また、ｊとｋは異なる値を取る（ｊ！＝ｋ）。

【００４４】

なお、本発明の目的は、前述した実施例の機能を実現するソフトウェアのプログラムコードを記録した記憶媒体を、システムあるいは装置に供給し、そのシステムあるいは装置のコンピュータ（またはＣＰＵやＭＰＵ）が記憶媒体に格納されたプログラムコードを読み出し実行することによっても達成されることは言うまでもない。

【００４５】

この場合、記憶媒体から読み出されたプログラムコード自体が前述した実施形態の機能を実現することになり、そのプログラムコードを記憶した記憶媒体は本発明を構成することになる。

【００４６】

プログラムコードを供給するための記憶媒体としては、例えば、フレキシブルディスク、ハードディスク、光ディスク、光磁気ディスク、ＣＤ－ＲＯＭ、ＣＤ－Ｒ、磁気テープ、不揮発性のメモリカード、ＲＯＭなどを用いることができる。

【００４７】

また、コンピュータが読出したプログラムコードを実行することにより、前述した実施形態の機能が実現されるだけでなく、そのプログラムコードの指示に基づき、コンピュータ上で稼働しているＯＳ（オペレーティングシステム）などが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【００４８】

さらに、記憶媒体から読出されたプログラムコードが、コンピュータに挿入された機能拡張ボードやコンピュータに接続された機能拡張ユニットに備わるメモリに書込まれた後、そのプログラムコードの指示に基づき、その機能拡張ボードや機能拡張ユニットに備わるＣＰＵなどが実際の処理の一部または全部を行い、その処理によって前述した実施形態の機能が実現される場合も含まれることは言うまでもない。

【図面の簡単な説明】

【００４９】

【図１】実施例に係る音声認識結果修正方法を搭載した情報機器のハードウェア構成を示したブロック図である。

【図２】実施例に係る音声認識結果修正方法のモジュール構成を示したブロック図である。

【図３】１発声で２つまでのコマンドを同時に認識可能な場合の入力されるコマンド（音声入力コマンド）と出力されるコマンド（認識コマンド）の正誤の全組み合わせを示す図である。

【図４】認識結果を修正する物理キーの一例である。

【図５】図３の組み合わせに対して認識結果の正解部分を指定する際の物理キーの押下例を示す図である。

【図６】認識結果の正解部分を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートである。

【図７】図３の組み合わせに対して認識結果の誤り部分を指定する際の物理キーの押下例を示す図である。

【図８】認識結果の誤り部分を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートである。

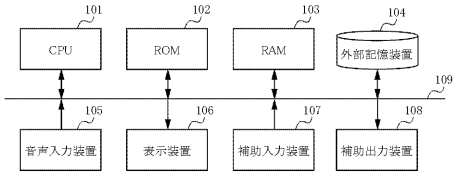
【図９】図３の組み合わせに対して認識結果の正誤を指定する際の物理キーの押下例を示す図である。

【図１０】認識結果の正誤を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートである。

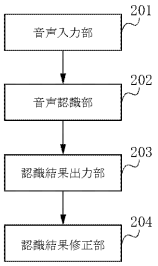
【図１１】認識結果の正誤を認識単位ごとに逐次的に指定する際の音声認識結果修正方法

の全体の処理を示したフローチャートである。
【図１２】図３の組み合わせに対して認識結果の誤り部分と誤りの種類を指定する際の物理キーの押下例を示す図である。
【図１３】認識結果の誤り部分と誤りの種類を指定する際の音声認識結果修正方法の全体の処理を示したフローチャートである。
【図１４】１発声で３つまでのコマンドを同時に認識可能な場合の入力されるコマンド（音声入力コマンド）と出力されるコマンド（認識コマンド）の正誤の全組み合わせを示す図である。
【図１５】図１４の組み合わせに対して認識結果の正解部分を指定する際の物理キーの押下例を示す図である。
【図１６】図１４の組み合わせに対して認識結果の誤り部分を指定する際の物理キーの押下例を示す図である。
【図１７】図１４の組み合わせに対して認識結果の正誤を指定する際の物理キーの押下例を示す図である。
【図１８】図１４の組み合わせに対して認識結果の誤り部分と誤りの種類を指定する際の物理キーの押下例を示す図である。

【図１】



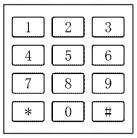
【図２】



【図３】

音声入力コマンド数＼認識コマンド数	1	2
1	(C) (S)	(C,I) (S,I)
2	(C,D) (S,D)	(C,C) (C,S) (S,S)

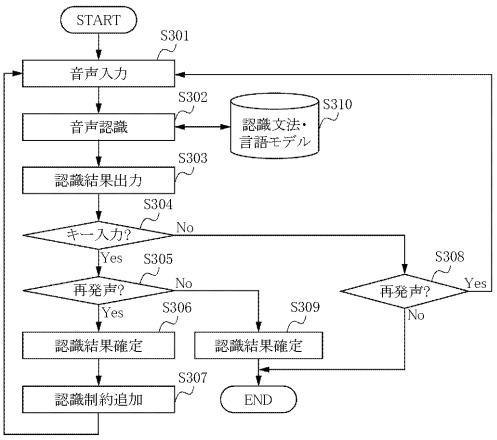
【図４】



【図５】

音声入力コマンド数＼認識コマンド数	1	2
1	(C):1 (S):R	(C,I):m (S,I):R
2	(C,D):1,R (S,D):R	(C,C):1,2 (C,S):m,R (S,S):R

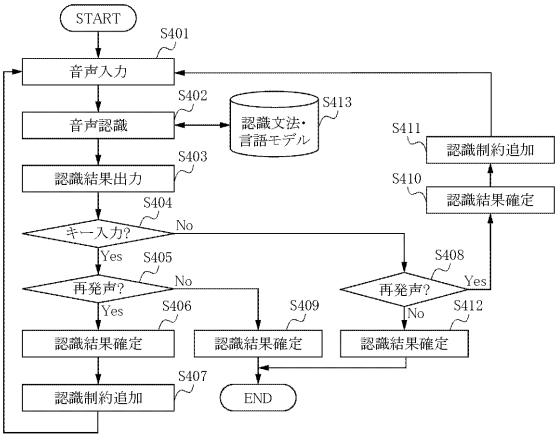
【図６】



【図７】

認識コマンド数 / 音声入力コマンド数	1	2
1	(C):N/A (S):1,R	(C,I):m (S,I):1,2,R
2	(C,D):R (S,D):1,R	(C,C):N/A (C,S):m,R (S,S):1,2,R

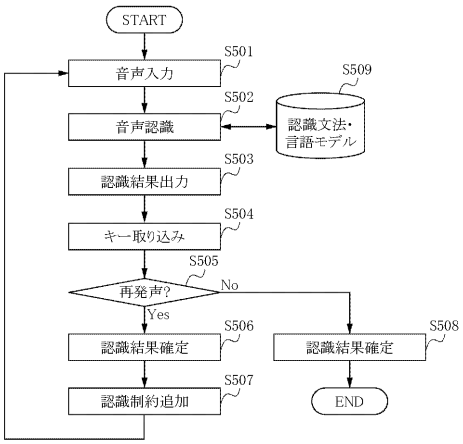
【図８】



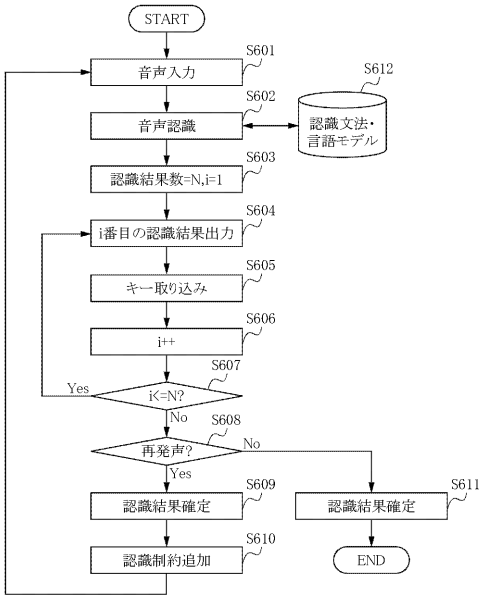
【図９】

認識コマンド数 / 音声入力コマンド数	1	2
1	(C):1 (S):2,R	(C,I):1,2 (S,I):2,2,R
2	(C,D):1,R (S,D):2,R	(C,C):1,1 (C,S):1,2,R (S,S):2,2,R

【図１０】



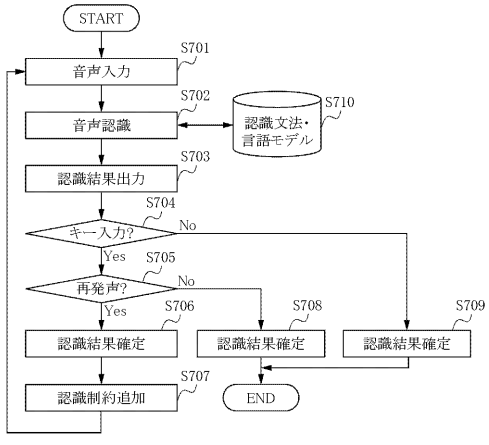
【図１１】



【図１２】

認識コマンド数 / 音声入力コマンド数	1	2
1	(C):N/A (S):1, 1, R	(C, I):m (S, I):1, 1, R
2	(C, D):2, R (S, D):2, 2, R	(C, C):N/A (C, S):m, R (S, S):2, 2, R

【図１３】



【図１４】

認識コマンド数 / 音声入力コマンド数	1	2	3
1	(C) (S)	(C, I) (S, I)	(C, I, I) (S, I, I)
2	(C, D) (S, D)	(C, C) (C, S) (S, S)	(C, C, I) (C, S, I) (S, S, I)
3	(C, D, D) (S, D, D)	(C, C, D) (C, S, D) (S, S, D)	(C, C, C) (C, C, S) (C, S, S) (S, S, S)

【図１５】

音声入力コマンド数＼認識コマンド数	1	2	3
1	(C):1 (S):R	(C,I):m (S,I):R	(C,I,I):j (S,I,I):R
2	(C,D):1,R (S,D):R	(C,C):1,2 (C,S):m,R (S,S):R	(C,C,I):j,k (C,S,I):j,R (S,S,I):R
3	(C,D,D):1,R (S,D,D):R	(C,C,D):1,2,R (C,S,D):m,R (S,S,D):R	(C,C,C):1,2,3 (C,C,S):j,k,R (C,S,S):j,R (S,S,S):R

【図１６】

音声入力コマンド数＼認識コマンド数	1	2	3
1	(C):N/A (S):1,R	(C,I):m (S,I):1,2,R	(C,I,I):j,k (S,I,I):1,2,3,R
2	(C,D):R (S,D):1,R	(C,C):N/A (C,S):m,R (S,S):1,2,R	(C,C,I):j (C,S,I):j,k,R (S,S,I):1,2,3,R
3	(C,D,D):R (S,D,D):1,R	(C,C,D):R (C,S,D):m,R (S,S,D):1,2,R	(C,C,C):N/A (C,C,S):j,R (C,S,S):j,k,R (S,S,S):1,2,3,R

【図１７】

音声入力コマンド数＼認識コマンド数	1	2	3
1	(C):1 (S):2,R	(C,I):1,2 (S,I):2,2,R	(C,I,I):1,2,2 (S,I,I):2,2,2,R
2	(C,D):1,R (S,D):2,R	(C,C):1,1 (C,S):1,2,R (S,S):2,2,R	(C,C,I):1,1,2 (C,S,I):1,2,2,R (S,S,I):2,2,2,R
3	(C,D,D):1,R (S,D,D):2,R	(C,C,D):1,1,R (C,S,D):1,2,R (S,S,D):2,2,R	(C,C,C):1,1,1 (C,C,S):1,1,2,R (C,S,S):1,2,2,R (S,S,S):2,2,2,R

【図１８】

音声入力コマンド数＼認識コマンド数	1	2	3
1	(C):N/A (S):1,1,R	(C,I):m (S,I):1,1,R	(C,I,I):j,k (S,I,I):1,1,R
2	(C,D):2,R (S,D):2,2,R	(C,C):N/A (C,S):m,R (S,S):2,2,R	(C,C,I):j (C,S,I):j,k,3,R (S,S,I):2,2,R
3	(C,D,D):2,3,R (S,D,D):3,3,R	(C,C,D):3,R (C,S,D):m,3,R (S,S,D):3,3,R	(C,C,C):N/A (C,C,S):j,R (C,S,S):j,k,R (S,S,S):3,3,R

フロントページの続き

(56)参考文献 特開平１０－００３２９５（ＪＰ，Ａ）
特開平０１－１５４１００（ＪＰ，Ａ）
特開平０８－０６３１８５（ＪＰ，Ａ）
特開平０１－１８９６９９（ＪＰ，Ａ）
特開平０２－０１４０００（ＪＰ，Ａ）
特開２００２－１４００９４（ＪＰ，Ａ）
特開２００３－０２９７７９（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　１１／００－２１／０６