

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第3980791号
(P3980791)

(45) 発行日 平成19年9月26日(2007.9.26)

(24) 登録日 平成19年7月6日(2007.7.6)

(51) Int. Cl.

F I

G 1 0 L 15/06 (2006.01)
 G 1 0 L 15/22 (2006.01)
 G 1 0 L 15/00 (2006.01)
 G O 1 C 21/00 (2006.01)
 G O 8 G 1/0969 (2006.01)

G 1 0 L 15/06 4 0 0 W
 G 1 0 L 15/22 4 7 0 Z
 G 1 0 L 15/00 2 0 0 Q
 G O 1 C 21/00 H
 G O 8 G 1/0969

請求項の数 8 (全 24 頁)

(21) 出願番号 特願平11-161474
 (22) 出願日 平成11年5月3日(1999.5.3)
 (65) 公開番号 特開2000-315096(P2000-315096A)
 (43) 公開日 平成12年11月14日(2000.11.14)
 審査請求日 平成16年5月6日(2004.5.6)

(73) 特許権者 000005016
 パイオニア株式会社
 東京都目黒区目黒1丁目4番1号
 (74) 代理人 100063565
 弁理士 小橋 信淳
 (72) 発明者 岩田 孝洋
 埼玉県川越市大字山田字西町25番地1
 パイオニア株式会社川越工場内
 審査官 涌井 智則

最終頁に続く

(54) 【発明の名称】 音声認識装置を備えたマンマシンシステム

(57) 【特許請求の範囲】

【請求項1】

1又は2以上の処理機能を有し、音声情報を情報伝達媒体として対話的に一の前記処理機能の動作を行う音声認識装置を備えたマンマシンシステムにおいて、

上位概念の属性から下位概念の属性の範疇に分類した階層化構造に基づいて、予め複数の照合情報を記憶すると共に、入力される音声を音声認識して得られる語彙を保持する情報保持部を有する記憶部と、

前記入力される音声を音声認識し、前記音声認識した結果と前記記憶部に記憶されている前記照合情報とを照合して、所定の類似度基準より類似度の高い1又は複数の照合情報を認識情報として保持する認識結果保持領域と、前記各処理機能毎に対応付けられた制御語とを有する制御部と、を具備し、

前記認識結果保持領域は、前記認識情報を前記上位概念の属性から下位概念の属性の範疇に分類した各属性ごとに保持する複数の保持領域を持ち、

前記制御部は、前記一の処理機能を提示した後、前記提示に応じて前記一の処理機能を指定する指令情報を有する音声が入力されると前記音声認識を行い、前記音声認識により得られる前記指令情報に対応する前記制御語に従って、前記分類に応じた保持領域中に保持している前記認識情報に対して前記一の処理機能の動作を行うこと、

を特徴とする音声認識装置を備えたマンマシンシステム。

【請求項2】

前記認識情報は、前記音声認識した結果と前記照合情報との照合により得られる1又は

10

20

複数の語彙情報であり、

前記制御語は、前記語彙情報を訂正する指令情報であること、
を特徴とする請求項 1 に記載の音声認識装置を備えたマンマシンシステム。

【請求項 3】

前記語彙情報を訂正する指令情報は、前記 1 又は複数の語彙情報の一を指定して訂正する制御情報であること、

を特徴とする請求項 2 に記載の音声認識装置を備えたマンマシンシステム。

【請求項 4】

前記語彙情報を訂正する指令情報は、前記 1 又は複数の語彙情報を順次に次候補選択して訂正する制御情報であること、

を特徴とする請求項 2 に記載の音声認識装置を備えたマンマシンシステム。

【請求項 5】

前記制御語は、前記処理機能に相当する訂正の動作を起動させる制御命令語と、前記制御命令語の被処理対象となる前記語彙情報に相当する制御対象語とを組み合わせる構成される前記指令情報であり、

前記制御部は、前記指令情報を有する前記音声が入力されると、前記指令情報に対応する前記制御対象語と制御命令語から成る前記制御語に従って、前記訂正の動作を行うこと、

を特徴とする請求項 2 ～ 4 のいずれか 1 項に記載の音声認識装置を備えたマンマシンシステム。

【請求項 6】

前記制御部は、前記 2 以上の属性に対応する前記処理機能を提示して、前記提示に応じて前記 2 以上の属性に該当する前記語彙を各属性毎の前記認識情報として前記認識結果保持領域に記憶した後、前記指令情報により上位概念の属性の前記認識情報の訂正が指令されると、前記上位概念の属性の前記認識情報の訂正処理を行うと共に、前記情報保持部に保持されている前記上位概念より下位概念の属性の語彙を、前記記憶部に記憶されている前記上位概念に従属する下位概念の属性の前記照合情報と照合して、所定の類似度基準より類似度の高い 1 又は複数の照合情報を認識情報として再度前記認識結果保持領域に保持すること、

を特徴とする請求項 2 ～ 5 のいずれか 1 項に記載の音声認識装置を備えたマンマシンシステム。

【請求項 7】

前記制御部は、前記音声入力情報を再度照合させる際、前記入力される音声を音声認識して得られる語彙と、前記下位概念の認識情報の属性に属している、前記下位概念の認識情報と同一の照合情報を除外した照合情報とを照合させること、

を特徴とする請求項 6 に記載の音声認識装置を備えたマンマシンシステム。

【請求項 8】

前記制御部は、前記 2 以上の属性に対応する前記処理機能を提示して、前記提示に応じて前記 2 以上の属性に該当する前記語彙を各属性毎の前記認識情報として前記認識結果保持領域に記憶した後、前記指令情報により下位概念の属性の前記認識情報を順次に次候補選択して訂正する訂正が指令されると、次候補選択した認識情報を新たな認識情報として提示すること、

を特徴とする請求項 4 に記載の音声認識装置を備えたマンマシンシステム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、音声認識機能を有するマンマシンシステム (man-machine system) に関し、特に、ユーザーにとっては意志を容易に指示することができ、マンマシンシステム側ではユーザーの意志に従って的確な処理を行うことを可能にするマンマシンシステムに関するものである。

10

20

30

40

50

【 0 0 0 2 】

【従来の技術】

マンマシンシステムの初期の概念は、人間と電子計算機との長所を生かすべく構成されたシステム（装置）というものであったが、今では、この概念は電子計算機のみならず、更に一般的な機械（広義の機械）と人間との間での関係を円滑にすることを可能にしたシステムについても言われるようになっている。

【 0 0 0 3 】

このようなマンマシンシステムとして、話者（ユーザー）が音声入力することで自分の意志を指示（命令）することを可能にする音声認識装置を備えたものが知られている。一例として、音声入力によりユーザーが目的地等を指示することを可能にした、GPS（Global Positioning System）航法を用いた車載用ナビゲーションシステムが知られている。

10

【 0 0 0 4 】

このナビゲーションシステムは、ユーザーが所望の目的地を発話すると、その目的地を音声認識し、現在地から目的地までの走行経路を探索して、表示装置にその走行経路を地図表示する機能を有している。

【 0 0 0 5 】

例えば、ナビゲーションシステムは、図13に示すような所定ステップの処理を行うことで、ユーザーの指示する目的地を認識する。ユーザーが目的地である「目黒駅」までの走行経路を調べたい場合を例示すれば、まず、第1ステップにおいて、音声認識装置に組み合わされた音声合成装置が、『名称をどうぞ』という合成音を発声し、ユーザーに対して具体的な目的地の名称を音声入力（発話）するように示唆する。

20

【 0 0 0 6 】

これに対し、ユーザーが「目黒駅」と発話すると、音声認識装置が、この「目黒駅」の声の特徴を抽出し、抽出した特徴パラメータD1を記憶部等に一時的に保持する。すなわち、第1ステップでは、音声認識装置は、ユーザーの発話した「目黒駅」という声の特徴抽出を行うだけに止めておき、最終的な認識までは行わない。

【 0 0 0 7 】

次に、第2ステップにおいて、再び音声合成装置が『ジャンルをどうぞ』という合成音を発声することにより、ユーザーの欲する具体的な目的地よりも上位概念であるジャンルを発話するように示唆する。

30

【 0 0 0 8 】

これに対し、ユーザーが「駅名」と発話すると、音声認識装置は、この「駅名」の声の特徴を抽出することにより、その特徴パラメータD2を生成する。更に、音声認識装置に予め備えられている認識語辞書中の認識対象語彙と、特徴パラメータD2とを照合し、特徴パラメータD2に最も類似した認識対象語彙LD2を選択することで、ユーザーの発話した「駅名」の声についての音声認識を行う。

【 0 0 0 9 】

次に、ステップ3において、音声合成装置が『県名をどうぞ』という合成音を発声することにより、ユーザーに対して地域名を発話するように示唆する。

40

【 0 0 1 0 】

これに対して、ユーザーが「東京都」と発話すると、音声認識装置は、この「東京都」の声の特徴を抽出することにより、その特徴パラメータD3を生成する。更に、上記認識語辞書中の認識対象語彙と、特徴パラメータD3とを照合し、特徴パラメータD3に最も類似した認識対象語彙LD3を選択することで、ユーザーの発話した「東京都」の声についての音声認識を行う。

【 0 0 1 1 】

次に、ステップ4において、音声認識装置は、認識語辞書中の認識対象語彙のうち、認識対象語彙LD2及びLD3の範疇に属する認識対象語彙を絞り込む。更に、その絞り込んだ認識対象語彙と上記の特徴パラメータD1とを照合し、特徴パラメータD1に最も類似

50

した認識対象語彙LD1を選択することにより、第1ステップで発話された「目黒駅」の声についての音声認識を行う。

【0012】

すなわち、上記の第1ステップにおいて下位概念の名称「目黒駅」が発話された時点では、認識語辞書中に格納されている膨大な数の認識対象語彙の中から、ユーザーの所望する地域内に存在する「目黒駅」の認識対象語彙LD1を特定することは一般的には困難である。

【0013】

このため、発話された「目黒駅」の声の特徴パラメータD1を記憶部等に保持しておき、第2～第4ステップにおいてユーザーからのジャンルと地域名に関する発話を受けて、認識語辞書中の認識対象語彙の検索範囲を絞り込んだ後、その絞り込んだ認識対象語彙と特徴パラメータD1とを照合することで、「目黒駅」の認識対象語彙LD1を特定し易くしている。

10

【0014】

そして、選択した認識対象語彙LD3とLD1に基づいて『 内の ですね』という合成音を発声することにより、認識結果をユーザーに提示する。すなわち、認識対象語彙LD3とLD1が適切に「東京都」「目黒駅」と認識された場合には、『 内の ですね』は『東京都内の目黒駅ですね』という合成音となってユーザーに提示される。

【0015】

20

つまり、音声認識装置はあくまでも認識語辞書の中から最も類似した認識対象語彙LD1～LD3を選択するに過ぎないため、ユーザーの発話した声が不明瞭であった場合等では、「目黒駅」を「目白駅」と誤認識したり、「東京都」を「京都府」と誤認識する場合が可能性としてあり得る。このように誤認識した場合には、『京都府内の目白駅ですね』という合成音をユーザーに提示することになる。したがって、認識結果をユーザーに提示し確認を仰ぐために、認識対象語前LD3とLD1に基づいて合成音を発声することとしている。

【0016】

そして、ユーザーがこの提示された合成音を聴き、正確な音声認識がなされたと判断して、例えば「検索開始」等と発話すると、音声認識装置がこれを認識し、更にナビゲーションシステムがその確認の指示を受けて現在地から東京都内の目黒駅までの走行経路を探索し、その探索した走行経路の地図を表示装置に表示する。

30

【0017】

一方、ユーザーが誤認識と判断した場合には、「戻れ」と発話して指示すると、音声認識装置は、その指示を受けて再び音声認識を開始し、再び提示した認識結果についてユーザーからの「検索開始」の指示を受けるまで、その音声認識を繰り返すようになっている。

【0018】

このように、上記のナビゲーションシステムは、音声認識装置と音声合成装置との組み合わせにより、ユーザーに対して対話操作を可能にしたものである点で優れた機能を有している。

40

【0019】

また、ユーザーの思考特性に合わせた順番でキーワードとなる言葉を発話させるようにしているので、ユーザーに対して優れた利便性を提供するものである。つまり、ユーザーは所望の目的地を指示する際に、最も具体的な目的地（上記例で言えば、目黒駅）を最初に指示し、次にその目的地の存在するジャンルや地域名を指示するという思考特性を有していることから、ユーザーの思考特性に合わせたマンマシンシステムとなっている。

【0020】

具体例を述べれば、情報検索システムでは、合理的な情報管理手法として、最も上位概念のカテゴリを決めておき、その上位概念のカテゴリに関連する中位概念と下位概念の情報を階層化して管理するという手法が採られている。こうした階層化構造を採用することで

50

、ユーザーが膨大な下位概念の情報の中から特定情報を検索する場合に、上位概念と中位概念の情報を利用して所謂情報の絞り込みを行い、所望の特定情報にいち早く到達できるようにしている。

【0021】

ところが、かかる情報検索システムと同様の検索手法を用いてマンマシンシステムを構築することとすると、ユーザーの思考特性とは合わなくなる場合がある。すなわち、上記のナビゲーションシステムを引用して例示すれば、最初に上位概念のカテゴリ『ジャンル』を質問し、それに対してユーザーに「駅名」と発声してもらうことにし、次に、中位概念の『県名』を質問して、それに対してユーザーに「東京都」と発声してもらうことにし、最後に、下位概念の『具体的な駅の名称』を質問して、それに対して「目黒駅」と発声し

10

【0022】

このような観点からして、上記従来のナビゲーションシステムは、ユーザーの所望する事項を違和感のない順序で入力させるので、ユーザーに対して優れた利便性を提供するものであるということが出来る。

【0023】

【発明が解決しようとする課題】

しかしながら、上記従来のナビゲーションシステムは、ユーザーの思考特性に合わせた音声認識手法を講じたことで、次のような不具合を生じる場合があった。

20

【0024】

例えば図13の場合には、上記したように、第1ステップで発話された「目黒駅」の音を音声認識せず、第2～第4ステップで上記の絞り込みを行った後に初めて「目黒駅」の音を音声認識して、その認識結果を提示することとしている。

【0025】

そして、誤認識があった場合には、「戻れ」の指示を受けて再度の音声認識を行うことによって対応することとしている。

【0026】

ところが、この「戻れ」の意味は、「1ステップ前の処理に戻ってから再度処理を開始しろ」という命令を意味している。このため、目的地「目黒駅」が誤認識であった場合には、ユーザーは第4ステップから第1ステップに処理を戻すために、3回「戻れ」と発話し、更に図13に示した第1～第4ステップの操作を繰り返さなければならぬため、ユーザーに対して煩雑な操作を強いるという問題があった。また、ジャンル「駅名」が誤認識であった場合にも同様に、ユーザーは第4ステップから第2ステップに処理を戻すために、2回「戻れ」と発話し、更に図13に示した第2～第4ステップの操作を繰り返さなければならぬため、ユーザーに対して煩雑な操作を強いるという問題があった。

30

【0027】

このように、従来のナビゲーションシステムは、「戻れ」という発話がなされた場合に、先に音声入力された情報を新たに音声入力された情報に書き換える（置き換える）という機能を用意しておくことで、誤認識に対応することとしているが、この機能は再度の音声認識を行うためのもの過ぎないため、ユーザーが簡単な操作で訂正の指示を行うことを可能にする機能ではなかった。このため、ユーザーに対し煩雑な操作を強いることになるという問題があった。

40

【0028】

本発明は上記従来の問題点を克服するためになされたものであり、ユーザーにとって容易な対話操作（例えば、訂正操作等）を行うことができる音声認識装置を備えたマンマシンシステムを提供することを目的とする。

【0029】

【課題を解決するための手段】

請求項1に記載の発明は、1又は2以上の処理機能を有し、音声を情報伝達媒体として

50

対話的に一の前記処理機能の動作を行う音声認識装置を備えたマンマシンシステムにおいて、上位概念の属性から下位概念の属性の範疇に分類した階層化構造に基づいて、予め複数の照合情報を記憶すると共に、入力される音声を音声認識して得られる語彙を保持する情報保持部を有する記憶部と、前記入力される音声を音声認識し、前記音声認識した結果と前記記憶部に記憶されている前記照合情報とを照合して、所定の類似度基準より類似度の高い1又は複数の照合情報を認識情報として保持する認識結果保持領域と、前記各処理機能毎に対応付けられた制御語とを有する制御部と、を具備し、前記認識結果保持領域は、前記認識情報を前記上位概念の属性から下位概念の属性の範疇に分類した各属性ごとに保持する複数の保持領域を持ち、前記制御部は、前記一の処理機能を提示した後、前記提示に応じて前記一の処理機能を指定する指令情報を有する音声が外部より入力されると前記音声認識を行い、前記音声認識により得られる前記指令情報に対応する前記制御語に従って、前記分類に応じた保持領域中に保持している前記認識情報に対して前記一の処理機能の動作を行うこと、を特徴とする。

10

【0034】

この構成によっても、上記同様に制御語に各処理機能を制御する種々の制御命令語を登録しておくことで、指令情報に対応した多様な処理が可能となる。

【0035】

請求項2に記載の発明は、請求項1に記載の音声認識装置を備えたマンマシンシステムにおいて、前記認識情報は、前記音声認識した結果と前記照合情報との照合により得られる1又は複数の語彙情報であり、前記制御語は、前記語彙情報を訂正する指令情報であること、を特徴とする。

20

【0036】

請求項3に記載の発明は、請求項2に記載の音声認識装置を備えたマンマシンシステムにおいて、前記語彙情報を訂正する指令情報は、前記1又は複数の語彙情報の一を指定して訂正する制御情報であること、を特徴とする。

【0037】

請求項4に記載の発明は、請求項2に記載の音声認識装置を備えたマンマシンシステムにおいて、前記語彙情報を訂正する指令情報は、前記1又は複数の語彙情報を順次に次候補選択して訂正する制御情報であること、を特徴とする。

【0038】

30

請求項5に記載の発明は、請求項2～4のいずれか1項に記載の音声認識装置を備えたマンマシンシステムにおいて、前記制御語は、前記処理機能に相当する訂正の動作を起動させる制御命令語と、前記制御命令語の被処理対象となる前記語彙情報に相当する制御対象語とを組み合わせる構成される前記指令情報であり、前記制御部は、前記指令情報を有する前記音声が外部より入力されると、前記指令情報に対応する前記制御対象語と制御命令語から成る前記制御語に従って、前記訂正の動作を行うこと、を特徴とする。

【0039】

請求項6に記載の発明は、請求項2～5のいずれか1項に記載の音声認識装置を備えたマンマシンシステムにおいて、前記制御部は、前記2以上の属性に対応する前記処理機能を提示して、前記提示に応じて前記2以上の属性に該当する前記語彙を各属性毎の前記認識情報として前記認識結果保持領域に記憶した後、前記指令情報により上位概念の属性の前記認識情報の訂正が指令されると、前記上位概念の属性の前記認識情報の訂正処理を行うと共に、前記情報保持部に保持されている前記上位概念より下位概念の属性の語彙を、前記記憶部に記憶されている前記上位概念に従属する下位概念の属性の前記照合情報と照合して、所定の類似度基準より類似度の高い1又は複数の照合情報を認識情報として再度前記認識結果保持領域に保持すること、を特徴とする。

40

【0040】

請求項7に記載の発明は、請求項6に記載の音声認識装置を備えたマンマシンシステムにおいて、前記制御部は、前記音声入力情報を再度照合させる際、前記入力される音声を音声認識して得られる語彙と、前記下位概念の認識情報の属性に属している、前記下位概

50

念の認識情報と同一の照合情報を除外した照合情報とを照合させること、を特徴とする。

【0041】

請求項8に記載の発明は、請求項4に記載の音声認識装置を備えたマンマシンシステムにおいて、前記制御部は、前記2以上の属性に対応する前記処理機能を提示して、前記提示に応じて前記2以上の属性に該当する前記語彙を各属性毎の前記認識情報として前記認識結果保持領域に記憶した後、前記指令情報により下位概念の属性の前記認識情報を順次に次候補選択して訂正する訂正が指令されると、次候補選択した認識情報を新たな認識情報として提示すること、を特徴とする。

【0042】

これらの構成によると、外部入力される訂正の指令情報を、種々の訂正処理に関する情報を有する制御語を介して受け付ける。そして、その制御語に基づいて各種の訂正処理が行われる。特に、認識情報が階層化されて初めて意味をなす情報である場合に、上位概念の属性の認識情報を訂正すると、下位概念の属性に属する認識情報に影響を及ぼし、逆に、下位概念の属性の認識情報を訂正すると、上位概念の属性に属する認識情報に影響を及ぼすといった相互に影響を及ぼすことになり、適切な訂正処理を行う必要が生じる。このような場合に、制御語に予め適切な訂正処理を可能にする機能を設定しておくことで、迅速且つ適切な訂正処理が可能となる。

【0043】

例えば、ナビゲーションシステムに目的地を入力する際に、認識結果である認識情報に誤りがあったり、別の目的地に変更するような場合には、迅速な訂正処理が望まれるが、このような場合に制御語を用いることで対話形式の操作を迅速に行うことが可能となる。

【0044】

【発明の実施の形態】

以下本発明の実施の形態を図面を参照して説明する。尚、一実施形態として、音声認識機能を有する車載用ナビゲーションシステムについて説明する。

【0045】

図1は、本実施形態の車載用ナビゲーションシステムに備えられた音声認識装置の構成を示すブロック図である。より厳密に言えば、車載用ナビゲーションシステムにマンマシンインターフェース(man-machine interface)装置として設けられた音声認識装置の部分の構成を示している。

【0046】

図1において、本音声認識装置は、マイクロフォンを有する音声入力部1と、信号処理部2と、記憶部3と、マイクロプロセッサ(MPU)を備えた制御部4、キーボード等の操作部5、音声合成用集積回路装置を有する音声出力部6、液晶ディスプレイ等で形成された表示部7とを備えて構成されている。

【0047】

ただし、操作部5と音声出力部6及び表示部7は、ナビゲーションシステムに備えられている構成要素を供用したり、外部装置として別途接続してもよい。

【0048】

記憶部3は、分析データ記憶部10と認識語記憶部11を備えて構成されている。信号処理部2は、デジタルシグナルプロセッサ(Digital Signal Processor: DSP)等で形成され、制御部4からの制御信号に従って動作する音声分析部8及び類似度演算部9が備えられている。

【0049】

音声入力部1は、ユーザーが発話した音声を集音し、デジタルデータの音声データD_{in}に変換して音声分析部8に供給する。

【0050】

音声分析部8は、音声データD_{in}に基づいてユーザーが発話した音声の特徴抽出を行い、その特徴抽出結果である特徴パラメータのデータ(以下、分析データという)D_aを出力する。

10

20

30

40

50

【 0 0 5 1 】

分析データ記憶部 10 は、再記憶可能なメモリ (R A M) で形成され、音声分析部 8 より出力される分析データ D a を記憶する。また、記憶した分析データ D a ' を制御部 4 からの制御信号に応じて類似度演算部 9 に供給する。

【 0 0 5 2 】

ここで、本音声認識装置は、ユーザーに対して発話してもらうための語彙の範疇を示唆し、その範疇内の語彙をユーザーが発話するという対話形式を採用している。そこで、分析データ記憶部 10 には、示唆した範疇に対応付けて分析データ D a を記憶するようにしている。

【 0 0 5 3 】

認識語記憶部 11 は、読出し専用メモリ (R O M) で形成され、ユーザーの発話した語彙を認識する際に分析データ D a , D a ' と照合される多数の参照データ (以下、認識対象語彙という) が記憶されている。これらの認識対象語彙は、複数の認識辞書中に分類して格納されている。認識辞書として、図 2 に示すように、階層化されたジャンル辞書 M J と地域名辞書 M K と施設名称辞書 M L が備えられている。

10

【 0 0 5 4 】

最上層に位置付けられたジャンル辞書 M J には、「駅名」「病院名」「遊園地名」等の認識対象語彙が割り当てられている。ジャンル辞書 M J に従属された地域名辞書 M K には、「北海道」「東京都」等の地域名の認識対象語彙が割り当てられている。地域名辞書 M K に従属された施設名称辞書 M L には、具体的な「目黒駅」「東京駅」等の施設名称の認識対象語彙が割り当てられている。

20

【 0 0 5 5 】

つまり、ジャンル辞書 M J 中の「駅名」の認識対象語彙前には、「北海道」「東京都」等の全国の地域名の認識対象語彙がツリー構造で従属され、各地域名の認識対象語彙には、各地域に存在する多数の駅の施設名称がツリー構造で従属されている。また、ジャンル辞書 M J 中の「病院名」の認識対象語彙にも、全国の地域名が従属し、それぞれの地域名には多数の病院の施設名称が従属している。「遊園地名」の認識対象語彙についても同様に全国の地域名と多数の遊園地の施設名称が従属している。

【 0 0 5 6 】

更に、本音声認識装置には、ユーザーが音声入力により、又は操作部 5 から訂正等の命令をするため複数の命令情報 (以下、命令語彙という) が準備されており、これらの命令語彙 W も命令語辞書 M C として認識語記憶部 11 に記憶されている。

30

【 0 0 5 7 】

尚、ジャンル辞書 M J の範疇に属する「駅名」「病院名」「遊園地」等の認識対象語彙はジャンル認識対象語彙、地域名辞書 M K の範疇に属する「北海道」「東京都」等の認識対象語彙は地域名認識対象語彙、施設名称辞書 M L の範疇に属する認識対象語彙は施設名称認識対象語彙と呼ばれ、ユーザーが命令語彙 W を用いて本音声認識装置に命令する操作はコマンド操作と呼ばれている。

【 0 0 5 8 】

類似度演算部 9 は、音声分析部 8 より供給される分析データ D a を辞書 M J , M K , M L , M C 中の認識対象語彙 D b と照合し、分析データ D a に最も類似している認識対象語彙 D b と、所定の類似度判定基準より高い類似度の認識対象語彙 (1 又は複数個の認識対象語彙) D b とを選択し、その選択した認識対象語彙 D b を認識結果語彙 D c として制御部 4 へ出力する。また、分析データ記憶部 10 から供給された分析データ D a ' についても同様に処理して、分析データ D a ' に対応する認識結果語彙 D c を制御部 4 へ出力する。

40

【 0 0 5 9 】

これにより、ユーザーが発話した「駅名」「東京都」「目黒駅」等の声や、コマンド操作により発話された「訂正」等の声を認識した結果が、認識結果語彙 D c として制御部 4 に供給される。

【 0 0 6 0 】

50

尚、分析データ $D a$, $D a'$ を照合するための辞書 $M J$, $M K$, $M L$, $M C$ は、制御部 4 から供給される制御信号 C によって選択的に切り替えられる。例えば、ユーザーに対してジャンルの範疇に属する語彙を発話してもらうように示唆するときは、制御信号 C によってジャンル辞書 $M J$ を選択する。そして、示唆に応じてユーザーが発話し、それによって得られる分析データ $D a$ (又は $D a'$) を類似度演算部 9 がジャンル辞書 $M J$ 中の認識対象語彙 $D b$ と照合する。このように、ユーザーに対して示唆する内容と辞書 $M J$, $M K$, $M L$, $M C$ の選択との関係が対応付けられている。また、分析データ記憶部 10 から分析データ $D a'$ を読み出すときも、辞書 $M J$, $M K$, $M L$, $M C$ の選択と関連付けて行われる。

【0061】

10

制御部 4 は、予め設定されたシステムプログラムを実行することにより、本音声認識装置全体の動作を制御するための上記制御信号を出力すると共に、音声認識した結果を制御出力としてナビゲーションシステム内の中央制御部(図示省略)へ転送する。

【0062】

更に、制御部 4 は、操作部 5 からユーザーの指示を受けたり、ユーザーに提示すべきデータや音声認識した結果を音声出力部 6 又は表示部 7 に供給して、合成音や表示を提示させることにより、聴覚的・視覚的な対話を行うようになっている。

【0063】

また、制御部 4 には、類似度演算部 9 から供給される認識結果語彙 $D c$ を記憶するための認識結果記憶領域 $M R$ が備えられている。図 3 に示すように、認識結果記憶領域 $M R$ には、ジャンルファイル $F J$ 、地域名ファイル $F K$ 、施設名称ファイル $F L$ が備えられ、これらのファイル $F J$, $F K$, $F L$ に認識結果語彙 $D c$ を分類して格納する。

20

【0064】

すなわち、上述したように、本音声認識装置は、語彙の範疇をユーザーに示唆し、それに応じてユーザーがその範疇内の語彙を発話するという対話形式を採用していることから、『ジャンルをどうぞ』と示唆しそれに応じて発話されたときに得られ認識結果語彙 $D c$ をジャンルファイル $F J$ に格納する。『地域名をどうぞ』『県名をどうぞ』等と示唆しそれに応じて発話されたときに得られ認識結果語彙 $D c$ を地域名ファイル $F K$ に格納する。『名称をどうぞ』と示唆しそれに応じて発話されたときに得られ認識結果語彙 $D c$ を施設名称ファイル $F L$ に格納する。

30

【0065】

更に、類似度演算部 9 から制御部に類似度の高い順に複数個の認識結果語彙 $D c$ が転送されてくる。そこで、図 3 に示すように、ファイル $F J$, $F K$, $F L$ には、複数個ずつ転送されてくる認識結果語彙 $D c11$, $D c12 \sim$ 、 $D c21$, $D c22 \sim$ 、 $D c31$, $D c32 \sim$ を類似の高い順に揃えて格納する。

【0066】

そして、制御部 4 は、基本的に、これらのファイル $F J$, $F K$, $M L$ 中に記憶した最も類似度の高い認識結果語彙 $D c11$, $D c21$, $D c31$ を調べることで、ユーザーの指示した目的地(施設名称)等を特定する。ただし、次候補に関するコマンド操作がなされた場合には、次候補となる認識結果語彙に基づいて、ユーザーの指示した目的地(施設名称)等を特定する。

40

【0067】

更に、制御部 4 には、複数の制御語 $W W$ が予めファイリングされて格納されている。これらの制御語 $W W$ は、コマンド操作により指令された命令の内容(意味)を判定して迅速な対応を行うために設けられている。この制御語 $W W$ は、2つの基本的な形態を有している。

【0068】

第 1 の基本型の制御語 $W W$ は、図 4 (a) に模式的に示すように、制御対象語 X と制御命令語 Y の対で構成されている。制御命令語 Y は、ユーザーの指令に応じて処理すべき作業事項、制御対象語 X は、制御命令語 Y に従って処理すべき対象事項である。例えば、本音

50

声認識装置が音声認識した結果を『東京都内の駅名ですね』と提示し、ユーザーが「駅名を訂正」と発話した場合には、 $(X + Y) = (\text{駅名} + \text{訂正})$ の構成からなる制御語WWに基づいて、駅名を訂正するための処理を開始する。そして、新たに音声入力された結果が「病院名」となった場合には、訂正結果を『東京都内の病院名ですね』という合成音声等で再提示する。

【0069】

つまり、被制御対象である制御対象語Xは、上記のジャンルと地域名と施設名称の各属性（範疇）を表す認識対象語彙等と同じ内容となっており、これらの認識対象語彙等に制御命令語Yが予め付加されて制御語WWが構成されている。

【0070】

上記の例は、制御命令語Yが「訂正」になっている場合であり、その他にも、図5に示すような各種形態の制御語WWが予め準備されている。図5中の左欄に命令語の形態、右欄に命令語の意味、中欄に制御命令語Yの識別子を示し、「～」の部分が制御対象語X、それに続く語彙が制御命令語Yである。

【0071】

次に、第2の基本型の制御語WWは、図4(b)に模式的に示すように、制御命令語Yだけを予めファイリングして記憶しておき、ユーザーのコマンド操作で指摘された認識結果語彙Dcに、指令された制御命令語Yを後から付加することで、制御語WWを形成するという構成になっている。つまり、図4(a)に示した制御語WWは、制御対象語Xが予め定められているのに対し、図4(b)に示した制御語WWは、制御対象語Xの部分に融通

【0072】

本音声認識装置が、ジャンルに属する「駅名」、地域名に属する「東京都」の順に音声認識した結果を『東京都内の駅名ですね』と提示し、ユーザーが「訂正」と発話した場合には、最後（直近）に認識した地域名に属する「東京都」を訂正すべきと判断する。

【0073】

例えば、最後に認識した結果が地域名ファイルFK中の認識結果語彙Dc21であった場合であって、認識結果語彙Dc21に基づいて「東京都」の提示をした場合に、訂正の指示がなされたときには、最後の認識結果である認識結果語彙Dc21を図4(b)の認識結果語彙Dcとする。この認識結果語彙Dc21に、図5中に示す「訂正(CW)」の制

【0074】

したがって、第2の基本構成の制御語WWは、ユーザーが音声入力した結果を直ぐに訂正したいような場合に、迅速な対応ができるようにしたものである。例えば、ユーザーが目的地等を頻繁に切り替えたいと欲した場合等に、目的地等の音声入力と、「訂正」のコマンド操作を交互に繰り返すだけで、最終的に指定したい目的地等を簡単且つ迅速に指令することができるようになっている。つまり、ユーザーにとっては、制御命令語Yに該当する「訂正」の語彙を発話するだけで、直近の認識対象語彙を指定したことになるため、簡単な訂正が可能となり、優れた操作性が実現されている。

【0075】

一方、上記第1の基本構成の制御語WWは、最後に音声認識された認識結果語彙だけでなくそれより前に音声認識された認識結果語彙を訂正しようとする場合に、適切な対応処理が可能となる。

【0076】

図4(c)～(e)は、制御語WWの発展型の具体例を示している。ユーザーのコマンド操作の仕方に応じて、様々な長さや組み合わせの制御語WWが形成されることを示している。これらの発展型によれば、複数の認識結果語彙を一括して訂正したり(図4(d)) (

10

20

30

40

50

e))、訂正の指令を行っても即座に訂正を行わせるのではなく、何らかの処理を行った後に、実質的な訂正を指示することができる(図4(c)~(e))。

【0077】

特に、ユーザーが複数の語彙や命令語彙を連続した音声として入力する場合に、この発展型の制御語WWに基づいてユーザーの意志を反映した処理を行うことが可能である。

【0078】

このように、本音声認識装置は、固定の又は融通性を持たせた制御対象語Xと制御命令語Yとを組み合わせた制御語WWを介して、ユーザーの指示と指令を受け付けるようにしたので、優れた対話操作を可能にしている。

【0079】

次に、かかる構成を有する本音声認識装置の動作例を図6ないし図9に示すフローチャートを参照して説明する。尚、代表例として、ユーザーが既に音声入力した語彙を訂正する場合、すなわち図5に示した「~を訂正」、「訂正」、「~の次候補」、「次候補」のコマンド操作した場合の動作例について説明する。

【0080】

図6は、本音声認識装置における音声認識処理の基本動作を示している。同図において、本音声認識装置は、基本的には、施設名称の語彙 ジャンルの語彙 地域名の語彙の順でユーザーに発話してもらい、ジャンルと地域名の語彙に基づいて認識語記憶部11中の照合すべき認識対象語彙を絞り込んだ後、分析データ記憶部10に保存しておいた施設名称の分析データDa'と絞り込んだ認識対象語彙Dbとを照合することで、ユーザーの発話した施設名称の語彙を特定する。これを直接発話方式と呼び、ユーザーの思考特性に合わせた対話操作を可能にしている。

【0081】

尚、ジャンルの語彙を音声認識し、その認識結果に基づいて施設名称の語彙を特定できる場合には、地域名の語彙を発話してもらう必要がないので、地域名の語彙を発話してもらうための示唆を行わない。

【0082】

すなわち、図6において、まず本音声認識装置が、施設名称の語前を音声入力するようにユーザーに対して示唆し、これに応じて、ユーザーが所望の目的地である施設名称の語彙を音声入力すると、この分析データDaを分析データ記憶部10に記憶(保存)する。

【0083】

次に、上位概念であるジャンルの範疇に属する語彙を音声入力するように示唆し、これに対してユーザーがジャンルの範疇に属する語彙(例えば、「駅名」等の語彙)を音声入力すると、その音声の分析データDaを生成して、ジャンル辞書M中の認識対象語彙Dbと照合し、照合結果の認識結果語彙Dcを記憶結果記憶領域MR中の該当ファイルFに記憶する。そして、分析データ記憶部10に保存されている分析データ(ユーザーが発話した施設名称の分析データ)Da'を音声認識できる条件が揃った場合には、分析データDa'を該当する施設名称辞書ML中の認識対象語彙Dbと照合し、照合結果の認識結果語彙Dcを施設名称ファイルFLに記憶する。

【0084】

そして、認識結果である施設名称の認識結果語彙Dcをユーザーに提示して誤認識等の有無を確認させる。これに応じて目的地までの走行経路の探索を開始すべきコマンド操作が行われると、制御語WWに基づいて施設名称の目的地までの走行経路を探索して、探索した走行経路を表示する。

【0085】

ただし、上位概念の語彙が1回入力されただけでは、目的地である施設名称を特定(絞り込み)できなかった場合には、更に中位概念の地域名を音声入力すべき趣旨の示唆をし、地域名の範疇に属する語彙(例えば、「東京都」等の語彙)を発話してもらい、ジャンルの場合と同様の処理を繰り返す。

【0086】

10

20

30

40

50

図7は、本音声認識装置が音声入力された語彙を音声認識し、その認識結果をユーザーに提示したときに、ユーザーから訂正等のコマンド操作がなされた場合の基本動作を示している。

【0087】

まず、ユーザーに対して訂正等のコマンド操作をするか否か示唆し、それに応じて音声入力により又は操作部5から訂正等のコマンド操作がなされると、その指令された内容を認識する。そして、制御語WWに基づいて、例えば訂正対象を訂正するための処理を行うようになっている。尚、図7に示した処理は、図6に示した音声認識処理の各ステップにおいて適宜に行われる。

【0088】

次に、より具体的な動作例を図8及び図9を参照して説明する。

図8において、ユーザーが音声入力により又は操作部5によって音声認識開始のコマンド操作をすると、制御部5の制御下でステップ100以降の処理を開始する。

【0089】

ステップ100では、制御部4の指示に従って音声出力部6が『名称をどうぞ』という合成音を発声することにより、ユーザーに対して所望の目的地の名称（施設名称）を音声入力するように示唆（質問）する。

【0090】

これに対して、ユーザーが施設名称である例えば「目黒駅」と発話すると（ステップS102）、この「目黒駅」の音声を集音して音声分析部8が音声分析し、「目黒駅」の音声の分析データDaを生成する（ステップS104）。

【0091】

次に、分析データ記憶部10が、施設名称「目黒駅」の分析データDaを記憶する（ステップS106）。すなわち、施設名称に該当する認識対象語彙は膨大な数に上るため、後述する絞り込み処理を完了して認識辞書の照合範囲が絞られるまで、施設名称の分析データDaを分析データ記憶部10に記憶する。

【0092】

次に、ステップS108において、認識語記憶部11中のジャンル辞書MJを選択（設定）し、ジャンルに関する音声入力可能状態にする。

【0093】

次に、ステップS110において、類似度演算部9がジャンル辞書MJ中の認識対象語彙Dbに基づいて入力音声の認識を開始する。更に、音声出力部6が『ジャンルをどうぞ』という合成音を発声することにより、ユーザーに対してジャンルの範疇に属する語彙を発話するように示唆する。

【0094】

この示唆に応じてユーザーがジャンルに属する語彙（例えば、「駅名」）と発声すると（ステップS112）、この「駅名」の音声を集音して音声分析部8が音声の特徴抽出を行うことにより、「駅名」の音声の分析データDaを生成して、類似度演算部9へ供給する（ステップS114）。また、この分析データDaを分析データ記憶部10に記憶する。

【0095】

次に、ステップS116において、類似度演算部9が「駅名」の音声の分析データDaをジャンル辞書MJ中の認識対象語彙Dbと照合し、類似度判断基準より高い類似度の1又は複数個の認識対象語彙を選択する。そして、選択した認識対象語彙を類似度の高い順番に配列して、認識結果語彙Dcとして制御部4へ供給し、図3に示したジャンルファイルFJに記憶させる。

【0096】

次に、ステップS118において、ジャンルファイルFJに記憶した認識結果語彙Dcのうち最も類似度が高いもの（図3中のDc11が該当する）を音声出力部6と表示部7に供給し、合成音声と文字表示等によって認識結果を提示する。例えば、認識結果語彙Dc11が「駅名」であれば、『駅名ですね』等の合成音を提示する。

10

20

30

40

50

【 0 0 9 7 】

次に、この提示に対応して、ユーザーが発話によって「駅名を訂正」とコマンド操作すると、ステップ S 1 2 0 において訂正指令を認識して「YES」と判断し、ステップ S 1 2 2 へ移行する。

【 0 0 9 8 】

ここで、ステップ S 1 2 0 では、「駅名を訂正」とのコマンド操作がなされたので、図 4 (a) に示した第 1 の基本型の制御語 WW に基づいて、訂正の判断をする。つまり、制御対象語 X が「駅名」の認識結果語彙であり、且つ制御命令語 Y が「訂正」になっている制御語 WW を調べて、「駅名」をそれと同じ階層に在る他の認識対象語彙（すなわち、ジャンルの範疇に属する他の認識対象語彙）で訂正すべきと判断する。

10

【 0 0 9 9 】

そして、ステップ S 1 2 2 において、制御語 WW に基づいて制御信号 C を認識語記憶部に供給して、ジャンル辞書 M J を設定する。ただし、制御対象語 X が「駅名」となっているのに応じてジャンル辞書 M J 中の「駅名」の認識対象語彙は除外し、残りの認識対象語彙を設定する。

【 0 1 0 0 】

次に、ステップ S 1 1 0 - S 1 1 8 の処理を繰り返す。これにより、ステップ S 1 1 0 において『ジャンルをどうぞ』という合成音声でユーザーに示唆し、ステップ S 1 1 2 においてユーザーが「駅名」以外の音声入力をする、ステップ S 1 1 4 , S 1 1 6 でこれを分析且つ認識して、新たに指示されたジャンルの認識結果語彙 D c が制御部 4 に供給される。

20

【 0 1 0 1 】

そして、ステップ S 1 1 8 において制御部が、ジャンルファイル F J 中の古い認識結果語彙 D c11 , D c12 ~ を、新たな認識結果語彙 D c11 , D c12 ~ によって訂正する。例えば、ユーザーが「病院名」を音声入力すると、「病院名」に類似した認識結果語彙 D c がジャンルファイル F J に記憶されることになり、ユーザーの指令した訂正処理がなされることとなる（ステップ S 1 1 2 ~ S 1 1 8 ）。

【 0 1 0 2 】

また、上記のステップ S 1 2 0 において、単に「訂正」のコマンド操作がなされた場合には、図 4 (b) に示した第 2 の基本型の制御語 WW に基づいて、訂正の判断をする。つまり、ジャンルファイル F J 中の最も類似度の高い認識対象語彙 D c 1 1 に、「訂正」の命令制御語 Y を付加した制御語 WW に基づいて、「駅名」をそれと同じ階層に在る他の認識対象語彙（すなわち、ジャンルの範疇に属する他の認識対象語彙）で訂正すべきと判断する。

30

【 0 1 0 3 】

そして、ステップ S 1 2 2 を通って上記ステップ S 1 1 0 ~ S 1 1 8 の処理を繰り返すことで、ジャンルファイル F J 中の古い認識結果語彙 D c 1 1 , D c 1 2 ~ を、新たな認識結果語彙 D c 1 1 , D c 1 2 ~ によって訂正する。

【 0 1 0 4 】

尚、ステップ S 1 2 0 において、上記の第 1 , 第 2 の基本型の制御語 WW に対応する訂正コマンド操作が繰り返し行われた場合には、今までに訂正の指示がなされた複数の制御対象語 X（又は制御結果語 D c）の全てを、ジャンル辞書 M J から除外して、ジャンル辞書 M J 中の残余の認識対象語彙 D b に基づいて類似度の判定が行われる。つまり、訂正のコマンド操作がなされる度に、除外する認識対象語彙を広げていって、照合すべき認識対象語彙 D b の範囲を絞り込んでいく。このため、照合範囲が絞られて迅速な訂正処理が可能となる。

40

【 0 1 0 5 】

更にまた、ステップ S 1 2 0 において、図 5 中に示されている「～の次候補」の制御語（第 1 の基本型の制御語）WW に該当する例えば「駅名の次候補」のコマンド操作がなされた場合には、ジャンルファイル F J に記憶されている次の認識結果語彙 D c 1 2 を、ステ

50

ップS 1 1 8において提示する。例えば、認識結果語彙D c 1 2が「病院名」であった場合には、『病院名ですね』と提示する。

【0106】

すなわち、「駅名の訂正」や「訂正」のコマンド操作がなされた場合とは異なり、上記したステップS 1 2 2及びステップS 1 1 0～S 1 1 6における再度の処理では実質的に訂正処理は行わずに、ステップS 1 1 8において認識結果語彙D c 1 2を最も類似度の高い認識結果語彙として提示する。

【0107】

その提示に対してユーザーが再び「～の次候補」のコマンド操作をすると、ジャンルファイルF J中の更に次候補である認識結果語彙D c 1 3を提示する。また、再び次候補のコマンド操作がなされた場合には、ジャンルファイルF J中の更に次候補である認識結果語彙D c 1 4を提示する。したがって、ジャンルファイルF J中に存在する認識結果語彙D cの個数分だけ、ユーザーは次候補のコマンド操作を行うことができるようになっている。

10

【0108】

かかる次候補のコマンド操作を行うことで、ユーザーはジャンルの語彙を再度音声入力する手間を省くことができる。更に、本音声認識装置における訂正処理が高速化する。

【0109】

更にまた、ステップS 1 2 0において、図5中に示されている「次候補」の制御語(第2の基本形の制御語)WWに該当するコマンド操作がなされた場合にも、「～の次候補」と同様の訂正処理を行う。つまり、ユーザーが単に「次候補」と発話すると、ジャンルファイルF J中に記憶されている次候補の認識結果語彙D c 1 2が制御対象語となり、この認識結果語彙D c 1 2を提示する。そして、再び次候補のコマンド操作がなされた場合には、ジャンルファイルF J中に記憶されている更に次候補の認識結果語彙D c 1 3が制御対象語となり、この認識結果語彙D c 1 3を提示する。したがって、ジャンルファイルF J中に存在する認識結果語彙D cの個数分だけ、ユーザーは次候補のコマンド操作を行うことができるようになっている。

20

【0110】

この「次候補」とのみ発話するコマンド操作によれば、ユーザーは制御対象語Xに該当する語彙を発話しなくて済むので、「～の次候補」のコマンド操作に比してより迅速な訂正が可能となる。

30

【0111】

ただし、いずれの次候補コマンド操作を行うかはユーザーが適宜に選択することができるので、本音声認識装置は、融通性の高いシステムとなっている。

【0112】

このように、ステップS 1 2 0において、「～を訂正」「訂正」「～の次候補」「次候補」のコマンド操作を行うことで、ユーザーは簡単に訂正を行うことができ、所望のジャンルを確定することができる。また、説明をするまでもなく、これらのコマンド操作を混在して行うことも当然に可能である。

【0113】

次に、ステップS 1 1 8で提示した認識結果に対して、ユーザーがジャンルの語彙が確定したと判断すると、ステップS 1 2 0からステップS 1 2 4の処理に移行することになる。

40

【0114】

ステップS 1 2 4では、ファイルF J, F Kを調べ、記憶されている認識結果語彙に基づいて施設名称を確定するのに十分な絞り込み条件が揃ったか否か判断する。ここで、未だ十分な絞り込み条件が揃っていなければ、ステップS 1 2 6に移行して、次の地域名辞書M Kを設定して、ステップS 1 1 0からの処理を繰り返す。

【0115】

例えば、ジャンルファイルF J中に、既に認識結果語彙D c 1 1, D c 1 2～が記憶され

50

ていても、地域名ファイル F K 中には未だ認識結果語彙が記憶されていないために絞り込みができない場合には、ステップ S 1 2 4 において、未だ絞り込み条件が揃っていないと判断し、更にステップ S 1 2 6 において、地域名辞書 M K を設定してから、ステップ S 1 1 0 の処理へ移行する。

【 0 1 1 6 】

このようにステップ S 1 1 0 からの処理を繰り返すと、ステップ S 1 1 0 では、ユーザーに対して地域名の範疇に属する語彙を発話してもらうために例えば『県名をどうぞ』という合成音を発声する。

【 0 1 1 7 】

これに応じてユーザーが「東京都」等の地域名を音声入力すると（ステップ S 1 1 2 ）、上記したステップ S 1 1 4 ~ S 1 1 8 の処理により、地域名ファイル F K に、類似度の高い 1 又は複数の認識結果語彙 D c 2 1 , D c 2 2 ~ を記憶する。

10

【 0 1 1 8 】

更に、ステップ S 1 1 8 において、ジャンルファイル F J と地域名ファイル F K に記憶された認識結果語彙に基づいて例えば『東京都の駅名ですね』という合成音を提示する。

【 0 1 1 9 】

これに対してユーザーが、音声で「東京都を訂正」とコマンド操作すると（S 1 2 0 ）、前述したのと同様に、「~を訂正」の制御語 W W に基づいて地域名ファイル F K 中の認識結果語彙を訂正する。

【 0 1 2 0 】

20

尚、地域名に関する訂正処理も上記したジャンルに関する訂正処理と同様に行われるので、詳細な説明は割愛するが、ステップ S 1 2 0 においてユーザーが単に「訂正」と発話すれば、直近の認識結果である認識結果語彙 D c 2 1 を訂正するための処理が行われる。

【 0 1 2 1 】

また、「東京都の次候補」と発話すれば、地域名ファイル F K 中の次の候補である認識結果語前 D c 2 2 を提示して、提示結果に対して再度「~の次候補」のコマンド操作がなされる度に、更に次の次候補を提示する。このように、「~の次候補」のコマンド操作がなされた場合には、ユーザーは訂正のための語彙を発話する必要がなく簡単な訂正が可能となる。

【 0 1 2 2 】

30

また、単に「次候補」と発話された場合にも「~の次候補」の場合と同様に、次候補の認識結果語彙を提示しつつ訂正処理を行う。よって、「次候補」のコマンド操作によっても、ユーザーは訂正のための語彙を発話する必要がなく、簡単な訂正が可能となる。

【 0 1 2 3 】

ただし、ジャンルと地域名に関する認識結果語彙 D c がジャンルファイル F J 及び地域名ファイル F K に記憶された後、ジャンルファイル F J 中の認識結果語彙を訂正すべきコマンド操作がなされた場合には、次に述べる訂正処理が行われる。

【 0 1 2 4 】

まず、上記の「~を訂正」、「~の次候補」のいずれかのコマンド操作に応じて、ジャンルファイル F J 中の認識結果語彙 D c 1 1 , D c 1 2 ~ を訂正することになる。次に、ジャンルファイル F J 中の最も類似度の高い認識結果語彙（訂正後の認識結果語彙）D c 1 1 に従属する地域名辞書 M J 中の認識対象語彙 D b を選択設定する。そして、分析データ記憶部 1 0 に既に記憶されているジャンルの分析データ D a ' と選択設定した認識対象語彙 D b とを照合することにより、新たな認識結果語彙 D c を求めて地域名ファイル F K に記憶する。

40

【 0 1 2 5 】

したがって、地域名より上位概念のジャンルを訂正すべきコマンド操作がなされた場合には、既に記憶されている上記分析データ D a ' を利用して再度の音声認識を自動的に行うので、ユーザーは地域名の語彙を発話しなくとも、地域名ファイル F K 中の認識結果語彙 D c 2 1 , D c 2 2 ~ を訂正することができる。

50

【0126】

尚、このように自動的に地域名ファイルF K中の認識結果語彙を訂正した場合でも誤認識等が生じる場合が可能性としては存在するが、地域名の範疇に属する語彙をユーザーが発話した時に得られた分析データD a'に基づいて上記再度の音声認識を行うので、ユーザーの所望する地域名を高確率で音声認識することができる。

【0127】

この結果、既に音声入力された分析データD a'を有効利用することで、訂正処理の高速化が実現されると共に、ユーザーに対して操作性の向上を図ることができる。

【0128】

このようにして認識結果記憶領域M R中に絞り込み条件を満足する認識結果語彙D cが揃うと、ステップS 1 2 4における判断が「Y E S」となり、図9に示すステップS 1 2 7の処理へ移行する。 10

【0129】

ステップS 1 2 7では、図3に示したジャンルファイルF Jと地域名ファイルF K中に記憶されている絞り込み条件に適合した認識結果語彙D cに該当する施設名称辞書M Lを設定する。例えば、絞り込み条件に適合した認識結果語彙D c 1 1とD c 2 1が、それぞれ「駅名」と「東京都」であった場合には、図2に示した「東京都内の駅名」に従属する「駅名称の一覧」の施設名称辞書M Lが絞り込まれて設定されることになる。また、「～の次候補」のコマンド操作がなされた場合には、次候補で確定した認識結果語彙に基づいて絞り込みが行われることになる。 20

【0130】

次に、ステップS 1 2 8において、分析データ記憶部10に既に記憶されていた、施設名称（例えば、「目黒駅」）の分析データD a'を類似度演算部9に導入し、この分析データD a'と、上記の「駅名称の一覧」の施設名称辞書M L中の認識対象語彙D dとの類似度を判定する。そして、類似度の高い認識結果語彙D cを施設名称ファイルF Lに記憶（格納）する。

【0131】

次に、ステップS 1 3 0において、施設名称ファイルF Lに記憶した最も類似度の高い認識結果語彙（図3中に示すD c 3 1）を音声出力部6から合成音声によって提示したり表示部7で表示する。例えば、認識結果語彙D c 3 1が「目黒駅」となった場合には、最終的な判定結果を『目黒駅ですね』という合成音声等で提示する。 30

【0132】

そして、ステップS 1 3 2以降の処理に移行して、最終的な判定結果をユーザーに確認させるための確認ルーチンとなる。

【0133】

まず、ステップS 1 3 2において、最終的な認識結果を更に訂正すべきか否かを示唆するために、『コマンドをどうぞ』という合成音を発声する。

【0134】

これに対してユーザーが、ジャンル又は地域名の訂正をすべきと判断して、「～を訂正」、「～の次候補」の何れかのコマンド操作を行うと、入力された音声の分析データD aを認識し、ステップS 1 3 4において訂正のためのコマンド操作がなされたことを、制御語WWに基づいて判断（「Y E S」と判断）して、ステップS 1 3 6へ移行する。また、音声入力に代えて操作部5から「訂正」等の指示がなされた場合にも同様にステップS 1 3 6へ移行する。 40

【0135】

ステップS 1 3 6では、ステップS 1 2 2（図8参照）と同様の処理により、訂正命令に該当する認識語記憶部11中の認識辞書を選択設定する。ただし、この選択設定した認識辞書中、訂正指示のなされた認識対象語彙は除外して設定する。そして、図8に示したステップS 1 1 0からの処理を繰り返すことにより、ジャンルファイルF J又は地域名ファイルF K中の認識結果語彙を新たな認識結果語彙に訂正する。 50

【 0 1 3 6 】

尚、ステップ S 1 1 0 からの訂正処理については、先に説明したので、ここではその説明を割愛することとする。

【 0 1 3 7 】

ステップ S 1 3 4 においてジャンル又は地域名を訂正するためのコマンド操作がなされず、「施設名称」を訂正するためのコマンド操作がなされた場合には、ステップ S 1 3 8 においてこの「施設名称」を訂正するためのコマンド操作を、制御語 WW に基づいて判断（「YES」と判断）して、ステップ S 1 4 2 へ移行する。

【 0 1 3 8 】

一方、ジャンルと地域名及び施設名称のコマンド操作がなされない場合、例えば訂正コマンド以外の「ルート探索」等のコマンド操作がなされた場合には、ステップ S 1 3 8 において、制御語 WW に基づいて訂正なしと判断（「NO」と判断）して、ステップ S 1 4 0 へ移行する。

10

【 0 1 3 9 】

ステップ S 1 4 0 に移行すると、コマンド操作で指令された動作を開始する。上記例の「ルート検索」のコマンド操作がなされたときは、その「コマンド」の意味する制御語 WW に基づいてファイル F L 中の確定した認識結果語彙 D c を抽出し、その認識結果語彙 D c の施設名称に該当する目的地までの走行経路を探索し、その探索した走行経路を表示部 7 に地図表示する。そして、一の音声認識処理を終了し、再び音声認識開始の指示がなされるまで待機する。

20

【 0 1 4 0 】

一方、ステップ S 1 3 8 において「施設名称」の訂正コマンド操作、すなわち、「～を訂正」又は「訂正」のコマンド操作がなされると、ステップ S 1 4 2 へ移行して、『施設名称をどうぞ』という合成音を発声して、施設名称の範疇に属する新たな語彙を発話するように示唆する。

【 0 1 4 1 】

これに応じてユーザーが新たな施設名称の語前を発話すると（ステップ S 1 4 4 ）、その分析データ D a を生成し（ステップ S 1 4 6 ）、その分析データ D a を分析データ記憶部 1 0 に記憶して（ステップ S 1 4 8 ）、ステップ S 1 2 7 の処理に移行する。

【 0 1 4 2 】

ステップ S 1 2 7 では、今までにジャンルファイル F J と地域名ファイル F K に記憶され確定している認識結果語彙 D c に基づいて絞り込みが行われて、認識語記憶部 1 1 中の施設名称辞書 M L が設定される。ただし、訂正の指示がなされた認識対象語彙は除外して設定する。

30

【 0 1 4 3 】

そして、ステップ S 1 2 8 において、分析データ記憶部 1 0 に記憶されている施設名称の分析データ D a ' と上記の施設名称辞書 M L 中の認識対象語彙 D b を照合して類似度を判定し、類似度の高い認識結果語彙 D c が施設名称ファイル F L に記憶される。

【 0 1 4 4 】

このように、「施設名称」の訂正指示がなされた場合には、ジャンルファイル F J と地域名ファイル F K 中に既に記憶されている認識結果語彙 D c 、すなわち、絞り込み条件を満足している認識結果語彙 D c に基づいて、施設名称辞書 M L を選択設定するので、最初からジャンルや地域名を再設定（訂正や変更）をするのに比して、迅速な訂正処理が可能となる。この結果、ユーザーに対して煩雑な操作を強いることがなく、操作性を向上させることができる。

40

【 0 1 4 5 】

従来技術では、「施設名称」の訂正を行う場合に、上位概念のジャンルや地域名を再度音声入力しなければならなかったのに比べると、大幅な操作性の向上が図れる。また、絞り込み処理が適切になされるように訂正処理が行われるので、音声認識率が向上する。

【 0 1 4 6 】

50

更に、ユーザーが施設名称を訂正する際に、「次候補」又は「～の次候補」のコマンド操作をすると、ステップS 1 4 2～S 1 4 8において、施設名称ファイルF L中に次候補の認識結果語彙が存在するか否か判断し、存在していればその次候補の認識結果語彙を認識結果としてステップS 1 2 7の処理を行うことで、上位概念のジャンルや地域名を再度音声入力しなくとも、絞り込み処理が適切になされるように訂正処理が行われる。また、ユーザーに対して煩雑な操作を強いることがなく、操作性を向上させることができる。

【0 1 4 7】

以上に説明したように、本実施形態によれば、ユーザーの思考特性に合わせた音声認識を行うことができるのに加えて、ユーザーに対して煩雑な操作を強いることなく、簡単且つ迅速な訂正操作（訂正コマンド操作）を提供することができる。

10

【0 1 4 8】

次に、「次候補」又は「～の次候補」のコマンド操作がなされた場合の訂正処理のより具体的な動作例を、図10ないし図12に示すフローチャートを参照して説明しておく。

【0 1 4 9】

尚、図10ないし図12において図8及び図9と同一又は相当する処理については同一のステップ符号で示している。

【0 1 5 0】

ただし、ステップS 1 2 0において、ユーザーからの訂正指示がなされたと判断（「YES」と判断）すると、図12（a）の処理を行った後、ノードD又はEを介して、図10中のステップS 1 2 2又はS 1 1 8の処理を開始する点がより具体的に示されている。更に、図11において、ステップS 1 3 4でユーザーからの訂正指示がなされたと判断（「YES」と判断）すると、図12（b）の処理を行った後、ノードD'又はAを介してステップS 1 3 6又はS 1 2 7の処理を開始する点がより具体的に示されている。

20

【0 1 5 1】

すなわち、「次候補」又は「～の次候補」のコマンド操作に該当する次候補の認識結果語彙の決定は、図12（a）（b）中のステップS 2 0 4とS 3 0 4で行われるようになっている。

【0 1 5 2】

まず、図10中のステップS 1 2 0において、訂正をすべきコマンド操作が行われると、図12（a）に示すステップS 2 0 0の処理に移行する。

30

【0 1 5 3】

ステップS 2 0 0では、単に「次候補」と発話するコマンド操作がなされたか、又は「～の次候補」のコマンド操作がなされたか否か判断し、これらいずれかの次候補のコマンド操作がなされた場合にはステップS 2 0 2において、制御語WWに基づいて、指令されたファイル（F J，F K，F Lのうちの指令されたファイル）中に次候補の認識結果語彙が存在するか否か判断する。

【0 1 5 4】

そして、次候補に該当する認識結果語彙が存在すれば、ステップS 2 0 4に移行して、次候補を認識結果とする。更にノードEを介して図10中のステップS 1 1 8に移行して、その認識結果とした認識結果語彙を提示する。

40

【0 1 5 5】

したがって、「次候補」又は「～の次候補」のコマンド操作がなされた場合には、ステップS 1 1 0～S 1 1 6の処理を行うことなく、ノードEを通じて直接にステップS 1 1 8の処理を行う。これにより、処理の簡素化を実現している。

【0 1 5 6】

また、図12（a）中のステップS 2 0 0とS 2 0 2において「NO」と判断した場合には、ノードDを介してステップS 1 2 2に処理が移行する。これにより、ユーザーは、「次候補」又は「～の次候補」のコマンド操作以外の訂正コマンド操作、すなわち「訂正」又は「～を訂正」のコマンド操作によって、訂正の指令を行うことになる。

【0 1 5 7】

50

次に、図 1 1 中のステップ S 1 3 8 において、施設名称を訂正するための「次候補」又は「～の次候補」のコマンド操作がなされたと判断（「YES」と判断）すると、ステップ S 1 4 2 ～ S 1 4 8 の処理を行ってステップ S 1 2 7 以降の処理を繰り返す。ただし、施設名称の訂正指令が「次候補」又は「～の次候補」のコマンド操作によってなされた場合には、ステップ S 1 4 2 ～ S 1 4 8 の処理において、図 1 2 (b) の処理を行って、ステップ S 1 2 7 以降の処理を繰り返す。

【 0 1 5 8 】

このように、図 1 2 (a) 又は (b) の処理を行うことで、「次候補」又は「～の次候補」のコマンド操作がなされた場合に、次候補の認識結果語彙を決定するようになっている。

10

【 0 1 5 9 】

以上説明したように本実施形態によれば、特に、「次候補」と「～の次候補」のコマンド操作を備えたことで、ユーザーにとって極めて簡単且つ迅速な訂正指令の操作をすることができるという優れた機能を有している。つまり、認識結果記憶領域 MR に記憶されている複数の認識結果語彙 D c は、ユーザーが発話した施設名称やジャンルや地域名の範疇に属する語彙と同一又は類似した認識結果語彙であるため、最も類似度の高い認識結果語彙が誤認識であったとしても、残余の（次候補の）認識結果語彙の中に、ユーザーが発話した語彙と同一の認識結果語彙が存在する可能性が極めて高い。このため、ユーザーが「次候補」又は「～の次候補」のコマンド操作により訂正の指令をすれば、高確率で所望の認識結果語彙を設定することができ、ひいては簡単に訂正の操作を行うことができる。

20

【 0 1 6 0 】

尚、以上に説明した本実施形態では、典型例として、本音声認識装置がユーザーに対して発話すべき語前を示唆し、それに従って、ユーザーが語前を 1 つ 1 つ音声入力する場合を説明した。しかし、本発明はこれに限定されるものではない。本音声認識装置がユーザーに対して発話すべき複数の語彙を示唆し、それに従って、ユーザーが複数の語彙を連続音声で音声入力することが可能である。

【 0 1 6 1 】

すなわち、複数の語彙を示唆する際に、認識語記憶部 1 1 中の認識辞書のうち、それらの語彙の属性（範疇）に該当する認識辞書を設定しておき、その設定した認識辞書中の認識対象語彙 D b と複数の分析データ D a を照合して、図 4 に例示した制御語 WW を用いてユーザーからの指示や指令を受けることにより、複数の語彙を制御語 WW によって一括して受け付けることができる。

30

【 0 1 6 2 】

つまり、制御語 WW は、図 4 (c) ～ (e) に示したように、基本型を複数個繋げることで、実現できるので、複数の語彙が音声入力された場合に、これら複数の語彙を一括して制御語 WW とすることができ、ユーザーから指令された処理をその制御語 WW に基づいて一括処理したり、分割処理することが可能である。このように、制御語 WW を用いてユーザーの音声入力を受け付けるようにしたことによって、極めて融通性の高い処理が可能となっている。

【 0 1 6 3 】

また、以上に説明した制御命令語 Y の語彙は便法として示したものである。よって、他の語彙でもよい。例えば、「訂正」の語彙は、「違う」「直せ」等の語彙でもよいし、「次候補」の語彙は、「次ぎ」「候補」「候補変更」等の語彙でもよい。

40

【 0 1 6 4 】

【発明の効果】

以上説明したように本発明によれば、各処理機能毎に対応付けられた制御語を予め備え、前記一の処理機能を提示した後、前記提示に応じて前記一の処理機能を指定する指令情報を有する音声入力情報が外部より入力されると、前記音声入力情報を認識し、前記指令情報に対応する前記制御語に基づいて、前記一の処理機能の動作を行うようにしたので、予め制御語に多様な処理機能を設定しておくことで、多様な処理を迅速・適切に実施するこ

50

とが可能となる。この結果、対話型操作を行うのに優れた効果が得られる。

【図面の簡単な説明】

【図 1】本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。

【図 2】認識辞書の階層化構造を示す説明図である。

【図 3】認識結果記憶領域中のファイル構造を示す説明図である。

【図 4】制御語の構造を示す説明図である。

【図 5】制御語の種類を例示した説明図である。

【図 6】本音声認識装置の基本的な動作例を説明するためのフローチャートである。

【図 7】本音声認識装置の訂正処理の際の基本的な動作例を説明するためのフローチャートである。

10

【図 8】本音声認識装置のより具体的な動作例を説明するためのフローチャートである。

【図 9】本音声認識装置のより具体的な動作例を更に説明するためのフローチャートである。

【図 10】本音声認識装置のより具体的な動作例を更に説明するためのフローチャートである。

【図 11】本音声認識装置のより具体的な動作例を更に説明するためのフローチャートである。

【図 12】本音声認識装置のより具体的な動作例を更に説明するためのフローチャートである。

【図 13】従来の音声認識装置の動作を説明するための説明図である。

20

【符号の説明】

1 ... 音声入力部

2 ... 信号処理部

3 ... 記憶部

4 ... 制御部

5 ... 操作部

6 ... 音声出力部

7 ... 表示部

8 ... 音声分析部

9 ... 類似度演算部

30

10 ... 分析データ記憶部

11 ... 認識語記憶部

M R ... 認識結果記憶領域

M J ... ジャンル辞書

M K ... 地域名辞書

M L ... 施設名称辞書

M C ... 制御語辞書

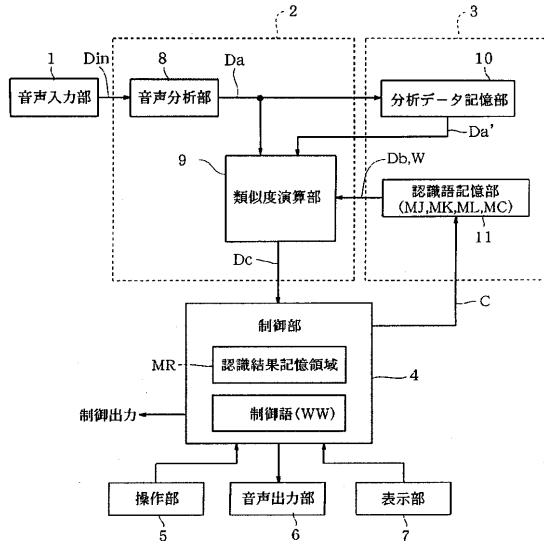
F J ... ジャンルファイル

F K ... 地域名ファイル

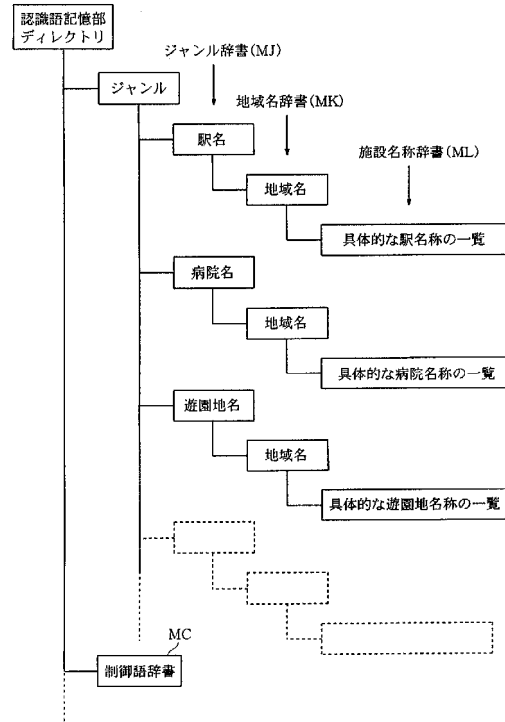
F L ... 施設名称ファイル

40

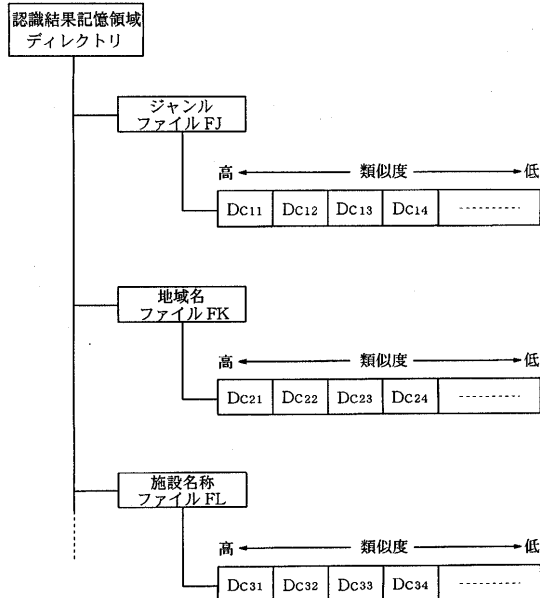
【図 1】



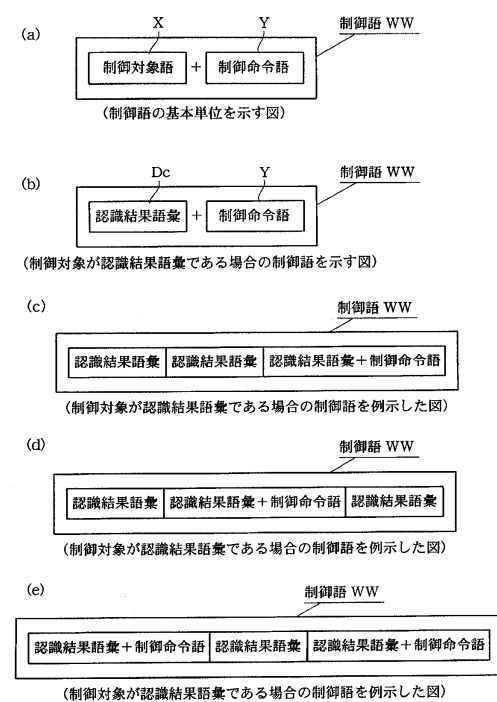
【図 2】



【図 3】



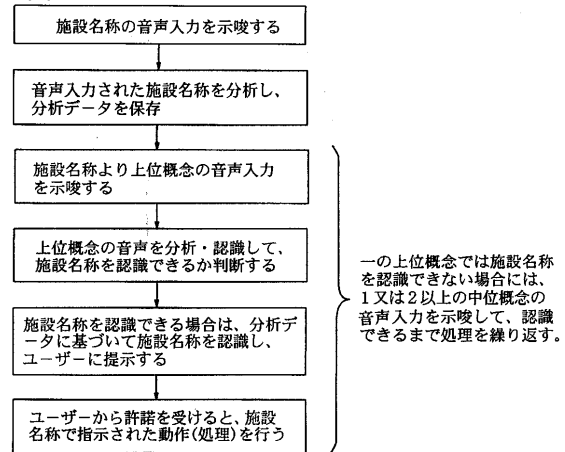
【図 4】



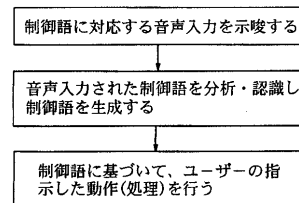
【図 5】

制御語	識別子	制御命令の意味
～の詳細情報	CD	「～」の詳細な情報を提示
～のルート表示	RD	「～」までの走行経路を探索して表示
～の拡大表示	LD	「～」を拡大して地図表示
～の縮小表示	SD	「～」を縮小して地図表示
～の変更	CG	「～」の認識結果語彙を変更
～を削除	DW	「～」の認識結果語彙を削除
～を追加	AW	「～」の語彙を新たに追加
～を訂正	CW	「～」の認識結果語彙を訂正
訂正	CW	直前の認識結果語彙を訂正
～の次候補	NW	「～」の認識結果語彙を次候補の語彙に変更
次候補	NW	直前の認識結果語彙を次候補の語彙に変更

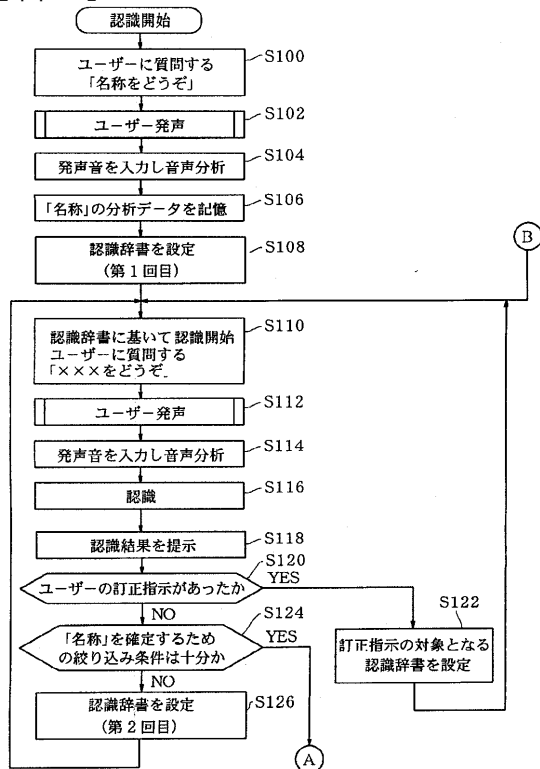
【図 6】



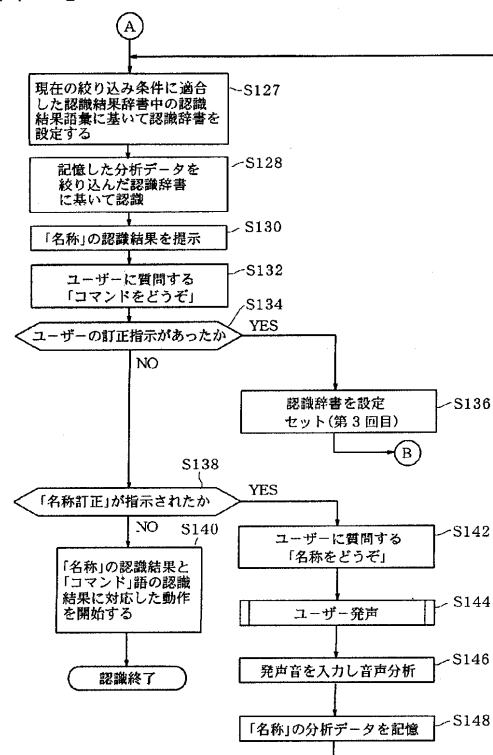
【図 7】



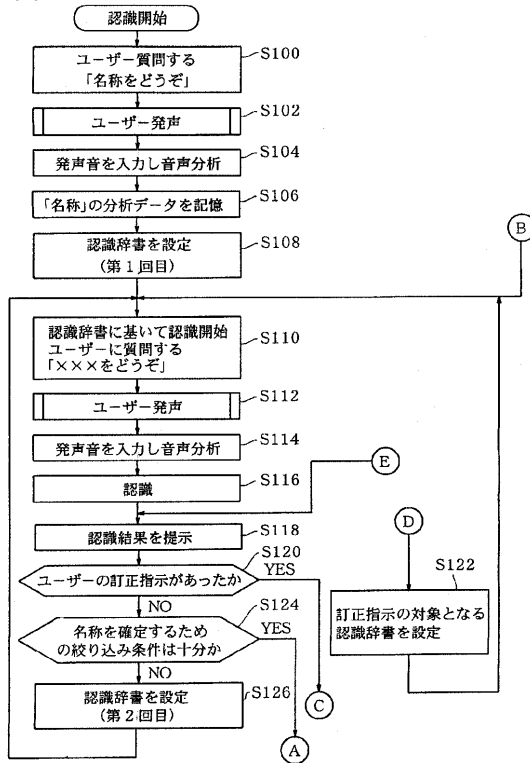
【図 8】



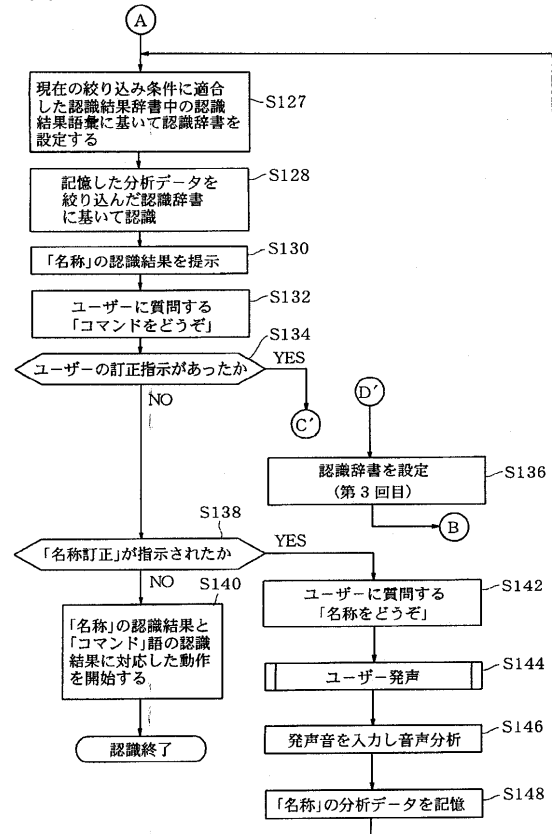
【図 9】



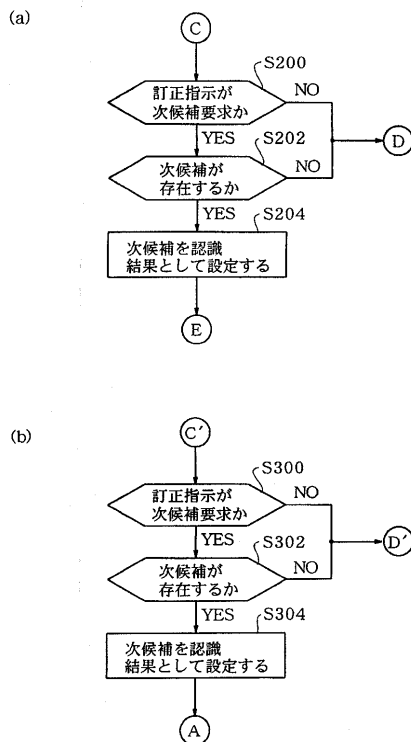
【図 10】



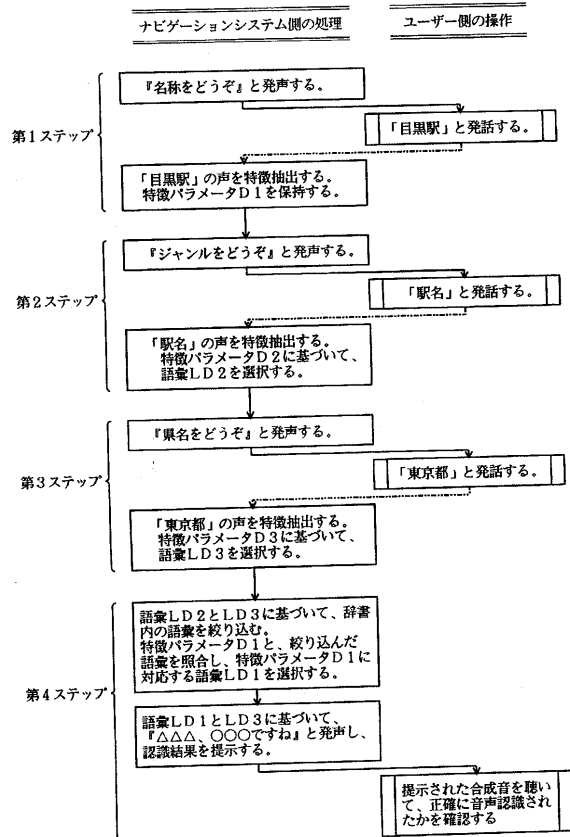
【図 11】



【図 12】



【図 13】



フロントページの続き

- (56)参考文献 特開平11-338494(JP,A)
特開2000-259178(JP,A)
特開平02-126300(JP,A)
特公平05-052518(JP,B2)
特開平07-248792(JP,A)
特開昭63-038996(JP,A)
特開平08-054894(JP,A)
特開平05-108091(JP,A)
特開昭62-209597(JP,A)
特開平10-282989(JP,A)
特開平10-312193(JP,A)
特開平09-274497(JP,A)
特開平09-198087(JP,A)
特開2000-259177(JP,A)
特開2000-293195(JP,A)
特開平11-175094(JP,A)

- (58)調査した分野(Int.Cl., DB名)

G10L 11/00-21/06