

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4201411号
(P4201411)

(45) 発行日 平成20年12月24日(2008.12.24)

(24) 登録日 平成20年10月17日(2008.10.17)

(51) Int.Cl. F I
G 1 O L 13/00 (2006.01)
 G 1 O L 13/00 1 O O C
 G 1 O L 13/00 1 O O H
 G 1 O L 13/00 1 O O M

請求項の数 11 (全 14 頁)

(21) 出願番号	特願平10-359211	(73) 特許権者	000004260
(22) 出願日	平成10年12月17日(1998.12.17)		株式会社デンソー
(65) 公開番号	特開2000-181488(P2000-181488A)		愛知県刈谷市昭和町1丁目1番地
(43) 公開日	平成12年6月30日(2000.6.30)	(74) 代理人	100096998
審査請求日	平成17年1月7日(2005.1.7)		弁理士 碓水 裕彦
審判番号	不服2007-7031(P2007-7031/J1)	(74) 代理人	100123191
審判請求日	平成19年3月8日(2007.3.8)		弁理士 伊藤 高順
		(72) 発明者	宮内 英夫
			愛知県刈谷市昭和町1丁目1番地 株式会
			社デンソー内
		(72) 発明者	赤堀 一郎
			愛知県刈谷市昭和町1丁目1番地 株式会
			社デンソー内

最終頁に続く

(54) 【発明の名称】 音声認識装置及びナビゲーションシステム

(57) 【特許請求の範囲】

【請求項1】

音声を入力するための音声入力手段と、

該音声入力手段を用いて認識対象となる音声を入力させる期間の開始及び終了を利用者自身が指定するための所定の操作を行うための入力期間指定手段と、

該入力期間指定手段によって指定された入力期間内の前記音声入力手段を介して入力された音声を、予め辞書手段に記憶されている複数の比較対象パターン候補と比較して一致度合の高いものを認識結果とする認識手段と、

該認識手段による認識結果を報知する報知手段と、

該報知手段によって認識結果が報知された後に所定の確定指示がなされた場合には、当該認識結果を確定したものとして所定の確定後処理を実行する確定後処理手段と、

を備える音声認識装置であって、

前記入力期間指定手段には、前記音声入力期間の開始及び終了の指定操作とは異なる2つ以上のガイド要求操作が設定され、当該ガイド要求操作それぞれに対して、熟知度合に応じたガイド要求レベルが設定されており、

前記入力期間指定手段によるガイド要求操作がされた場合には、対応するガイド要求レベルに応じた所定のガイド内容を報知するガイド手段を備え、

前記ガイド手段は、前記ガイド内容として、前記音声入力手段を介した指示が可能な設定項目自体を報知すること、を特徴とする音声認識装置。

【請求項2】

10

20

請求項 1 に記載の音声認識装置において、

前記入力期間指定手段による音声入力期間の開始及び終了の指定操作は、スイッチの操作開始から終了までが所定時間以上となる場合の、その操作開始及び終了にそれぞれ対応しており、

一方、前記ガイド要求レベルとしては、初心者レベル及び中級者レベルの 2 つがあり、当該 2 つのレベルに対するガイド要求操作は、スイッチの操作開始及び終了が所定のクリック時間内に行われるクリック操作と、前記クリック操作が所定間隔以内に 2 回連続して行われるダブルクリック操作が、各々いずれかのレベルの操作に対応していること、を特徴とする音声認識装置。

【請求項 3】

10

請求項 2 に記載の音声認識装置において、

前記ガイド手段は、前記初心者用のガイドにおいて、所定のガイド内容を報知した後、音声入力が行われることなくさらにクリック操作がなされた場合、従前に報知したガイド内容よりも詳細なガイド内容を報知すること、を特徴とする音声認識装置。

【請求項 4】

請求項 1 ～ 3 のいずれかに記載の音声認識装置において、

前記報知手段が前記認識結果を少なくとも音声にて報知するよう構成されていると共に、前記ガイド手段が前記所定のガイド内容を少なくとも音声にて報知するよう構成されていること、を特徴とする音声認識装置。

【請求項 5】

20

請求項 1 ～ 4 のいずれかに記載の音声認識装置において、

前記ガイド手段が報知する前記所定のガイド内容は、本音声認識装置による音声認識結果が利用されるシステムでの処理に関する内容であること、を特徴とする音声認識装置。

【請求項 6】

請求項 5 に記載の音声認識装置において、

前記ガイド手段は、前記所定のガイド内容を報知する時点での前記システムでの処理実行段階に対応したガイド内容を報知するよう構成されていること、を特徴とする音声認識装置。

【請求項 7】

音声を入力するための音声入力手段と、

30

該音声入力手段を用いて認識対象となる音声を入力させる期間の開始及び終了を利用者自身が指定するための所定の操作を行うための入力期間指定手段と、

該入力期間指定手段によって指定された入力期間内の前記音声入力手段を介して入力された音声を、予め辞書手段に記憶されている複数の比較対象パターン候補と比較して一致度合の高いものを認識結果とする認識手段と、

該認識手段による認識結果を報知する報知手段と、

該報知手段によって認識結果が報知された後に所定の確定指示がなされた場合には、当該認識結果を確定したものであるとして所定の確定後処理を実行する確定後処理手段と、

を備える音声認識装置であって、

前記入力期間指定手段には、前記音声入力期間の開始及び終了の指定操作とは異なる 2 つ以上のガイド要求操作が設定され、当該ガイド要求操作それぞれに対して、熟知度合に応じたガイド要求レベルが設定されており、

40

前記入力期間指定手段によるガイド要求操作がされた場合には、対応するガイド要求レベルに応じた所定のガイド内容を報知するガイド手段を備え、

前記ガイド手段が報知する前記所定のガイド内容は、本音声認識装置による音声認識結果が利用されるシステムでの処理に関する内容であり、

前記ガイド手段は、前記所定のガイド内容を報知する時点での前記システムでの処理実行段階に対応したガイド内容を報知するよう構成され、

前記ガイド手段は、前記システムでの処理実行段階に対応したガイド内容として、その処理実行段階において前記音声入力手段を介した指示が可能な設定項目自体を報知するこ

50

と、を特徴とする音声認識装置。

【請求項 8】

請求項 7 に記載の音声認識装置において、

前記入力期間指定手段による音声入力期間の開始及び終了の指定操作は、スイッチの操作開始から終了までが所定時間以上となる場合の、その操作開始及び終了にそれぞれ対応しており、

一方、前記ガイド要求レベルとしては、初心者レベル及び中級者レベルの 2 つがあり、当該 2 つのレベルに対するガイド要求操作は、スイッチの操作開始及び終了が所定のクリック時間内に行われるクリック操作と、前記クリック操作が所定間隔以内に 2 回連続して行われるダブルクリック操作が、各々いずれかのレベルの操作に対応していること、を特徴とする音声認識装置。

10

【請求項 9】

請求項 8 に記載の音声認識装置において、

前記ガイド手段は、前記初心者用のガイドにおいて、所定のガイド内容を報知した後、音声入力が行われることなくさらにクリック操作がなされた場合、従前に報知したガイド内容よりも詳細なガイド内容を報知すること、を特徴とする音声認識装置。

【請求項 10】

請求項 7 ～ 9 のいずれかに記載の音声認識装置において、

前記報知手段が前記認識結果を少なくとも音声にて報知するよう構成されていると共に、前記ガイド手段が前記所定のガイド内容を少なくとも音声にて報知するよう構成されていること、を特徴とする音声認識装置。

20

【請求項 11】

請求項 1 ～ 10 のいずれかに記載の音声認識装置と、ナビゲーション装置とを備え、

前記音声認識装置の前記音声入力手段は、少なくとも前記ナビゲーション装置がナビゲート処理をする上で指定される必要のある所定のナビゲート処理関連データの指示を利用者が音声にて入力するために用いられるものであり、前記確定後処理手段は、前記認識手段による認識結果を前記ナビゲーション装置に出力するよう構成されていること、を特徴とするナビゲーションシステム。

【発明の詳細な説明】

【0001】

30

【発明の属する技術分野】

本発明は、例えばナビゲーションシステムにおける目的地の設定などを音声入力できるようにする場合などに有効な音声認識装置及びその音声認識装置を備えたナビゲーションシステムに関する。

【0002】

【従来の技術】

従来より、入力された音声を予め記憶されている複数の比較対象パターン候補と比較し、一致度合の高いものを認識結果とする音声認識装置が既に実用化されており、例えばナビゲーションシステムにおいて設定すべき目的地を利用者が地名を音声で入力するためなどに用いられている。特に車載ナビゲーションシステムを運転手自身が利用する場合、音声入力であればボタン操作や画面注視が伴わないため、車両の走行中に行っても安全性が高いため有効である。

40

【0003】

また、音声認識装置側が常に音声入力に備えるように準備しておくのは負担が大きいため、認識対象となる音声を入力させる期間の開始及び終了タイミングについては利用者自身が指定するような構成が採用されることが多い。この入力期間を指定するための手段としては、例えば P T T (Push-To-Talk) スイッチなどが利用される。つまり、P T T スイッチが押されている場合にだけ入力音声に対する認識処理を実行するようにして、実際上必要がない入力音声については処理をしないようにしている。

【0004】

50

このように利用価値の高い音声入力の手法であるが、利用者が音声認識のシステム自体に慣れていない初心者であると、次のような問題点が生じる。例えば、利用方法が判らないことに起因して、装置にとっては不適切な音声を適当に入力してしまい実効性が上がらないという問題や、あるいは、極端な場合には装置自体を使わないという問題がある。また、利用方法自体をある程度は理解しているが具体的な利用方法に慣れていないため、例えば入力すべき言葉がすぐに出てこなくて沈黙してしまうという問題もある。

【0005】

このような初心者のための対策として、例えば所定のガイド内容を報知することも考えられる。例えば入力できるコマンドとしてどのようなものがあるのかをガイドすることが考えられる。具体的には、「目的地の設定、地図拡大・縮小が行えます」というようなガイド内容を報知する。あるいは、コマンドの具体的な入力方法を知らない場合に「目的地を設定するときは都道府県名から入力して下さい」という入力方法の説明をガイド内容として報知することも考えられる。

10

【0006】

【発明が解決しようとする課題】

しかしながら、このように初心者向けの対策を講じた場合、初心者にとっては有益なガイド内容も、熟練者にとっては不要なものとなる。つまり、そのようなガイド内容を聞かなくても入力できるコマンドや入力方法自体を熟知している場合には、そのガイド内容が出力されている間は音声入力ができなくなるなどの不都合が生じてくる。

【0007】

20

また、全くガイド内容を聞かなくてよい熟練者のレベルまではいかないが、概略的なガイドだけしてもらえば後は自分で対処できる「中級者」レベルの場合には、初心者と同じ内容のガイドは不要である。そして、これら初心者、中級者、熟練者は必ずしも別人ではなく、同じ人が、最初は初心者であり、その後、中級者、熟練者とステップアップしていく。したがって、装置側としても、初心者用、中級者用、熟練者用というように使用者のレベルに合わせて固定的にシステム設定することは好ましくなく、使用時の状況に応じて、初心者には初心者用の入力方法、中級者には中級者用の入力方法、熟練者には熟練者用の入力方法を任意に選択できるようなシステム設定とすることが期待される。

【0008】

本発明は、このような問題を解決し、音声認識装置の利用方法に対する熟知度合が異なる各利用者のいずれにとっても使い勝手の良い音声認識装置及びその音声認識装置を備えたナビゲーションシステムを提供することを目的とする。

30

【0009】

【課題を解決するための手段及び発明の効果】

上記目的を達成するためになされた請求項1、7に記載の音声認識装置によれば、利用者は例えばマイクロフォンなどの音声入力手段を介して音声を入力するのであるが、認識対象の音声を入力させる期間の開始及び終了を利用者自身が指定するために入力期間指定手段が設けられており、この入力期間指定手段によって指定された入力期間内に入力された音声は認識対象となる。そして、認識手段が、その入力された音声を予め辞書手段に記憶されている複数の比較対象パターン候補と比較して一致度合の高いものを認識結果とし、報知手段によって認識結果を報知する。そして、認識結果が報知された後に所定の確定指示がなされた場合には、確定後処理手段が、その認識結果を確定したものとして所定の確定後処理を実行する。

40

【0010】

そして、本発明の音声認識装置においては、2つ以上のガイド要求操作が設定されている。このガイド要求操作は、入力期間指定手段には、前記音声入力期間の開始及び終了の指定操作とは異なる2つ以上のガイド要求操作が設定されており、当該ガイド要求操作それぞれに対して、熟知度合に応じたガイド要求レベルが設定されている。そして、ガイド手段は、入力期間指定手段によるガイド要求操作がされた場合には、対応するガイド要求レベルに応じた所定のガイド内容を報知する。ここで請求項1の発明では、前記ガイド手

50

段は、前記ガイド内容として、前記音声入力手段を介した指示が可能な設定項目自体を報知する。また、請求項7の発明では、前記ガイド手段が報知する前記所定のガイド内容は、本音声認識装置による音声認識結果が利用されるシステムでの処理に関する内容であり、前記ガイド手段は、前記所定のガイド内容を報知する時点での前記システムでの処理実行段階に対応したガイド内容を報知するよう構成され、前記ガイド手段は、前記システムでの処理実行段階に対応したガイド内容として、その処理実行段階において前記音声入力手段を介した指示が可能な設定項目自体を報知する。

【0011】

ここで、「音声入力期間の開始及び終了の指定操作とは異なるガイド要求操作」とは、例えば次のような場合を言う。例えば、入力期間指定手段による音声入力期間の開始及び終了の指定操作を、スイッチの操作開始から終了までが所定時間以上となる場合の、その操作開始及び終了にそれぞれ対応させた場合を想定する。この場合、スイッチ操作開始後に音声入力を行い、その後に操作終了となるため、スイッチ操作が継続している時間が例えば1秒以下となるようなことは想定しにくい。したがって、例えばスイッチ操作開始から終了までの時間が例えば0.5秒以下というような比較的短いクリック時間となる、いわゆる「クリック操作」を上述の「音声入力期間の開始及び終了の指定操作とは異なるガイド要求操作」としておけばよい。もちろん、装置側において通常の音声入力期間の開始及び終了の指定ではありえない様な特別の操作であれば、これ以外の操作もガイド要求のための操作であるとしてよい。例えば、クリック操作が所定間隔以内に2回連続して行われる「ダブルクリック操作」であってもよい。

【0012】

このようにすれば、当該ガイド要求操作が2つ以上設定されている場合には、ガイドが必要な利用者は、自分の熟知度合に応じたガイド要求レベルでのガイド内容となるよう、対応するガイド要求操作をすればよく、一方、ガイドが必要ない利用者は、ガイド要求操作をせずに通常通り、入力期間指定手段による音声入力期間の開始及び終了の指定操作をしながら所定の音声入力をすればよい。

【0013】

この場合の所定のガイド内容としては種々考えられるので、理解を容易にするため、ここでは、ナビゲーションシステムにおける目的地などを音声入力するために音声認識装置を適用した場合を想定していくつか説明する。

▲1▼まず、ナビゲーションシステムに対して自分が音声で何を入力することができるのか自体を知らない利用者が想定される。この場合にはクリック操作をすることで、例えば「目的地の設定が行えます」というように、音声入力によって指示が可能な設定項目自体を案内するものが考えられる。なお、音声入力によって指示する対象のシステム（この場合にはナビゲーションシステム）における一連の処理のどの段階にあるかで内容を変更してもよい。例えば、目的地までの経路設定を一連の処理と考えると、上述した目的地の設定がされた後、例えば経由地の指定はするのか、というような経路を設定する上での条件をさらに指定する場合がある。したがって、目的地の設定が済んでいる状況で、上述した音声入力期間の開始が指定された時点から所定時間経過しても音声入力がない場合には、今度は「経由地の指定が行えます」というようなガイド内容を報知すればよい。

【0014】

▲2▼また、ナビゲーションシステムに対して自分が音声で何を入力することができるのか自体は知っているが、具体的な入力方法を熟知していない利用者也想定される。この場合には、「目的地を設定するときは都道府県名から入力して下さい」というように、入力方法の説明をガイド内容として報知することが考えられる。また、これでも利用者がどのように入力するかを完全には理解できず、入力できずにいる場合には、「例えば愛知県刈谷市昭和町と入力して下さい」というように、具体的な入力例をガイド内容として報知することが考えられる。これによって利用者は具体的な入力方法が判り、また、具体例まで報知してもらえばそれに倣って自分が希望する目的地を容易に入力することができる。

【0015】

もちろん、上記▲1▼及び▲2▼で説明したガイド内容は一例であり、適用するナビゲーションシステムの処理に適合した適切なガイド内容を出せばよいし、またナビゲーションシステム以外に適用するのであれば、そのシステムの処理に適合するように工夫すればよい。

【0016】

また、ガイドが必要な対象者として、初心者レベル及び中級者レベルの2つを持つようにした場合に、初心者レベルの場合にどのようなガイド内容にし、中級者の場合にどのようなガイド内容とするか、などについても、当該システムの使用方法の難易性に応じて適宜設定すればよい。例えば、初心者レベルではガイド要求操作がある度に、A→B→Cという順番でガイド内容を報知するようにされている場合に、中級者レベルではA、B、Cのガイド内容をまとめて報知してしまうことも考えられる。もちろん、適用するシステムによっては、ガイドが必要な対象者を3つ以上のレベルに分類してもよい。

10

【0017】

一方、ガイド要求操作をしない限りガイド内容が勝手に報知されてしまうことはないため、ガイドが必要でない利用者にとっては、入力期間指定手段による音声入力期間の開始及び終了の指定操作をしながら所定の音声入力をすれば、相対的に（つまりガイド内容を報知させている状態よりも）短時間で所望の音声入力を行うことができる。

【0018】

このように、本発明の音声認識装置によれば、音声認識装置の利用方法に対する熟知度合が異なる各利用者のいずれにとっても使い勝手が良くなる。その上、ガイド要求操作を受け付けるために特別な構成を必要とするのではなく、音声入力のための基本的な処理に必要な入力期間指定手段をガイド要求操作にも利用できるため、構成の簡略化にも寄与する。

20

【0019】

なお、ガイドが必要な対象者として、初心者レベル及び中級者レベルの2つを持つようにした場合のそれぞれのガイド要求操作としては、請求項2、8に示すようにすることが考えられる。すなわち、入力期間指定手段による音声入力期間の開始及び終了の指定操作を、スイッチの操作開始から終了までが所定時間以上となる場合の、その操作開始及び終了にそれぞれ対応させることを前提とし、初心者レベル及び中級者レベルに対するガイド要求操作として、クリック操作とダブルクリック操作を、各々いずれかのレベルの操作に対応させて設定するのである。

30

【0020】

なお、請求項3、9に示すように、初心者用のガイドにおいて、所定のガイド内容を報知した後、音声入力が行われることなくさらにクリック操作がなされた場合、従前に報知したガイド内容よりも詳細なガイド内容を報知するようにしてもよい。

【0021】

もちろん、これらは一例であり、操作内容の違いが装置側で区別できればよい。なお、報知手段による認識結果の報知形態、あるいはガイド手段によるガイド内容の報知形態については、利用者に対して報知が可能であればどのような形態でもよいが、請求項4、10に示すように、少なくとも音声による報知を行なうことが好ましいと考えられる。これは、例えばカーナビゲーションシステムなどの車載機器用として用いる場合には、音声で出力されればドライバーは視点を表示装置にずらしたりする必要がないので、安全運転のより一層の確保の点では有利であることなどの理由からである。また、認識結果の報知を音声で行ない、かつガイド内容の報知も音声で行なえば、それらのためのハード構成を共通化することができる。但し、音声に加え例えば画像で報知するようにしてもよい。車載機器として適用する場合に音声出力が有利であることを述べたが、もちろん車両が走行中でない状況もあるので、音声及び画像の両方で報知すれば、ドライバーは表示による確認と音声による確認との両方が可能となる。

40

【0022】

また、請求項1～10のいずれかに記載の音声認識装置をナビゲーションシステム用と

50

して用いる場合には、請求項 11 に示すように構成することが考えられる。すなわち、請求項 1 ～ 10 のいずれかに記載の音声認識装置と、ナビゲーション装置とを備え、音声認識装置の音声入力手段は、少なくともナビゲーション装置がナビゲート処理をする上で指定される必要のある所定のナビゲート処理関連データの指示を利用者が音声にて入力するために用いられるものであり、確定後処理手段は、認識手段による認識結果をナビゲーション装置に出力するよう構成されているのである。この場合の「所定のナビゲート処理関連データ」としては、目的地が代表的なものとして挙げられるが、それ以外にもルート探索に関する条件選択など、ナビゲート処理をする上で指定の必要のある指示が含まれる。

【0023】

なお、音声認識装置の適用先としては、上述したナビゲーションシステムには限定されない。例えば音声認識装置を空調システム用として用いる場合には、設定温度の調整、空調モード（冷房・暖房・ドライ）の選択、あるいは風向モードの選択を音声入力によって行うようにすることが考えられる。そして、この場合には、その設定項目（温度・空調モード・風向モードなど）自体をガイド内容として報知したり、あるいは、「設定温度を 25 度にする」と言えばよいのか「設定温度を 5 度下げる」というように言えばよいのか、などをさらにガイド内容として報知することが考えられる。空調モードや風向モードなどについても同様である。

【0024】

なお、上述のナビゲーションシステム及び空調システムは、車載機器として用いられる場合だけではなく、例えば携帯型ナビゲーション装置や屋内用空調装置などでもよい。但し、これまで説明したように車載機器用として用いる場合には利用者がドライバーであることが考えられ、その場合には運転自体が最重要であり、それ以外の車載機器については、なるべく運転に支障がないことが好ましい。したがって、車載機器としてのナビゲーションシステムや空調システムを前提とした音声認識装置の場合には、より一層の利点がある。もちろん、このような視点で考えるならば、ナビゲーションシステムや空調システム以外の車載機器に対しても同様に利用することができる。例えば、カーオーディオ機器などは有効である。また、いわゆるパワーウィンドウの開閉やミラー角度の調整などを音声によって指示するような構成を考えれば、そのような状況でも有効である。

【0025】

また、車載機器用とした場合にはそれ特有の利点があることは述べたが、本発明の音声認識装置の適用先としては、利用者による音声入力指示にしたがって所定の処理を実行するものであれば同様に考えられる。例えば、携帯用の情報端末装置、あるいは街頭やパーキングエリアなどに設定される情報端末装置などにも同様に適用できる。

【0026】

【発明の実施の形態】

図 1 は本発明の実施形態の音声認識装置 30 を適用したカーナビゲーションシステム 2 の概略構成を示すブロック図である。本カーナビゲーションシステム 2 は、位置検出器 4、地図データ入力器 6、操作スイッチ群 8、これらに接続された制御回路 10、制御回路 10 に接続された外部メモリ 12、表示装置 14 及びリモコンセンサ 15 及び音声認識装置 30 を備えている。なお制御回路 10 は通常のコンピュータとして構成されており、内部には、周知の CPU、ROM、RAM、I/O 及びこれらの構成を接続するバスラインが備えられている。

【0027】

位置検出器 4 は、周知の地磁気センサ 16、ジャイロスコープ 18、距離センサ 20、及び衛星からの電波に基づいて車両の位置を検出する GPS (Global Positioning System) のための GPS 受信機 22 を有している。

これらのセンサ等 16, 18, 20, 22 は各々が性質の異なる誤差を持っているため、複数のセンサにより、各々補間しながら使用するように構成されている。なお、精度によっては上述した内の一部で構成してもよく、更に、ステアリングの回転センサ、各転動輪の車輪センサ等を用いてもよい。

10

20

30

40

50

【0028】

地図データ入力器6は、位置検出の精度向上のためのいわゆるマップマッチング用データ、地図データ及び目印データを含む各種データを入力するための装置である。媒体としては、そのデータ量からCD-ROMを用いるのが一般的であるが、メモリカード等の他の媒体を用いても良い。

【0029】

表示装置14はカラー表示装置であり、表示装置14の画面には、位置検出器4から入力された車両現在位置マークと、地図データ入力器6より入力された地図データと、更に地図上に表示する誘導経路や後述する設定地点の目印等の付加データとを重ねて表示することができる。

10

【0030】

また、本カーナビゲーションシステム2は、リモートコントロール端末（以下、リモコンと称する。）15aを介してリモコンセンサ15から、あるいは操作スイッチ群8により目的地の位置を入力すると、現在位置からその目的地までの最適な経路を自動的に選択して誘導経路を形成し表示する、いわゆる経路案内機能も備えている。このような自動的に最適な経路を設定する手法は、ダイクストラ法等の手法が知られている。操作スイッチ群8は、例えば、表示装置14と一体になったタッチスイッチもしくはメカニカルなスイッチ等が用いられ、各種入力に使用される。

【0031】

そして、音声認識装置30は、上記操作スイッチ群8あるいはリモコン15aが手動操作により目的地などを指示するために用いられるのに対して、利用者が音声で入力することによっても同様に目的地などを指示することができるようにするための装置である。

20

【0032】

この音声認識装置30は、「認識手段」としての音声認識部31と、「確定後処理手段」としての対話制御部32と、音声合成部33と、音声抽出部34と、「音声入力手段」としてのマイク35と、「入力期間指定手段」としてのPTT（Push-To-Talk）スイッチ36と、スピーカ37と、制御部38と、表示制御部39と、表示部40とを備えている。

【0033】

音声認識部31は、音声抽出部34から入力された音声データを、対話制御部32からの指示により入力音声の認識処理を行い、その認識結果を対話制御部32に返す。すなわち、音声抽出部34から取得した音声データに対し、記憶している辞書データを用いて照合を行ない、複数の比較対象パターン候補と比較して一致度の高い上位比較対象パターンを対話制御部32へ出力する。入力音声の単語系列の認識は、音声抽出部34から入力された音声データを順次音響分析して音響的特徴量（例えばケプストラム）を抽出し、この音響分析によって得られた音響的特徴量時系列データを得る。そして、周知のDPマッチング法、HMM（隠れマルコフモデル）あるいはニューラルネットなどによって、この時系列データをいくつかの区間に分け、各区間が辞書データとして格納されたどの単語に対応しているかを求める。

30

【0034】

対話制御部32は、音声認識部31における認識結果や制御部38からの指示に基づき、音声合成部33への応答音声の出力指示や表示制御部39への応答表示の出力指示、あるいは、システム自体の処理を実行する制御回路10に対して例えばナビゲート処理のために必要な目的地を通知して設定処理を実行させるよう指示する処理を実行する。このような処理が確定後処理であり、結果として、この音声認識装置30を利用すれば、上記操作スイッチ群8あるいはリモコン15aを手動しなくても、音声入力によりナビゲーションシステムに対する目的地の指示などが可能となるのである。

40

【0035】

なお、音声合成部33は、波形データベース内に格納されている音声波形を用い、対話制御部32からの応答音声の出力指示に基づく音声を合成する。この合成音声はスピーカ37から出力されることとなる。また、表示制御部39が、対話制御部32からの応答表示

50

の出力指示に基づく表示画像を生成する。この生成画像が表示部 40 に出力されることとなる。

【0036】

音声抽出部 34 は、マイク 35 にて取り込んだ周囲の音声をデジタルデータに変換して音声認識部 31 に出力するものである。詳しくは、入力した音声の特徴量を分析するため、例えば数 10 ms 程度の区間のフレーム信号を一定間隔で切り出し、その入力信号が、音声の含まれている音声区間であるのか音声の含まれていない雑音区間であるのか判定する。マイク 35 から入力される信号は、認識対象の音声だけでなく雑音も混在したものであるため、音声区間と雑音区間の判定を行なう。この判定方法としては従来より多くの手法が提案されており、例えば入力信号の短時間パワーを一定時間毎に抽出していき、所定の閾値以上の短時間パワーが一定以上継続したか否かによって音声区間であるか雑音区間であるかを判定する手法がよく採用されている。そして、音声区間であると判定された場合には、その入力信号が音声認識部 31 に出力されることとなる。

10

【0037】

また、本実施形態においては、利用者が PTT スイッチ 36 を押しながらマイク 35 を介して音声を入力するという使用方法である。具体的には、制御部 38 が PTT スイッチ 36 が押されたタイミングや戻されたタイミング及び押された状態が継続した時間を監視しており、PTT スイッチ 36 が押された場合には音声抽出部 34 及び音声入式部 31 に対して処理の実行を指示する。一方、PTT スイッチ 36 が押されていない場合にはその処理を実行させないようにしている。したがって、PTT スイッチ 36 が押されている間にマイク 35 を介して入力された音声データが音声認識部 31 へ出力されることとなる。

20

【0038】

また、制御部 38 は、PTT スイッチ 36 がクリック操作されたこと、及びダブルクリック操作されたことも判断できるようにされている。具体的には、PTT スイッチ 36 がオンされた後の比較的短い時間（例えば 0.5 秒以内）にオフされた場合にはそれをクリック操作とみなす。そして、そのクリック操作が所定間隔以内（例えば 0.5 秒以内）に 2 回連続して行われた場合にクリック操作とみなす。

【0039】

次に、本実施形態 1 のカーナビゲーションシステム 2 の動作について説明する。なお、音声認識装置 30 に関係する部分が特徴であるので、カーナビゲーションシステム 2 としての一般的な動作を簡単に説明した後、音声認識装置 30 に関係する部分の動作について詳しく説明することとする。

30

【0040】

カーナビゲーションシステム 2 の電源オン後に、表示装置 14 上に表示されるメニューから、ドライバーがリモコン 15 a（操作スイッチ群 8 でも同様に操作できる。以後の説明においても同じ）により、案内経路を表示装置 14 に表示させるために経路情報表示処理を選択した場合、あるいは、音声認識装置 30 を介して希望するメニューをマイク 35 を介して音声入力することで、対話制御部 32 から制御回路 10 へ、リモコン 15 a を介して選択されるのを同様の指示がなされた場合、次のような処理を実施する。

【0041】

すなわち、ドライバーが表示装置 14 上の地図に基づいて、音声あるいはリモコンなどの操作によって目的地を入力すると、GPS 受信機 22 から得られる衛星のデータに基づき車両の現在地が求められ、目的地と現在地との間に、ダイクストラ法によりコスト計算して、現在地から目的地までの最も短距離の経路を誘導経路として求める処理が行われる。そして、表示装置 14 上の道路地図に重ねて誘導経路を表示して、ドライバーに適切なルートを案内する。このような誘導経路を求める計算処理や案内処理は一般的に良く知られた処理であるので説明は省略する。

40

【0042】

次に、音声認識装置 30 における動作について、上述の経路案内のための目的地を音声入力する場合を例にとって説明する。

50

まず、最初のステップ S 1 1 0 では、P T T スイッチ 3 6 がオンされたか（押下されたか）否かを判断する。この判断は制御部 3 8 で行われる。そして、P T T スイッチ 3 6 がオンされた場合には（S 1 1 0：Y E S）、続く S 1 2 0 でクリック操作かどうかを判断する。この判断も制御部 3 8 にて行われるが、具体的には、P T T スイッチ 3 6 がオンされた後の比較的短い時間（例えば 0. 5 秒以内）にオフされた場合にはそれをクリック操作とみなす。クリック操作がされた場合には（S 1 2 0：Y E S）、続く S 1 3 0 でダブルクリック操作かどうかを判断する。この判断も制御部 3 8 にて行われるが、具体的には、クリック操作が所定間隔以内（例えば 0. 5 秒以内）に 2 回連続して行われた場合にはそれをダブルクリック操作とみなす。

【0 0 4 3】

10

S 1 3 0 にて否定判断であればクリック操作がなされたと判断し、S 1 4 0 へ移行して、初心者用のガイド内容を報知する。一方、S 1 3 0 にて肯定判断であればダブルクリック操作がなされたと判断し、S 1 5 0 へ移行して、中級者用のガイド内容を報知する。

【0 0 4 4】

これら初心者用及び中級者用のガイド内容については、後述することとして、図 2 のフローチャートの説明を続ける。

S 1 2 0 にて否定判断、つまりクリック操作でない場合には、P T T スイッチ 3 6 が所定時間以上押し続けられている状態であるため、S 1 6 0 へ移行して音声抽出及び認識の処理を実行する。音声抽出部 3 4 はマイク 3 5 を介して入力された音声データに基づき音声区間であるか雑音区間であるかを判定し、音声区間のデータを音声認識部へ出力する。音声認識部 3 1 では、音声抽出部 3 4 から取得した音声データに対し、記憶されている辞書データを用いて照合を行なう。そして、その照合結果によって定まった上位比較対象パターンを認識結果として対話制御部 3 2 に出力することとなる。

20

【0 0 4 5】

P T T スイッチ 3 6 のオン状態が継続している間は（S 1 7 0：Y E S）、このような音声抽出・認識（S 1 6 0）を実行し、P T T スイッチ 3 6 がオフされた場合には（S 1 7 0：N O）、続く S 1 8 0 にて、その認識結果をトークバック及び表示する。つまり、対話制御部 3 2 が音声合成部 3 3 及び表示制御部 3 9 を制御し、認識した結果を音声によりスピーカ 3 7 から出力させると共に、表示部 4 0 に認識結果を示す文章を表示させる。

【0 0 4 6】

30

その後、S 1 9 0 にて正しい認識であるかどうかを判断する。これは、利用者からの指示に応じて判断することとなるが、例えば正しい認識であればマイク 3 5 から「はい」を音声入力し、間違った認識であれば「いいえ」を音声入力するようにしておくことが考えられる。もちろん、操作スイッチ群 8 を介してこれらの指示を入力するようにしてもよい。

【0 0 4 7】

そして、誤った認識であれば（S 1 9 0：N O）、S 1 1 0 へ戻るが、正しい認識である場合には（S 1 9 0：Y E S）、S 2 0 0 へ移行して認識結果を確定する。そして続く S 2 1 0 にて、所定の確定後処理を実行する。この場合の確定後処理とは、認識結果としての「経路案内のための目的地」に関するデータを、制御回路 1 0（図 1 参照）へ出力する処理などである。

40

【0 0 4 8】

このような確定後処理（S 2 1 0）が終了した後は、S 1 1 0 へ移行して、P T T スイッチ 3 6 がオンされたか（押下されたか）否かの判断処理へ戻る。

なお、本実施形態の場合には、対話制御部 3 2、音声合成部 3 3、制御部 3 8、表示制御部 3 9 及び表示部 4 0 が「ガイド手段」に相当する。

【0 0 4 9】

ここで、S 1 4 0 及び S 1 5 0 にて報知する初心者用及び中級者用のガイド内容について説明する。

まず、初心者用のガイドについて説明する。

▲ 1 ▼ このガイドは、ナビゲーションシステムに対して自分が音声で何を入力することが

50

できるのか自体を知らない利用者を対象とした内容とされている。そのため、例えば「目的地の設定が行えます」というように、音声入力によって指示が可能な設定項目自体を案内するものが考えられる。

【0050】

なお、音声入力によって指示する対象のシステム（この場合にはナビゲーション装置）における一連の処理のどの段階にあるかで内容を変更してもよい。例えば、目的地までの経路設定を一連の処理と考えると、上述した目的地の設定がされた後、例えば経由地の指定はするのか、というような経路を設定する上での条件をさらに指定する場合がある。したがって、目的地の設定が済んでいる状況でクリック操作がされた場合、今度は「経由地の指定が行えます」というようなガイド内容を報知すればよい。

10

【0051】

▲2▼また、上述したガイドがされて、ナビゲーションシステムに対して何を入力すればよいかが判った利用者であっても、具体的な入力方法を知らない場合も想定される。したがって、上述の「目的地の設定が行えます」というガイドをした後に、音声入力がされることなくさらにクリック操作がされた場合には、「目的地を設定するときは都道府県名から入力して下さい」というように、入力方法の説明をガイド内容として報知することが考えられる。

【0052】

なお、これでも利用者がどのように入力するかを完全には理解できず、入力できずにいる場合も想定されるので、上述の「目的地を設定するときは都道府県名から入力して下さい」というガイドをした後に、音声入力がされることなくさらにクリック操作がされた場合には、「例えば愛知県刈谷市昭和町と入力して下さい」というように、具体的な入力例をガイド内容として報知することが考えられる。これによって利用者は具体的な入力方法が判り、また、具体例まで報知してもらえばそれに倣って自分が希望する目的地を容易に入力することができる。

20

【0053】

次に、中級者用のガイドについて説明する。

上述の初心者用ガイドでは、目的地設定に際して、1回クリック操作をする毎に、「目的地の設定が行えます」→「経由地の指定が行えます」→その他、というように段階的にガイドした。これに対して、中級者用ガイドとしては、それらをまとめて報知してしまうことも考えられる。また、目的地設定を行えるという基本的なガイドはせずに、その設定の際の指定可能条件（つまり経由地の指定や道路種類の指定など）のみをガイドしてもよい。

30

【0054】

また、上述した▲2▼の内容のような初心者用ガイドは、中級者用としては報知する必要はない。

このように、本実施形態の音声認識装置30を用いれば、何を音声入力してよいのかまで判っていない初心者にとっては、クリック操作をすれば、その場面場面で「初心者」として必要なガイドがなされるため、非常に使い勝手が良くなる。また、基本的なことは判っているが、細かい点については判っていない中級者にとっては、ダブルクリック操作をすれば、その場面場面で「中級者」に必要なガイドがなされるため、やはり非常に使い勝手が良くなる。また、初心者用のガイドのような逐次的なものでもないため、まどろっこしいこともない。

40

【0055】

一方、クリックあるいはダブルクリック操作というガイドを要求する操作をしない限りガイド内容が勝手に報知されてしまうことはないため、ガイドが必要でない利用者にとっては、通常通り、P T Tスイッチ36を押している間に所定の音声入力をすれば、相対的に（つまりガイド内容を報知させている状態よりも）短時間で所望の音声入力を行うことができる。

【0056】

50

このように、本実施形態の音声認識装置 30 によれば、音声認識装置 30 の利用方法に対する熟知度合が異なる各利用者のいずれにとっても使い勝手が良くなる。その上、ガイド要求操作を受け付けるために特別な構成を必要とするのではなく、音声入力のための基本的な処理に必要な P T T スイッチ 36 及び制御部 38 をガイド要求操作にも利用できるため、構成の簡略化にも寄与する。

【0057】

以上、本発明はこのような実施形態に何等限定されるものではなく、本発明の主旨を逸脱しない範囲において種々なる形態で実施し得る。

(1) 例えば、上記実施形態では、目的地設定に関するガイド内容を取り上げて説明したが、他にも地図の拡大、縮小や向き指定（ノースアップ・ヘディングアップなど）などについてガイド内容、あるいはその他、カーナビゲーションシステムにて音声入力可能な項目であれば、同様にガイド内容を設定することができる。

10

【0058】

(2) 上記実施例では、P T T スイッチ 36 の長押しを音声入力の期間を指示する操作として採用したが、その音声入力期間の開始及び終了の指定操作にクリック操作又はダブルクリック操作を採用してもよい。その場合には、ガイド要求操作として、スイッチの操作開始から終了までが所定時間以上となる操作を採用する。そしてこの場合は、スイッチ操作が所定時間継続した時点からガイド内容を報知し始め、スイッチ操作が終了した時点でそのガイド内容の報知も終了するようにすればよい。そして、そのガイド内容の報知後に入力された利用者の音声認識すればよい。

20

【0059】

(3) また、上記実施形態では、ガイド対象の利用者レベルとして初心者及び中級者という 2 種類としたが、3 種類以上設定しても構わない。

(4) また、上記実施形態では、認識結果の報知やガイド内容の報知を音声と画像の 2 種類で行なうようにしたが、利用者に対して報知が可能であればどのような形態でもよく、仮に 1 つだけ採用するならば、音声による報知を行なうことが好ましい。これは、カーナビゲーションシステム 2 という車載機器として用いていることを考慮したものであり、音声で出力されればドライバーは視点を表示装置にずらしたりする必要がなく、安全運転のより一層の確保の点では有利であることなどの理由からである。もちろん上記実施形態のように、音声及び画像の両方で報知すれば、ドライバーは表示による確認と音声による確認との両方が可能となる。

30

【0060】

(5) また、音声認識装置 30 の適用先としては、上述したナビゲーションシステムには限定されない。例えば音声認識装置を空調システム用として用いる場合には、設定温度の調整、空調モード（冷房・暖房・ドライ）の選択、あるいは風向モードの選択を音声入力によって行うようにすることが考えられる。そして、この場合には、その設定項目（温度・空調モード・風向モードなど）自体をガイド内容として報知したり、あるいは、「設定温度を 25 度にする」と言えばよいのか「設定温度を 5 度下げる」というように言えばよいのか、などをさらにガイド内容として報知することが考えられる。空調モードや風向モードなどについても同様である。

40

さらに、ナビゲーションシステムや空調システムを車載機器として用いる場合には限定されず、例えば携帯型ナビゲーション装置や屋内用空調装置などのように車載機器以外に用いてもよい。但し、これまで説明したように車載機器用として用いる場合には利用者がドライバーであることが考えられ、その場合には運転自体が最重要であり、それ以外の車載機器に対する操作については、なるべく運転に支障がないことが好ましい。したがって、車載機器としてのカーナビゲーションシステム 2 や空調システムを前提とした音声認識装置 30 の場合には、より一層の利点がある。

【0061】

もちろん、このような視点で考えるならば、ナビゲーションシステムや空調システム以外の車載機器に対しても同様に利用することができ、例えばカーオーディオ機器などは有効

50

である。また、それ以外にも、いわゆるパワーウィンドウの開閉やミラー角度の調整などを音声によって指示するような構成を考えれば、そのような制御対象についても同様に適用でき、やはり有効である。

【0062】

また、車載機器用とした場合にはそれ特有の利点があることは述べたが、音声認識装置30の適用先としては、利用者による音声入力指示にしたがって所定の処理を実行するものであれば同様に考えられる。例えば、携帯用の情報端末装置、あるいは街頭やパーキングエリアなどに設定される情報端末装置などにも同様に適用できる。

【図面の簡単な説明】

【図1】 本発明の実施形態としてのカーナビゲーションシステムの概略構成を示すブロック図である。

10

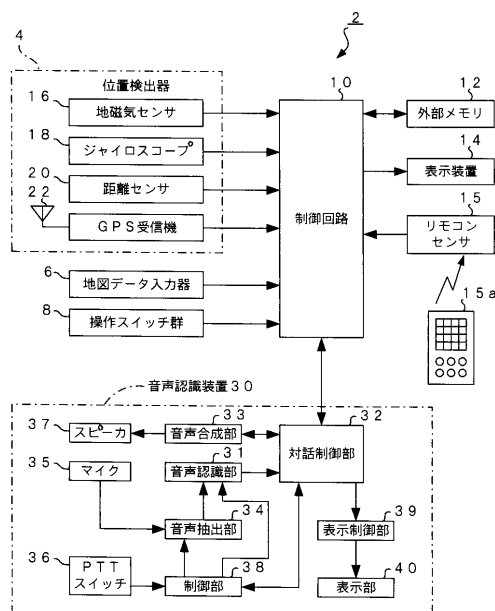
【図2】 音声認識装置が実行する処理を示すフローチャートである。

【符号の説明】

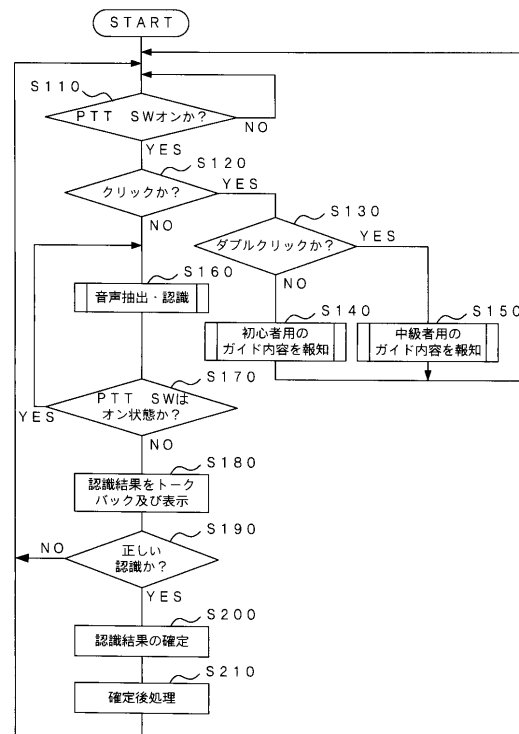
- | | |
|-----------------|------------|
| 2…カーナビゲーションシステム | 4…位置検出器 |
| 6…地図データ入力器 | 8…操作スイッチ群 |
| 10…制御回路 | 12…外部メモリ |
| 14…表示装置 | 15…リモコンセンサ |
| 15a…リモコン | 16…地磁気センサ |
| 18…ジャイロ스코プ | 20…距離センサ |
| 22…GPS受信機 | 30…音声認識装置 |
| 31…音声認識部 | 32…対話制御部 |
| 33…音声合成部 | 34…音声入力部 |
| 35…マイク | 36…PTTスイッチ |
| 37…スピーカ | 38…制御部 |
| 39…表示制御部 | 40…表示部 |

20

【図1】



【図2】



フロントページの続き

(72)発明者 北岡 教英
愛知県刈谷市昭和町1丁目1番地 株式会社デンソー内

合議体

審判長 原 光明

審判官 井上 健一

審判官 加藤 恵一

(56)参考文献 特開平9-113287 (JP, A)
特開平8-305391 (JP, A)
特開昭63-153594 (JP, A)

(58)調査した分野(Int.Cl., DB名)
G10L 13/00