

(19) 日本国特許庁(JP)

(12) 公開特許公報(A)

(11) 特許出願公開番号

特開2012-128440
(P2012-128440A)

(43) 公開日 平成24年7月5日(2012.7.5)

(51) Int.Cl.	F I	テーマコード (参考)
G10L 13/02 (2006.01)	G10L 13/02 130Z	
G10L 15/22 (2006.01)	G10L 15/22 300Z	
G06F 3/16 (2006.01)	G06F 3/16 340W	
	G06F 3/16 330H	

審査請求 有 請求項の数 3 O L (全 9 頁)

(21) 出願番号 特願2012-22981 (P2012-22981)
 (22) 出願日 平成24年2月6日 (2012.2.6)
 (62) 分割の表示 特願2006-196871 (P2006-196871)
 の分割
 原出願日 平成18年7月19日 (2006.7.19)

(71) 出願人 000004260
 株式会社デンソー
 愛知県刈谷市昭和町1丁目1番地
 (74) 代理人 110000567
 特許業務法人 サトー国際特許事務所
 (72) 発明者 浅田 博重
 愛知県刈谷市昭和町1丁目1番地 株式会
 社デンソー内

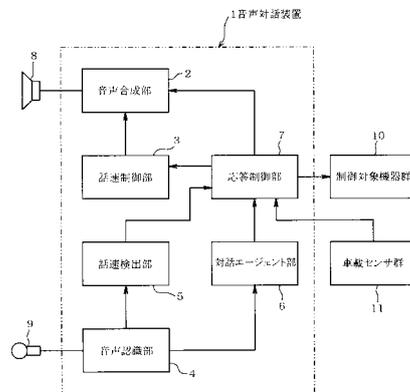
(54) 【発明の名称】 音声対話装置

(57) 【要約】

【課題】 音声対話装置において、話者の感性に即した円滑な対話を行うこと。

【解決手段】 応答制御部7は、音声認識部4が認識した話者の発話内容に応じた応答音声を、音声合成部2からスピーカ8を通じて出力する。この場合、通常状態では、音声認識部4が話者の発話の終了を検出した時点から応答音声の提供を開始するまでの応答時間を、話速検出部5により検出された発話速度に応じて変化させる。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

話者の発話内容を認識する音声認識手段、及びその認識結果に応じた応答音声を合成音声により提供する応答制御手段を備えた音声対話装置において、

前記話者による発話速度を検出する話速検出手段を備え、

前記応答制御手段は、前記音声認識手段が話者の発話の終了を検出した時点から前記応答音声の提供を開始するまでの応答時間を、前記話速検出手段により検出された発話速度に応じて変化させることを特徴とする音声対話装置。

【請求項 2】

前記応答制御手段は、環境条件検出センサからの検出信号を取り込むように構成され、その検出信号が予め決められた環境条件を満たすものであった場合に、その環境条件が解除されるまでの期間だけ前記応答音声の提供を見合わせた待機状態を呈することを特徴とする請求項 1 記載の音声対話装置。 10

【請求項 3】

前記応答制御手段は、制御対象機器の動作を制御可能に構成されると共に、環境条件検出センサからの検出信号を取り込むように構成され、前記音声認識手段による認識結果に前記制御対象機器に係る動作開始コマンド或いは動作停止コマンドが含まれていた場合に、当該コマンドの実行タイミングを、前記環境条件検出センサからの検出信号により示される環境条件に応じた最適なタイミングとなるように調整することを特徴とする請求項 1 または 2 に記載の音声対話装置。 20

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は、話者の発話内容の認識結果に応じた応答音声を合成音声により提供するようにした音声対話装置に関する。

【背景技術】**【0002】**

例えば、カーナビゲーションシステムやハンズフリー電話システムなどの車載システムにおいては、ユーザからの発話音声を、予め記憶された認識対象語彙の標準音声発生パターンと比較することにより、ユーザが発話した音声コマンドの認識を行い、その音声コマンドに応じた制御処理を行うものが普及している。この種の音声認識技術を利用する場合、ユーザからの 1 回の発話音声を解析するだけでは、その発話内容を一意的に特定できない場合が多く、誤って特定した場合には不要な制御処理が行われてしまう。そこで、このような事態への対策として、車載システム側でユーザによる発話内容を特定するための質問（応答音声）を出力して返答を求めることで対話的に発話内容を理解する音声対話装置が考えられている。 30

【0003】

一般的な音声対話装置は、話者（ユーザ）の発話内容に対する応答音声の合成速度や応答時間（応答開始までの時間）が画一的な構成となっている。ところが、話者の発話速度は、その個人毎の特性などに応じて異なってくるという事情がある。このため、話者側で、応答音声の出力が終了するまで不要に待たされると感じたり、質問に対する返答を急がされるといった感じを持ったりするなど、話者の感性に即した円滑な対話（話者が苛立つことが少ない対話）が困難になる状態が発生しやすいという問題点があった。このような問題点を解決するために、従来では、例えば特許文献 1 に見られるように、話者による発話速度を測定し、測定された発話速度に応じて応答音声の出力速度を変化させるようにした音声対話装置が考えられている。 40

【特許文献 1】特公平 7 - 2 1 7 5 9 号公報**【発明の開示】****【発明が解決しようとする課題】****【0004】**

10

20

30

40

50

上記従来の音声対話装置では、応答音声の出力速度を話者の発話速度に応じて単純に変化させているに過ぎないため、その出力が完了するまでの期間は、応答音声の速度が一定に保持されることになる。このため、応答音声の出力途中の段階で、話者側で応答音声は早期に終わって欲しいと感じるイベント（例えば、電話の着信）が発生した場合でも、その応答音声は終了するまで不要に待たされることがあり、これにより、話者側の苛立ちが増大するなど、話者の感性に即した円滑な対話が困難になる恐れがあった。また、音声対話装置では、会話における所謂「間」が、話者の感性に合った円滑な対話を行う上で非常に重要な要素となるものであるが、従来では、この「間」が話者の発話速度と無関係に一定であったため、話者の感性に即した円滑な対話が困難になるという状況下にあった。

【0005】

本発明は上記事情に鑑みてなされたものであり、その目的は、話者の感性に即した円滑な対話を行うことができる音声対話装置を提供することにある。

【課題を解決するための手段】

【0006】

請求項1記載の手段によれば、話者の発話内容が音声認識手段により認識された場合には、応答制御手段が、その認識結果に応じた応答音声を作成音声により提供されるようになる。この場合、話者による発話速度を検出する話速検出手段が設けられており、応答制御手段は、音声認識手段が話者の発話の終了を検出した時点から応答音声を提供するまでの応答時間、つまり、会話における所謂「間」を、前記話速検出手段により検出された発話速度に応じて変化させるようになる。このように、話者の感性に合った円滑な対話を行う上で非常に重要な要素となる「間」が、その話者の発話速度に応じた長さに制御される結果、話者の感性に即した円滑な対話を実現できるようになる。

【0007】

請求項2記載の手段によれば、応答制御手段は、環境条件検出センサから取り込んだ検出信号が予め決められた環境条件を満たすものであった場合に、その環境条件が解除されるまでの期間だけ応答音声の提供を見合わせた待機状態を呈するようになる。このため、例えば、話者が応答音声に反応しない方が良くと考えられる環境条件（音声対話装置が車載システムとして搭載されていた場合、車両の右左折などに伴うステアリング操作を実施中の状態や、急ブレーキによる減速操作を実施中の状態などが考えられる）の下では、応答音声の提供が見合わされることになって、対話の信頼性が向上するようになる。

【0008】

請求項3記載の手段によれば、応答制御手段は、音声認識手段による認識結果に、制御対象機器に係る動作開始コマンド或いは動作停止コマンドが含まれていた場合に、当該コマンドの実行タイミングを、環境条件検出センサからの検出信号により示される環境条件に応じた最適なタイミングとなるように調整するようになる。このため、話者からの動作開始コマンド或いは動作停止コマンドに基づいた制御対象機器の制御動作を実行しようとする際に、その実行を直ちに行わない方が良くと考えられる環境条件（制御対象機器が車両用のものであった場合、車両の右左折などに伴うステアリング操作を実施中の期間や、車両の後退走行期間などが考えられる）の下では、制御対象機器の制御動作の開始が見合わされることになるから、制御対象機器の制御に係る信頼性が向上するようになる。

【図面の簡単な説明】

【0009】

【図1】本発明の一実施例の基本構成を示す機能ブロック図

【図2】応答制御部による制御内容の一例を説明するためのシーケンス図

【発明を実施するための形態】

【0010】

以下、本発明の一実施例について図面を参照しながら説明する。

図1には、音声対話装置1及びこれに関連した部分の基本的な構成例が機能ブロックの組み合わせにより概略的に示されている。この図1において、音声対話装置1は、車両に搭載されて、話者（車両運転者）との間での音声応答処理及びその話者による音声コマン

10

20

30

40

50

ドに応じた制御処理を行うように構成されたもので、具体的に図示しないが、カーナビゲーション装置やエンジン制御 ECU、移動体通信用 ECU などとの間で例えば車内 LAN を介してデータの授受を行い得るように構成されている。

【0011】

音声対話装置 1 を構成する音声合成部 2、話速制御部 3（話速制御手段に相当）、音声認識部 4（音声認識手段に相当）、話速検出部 5（話速検出手段に相当）、対話エージェント部 6、応答制御部 7（応答制御手段に相当）は、実際には、音声対話装置用 ECU のプログラムにより実現されたものであり、それぞれには以下のような機能が設定されている。

【0012】

音声合成部 2 は、応答制御部 7 からの指示に基づいた音声信号を、周知の音声合成処理によってリアルタイムに生成し、その音声信号を応答音声として車載スピーカ 8 から出力する。尚、このスピーカ 8 は専用のもので良いが、カーオーディオ装置のスピーカを兼用できるものである。

【0013】

話速制御部 3 は、音声合成部 2 から出力される応答音声の発話速度を、応答制御部 7 からの指示速度に基づいた速度となるようにリアルタイム且つ連続的に変更する機能を備えたものである。この場合、話速制御部 3 は、その発話速度の変更アルゴリズムとして、例えば、応答音声の時間軸を圧縮伸長するという周知の TDHS（Time Domain Harmonic Scaling）法を採用しており、応答音声のピッチが一定のまま発話速度が変更されることになる。

【0014】

音声認識部 4 は、車載マイクロホン 9 から入力された話者の音声を取り込み、その音声（つまり、話者の発話内容）を、例えばキーワードスポッティングを使用した音声認識処理方式より解析することにより、対話に必要な語彙を認識する。

話速検出部 5 は、音声認識部 4 が認識した語彙とその発声時間に基づいて、話者による発話速度を検出（予測）する。

対話エージェント部 6 は、対話のマネジメントのために設けられたもので、音声認識部 4 が認識した語彙中に含まれるコマンドを選別して応答制御部 7 に与える構成となっている。

【0015】

応答制御部 7 は、上述したような音声合成部 2 及び話速制御部 3 の制御機能の他に、車両に搭載された移動体通信機器（携帯電話、自動車電話、データコミュニケーションモジュール（DCM）など）、カーナビゲーション装置、カーオーディオ装置などの制御対象機器群 10 の動作制御を実行する機能を備えたもので、車載センサ群 11（環境条件検出センサに相当）からの各種の検出信号が入力されるようになっている。この場合、上記車載センサ群 11 の構成要素としては、車両の走行状態や操作状態を検出するための車速センサ、加速度センサ、操舵角センサの他に、移動体通信機器の動作状態などを検出するためのアダプタなどが含まれるものであり、また、カーナビゲーション装置における経路案内情報、渋滞情報のような運転者支援情報などを出力するためのインタフェース部や、所定の車載機器の操作状況をモニタする手段なども含まれ、必要に応じて、車両運転者の視線方向を検出するための視線認識装置や車室内の騒音や会話をモニタするためのマイクロホンも含まれるものである。尚、車載センサ群 11 からの検出信号は、エンジン制御 ECU、移動体通信用 ECU などから車内 LAN を通じて入力する構成であっても良い。

【0016】

さて、以下においては、応答制御部 7 による制御内容のうち、本発明の要旨に関係した部分について関連した作用と共に説明する。

即ち、応答制御部 7 は、通常状態ではノンアクティブ状態を呈しているが、例えば話者（車両運転者）が操作可能な位置に設けられた対話開始スイッチがオンされたとき、或いは、音声応答装置 1 の動作開始を要求する特定音声コマンドが音声認識部 4 及び対話エー

10

20

30

40

50

ジェント部 6 を通じて与えられたとき（話者がマイクロホン 9 を通じて特定音声コマンドを入力したとき）に、アクティブ状態に切り換えられて話者との対話がスタートする構成となっている。

【 0 0 1 7 】

話者との対話がスタートしたときには、マイクロホン 9 を通じて入力された話者の発話内容が音声認識部 4 にて解析されて対話に必要な語彙が認識される共に、その語彙中に含まれるコマンドが対話エージェント部 6 により選別されて応答制御部 7 に与えられる。

【 0 0 1 8 】

上記のようなコマンドを受けた応答制御部 7 は、以下（ 1 ）、（ 2 ）のような制御を連続的に実行する。

（ 1 ）話速検出部 5 により検出される話者の発話速度に応じた長さのポーズ時間を決定すると共に、そのポーズ時間が経過したときに、音声合成部 2 に対して与えられたコマンドに応じた応答を行うための音声信号の作成を指示する制御。

【 0 0 1 9 】

（ 2 ）話速制御部 3 に対して、上記音声合成部 2 での音声信号の生成速度、つまり、スピーカ 8 を通じて出力される応答音声の発話速度を、話速検出部 5 により検出される話者の発話速度に応じた指示速度となるように決定し、その指示速度を話速制御部 3 に与える制御。

尚、本実施例では、上記ポーズ時間は、話者の発話速度が速い状態時ほど比例的に短くなるように制御され、また、上記指示速度は、後述するようなイベントが発生していない通常状態では、話者の発話速度が速い状態時ほど比例的に高速化するように制御される。

【 0 0 2 0 】

応答制御部 7 には、車載センサ群 1 1 からの検出信号をモニタすることにより、予め決められた所定のイベント、例えば、話者側で応答音声の早期に終わって欲しいと感じるイベント（電話の着信など）の発生の有無を監視する機能が設定されている。応答制御部 7 は、上記のような応答音声の出力途中の段階で、当該イベントが発生したときには、以下（ 3 ）のような制御を実行する。

【 0 0 2 1 】

（ 3 ）音声合成部 2 での音声信号の生成速度、つまり、応答音声の発話速度を、それまでの速度より連続的に高速化する制御（応答音声のピッチは変化しない）。

また、応答制御部 7 は、車載センサ群 1 1 から取り込んだ検出信号が予め決められた環境条件（例えば、車両の右左折などに伴うステアリング操作を実施中の状態、車両の急制動を実施中の状態）を満たすものであった場合には、以下（ 4 ）、（ 5 ）のような制御を実行する。

【 0 0 2 2 】

（ 4 ）上記のような環境条件が解除されるまでの期間だけ、音声合成部 2 による音声合成処理を禁止した状態、要するに、スピーカ 8 からの応答音声の提供を見合わせた待機状態を呈する制御。

（ 5 ）応答音声の出力途中の段階で上記環境条件が満たされた場合には、例えば、応答音声の出力を一旦停止し、その後に環境条件が解除されたときに、当該応答音声を最初から出力し直すという制御。

【 0 0 2 3 】

さらに、応答制御部 7 は、音声認識部 4 による認識結果に、制御対象機器群 1 0 に係る動作開始コマンド或いは動作停止コマンドが含まれていた場合に、以下（ 6 ）のような制御を行う。

（ 6 ）動作開始コマンド或いは動作停止コマンドの実行タイミングを、車載センサ群 1 1 から取り込んだ検出信号により示される環境条件（例えば、車両の右左折などに伴うステアリング操作を実施中の状態、車両の急制動を実施中の状態）に応じた最適なタイミングとなるように調整する制御（具体的には、例えば、上記動作開始コマンドが、移動体通信機器によるハンズフリー電話に対する発信コマンドであった場合には、上記環境条件が

10

20

30

40

50

満たされている状態では当該発信コマンドの実行を見合わせ、その後に当該環境条件が解除されたときに始めて発信コマンドを実行することになる。

【 0 0 2 4 】

図 2 のシーケンス図には、上記 (1)、(2)、(4) ~ (6) に係る制御内容についての具体例が示されており、以下これについて説明する。即ち、この図 2 の例は、話者である車両運転者が「 A 」さん宛てにハンズフリー電話をかけるというシチュエーションに対応したものであり、車両の運転状況が、直進 右折 (或いは左折) 直進 急制動 直進再開というように変化した場合における、車載センサ群 1 1、話者、音声認識部 4、音声合成部 2、話速検出部 5、話速制御部 3、応答制御部 7、対話エージェント部 6、制御対象機器群 1 0 の動作内容の推移が概略的に示されている。

10

【 0 0 2 5 】

S 1 : 車両の直進状態で、話者が「電話したい」と発話する。

S 2 : 音声認識部 4 が「電話したい」という発話内容を認識し、話速検出部 5 が話者による発話速度を検出し、対話エージェント部 6 が音声認識部 4 による音声認識結果に応じたコマンド (電話発信を要求するコマンド) を選別して応答制御部 7 へ出力する。

【 0 0 2 6 】

S 3 : 応答制御部 7 が、ポーズ時間・指示速度決定及び音声出力処理を行う。この処理時には、話者の発話速度が速い状態時ほど比例的に短くなるポーズ時間と、話者の発話速度が速い状態時ほど比例的に高速化する指示速度とを決定し、そのポーズ時間経過後に上記指示速度を話速制御部 3 に与えると共に、音声合成部 2 に対し、所定の音声信号 (例えば、話者の発話を受け付けたことを示すと共に、電話の発信先に質問するための「はい、どちらですか」の音声信号の生成を指示する制御が行われる。

20

【 0 0 2 7 】

S 4 : 音声合成部 2 が、指示された音声信号 (「はい、どちらですか」) を生成し、その音声信号を応答音声としてスピーカ 8 から出力すると共に、話速制御部 3 が、応答音声の発話速度を応答制御部 7 からの指示速度となるように制御する。

S 5 : 話者が「 A さんの携帯」と発話する (車両は直進状態のまま) 。

S 6 : 音声認識部 4 が「 A さんの携帯」という発話内容を認識し、話速検出部 5 が話者による発話速度を検出し、対話エージェント部 6 が音声認識部 4 による音声認識結果に応じたコマンド (電話発信左記を指示するコマンド) を選別して応答制御部 7 へ出力する。

30

【 0 0 2 8 】

S 7 : 応答制御部 7 が、ポーズ時間・指示速度決定及び音声出力処理を行う。この処理時には、ポーズ時間及び指示速度を決定し、そのポーズ時間経過後に上記指示速度を話速制御部 3 に与えると共に、音声合成部 2 に対し、所定の音声信号 (例えば、電話の発信先に報知するための「 A さんの携帯に電話します」の音声信号の生成を指示する制御が行われる。

【 0 0 2 9 】

S 8 : 話者が車両を右折 (或いは左折) させるためのステアリング操作を行う。

S 9 : 車載センサ群 1 1 (特には操舵角センサ) が右折 (或いは左折) のためのステアリング操作を検出する。尚、車載センサ群 1 1 に車両運転者の視線方向を検出するための視線認識装置が含まれていた場合には、その視線認識装置による検出出力を車両の右折或いは左折操作の判定に利用しても良い。

40

【 0 0 3 0 】

S 1 0 : 応答制御部 7 が、音声合成部 2 による音声合成 (音声信号の生成) をストップさせる。

S 1 1 : 話者が車両を直進状態に戻すためのステアリング操作を行う。

S 1 2 : 車載センサ群 1 1 (特には操舵角センサ) が車両を直進状態へ戻すためのステアリング操作を検出する。

S 1 3 : 応答制御部 7 が、音声合成部 2 による音声合成 (音声信号の生成) をスタートさせる。尚、車両の右折 (或いは左折) が、前記ポーズ時間が経過する前に行われたとき

50

には、音声合成が最初から行われることは当然であるが、車両の右折（或いは左折）がポーズ時間経過したタイミングであって応答音声の出力中に行われた場合にも、音声合成を最初からスタートさせる構成となっている。

【0031】

S14：音声合成部2が、指示された音声信号（「Aさんの携帯に電話します」）を生成し、その音声信号を応答音声としてスピーカ8から出力すると共に、話速制御部3が、応答音声の発話速度を応答制御部7からの指示速度となるように制御する。

S15：話者が車両に急ブレーキをかける操作を所定期間だけ行う。

S16：車載センサ群11が急ブレーキ操作を検出する。尚、この検出は、車載センサ群11の車速センサ或いは加速度センサの出力に基づいて行うことができる。

10

【0032】

S17：車載センサ群11（特には、車速センサ）が、急ブレーキの解除に伴う直進状態の再開を検出する。

S18：応答制御部7が、制御対象機器群10（特には、移動体通信機器）に対して、Aさんへの電話発信コマンドを送出する。

【0033】

S19：制御対象機器群10（特には、移動体通信機器）が、Aさんへの電話発信を行う。尚、この電話発信に必要な電話番号は、例えば音声対話装置1内或いは移動体通信機器側に予め構築された電話番号データベースから取得する構成になっている。また、電話発信時には、その発信音が例えば移動体通信機器から報知されるものであるが、例えば、音声合成部2にて電話発信音を生成し、その発信音をスピーカ8から出力する構成としても良い。

20

【0034】

要するに、上記した実施例の構成によれば、話者の発話内容が音声認識部4により認識された場合には、応答制御部7が、その認識結果に応じた応答音声を、音声合成部2による合成音声により提供するようになる。この場合、話者による発話速度を検出する話速検出部5及び上記応答音声の発話速度を調整するための話速制御部3が設けられており、通常状態では、当該応答音声の発話速度が、そのピッチを変化させることなく、話者の発話速度が速い状態時ほど比例的に高速化された速度となるように制御される。これにより、話者の感性に即した円滑な対話が可能になる。

30

【0035】

また、応答音声は、話者の発話内容が音声認識部4により認識されたときに直ちに出力されるのではなく、話者の発話速度が速い状態時ほど比例的に短くなるように制御されるポーズ時間が経過したときに始めて出力される構成、換言すれば、音声認識部4が話者の発話内容を認識した時点から応答音声を提供するまでの応答時間、つまり、会話における所謂「間」を、話者の実際の発話速度に応じて変化させる構成となっている。このように、話者の感性に合った円滑な対話を行う上で非常に重要な要素となる「間」が、その話者の発話速度に応じた長さに制御される結果、話者の感性に即した円滑な対話を実現できるようになる。

【0036】

40

一方、応答音声の出力期間中に、話者側で応答音声が早期に終わって欲しいと感じるイベント（電話の着信など）が発生したときには、応答音声の発話速度が、そのピッチを変化させることなく、それまでの速度より連続的に高速化するように制御される。このため、話者側で応答音声が早期に終わって欲しいと感じるような範疇のイベントが発生した場合に、その応答音声を終了するまで不要に待たされる事態を防止できるようになり、以て話者側の苛立ちが増大する恐れがなくなるなど、話者の感性に即した円滑な対話が可能になるものである。また、この場合には、イベントの発生に応じて応答音声の発話速度が変化することになるから、話者側では、イベントが発生したことを応答音声の発話速度の変化に基づいて間接的に認知可能になるという利点もある。

【0037】

50

尚、例えば、話者側で、応答音声の内容を確実に認識したいと感ずるような範疇のイベントが発生した場合には、応答音声、ピッチを変化させることなく連続的に低速化する制御を行う構成とすれば良く、このような構成とした場合には、その応答音声内容を話者側で確実に認識できるようになり、以て話者の要求に応じた円滑な対話が可能になるものである。

【0038】

応答制御部7は、車載センサ群11から取り込んだ検出信号が予め決められた環境条件を満たすものであった場合に、その環境条件が解除されるまでの期間だけ応答音声の提供を見合わせた待機状態を呈するようになる。このため、例えば、話者が応答音声に反応しない方が良くと考えられる環境条件（実施例中で説明したような車両の右左折などに伴うステアリング操作を実施中の状態、或いは、車両の急制動を実施中の状態など）の下では、応答音声の提供が見合わされることになって、対話の信頼性が向上するようになる。

10

【0039】

また、応答制御部7は、音声認識部4による認識結果に、制御対象機器群10の動作に係るコマンドが含まれていた場合に、当該コマンドの実行タイミングを、車載センサ群11からの検出信号により示される環境条件に応じた最適なタイミングとなるように調整するようになる。このため、話者からのコマンドに基づいた制御対象機器群10の制御動作を実行しようとする際に、その実行を直ちに行わない方が良くと考えられる環境条件（実施例中で説明したような車両の急制動を実施中の期間や、車両の後退走行期間などが考えられる）の下では、制御対象機器群10の制御動作の開始が見合わされることになり、その制御対象機器群10の制御に係る信頼性が向上するようになる。

20

【0040】

（その他の実施の形態）

本発明は上記した実施例に限定されるものではなく、例えば以下のような変形または拡張が可能である。

音声対話装置1を車両に搭載した場合の実施例について説明したが、所謂デジタル家電の動作を制御するための音声対話装置などに適用しても良い。

話速制御部3は、応答音声の発話速度をTDHS法により変更する構成としたが、他の方式の音声時間軸圧縮伸長アルゴリズムを採用しても良く、また、音声信号の生成スピードを変えることで応答音声の発話速度を変える手法を採用しても良い。

30

【0041】

応答制御部7は、話者の発話に応答するための応答音声の発話速度を、その応答直前の話者の発話速度に応じた指示速度となるように決定する構成となっているが、対話が始まった後における話者側の一連の発話速度を順次平均し、その平均発話速度に応じた指示速度を決定する構成としても良い。この構成によれば、話者の発話速度が一時的要因により変化した場合であっても、最終的には、応答音声の発話速度が当該話者の感性にあった状態に収束するようになるから、円滑な対話を行う上で有益になる。

【0042】

音声対話装置1を複数の話者が利用する場合には、話速検出部5により検出される話者の発話速度に応じた指示速度のデフォルト値を、実際に音声対話装置1を利用する話者毎に変更する構成としても良い。このような構成とする場合には、各話者の発話速度の平均値を別途に記憶しておき、その平均値に応じて上記指示速度のデフォルト値を変更することになる。

40

【0043】

話者の発話に応答した応答音声の出力が終了する前に、その話者が次の発話を行った場合には、その発話に対する応答音声に係るポーズ時間や発話速度を早くする制御を行う構成としても良い。この構成によれば、例えば、話者側に対話を急ぎたい事由がある場合や、所謂せっかちな話者に対するケアが可能になるものである。

【0044】

請求項1、請求項2に関しては、音声対話装置のみならず、ETCシステムやカーナビ

50

