(19)日本国特許庁(JP)

(12)特許公報(B2)

(11)特許番号
特許第7556202号
(P7556202)

(45)発行日　令和6年9月26日(2024.9.26)

(24)登録日　令和6年9月17日(2024.9.17)

(51)国際特許分類			FI		
G10L	*15/18*	*(2013.01)*	G10L	15/18	400
G10L	*15/00*	*(2013.01)*	G10L	15/00	200Q
B60R	*16/02*	*(2006.01)*	B60R	16/02	655K
G06F	*3/16*	*(2006.01)*	G06F	3/16	630

請求項の数　10　（全16頁）

(21)出願番号　特願2020-49179(P2020-49179)
(22)出願日　令和2年3月19日(2020.3.19)
(65)公開番号　特開2021-148971(P2021-148971A)
(43)公開日　令和3年9月27日(2021.9.27)
審査請求日　令和5年1月11日(2023.1.11)

(73)特許権者　000003997
日産自動車株式会社
神奈川県横浜市神奈川区宝町２番地
(74)代理人　110000486
弁理士法人とこしえ特許事務所
(72)発明者　杉本 美香
神奈川県横浜市神奈川区宝町２番地 日産自動車株式会社内
審査官　中村 天真

最終頁に続く

(54)【発明の名称】　音声認識方法及び音声認識装置

(57)【特許請求の範囲】

【請求項１】

ユーザの発した音声による指示に基づいて、移動体に搭載される装置に対する制御要求を受け付ける音声認識装置を用いた音声認識方法であって、

前記音声認識装置は、

前記音声を取得し、

取得した前記音声を音声データに変換し、

前記移動体の状態情報を取得し、

前記移動体の前記状態情報に基づいて、前記音声が取得される集音環境における前記音声の誤認識の可能性を示す誤認識レベルを算出し、

算出された前記誤認識レベルが予め定められた基準レベル未満である場合は、前記集音環境が良好であると判定し、前記音声データに対応する前記制御要求を特定するための検出モードを、自然言語認識によって前記音声データを解析する第１検出モードに設定し、

算出された前記誤認識レベルが前記基準レベル以上である場合は、前記集音環境が良好でないと判定し、前記検出モードを、自然言語認識とは異なる手法の音声認識によって前記音声データを解析する第２検出モードに設定し、

設定された前記検出モードに基づいて、前記音声データを解析して、制御の対象である操作対象装置と前記操作対象装置に対する前記制御要求とを特定し、

各々の前記操作対象装置及び前記制御要求に対応して認識することができる前記音声データの範囲は、前記第２検出モードの方が、前記第１検出モードよりも狭い、音声認識方

法。
【請求項２】
　前記検出モードが前記第２検出モードに設定されている場合に、
　前記音声認識装置は、
　前記音声データが、予め定義された制御要求テキストに一致するか否かを判定し、
　前記音声データが前記制御要求テキストに一致すると判定した場合に、前記制御要求テキストに対応する前記操作対象装置及び前記制御要求を特定する、請求項１に記載の音声認識方法。
【請求項３】
　前記検出モードが前記第２検出モードに設定されている場合であって、前記音声による指示に基づく前記音声データが前記制御要求テキストに一致しない場合に、
　前記音声認識装置は、前記ユーザに対して、前記音声データが前記制御要求テキストに一致するように前記指示をやり直すことを案内する案内情報を、前記移動体に設けられた情報出力部に出力する、請求項２に記載の音声認識方法。
【請求項４】
　前記音声認識装置は、前記移動体に設けられた空調装置から吹き出される空気の風量を前記状態情報として取得する、請求項１～３のいずれか一項に記載の音声認識方法。
【請求項５】
　前記音声認識装置は、前記移動体に設けられた開口部の開閉状態と、前記開口部が開状態である場合における前記移動体の走行速度とを前記状態情報として取得する、請求項１～４のいずれか一項に記載の音声認識方法。
【請求項６】
　前記音声認識装置は、前記移動体に設けられた開口部の開閉状態と、前記開口部が開状態である場合における前記移動体の周囲の風速を前記状態情報として取得する、請求項１～５のいずれか一項に記載の音声認識方法。
【請求項７】
　前記開口部は、ドアウィンドウ又はルーフウィンドウである、請求項５又は６に記載の音声認識方法。
【請求項８】
　前記音声認識装置は、前記移動体の内部の騒音レベルを前記状態情報として取得する、請求項１～７のいずれか一項に記載の音声認識方法。
【請求項９】
　前記音声認識装置は、前記集音環境に応じて前記検出モードを変更した場合に、前記検出モードが変更されたことを示す情報を、前記移動体に設けられた情報提供装置に出力する、請求項１～８のいずれか一項に記載の音声認識方法。
【請求項１０】
　音声による指示に基づいて、移動体に搭載される装置に対する制御要求を受け付ける音声認識装置であって、
　前記音声を取得する音声取得部と、
　前記音声取得部により取得された前記音声を音声データに変換する音声データ変換部と、
　前記移動体の状態情報を取得する移動体情報取得部と、
　前記移動体情報取得部により取得された前記移動体の前記状態情報に基づいて、前記音声が取得される集音環境における前記音声の誤認識の可能性を示す誤認識レベルを算出し、算出された前記誤認識レベルが予め定められた基準レベル未満である場合は、前記集音環境が良好であると判定し、算出された前記誤認識レベルが前記基準レベル以上である場合は、前記集音環境が良好でないと判定する集音環境推定部と、
　前記集音環境が良好であると判定された場合、前記音声データに対応する前記制御要求を特定するための検出モードを、自然言語認識によって前記音声データを解析する第１検出モードに設定し、前記集音環境が良好でないと判定された場合、前記検出モードを、自然言語認識とは異なる手法の音声認識によって前記音声データを解析する第２検出モード

に設定する検出モード設定部と、

設定された前記検出モードに基づいて、前記音声データを解析して、制御の対象である操作対象装置と前記操作対象装置に対する前記制御要求とを特定する制御要求特定部とを備え、

各々の前記操作対象装置及び前記制御要求に対応して認識することができる前記音声データの範囲は、前記第２検出モードの方が、前記第１検出モードよりも狭い、音声認識装置。

【発明の詳細な説明】

【技術分野】

【０００１】

本発明は、音声認識方法及び音声認識装置に関するものである。

【背景技術】

【０００２】

音声認識装置は、ユーザが発した音声を音声認識し、音声認識結果に基づいて、各種の処理を行う。音声認識装置が音声を正しく認識するためには、集音環境が適正であることが必要である。そこで、特許文献１の情報処理装置は、受信された音声に基づいて集音の適正度を算出し、集音の適正度に応じて、集音部の位置や姿勢を調整し、音声認識の精度を向上させる。

【先行技術文献】

【特許文献】

【０００３】

【文献】国際公開第２０１７／１９５４１２号

【発明の概要】

【発明が解決しようとする課題】

【０００４】

しかしながら、車両等の移動体の内部で音声認識装置を用いて装置の制御が行われる場合、特許文献１の情報処理装置では、移動体の状態が集音環境に及ぼす影響が考慮されていない。そのため、移動体の状態によっては、集音環境にノイズが多くなり、情報処理装置が音声を適切に取得することができずに、ユーザの指示内容が誤認識されてしまうおそれがあった。

【０００５】

本発明が解決しようとする課題は、移動体の状態に応じて集音環境を推定し、指示内容の誤認識を防止することができる音声認識装置を提供することである。

【課題を解決するための手段】

【０００６】

本発明に係る音声認識方法及び音声認識装置は、移動体の状態情報に基づいて集音環境における音声の誤認識の可能性を示す誤認識レベルを算出し、誤認識レベルが基準レベル未満である場合は、検出モードを、自然言語認識によって音声データを解析する第１検出モードに設定し、誤認識レベルが基準レベル以上である場合は、検出モードを、自然言語認識とは異なる手法の音声認識によって音声データを解析する第２検出モードに設定し、設定された検出モードに基づき、制御の対象である操作対象装置と操作対象装置に対する制御要求とを特定することによって、上記課題を解決する。

【発明の効果】

【０００７】

本発明によれば、集音環境に応じて制御要求の検出モードを設定するので、操作対象装置及び制御要求を含む音声指示の内容を誤って特定することを防止することができるという効果を奏する。

【図面の簡単な説明】

【０００８】

【図１】本発明の実施の形態に係る音声認識装置を含む車両の構成を示すブロック図であ

る。
【図２】図１に示す音声認識装置を用いた音声認識方法の概要を示すフローチャートである。
【図３】図２に示す音声認識方法において、各々の車両状態に応じて加算される誤認識レベルの点数を示す表である。
【図４】図２に示す音声認識方法に用いられるルールベース認識に基づく第２検出モードと自然言語認識に基づく第１検出モードとの音声の認識範囲の違いを例示する表である。
【図５】図２に示す音声認識方法を用いてエアコンの電源をＯＮにする操作を行う場合において、誤認識レベルに対応して設定される検出モードと、各々の検出モードで認識することができる発話例との関係を示す表である。
【図６】図２に示す音声認識方法を用いてルーフウィンドウの開閉の操作をする場合において、誤認識レベルに対応して設定される検出モードと、各々の検出モードで認識することができる発話例との関係を示す表である。
【発明を実施するための形態】
【０００９】
以下、本発明の実施形態を図面に基づいて説明する。
図１に示すように、移動体である車両１は、音声入力によって、ディスプレイ１２，オーディオ機器１３，ナビゲーション機器１４及び通話装置１５のうちの何れか一つ以上又はこれらの全てを制御するＩＶＩシステム等のナビゲーションシステム１０を有している。ナビゲーションシステム１０は、車両１に乗車したユーザの発した音声による指示に基づいて、車両１に搭載される各種の装置に対する制御要求を受け付ける音声認識装置１００を有する。またさらに、ナビゲーションシステム１０は、音声認識装置１００との間で通信可能な機器制御部１１を有する。機器制御部１１は、無線又は有線で信号を出力することにより、ディスプレイ１２，オーディオ機器１３，ナビゲーション機器１４及び通話装置１５の各々を制御する。なお、ディスプレイ１２は、ユーザのタッチ操作によって、オーディオ機器１３，ナビゲーション機器１４，通話装置１５及びその他の車載装置への制御指示を入力することができる入力機能を備えるタッチパネル式のディスプレイである。また、ディスプレイ１２は、通話装置１５への着信情報、通話装置１５を介して会話をしている通話相手の情報、ナビゲーション機器１４による案内情報などを、文字情報やアイコンを含む映像として出力することができる。また、オーディオ機器１３は、通話装置１５に着信があったことを通知する音声、通話装置１５を介して会話をしている通話相手の発話音声、ナビゲーション機器による案内情報を示す音声などを出力することができる。
ディスプレイ１２及びオーディオ機器１３は、情報出力部を構成する。
【００１０】
また、車両１は、車載装置として、例えば、室内灯３、エアコン４、ドアウィンドウ５、ルーフウィンドウ６及びバックドア７を有している。ナビゲーションシステム１０の機器制御部１１は、無線又は有線で信号を出力することにより、室内灯３の電源のＯＮ／ＯＦＦ、エアコン４の電源のＯＮ／ＯＦＦ、ドアウィンドウ５の開閉、ルーフウィンドウ６の開閉及びバックドア７の開閉の各々を制御することができる。また、車両１は、車両１の走行速度を検出する車速センサ２を有している。車速センサ２は、例えば、車両１の車輪の回転速度（車輪速）から、車両１の走行速度を検出する。さらに、車両１は、車両１の周囲の風速を検出する風速センサ８を有している。
なお、ドアウィンドウ５、ルーフウィンドウ６及びバックドア７は、車両１の開口部を構成する。また、エアコン４は、空調装置を構成する。
【００１１】
ナビゲーションシステム１０の音声認識装置１００は、音声取得部１０１、音声データ変換部１０２、車両情報取得部１０３、集音環境推定部１０４、検出モード設定部１０５、制御要求特定部１０６及び通信部１０７を有する。車両情報取得部１０３，検出モード設定部１０５及び制御要求特定部１０６は、通信部１０７を介して、機器制御部１１とデータ又は信号のやり取りを行うことができる。また、車両情報取得部１０３は、車速セン

サ２及び風速センサ８に無線又は有線で接続されている。
　車両情報取得部１０３は、移動体情報取得部を構成する。
【００１２】
　音声取得部１０１は、車両１の室内の音声を取得するマイクである。音声データ変換部１０２は、音声取得部１０１が取得した音声を音声データに変換する。なお、音声データは、音声に基づくテキストデータである。
【００１３】
　車両情報取得部１０３は、車速センサ２，風速センサ８，音声取得部１０１及び通信部１０７を介して、車両１の状態情報を取得する。車両１の状態情報とは、車両１の室内の音環境に影響を与える要素に関する情報である。例えば、車両１の状態情報は、車両１の室内の騒音の大きさ、ドアウィンドウ５，ルーフウィンドウ６又はバックドア７の開閉情報、車両１の走行速度、車両１の周囲の風速、オーディオ機器１３の設定音量、エアコン４から吹き出される空気の風量等である。車両１の室内の騒音の大きさは、音声取得部１０１によって検出されるノイズ、すなわち、ユーザの発話による音声以外の音の大きさ（例えば計測値）であり、ｄＢ（デシベル）を単位として算出される。また、車両情報取得部１０３は、通信部１０７及び機器制御部１１を介して、オーディオ機器１３の設定音量、エアコン４から吹き出される空気の風量、ドアウィンドウ５，ルーフウィンドウ６又はバックドア７の開閉情報等を、車両１の状態情報として取得する。また、車両情報取得部１０３は、車両１の室内の騒音の大きさとオーディオ機器１３の設定音量とに基づいて、車両１の内部のうち、例えば、室内の騒音レベルを算出し、算出された騒音レベルを車両１の状態情報として取得することもできる。具体的には、車両１の室内の騒音の大きさに応じて割り振られた点数と、オーディオ機器１３の設定音量に応じて割り振られた点数とを合計することで、騒音レベルが算出される。騒音レベルは、車両１の車室内に設けられたマイクによる計測値を用いて算出してもよいし、車両１の車体に設けられたマイクや振動センサによる計測値を用いて算出してもよい。
【００１４】
　集音環境推定部１０４は、車両情報取得部１０３によって取得された車両１の状態情報に基づいて、車両１の室内の集音環境を推定する。車両１の室内の集音環境とは、音声認識装置１００の音声取得部１０１がユーザの発した音声が取得される環境である。車両１の室内の集音環境が良好である程、音声取得部１０１が取得する音声に含まれるノイズ（ユーザの発話による音声以外の音）が少なく、音声認識装置１００がユーザの音声による指示を正確に認識できる可能性が高くなる。一方、車両１の室内の集音環境が良好でない程、音声取得部１０１が取得する音声に含まれるノイズが多く、音声認識装置１００がユーザの音声による指示を、誤って特定する、つまり、誤認識してしまう可能性が高くなる。なお、ユーザの音声による指示の認識処理には、制御要求の内容の特定及び／又は制御要求の対象となる操作対象装置の特定処理を含む。
【００１５】
　検出モード設定部１０５は、音声データ変換部１０２が変換した音声データに対応する制御要求を特定するための検出モードを、集音環境推定部１０４が推定した集音環境に応じて、設定する。具体的には、検出モード設定部１０５は、検出モードを、集音環境に応じて、後述する自然言語認識に基づく第１検出モード、又は、後述するルールベース認識に基づく第２検出モードのいずれかに設定する。検出モード設定部１０５によって設定された検出モードの情報は、通信部１０７及び機器制御部１１を介してディスプレイ１２及びオーディオ機器１３に出力される。すなわち、現在の検出モードが、第１検出モード又は第２検出モードのいずれに設定されているかを示す情報が、ディスプレイ１２に映像やテキストを含む画像として表示され、又は、オーディオ機器１３により音声によって出力される。また、集音環境の変化によって、検出モード設定部１０５が検出モードを、第１検出モードから第２検出モードに、又は、第２検出モードから第１検出モードに切り替えた場合、検出モードが変更されたことを示す情報がディスプレイ１２又はオーディオ機器１３に出力される。

【００１６】
制御要求特定部１０６は、検出モード設定部１０５によって設定された検出モードに基づいて、音声データ変換部１０２が変換した音声データを解析し、制御の対象である操作対象装置と、操作対象装置に対する制御要求の内容とを特定する。なお、制御要求特定部１０６は、操作対象装置及び制御要求を特定する際に参照するための辞書データ及び文脈データを有している。そして、制御要求特定部１０６は、操作対象装置の情報及び操作対象装置に対する制御要求の内容を、通信部１０７を介して、機器制御部１１に送信する。機器制御部１１は、受信した操作対象装置の情報及び操作対象装置に対する制御要求の内容に基づいて、操作対象装置を制御する。
【００１７】
次に、音声認識装置１００による音声認識方法の手順について、図２に示すフローチャートを用いて説明する。
まず、ステップＳ１において、音声認識装置１００の音声取得部１０１は、ユーザが発した音声を取得する。具体的には、音声取得部１０１は、周波数によってユーザが発話したことを検出し、ユーザの発話が検出された場合に、ユーザが発した音声を含む音声情報を音声認識対象として取得する。このように、音声取得部１０１が、ユーザが発話している間にのみ音声情報を取得することにより、常に音声情報を取得している場合に比べて、音声認識装置１００の制御負荷を減らすことができる。また、音声による指示の開始を示す特定のキーワードが予め決められている場合は、音声取得部１０１は、ユーザが特定のキーワードを発話した直後から、音声認識対象としての音声情報を取得してもよい。
【００１８】
次に、ステップＳ２において、音声データ変換部１０２は、取得した音声を、テキストとして認識可能な音声データに変換する。また、ステップＳ３において、車両情報取得部１０３が、車両の状態情報を取得する。そして、ステップＳ４において、集音環境推定部１０４が車両の状態情報に基づいて、集音環境を推定し、集音環境が良好であるか否かを判定する。
なお、ステップＳ３の処理は、ステップＳ１又はステップＳ２の処理と並行して実行してもよい。ステップＳ３の処理は、本制御手順のスタート以降、継続的又は周期的に行ってもよい。また、車両情報取得部１０３は、本制御手順のスタート前から車両の状態情報を継続的又は周期的に取得しており、ステップＳ３の処理で、予め継続的又は周期的に取得していた状態情報の中から、音声取得後の車両の状態情報を抽出して取得してもよい。
【００１９】
ここで、車両の状態情報に基づいて集音環境が良好であるか否かを判定する基準について、図３を用いて説明する。
集音環境が良好であるか否かは、車両１の状態情報に応じた誤認識レベルによって判定される。誤認識レベルとは、音声認識装置１００がユーザの発した音声を正確に認識することができず、正しい制御要求を特定することができない可能性を示す指標である。図３に示す例では、誤認識レベルを定量的な指標値として表現する。すなわち、誤認識レベルの指標値が高いほど、車両１の室内にノイズが多く、音声データ変換部１０２が音声を正確に音声データに変換することができないため、ユーザの音声による指示の誤認識の可能性が高くなる。従って、誤認識レベルの指標値が高いほど、集音環境は良好ではない。図３に示す表では、車両１の状態情報に応じた誤認識レベル（指標値）の加算点数が規定されている。なお、図３は、誤認識レベルの算出手法の一例を示すものであり、加算点数や加算の事項は、車両特性、音声認識装置の特性等に応じて予め定義することができる。
【００２０】
図３に示すように、ルーフウィンドウ６又はドアウィンドウ５が開状態である場合の誤認識レベルの加算点数は１点である。また、エアコンの風量が予め定められた基準風量以上である場合の誤認識レベルの加算点数は２点である。また、ルーフウィンドウ６又はドアウィンドウ５が開状態であって、車両１の走行速度が２０ｋｍ／ｈ以上８０ｋｍ／ｈ未満である場合の誤認識レベルの加算点数は２点である。また、ルーフウィンドウ６又はド

アウィンドウ５が開状態であって、車両１の走行速度が８０ｋｍ／ｈ以上である場合の誤認識レベルの加算点数は３点である。さらに、車両１の室内の騒音レベルが予め定められた基準騒音レベル以上である場合の誤認識レベルの加算点数は３点である。

【００２１】

集音環境の誤認識レベル（指標値）は、車両１の状態情報に応じて、図３の表の右側に示す点数を加算することにより算出される。例えば、ルーフウィンドウ６が開状態であり、車速が２０ｋｍ／ｈ以上８０ｋｍ／ｈ未満である場合の誤認識レベルを示す点数は、１点と２点とを加算して３点と算出される。また、ルーフウィンドウ６が開状態であり、エアコン風量が基準風量以上であり、車速が８０ｋｍ／ｈ以上である場合の誤認識レベルを示す点数は、１点と２点と３点とを加算して６点と算出される。

ステップＳ４において、集音環境推定部１０４は、誤認識レベルの基準レベルを４点と定め、誤認識レベルが基準レベルである４点未満である場合に集音環境が良好であると判定し、誤認識レベルが４点以上である場合に集音環境が良好でないと判定する。なお、基準レベルの点数は、車両特性、音声認識装置の特性等に応じて予め定義することができる。また、基準レベルの点数は、実験又はシミュレーションに基づいて予め定義することもできる。

【００２２】

ステップＳ４において、集音環境が良好であると判定された場合、ステップＳ５において、検出モード設定部１０５は、検出モードを第１検出モードに設定する。そして、ステップＳ６において、制御要求特定部１０６は、ステップＳ２で変換された音声データから、自然言語認識によって、ユーザの指示の内容が読み取り可能か否かを判定する。指示の内容が読み取り可能か否かは、音声データ変換部１０２によって変換された音声データを、制御要求特定部１０６が、辞書データ及び文脈データを参照して自然言語認識に基づき解釈し、操作対象装置及び制御要求の内容を特定することができるか否かによって判定する。音声データから指示を読み取ることができない場合は、ステップＳ１２において、音声認識装置１００は、ユーザに聞き直しを行う。聞き直しは、ディスプレイ１２に表示されるか、又は、オーディオ機器１３が音声案内を行うことによって、ユーザに提示される。一方、音声データから指示内容が読み取り可能であった場合、ステップＳ７において、制御要求特定部１０６は、操作対象装置及び制御要求を特定する。すなわち、制御要求特定部１０６は、音声データ変換部１０２によって変換された音声データを、辞書データ及び文脈データを参照して自然言語認識に基づき解釈し、操作対象装置及び制御要求の内容を特定する。特定された制御要求は、ステップＳ１１において、操作対象装置を制御するために、機器制御部１１へ送信され、フローは終了する。

【００２３】

ステップＳ４において、集音環境が良好でないと判定された場合、ステップＳ８において、検出モード設定部１０５は、検出モードを第２検出モードに設定する。そして、ステップＳ９において、ステップＳ２で変換された音声データが、予め定義された制御要求テキストに一致するか否かを判定し、一致していない場合は、ステップＳ１２において、ユーザに聞き直しを行う。ユーザへの聞き直しとは、音声入力（指示の発話）のやり直しを提案又は要求する処理である。このとき、制御要求特定部１０６は、ユーザに対して、音声データが制御要求テキストに一致するように指示（音声入力／発話）をやり直すことを案内する案内情報を、通信部１０７及び機器制御部１１を介して、ディスプレイ１２又はオーディオ機器１３に出力する。

【００２４】

一方、音声データが制御要求テキストに一致している場合、ステップＳ１０において、制御要求特定部１０６は、ルールベース認識に基づく第２検出モードにより、制御要求テキストに対応する操作対象装置及び制御要求を特定する。ステップＳ１１において、特定された制御要求は、操作対象装置を制御するために、機器制御部１１へ送信され、フローは終了する。

なお、ルールベース認識における音声データが制御要求テキストに一致しているか否か

の判断は、ステップＳ１において入力された音声に基づく音声データ（テキストデータ）の全部と、制御要求テキストとを比較することによって実行する。音声データの全部とは、発話開始から発話終了までに入力された音声に基づく音声データである。つまり、制御要求特定部は、発話が開始されてから終了するまでの一連の音声に基づく音声データと制御要求テキストとが一致するか否かを判断する。

【００２５】

ここで、自然言語認識に基づく第１検出モードと、ルールベース認識に基づく第２検出モードとの違いについて、図４を用いて説明する。

図４の表に示す例では、制御対象装置がエアコンであり、制御要求の内容が「エアコンの電源をＯＮにしたい」というものである場合、この制御対象装置及び制御要求の内容に対応する制御要求テキストは、「エアコンかけて」であると予め定義されている。

ルールベース認識に基づく第２検出モードでは、ユーザの発した音声に基づく音声データが制御要求テキストに一致すると判定した場合に、操作対象装置及び制御要求を特定する（図２のステップＳ１０参照）。すなわち、図４に示す例では、検出モード設定部１０５が検出モードを第２検出モードに設定した場合には、ユーザが、定義された制御要求テキストに従って「エアコンかけて」と発話した時にのみ、機器制御部１１は、エアコン４の電源をＯＮにする。従って、ユーザが制御要求テキストに一致しない発話をした場合、例えば、「エアコンつけて」、「えーと、エアコンかけたい」又は「エアコンかけると寒くなっちゃうよね」と発話した場合は、機器制御部１１は、エアコン４の電源をＯＮにせず、ＯＦＦの状態を維持する。

【００２６】

自然言語認識に基づく第１検出モードでは、ユーザが、制御要求テキストである「エアコンかけて」を発話する場合以外にも、例えば、「エアコンつけて」又は「えーと、エアコンかけたい」と発話した場合には、機器制御部１１は、エアコン４の電源をＯＮにする。具体的には、ユーザが「エアコンつけて」と発話した場合、制御要求特定部１０６は、音声データを「エアコン」と「つけて」とに分節し、辞書データを参照し、「つけて」が「かけて」の類語であると判断して、「エアコンの電源をＯＮにしたい」という制御要求を特定する。なお、第１検出モードで用いる辞書データは、登録された一の単語と、その単語と意味内容が共通する他の単語とを、類語として対応づけて記憶する。また、ユーザが「えーと、エアコンかけたい」と発話した場合は、制御要求特定部１０６は、「えーと」が間投詞であり、意味を持たない単語であると判断するとともに、「えーと、エアコンかけたい」は、「エアコンかけて」と同じ要求を意味する音声データであると判断する。これによって、制御要求特定部１０６は、「えーと、エアコンかけたい」という音声データに基づいて「エアコンの電源をＯＮにしたい」という制御要求を特定する。

すなわち、各々の操作対象装置及び制御要求に対応して認識することができる音声データの範囲は、ルールベース認識に基づく第２検出モードの方が、自然言語認識に基づく第１検出モードよりも狭い。すなわち、ルールベース認識に基づく第２検出モードは、制御要求テキストが定義づけられている制御要求のみを特定するので、制御要求の内容を特定することができる音声データの範囲は、定義された制御要求の数に応じて制限される。また、ルールベース認識に基づく第２検出モードは、定義された制御要求テキストと一致する制御要求のみを特定するので、第２検出モードで特定することができる（第２検出モードで入力できる）音声データは制限される。すなわち、検出モードが第２検出モードに設定された場合の音声入力の自由度は、検出モードが第１検出モードに設定された場合よりも低い。従って、第２検出モードの方が第１検出モードよりも、操作対象装置及び制御要求を特定するための条件が厳しい。

言い換えると、各々の操作対象装置及び制御要求に対応して認識することができる音声データの範囲は、自然言語認識に基づく第１検出モードの方が、ルールベース認識に基づく第２検出モードよりも広い。すなわち、第１検出モードの方が第２検出モードよりも、操作対象装置及び制御要求を特定するための条件が緩和されている。

【００２７】

一方、ユーザが「エアコンかけると寒くなっちゃうよね」と発話した場合、制御要求特定部１０６は、単語の組み合わせ及び順序を辞書データ及び文脈データに沿って解析し、「エアコンの電源をＯＮにした場合は、車両の室温が適温を下回るため、エアコンの電源をＯＮにしたくない」というユーザによる音声入力の意味を推測する。従って、制御要求特定部１０６は、機器制御部１１に対して、エアコン４の電源をＯＮにする指示を送信しない。すなわち、ユーザが「エアコンかけると寒くなっちゃうよね」と発話した場合は、検出モードが第１検出モード又は第２検出モードのいずれに設定されていても、エアコン４の電源はＯＦＦの状態に維持され、ＯＮにはならない。

なお、音声認識装置１００の制御要求特定部１０６は、音声データとユーザの実際の指示内容との対応関係のパターンを学習し、辞書データ及び文脈データを定期的に更新することができる。

【００２８】

次に、集音環境の誤認識レベルと検出モードとの関係について、図５及び６を用いて説明する。

図５は、制御対象装置がエアコンであり、制御要求が「エアコンの電源をＯＮにしたい」である場合の、誤認識レベルに対応して設定される検出モードと、各々の検出モードに対応する発話例とを示す表である。

まず、図５の表の下段では、図３の表に基づいて算出される誤認識レベルが４以上であり、集音環境が良好ではないと判定される。従って、検出モード設定部１０５は、検出モードを、ルールベース認識に基づく第２検出モードに設定する（図２のステップＳ８参照）。この場合、「エアコンの電源をＯＮにしたい」という制御要求に対応する音声は、前述の通り、予め定められた制御要求テキストと一致する「エアコンかけて」のみである。

【００２９】

次に、図５の表の中段では、図３の表に基づいて算出される誤認識レベルが２～３であり、集音環境は良好であると判定される。従って、検出モード設定部１０５は、検出モードを、自然言語認識に基づく第１検出モードに設定する（図２のステップＳ５参照）。ただし、誤認識レベルが２～３である場合の第１検出モードでは、後述する誤認識レベルが０～１である場合の第１検出モードに比べて、制御要求に対応して認識することができる音声データの範囲が狭く、音声入力の自由度が小さい。図５の表の中段に示す第１検出モードの例では、「エアコンの電源をＯＮにしたい」という制御要求に対応する音声は、制御要求テキストに一致する「エアコンかけて」に加えて、「エアコンつけて」又は「Ａ／Ｃ（エーシー）かけて」という発話例も含む。

【００３０】

さらに、図５の表の上段では、図３の表に基づいて算出される誤認識レベルが０～１であり、集音環境はさらに良好であると判定される。従って、検出モード設定部１０５は、検出モードを、自然言語認識に基づく第１検出モードに設定する（図２のステップＳ５参照）。前述の通り、誤認識レベルが０～１である場合の第１検出モードは、誤認識レベルが２～３である場合の第１検出モードに比べて、制御要求に対応して認識することができる音声データの範囲が広く、音声入力の自由度が大きい。図５の表の上段に示す第１検出モードの例では、「エアコンの電源をＯＮにしたい」という制御要求に対応する音声は、「エアコンかけて」、「エアコンつけて」又は「Ａ／Ｃかけて」に加えて、「暑いなぁ」又は「もっと涼しくして」という発話例も含む。すなわち、制御要求特定部１０６は、ユーザが「暑いなぁ」又は「もっと涼しくして」と発話した場合に、「車両１の室内の温度が高いため、室温を低下させるために、エアコンの電源をＯＮにしたい」という制御要求を特定する。このように、音声入力の自由度が大きい第１検出モードが設定された場合、制御要求特定部１０６は、ユーザの発した音声のみならず、車両１の室内の温度や外気温等の情報も、制御要求を特定するための条件の一部として記憶し、音声データを分析することができる。

【００３１】

一方、図６は、制御対象装置がルーフウィンドウであり、制御要求が「ルーフウィンド

ウを開けたい」又は「ルーフウィンドウを閉じたい」である場合の、誤認識レベルに対応して設定される検出モードと、各々の検出モードに対応する発話例とを示す表である。

まず、図６の表の下段では、図３の表に基づいて算出される誤認識レベルが４以上であり、集音環境が良好ではないと判定される。従って、検出モード設定部１０５は、検出モードを、ルールベース認識に基づく第２検出モードに設定する（図２のステップＳ８参照）。この場合、「ルーフウィンドウを開けたい」という制御要求に対応する音声は、予め定められた制御要求テキストと一致する「サンルーフをあけて」のみである。

【００３２】

次に、図６の表の中段では、図３の表に基づいて算出される誤認識レベルが２～３であり、集音環境は良好であると判定される。従って、検出モード設定部１０５は、検出モードを、自然言語認識に基づく第１検出モードに設定する（図２のステップＳ５参照）。ただし、図５の例と同様に、誤認識レベルが２～３である場合の第１検出モードでは、誤認識レベルが０～１である場合の第１検出モードに比べて、制御要求に対応して認識することができる音声データの範囲が狭く、音声入力の自由度が小さい。図６の表の中段に示す第１検出モードの例では、「ルーフウィンドウを開けたい」という制御要求に対応する音声は、制御要求テキストに一致する「サンルーフをあけて」に加えて、「上の窓あけて」又は「屋根あけて」という発話例も含む。

なお、制御要求特定部１０６は、車両１の運転中にユーザが発話する「上の窓」又は「屋根」は、「ルーフウィンドウ」を意味しているということを、予め第１検出モードの辞書データに記憶させておいてもよいし、ユーザの発話とユーザの操作との対応関係の履歴を用いた学習機能によって、辞書データ及び文脈データに登録してもよい。

【００３３】

さらに、図６の表の上段では、図３の表に基づいて算出される誤認識レベルが０～１であり、集音環境はさらに良好であると判定される。従って、検出モード設定部１０５は、検出モードを、自然言語認識に基づく第１検出モードに設定する（図２のステップＳ５参照）。前述の通り、誤認識レベルが０～１である場合の第１検出モードは、誤認識レベルが２～３である場合の第１検出モードに比べて、制御要求に対応して認識することができる音声データの範囲が広く、音声入力の自由度が大きい。図６の表の上段に示す第１検出モードの例では、「ルーフウィンドウを開けたい」という制御要求に対応する音声は、「サンルーフをあけて」、「上の窓あけて」又は「屋根あけて」に加えて、「星がみたい」という発話例も含む。すなわち、制御要求特定部１０６は、ユーザが「星がみたい」と発話した場合に、「上空の星空を視認することが可能となるように、ルーフウィンドウを開けたい」という制御要求を特定する。

【００３４】

また、図６の表の上段に示す第１検出モードでは、「ルーフウィンドウを閉じたい」という制御要求に対応する発話例は、「雨がふってきた」である。すなわち、制御要求特定部１０６は、ユーザが「雨がふってきた」と発話した場合に、「自分や車内が雨で濡れてしまわないように、ルーフウィンドウを閉めたい」という制御要求を特定する。

このように、音声認識の自由度が大きい第１検出モードが設定された場合、制御要求特定部１０６は、ユーザの発した音声のみならず、時刻情報や気象情報等も制御要求を特定するための条件の一部として記憶し、音声データを分析することができる。

【００３５】

以上より、本実施の形態に係る音声認識装置１００及び音声認識装置１００を用いた音声認識方法は、集音環境に応じて、音声データに対応する制御要求を特定するための検出モードを設定し、設定された検出モードに基づいて操作対象装置と制御要求とを特定する。従って、ノイズが多い集音環境であっても、集音環境に適した検出モードが設定されるため、音声認識装置１００は、音声に基づく制御要求の特定に係る誤認識を防止することができる。

【００３６】

また、音声認識装置１００は、集音環境に応じて、検出モードを、自然言語認識によっ

て音声データを解析する第１検出モード、又は、自然言語認識とは異なる手法の音声認識によって音声データを解析する第２検出モードのいずれかに設定することができる。また、各々の操作対象装置及び制御要求に対応して認識することができる音声データの範囲は、第２検出モードの方が、第１検出モードよりも狭く、限定的である。従って、音声認識装置１００は、集音環境に応じて、音声による曖昧な指示（自由度の高い指示）を受け付ける自然言語認識に基づく第１検出モードと、自然言語認識よりも自由度が低いものの、確度の高い音声認識に基づく第２検出モードとを使い分けることができる。よって、音声認識装置１００は、音声認識に適した集音環境では、音声認識ができる範囲が広く、入力が許容される音声の自由度が高い第１検出モードを選択することができ、一方で、自然言語認識による音声認識が適していない集音環境では、検出モードとして、高い確度の音声認識が期待できる第２検出モードを選択することができる。これにより、音声認識装置１００は、集音環境に関わらず、操作対象装置及び制御要求の特定に関する誤認識を防止することができる。

【００３７】

また、検出モードが第２検出モードに設定されている場合に、音声認識装置１００は、音声データが、予め定義された制御要求テキストに一致するか否かを判定する。そして、音声認識装置１００は、音声データが制御要求テキストに一致すると判定した場合に、制御要求テキストに対応する操作対象装置及び制御要求を特定する。すなわち、第２検出モードは、ルールベース認識によって音声データを解析する検出モードである。従って、自然言語認識による音声認識が適していない、例えばノイズの大きい集音環境では、音声認識装置１００は、ルールベース認識によって音声データを解析する第２検出モードを選択することができる。よって、音声認識装置１００は、集音環境に応じて、音声認識の確度を向上させ、ユーザの音声による指示の誤認識を防止することができる。

【００３８】

また、検出モードが第２検出モードに設定されている場合であって、音声による指示に基づく音声データが制御要求テキストに一致しない場合に、音声認識装置１００は、ユーザに対して、音声による指示をやり直すように案内する。これにより、検出モードが第２検出モードに設定されている場合に、音声認識装置１００は、ユーザに対して、ルールベース認識が可能な態様で、音声による指示を出すように促すことができる。

【００３９】

さらに、音声認識装置１００は、車両１の状態情報に基づいて集音環境が良好か否かを判定する。そして、音声認識装置１００は、集音環境が良好であると判定した場合は、検出モードを第１検出モードに設定し、集音環境が良好でないと判定した場合は、検出モードを第２検出モードに設定する。これにより、音声認識装置１００は、集音環境が良好である場合、すなわち、ノイズが少ない集音環境の場合は、自然言語認識によって曖昧な音声指示も受け付けて、ユーザによる音声指示の意味を推測した上で、操作対象装置及び制御要求を特定することができる。一方、集音環境が良好でない場合、すなわち、ノイズが多い集音環境の場合にも、自然言語認識によって曖昧な音声指示を受け付けてしまうと、指示の内容の誤認識の可能性が高くなる。そのため、集音環境が良好でない場合に、検出モードとして第１検出モードを設定してしまうと、音声認識装置１００が、ユーザの意図に反した操作対象装置及び制御要求を特定してしまうおそれがある。従って、音声認識装置１００は、集音環境が良好でない場合、すなわち、ノイズが多い集音環境の場合には、ルールベース認識により、予め決められた制御要求テキストに一致する指示のみを受け付ける。これにより、音声認識装置１００は、集音環境が良好でない場合において、音声による指示の誤認識を防止することができる。

【００４０】

また、音声認識装置１００は、車両１の状態情報に基づいて、音声の誤認識の可能性を示す指標値である誤認識レベルを算出する。そして、音声認識装置１００は、誤認識レベルが予め定められた基準レベル未満である場合は、検出モードを第１検出モードに設定し、誤認識レベルが基準レベル以上である場合は、検出モードを第２検出モードに設定する

。これにより、音声認識装置１００は、音声の誤認識の可能性を誤認識レベルとして具体的に算出することができる。さらに、音声認識装置１００は、誤認識の可能性が基準より低い場合は、自然言語認識によって曖昧な音声指示を受け付け、誤認識の可能性が基準以上の場合は、ルールベース認識によって予め定義された制御要求テキストに沿った音声指示のみを受け付ける。これにより、音声認識装置１００は、音声の誤認識の可能性が高い集音環境での、音声による指示の内容の誤認識を防止することができる。

【００４１】

音声認識装置１００は、空調装置であるエアコン４から吹き出される空気の風量を状態情報として取得する。エアコン４の風量が多い程、音声取得部１０１が取得する音声にノイズが混じりやすくなるため、エアコン４から吹き出される空気の風量を状態情報として取得することによって、音声認識装置１００は、集音環境をより適切に推定することができる。

【００４２】

また、音声認識装置１００は、車両１に設けられた開口部であるドアウィンドウ５又はルーフウィンドウ６の開閉状態と、ドアウィンドウ５又はルーフウィンドウ６が開状態である場合における車両１の走行速度を状態情報として取得する。ドアウィンドウ５又はルーフウィンドウ６が開けられた状態で車両１が走行している場合、車両１の走行速度が速い程、風音や対向車の走行音等、外から車両１の室内に入ってくるノイズが大きくなるからである。そのため、ドアウィンドウ５又はルーフウィンドウ６の開閉状態と、車両１の走行速度とを状態情報として取得することにより、音声認識装置１００は、集音環境をより適切に推定することができる。

【００４３】

さらに、音声認識装置１００は、開口部であるドアウィンドウ５、ルーフウィンドウ６又はバックドア７の開閉状態と、ドアウィンドウ５、ルーフウィンドウ６又はバックドア７が開状態である場合における車両１の周囲の風速を状態情報として取得する。ドアウィンドウ５、ルーフウィンドウ６又はバックドア７が開状態にある場合には、車両１の周囲の風速が速い程、車両１の室内に吹き込む風量が多く、ノイズが大きくなるからである。なお、車両１が走行しておらず停止している場合であっても、ドアウィンドウ５、ルーフウィンドウ６又はバックドア７が開状態にあれば、車両１の周囲の風速は、車両１の室内の集音環境に影響を及ぼす。そのため、ドアウィンドウ５、ルーフウィンドウ６又はバックドア７の開閉状態と、車両１の周囲の風速とを状態情報として取得することにより、音声認識装置１００は、集音環境をより適切に推定することができる。

なお、車両１が走行中である時は、バックドア７は常に閉状態であるため、音声認識装置１００は、ドアウィンドウ５又はルーフウィンドウ６の開閉状態と、車両１の周囲の風速とを状態情報として取得する。

また、車両１が停止している場合には、音声認識装置１００は、バックドア７以外のドア（図示せず）の開閉状態も、車両１の状態情報として取得してもよい。

【００４４】

また、音声認識装置１００は、車両１の室内の騒音レベルを状態情報として取得する。車両１の室内の騒音レベルが高い程、音声取得部１０１が取得する音声にノイズが混じりやすくなるため、車両１の室内の騒音レベルを状態情報として取得することによって、音声認識装置１００は、集音環境をより適切に推定することができる。

【００４５】

また、音声認識装置１００は、集音環境に応じて検出モードを変更した場合に、検出モードが変更されたことを示す情報を、ディスプレイ１２又はオーディオ機器１３に出力する。これにより、検出モードが第１検出モードから第２検出モードに変更されたこと、又は、検出モードが第２検出モードから第１検出モードに変更されたことをユーザが確実に確認することができ、ユーザは、検出モードに合わせた音声指示を出しやすくなる。

【００４６】

また、音声認識装置１００は、音声認識の手法として自然言語認識又はルールベース認

識のいずれを用いるかの区別に関わらずに、集音環境に応じて、第１検出モード又は第２検出モードを設定してもよい。例えば、第１検出モード及び第２検出モードは、いずれも、自然言語認識に基づく検出モードであってもよい。この場合も、各々の操作対象装置及び制御要求に対応して認識することができる音声データの範囲は、第２検出モードの方が、第１検出モードよりも狭くなるように設定される。具体的には、音声認識装置１００は、図５又は図６の上段の例に示す検出モードを第１検出モードとして設定する一方で、図５又は図６の中段の例に示す検出モードを第２検出モードとして設定してもよい。これにより、音声認識装置１００は、常に、音声入力の自由度が高い自然言語認識によって音声データの分析を行うことができるとともに、集音環境に応じて、音声認識することができる音声データの範囲、すなわち音声入力の自由度を適宜変えることができる。そのため、音声認識装置１００は、音声入力の自由度を高めるとともに、音声による指示の誤認識を防止することができる。

【００４７】

なお、本実施の形態では、車両１に設けられた音声認識装置１００が、図２に示す音声認識方法のうち、ステップＳ１～Ｓ１２の全ての処理を実行しているが、これに限定されず、一部の処理を、音声認識装置１００と通信可能なサーバが行っていてもよい。

【００４８】

また、移動体は車両に限定されず、電車又は飛行機であってもよい。

【００４９】

上記の車両１は、本発明に係る移動体に相当する。上記のエアコン４は、本発明に係る空調装置に相当する。上記のドアウィンドウ５、ルーフウィンドウ６及びバックドア７は、本発明に係る開口部に相当する。上記のディスプレイ１２及びオーディオ機器１３は、本発明に係る情報出力部に相当する。上記の車両情報取得部１０３は、本発明の移動体情報取得部に相当する。

【符号の説明】

【００５０】

１…車両（移動体）
４…エアコン（空調装置）
５…ドアウィンドウ（開口部）
６…ルーフウィンドウ（開口部）
７…バックドア（開口部）
１２…ディスプレイ（情報出力部）
１３…オーディオ機器（情報出力部）
１００…音声認識装置
１０１…音声取得部
１０２…音声データ変換部
１０３…車両情報取得部（移動体情報取得部）
１０４…集音環境推定部
１０５…検出モード設定部
１０６…制御要求特定部

【図面】

【図１】

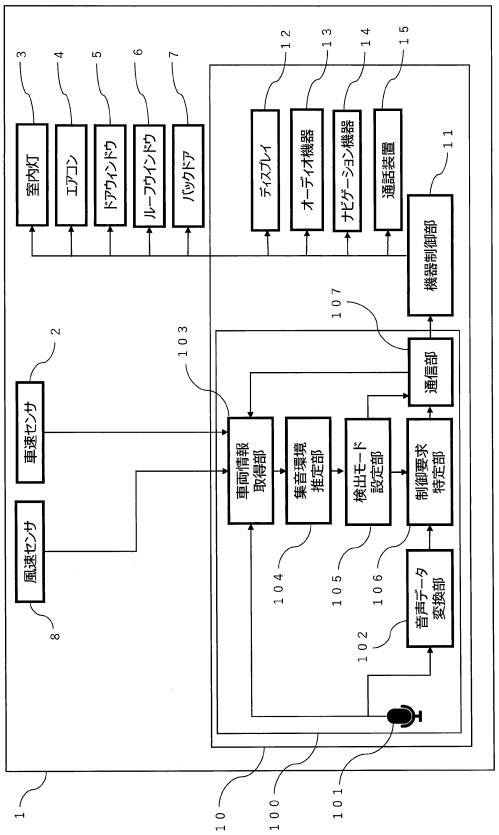
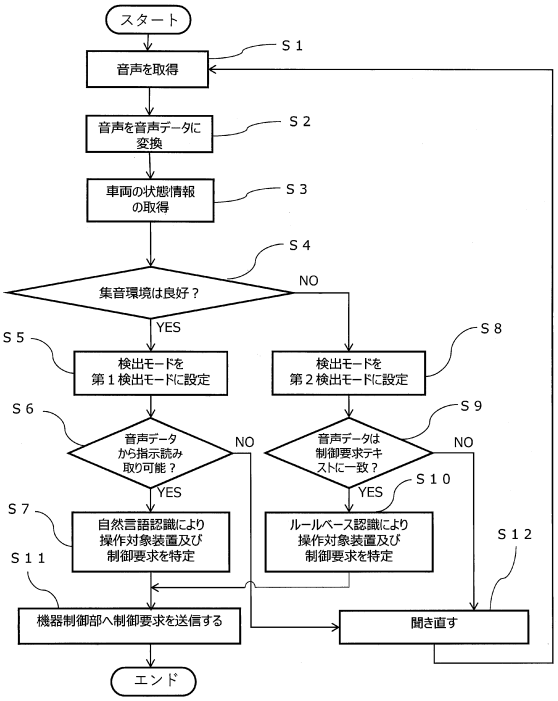


図１

【図２】

図２



【図３】

図３

車両の状態情報	誤認識レベル加算点数
ルーフウィンドウ開	1
ドアウィンドウ開	1
エアコン風量が基準風量以上	2
(ルーフウィンドウまたはドアウィンドウ開)車速20km/h以上	2
(ルーフウィンドウまたはドアウィンドウ開)車速80km/h以上	3
騒音レベルが基準騒音レベル以上	3

【図４】

図４

	第２検出モード（ルールベース認識）	第１検出モード（自然言語認識）
「エアコンかけて」（制御要求テキスト）	OK（エアコンON）	OK（エアコンON）
「エアコンつけて」	NG	OK（エアコンON）
「えーと、エアコンかけたい」	NG	OK（エアコンON）
「エアコンかけると寒くなっちゃうよね」	NG	NG

【図５】

図５

制御対象装置：エアコン
制御要求：エアコンの電源をＯＮにしたい

誤認識レベル	検出モード	発話例
0～1	**第１検出モード** 自然言語認識(自由度：大)	「暑いなあ」 「もっと涼しくして」
2～3	**第１検出モード** 自然言語認識(自由度：小)	「エアコンつけて」 「A/Cつけて」
4～	**第２検出モード** ルールベース認識	「エアコンかけて」 （※制御要求テキスト）

【図６】

図６

制御対象装置：ルーフウィンドウ
制御要求：ルーフウィンドウを開けたい or 閉じたい

誤認識レベル	検出モード	発話例
0～1	**第１検出モード** 自然言語認識(自由度：大)	「星がみたい」（開） 「雨がふってきた」（閉）
2～3	**第１検出モード** 自然言語認識(自由度：小)	「上の窓あけて」（開） 「屋根あけて」（開）
4～	**第２検出モード** ルールベース認識	「サンルーフをあけて」（開） （※制御要求テキスト）

フロントページの続き

(56)参考文献　特開２０１９－２１１７５２（ＪＰ，Ａ）
特開２００５－２０８１９８（ＪＰ，Ａ）
特開２００４－３５４７２２（ＪＰ，Ａ）
特開２００４－１８４８０３（ＪＰ，Ａ）
国際公開第２０１９／１４２４１８（ＷＯ，Ａ１）
特開２０１７－０９０６１２（ＪＰ，Ａ）
特開２００３－０１５６８９（ＪＰ，Ａ）
特開平０７－１４６６９８（ＪＰ，Ａ）

(58)調査した分野　(Int.Cl.，ＤＢ名)
Ｂ６０Ｒ　１６／０２
Ｇ０６Ｆ　３／１６
Ｇ１０Ｌ　１５／００－１５／３４