

(19) 日本国特許庁 (JP)

(12) 特 許 公 報 (B2)

(11) 特許番号

特許第4586566号  
(P4586566)

(45) 発行日 平成22年11月24日 (2010.11.24)

(24) 登録日 平成22年9月17日 (2010.9.17)

(51) Int.Cl.	F I
G 1 0 L 13/00 (2006.01)	G 1 0 L 13/00 1 0 0 M
G 1 0 L 15/22 (2006.01)	G 1 0 L 15/22 3 0 0 Z
G 1 0 L 21/04 (2006.01)	G 1 0 L 21/04 1 2 0 C
G 1 0 L 13/08 (2006.01)	G 1 0 L 13/08 1 3 2

請求項の数 4 (全 9 頁)

(21) 出願番号	特願2005-45861 (P2005-45861)	(73) 特許権者	000003207 トヨタ自動車株式会社
(22) 出願日	平成17年2月22日 (2005.2.22)		愛知県豊田市トヨタ町1番地
(65) 公開番号	特開2006-234953 (P2006-234953A)	(74) 代理人	100070150 弁理士 伊東 忠彦
(43) 公開日	平成18年9月7日 (2006.9.7)	(72) 発明者	金子 純也 愛知県豊田市トヨタ町1番地 トヨタ自動車株式会社内
審査請求日	平成19年9月7日 (2007.9.7)	審査官	清水 正一

最終頁に続く

(54) 【発明の名称】 音声対話システム

(57) 【特許請求の範囲】

【請求項 1】

音声による対話をユーザと行う音声対話システムにおいて、

無線通信手段を介して接続される通信相手と該音声対話システムとの間で行う対話で用いる第1の音声と車室内の乗員と該音声対話システムとの間で行う対話で用いる第2の音声と前記通信相手が発する第3の音声とを車室内に出力する音声出力手段を備え、

前記音声出力手段によって出力される第1乃至第3の音声の音色を互いに異ならせることを特徴とする音声対話システム。

【請求項 2】

前記音声出力手段によって出力されるそれぞれの音声の大きさを制御する音声制御手段を備える請求項1に記載の音声対話システム。

【請求項 3】

前記音声出力手段によって出力されるそれぞれの音声の高さを制御する音声制御手段を備える請求項1に記載の音声対話システム。

【請求項 4】

音声による対話をユーザと行う音声対話システムにおいて、

無線通信手段を介して接続される通信相手と該音声対話システムとの間で行う対話で用いる第1の音声と車室内の乗員と該音声対話システムとの間で行う対話で用いる第2の音声と前記通信相手が発する第3の音声とを車室内に出力する音声出力手段と、

前記通話相手と行う対話の成否を判断する判断手段とを備え、

10

20

前記音声出力手段によって出力される第 1 及び第 2 の音声の音色を互いに異ならせ、  
前記音声出力手段は、前記判断手段により対話が成立しないと判断された場合、第 1 及び第 3 の音声の少なくとも一つを出力することを特徴とする音声対話システム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声による対話をユーザと行う音声対話システムに関し、より詳細には、無線通信を介して接続される通信相手と行う対話を含むものに関する。

【背景技術】

【0002】

10

従来から、乗員に代わって運転中の通話を代行する車両用通信代行装置が知られている（例えば、特許文献 1）。本車両用通信代行装置は、音声対話機能を有しており、擬人化されたエージェントが通話相手と通話を行い、通話相手から取得した情報を乗員に伝えるものである。

【特許文献 1】特開 2003 - 32329 号公報

【発明の開示】

【発明が解決しようとする課題】

【0003】

上記の従来技術では、乗員が自分に話しかけていると誤解しないように、エージェントが通話相手と対話するときには、エージェントの音声 が車室内に出力されない（聞こえない）設定になっている。しかしながら、乗員はエージェントが通話相手に対して話す内容を聞くことができないという問題が生ずるため、乗員に不都合となる場合がある。

20

【0004】

そこで、本発明は、ユーザにとってより利便性の高い音声対話システムの提供を目的とする。

【課題を解決するための手段】

【0005】

上記課題を解決するため、本発明の一局面によれば、

音声による対話をユーザと行う音声対話システムにおいて、

無線通信手段を介して接続される通信相手と該音声対話システムとの間で行う対話で用いる第 1 の音声と車室内の乗員と該音声対話システムとの間で行う対話で用いる第 2 の音声とを車室内に出力する音声出力手段を備え、

30

前記音声出力手段によって出力される第 1 及び第 2 の音声の音色を互いに異ならせることを特徴とする音声対話システムが提供される。

【0006】

本局面によれば、乗員は、音声対話システムが通信相手に対して話す内容を聴取できるようになるとともに、音声対話システムが自分に対して話しかけているのか通信相手に対して話しかけているのかを音色の違いによって容易に識別できるようになり誤解もなくなる。

【0007】

40

また、前記音声出力手段は、前記通信相手が発する第 3 の音声を車室内に出力することが好ましい。これにより、乗員は通話相手の音声も聴取可能になり、音声対話システムと通話相手との間の対話内容が理解できるようになる。

【0008】

また、前記音声出力手段によって出力される第 1 乃至第 3 の音声の音色を互いに異ならせることが好ましい。これにより、通話相手が発する音声も音色の違いによって容易に識別可能である。特に、通話相手も本システムと同様の音声対話システムであれば、両システムが同じ音声であることも考えられるが、どちらの音声対話システムによる音声なのかを乗員は音色の違いによって容易に識別可能になる。

【0009】

50

また、音声出力の要否を指示する乗員からの入力に応じて、前記第 1 乃至第 3 の音声の中から出力すべき音声を選択する選択手段を備えることが好ましい。これにより、乗員は出力すべき音声を任意に選択でき、対話のやりとりの聞く・聞かないを選択することができる。

#### 【 0 0 1 0 】

また、前記通話相手と行う対話の成否を判断する判断手段を備え、前記音声出力手段は、前記判断手段により対話が成立しないと判断された場合、第 1 及び第 3 の音声の少なくとも一つを出力することが好ましい。これにより、音声対話システムが対話不能なやりとりでも、乗員が聞くことによって対話内容を把握することができるようになる。このとき、対話のやりとりを聞かないと選択していたとしても、同様に、乗員は対話内容を把握することができるようになる。

10

#### 【 0 0 1 1 】

また、前記音声出力手段によって出力されるそれぞれの音声の大きさを制御する音声制御手段を備えることが好ましい。これにより、音色を異ならせることによって音声の大きさに変化がでても、その大きさを調整することができるようになる。

#### 【 0 0 1 2 】

また、前記音声出力手段によって出力されるそれぞれの音声の高さを制御する音声制御手段を備えることが好ましい。これにより、音色を異ならせることによって音声の高さに変化がでても、その高さを調整することができるようになる。

#### 【 発明の効果 】

20

#### 【 0 0 1 3 】

本発明によれば、ユーザにとっての利便性をより高くすることができる。

#### 【 発明を実施するための最良の形態 】

#### 【 0 0 1 4 】

以下、本発明を実施するための最良の形態の説明を行う。

#### 【 0 0 1 5 】

##### [ 1 . 本実施形態の概要 ]

図 1 は、本実施形態の音声対話システムの概略動作を示すフローチャートである。本実施形態の音声対話システムは、運転者等の乗員の要求に基づいて乗員に代わって電話をかけ、通話相手と対話する中で乗員が要求する情報を取得し、その取得した情報を乗員に伝えるものである。図 1 における各ステップの詳細な説明については後述する。

30

#### 【 0 0 1 6 】

##### [ 2 . 本実施形態の詳細 ]

図 2 は、音声対話システムを取り巻く全体的な概略構成を示した図である。図 2 に示されるように、車両 1 0 0 には、音声対話システムとして、制御装置 1 0、入力装置 2 0、出力装置 3 0、記憶装置 4 0 及び通信装置 5 0 が搭載されている。また、通信装置 5 0 を介して、車両 1 0 0 から離れた場所に位置する施設 2 0 0 にアクセス可能なが示されている。

#### 【 0 0 1 7 】

制御装置 1 0 は、音声認識手段 1 1、音声合成手段 1 2、対話制御手段 1 3 及び音声制御手段 1 4 を備えている。

40

#### 【 0 0 1 8 】

音声認識手段 1 1 は、本音声対話システムと音声対話するユーザが話す音声を認識し、文字列に変換するものである。音声入力手段 2 2 から入力された音声について、特徴抽出、音素解析、単語解析及び構文解析を行うことによって、その入力された音声を文字列に変換する。

#### 【 0 0 1 9 】

音声合成手段 1 2 は、文字列から音声を生成するものである。音声合成手段 1 2 は、文字列解析、韻律制御及び音声波形生成を行うことによって、音声出力手段 3 1 によって出力される音声や携帯電話機 5 1 を介して送信される音声を文字列から生成する。

50

## 【 0 0 2 0 】

対話制御手段 1 3 は、本音声対話システムとそのユーザとの間の音声による対話の流れを制御するものである。対話制御手段 1 3 は、音声認識手段 1 1 によって入力音声から変換された文字列を認識し、その認識された文字列の内容に応じた制御を行う。また、対話制御手段 1 3 は、出力すべき音声がある場合には、音声合成手段 1 2 によって生成される音声の元になる文字列を生成する。

## 【 0 0 2 1 】

このような音声対話技術は、例えば、電話によるお客様窓口などに採用されている自動応答システム（IVR：Interactive Voice Response）、カーナビゲーションシステムのルート検索に用いる音声認識、通信機能を搭載した際の音声合成による電子メール読み上げ等がよく知られている。

10

## 【 0 0 2 2 】

音声制御手段 1 4 は、音声合成手段 1 2 によって生成された音声波形（音波）を制御するものである。音声波形を制御することによって、「音の大きさ（音圧、音圧レベル）」「音の高さ」「音色」に変化を与えることができる。「音の大きさ」は音声波形の振幅によって決まり、「音の高さ」は音声波形の周波数によって決まり、「音色」は音声波形の形状によって決まる。音の大きさと音の高さについてそれぞれ同一な二つの音声波形（つまり、振幅、周波数が同一）を比べた場合、その音声波形の形状が異なれば、その音声の音色は互いに異なる。

## 【 0 0 2 3 】

20

入力装置 2 0 は、スイッチ（S W）入力手段 2 1 及び音声入力手段 2 2 を備えている。S W 入力手段 2 1 には、例えば、プッシュ S W、レバー S W 及びタッチパネルディスプレイが挙げられる。制御装置 1 0 に対する乗員による操作入力を受け付けるインターフェースである。音声入力手段 2 2 は、車室内の音声を集音し、その集音された音声を音声認識手段 1 1 に供給する。また、音声入力手段 2 2 は、通信装置 5 0 を介して集音した通信相手の音声を音声認識手段 1 1 に供給する。

## 【 0 0 2 4 】

出力装置 3 0 は、音声出力手段 3 1 及び画像出力手段 3 2 を備えている。音声出力手段 3 1 は、上述の音声波形に基づいて車室内に実際の音声を出力する、いわゆるスピーカである。画像出力手段 3 2 は、乗員に対し視覚的な情報提供をする表示装置であって、より具体的には、T F T - L C D（薄膜トランジスタ方式液晶ディスプレイ）、自発光タイプの E L（Electro Luminescence）パネル、V F D（蛍光表示管）及びヘッドアップディスプレイ等がある。そして、画像出力手段 3 2 には、例えば、ナビゲーションシステム画面（メニュー画面や経路案内画面）、擬人化されたエージェントの像、音声対話のやりとりの中で認識され変換された文字列が表示される。

30

## 【 0 0 2 5 】

記憶装置 4 0 には、会話シナリオ 4 1、施設情報 4 2、選択施設情報 4 3 及び地図情報 4 4 が記憶されている。会話シナリオ 4 1 は、単語や構文の辞書データ及び文例であって、上述の音声対話の制御をする際に参照されるものである。施設情報 4 2 は、レストランや公園等の施設に関する位置情報や料金等の情報である。選定施設情報 4 3 は、後に詳述するが乗員が選定した施設に関する情報である。地図情報 4 4 は、G P S からの車両 1 0 0 の位置、地図データ、建物データ、道路データ等であって、ナビゲーションシステムが経路探索等をする際に参照されるものである。

40

## 【 0 0 2 6 】

通信装置 5 0 は、車外との無線通信を可能にする装置であって、より具体的には、携帯電話機 5 1 である。制御装置 1 0 は、携帯電話機 5 1 を介して車外の電話 6 0 を備える施設 2 0 0 に電話をかけ、電話の相手と対話することを可能にする。

## 【 0 0 2 7 】

## 〔 3 . 本実施形態の動作 〕

図を参照しながら、本実施形態の動作例について説明する。図 1 は、本実施形態の音声対

50

話システムの概略動作を示すフローチャートである。ここでは、擬人化されたエージェントによって音声対話システムの動作が表現されるものとする。エージェントは、乗員と対話を行ったり、乗員の好みや要求やそれらの学習結果に応じて最適な推奨案（例えば、乗員が好みそうなレストランが近くにあれば場所を案内してくれる）を提供してくれたりする。エージェントの容姿は、人間をはじめとして、動物、ロボット、漫画のキャラクター等、様々存在し、ユーザの好みによって選択可能なものである。エージェントは、ディスプレイ上を動くものであってもよいし、ホログラフィのようなものであってもよい。

【 0 0 2 8 】

以下、乗員が運転中に定食屋を探している状況において音声対話システム（エージェント）自らが定食屋に関する情報を取得するとともにその情報を乗員に提供するというケースに当てはめて、本実施形態の動作の一例を説明する。

10

【 0 0 2 9 】

図 1 において、まず、乗員は通話相手の選定を行う（ステップ 1 0 0）。ここでの通話相手は定食屋になる。図 3 を参照しながら、通信相手の選定について詳細に説明する。図 3 は、通信相手の選定についてのフローチャートである。乗員は、ナビゲーションシステムで施設一覧を画像出力手段 3 2 に表示させ、その施設一覧の中から定食屋を選択する（ステップ 1 1 0）。乗員は、情報取得ボタンを押して、メニューの中から取得したい情報、例えば、「営業時間」や「待ち時間」を選択する（ステップ 1 2 0）。ここで、優先度をつけて複数の定食屋を選択してもよい。優先度に応じて音声対話システムが電話をかける順番（情報を取得したい順番）が変わる。ただし、電話をかける順番は、記憶装置 4 0 に記憶されている施設情報 4 2（現在値からの距離や定食屋の料金設定等）に応じて自動的に設定されるようにしてもよい。そして、乗員が施設選択完了ボタンを押すことによって、通話相手の選定が完了する（ステップ 1 3 0）。

20

【 0 0 3 0 】

図 1 に戻り、エージェントは選択された定食屋に電話をかけ（ステップ 2 0 0）、相手が電話に出た場合に対話を開始する（ステップ 3 0 0）。図 4 は、エージェントと通話相手との間の音声対話についてのフローチャートである。エージェントは、相手が話した内容を認識し、電話に出た相手を特定する。特定できなかった場合、エージェントは通話相手に「もしもし、定食屋 さんでよろしいですね？」と尋ねることによって相手を特定する。

30

【 0 0 3 1 】

次に、エージェントは、「こちらはドライバーエージェントシステムです。ドライバーに代わり、教えてほしいことがあります。まず「営業時間」について教えて下さい」と話して、自己紹介するとともに取得したい情報について尋ねる（ステップ 3 1 0）。通話相手は、「営業時間は です」と答える。そして、エージェントは、自分の質問に対する通話相手の応答音声内容を受信し、その音声認識を行う（ステップ 3 2 0）。音声認識を行ったエージェントは、「営業時間は でよろしいですね？」と通話相手に確認する（ステップ 3 3 0）。OKであることを確認した場合、次の取得したい情報についての質問をする（ステップ 3 5 0；No）。

【 0 0 3 2 】

40

エージェントは、次の質問として「現在どれくらいの「待ち時間」ですか？」と尋ねる（ステップ 3 1 0）。尋ねられた通話相手は「待ち時間は です」と答える。その答えを音声認識したエージェントは、上記と同様に、「待ち時間は××でよろしいですね？」と通話相手に確認する（ステップ 3 2 0，3 3 0）。ここで、その確認内容が間違っていると通話相手から指摘された場合や通話相手の音声認識が不可能な場合、エージェントは、通話相手に対して少し待って欲しい旨を話す（ステップ 3 3 0；No）。

【 0 0 3 3 】

エージェントは、「待ち時間について認識できません。対話内容を聞きますか？」と、乗員に対して情報取得できないため対話内容を聞くか否かを確認する（ステップ 3 4 0）。乗員が「聞く」と答えると、その音声認識をしたエージェントは、車室内のスピーカを

50

通して、その対話内容を乗員が聞くことができるようにする。エージェントは、再び、「現在どれくらいの待ち時間ですか？」と尋ねると、通話相手は「待ち時間は です」と答える。今度は、車室内の乗員でも通話相手の音声を聞くことができるようになるので、乗員は相手の音声を直接聞くことによって待ち時間を知ることができるようになる。そのとき、エージェントももちろん通話相手の話しを聞いており、「待ち時間 」について音声認識をあらためて行っている。エージェントがそれでも音声認識できない場合は、「やっぱり、認識できませんでした。ごめんなさい。」と乗員に謝る。さらに、乗員の聞きたい情報が他にも出てきたならば、乗員が直接通話相手に話しかけ、その通話相手の応答する音声を聞くことができるようにしてもよい。要取得情報がすべて取得できたならば、エージェントは、「質問は以上です。ありがとうございました」と、通話相手に通話を終了する旨を伝える。

10

【 0 0 3 4 】

そして、エージェントは、「営業時間は で、待ち時間は です。次の候補も確認しますか？」と、取得した情報を乗員にディスプレイ表示とともに音声で伝え、次の候補の情報を取得するか否かを確認する（図 1 のステップ 4 0 0）。乗員は「もう、確認しなくてよい」と答えれば、その音声で認識され、情報取得が終了する。

【 0 0 3 5 】

このように、車室内のスピーカから流れる音声は、エージェントが通話相手に話しかける音声もあれば乗員に話しかける音声もあり、通話相手がエージェントに話しかける音声もあれば乗員に話しかける音声もあり、対話も入り組んでいる。したがって、乗員、特に運転中のドライバーは、誰が誰に話しかけているのかがわかりにくくなるおそれがある。エージェントや通話相手の話す言葉が質問形式であればなおさらである。そこで、本発明の音声対話システムでは、車室内のスピーカから出力される通話相手と行う対話で用いる音声と車室内の乗員と行う対話で用いる音声と通話相手が発する音声について、その音色を互いに異ならせるようにしている。

20

【 0 0 3 6 】

図 5 は、対話音声の音色制御を示すフローチャートである。音声合成手段 1 2 は、エージェントが話す文章の文字列から音声波形を生成する（ステップ 5 1 0）。生成された音声波形に基づいて、エージェントの音声で、車室内のスピーカを通して乗員に且つ携帯電話機 5 1 を介して通話相手に、出力される（ステップ 5 2 0）。音声制御手段 1 4 は、音声波形の振幅や振動数や形状を調整することによって、乗員に話しかける音声と通話相手に話しかける音声の音色を互いに異なるように変える（ステップ 5 3 0）。

30

【 0 0 3 7 】

また、携帯電話機 5 1 を介して通話相手の音声が入力された場合（ステップ 5 4 0）、車室内のスピーカを介して通話相手の音声出力されるとともに（ステップ 5 5 0）、音声認識手段 1 1 は通話相手の音声の音声認識を行う（ステップ 5 6 0）。音声制御手段 1 4 は、ステップ 5 3 0 で行われたエージェントの音声の音色と異なる音色になるように、通話相手の音声の音声波形の振幅や振動数や形状を調整する（ステップ 5 7 0）。

【 0 0 3 8 】

このように、車室内のスピーカから出力される通話相手と行う対話で用いる音声と車室内の乗員と行う対話で用いる音声の音色を互いに異ならせることによって、音声対話システムが自分（乗員）に対して話しかけているのか通信相手に対して話しかけているのかを容易に識別できるようになる。

40

【 0 0 3 9 】

以上、本発明の好ましい実施例について詳説したが、本発明は、上述した実施例に制限されることなく、本発明の範囲を逸脱することなく、上述した実施例に種々の変形及び置換を加えることができる。

【 0 0 4 0 】

例えば、音声対話システム側から通話を発信する場合について上述したが、受信する場合についても同様に音声対話することは可能であり、そのときの音声の音色も互いに異な

50

らせるようにすればよい。

【 0 0 4 1 】

また、エージェントが電話をかけたときに電話がつながらない場合には、一定時間後にあらためて電話をかけるようにすればよい。この場合、エージェント自らが一定時間後に自動的に電話をかけなおすようにしてもよいし、乗員にあらためて電話をかけるか否かの判断を問い合わせるようにしてもよい。乗員に問い合わせることによって、待ってでも情報取得するのか、待たずに次の施設の情報取得に移行させるのかを乗員は選択することができる。

【 0 0 4 2 】

また、通話相手の言語を推定する手段を備えてもよい。記憶装置 4 0 内の会話シナリオ 4 1 を言語別に記憶し、通話相手の言語に応じてエージェントが話す言語を切り替えるようにしてもよい。これにより、言語の不一致によって対話ができないと認識されれば、通話相手が理解できる言語に切り替えて対話することにより、情報取得が可能となる。なお、この場合の言語は、人間が話す日本語や英語等であるが、上記例示の IVR 等の通信相手が機械の場合にはその通信プロトコル等を相手に合わせて切り替えるようにしてもよい。

【図面の簡単な説明】

【 0 0 4 3 】

【図 1】本実施形態の音声対話システムの概略動作を示すフローチャートである。

【図 2】音声対話システムを取り巻く全体的な概略構成を示した図である。

【図 3】通信相手の選定についてのフローチャートである。

【図 4】エージェントと通話相手との間の音声対話についてのフローチャートである。

【図 5】対話音声の音色制御を示すフローチャートである。

【符号の説明】

【 0 0 4 4 】

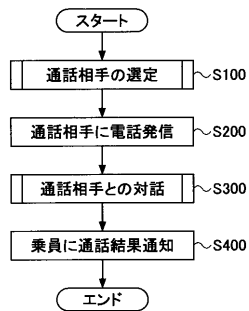
- 1 1 音声認識手段
- 1 2 音声合成手段
- 1 3 対話制御手段
- 1 4 音声制御手段
- 1 0 0 車両
- 2 0 0 施設

10

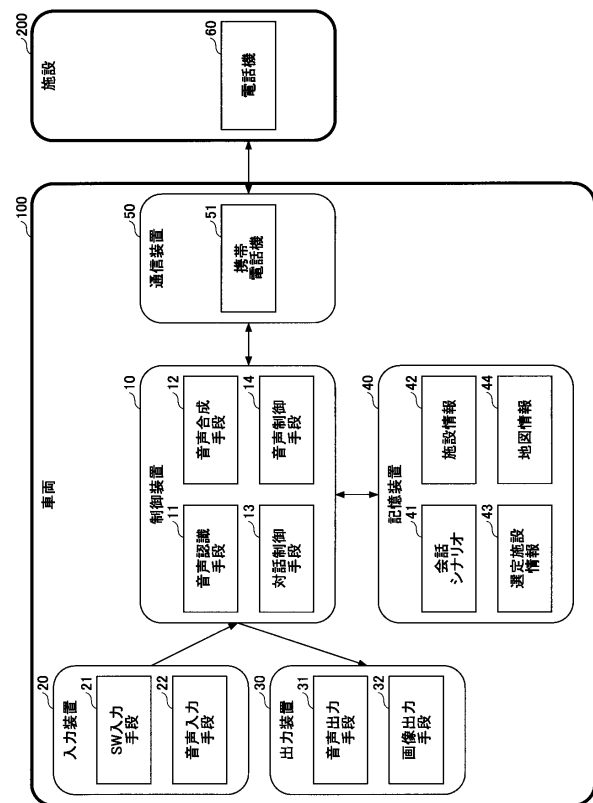
20

30

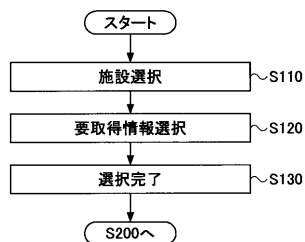
【図 1】



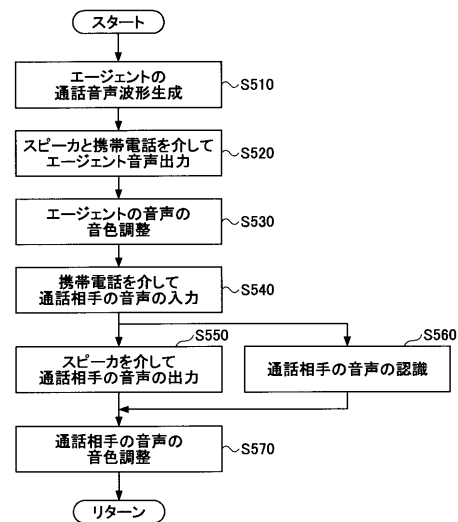
【図 2】



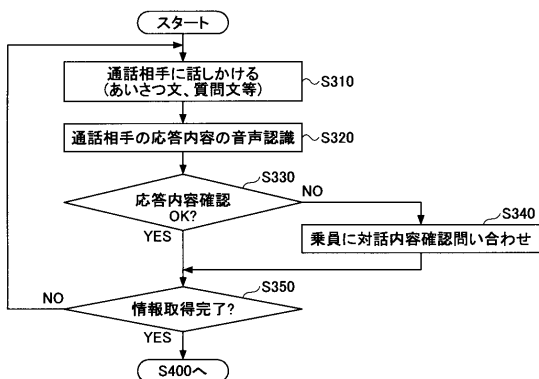
【図 3】



【図 5】



【図 4】





---

フロントページの続き

(56)参考文献 特開2003-281652(JP,A)  
特開2005-012833(JP,A)  
特開平06-209475(JP,A)

(58)調査した分野(Int.Cl., DB名)

G10L 13/00  
G10L 15/00、15/22  
G10L 21/04