

(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2007-11380

(P2007-11380A)

(43) 公開日 平成19年1月18日(2007.1.18)

(51) Int. Cl.	F I	テーマコード (参考)
G10L 15/06 (2006.01)	G10L 15/06 310S	2F129
G10L 15/00 (2006.01)	G10L 15/00 200Q	3D020
G01C 21/00 (2006.01)	G01C 21/00 A	5D015
B60R 11/02 (2006.01)	B60R 11/02 M	
B60R 16/02 (2006.01)	B60R 16/02 655A	

審査請求 未請求 請求項の数 46 OL (全 21 頁)

(21) 出願番号 特願2006-183890 (P2006-183890)
 (22) 出願日 平成18年7月3日(2006.7.3)
 (31) 優先権主張番号 11/173,736
 (32) 優先日 平成17年7月1日(2005.7.1)
 (33) 優先権主張国 米国 (US)

(71) 出願人 591009509
 ボーズ・コーポレーション
 BOSE CORPORATION
 アメリカ合衆国マサチューセッツ州017
 01, フラミンガム, ザ・マウンテン (番地なし)
 (74) 代理人 100106909
 弁理士 棚井 澄雄
 (74) 代理人 100089037
 弁理士 渡邊 隆
 (72) 発明者 ユー・チャン
 アメリカ合衆国・マサチューセッツ・01
 581・ウエストボロー・トーマス・ニュー
 ートン・ドライブ・36

最終頁に続く

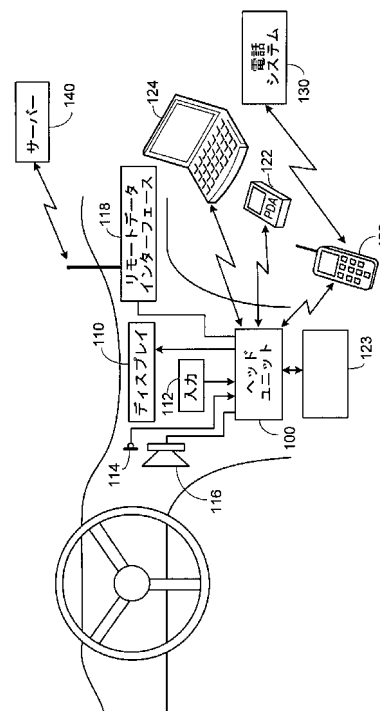
(54) 【発明の名称】 自動車インターフェース

(57) 【要約】

【課題】自動車システムは、自動車または他のタイプの車両内での制御および通信機能のための統合されたユーザーインターフェースを提供する。

【解決手段】ユーザーインターフェースは、制御盤、例えばダッシュボードまたはステアリングホイールに取り付けられた制御盤を用いて、ボイスイネーブルインタラクションをサポートするだけでなく、インタラクションの他のモード、例えばマニュアルインタラクションもサポートする。システムは、車両内のデバイスに対するインターフェース、例えば車両内に持ち込まれるモバイル機器に対する無線インターフェースも備えている。システムは、例えば、情報にアクセスするための、リモートサーバーのような情報ソースに対するインターフェースも提供する。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

ボイスインネーブルユーザーインターフェースを提供する方法において、
インターフェースを設定して、音声入力の発声処理するステップを有していることを特徴とする方法。

【請求項 2】

前記音声入力の発声処理するステップは、識別子を、識別子と関連する、冗長な、または曖昧さをなくす情報と共に処理するステップを有していることを特徴とする請求項 1 に記載の方法。

【請求項 3】

追加の、または曖昧さをなくす情報を、識別子を前記情報と関連付けているデータベースから検索するステップを更に有していることを特徴とする請求項 1 に記載の方法。

【請求項 4】

前記データベースは、コンタクトデータベースを有していることを特徴とする請求項 3 に記載の方法。

【請求項 5】

前記冗長な、または曖昧な情報は、識別子と直接連結されることを特徴とする請求項 2 に記載の方法。

【請求項 6】

前記識別子は、名前を含んでいて、かつ、前記冗長な、または曖昧さをなくす情報は、名前の頭文字を含んでいることを特徴とする請求項 2 に記載の方法。

【請求項 7】

インターフェースのための一般的な設定情報を記憶するステップと、
インターフェースのためのユーザー固有の設定情報を記憶するステップと、
一般的な設定情報およびユーザー固有の設定情報を用いて、ユーザーからの音声入力の処理を可能にするステップと、
音声入力処理の結果に基づいて、ユーザー固有の設定情報を選択的に更新するステップとを更に有していることを特徴とする請求項 1 に記載の方法。

【請求項 8】

前記更新は、入力と関連する得点が、不正確な認識仮説が予め決められた閾値以内の得点を有していることを示すとき、入力の正しい認識の上で、実行されることを特徴とする請求項 7 に記載の方法。

【請求項 9】

前記更新は、ユーザーからの更なる音声入力を必要とせずに、実行されることを特徴とする請求項 8 に記載の方法。

【請求項 10】

ユーザー固有の設定情報は、辞書項目に対するユーザー固有の発音情報を含んでいて、そのための一般的な発音情報は、一般的な設定情報の中に含まれていることを特徴とする請求項 7 に記載の方法。

【請求項 11】

ユーザー固有の設定情報の更新は、インターフェースの使用中に、実行されることを特徴とする請求項 7 に記載の方法。

【請求項 12】

前記更新は、ユーザーからの音声入力の誤認識に基づくことを特徴とする請求項 11 に記載の方法。

【請求項 13】

ユーザー固有の設定情報の更新は、ユーザーからの音声入力の認識結果に基づくことを特徴とする請求項 11 に記載の方法。

【請求項 14】

10

20

30

40

50

インターフェースの複数のユーザーの各々のための設定情報を提供するステップと、
インターフェースのユーザーを決定するステップと、
決定されたユーザーに対する設定情報を用いて、インターフェースを適合させるステップとを更に有していることを特徴とする請求項 1 に記載の方法。

【請求項 15】

前記ユーザーを決定するステップは、ユーザーからの口頭での入力に対して話者識別プロセスを適用するステップを含んでいることを特徴とする請求項 14 に記載の方法。

【請求項 16】

音声入力の発声処理のためにインターフェースを設定する前に、
語彙項目のテキストの明細を受け入れるステップと、
語彙項目に対する発音を決定するステップとを有して、
音声入力の発声処理のためにインターフェースを設定した後に、
決定された発音を用いて、口頭での入力の中の語彙項目を自動的に認識するステップと

10

もし、認識された語彙項目と関連する得点が、予め決められた範囲内に入るならば、受け入れられた口頭での入力からの 1 つ以上の語彙項目の発音を決定するステップとを有していることを特徴とする請求項 1 に記載の方法。

【請求項 17】

音声入力は、ユーザーからの通信コマンドを含んでいて、
更に、通信コマンドから宛先情報を決定するステップと、
決定された宛先情報に基づいて、ナビゲーションシステムにコマンドを発するステップとを有していることを特徴とする請求項 1 に記載の方法。

20

【請求項 18】

前記通信コマンドは、テレフォングダイヤリングコマンドを含んでいることを特徴とする請求項 17 に記載の方法。

【請求項 19】

前記通信コマンドから決定される宛先へのルートを決めるステップを更に有していることを特徴とする請求項 17 に記載の方法。

【請求項 20】

前記通信コマンドは、口頭でのコマンドを含んでいることを特徴とする請求項 17 に記載の方法。

30

【請求項 21】

前記決定するステップは、通信コマンドと、コンタクトデータベース内の宛先情報との関連に基づくことを特徴とする請求項 17 に記載の方法。

【請求項 22】

第 1 のデバイスで、第 2 のデバイスからコンタクトデータを受け取るステップと、
前記第 1 のデバイスで、データソースからの追加データを要求して、コンタクトデータを増やすステップと、
第 2 のデバイスからの受信コンタクトデータおよびデータソースからの追加データを用いて、ユーザーインターフェースを設定するステップとを更に有していることを特徴とする請求項 1 に記載の方法。

40

【請求項 23】

前記第 1 のデバイスは、車両に基づくデバイスを含んでいて、第 2 のデバイスは、モバイル機器を含んでいることを特徴とする請求項 22 に記載の方法。

【請求項 24】

前記データソースは、リモートサーバーに常駐していることを特徴とする請求項 22 に記載の方法。

【請求項 25】

リモートサーバーから要求される追加データは、コンタクトデータ内の項目と関連する情報を含んでいることを特徴とする請求項 22 に記載の方法。

50

【請求項 26】

コンタクトデータ内の項目は、名前および関連する電話番号を含んでいて、要求される追加データは、この項目と関連する住所情報を含んでいることを特徴とする請求項 25 に記載の方法。

【請求項 27】

ナビゲーションシステムからテキストのストリート情報を得るステップと、
テキストのストリート情報を、その口頭での等価物に変換するステップと、
ユーザーに口頭での等価物を提示するステップとを更に有していることを特徴とする請求項 1 に記載の方法。

【請求項 28】

グラフィックディスプレイ上でユーザーに視覚的なキューを提示するステップを更に有して、

この視覚的なキューは、ユーザーが直接ディスプレイを見る必要がなく、ユーザーの周辺視野内でのユーザーによる検出に適していることを特徴とする請求項 1 に記載の方法。

【請求項 29】

前記視覚的なキューは、口頭インターフェースの状態に関連していることを特徴とする請求項 28 に記載の方法。

【請求項 30】

前記視覚的なキューを提示するステップは、口頭インターフェースが、口頭でのコマンドを受け入れる状態にあることの表示を提供するステップを含んでいることを特徴とする請求項 29 に記載の方法。

【請求項 31】

前記視覚的なキューを提示するステップは、口頭インターフェースが、ユーザーから口頭でのコマンドを受け入れたことの表示を提供するステップを含んでいることを特徴とする請求項 29 に記載の方法。

【請求項 32】

調節の前に制御設定を記録することを含む、音声入力対話の開始で、音響環境と関連する 1 つ以上の制御設定を調節するステップと、

口頭での対話の後に、制御設定を元に戻すステップとを更に有していることを特徴とする請求項 1 に記載の方法。

【請求項 33】

前記制御設定は、音量設定を含んでいることを特徴とする請求項 32 に記載の方法。

【請求項 34】

前記音量設定は、音声の音量設定を含んでいることを特徴とする請求項 33 に記載の方法。

【請求項 35】

前記制御設定は、ノイズを発生するデバイスに対する設定を含んでいることを特徴とする請求項 32 に記載の方法。

【請求項 36】

ノイズを発生するデバイスに対する設定は、ファンの設定を含んでいることを特徴とする請求項 35 に記載の方法。

【請求項 37】

マルチモード車内ユーザーインターフェースを提供する方法において、
第 1 の対話モードにおいて、コマンドの第 1 の部分を受け入れるステップと、
ユーザーインターフェースに関する状態を検出するステップと、
検出された状態に基づいて、第 2 の対話モードを決定するステップと、
第 2 の対話モードにおいて、コマンドの第 2 の部分を受け入れるステップとを有して、

前記第 1 の対話モードおよび第 2 の対話モードは、口頭対話モードおよびマニュアル対話モードのそれぞれであることを特徴とする方法。

10

20

30

40

50

【請求項 38】

ユーザーに、第1の対話モードおよび第2の対話モードのうちの一つだけに実質的に関連する情報を提示するステップを更に有していることを特徴とする請求項37に記載の方法。

【請求項 39】

前記ユーザーインターフェースに関する状態を検出するステップは、ユーザーインターフェースの環境内でのノイズレベルを検出するステップを含んでいることを特徴とする請求項37に記載の方法。

【請求項 40】

コマンドの第1の部分の受け入れおよびコマンドの第2の部分の受け入れが実行され、コマンドの第1の部分の反復を必要とせずに、コマンドの連続的な入力を提供することを特徴とする請求項37に記載の方法。

10

【請求項 41】

通信媒体を通じて第1のデバイスと第2のデバイスを関連付ける方法において、
第1のデバイスでデータを生成するステップと、
第1のデバイスで生成されたデータをユーザーに提示するステップと、
第2のデバイスに関連付け要求を送信するステップと、
第1および第2のデバイスの関連付けを許可するユーザーからの入力時に、第2のデバイス上で、データを受け入れるステップとを有していることを特徴とする方法。

【請求項 42】

前記第1のデバイスは、車両に基づくデバイスを含んでいて、第2のデバイスは、モバイル機器を含んでいることを特徴とする請求項41に記載の方法。

20

【請求項 43】

前記モバイル機器は、携帯電話およびPDAのうちの一つを含んでいることを特徴とする請求項42に記載の方法。

【請求項 44】

第1のデバイス上でボタンを押すことによる開始ステップを更に有していることを特徴とする請求項41に記載の方法。

【請求項 45】

ブルートゥース通信アプローチに従って、第1のデバイスと第2のデバイスを関連付けるステップを更に有していることを特徴とする請求項41に記載の方法。

30

【請求項 46】

第1のデバイスで関連付け手続を開始するステップと、
第1のデバイスで1つ以上の候補デバイスを検出するステップと、
ユーザーから候補デバイスの中の第2のデバイスの選択を受け入れるステップと、
第1のデバイスからユーザーに関連付け情報を提供するステップと、
第1のデバイスから第2のデバイスに要求を送信するステップと、
第2のデバイスでユーザーからの関連付け情報を受け入れるステップとを更に有していることを特徴とする請求項41に記載の方法。

【発明の詳細な説明】

40

【技術分野】

【0001】

本発明は、自動車インターフェースに関する。

【背景技術】

【0002】

自動車は、今日、ユーザーとの対話のための様々なインターフェースをサポートしている。このインターフェースは、ユーザーに、情報（例えば、ナビゲーション、車両の状態）、通信（例えば、セルラー電話）、または制御（例えば、環境制御のような車両システムの制御）に対するアクセスを提供する。このようなシステムは、自動車内でホストとして働いてもよいし、自動車から離れてホストとして働いて、通信システムを通じてアクセ

50

スされてもよい。音声認識は、自動車内の人と、このようなシステムとの間で、口頭での対話を可能にするために、自動車内で用いられてきた。

【0003】

自動車は、今日、セルラー電話のような外部デバイスの統合のために、いくつかのインターフェースを提供する。例えば、外部デバイスとのこのようなインターフェースは、外部デバイスと自動車システム間のブルートゥースのような無線リンクを使用することができる。

【発明の開示】

【課題を解決するための手段】

【0004】

一態様において、一般に、ボイスインテリジェントユーザーインターフェースは、インターフェースに対する一般的な設定情報を記憶するステップと、インターフェースに対するユーザー固有の設定情報を記憶するステップとを備えている。ユーザーからの音声入力の処理は、一般的な設定情報およびユーザー固有の設定情報を用いて可能になる。ユーザー固有の設定情報は、音声入力の処理結果に基づいて、選択的に更新される。

【0005】

この態様は、以下の特徴のうちの1つ以上を有していてもよい。

【0006】

更新は、入力と関連する得点が、不正確な認識仮説が予め決められた閾値以内の得点を有していることを示すとき、入力の正しい認識の上で、実行される。

【0007】

更新は、ユーザーからの更なる音声入力を必要とせずに、実行される。

【0008】

ユーザー固有の設定情報は、辞書項目に対するユーザー固有の発音情報を含んでいて、そのための一般的な発音情報は、一般的な設定情報に含まれている。

【0009】

ユーザー固有の設定情報の更新は、インターフェースの使用中に、実行される。

【0010】

更新は、ユーザーからの音声入力の誤認識に基づく。

【0011】

ユーザー固有の設定情報の更新は、ユーザーからの音声入力の認識結果に基づく。

【0012】

別の態様において、一般に、ボイスインテリジェントユーザーインターフェースは、識別子と関連する、冗長な、または曖昧さをなくす情報と共に、識別子の発声を処理するように設定される。インターフェースは、識別子を前記情報と関連付けているデータベースから、追加の、または曖昧さをなくす情報を検索することができる。例えば、このデータベースは、コンタクトデータベースを含んでいる。冗長な、または曖昧な情報は、識別子と直接連結され得る。一例として、識別子は、名前を含んでいて、冗長な、または曖昧さをなくす情報は、名前の頭文字を含んでいる。

【0013】

別の態様において、一般に、ボイスインテリジェントユーザーインターフェースを提供する方法は、インターフェースのユーザーの組の各々に対する設定情報を提供するステップを有している。インターフェースのユーザーが決定され、このインターフェースは、決定されたユーザーのための設定情報を用いて適合される。ユーザーを決定するステップは、ユーザーからの口頭での入力に対して、話者識別プロセスを適用するステップを有していてもよい。

【0014】

別の態様において、一般に、通信媒体を通じて第1のデバイスと第2のデバイスを関連付ける方法は、第1のデバイスでデータを生成するステップと、第1のデバイスで生成されたデータをユーザーに提示するステップと、第2のデバイスに関連付け要求を送信する

10

20

30

40

50

ステップとを有している。第1および第2のデバイスの関連付けを許可するユーザーからの入力時に、第2のデバイス上で、データは受け入れられる。

【0015】

この態様は、以下の特徴のうちの1つ以上を有していてもよい。

【0016】

第1のデバイスは、車両に基づくデバイスであり、第2のデバイスは、モバイル機器である。例えば、モバイル機器は、携帯電話およびPDAのうちの1つである。

【0017】

この方法は、第1のデバイス上のボタンを押すことによる開始ステップを更に有している。

【0018】

第1のデバイスと第2のデバイスの関連付けは、ブルートゥース通信アプローチに従う。

【0019】

別の態様において、一般に、ボイスイネーブルユーザーインターフェースを設定する方法は、第1のデバイスで、第2のデバイスからコンタクトデータを受け取り、かつデータソースからの追加データを要求して、コンタクトデータを増やすステップを有している。ユーザーインターフェースは、第2のデバイスから受信したコンタクトデータおよびデータソースからの追加データを用いて設定される。

【0020】

この態様は、以下の特徴のうちの1つ以上を有していてもよい。

【0021】

第1のデバイスは、車両に基づくデバイスであり、第2のデバイスは、モバイル機器である。

【0022】

データソースは、リモートサーバーに常駐している。

【0023】

リモートサーバーから要求される追加データは、コンタクトデータ内の項目に関連する情報を含んでいる。例えば、コンタクトデータ内の項目は、名前および関連する電話番号を含んでいて、要求される追加データは、この項目と関連する住所情報を含んでいる。

【0024】

別の態様において、一般に、車内ボイスイネーブルユーザーインターフェースを提供する方法は、ユーザーから通信コマンドを受け入れるステップを有している。そして、宛先情報が、通信コマンドから決定される。そして、ナビゲーションシステムが、決定された宛先情報に基づいて、コマンドを発せられる。

【0025】

この態様は、以下の特徴のうちの1つ以上を有していてもよい。

【0026】

通信コマンドは、テレフォンダイヤリングコマンドを含んでいる。

【0027】

宛先へのルートは、通信コマンドから決定される。

【0028】

通信コマンドは、口頭でのコマンドを含んでいる。

【0029】

決定は、通信コマンドと、コンタクトデータベース内の宛先情報との関連に基づく。

【0030】

別の態様において、一般に、車内ボイスイネーブルユーザーインターフェースを提供する方法は、ナビゲーションシステムからテキストのストリート情報を得るステップと、テキストのストリート情報を、その口頭での等価物に変換するステップとを有している。そして、口頭での等価物が、ユーザーに提示される。

10

20

30

40

50

【0031】

別の態様において、一般に、車内ボイスインナーブルインターフェースを提供する方法は、グラフィックディスプレイ上でユーザーに視覚的なキューを提示するステップを有している。視覚的なキューは、ユーザーが直接ディスプレイを見ることを必要とせずに、ユーザーの周辺視野内でのユーザーによる検出に適している。

【0032】

この態様は、以下の特徴のうちの1つ以上を有していてもよい。

【0033】

視覚的なキューは、口頭インターフェースの状態に関連している。

【0034】

視覚的なキューの提示は、口頭インターフェースが、口頭コマンドを受け入れる状態にあることの表示を提供するステップ、例えば、口頭インターフェースが、ユーザーから口頭でのコマンドを受け入れたことの表示を提供するステップを有している。

【0035】

別の態様において、一般に、車内ボイスインナーブルインターフェースを提供する方法は、口頭での対話の開始時に、音響環境に関する1つ以上の制御設定を調節するステップを有していて、調節の前に、制御設定を記録するステップと、口頭での対話の後に、制御設定を元に戻すステップとを有している。

【0036】

制御設定は、音量設定を含んでいてもよく、これは、例えば、音声の音量設定またはノイズを発生するデバイスに対する設定を含んでいる。ノイズを発生するデバイスは、ファンを含んでいてもよい。

【0037】

別の態様において、一般に、マルチモード車内ユーザーインターフェースを提供する方法は、第1の対話モードにおいて、コマンドの第1の部分を受け入れるステップと、このユーザーインターフェースに関する状態を検出するステップと、検出された状態に基づいて、第2の対話モードを決定するステップと、第2の対話モードにおいて、コマンドの第2の部分を受け入れるステップとを有している。第1の対話モードおよび第2の対話モードは、口頭対話モードおよびマニュアル対話モードのそれぞれである。

【0038】

この態様は、以下の特徴のうちの1つ以上を含んでいてもよい。

【0039】

この方法は、ユーザーに、第1の対話モードおよび第2の対話モードのうちの1つだけに実質的に関連する情報を提示するステップを更に有している。

【0040】

ユーザーインターフェースに関する状態を検出するステップは、ユーザーインターフェースの環境内でのノイズレベルを検出するステップを有している。

【0041】

コマンドの第1の部分の受け入れおよびコマンドの第2の部分の受け入れが実行され、コマンドの第1の部分の反復を必要とせずに、コマンドの連続的な入力を提供する。

【0042】

別の態様において、一般に、ボイスインナーブルユーザーインターフェースを提供する方法は、語彙項目のテキストの明細を受け入れるステップと、語彙項目に対する発音を決定するステップと、ユーザーからの口頭での入力を受け入れるステップと、決定された発音を用いて、口頭での入力の中で語彙項目を自動的に認識するステップとを有している。もし、認識された語彙項目と関連する得点が、予め決められた範囲内に入るならば、語彙項目のうちの1つ以上の発音が、受け入れられた口頭での入力から決定される。

【0043】

別の態様において、一般に、デバイス間の関連付けを確立する方法は、第1のデバイスで関連付け手続を開始するステップと、第1のデバイスで1つ以上の候補デバイスを検出

10

20

30

40

50

するステップと、ユーザーから候補デバイスのうちの第2のデバイスの選択を受け入れるステップと、第1のデバイスからユーザーに関連付け情報を提供するステップと、第1のデバイスから第2のデバイスに要求を伝送するステップと、第2のデバイスでユーザーから関連付け情報を受け入れるステップとを有している。

【0044】

他の特徴および利点は、以下の説明および請求項から明らかである。

【発明を実施するための最良の形態】

【0045】

図1は、ブロック図である。

【0046】

図2は、ブロック図である。

【0047】

図3は、フローチャートである。

【0048】

図4は、フローチャートである。

【0049】

図5は、タイミング図である。

【0050】

図1を参照すると、自動車システムは、自動車または他のタイプの車両内で、制御および通信機能のための統合されたユーザーインターフェースを提供する。このユーザーインターフェースは、ボイスイネーブルインタラクションをサポートするのに加えて、他のモードのインタラクション、例えば、制御盤を用いるマニュアルインタラクションもサポートする。制御盤とは、例えば、ダッシュボードまたはステアリングホイールに取り付けられた制御盤（例えば、ボタン、ノブ）である。このシステムは、車両内のデバイスに対するインターフェースも備えている。インターフェースとは、例えば、車両内に持ち込まれるモバイル機器に対する無線インターフェースである。このシステムは、情報ソースに対するインターフェースも提供する。情報ソースとは、例えば、情報にアクセスするためのリモートサーバー（すなわち、自動車の中でホストとして働くわけではないシステム、または自動車と共に移動するわけではないシステム）である。

【0051】

このシステムのヘッドユニット100は、自動車の中で、ユーザー（運転手または他の人）との対話を可能にする他の部品に接続されている。図1に示したシステムの具体例において、グラフィカルディスプレイ110は、ユーザーに対してグラフィカルな出力（または他の視覚的なキュー）を提供する。ユーザーは、入力装置112を用いてマニュアル入力を行うことができる。入力装置112は、（例えば、センターコンソール、ステアリングホイール上に）ボタンを備えていてもよいし、ディスプレイ上に可変（“ソフト”）ボタンを備えていてもよいし、等々。マイクロフォンまたはマイクロフォンアレイ114およびスピーカー116は、それぞれ、このシステムへの音声入力およびこのシステムからの音声出力を可能にする。ヘッドユニット100は、外部デバイスとの通信インターフェースを備えている。外部デバイスとは、例えば、携帯（例えば、セルラー）電話120、および携帯情報端末（PDA）122、組み込み電話123、および携帯用パーソナルコンピュータ124である。このバージョンのシステムにおいて、通信リンクは、無線アプローチを使用し、それは、ブルートゥース規格を使用する。代替案として、他の有線または無線アプローチも、このような通信リンクのために用いられ得る。外部デバイスは、別々のユーザーインターフェースを備えることができ、このユーザーインターフェースを通して、ユーザーは、それらのデバイスと対話することができる。いくつかの外部デバイス、例えば携帯電話120は、無線ネットワークのような外部通信システムとの通信リンクを提供する。このシステムは、リモートデータインターフェース118、例えば無線データトランシーバも備えていて、これは、リモートサーバー140との通信リンクを提供する。

10

20

30

40

50

【 0 0 5 2 】

図 2 を参照すると、ヘッドユニット 1 0 0 は、1 人以上のユーザーに対するプロフィールデータ 2 4 0 を有している。プロフィールデータは、システムを、個々のユーザーに適合またはカスタマイズさせることを可能にする。ヘッドユニット 1 0 0 は、ボイスインテリジブルサブシステム 2 1 0 も有していて、これは、例えば自動音声認識 (A S R) およびテキスト - 音声合成 (T T S) を用いて、システムとユーザー間での音声に基づく対話を可能にする。ボイスインテリジブルサブシステム 2 1 0 は、話者に依存しない (S I) 音声設定データ 2 1 2 を使用するが、これは、ユーザーによって発され得る口頭でのコマンドおよび特徴に関する情報を含んでいる。特徴とは、例えば、それらのコマンドに関する単語の発音、および他の構成要素であり、それは、一般に、特定のユーザーに固有というわけではない。ボイスインテリジブルサブシステム 2 1 0 および S I 音声設定データ 2 1 2 については、以下でより詳細に述べる。ヘッドユニットは、自然に聞こえる音声出力を達成するために、言語の規則および辞書を用いて、言語の普通のテキスト表現を、その口頭での等価物に変換するための、テキストを音声に変換するサブシステム 2 2 0 も有している。入力テキストは、様々な方法、例えばキーボードまたはダッシュボードもしくはステアリングホイール上の制御盤 (ボタン、ノブなど) を介して入力され得る。または、以下のような様々なタイプのファイルから読み出され得る。ワープロファイル、ウェブページ、データベースレコード、電子メールメッセージなど。ヘッドユニットは、ユーザーの口頭での入力に基づいてユーザーを認証するための話者確認サブシステム 2 2 2 も有している。

10

【 0 0 5 3 】

この具体例としてのシステムにおいては、プロフィールデータ 2 4 0、ボイスインテリジブルサブシステム 2 1 0、S I 音声設定データ 2 1 2 は、全てヘッドユニット 1 0 0 内に設置されていることが示されているが、上記のサブシステムおよびデータ (例えば、プロフィール、音声設定データ) は、代替案として、様々な異なる物理的な構成要素の中に設置され得ることに留意する必要がある。

20

【 0 0 5 4 】

一例として、ボイスインテリジブルサブシステム 2 1 0 は、Burlington, MA 01803 に本社がある ScanSoft 社からの A S R および T T S ソフトウェアを使用する。ScanSoft A S R エンジン (1 つの単語の不連続な発声を必要とするのと比較して) 連続的な入力音声をサポートし、(拘束を受けていない口頭でのテキストの口述と比較して) コマンドおよび制御文法の使用をサポートし、話者に依存しない設定をサポートし、設定をユーザーに適合させる機能、例えば、追加の発声を用いて特定の単語のための設定情報を訓練する機能をサポートする。ScanSoft A S R エンジン上で構築された具体例としてのボイスインテリジブルサブシステム 2 1 0 は、以下の構成要素を有している。

30

- ・ 音声入力を一連の音響特性に変換し、それぞれは、限られた時間の範囲内の音声入力の特徴を表す数値ベクトルとして表される特徴抽出器。

- ・ 例えば、特定の音に対する音響特性の統計的分布に基づいて、言語内の様々な音声の数学的表現を提供する音響モデル。1 つの音響モデルは、1 つのクラスの人々にとって一般的であり得る。または、1 つまたは小さいグループの人々に固有であり得る。例えば、話者に依存しない (S I) モデルは、多くの人々の声のサンプルから集められたデータに基づく一方で、話者に依存するモデルは、1 人の人の声のサンプルから集められたデータに基づく。S I モデルは、例えば、性または年齢に依存している 1 つのクラスの人々を表し得る (例えば、男性モデル、女性モデル、または子供モデルなど)。この音響モデルで表される音は、(例えば、必ずしも単語の言語構造に基づかなくても)、音節に基づいて、音素に基づいて、または文脈に依存する音素に基づいて、特定の単語と関連付けられ得る。

40

- ・ 話され得る妥当なフレーズを記述する文法、または、どちらの文が話され得るかを決定する規則。例えば、文法は、テキスト表現 (例えば、B N F、バックス - ナウア記法) を用いて表され得る。

- ・ 音響モデルで表される音に関して、単語と、それらの単語または全部のフレーズの表現

50

との間のマッピングを提供する辞書。例えば、音素に基づく音響モデルの場合、辞書は、単語またはフレーズの音声の筆写をリストアップして、各々は、音素のシーケンスとして、代替シーケンスのリストとして、または（例えば、音素を有するネットワーク内の弧にラベルを付けることによって）各経路が特定のシーケンスを表すネットワークとして表される。辞書は、音響モデルと共にあるので、一般的または固有であり得る。例えば、1つのバージョンのシステムの中で、S I 辞書が用いられ、加えて、最初に見つけられなかった単語または辞書の中で適切に表現されていなかった単語のために、S D 例外辞書も用いられる。

- ・書記素 - 音素 (G 2 P) モジュール。これは、自動化されたシステムであり、書かれたテキストストリングを辞書のような筆写に変換する。

- ・文法と、1つ以上の辞書（例えば、S I 辞書および例外辞書）とを結合して、実行時設定データを計算する実行時設定エンジン。

- ・実行時設定データに従って、音声入力の特徴を音響モデルと比較して、結果として最適な組み合わせを提供する認識エンジン。

【0055】

上記の具体例としてのボイスイネーブルサブシステムにおいては、S I 音声設定データ 212 は、辞書および文法を含んでいる。ボイスイネーブルサブシステム 210 の具体的な実施に応じて、S I 音声設定データ 212 は、より多いか、またはより少ない構成要素を含んでいてもよい。

【0056】

1人のユーザーのためのプロフィールデータ 240 は、コンタクトデータ 242 を含んでいて、例えば、それは、人々の名前（または同等に、場所、企業など）および他の関連する項目、例えば電話番号、住所、関心がある地点 (P O I)、カレンダーまたはスケジュール、および人または P O I に関する情報を提供するための任意の項目、例えば P O I のイベントおよび過去の訪問の経験を記述しているメモを含んでいる。プロフィールデータ 240 は、話者に依存する (S D) 音声設定データ 244 も含んでいて、これは、そのユーザーに特有のものである。以下で更に詳細に述べるが、S D 音声設定データ 244 は、精度を改善することができる情報を含んでいて、これによって、例えば、ユーザーによって話される名前のユーザー固有の発音を含むことによって、ボイスイネーブルサブシステム 210 は、ユーザーの発声を認識することができる。一般に、S D 音声設定データ 244 は、訓練のために、A S R エンジンからの設定データの修正されたか更新された部分、例えば辞書内のユーザー固有の単語筆写項目および具体例としてのボイスイネーブルサブシステム内の再編集された文法を含んでいる。

【0057】

ヘッドユニット 100 は、無線サブシステム 230 も有していて、これは、外部デバイス、例えば携帯電話 120 との通信を確立かつ維持するための通信サービスを提供する。無線サブシステム 230 および携帯電話 120 は、それぞれペアリングモジュール 232 および 252 を有していて、これは、外部デバイス（例えば、電話 120、P D A 122、またはコンピュータ 124）がヘッドユニット 100 に近接しているとき、通信を確立するための手続を実現する。ペアリング手続の詳細は、この説明の中で後述される。

【0058】

ヘッドユニット 100 に接続され得る外部デバイスは、オプションとして、ヘッドユニット 100 によって用いられるデータを有している。例えば、図 2 に示したように、携帯電話 120 は、コンタクトデータ 250 を有していて、これは、名前および対応する電話番号、およびオプションの追加情報、例えばコンタクト（例えば、そのユーザーに対して知られている人、場所、または実体）のストリートアドレスを含んでいる。1つの動作モードにおいて、携帯電話がヘッドユニット 100 と通信しているとき、このようなコンタクトデータ 250 は、無線サブシステム 230 を通してヘッドユニット 100 に転送され、ユーザーのプロフィールデータ 240 のコンタクトデータ 242 を更新またはポピュレートするために用いられる。

10

20

30

40

50

【 0 0 5 9 】

ヘッドユニット 1 0 0 内のコンタクトデータ 2 4 2 は、様々な方法で、例えば C D / D V D、ハードドライブ、リモートサーバー、P D A、セルフォン等から、ユニットに与えられ、タスク、例えばハンズフリーネームダイヤリングのために用いられる。ヘッドユニットは、コンタクトデータの音声登録を受け入れることができるのに加えて、テキストに基づく登録、例えば、外部デバイスから提供されるデータに基づく登録も受け入れることができる。音声登録においては、ユーザーは、自動車内で、名前および番号を話すことができ、ヘッドユニットは、それらの名前および対応する番号を録音して（例えば、処理して、処理された表現を記憶して）、登録する。ヘッドユニットのボイスインイーネーブルサブシステム 2 1 0 は、それから、タスク、例えば、ユーザーが以前に登録した名前を後に話すボイスインイーネーブルダイヤリングのために、これらの録音された名前を使用し、システムは、ヘッドユニットに接続されたセルラー電話を通して、対応する番号にダイヤルする。

10

【 0 0 6 0 】

図 3 を参照すると、テキストに基づく登録において、名前（例えば、人、場所、実体、または他の項目（例えば、歌のタイトルまたはオーディオディスク））が、入力 3 1 0 として（すなわち、テキスト形式で）ヘッドユニットに提供され、システムは、S I 音声設定データ 2 1 2 を増やして、それらの名前を含む口頭での発声を受け入れて認識する。ボイスインイーネーブルサブシステム 2 1 0 は、それから、S I 音声設定データ 2 1 2 を使用して、口頭ユーザー入力 3 3 0 を処理する。ボイスインイーネーブルサブシステムは、また、口頭ユーザー入力 3 3 0 を使用して、S D 音声設定データ 2 4 4 を決定することができる。この一例は、音声登録モードであり、ここで、テキスト入力 3 1 0 の提供よりはむしろ、ユーザーは、システムによって認識されるべきコマンドまたは名前を話す。他の例において、その幾つかは、更に以下で論じられるが、口頭ユーザー入力 3 3 0 は、S D 音声設定データ 2 4 4 を更新するために用いられ、それは、S I 音声設定データ 2 1 2 と共に用いられる。1 つのシナリオにおいて、S I 音声設定データ 2 1 2 のみを用いる認識が成功しなかったとき、S D 音声設定データ 2 4 4 を更新するために、音声登録が開始される。このように、設定データの選択された部分だけが、S D 構成要素を持つ必要があり、テキスト登録によって提供される便宜、および音声登録を用いて得られた S D 設定データによって提供される精度を提供する。

20

【 0 0 6 1 】

システムによって認識され得る名前またはコマンドのテキストに基づく明細の一例として、ヘッドユニット 1 0 0 が外部デバイスと通信するとき、テキストに基づく情報（例えば、名前または他の単語のつづりを略さずに書いた表現）は、例えば携帯電話 1 2 0 内のコンタクトデータ 2 5 0 から、ヘッドユニットにダウンロードされる。ヘッドユニットのボイスインイーネーブルサブシステムは、それから、テキスト情報の発音（「音声のつづり」）を決定する自動的な方法、例えば、辞書検索 3 2 0 または自動書記素 - 音素手続 3 2 2 を使用して、システムが、テキスト登録からの名前または他の単語を含む口頭コマンドを受け入れて処理することを可能にする。例えば、従来 of S I 音声認識技術は、テキスト情報を処理するため、およびテキスト情報に基づく単語（語彙）を含むユーザーの発声を認識するために用いられる。

30

40

【 0 0 6 2 】

登録のための音声に基づくアプローチおよびテキストに基づくアプローチに加えて、ヘッドユニット 1 0 0 は、ユーザーが口頭コマンドの中で用いることができる、語彙の登録に対する更なるアプローチを使用する。これらの更なるアプローチのうちのいくつかは、システムによって受け入れられる単語のための音声サブシステムによって用いられる、発音の精度に関する問題に対処することに狙いを付けている。例えば、これらの問題は、予測が難しい標準の発音、または、例えば、外国風のアクセントによる、もしくは、より一般的に、ユーザーの国家的、地域的、文化的、または言語学的バックグラウンドに基づく、ユーザーに固有の発音に関する。

【 0 0 6 3 】

50

テキスト表現から単語の発音を決定するために、ボイスイネーブルサブシステム 2 1 0 によって用いられる 1 つのアプローチは、S I 辞書にアクセスすることを含んでいる。S I 辞書内の単語に対して、発音は、それらの単語を含んでいる発音が話されたかどうかを決定するために、ユーザーからの音声入力を処理するとき、音声認識エンジンによって用いられる。例えば、外部デバイスからのダウンロードの結果として、システムに動的に追加される、全ての名前または新しい単語をカバーするのに十分大きい辞書を有することは、実際的ではないだろう。ボイスイネーブルサブシステム 2 1 0 によって用いられる第 2 のアプローチは、単語のテキスト表現を受け入れるために、かつ辞書の中で用いられるタイプの表現を出力するために、自動化された書記素 - 音素 (G 2 P) 手続を実行することを含んでいる。G 2 P 手続のための辞書および音声設定データは、S I 音声設定データ 2 1 2 の中に含まれていて、それは、一般にいかなる特定のユーザーにも固有ではない。

【 0 0 6 4 】

名前の認識は、名前の発音が必ずしも予め正確に予測されることができないという事実のために、挑戦していてもよい。同様に、G 2 P 手続は、一般に、ユーザーの特徴、例えば彼または彼女の自国語および文化に固有ではない。

【 0 0 6 5 】

ボイスイネーブルサブシステム 2 1 0 は、特定のユーザーのプロフィールデータ 2 4 0 からの S D 音声設定データ 2 4 4 と同様に、必ずしも特定のユーザーに固有であるというわけではない、S I 音声設定データ 2 1 2 の組み合わせを用いることができる。例えば、2 つのタイプの設定データは、同時に用いられて、口頭ユーザー入力を認識する。代替案として、S I 設定データ 2 1 2 が最初に用いられて、その使用の中で決定される組み合わせまたは得点に応じて、S D データ 2 4 4 が次に用いられてもよい。または、S I 音声設定データ 2 1 2 および S D 設定データ 2 4 4 に基づく認識の得点が比較され、比較に基づいて 1 つが選択される。

【 0 0 6 6 】

音声認識に対する 1 つのアプローチは、2 つのステップを用いて、ユーザー固有の発音を決定する。第 1 に、変化を含む辞書から、または、例えば、ネットワーク表現における、複数の変化を生成する G 2 P 手続から、1 つの単語に対する複数の発音の組が決定される。第 2 のステップにおいて、例えば、ユーザーからの口頭入力に基づいて、一つの最善の (または、より一般的に、複数の最も一致しているサブセットの) 一致している発音が選択される。

【 0 0 6 7 】

システムのいくつかのバージョンにおいて、システムが特定のユーザーのために S D 音声設定データ 2 4 4 を使用するとき、それは、S I 設定データ 2 1 2 に対するサポートも維持する。このようにすれば、もし、システムがあるユーザーのプロフィールデータを用いていながら、他のユーザーがコマンドを話したとしても、他のユーザーの発音は認識され得る。これは、たとえ、プロフィールデータと関連があるユーザーが、非常に特殊な設定を有していたとしても、可能である。

【 0 0 6 8 】

S D 音声設定データ 2 4 4 を生成するための他のアプローチは、ユーザーの口頭での入りに依存する。例えば、ボイスイネーブルサブシステム 2 1 0 は、不正確な、話者に依存しない音声の筆写のいくつかを訂正する機能を提供する。不正確な、話者に依存しない音声の筆写は、話者に依存しない辞書のみを使用することによって起こり得る。この機能は、自動化された音声認識の中で用いる発音を決定する際に、ユーザーの発音を使用する。1 つのこのようなアプローチは、認識の中で用いる発音を決定するために、ユーザーが全ての新しい名前を話すことを必要とする。好ましくは、システムは、自動的にユーザーの名前の発音を訓練して、ユーザーの介入を必要とせずに、すなわちプロセスを認識して、S I 辞書 / 規則を用いるか、またはそのユーザーのためのユーザー固有の発音を決定するかどうかに関する決定を下す。

【 0 0 6 9 】

10

20

30

40

50

システムの1つのバージョンにおいて、S I 辞書およびG 2 P 手続は、認識文法、例えば特別な音声タスクのための文法における全ての単語の発音表記を形成するために用いられる。このようなタスクの例は、音声ダイアリングであり、この中で、ユーザーは、長いコマンドの文脈の中で名前を言うことができる（例えば、「ジョンスミスに電話をして下さい」）。加えて、システムは、特定のユーザーに固有な例外辞書の使用をサポートする。例外辞書の1つの用途は、特別な単語の追加の表現を含ませることである。例えば、もしユーザーが「スミス」を発音する特別な方法を有するならば、例外辞書は、その単語のための例外項目を含むように、システムによって設定され得る。実行時設定エンジンが、エンジン用に、実行時設定データを計算するときに、認識文法の中で単語「スミス」を表現する際に、実行時設定エンジンは、S I 発音および例外発音を含む。代替案として、例外発音だけが、実行時設定データを生成するために用いられる。代替案として、名前のユーザー固有の発音を適応させるために、システムは、例外辞書の中に、例えばスミス - 1 のような、新しい単語を導入し、かつ文法を修正して、元のスミスという単語に加えて、スミス - 1 という単語も、両方とも可能にする。従って、たとえ例外辞書がS I 辞書より優先されるとしても、単語の両方の形式は、入力発声を処理するとき、認識エンジンによって受け入れられるであろう。以下の説明において、辞書を更新するとは、上記のように、例外辞書の中に項目を追加または更新することに相当し得る。

10

【0070】

図4を参照すると、ボイスイネーブルサブシステム210（図2参照）によって実現されるプロセスは、動作中に（すなわち、必ずしも訓練モード中ではない）口頭ユーザー入力330（図3参照）を受け入れるステップを含んでいる。そして、この入力に応じて、おそらくは、発声の中の1つ以上の単語のユーザー固有の発音を含む、ユーザー固有の辞書を更新する。このプロセスにおいて、単語またはフレーズを含む発声、例えばユーザー音声コマンドからの名前は、ボイスイネーブルサブシステムによって認識される（ステップ410）。エンジンは、認識された仮説が発声と一致する程度を示す得点を、仮説と関連する信頼度C1の形（または得点の他の形式）で提供する。もし信頼度が前もってセットされた閾値T1を上回るならば（ステップ412）、システムは、辞書を更新することなく進む（ステップ440への分岐）。もし信頼度が第2のより低い閾値T2の下にあるならば（ステップ414）、システムは、辞書を更新することなく、発声を拒絶する（ステップ430）。しかし、信頼度C1が2つの閾値によって形成される範囲内にあるとき、システムは「未決定」状態420に入る。

20

30

【0071】

「未決定状態」において、システムはユーザーからの正しい答えを確認する（ステップ422）。例えば、ユーザーからの更なるプロンプトおよび口頭または他のモードの入力を用いて、仮説を確かめる。または、暗黙のうちに行う。なぜなら、ユーザーは、認識された答えに基づいて、動作をキャンセルするよりはむしろ受け入れるからである。ユーザーからの応答が、システムが単語/フレーズを正しく認識しなかったことを確認したとき、ユーザーの個人の発音を更新する動作はとられない。ユーザーからの応答が、システムが正しい単語/フレーズを正しく認識したことを確認したが、信頼度が高くなかったとき（ステップ424）、システムは訓練モードに入り、この中でSD設定データが更新される（ステップ426）。この訓練モードの中で、システムは発声を処理し、この発声に基づいてユーザー固有の辞書を更新する。ScanSoft ASRエンジン上で構築された具体例としてのシステムに対して、訓練は、以下のステップを含んでいる。

40

- ・ユーザーの発声を集める、または以前に保存された発声を使用する。
- ・ユーザー単語の筆写を生成する。
- ・ユーザー単語の筆写を（例外）辞書の中に追加する。
- ・この単語を使用する文法を再編集する。
- ・エンジンは、新しい文法を用いて、音声入力を認識する。

【0072】

いくつかの実施形態において、ユーザーが後に同じ単語またはフレーズを発音するとき

50

、もし更新されたSD辞書C2を用いる信頼度と、SI辞書C1を用いる信頼度との間の差が、予め決められた閾値(T3)より大きいならば、更新されたSD辞書が用いられるだろうし、さもなければ、SI辞書が引き続き用いられるだろう。

【0073】

音声イネーブルシステム210は、名前の口頭での入力精度を改善可能な他の機能を提供する。一般に、このアプローチは、名前を話すとき、ユーザーが追加情報を追加することを必要とする。これは、例えば、「ジョンズミス」.S.」のような名前の発声の後の頭文字のような追加情報を追加することによる。追加情報は、冗長であり得る。例えば、頭文字の場合、名前または名前から一意に決定される他の情報から、直接、導き出される。または、追加情報は、情報の曖昧さをなくし得る。例えば、その名前と関連がある複数の項目を有するコンタクトデータベース内の「ジョンズミス」の特定の例を識別する。入力タスクの中で、ユーザーは、コンタクトデータからコンタクトを選択する。このコンタクトデータは、認識文法を構築するために用いられる。認識文法は、このような追加情報と共に発声を受け入れる。ユーザーは、名前を呼ぶことによって、コンタクトに電話をかける。例えば、「ジョンズミスに電話せよ」。しかし、一般的ではなく、かつ/または混乱しやすい名前に対しては、エンジンは、それを正確に認識しないかもしれない。例えば2つの文字から成る追加情報が、音声認識サブシステムによって用いられ、正しい認識仮説を不正確な誤認識仮説から区別する。

10

【0074】

このアプローチに基づく他の例は、以下の通りである。

20

システム：名前と番号をどうぞ。

ユーザー：ボブスミスの自宅。

システム：申し訳ありませんが、もう一度言って下さい。

ユーザー：ボブスミスの自宅の電話。

システム：申し訳ありませんが、別の方法で試みましょう。人の名前に続けて姓名の頭文字を言って下さい。

ユーザー：ボブスミスBSの自宅。

システム：ボブスミスの自宅に電話します。

【0075】

上述したアプローチの中で、コンタクトデータは、外部デバイスによって提供されるが、このデータは、不完全でもよい。例えば、電話番号はあるが、住所情報(これはナビゲーションの構成要素が必要とする可能性がある)は、ユーザーのコンタクトデータから見つからないかもしれない。このシステムの特徴は、見つからない情報が、ヘッドユニット内のデータに、自動または手動で追加され得ることである。これは、外部の情報ソース、例えばリモートサーバー140(図1参照)にアクセスすることによって達成される。例えば、「ホワイトページ」ディレクトリは、無線ネットワークシステムを介したインターネットを通じてアクセスされる。追加される見つからない情報は、音声コマンドの中で用いられる可能性がある。例えば、「メープルストリートファーストに住んでいるジョンズミスに電話せよ」または「ジョンズミスに電話せよ。彼の職場の電話番号は508-555-1212だ。」このようなより複雑な口頭での入力は、自動化された認識の中で、より高い精度を達成することができる。難しい認識問題(例えば2人のジョンズミス)に対して、これらの追加情報の使用のおかげで、このシステムは、まだ話者に依存しないモードを提供することができ、冗長な話者に依存するモードに切り替える必要はない。例えば、「オークストリートのジョンズミス」に電話せよに対して「ウォータードライブのジョンズミス」に電話せよ。

30

40

【0076】

見つからない情報は、他のサブシステム、例えばナビゲーションシステムによっても用いられ得る。例えば、ユーザーは、POI(例えばレストラン)の名前を言うことができる。システムは、自動的に、コンタクトデータから住所情報を抽出し、抽出された住所情報を用いて、ナビゲーションサブシステムにコマンドを発する。各コンタクトデータレコ

50

ードのための複数の項目（セル、自宅、職場の電話番号、好ましいコンタクトリストなど）は、選択するユーザーまたは利用するシステムのために、全てユーザーに提示され得る。いくつかの態様において、例えば音声（例えば口頭での名前または数字の列）によって、またはマニュアル入力によって、ユーザーがテレフォングダイヤリングコマンド（または他のタイプの特定の宛先との通信に関するコマンド）を提供するとき、ナビゲーションシステムは、自動的に、テレフォングダイヤリングコマンドと関連する場所へのナビゲーションルートを決定することができる。または、テレフォングダイヤリングコマンドと関連する場所を示している地図を決定することができる。

【0077】

システムの他の態様において、話者識別機能が、記憶されたセットのユーザーのうちどのユーザーがシステムと対話しているかを決定するために用いられる。この話者識別は、ユーザー固有のシステムの基本設定を設定する。この基本設定は、ユーザー固有のコンタクトデータ、ユーザー固有の発音、または他の基本設定を含んでいてもよい。様々な技術が、どのユーザーがシステムと対話しているかを決定するために用いられ得る。例えば、ヘッドユニットとペアを組んでいる特定の外部デバイス（例えば携帯電話）が、ユーザーを識別するために用いられ得る。話者確認サブシステム222もまた、ユーザー識別のために用いられ得る。他の適切な技術、例えばパスワード、指紋、顔の認識などが、代替案として用いられる。

【0078】

話者識別は、このシステムをよりユーザーフレンドリーにすることができ、かつパーソナライズされたサービスを提供することができる。さらに、音声インターフェースは、制御されているシステムと関連する情報を使用することができる。このように、ユーザーインターフェースは、より「知的」にすることができる。他の例として、音声認識サブシステムが、ユーザーが言った仕事またはフレーズを認識しなかったとき、「もう一度言って下さい」を繰り返す代わりに、サブシステムは、ユーザーの基本設定および履歴に基づいて、可能な選択肢のリストを提供することができる。「ジョンスミスですか、またはジョンスティヴですか。」ユーザーからの正しい答えは、更に、上記のような話者に依存する文法および辞書を生成するために、エンジンを訓練するのに利用され得る。

【0079】

口頭でのユーザーインターフェースシステムにおいて、ユーザーにシステムの状態のフィードバックを提供することは望ましい。例えば、システムが「リスニング」状態にあることをフィードバックすることは、役に立ち得る。自動車への応用においては、ユーザーとの対話を実行するために、道から目を離す必要がないこともまた、運転手のために望ましい。1つのアプローチは、例えば、トーンおよびビープ音の形で、または、録音もしくは合成された助言の形で、音声フィードバックを提供することである。他のアプローチは、運転手が道から目を離す必要がない視覚的フィードバックを提供することである。より詳細には、運転手の周辺視野を用いて知覚され得る視覚的インジケータが、ヘッドユニットによって生成される。このような視覚的インジケータの一例は、グラフィカルディスプレイの比較的広い領域、例えば、ほぼ全領域を使用する。領域全体の視覚的な色または輝度の変化が、システムの状態を示すために用いられる。例えば、特定の色が、システムのマイクロフォンがユーザーの音声入力を聞く準備ができていることを示すために用いられる。領域および色は、ユーザーにとって十分に目立つものであり、ユーザーはディスプレイを直接見る必要はなく、例えば、彼または彼女の周辺視野を用いる。

【0080】

システムの他の態様において、ヘッドユニットは、音声認識エンジンが困難に遭遇しそのようなノイズ状態を検出したとき、ユーザーに警告する。システムは、ノイズレベルを測定し、音声認識が推奨されないとき、ユーザーに警告する。システムは、キャビン内の音響ノイズレベルを測定し、ユーザーが音声認識入力を用いようとするときに、キャビン内にノイズが多すぎるかどうかをユーザーに知らせる。警報は、音声、映像または上述したような視覚的インジケータの形であってもよい。システムは、更に、自動切換によって、

10

20

30

40

50

またはユーザーに選択を勧めることによって、コマンド発行の他のモードに移行することができる。一般に、モード移行は、イベントによって引き起こされ得る。例えば、音声インターフェースから、ディスプレイ上のタッチスクリーンを介するマニュアル入力への移行は、過剰なノイズが車両内で検出されたときに引き起こされ得る。システムの更に他の態様において、移行は、システムが、このシステムとのユーザーの対話の状態を追跡し、かつユーザーによるコマンド発行を続けることができるように、スムーズかつ連続的になされてもよい。例えば、ユーザーは、音声インターフェースを用いて、電話番号を言っていたが、ノイズによって引き起こされたモード移行のために、完了しなかった。音声インターフェースモードからマニュアル入力モードに移行するとき、システムは、完了しなかった口頭での電話番号を移植して、ユーザーのためにディスプレイ上にそれらを示し、コールを続けて完了することができる。

10

【0081】

一般に、ユーザーとの対話のためのインターフェースは、複数のモードの入力、例えば音声、タッチ、ディスプレイおよび他のセンサを有している。システムは、ユーザーを、従来技術のように現在使用中のモードにおけるコマンド発行のために必要でない莫大な量の情報に埋没させることを避けるために、使用中のモードに関する情報を、表示または示すことによって、提示するのみである。例えば、音声インターフェースモードにあるとき、ディスプレイ上のソフトキーのようなタッチ入力は表示されない。ハードドライブのようなオーディオソースが選択されているとき、FMもAMもディスプレイ上に表示されない。このようなマルチモードユーザーインターフェースを提供する1つの例は、音声入力モードにおけるコマンドの最初の部分を受け入れることを必要とする。ユーザーインターフェースに関する状態が検出される。例えば、高いノイズレベルが検出される。システムは、このノイズレベルに基づいて、マニュアルモードが用いられるべきであることを決定する。次に、コマンドの2番目の部分が、マニュアルモードにおいて受け入れられる。コマンドの最初および2番目の部分は、例えば、コマンドの最初の口頭での部分がマニュアル入力を用いて繰り返されることを必要とせずに、スムーズかつ連続的な移行を形成することができる。

20

【0082】

システムの他の態様において、ヘッドユニットは、「ボイスコマンド」シーケンスの間、変更されていた環境設定を元に戻す。このような環境設定は、ファンの速度、または音楽の音量を含んでいてもよい。例えば、ヘッドユニットは、音声入力の間、音楽の音量またはファンの速度を下げて、口頭での対話が終わったら、それらを元に戻してもよい。

30

【0083】

ヘッドユニット100の他の態様は、外部デバイスにインターフェースを提供することに関し、特に、Bluetooth無線通信規格を用いる。この規格によって規定されるプロトコルによれば、2つのBluetoothデバイスが「ペアになる」とき、それらは、データを共有し、かつ制御信号をお互いに送信することができる。例えば、車内音声インターフェースユニットは、Bluetoothを介してPDAまたはセルラー電話と通信して、このデバイスに記憶されたコンタクトリストをダウンロードすることができる。ペアリングプロセスに対する典型的なアプローチは、安全な方法での2つのデバイスのお互いに対する紹介を必要とする。1つのデバイスが他方を「発見」し、共通のコードが2つのデバイス間で渡される。

40

【0084】

ヘッドユニットは、ペアリングの効率を改善可能な単純化されたペアリングプロセスを実行する。一般に、知能をデバイスのうちの中に入れてことによって、ステップを単純化して、ユーザーへの負担を軽くする。図5を参照すると、このような単純化されたペアリングプロセスの一例は、以下のような対話のシーケンスを有している。この例では、「デバイスA」はヘッドユニットであり、「デバイスB」は、例えば、携帯電話である。

1) ユーザーは、デバイスA上でペアリングを開始する(ステップ510)。ボタンを押

50

すことによって、デバイスAは、ブルートゥースデバイスとの対話を開始する。例えば、ハンズフリー電話への応用においては、ヘッドユニットは、携帯用モバイル電話とペアにされるべきであり、ユーザーは、ヘッドユニットに接続されているトークボタンを押す。ユーザーは、特別なペアリングシーケンスについて知っている必要はない。

2) デバイスAは、ペアリングが必要であることを自動的に検出し(ステップ520)、利用可能なデバイスを探して、その範囲を検索し、かつユーザーに選択肢を提示する。例えば、ヘッドユニットは、車両内でデバイス(例えば携帯電話およびPDA)のリストを表示する。

3) ユーザーは、デバイスAのユーザーインターフェース上で、デバイスBを選択する(ステップ530)。

4) デバイスAは、ランダムコードを自動的に生成し、ユーザーに、このコードを、例えばディスプレイを介して提示する。それは、デバイスAのディスプレイまたは別のディスプレイ、例えばヘッドユニットのディスプレイであってもよい。そして、デバイスBに接続要求を送信する(ステップ540)。

5) デバイスBは、デバイスAから接続要求を受信する。ユーザーは、ランダムコードを用いて、この要求を受け入れる選択肢を有している(ステップ550)。

6) ユーザーは、デバイスB上で、この要求を受け入れる(ステップ560)。

【0085】

このアプローチは、両方のデバイスをペアリング/発見モードに設定する必要性を避けることができる。

【0086】

ヘッドユニットは、自動車内の多数のシステムに対するインターフェースを提供する。2つのこのようなシステムは、ナビゲーションシステムおよび携帯電話システムを含んでいる。このシステムの態様は、ダウンロードされたコンタクト情報を、ユーザーインターフェースおよび関連するシステムを設定する際に用いることができることである。この設定は、コンタクト情報に基づいてナビゲーション宛先情報を登録する機能を提供することを含んでいてもよい。例えば、人の名前は、ナビゲーションのための宛先の明細として用いられ得る。いくつかのシナリオにおいて、対応する住所情報は、外部デバイスからダウンロードされた情報の中から利用可能ではなく、むしろ遠隔の情報プロバイダー、例えば電話「ホワイトページ」システムからダウンロードする。コンタクト情報の有効性に関する他の機能は、名前と他の関連する情報の組み合わせに基づく、電話する相手の明細に関する。例えば「メインストリート上のジョンズミス」または「オフィスにいるジョンズミス」。追加の関連情報は、電話する番号の曖昧さをなくすために用いることができ、名前のみに基づいても曖昧さがなくとも、認識精度を改善することができる。

【0087】

システムの他の態様は、ディスプレイ上に表示されるナビゲーションデータベースからのテキストのような情報である。ストリート名、方向(右、左、直進など)が、テキストから音声への変換を介して、音声で運転手に提供され得る。例えば、ナビゲーションマップ上でストリート名およびそこからの距離を表示することの代わりに、かつ/または、このことの他に、ボイスインネーブルサブシステムが「次の4分の1マイル以内にメインストリートを左へ曲がって下さい」と言うことができる。一実施形態において、強調された「メインストリート」は、TTSを介して得ることができる一方で、他は、予め録音しておくことができる。

【0088】

上述した態様の1つ以上が、ヘッドユニットの様々なバージョンの中で実現され得る。すなわち、機能の組み合わせが、様々な実施形態の中でなされ得る。このシステムは、様々なタイプの車両および他の車両ではないユーザー環境に適用できる。ヘッドユニットは、ハードウェア、ソフトウェア、またはハードウェアとソフトウェアの組み合わせを用いて実現することができる。ソフトウェアは、インストラクションを有していて、これは、媒体、例えばリードオンリーメモリ内またはディスクドライブ上に記憶され、かつヘッド

10

20

30

40

50

ユニット内に組み込まれたプロセッサ上で実行され得る。

【0089】

前述の説明は、例示を目的とするものであって、本発明の範囲を限定するためのものではないことは理解されるべきである。本発明の範囲は、添付の請求項の範囲によって定められる。他の実施形態も、この請求項の範囲内である。

【図面の簡単な説明】

【0090】

【図1】ブロック図である。

【図2】ブロック図である。

【図3】フローチャートである。

【図4】フローチャートである。

【図5】タイミング図である。

【符号の説明】

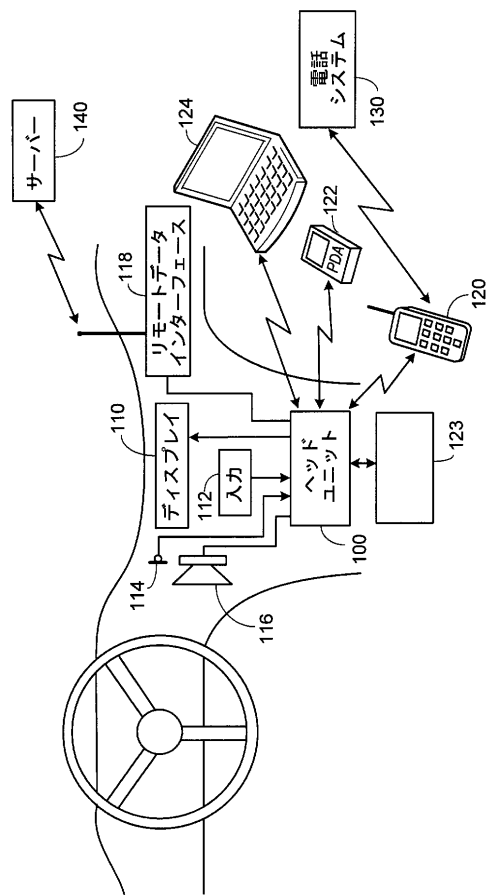
【0091】

- 100 ヘッドユニット
- 110 グラフィカルディスプレイ
- 112 入力装置
- 114 マイクロフォンまたはマイクroフォンアレイ
- 116 スピーカー
- 118 リモートデータインターフェース
- 120 携帯電話
- 122 携帯情報端末 (P D A)
- 123 組み込み電話
- 124 携帯用パーソナルコンピュータ
- 140 リモートサーバー

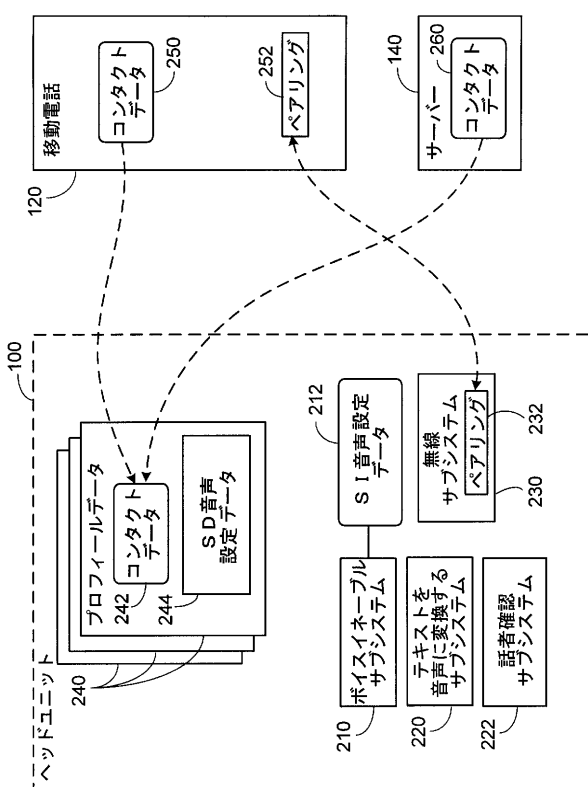
10

20

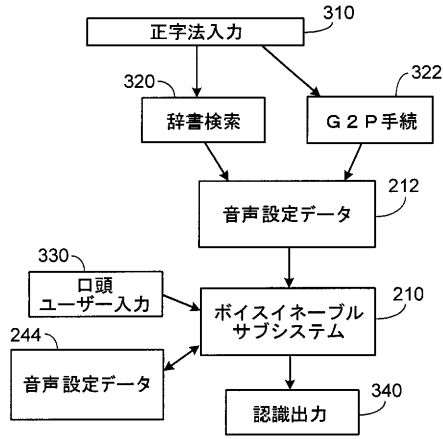
【図1】



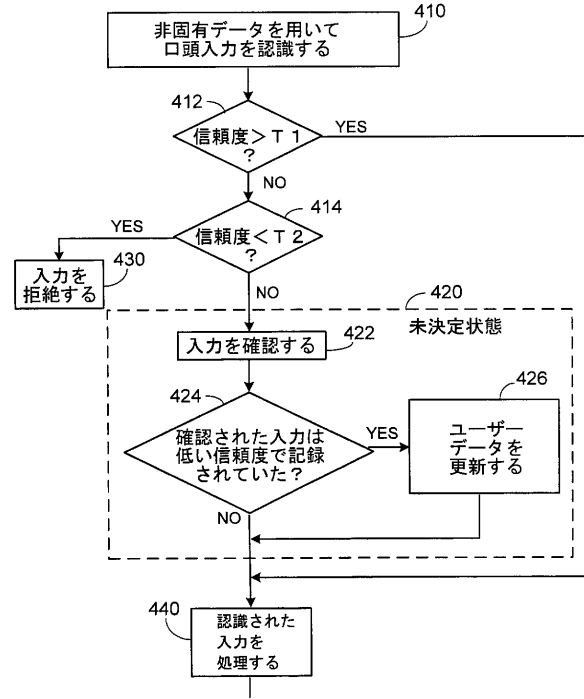
【図2】



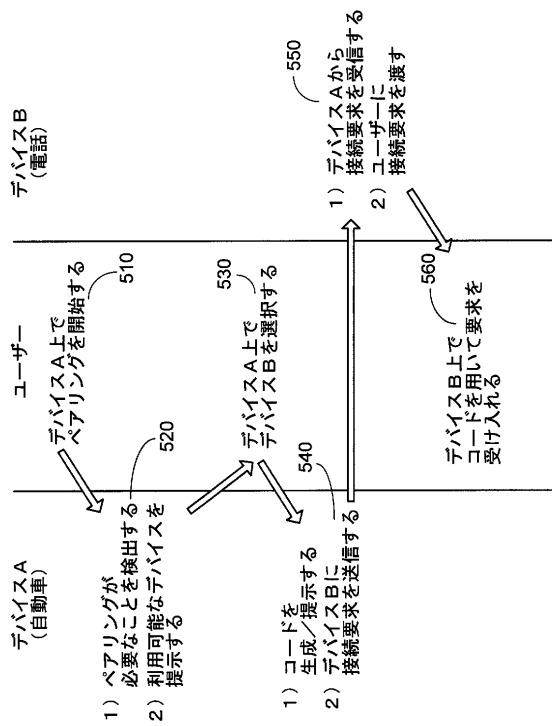
【 図 3 】



【 図 4 】



【 図 5 】



フロントページの続き

- (72)発明者 ジェフリー・ジェイ・ファネフ
アメリカ合衆国・マサチューセッツ・01721・アシュランド・アイヴィー・レーン・6
- (72)発明者 ウィリアム・ヒドゥン
アメリカ合衆国・マサチューセッツ・02038・フランクリン・カーディナル・ドライブ・9
- (72)発明者 ジェイムス・ティー・ホタリー
アメリカ合衆国・マサチューセッツ・01541・プリンストン・オーク・サークル・17
- (72)発明者 スティーヴン・シー・リー
アメリカ合衆国・コネチカット・06905・スタンフォード・ブリッジ・ストリート・287
- (72)発明者 ヴァス・アイエンガー
アメリカ合衆国・マサチューセッツ・01545・シュルーズベリー・コロニアル・ドライブ・80

Fターム(参考) 2F129 AA03 CC03 EE26 EE30 EE43 FF11 FF12 HH12 HH29
3D020 BA03 BA04 BA10 BA11
5D015 KK01