

(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2009-204872
(P2009-204872A)

(43) 公開日 平成21年9月10日(2009.9.10)

(51) Int.Cl.	F 1	テーマコード (参考)
G 1 0 L 15/06 (2006.01)	G 1 0 L 15/06 2 0 0 Z	5 D 0 1 5
G 1 0 L 15/00 (2006.01)	G 1 0 L 15/00 2 0 0 G	

審査請求 未請求 請求項の数 6 O L (全 15 頁)

(21) 出願番号 特願2008-46963 (P2008-46963)
(22) 出願日 平成20年2月28日 (2008. 2. 28)

(71) 出願人 000101732
アルパイン株式会社
東京都品川区西五反田1丁目1番8号
(74) 代理人 100111947
弁理士 木村 良雄
(72) 発明者 大谷 教明
東京都品川区西五反田1丁目1番8号
アルパイン株式会
社内
Fターム(参考) 5D015 GG00 KK01

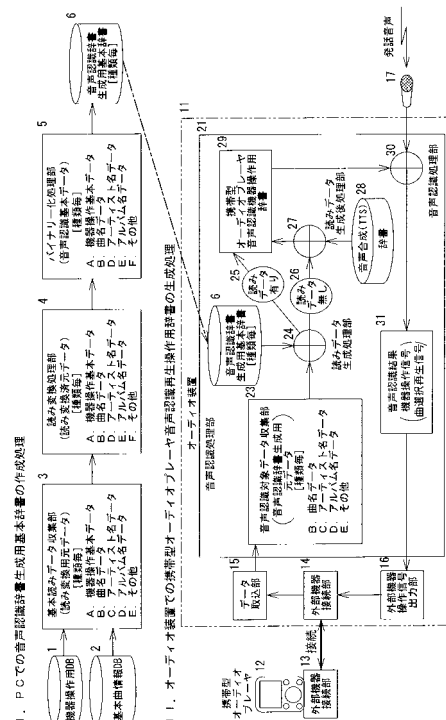
(54) 【発明の名称】 音声認識用辞書生成システム

(57) 【要約】

【課題】短時間で正しい音声認識用辞書を作成し、これを用いて正確に、且つ高速で音声認識を行うことができる「音声認識用辞書生成システム」とする。

【解決手段】、利用者の発話音声を認識して機器操作を行うために用いる音声認識機器操作用辞書を、音声認識対象データ収集部で収集した音声認識辞書生成用元データに読み仮名を付与することにより生成する音声認識用辞書生成システムにおいて、音声認識対象データ収集部では、音声認識辞書生成用元データをデータの種類毎に分けて収集し、別途作成した音声認識辞書生成用基本辞書を用いて、音声認識辞書生成用元データに読み仮名を付与することにより音声認識機器操作用辞書を生成する。音声認識辞書生成用基本辞書は、予め前記音声認識装置で用いると予測される読み変換用元データを収集して読み仮名を付与し、音声認識辞書生成用元データを分ける種類と同じ種類に分けて作成する。

【選択図】 図 1



【特許請求の範囲】**【請求項 1】**

利用者の発話音声認識して機器操作を行うために用いる音声認識機器操作用辞書を、音声認識対象データ収集部で収集した音声認識辞書生成用元データに読み仮名を付与することにより生成する音声認識用辞書生成システムにおいて、

前記音声認識対象データ収集部では、音声認識辞書生成用元データをデータの種類毎に分けて収集し、

別途作成した音声認識辞書生成用基本辞書を用いて、前記音声認識辞書生成用元データに読み仮名を付与することにより前記音声認識機器操作用辞書を生成し、

前記音声認識辞書生成用基本辞書は、予め前記音声認識装置で用いると予測される読み変換用元データを収集して読み仮名を付与し、前記音声認識辞書生成用元データを分ける種類と同じ種類に分けて作成されたものであり、

前記音声認識機器操作用辞書の生成に際しては、前記音声認識対象データの種類に対応した前記音声認識辞書生成用基本辞書内の種類のデータを用いて読み仮名を付与して生成することを特徴とする音声認識用辞書生成システム。

【請求項 2】

前記音声認識により操作する機器はオーディオ装置であり、

前記音声認識対象データ収集部では、オーディオ装置の再生操作に必要なデータをデータの種類毎に収集し、

前記音声認識辞書生成用基本辞書は、前記音声認識対象データ収集部で収集するデータを予測して基本読みデータをデータの種類毎に収集し、読みを付与して作成することを特徴とする請求項 1 記載の音声認識用辞書生成システム。

【請求項 3】

前記音声認識対象データ収集部では、前記オーディオ装置に他のオーディオプレーヤを接続したとき、該オーディオプレーヤのデータ記録媒体に記録されている曲の曲情報を取り込むことによって収集することを特徴とする請求項 2 記載の音声認識用辞書生成システム。

【請求項 4】

前記音声認識辞書生成用基本辞書によって音声認識辞書生成用元データに読み仮名を付与できない単語は、音声合成手段によって読み仮名を付与して音声認識機器操作用辞書を生成することを特徴とする請求項 1 記載の音声認識用辞書生成システム。

【請求項 5】

前記別途作成した音声認識辞書生成用基本辞書は、バイナリー処理することを特徴とする請求項 1 記載の音声認識用辞書生成システム。

【請求項 6】

前記音声認識辞書生成用基本辞書は、前記音声認識対象データ収集部で収集したデータに読まれない記号を除いた処理を行うとき、読み変換用元データに読みを付与するときも同じ処理をして作成することを特徴とする請求項 1 記載の音声認識用辞書生成システム。

【発明の詳細な説明】**【技術分野】****【0001】**

本発明は音声認識に用いる辞書を作成するに際して、例えば英語のような読み仮名情報を持たない言語の音声をできる限り正確に、且つ高速で認識できるようにした、音声認識用辞書生成システムに関する。

【背景技術】**【0002】**

従来より例えば車両用ナビゲーション装置において、運転者が目的地を検索する場合等に、特に車両の運転中にでも入力を可能とするため、或いは通常の目的地の設定に際して手作業を行うことなく容易に入力できるように、運転者等が発声する音声を認識し、入力データを得る音声認識技術が開発され、広く用いられるようになってきている。また、このよ

10

20

30

40

50

うなナビゲーション装置に限らず、車両に搭載した各種機器を運転者が安全に操作ができるように、オーディオ装置を初めエアコン等も、利用者の発声する音声を認識し、各種作動を行うことができるようにしたものも提案されている。

【0003】

前記車両用オーディオ装置においては、近年ハードディスク等に大量のオーディオデータを蓄積したオーディオ機器が用いられるようになっており、それらのオーディオデータの中から所望のアーティストやアルバム、或いは曲を選択して聞くために検索を行い、検索した結果得られるプレイリストに従って再生することが行われる。その際にも運転者でも容易に操作ができるように、音声によって操作するため音声認識装置を用いることも提案されている。

10

【0004】

特に近年は携帯型オーディオプレーヤにおいて、大容量のメモリチップや小型ハードディスク等をデータ記録媒体として内蔵し、MP3等で圧縮した大量のオーディオデータをこのデータ記録媒体に記録して自由に持ち運び、いつでもどこでも自分の好みの曲を聴くことができるようになっている。このような大量のオーディオデータを記録したデータ記録媒体を内蔵した携帯型オーディオプレーヤは、車両にも持ち込まれることが多く、その際には車両に搭載した高性能のオーディオ装置で再生し出力して聞くことが望まれる。そのため車両用オーディオ装置では、これらの携帯型オーディオ装置を接続して内蔵しているオーディオデータを読み出し、これを再生して車両用オーディオ装置から出力する手段を備えるようになっている。

20

【0005】

このように車両用オーディオ装置に携帯型オーディオプレーヤを接続して、携帯型オーディオ装置が内蔵しているデータ記録媒体のオーディオデータを入力し再生して出力する際には、車両用オーディオ装置において携帯型オーディオ装置のデータ記録媒体に記録されているオーディオデータを検索して、任意のアーティストやアルバム、或いは曲を選択してプレイリストを作成し、再生することとなる。このときにも前記と同様に、運転者が容易に曲の選択を行うことができるように、音声によって選択操作を行うことが望まれる。

【0006】

前記のように車両で使用する種々の機器は運転者が操作することが多いため、できる限り運転者が前方の安全の確認を妨げないように操作できるようにすることが求められ、そのために音声によって操作することが望まれる。このような音声による操作に際しては、利用者が発声した音声を認識するため、利用者が発声した音声データと、あらかじめ認識用辞書として登録している音声データとを比較し、最も適合する音声データの言葉を利用者が発声により指示した操作信号として出力することとなる。その際に用いる認識用辞書としては、利用者にあらかじめ音声認識で用いる言葉を発声させ、その音声データを辞書として蓄積することによっても作成することができる。

30

【0007】

しかしながらこの手法は、個人の住所録や電話帳のような、件数が比較的少ないもの場合には使用できるが、例えばハードディスクに録音した曲を音声認識で検索し再生するためにハードディスクに記録されている例えば10,000曲のアーティスト名やアルバム名、或いは曲名をあらかじめ登録しておくことは、登録の手間がかかり過ぎ、不可能に近い。特に車両用オーディオ装置に携帯型オーディオプレーヤを接続するときのように、時々異なったオーディオ記録媒体のオーディオデータを用いるときには、それぞれ収録している曲が異なるため、それらの曲を音声認識で検索するときには実質的に利用することができない。

40

【0008】

音声認識用辞書の作成に際してはそのほか、文字列から音声合成(TTS: text-to-speech)技術を用いて読みデータをあらかじめ作成し、利用者が発声した音声と比較することにより認識を行うことも可能である。即ち、例えばナビゲーション装置

50

において、利用者に案内を行うに際して音声で右左折の案内等を行っているが、その際にはテキストデータを音声に変換する音声合成技術が用いられている。この技術を用いて各曲に記録されているアーティスト名やアルバム名、或いは曲名等の文字列からなるテキストデータを読みデータとし、これを音声データ化して登録し、音声認識辞書を作成することが考えられる。

【0009】

この手法は先の手法と比較して利用者が操作する手間が省ける利点はあるが、これらのデータには読み仮名がふられていないとき、読み文字の付与はTTS任せとなり、利用者が意図しない読みデータが付与されてしまう可能性がある。即ち、例えば「110」が本来「ワンテン(one ten)」という読みであって読み仮名が付与されていないときは、これを「ワンハンドレッドテン(one hundred ten)」と付与するなど、本来の読みが付与されないことがある。このことは特に曲に関する名称には、販売のアピール効果を高めるため特異な読み方にすることが多く、それに対して利用者は単に発音だけで覚えていることが多いため、利用者が発音した音声に対応した適切な曲が検索されないことが多くなる。

10

【0010】

なお、「-」「?」「+」「/」等の読まれない記号を含むデータから適切な音声認識辞書を作成するため、利用者の発声に整合するように発音データを蓄積した音声認識辞書を生成する技術は特開2004-53978号に開示されており、また、言い換え語彙の発生状況を検出して、発声した言い換え語彙を登録して利用することができるようにした技術は特開2007-213005号公報に開示されている。

20

【特許文献1】特開2004-53978号公報

【特許文献1】特開2007-213005号公報

【発明の開示】

【発明が解決しようとする課題】

【0011】

前記のように、特に車両用機器の操作に際しては、利用者の発声した音声を認識して種々の操作を行うことが望まれ、オーディオ装置においても音声認識により各種操作を行うことが望まれるのに対して、ハードディスク等のオーディオ記録媒体にオーディオデータと共に記録したアーティスト名やアルバム名、或いは曲名等の曲情報に基づき、TTSによって音声認識辞書を作成すると、本来の読みのおりには読まれず、別異の音声認識辞書が作成されてしまうため、利用者が発声した曲に関する音声を正しく認識することができない場合が多くなる。

30

【0012】

これらの曲データにオーディオ機器の利用者が予め「読み仮名」を別途入力しておくこともあり、この場合にはその読み仮名のデータを用いることができるが、多くの場合このような「読み仮名」のデータが入力されておらず、その場合には特に前記のような問題を生じる。

【0013】

更に、前記ハードディスクのような大容量のデータ記憶媒体にオーディオデータを記録するときのように、膨大なデータを取り扱うときには、TTSによって音声認識辞書を作成すると長時間かかることとなる。したがって、例えば車両用オーディオ装置に大量のオーディオデータを記録した携帯型オーディオプレーヤを接続し、音声認識により任意の曲を選択して聞こうとするときには、車両用オーディオ装置に携帯型オーディオプレーヤを接続したとき直ちに、自動的に音声認識用辞書を作成する処理を行うような場合には、その処理に多くの時間を要することとなり、音声認識辞書が作成されるまでは音声による機器操作が行われないため、利用性の悪い装置となり、利用者には不快感や不信感を与えることとなる。

40

【0014】

このことは前記のような車両に搭載したオーディオ装置に限らず、例えばナビゲーション

50

ン装置において新しい地図データ、或いは新しい地図の差分のデータをダウンロードして地図データの更新を行うときに提供される、新しい地名のデータを含んでいる際には、これを音声認識により検索を行うときも同様であり、単に地名のテキストデータから音声認識用の辞書を作成するときには本来の読みを付与することができず、利用者が発声する特有の本来の読みに対応することができず、適切な認識を行うことができないという問題を生じる。

【0015】

更に、近年は車両に携帯電話を持ち込むとき、これをナビゲーション装置と接続し、携帯電話を利用してインターネット網に接続し、各種情報の取り込んで表示し、また利用することができるようになっており、更に音声認識機能を用いて携帯電話の操作を行うことも提案されている。その際に携帯電話の電話帳を利用して電話をかけるとき、電話帳に登録されている氏名、社名等について、音声によって検索を行い、電話番号を出力し、電話をかける機能を備えることも提案されている。そのような場合にも、携帯電話が接続されたとき、直ちに電話帳の音声認識による検索が行われることを考慮して音声認識辞書を作成する場合にも、電話帳に存在するテキストデータでは特有の読みがわからないため、作成される音声認識辞書は必ずしも適切な辞書とはなっておらず、したがって適切な音声認識による検索を行異、電話をかけることができないこととなる。

10

【0016】

これらの問題は必ずしも車両用の機器に限らず、利用する装置に蓄積されている読み仮名の付与されていないデータを用いて音声認識用辞書を作成し、それを用いて音声認識を行うときには同様の問題を生じる。

20

【0017】

したがって本発明は、利用する機器に蓄積されている読み仮名の付与されていないデータを用いて音声認識用辞書を作成し、それを用いて音声認識を行って各種の機器操作を行うとき、短時間で正しい音声認識用辞書を作成し、これを用いて正確に、且つ高速で音声認識を行うことができるようにした音声認識用辞書生成システムを提供することを主たる目的とする。

【課題を解決するための手段】

【0018】

本発明に係る音声認識用辞書生成システムは、上記課題を解決するため、利用者の発話音声認識して機器操作を行うために用いる音声認識機器操作用辞書を、音声認識対象データ収集部で収集した音声認識辞書生成用元データに読み仮名を付与することにより生成する音声認識用辞書生成システムにおいて、前記音声認識対象データ収集部では、音声認識辞書生成用元データをデータの種類毎に分けて収集し、別途作成した音声認識辞書生成用基本辞書を用いて、前記音声認識辞書生成用元データに読み仮名を付与することにより前記音声認識機器操作用辞書を生成し、前記音声認識辞書生成用基本辞書は、予め前記音声認識装置で用いると予測される読み変換用元データを収集して読み仮名を付与し、前記音声認識辞書生成用元データを分ける種類と同じ種類に分けて作成されたものであり、前記音声認識機器操作用辞書の生成に際しては、前記音声認識対象データの種類に対応した前記音声認識辞書生成用基本辞書内の種類のデータを用いて読み仮名を付与して生成することを特徴とする。

30

40

【0019】

また、本発明に係る他の音声認識用辞書生成システムは、前記音声認識用辞書生成システムにおいて、前記音声認識により操作する機器はオーディオ装置であり、前記音声認識対象データ収集部では、オーディオ装置の再生操作に必要なデータをデータの種類毎に収集し、前記音声認識辞書生成用基本辞書は、前記音声認識対象データ収集部で収集するデータを予測して基本読みデータをデータの種類毎に収集し、読みを付与して作成することを特徴とする。

【0020】

また、本発明に係る他の音声認識用辞書生成システムは、前記音声認識用辞書生成シス

50

テムにおいて、前記音声認識対象データ収集部では、前記オーディオ装置に他のオーディオプレーヤを接続したとき、該オーディオプレーヤのデータ記録媒体に記録されている曲の曲情報を取り込むことによって収集することを特徴とする。

【0021】

また、本発明に係る他の音声認識用辞書生成システムは、前記音声認識用辞書生成システムにおいて、前記音声認識辞書生成用基本辞書によって音声認識辞書生成用元データに読み仮名を付与できない単語は、音声合成手段によって読み仮名を付与して音声認識機器操作作用辞書を生成することを特徴とする。

【0022】

また、本発明に係る他の音声認識用辞書生成システムは、前記音声認識用辞書生成システムにおいて、前記別途作成した音声認識辞書生成用基本辞書は、バイナリー処理することを特徴とする。

【0023】

また、本発明に係る他の音声認識用辞書生成システムは、前記音声認識用辞書生成システムにおいて、前記音声認識辞書生成用基本辞書は、前記音声認識対象データ収集部で収集したデータに読まれない記号を除いた処理を行うとき、読み変換用元データに読みを付与するときも同じ処理をして作成することを特徴とする。

【発明の効果】

【0024】

本発明は上記のように構成したので、利用する装置に蓄積されている読み仮名の付与されていないデータを用いて音声認識機器操作作用辞書を作成し、それを用いて音声認識を行って各種の機器操作を行うとき、短時間で正しい音声認識用辞書を容易に作成し、これを用いて正確に、且つ高速で音声認識を行うことができるようにした音声認識用辞書生成システムとすることができる。

【発明を実施するための最良の形態】

【0025】

本発明は、利用する装置に蓄積されている読み仮名の付与されていないデータを用いて音声認識用辞書を作成し、それを用いて音声認識を行って各種操作を行うとき、短時間で正しい音声認識用辞書を作成し、これを用いて正確に、且つ高速で音声認識を行うことができるようにするという目的を、利用者の発話音声認識して機器操作を行うために用いる音声認識機器操作作用辞書を、音声認識対象データ収集部で収集した音声認識辞書生成用元データに読み仮名を付与することにより生成する音声認識用辞書生成システムにおいて、前記音声認識対象データ収集部では、音声認識辞書生成用元データをデータの種類毎に分けて収集し、別途作成した音声認識辞書生成用基本辞書を用いて、前記音声認識辞書生成用元データに読み仮名を付与することにより前記音声認識機器操作作用辞書を生成し、前記音声認識辞書生成用基本辞書は、予め前記音声認識装置で用いると予測される読み変換用元データを収集して読み仮名を付与し、前記音声認識辞書生成用元データを分ける種類と同じ種類に分けて作成されたものであり、前記音声認識機器操作作用辞書の生成に際しては、前記音声認識対象データの種類に対応した前記音声認識辞書生成用基本辞書内の種類のデータを用いて読み仮名を付与して生成することにより実現した。

【実施例1】

【0026】

本発明の実施例を図面に沿って説明する。図1は本発明をオーディオ装置に適用した実施例における機能ブロック図であり、本発明は図示するように、第1にオーディオ装置のメーカー等が行う、PCでの音声認識辞書生成用基本辞書6を作成する処理と、第2にオーディオ装置でこの音声認識辞書生成用基本辞書6を用いて、携帯型オーディオプレーヤ音声認識機器操作作用辞書を生成する処理とに大別される。

【0027】

第1のPCでの音声認識辞書生成用基本辞書を作成する処理に際しては、図示の例においては、機器を操作する機器操作作用データベース(DB)1と基本曲情報データベース(

10

20

30

40

50

DB) 2 とを用い、基本読みデータ収集部 3 において読み変換用元データを収集している。この作業は原則として、オーディオ装置のメーカー等が、オーディオ装置の付加価値を高めるサービスとして行う。その収集に際してはデータの種類毎に収集を行い、オーディオ装置で曲の選択再生を行うためには図示するように、機器を操作するための機器操作基本データ、曲名データ、アーティスト名データ、アルバム名データ、その他必要に応じて追加される例えばジャンル名データ等を種類毎に収集する。

【0028】

このような基本読みデータ収集部 3 における機器操作基本データの収集に際しては、機器を操作するために必要とされる単語を予め機器操作データベース 1 として記憶させているときにはこれを用いることができ、このようなデータが存在しないときには、機器操作に必要とされると推定する単語をパソコンを用いて入力することにより行う。

10

【0029】

その外の曲名データ、アーティスト名データ、アルバム名データは、基本曲情報データベース 2 に予め蓄積しているデータを用いる。基本曲情報データベース 2 としては種々のものを用いることができるが、例えば CD の曲情報を収集して公開している CDDB (CD データベース) を用いることができる。このデータには CD の TOC データが含まれているので、CD に記録された曲名、アーティスト名、アルバム名、ジャンル名、発売年月等のデータを容易に収集することができる。

【0030】

図 1 の基本読みデータ収集部 3 における読み変換用元データは、例えば図 2 に示すようなデータである。即ち図 2 に示す読み変換用元データ収集例には、機器操作データベースから機器操作基本データとして、「Play」「Vol」「Vol.」「Artist」「Song」「By」等を抽出した例を示している。ここでは音声認識による操作を行う利用者が、オーディオ装置に対して発声するこれらの音声を認識することができるように、また、後述するようなプログラムによって音声認識辞書を作成する処理を行うことができるように、使用される単語を予め調査し、更に推定して収集する。

20

【0031】

基本曲情報データベースから収集する曲名の読みデータとしては、例えば図 2 においては「Anything for you」「Black & Blue」「Crazy 4 U」「DANCE 2」等を抽出した例を示している。またアーティスト名の読みデータとしては「Bonnie Pink」「Cocco」「hide」「Mr. Children」等を抽出し、アルバム名読みデータとしては「Best-first-things [disc 1]」「Best-first-things [disc 2]」「Crispy!」「ULTRA BULE」等を抽出し、その他のデータとして「AC/DC」等を抽出した例を示している。

30

【0032】

図 1 における読み変換処理部 4 においては、基本読みデータ収集部 3 で収集した前記のように収集した読み変換用元データについて、読み仮名を付与して読み変換を行う処理をし、読み変換済元データを得る。ここでもデータの種類毎に作成するが、その読み仮名の付与に際しては、CDDB で読み仮名情報が付与されている場合はそれを利用することができる、付与されていないときにはこの音声認識辞書生成用基本辞書を作成する部署の人が適切なデータを入力することにより行う。その際に入力するデータは、必ずしも現在提供されている全ての楽曲に対して付与する必要はなく、有名なアーティストの曲で読みが特殊な曲名、アーティスト名、アルバム名等について付与するのみでも本発明を実施することができる。

40

【0033】

この処理の結果、例えば図 3 に示すような読み変換済元データを作成することとなる。即ち図 3 に示す例においては、機器操作基本データにおいて表記文字列が「Play」である文字を「プレイ」と読むものとし、以下同様に「Vol」及び「Vol.」を同じ「ボリューム」、「Artist」を「アーティスト」、「Song」を「ソング」、「B

50

y」を「バイ」と読むようにデータを作成した例を示している。曲名読みデータについては「Anything for you」を「エニシング フォー ユー」、「Black & Blue」を「ブラック アンド ブルー」、「Crazy 4 U」を「クレイジー フォー ユー」、「DANCE 2」を「ダンス ダンス」と読むものとした例を示している。

【0034】

このように、通常「DANCE 2」は「ダンス ツー」と読まれることが多いのに対して、正式には「ダンス ダンス」と読むことを入力しておくことにより、従来のオーディオ装置では困難であった読み仮名が振られることが少ない英語文字についても、正確な読み仮名を予め付与することができ、その後の音声認識処理において利用者が「ダンス ダンス」と発生したとき、正しく「DANCE 2」の曲であることを認識し、直ちにその曲の再生が可能となる。

10

【0035】

またアーティスト名の読みデータとしては「Bonnie Pink」を「ボニーピンク」、「Cocco」を「コッコ」、「hide」を「ヒデ」、「Mr. Children」を「ミスターチルドレン」と読むものとし、特に「Mr. Children」については、「ミスチル」と略称や愛称で呼ばれることが多いことを考慮してこの読みも別途入力している。このような略称や愛称も入力することにより、利用者が音声認識に際して同一のアーティストを種々の態様で発声することに柔軟に対応することができるようになる。

20

【0036】

アルバム名読みデータとしては「Best-first-things [disc 1]」を「ベスト ファースト シングス ディスク ワン」、「Best-first-things [disc 2]」を「ベスト ファースト シングス ディスク ツー」、「Crispy!」を「クリスピー」、「ULTRA BLUE」を「ウルトラブルー」と読むものとした例を示している。更にその他のデータとして「AC/DC」は「エーシーデーシー」と読み仮名を付与した例を示している。

【0037】

バイナリー化処理部5では、読み変換処理部4で作成した読み変換済元データについて、前記の種類毎にバイナリーデータとする処理を行う。ここでバイナリー化するのは、このデータを直接利用することができるようにするためであり、これによりここで作成した音声認識辞書生成用基本辞書をオーディオ装置、或いはこのオーディオ装置と接続したナビゲーション装置において、他の各種処理を行うとき、一般のコンパイル処理がされないようにし、オーディオ装置やこれと接続するナビゲーション装置で読み込み以外の処理が発生しないようにすることができる。なお、バイナリー処理するに際しては、より正確には周知のように、コンパイル処理をしてバイナリー化を行うこととなる。

30

【0038】

このようにしてバイナリー処理化された、前記種類毎の音声認識辞書生成用基本辞書6は、オーディオ装置11として示している車両用オーディオ装置、或いはこれに接続したナビゲーション装置（以下オーディオ装置と略称する）における音声認識処理部21において用いられる。その際には、オーディオ装置等に備えたメモリに入力し、HDD等のデータ記録媒体にダウンロードし、或いはデータを記録したメモリを移動し、更には回路のチップとして供給することができる。

40

【0039】

図1に示すオーディオ装置での携帯型オーディオプレーヤ音声認識機器操作辞書29の作成処理の例においては、車両用等のオーディオ装置11に携帯型オーディオプレーヤ12を、携帯型オーディオプレーヤの外部機器接続部13と、オーディオ装置11の外部機器接続部14とを有線或いは無線で接続しており、それによりオーディオ装置11における外部機器操作信号出力部16の信号によって携帯型オーディオプレーヤ12を再生操作等の操作を行い、またその操作指示に従って、携帯型オーディオプレーヤの任意のデー

50

タをオーディオ装置 1 1 のデータ取込部 1 4 から取り込むことができるようになっている。

【 0 0 4 0 】

図 1 に示すオーディオ装置 1 1 の音声認識処理部 2 1 は、音声認識対象データ収集部 2 3 を備えており、ここではオーディオ装置 1 1 に携帯型オーディオプレーヤを前記のように接続し、オーディオ装置 1 1 と携帯型オーディオプレーヤ間での前記のような通信が可能となったときに、自動的に音声認識対象データ収集部 2 3 が携帯型オーディオプレーヤ 1 2 に内蔵したメモリチップやハードディスク等のデータ記録媒体から、そこに記録しているオーディオデータについて、曲情報を取り込む。この曲情報の中には楽曲自体を記録したオーディオデータの取り込みは必要としない。これらの曲情報は、携帯型オーディオプレーヤのデータ記録媒体に M P 3 の形式でオーディオデータが記録されているときには、そのデータの中の曲情報を記録したタグ部分から抽出して収集することができる。

10

【 0 0 4 1 】

ここで収集する曲情報は、利用者が曲の選択を行うときに指示する、例えば曲名、アーティスト名、アルバム名、更にはジャンル名等の種類に分けて取り込む。これらの種類は全てデータ記録媒体に M P 3 等で記録している曲情報のデータから取り込むことができるが、それらのデータが存在しないものについては、予めオーディオ装置 1 1 に C D D B のデータを備えているときには、そのデータを検索して取り込むこともでき、更にはオーディオ装置 1 1 にインターネット等の通信機能を備えているときには、直接 C D D B データ提供サイトと接続し、データを取り込むこともできる。なお、前記のように P C 上で作業を行う音声認識辞書生成用基本辞書 6 を作成する際に収集する曲情報の種類は、音声認識辞書用データ収集部 2 3 の種類毎に収集する音声認識辞書生成用元データと同じ種類分けとし、両者のミスマッチを防止する。

20

【 0 0 4 2 】

音声認識処理部 2 1 には前記 P C 上で作成した音声認識辞書生成用基本辞書 6 のデータを、オーディオ装置 1 1 の音声認識処理部でアクセスするデータ記録媒体にダウンロードし、或いは予めチップ等の形式で装備し、或いはメモリチップとして挿入する等により、この音声認識辞書生成用基本辞書 6 のデータを利用することができるようになる。なお、このようにして音声認識処理部 2 1 に音声認識辞書生成用基本辞書 6 が存在するとき、その後このデータを更新することができるようにし、このオーディオ装置 1 1 のメーカー等が更新データを提供して、年々多数の曲が作られることに対応することがより好ましい。

30

【 0 0 4 3 】

音声認識処理部 2 1 の読みデータ生成後処理部 2 4 では、音声認識対象データ収集部 2 3 で収集した曲情報について、音声認識辞書生成用基本辞書 6 を用いて読みデータを作成する。その際には音声認識対象データ収集部 2 3 で収集した曲情報の種類が例えばアーティスト名データであるとき、この読みデータを生成するに際して用いる音声認識辞書生成用基本辞書 6 についても同じ種類であるアーティスト名データ部分を検索する。このような処理を行うことにより、少ないデータから高速で、且つ正確な読みデータを付与することができる。

40

【 0 0 4 4 】

この読みデータ生成後処理部 2 4 で、読みデータが音声認識辞書生成用基本辞書 6 にあったときには、読みデータ有り 2 5 の読みデータ付曲情報データとして携帯型オーディオプレーヤ音声認識機器操作作用辞書 2 9 に記録する。また、読みデータ生成後処理部 2 4 で、読みデータが音声認識辞書生成用基本辞書 6 になかったときには、読みデータ無し 2 6 として次に行う後処理部としての、読みデータ生成後処理部 2 7 に出力する。読みデータ生成後処理部 2 7 においては、例えばナビゲーション装置等で広く用いている音声案内のための音声合成 (T T S : t e x t - t o - s p e e c h) 技術における音声合成辞書 2 8、及び音声合成処理技術を用いて、例えばアーティスト名に対応する音声合成データがあるときにはそのデータを直接用い、無いときには通常読みと推定される読み方の読みデータを生成する。このようにして読みデータ生成後処理部 2 7 で生成した読みデータ付曲

50

情報データとして携帯型オーディオプレーヤ音声認識機器操作用辞書 29 に記録する。なお、図 1 の例においてはオーディオ装置 11 の音声認識辞書部 21 において、携帯型オーディオプレーヤ 12 を音声認識により操作する例を示すため「携帯型オーディオプレーヤ音声認識機器操作用辞書 29」として示しているが、各種機器を操作するときには単に「音声認識機器操作用辞書 29」と言い換えることができる。

【0045】

携帯型オーディオプレーヤ音声認識機器操作用辞書 29 においては、前記のように読みデータ生成処理部 24 において音声認識辞書生成用基本辞書 6 を用いて生成した読みデータ付曲情報データと、読みデータ生成後処理部 27 で生成した読みデータ付曲情報データとにより、オーディオ装置 11 に接続した携帯型オーディオプレーヤ 12 が蓄積している曲について、携帯型オーディオプレーヤをオーディオ装置に接続したとき直ちに音声認識用の辞書を生成することができる。

10

【0046】

そのため、その後マイク 17 に対して利用者が、曲を再生するために「プレイ」「アーティスト」「ボニーピンク」と予め定めた順序で発音することにより、音声認識処理部 30 がこれらの音声について携帯型オーディオプレーヤ音声認識機器操作用辞書 29 を順に検索して認識し、「ボニーピンク」の曲を再生する、という音声認識結果 31 を得ることができる。この音声認識結果 31 により、外部機器操作信号出力部 16 では携帯型オーディオプレーヤ 12 に対して「ボニーピンク」の曲を選択して出力する指示の出力を両機器の外部機器接続部を介して行い、出力されたオーディオデータをオーディオ装置 11 が取り込んで再生処理を行う。

20

【0047】

前記のような機能ブロックで構成される本発明の音声認識用辞書生成システムにおいては、例えば図 4 ~ 図 6 に示す作動フローにより順に作動させることによって実施することができる。図 4 には音声認識辞書生成用基本辞書の作成処理の作動フローを示し、この作動は図 1 の PC 上での音声認識辞書生成用基本辞書 6 の作成処理部分で行うものであり、最初音声認識辞書生成用基本辞書作成用の単語の収集を行う（ステップ S1）。次いで収集した単語を、A. 機器操作基本データ、B. 曲名データ、C. アーティスト名データ、D. アルバム名データ、E. その他等の種類毎に分類した読み変換用元データを作成する（ステップ S2）。これらの作動は、図 1 における基本読みデータ収集部 3 において、機器操作の単語を機器操作データベース 1 から、また各種の曲情報を基本曲情報データベースから収集することにより行う。

30

【0048】

その後元データの単語の読みを入力する（ステップ S3）。この処理は図 1 の読み変換処理部 4 において、先に述べた手法により行うことができる。次いで、このようにして得られた読み変換用元データ、及びこれに対して付与された読みデータとをバイナリー化し（ステップ S4）、単語の種類毎に分けた音声認識辞書生成用基本辞書を作成する（ステップ S5）。

【0049】

このようにして得られた音声認識辞書生成用基本辞書を用いて音声認識機器操作用辞書を生成するには、図 5 に示す作動フローによって行うことができる。図 5 に示す音声認識機器操作用辞書生成処理においては、最初に携帯型オーディオプレーヤをオーディオ装置に接続する（ステップ S11）。次いでオーディオ装置で携帯型オーディオプレーヤの曲情報をデータの種類毎に取得する（ステップ S12）。この処理は図 1 のオーディオ装置 11 における音声認識処理部 21 の音声認識対象データ収集部 23 で行う。

40

【0050】

このデータ収集によって音声認識辞書生成用元データの作成がなされ（ステップ S13）、次いで元データの認識用単語を順に選択出力し（ステップ S14）、その際に元データの種類を判別する（ステップ S15）。その後元データの種類に応じた音声認識辞書生成用基本辞書の種類を選択し（ステップ S16）、選択した種類の音声認識辞書生成用基

50

本辞書に読みデータはあるか否かを判別する（ステップS17）。その結果音声認識辞書生成用基本辞書に読みデータがないと判別したときには、音声合成（TTS）用辞書及びその処理技術によって読みデータを生成する。これらの処理は図1において読みデータ生成処理部24において音声認識辞書生成用基本辞書6を用いて、同じデータ種別の部分の読みデータを検索し、ここに読みデータがないと読みデータ生成後処理部27において音声合成（TTS）辞書28、及びその処理技術を用いて読みデータを得ることによって行う。

【0051】

ステップS17において、選択した種類の音声認識辞書生成用基本辞書に読みデータがあると判別したときには、音声認識辞書生成用基本辞書により読みデータを生成し（ステップS19）、その後ステップS18において読みデータの生成が行われた場合と共に、全ての音声認識用元データの読みデータを生成したか否かを判別し（ステップS20）、未だ音声認識用元データにおいて読みデータを生成していないものが存在すると判別したときにはステップS14に戻り、元データの認識用単語を順に選択出力する作動から以下同様の作動を繰り返す。最終的にステップS20において全ての音声認識用元データの読みデータを生成したと判別したときには、このデータを音声認識機器操作作用辞書が完成し、図1の携帯型オーディオプレーヤ音声認識機器操作作用辞書29が完成することとなる（ステップS21）。

【0052】

図5のようにして得られた携帯型オーディオプレーヤ音声認識機器操作作用辞書を用いて行う、音声認識によるオーディオ機器の再生操作は、図6に示す作動フローにより行うことができる。即ち図6に示す音声認識によるオーディオ機器再生操作処理においては、最初に曲再生操作作用音声の発声がなされ（ステップS31）、その後発声した音声を先に生成した音声認識機器操作作用辞書のデータの検索を行い（ステップS32）、この検索によって音声認識処理がなされる（ステップS33）。

【0053】

この検索及び音声認識処理に際しては、前記のように予め携帯型オーディオプレーヤの機器操作を音声認識により行う、音声認識機器操作作用辞書を生成する処理を行っている結果、オーディオ装置に接続した携帯型オーディオプレーヤが蓄積している曲情報は全て音声認識機器操作作用辞書に存在することとなり、利用者が発声する曲の再生に関する音声において、携帯型オーディオプレーヤに存在する曲はほぼ確実に認識することができる。

【0054】

その後曲再生操作の発声は終了したか否かの判別を行い（ステップS34）、例えば3秒間発声が途切れたか否かを検出することによりこの判別を行って、未だ終了していないと判別したとき、即ち続いて音声が発声されたときには再びステップS32に戻って、発声した音声を先に生成した音声認識機器操作作用辞書を検索し、以下同様の作動を繰り返す。ステップS34で前記のように所定時間次の発声がなされないとき、或いは携帯型オーディオプレーヤの再生操作の音声ではないと判別したときのような場合は、曲再生操作の発声が終了したと判別し、認識した言葉により機器を操作し、指示した曲の再生操作を行う（ステップS35）。

【0055】

これらの処理は図1において音声認識処理部30がマイク17から入力した利用者の発話音声を入力し、携帯型オーディオプレーヤ音声認識機器操作作用辞書29を検索することによって音声認識を行い、その処理による音声認識結果31を外部機器操作信号出力部16から携帯型オーディオプレーヤ12に出力し、所定の曲を検索してデータの出力を行い、オーディオ装置11ではこれをデータ取込部15で取り込んで再生処理を行うことによって実行する。

【0056】

前記のような音声認識辞書生成用基本辞書を用いる結果、認識辞書の元データが図7（a）に示すようなプログラムのデータであるとき、図3に示すような読み変換済みデータ

10

20

30

40

50

が得られた場合には、これをバイナリー処理して音声認識辞書生成用基本辞書として作成し、図1のオーディオ装置11における音声認識装置21で用いるとき、図7(b)の太字で示す部分のデータについて、この辞書により読みを付与することができる。

【0057】

図7(a)に示す例では、「Play by Song」の機器操作、即ち曲名を入力することによって作動する、という機器操作指示の元に、その曲名を順に取り込むとき、「Play」、「by」、「Song」の機器操作基本データについては、図3Aの読みデータによって読み仮名を付与することができ、曲名データ部分における「Anything for you」、「Black & Blue」、「Crazy 4 U」、「DANCE 2」については図3Bの読みデータによって読み仮名を付与し、その後音声認識機器操作用辞書とすることができる。

10

【0058】

このような処理が行われる結果、図7(b)に示すように、図中太字で示す機器操作基本データ部分の全て、及び曲名データ部分の図中8つの曲名中太字で示す4つの曲名の読みデータを取得できたことになる。それにより、ここで読みデータを取得できなかった残り4つの曲名についてのみ音声合成(TTS)辞書、及びTTS処理手法によって読みデータを取得する処理を行えば良くなる。その結果、例えば10,000件の曲データを対象に読みデータを付与した結果、全て音声合成(TTS)によって辞書生成処理を行うと作成時間が5分かかった場合、本発明の手法を用いると、読みデータが100%音声認識辞書生成用基本辞書から得られるとき(ヒット率100%)には十数秒で読み付与処理を終了

20

【図面の簡単な説明】

【0059】

【図1】本発明の実施例の機能ブロック図である。

【図2】同実施例における読み変換用元データの収集例を示す図である。

【図3】同収集した読み変換用元データについて読み変換を行った読み変換済み元データの例を示す図である。

【図4】同実施例において音声認識辞書生成用基本辞書を作成する処理を行う作動フロー図である。

30

【図5】同実施例において音声認識機器操作用辞書生成処理を行う作動フロー図である。

【図6】同実施例において音声認識によってオーディオ機器再生操作を行う作動フロー図である。

【図7】同実施例において音声認識辞書生成用基本辞書で読み取られた元データの例を示す図である。

【符号の説明】

【0060】

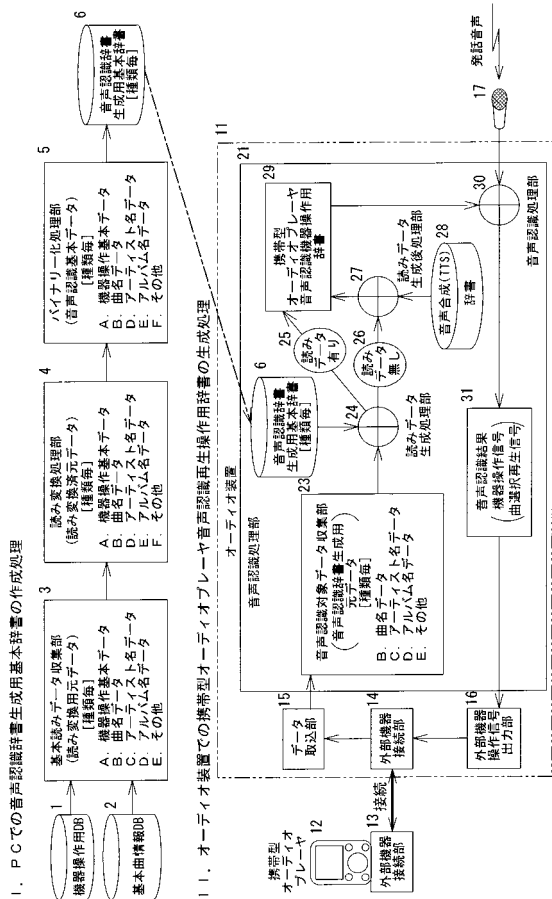
- 1 機器操作データベース
- 2 基本曲情報データベース
- 3 基本読みデータ収集部
- 4 読み変換処理部
- 5 バイナリー化処理部
- 6 音声認識辞書生成用基本辞書
- 11 オーディオ装置
- 12 携帯型オーディオプレーヤ
- 13 外部機器接続部
- 14 外部機器接続部
- 15 データ取込部
- 16 外部機器操作信号出力部
- 17 マイク

40

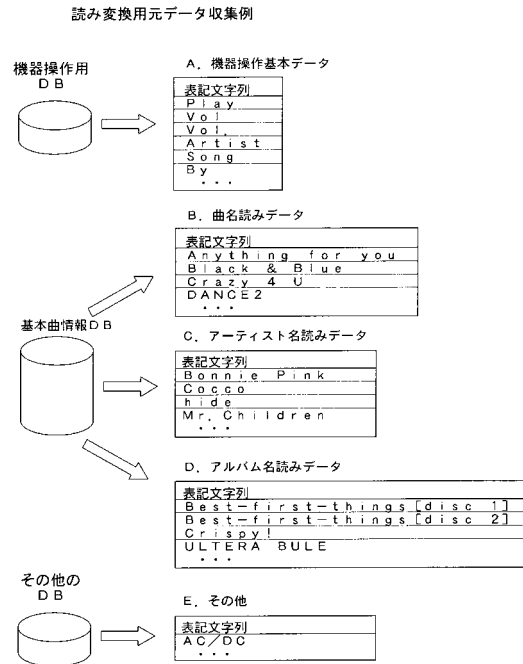
50

- 2 1 音声認識処理部
- 2 3 音声認識対象データ収集部
- 2 4 読みデータ生成処理部
- 2 5 読みデータ有り
- 2 6 読みデータ無し
- 2 7 読みデータ生成後処理部
- 2 8 音声合成(TTS)辞書
- 2 9 帯型オーディオプレーヤ音声認識機器操作作用辞書
- 3 0 音声認識処理部
- 3 1 音声認識結果

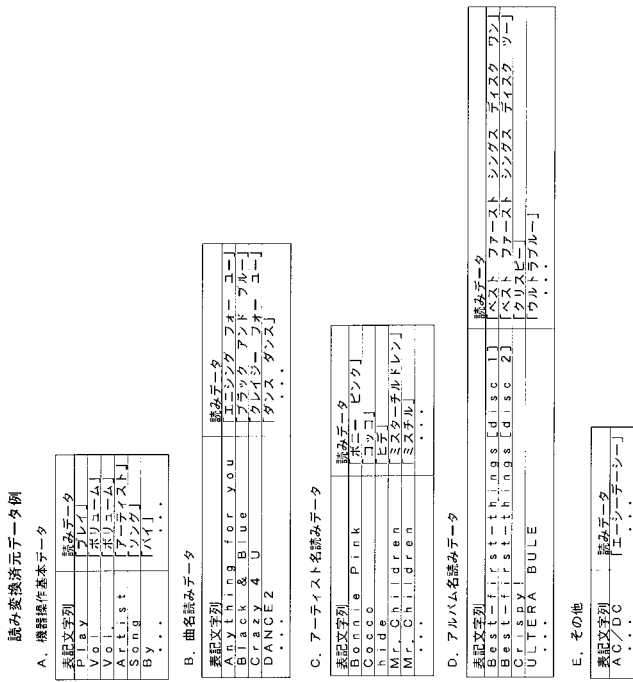
【 図 1 】



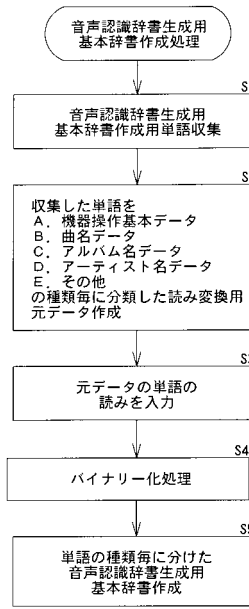
【 図 2 】



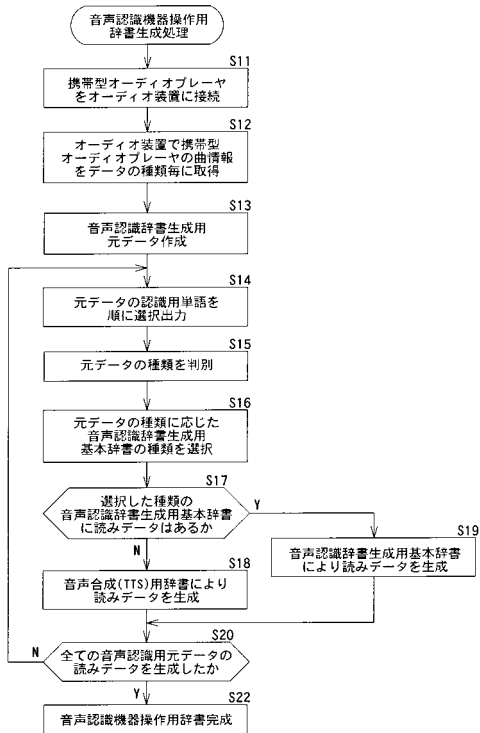
【 図 3 】



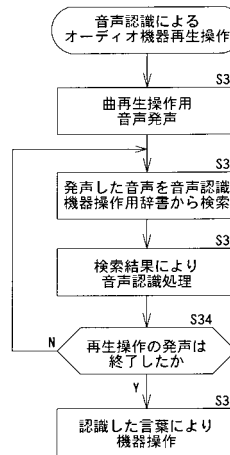
【 図 4 】



【 図 5 】



【 図 6 】



【 図 7 】

(a)

認識辞書の元データ

```
<root>=Play by Song<Song name>  
<Song name>=Anything for you |  
Black & Blue |  
Crazy 4U |  
DANCE2 |  
Thinking out load |  
Yes |  
Orange |  
Range |  
.
```

(b)

音声認識辞書生成用基本辞書で
読み取られた元データ (太字)

```
<root>=PLay by Song<Song name>  
<Song name>=  
"Anything for you" |  
"Black & Blue" |  
"Crazy 4U" |  
"DANCE2" |  
"Thinking out load " |  
"Yes" |  
"Orange" |  
"Range" |  
.
```