

(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2008-21235  
(P2008-21235A)

(43) 公開日 平成20年1月31日(2008.1.31)

(51) Int.Cl.	F I	テーマコード (参考)
<b>G06F 12/00 (2006.01)</b>	G06F 12/00 513D	5B009
<b>G06F 17/21 (2006.01)</b>	G06F 17/21 590E	5B075
<b>G06F 17/22 (2006.01)</b>	G06F 17/22 522K	5B082
<b>G06F 17/30 (2006.01)</b>	G06F 17/22 512V	
	G06F 17/30 170A	

審査請求 未請求 請求項の数 9 O L (全 17 頁) 最終頁に続く

(21) 出願番号 特願2006-194228 (P2006-194228)  
(22) 出願日 平成18年7月14日 (2006.7.14)

(71) 出願人 00004260  
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 100082500  
弁理士 足立 勉  
(72) 発明者 大矢 章博  
愛知県刈谷市昭和町1丁目1番地 株式会  
社デンソー内  
Fターム(参考) 5B009 MF02 MJ03  
5B075 ND03 NK02 NR02 NR05 PP07  
PP12 PP22 PQ02  
5B082 BA09 GA08

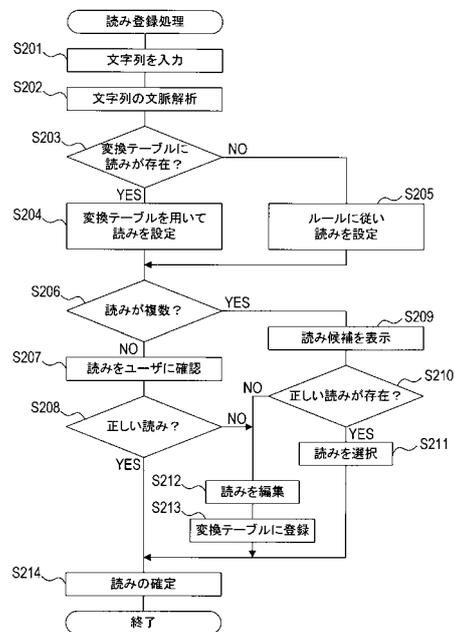
(54) 【発明の名称】 読み登録システム及び読み登録プログラム

(57) 【要約】

【課題】データベースを有効に利用して文字列の読みを設定する。

【解決手段】データベース(C D D B及び地図情報)に記憶されている文字列に含まれる単語とその読みとを対応づけた変換テーブルをあらかじめ生成しておき、読みを登録する対象となる単語をユーザから入力した場合に(S201)、入力された単語の読みを変換テーブルに基づき設定する(S204)。このため、ユーザから入力された単語自体が文字列としてデータベースに登録されていなくても、文字列の一部として登録されていればその単語の読みを抽出することが可能となる。

【選択図】図4



**【特許請求の範囲】****【請求項 1】**

読みを登録する対象となる単語を入力する入力手段と、  
前記入力手段により入力された単語の読みを記憶情報に基づき設定する設定手段と、  
前記設定手段により設定された読みを登録する登録手段と、  
を備え、

前記設定手段は、文字列及びその読みが記憶されたデータベースを元に、前記入力手段により入力された単語が含まれる文字列の読みから、その単語に対応する部分の読みを抽出することにより、前記入力手段により入力された単語の読みを設定すること  
を特徴とする読み登録システム。

10

**【請求項 2】**

前記設定手段は、文字列及びその読みが記憶されたデータベースに基づき、文字列に含まれる単語とその読みとを対応づけた記憶情報をあらかじめ生成し、前記入力手段により入力された単語の読みをその記憶情報に基づき設定すること  
を特徴とする請求項 1 に記載の読み登録システム。

**【請求項 3】**

前記設定手段は、前記入力手段により入力された単語の読みをデータベースに記憶されている情報に基づき設定できなかった場合には、その単語の読みをあらかじめ決められたルールに従い設定すること  
を特徴とする請求項 1 又は請求項 2 に記載の読み登録システム。

20

**【請求項 4】**

前記登録手段は、前記設定手段により設定された読みを登録すべきか否かをユーザからの指示に基づき判定し、登録すべきであると判定した場合にその読みを登録すること  
を特徴とする請求項 1 から請求項 3 までのいずれか 1 項に記載の読み登録システム。

**【請求項 5】**

前記登録手段は、前記設定手段により設定された読みを登録すべきでないと判定した場合には、前記入力手段により入力された単語の読みをユーザから入力してその読みを登録し、

前記設定手段は、ユーザから入力した単語とその読みとを対応づけた記憶情報を生成すること  
を特徴とする請求項 4 に記載の読み登録システム。

30

**【請求項 6】**

前記登録手段は、前記設定手段により複数の読みが設定された場合には、いずれの読みを登録すべきかをユーザからの指示に基づき判定し、登録すべきであると判定された読みを登録すること  
を特徴とする請求項 1 から請求項 5 までのいずれか 1 項に記載の読み登録システム。

**【請求項 7】**

前記設定手段は、文字列及びその読みが記憶された外部のデータベースと通信可能に構成され、その外部のデータベースを元に、前記入力手段により入力された単語の読みを設定すること  
を特徴とする請求項 1 から請求項 6 までのいずれか 1 項に記載の読み登録システム。

40

**【請求項 8】**

前記設定手段は、前記外部のデータベースを元に設定した単語の読みについての課金処理を実行すること  
を特徴とする請求項 7 に記載の読み登録システム。

**【請求項 9】**

請求項 1 から請求項 8 までのいずれか 1 項に記載の読み登録システムが備える前記入力手段、前記設定手段及び前記登録手段としてコンピュータシステムを機能させることを特徴とする読み登録プログラム。

**【発明の詳細な説明】**

50

## 【技術分野】

## 【0001】

本発明は、文字列の読みを設定するための読み登録システム及び読み登録プログラムに関するものである。

## 【背景技術】

## 【0002】

従来、例えば、車両に搭載されて用いられるナビゲーション装置、オーディオ装置、ハンズフリー装置などでは、運転者の運転操作を妨げないように、音声による指示を認識する音声認識機能や、音声による報知を行う音声出力機能が利用されている。このような機能では、文字列とその読みとの対応関係に基づき、入力された音声（読み）に対応する文字列を判断したり、文字列に対応する音声（読み）を出力したりする。また、こうした音声認識機能や音声出力機能以外にも、例えば文字列を五十音で検索したり五十音順に並べたりするような場合にも、文字列と読みとの対応関係に基づく処理が行われる。

10

## 【0003】

ところで、あらかじめ記憶されていない文字列について音声認識機能等を利用する場合には、その文字列の読みを新たに登録する必要がある。この場合、日本語のひらがなやカタカナ、ハングル文字、中国語の Pinyin のように文字列の読みが一義的に決まる表音文字であれば、文字列のみの入力により読みを入力を省略することも可能である。しかしながら、英語（A：「エイ」、「ア」）や漢字（音：「おん」、「おと」）などのように文字列の読みが一義的に決まらないものについては、ユーザに読み（例えば発音記号）を入力させる必要がある。

20

## 【0004】

なお、この種の関連技術として、例えば特許文献1には、リップングした楽曲データ等のコンテンツデータを、音声入力によって自由かつ簡単に指定できるようにした再生装置が開示されている。この再生装置は、楽曲データを再生させる場合に音声入力によって再生対象を指定することができるのであるが、アーティスト名、アルバム名、曲名という3種類の認識辞書と認識語とをそれぞれ比較した検索を行うため、これらいずれのカテゴリからも検索が可能である。そして、アーティスト名、アルバム名、曲名の中で同じ名称のものが存在する場合には、表示部に画面表示している内容の認識カテゴリ名に一意に決定するようにしている。

30

## 【0005】

また、例えば特許文献2には、データベースを適切なタイミングで、適切な内容に、しかも簡単かつ確実に更新できるようにした情報更新システムが開示されている。この情報更新システムにおいて、HD装置に形成される音声認識のための単語用データベースは、マイクロホン、音声認識処理部を通じて入力された音声情報に合致する単語データが無かった場合などにサーバ装置に対して更新要求を通信I/Fを通じて送信する。そして、外部から単語用データベースを更新するための更新データを得て、単語用データベースを更新し、入力された音声情報に合致する単語データが見つからないという状態を防止することにより、音声情報によって適切な処理を行うことができるようにする。

40

## 【0006】

さらに、例えば特許文献3には、個人がCD等より音楽コンテンツを大容量記憶装置に移植して逐次ライブラリを充実させる方式のミュージックサーバ装置が開示されている。このようなミュージックサーバ装置において、ライブラリに新しく受け入れる音楽コンテンツを加えるごとに当該曲のファイル名を入力する作業は、鑑賞ムードを壊し、手間が掛かる好ましくないプロセスである。そこで、ミュージックサーバ装置の大容量記憶装置にタイトル、アーティスト等を記述する「よみ」を索引できるデータベースを設置しておく。そして、新しく受け入れる音楽コンテンツはインターネットを介してTOC等を受け取って「よみ」を提供するデータベース・サービス事業者より「よみ集合」をダウンロードして当該コンテンツのファイル名を「よみ集合」中の例えばカタカナにより自動採番を行って前述データベースに加える。ディスプレイ装置には「よみ集合」を構成するカタカナ

50

、ひらがな、ローマ字、漢字混じ文字等複数の「よみ」のうちから可能で好適なものを表示する。

【特許文献1】特開2005-78705号公報

【特許文献2】特開2004-164466号公報

【特許文献3】特開2004-206648号公報

【発明の開示】

【発明が解決しようとする課題】

【0007】

上述したように、文字列の読みが一義的に決まらないものについては、ユーザに読みを入力させる（定義させる）必要があるが、ユーザにとっては面倒な作業となる。そこで、文字列及びその読みが記憶されたデータベースを利用して、入力された文字列の読みを設定することが考えられる。例えば、ユーザから入力された文字列の読みを、音楽CDに収録された楽曲に関する情報を提供するためのデータベースであるCDDB（CD Data Base）を利用して設定することが可能である。

10

【0008】

しかしながら、このような手法では、データベースに記憶されている文字列そのものが入力されない限り読みが設定されないため、データベースが有効に利用されないという問題がある。

【0009】

本発明は、こうした問題にかんがみてなされたものであり、データベースを有効に利用して文字列の読みを設定することを目的としている。

20

【課題を解決するための手段】

【0010】

上記目的を達成するためになされた本発明の請求項1に記載の読み登録システムでは、入力手段が、読みを登録する対象となる単語を入力し、設定手段が、入力手段により入力された単語の読みを記憶情報に基づき設定する。

【0011】

具体的には、設定手段は、文字列及びその読みが記憶されたデータベースを元に、入力手段により入力された単語が含まれる文字列の読みから、その単語に対応する部分の読みを抽出することにより、入力手段により入力された単語の読みを設定する。

30

【0012】

そして、登録手段が、設定手段により設定された読みを登録する。

つまり、本発明の読み登録システムでは、文字列及びその読みが記憶されたデータベースから、その文字列自体を検索するのではなく、その文字列の一部として含まれている単語を抽出し、その単語の読みを文字列の読みから抽出する。このようにすることで、単語自体が文字列として登録されていなくても、文字列の一部として単語が登録されていれば、その単語の読みを抽出することが可能となる。

【0013】

したがって、本発明の読み登録システムによれば、データベースを有効に利用して文字列の読みを設定することができる。

40

ところで、データベースに記憶されている情報の量が膨大である場合には、文字列に含まれる単語を検索する処理に時間がかかる。

【0014】

そこで、請求項2に記載の読み登録システムでは、設定手段は、文字列及びその読みが記憶されたデータベースに基づき、文字列に含まれる単語とその読みとを対応づけた記憶情報をあらかじめ生成する。そして、入力手段により入力された単語の読みをその記憶情報に基づき設定する。

【0015】

つまり、単語が入力された後に、その単語をデータベースを元に検索するのではなく、データベースに記憶されている文字列に含まれる単語とその読みとを対応づけた記憶情報

50

をあらかじめ生成しておき、その記憶情報に基づき単語を検索するのである。

【0016】

したがって、本発明の読み登録システムによれば、単語が入力されてからその読みを設定するまでに要する時間を短くすることができる。

ただし、入力された単語がデータベースに存在しないことも考えられる。

【0017】

そこで、請求項3に記載の読み登録システムでは、設定手段は、入力手段により入力された単語の読みをデータベースに記憶されている情報に基づき設定できなかった場合には、その単語の読みをあらかじめ決められたルールに従い設定する。ここで、あらかじめ決められたルールとしては、その単語（又は単語を構成する各文字）の一般的な読み方が挙げられる。

10

【0018】

このような構成の読み登録システムによれば、データベースに存在しない単語が入力された場合にも、ユーザによる読みの入力を省略させることが可能となる。

一方、設定された読みがユーザの意図する読みとは異なることも考えられる。

【0019】

そこで、請求項4に記載の読み登録システムでは、登録手段は、設定手段により設定された読みを登録すべきか否かをユーザからの指示に基づき判定し、登録すべきであると判定した場合にその読みを登録する。

【0020】

このような構成の読み登録システムによれば、ユーザの意図しない読みが自動的に登録されてしまうことを防ぐことができる。

20

具体的には、請求項5に記載の読み登録システムでは、登録手段は、設定手段により設定された読みを登録すべきでないとして判定した場合には、入力手段により入力された単語の読みをユーザから入力してその読みを登録する。そして、設定手段が、ユーザから入力した単語とその読みとを対応づけた記憶情報を生成する。

【0021】

このような構成の読み登録システムによれば、ユーザの意図する読みを登録することができる。また、ユーザから入力した単語とその読みとを対応づけた記憶情報を生成するため、以降に同じ単語が入力された場合には、記憶情報に基づきその読みを設定することができる。

30

【0022】

また、ユーザから入力された単語に対し、複数の読みが設定されることも考えられる。

そこで、請求項6に記載の読み登録システムでは、登録手段は、設定手段により複数の読みが設定された場合には、いずれの読みを登録すべきかをユーザからの指示に基づき判定し、登録すべきであると判定された読みを登録する。

【0023】

このような構成の読み登録システムによれば、ユーザの意図する読みを簡単な操作で登録することが可能となる。

ところで、本発明の読み登録システムは、例えば車両に搭載されて用いられる車載装置のように、装置単体として構成することができる。この場合、装置内で閉じた形で処理を行う構成とすることも可能であるが、例えば請求項7のように外部のデータベースを利用して処理を行う構成とすることも可能である。

40

【0024】

すなわち、請求項7に記載の読み登録システムでは、設定手段は、文字列及びその読みが記憶された外部のデータベースと通信可能に構成され、その外部のデータベースを元に、入力手段により入力された単語の読みを設定する。

【0025】

このような構成の読み登録システムによれば、装置内で閉じた形で処理を行う構成に比べ、最新のデータベースを利用しやすくすることができる。この結果、ユーザの意図する

50

読みが設定されやすくすることができる。

【 0 0 2 6 】

具体的には、例えば請求項 8 に記載のように、設定手段が、外部のデータベースを元に設定した単語の読みについての課金処理を実行するようにしてもよい。

このような構成の読み登録システムによれば、外部のデータベースから単語の読みを有料で提供するビジネスモデルを実現することが可能となる。

【 0 0 2 7 】

次に、請求項 9 に記載の読み登録プログラムは、請求項 1 から請求項 8 までのいずれか 1 項に記載の読み登録システムが備える入力手段、設定手段及び登録手段としてコンピュータシステムを機能させるものである。例えば、請求項 1 に記載の読み登録システムが備える入力手段、設定手段及び登録手段としてコンピュータシステムを機能させる読み登録プログラムは、次のような構成となる。

10

【 0 0 2 8 】

読みを登録する対象となる単語をユーザから入力する入力手段、前記入力手段により入力された単語の読みを記憶情報に基づき設定する設定手段、及び、前記設定手段により設定された読みを登録する登録手段、としてコンピュータシステムを機能させる読み登録プログラムであって、前記設定手段は、文字列及びその読みが記憶されたデータベースを元に、前記入力手段により入力された単語が含まれる文字列の読みから、その単語に対応する部分の読みを抽出することにより、前記入力手段により入力された単語の読みを設定することを特徴とする読み登録プログラム。

20

【 0 0 2 9 】

このようなプログラムによれば、コンピュータシステムを用いて前述した読み登録システムを構築することができ、これにより前述した効果を得ることができる。特に、プログラムはネットワーク等を利用して流通させることも可能である上、プログラムの入れ替えは部品の入れ替えに比較して容易であるため、読み登録システムの機能向上等を容易に行うことができる。

【 発明を実施するための最良の形態 】

【 0 0 3 0 】

以下、本発明が適用された実施形態について、図面を用いて説明する。

[ 第 1 実施形態 ]

30

図 1 は、第 1 実施形態の音声認識装置 10 の概略構成を表すブロック図である。

【 0 0 3 1 】

この音声認識装置 10 は、車載装置としてのナビゲーション装置、オーディオ装置、ハンズフリー装置等において音声認識機能及び音声出力機能を実現するためのものであり、制御部 20 と、記憶部 40 とを備えている。

【 0 0 3 2 】

制御部 20 は、CPU、ROM、RAM、I/O 及びこれらの構成を接続するバスラインなどからなるマイクロコンピュータを中心に構成されており、各種処理を実行する。なお、ROM には、後述する処理 ( 図 3 , 4 ) を実行するためのプログラムが記憶されている。

40

【 0 0 3 3 】

具体的には、制御部 20 は、認識エンジン 21、音声認識処理制御部 22、認識結果出力部 23、音声出力部 24、読み生成部 25、認識辞書生成部 26、合成エンジン 27 等として機能する。なお、これら各機能はそれぞれ独立したハードウェアにより実現することも可能であるが、本実施形態では共通のハードウェアにより実現している。

【 0 0 3 4 】

認識エンジン 21 は、音声を入力するためのマイクロホン 51 から入力した音声信号について、音響モデル 31 及び認識辞書群 32 を参照することによりパターンマッチングを行う。

【 0 0 3 5 】

50

音声認識処理制御部 2 2 は、音声認識を開始するためのスイッチ 5 2 が押されている間作動し、音声認識処理を管理する。具体的には、認識エンジン 2 1 において行われたパターンマッチングに従い、マイクロホン 5 1 から入力された音声信号の表す文字列を認識し、その認識した文字列を認識結果出力部 2 3 へ出力する。

【 0 0 3 6 】

認識結果出力部 2 3 は、音声認識処理制御部 2 2 から入力した文字列を、外部機器（例えばオーディオ装置や携帯電話機）6 1、表示装置 6 2 及び音声出力部 2 4 へ、各出力先に応じた形式で出力する。これにより、例えば、マイクロホン 5 1 を介して入力された音声信号の表す文字列に対応する処理を外部機器 6 1 に行わせたり、その文字列に対応する内容を表示装置 6 2 に表示させたり、その文字列に対応する内容をスピーカ 6 3 から出力させたりすることが可能となる。

10

【 0 0 3 7 】

音声出力部 2 4 は、認識結果出力部 2 3 又は合成エンジン 2 7 から入力された情報の表す音をスピーカ 6 3 から出力させる。なお、音声出力部 2 4 は、定型文の音声を出力させる場合には、あらかじめ記憶されている定型文の音データ 3 3 の中から該当する音データを抽出して、その音データの表す音をスピーカ 6 3 から出力させる。一方、あらかじめ音データが記憶されていない文字列の音声を出力させる場合には、その文字列を合成エンジン 2 7 により音データに変換させた後、その音データの表す音をスピーカ 6 3 から出力させる。

【 0 0 3 8 】

読み生成部 2 5 は、入力装置（入力操作の複数のキーを備えた装置）5 3 から入力された文字列について、後述する変換テーブル 4 1 ~ 4 4 を参照して読みを設定する。ここで、読み生成部 2 5 は、入力された文字列に対応する読みが複数存在する場合には、複数の読みを表す情報を表示装置 6 2 へ出力することにより表示させ、ユーザに選択させる。また、入力された文字列に対応する読みが存在しない場合には、読みの入力を促すメッセージを表す情報を表示装置 6 2 へ出力することにより表示させ、ユーザに読みを入力させる。そして、読み生成部 2 5 は、入力された文字列と、その文字列について設定した読みとを、認識辞書生成部 2 6 へ出力する。

20

【 0 0 3 9 】

認識辞書生成部 2 6 は、読み生成部 2 5 から入力した文字列及び読みを、認識辞書群 3 2 に登録する。

30

合成エンジン 2 7 は、読み生成部 2 5 又は音声出力部 2 4 から入力した文字列を音データに変換して音声出力部 2 4 へ出力する。

【 0 0 4 0 】

一方、記憶部 4 0 は、記憶情報の書き換えが可能な不揮発性の記憶媒体（本実施形態ではハードディスク）を備えている。そして、記憶部 4 0 には、ユーザによる読みの入力の手間を省くことができるように、単語と読みとの対応関係を表す変換テーブル 4 1 ~ 4 4 が記憶されている。

【 0 0 4 1 】

具体的には、記憶部 4 0 には、人名等の一般語彙についての変換テーブル 4 1 が記憶されている。このため、この変換テーブル 4 1 に記憶されている文字列が入力された場合には、その読みをユーザにわざわざ入力させることなく、読みを設定することが可能となる。

40

【 0 0 4 2 】

また、記憶部 4 0 には、音楽 CD に収録された楽曲に関する情報を提供するためのデータベースである CDDB 4 5 が記憶されている。CDDB 4 5 には、音楽 CD の情報（アルバム名、曲名、アーティスト名等）として、情報を表す複数の文字列とともに、各文字列の読みが記憶されている。このため、例えば音楽 CD を再生する場合に、その音楽 CD の情報を表示したり音声出力したりするといったことが可能となる。ここで、CDDB 4 5 は、CD-ROM や DVD-ROM 等の可搬型記憶媒体 7 1 から読み取ることにより、

50

音声認識装置 10 の記憶部 40 に記憶される。そして、音声認識装置 10 は、後述するように、可搬型記憶媒体 71 から C D D B を読み取った際に、その C D D B 45 に基づく変換テーブル 42 を生成するように構成されており、生成した変換テーブル 42 は記憶部 40 に記憶される。

【0043】

さらに、記憶部 40 には、地図に関する種々の情報を表すデータベースである地図情報 46 が記憶されている。この地図情報 46 には、地図に関する情報（地名、施設名等）として、情報を表す複数の文字列とともに、各文字列の読みが記憶されている。このため、例えばナビゲーション機能を実現する場合に、地図情報 46 を表示したり音声出力したりするといったことが可能となる。ここで、地図情報 46 も、C D D B 45 と同様、C D - R O M や D V D - R O M 等の可搬型記憶媒体（いわゆる地図ディスク）71 から読み取ることにより、音声認識装置 10 の記憶部 40 に記憶される。そして、音声認識装置 10 は、後述するように、可搬型記憶媒体 71 から地図情報 46 を読み取った際に、その地図情報 46 に基づく変換テーブル 43 を生成するように構成されており、生成した変換テーブル 43 は記憶部 40 に記憶される。

10

【0044】

加えて、記憶部 40 には、過去にユーザにより定義された情報を記憶した変換テーブル 44 が記憶されている。

以上のように、記憶部 40 には、一般語彙についての変換テーブル 41 に加え、C D D B 45 から生成した変換テーブル 42 と、地図情報 46 から生成した変換テーブル 43 と、ユーザにより定義された情報から生成した変換テーブル 44 とが記憶されている。このため、これらのうちのいずれかの変換テーブル 41 ~ 44 に記憶されている単語が入力された場合には、その読みをユーザにわざわざ入力させることなく、読みを設定することが可能となる。

20

【0045】

具体的には、各変換テーブル 41 ~ 44 には、図 2 に示すように、単語とその単語の読みとが対応づけて記憶されている。なお、図 2 は、過去にユーザにより定義された情報についての変換テーブル 44 を表す。

【0046】

次に、音声認識装置 10 の制御部 20 が実行する処理について説明する。

30

まず、C D D B 又は地図情報を記憶した可搬型記憶媒体 71 からデータベースが読み取られて記憶部 40 に記憶された場合に実行されるテーブル作成処理について、図 3 のフローチャートを用いて説明する。

【0047】

このテーブル作成処理が開始されると、まず、S 101 で、読み取ったデータベース（C D D B 45 又は地図情報 46）に含まれている文字列の中から、単語を抽出する。すなわち、C D D B 45 の場合、データベースに含まれている文字列とは、アーティスト名、アルバム名、曲名等であり、こうした文字列の中から単語を抽出する。例えば、「（アーティスト名）のシングルベスト」というアルバム名から、「」の部分を抽出することで、アーティスト名のみを表す単語が抽出されることになる。また、地図情報 46 の場合、データベースに含まれている文字列とは、地名や施設名等であり、こうした文字列の中から単語を抽出する。なお、複数の単語に分割できない文字列については、その文字列自体を単語として抽出すればよい。

40

【0048】

続いて、S 102 では、S 101 で抽出した単語に対応する部分の読みを抽出する。

続いて、S 103 では、S 101 で抽出した単語と S 102 で抽出した読みとを対応させて変換テーブル 42 又は変換テーブル 43 に登録する。

【0049】

続いて、S 104 では、データベース内のすべての単語を抽出したか否かを判定する。

そして、S 104 で、すべての単語を抽出していないと判定した場合には、S 101 へ

50

戻る。

【 0 0 5 0 】

一方、S 1 0 4 で、すべての単語を抽出したと判定した場合には、本テーブル作成処理を終了する。

以上のようなテーブル作成処理を実行することにより、複数の文字列についての読みが記憶されたデータベース（C D D B 4 5 又は地図情報 4 6 ）に基づき、文字に含まれる単語とその読みとを対応づけた変換テーブル 4 2 , 4 3 をあらかじめ生成しておく。

【 0 0 5 1 】

次に、文字列及びその読みを登録するためのユーザによる操作が入力装置 5 3 で行われることにより実行される読み登録処理について、図 4 のフローチャートを用いて説明する。

10

【 0 0 5 2 】

この読み登録処理が開始されると、まず、S 2 0 1 で、読みを登録する対象となる文字列をユーザから入力する。ここで、文字列の入力は、入力装置 5 3 により行われるが、例えば、SDカード等の可搬型の記憶媒体からデータの形で入力（コピー）してもよい。

【 0 0 5 3 】

続いて、S 2 0 2 では、S 2 0 1 で入力した文字列の文脈解析を行い、入力した文字列を単語単位に区切って認識する。

続いて、S 2 0 3 では、入力した文字列（単語）が変換テーブル 4 1 ~ 4 4 に存在するか否かを判定する。

20

【 0 0 5 4 】

そして、S 2 0 3 で、入力した文字列（単語）が変換テーブル 4 1 ~ 4 4 に存在すると判定した場合には、S 2 0 4 へ移行し、入力した文字列（単語）の読みを変換テーブル 4 1 ~ 4 4 に従い設定する。その後、S 2 0 6 へ移行する。

【 0 0 5 5 】

一方、S 2 0 3 で、入力した文字列（単語）が変換テーブル 4 1 ~ 4 4 に存在しないと判定した場合には、S 2 0 5 へ移行し、入力した文字列（単語）の読みをあらかじめ決められているルールに従い設定する。ここで、あらかじめ決められたルールとしては、その単語（又は単語を構成する各文字）の一般的な読み方が挙げられ、例えば、T T S ( T e x t t o s p e e c h ) の技術を利用することができる。その後、S 2 0 6 へ移行する。

30

【 0 0 5 6 】

S 2 0 6 では、設定された読みが複数であるか否かを判定する。

そして、S 2 0 6 で、設定された読みが複数でない（つまり、読みが一つである）と判定した場合には、S 2 0 7 へ移行し、その読みが正しいか否かを確認する内容の表示画面を表示装置 6 2 に表示させる。

【 0 0 5 7 】

続いて、S 2 0 8 では、S 2 0 7 で表示した表示画面に対してユーザにより行われる操作に基づき、設定した読みが正しいか否かを判定する。

そして、S 2 0 8 で、設定した読みが正しいと判定した場合には、そのまま S 2 1 4 へ移行する。

40

【 0 0 5 8 】

また、S 2 0 6 で、設定された読みが複数であると判定した場合には、S 2 0 9 へ移行し、複数の読みを表す内容の表示画面を表示装置 6 2 に表示させる。なお、複数の読みについての音声をスピーカ 6 3 から出力させるようにしてもよい。

【 0 0 5 9 】

続いて、S 2 1 0 では、S 2 0 9 で表示した表示画面に対してユーザにより行われる操作に基づき、正しい読みが存在するか否かを判定する。具体的には、表示画面に表示した複数の読みのうちのいずれかを選択する操作が行われた場合には正しい読みが存在すると判定し、いずれも正しくない旨の操作が行われた場合には正しい読みが存在しないと判定

50

する。

【0060】

そして、S210で、正しい読みが存在すると判定した場合には、S211へ移行し、ユーザにより選択された読みを正しい読みとして選択する。その後、S214へ移行する。

【0061】

一方、S208で、設定した読みが誤っていると判定した場合や、S210で、正しい読みが存在しないと判定した場合には、S212へ移行し、正しい読みを入力させる内容の表示画面を表示装置62に表示させ、その表示画面に対してユーザにより行われる操作に基づき正しい読みを入力する。

【0062】

続いて、S213では、S212で入力した読みを、ユーザにより定義された情報から生成した変換テーブル44に登録する。その後、S214へ移行する。

S214では、読みを確定し、認識辞書群32に登録する。その後、本読み登録処理を終了する。

【0063】

なお、本第1実施形態の音声認識装置10では、テーブル作成処理(図3)におけるS101~S104の処理と、読み登録処理(図4)におけるS203~S205の処理とを実行する制御部20が、本発明の設定手段に相当する。また、読み登録処理におけるS206~S214の処理を実行する制御部20が、本発明の登録手段に相当し、読み登録処理におけるS201, S202の処理を実行する制御部20が、本発明の入力手段に相当する。

【0064】

以上説明したように、本第1実施形態の音声認識装置10では、データベース(CDD B45及び地図情報46)に記憶されている文字列に含まれる単語とその読みとを対応付けた変換テーブル42, 43をあらかじめ生成する(S101~S104)。そして、読みを登録する対象となる単語をユーザから入力した場合に(S201)、入力された単語の読みを変換テーブル41~44に基づき設定する(S204)。

【0065】

したがって、音声認識装置10によれば、ユーザから入力された単語自体が文字列としてデータベースに登録されていなくても、文字列の一部として登録されていればその単語の読みを抽出することが可能となる。このため、既存のデータベースを有効に利用して文字列の読みを設定することができる。また、データベースに基づき変換テーブル42, 43をあらかじめ生成しておくようにしているため、データベースに記憶されている情報の量が膨大であっても、単語が入力されてからその読みを設定するまでに要する時間を短くすることができる。

【0066】

また、入力された単語がデータベースに存在しない場合には、その単語の読みをあらかじめ決められたルールに従い設定するようにしているため(S205)、このような場合にもユーザによる読みの入力を省略させることが可能となる。

【0067】

さらに、設定された読みがユーザの意図する読みであるか否かを確認するようにしているため(S207, S208)、ユーザの意図しない読みが自動的に登録されてしまうことを防ぐことができる。

【0068】

しかも、設定された読みがユーザの意図する読みでない場合には、その読みをユーザから入力して変換テーブル44に登録するようにしているため(S212, S213)、以降に同じ単語が入力された場合にはユーザの意図する読みを設定することができる。

【0069】

また、ユーザから入力された単語に対し、複数の読みが設定された場合には、いずれの

10

20

30

40

50

読みを登録すべきかをユーザからの指示に基づき判定して登録するようにしているため（S 2 0 9 ~ S 2 1 1）、ユーザの意図する読みを簡単な操作で登録することができる。

【 0 0 7 0 】

[ 第 2 実施形態 ]

次に、第 2 実施形態の音声認識装置 1 0 について説明する。

第 2 実施形態の音声認識装置 1 0 は、基本的には第 1 実施形態の音声認識装置 1 0（図 1）と同一であるが、上述した読み登録処理（図 4）に替えて、図 5 に示す読み登録処理を行う点異なる。具体的には、第 2 実施形態の音声認識装置 1 0 は、文字列の読みを有料で提供するサーバ 8 0 と通信可能に構成されており、文字列の読みをサーバ 8 0 へ問い合わせる機能を備えている。すなわち、サーバ 8 0 は、1 又は複数の音声認識装置 1 0 を対象として、文字列の読みを有料で提供するサービスを行うために用いられるものであり、音声認識装置 1 0 とサーバ 8 0 とにより音声認識システムが構築されている。以下、第 1 実施形態との相違点を中心に説明することとし、共通する構成については同一符号を用いて説明を省略する。

10

【 0 0 7 1 】

サーバ 8 0 は、制御部 8 1 と、記憶部 8 2 とを備えている。

制御部 8 1 は、CPU、ROM、RAM、I/O 及びこれらの構成を接続するバスラインなどからなるマイクロコンピュータを中心に構成されており、音声認識装置 1 0 からの要求に応じた応答処理を実行する。なお、ROM には、後述する処理（図 6）を実行するためのプログラムが記憶されている。

20

【 0 0 7 2 】

記憶部 8 2 には、文字列及びその読みが登録された最新のデータベースが記憶されている。また、記憶部 8 2 には、複数の音声認識装置 1 0 のそれぞれに割り当てられた識別情報である識別コードと、各音声認識装置 1 0 から徴収すべき料金を表す課金情報との対応関係を表す課金データベースも記憶されている。なお、課金データベースには、課金情報の他に個人情報（料金徴収の方法等）も記憶されるようにしてもよい。

【 0 0 7 3 】

次に、文字列及びその読みを登録するためのユーザによる操作が入力装置 5 3 で行われることにより音声認識装置 1 0 の制御部 2 0 が実行する読み登録処理について、図 5 のフローチャートを用いて説明する。なお、この読み登録処理（図 5）における S 3 0 1 ~ S 3 0 4, S 3 0 7 ~ S 3 0 9, S 3 1 1 ~ S 3 1 3, S 3 1 5, S 3 1 6, S 3 1 8 の各処理は、第 1 実施形態の読み登録処理（図 4）における S 2 0 1 ~ S 2 0 4, S 2 0 6 ~ S 2 1 4 の各処理と同一の内容である。

30

【 0 0 7 4 】

この読み登録処理が開始されると、まず、S 3 0 1 で、読みを登録する対象となる文字列をユーザから入力する。

続いて、S 3 0 2 では、S 3 0 1 で入力した文字列の文脈解析を行い、入力した文字列を単語単位に区切って認識する。

【 0 0 7 5 】

続いて、S 3 0 3 では、入力した文字列（単語）が変換テーブル 4 1 ~ 4 4 に存在するか否かを判定する。

40

そして、S 3 0 3 で、入力した文字列（単語）が変換テーブル 4 1 ~ 4 4 に存在すると判定した場合には、S 3 0 4 へ移行し、入力した文字列（単語）の読みを変換テーブル 4 1 ~ 4 4 に従い設定する。その後、S 3 0 7 へ移行する。

【 0 0 7 6 】

一方、S 3 0 3 で、入力した文字列（単語）が変換テーブル 4 1 ~ 4 4 に存在しないと判定した場合には、S 3 0 5 へ移行し、入力した文字列及び本音声認識装置 1 0 に固有の識別コードである車両識別コードをサーバ 8 0 へ送信する。

【 0 0 7 7 】

続いて、S 3 0 6 では、S 3 0 5 で送信した情報に対する応答としてサーバ 8 0 から送

50

信されてくる読みを受信し、入力した文字列の読みとして設定する。その後、S 3 0 7へ移行する。

【0078】

S 3 0 7では、設定された読みが複数であるか否かを判定する。

そして、S 3 0 7で、設定された読みが複数でない（つまり、読みが一つである）と判定した場合には、S 3 0 8へ移行し、その読みが正しいか否かを確認する内容の表示画面を表示装置62に表示させる。

【0079】

続いて、S 3 0 9では、S 3 0 8で表示した表示画面に対してユーザにより行われる操作に基づき、設定した読みが正しいか否かを判定する。

そして、S 3 0 9で、設定した読みが正しいと判定した場合には、S 3 1 0へ移行し、課金許可を表す情報をサーバ80へ送信することにより、サーバ80から受信した読みに対する課金をサーバ80に行わせる。その後、S 3 1 8へ移行する。なお、設定した読みがサーバ80から受信したものでない場合（記憶部40に記憶されている変換テーブル41～44に従い設定したものである場合）には、S 3 1 0の処理をスキップしてS 3 1 8へ移行する。

【0080】

また、S 3 0 7で、設定された読みが複数であると判定した場合には、S 3 1 1へ移行し、複数の読みを表す内容の表示画面を表示装置62に表示させる。

続いて、S 3 1 2では、S 3 1 1で表示した表示画面に対してユーザにより行われる操作に基づき、正しい読みが存在するか否かを判定する。

【0081】

そして、S 3 1 2で、正しい読みが存在すると判定した場合には、S 3 1 3へ移行し、ユーザにより選択された読みを正しい読みとして選択する。

続いて、S 3 1 4では、S 3 1 0と同様、課金許可を表す情報をサーバ80へ送信することにより、サーバ80から受信した読みに対する課金をサーバ80に行わせる。その後、S 3 1 8へ移行する。なお、設定した読みがサーバ80から受信したものでない場合には、S 3 1 4の処理をスキップしてS 3 1 8へ移行する。

【0082】

一方、S 3 0 9で、設定した読みが誤っていると判定した場合や、S 3 1 2で、正しい読みが存在しないと判定した場合には、S 3 1 5へ移行し、正しい読みを入力させる内容の表示画面を表示装置62に表示させ、その表示画面に対してユーザにより行われる操作に基づき正しい読みを入力する。

【0083】

続いて、S 3 1 6では、S 3 1 5で入力した読みを、ユーザにより定義された情報から生成した変換テーブル44に登録する。

続いて、S 3 1 7では、課金不可を表す情報をサーバ80へ送信することにより、サーバ80から受信した読みに対する課金をサーバ80に行わせないようにする。その後、S 3 1 8へ移行する。なお、設定した読みがサーバ80から受信したものでない場合には、S 3 1 7の処理をスキップしてS 3 1 8へ移行する。

【0084】

S 3 1 8では、読みを確定し、認識辞書群32に登録する。その後、本読み登録処理を終了する。

次に、音声認識装置10から文字列及び車両識別コードを受信した場合にサーバ80の制御部81が実行する読み提供処理について、図6のフローチャートを用いて説明する。

【0085】

この読み提供処理が開始されると、まずS 4 0 1で、受信した文字列に対応する読みを、データベースに基づき設定する。具体的には、受信した文字列を構成する単語について、その単語を含む文字列をデータベースから検索し、検索した文字列の読みからその単語に対応する部分の読みを抽出する。

10

20

30

40

50

## 【 0 0 8 6 】

続いて、S 4 0 2では、文字列の送信元である音声認識装置 1 0 に対し、S 4 0 1 で設定した読みを送信する。

続いて、S 4 0 3では、S 4 0 2で送信した読みについての課金が許可されたか否かを判定する。具体的には、S 4 0 2で読みを送信した送信先の音声認識装置 1 0 から課金許可を表す情報を受信した場合に、課金が許可されたと判定し、課金不可を表す情報を受信した場合に、課金が許可されなかったと判定する。

## 【 0 0 8 7 】

そして、S 4 0 3で、課金が許可されたと判定した場合には、S 4 0 4へ移行し、課金データベースにおいて、文字列の送信元である音声認識装置 1 0 に対応する課金情報に、文字列を送信した分の料金を課金する。その後、本読み提供処理を終了する。

10

## 【 0 0 8 8 】

一方、S 4 0 3で、課金が許可されなかったと判定した場合には、そのまま本読取提供処理を終了する。

なお、本第 2 実施形態の音声認識装置 1 0 では、テーブル作成処理 ( 図 3 ) における S 1 0 1 ~ S 1 0 4 の処理と、読み登録処理 ( 図 5 ) における S 3 0 3 ~ S 3 0 6 , S 3 1 0 , S 3 1 4 , S 3 1 7 の処理とを実行する制御部 2 0 が、本発明の設定手段に相当する。また、読み登録処理における S 3 0 7 ~ S 3 0 9 , S 3 1 1 ~ S 3 1 3 , S 3 1 5 , S 3 1 6 , S 3 1 8 の処理を実行する制御部 2 0 が、本発明の登録手段に相当し、読み登録処理における S 3 0 1 , S 3 0 2 の処理を実行する制御部 2 0 が、本発明の入力手段に相当する。

20

## 【 0 0 8 9 】

以上説明したように、本第 2 実施形態の音声認識装置 1 0 によれば、第 1 実施形態の音声認識装置 1 0 と同様の効果を得ることができる。

さらに、本第 2 実施形態の音声認識装置 1 0 では、サーバ 8 0 により管理されている最新のデータベースを利用することができるため、ユーザの意図する読みが設定されやすくすることができる。

## 【 0 0 9 0 】

加えて、本第 2 実施形態の音声認識装置 1 0 は、サーバ 8 0 から受信した読みが正しい読みであると判定した場合に、サーバ 8 0 に課金を行わせる ( S 3 1 0 , S 3 1 4 ) 。このような構成により、単語の読みを有料で提供するビジネスモデルを実現することができる。

30

## 【 0 0 9 1 】

なお、上記第 2 実施形態では、ユーザから入力された単語の読みが音声認識装置 1 0 に記憶されていない場合には、必ずサーバ 8 0 に問い合わせるようにしているが、これに限定されるものではない。例えば、サーバ 8 0 に問い合わせるか否かをユーザに確認し、問い合わせる旨の指示を受けた場合にのみサーバ 8 0 に問い合わせるようにしてもよい。このようにすれば、ユーザが課金を望まない場合にはサーバ 8 0 への問い合わせを行わないようにすることができる。

40

## 【 0 0 9 2 】

また、上記第 2 実施形態では、ユーザから入力された単語の読みが音声認識装置 1 0 に記憶されていない場合にサーバ 8 0 に問い合わせるようにしているが、これに限定されるものではない。例えば、ユーザから入力された単語の読みについてはすべてサーバ 8 0 に問い合わせるようにしてもよい。このようにすれば、音声認識装置 1 0 において変換テーブル 4 1 ~ 4 4 等を記憶しておく必要がなくなる。

## 【 0 0 9 3 】

さらに、上記第 2 実施形態では、サーバ 8 0 から受信した読みが正しい読みであると判定した場合に、サーバ 8 0 に課金を行わせるようにしているが、これに限定されるものではない。例えば、サーバ 8 0 から受信した読みが正しい読みであるか否かに関係なく、サーバ 8 0 に課金を行わせるようにしてもよい。

50

## 【 0 0 9 4 】

以上、本発明の一実施形態について説明したが、本発明は、種々の形態を採り得ることは言うまでもない。

例えば、上記各実施形態の音声認識装置 10 では、データベースに基づく変換テーブル 42, 43 をあらかじめ生成するようにしているが、これに限定されるものではなく、ユーザから入力された単語をデータベースから直接検索するようにしてもよい。このようにすれば、データベースに基づく変換テーブル 42, 43 を生成する必要がない。

## 【 0 0 9 5 】

また、上記各実施形態の音声認識装置 10 では、データベースとして C D D B 及び地図情報を利用した構成を例に挙げて説明したが、これに限定されるものではなく、例えば、電話帳のデータベースを利用した構成としてもよい。

10

## 【 0 0 9 6 】

さらに、上記各実施形態では、音声認識機能及び音声出力機能を実現するための音声認識装置 10 を例に挙げて説明したが、本発明は、音声認識や音声出力以外にも利用することができる。例えば、単語を五十音で検索したり五十音順に並べたりする処理において単語の読みを判断する場合に利用することができる。

## 【 図面の簡単な説明 】

## 【 0 0 9 7 】

【 図 1 】 第 1 実施形態の音声認識装置の概略構成を表すブロック図である。

【 図 2 】 変換テーブルの説明図である。

20

【 図 3 】 第 1 実施形態の音声認識装置が実行するテーブル作成処理のフローチャートである。

【 図 4 】 第 1 実施形態の音声認識装置が実行する読み登録処理のフローチャートである。

【 図 5 】 第 2 実施形態の音声認識装置が実行する読み登録処理のフローチャートである。

【 図 6 】 第 2 実施形態のサーバが実行する読み提供処理のフローチャートである。

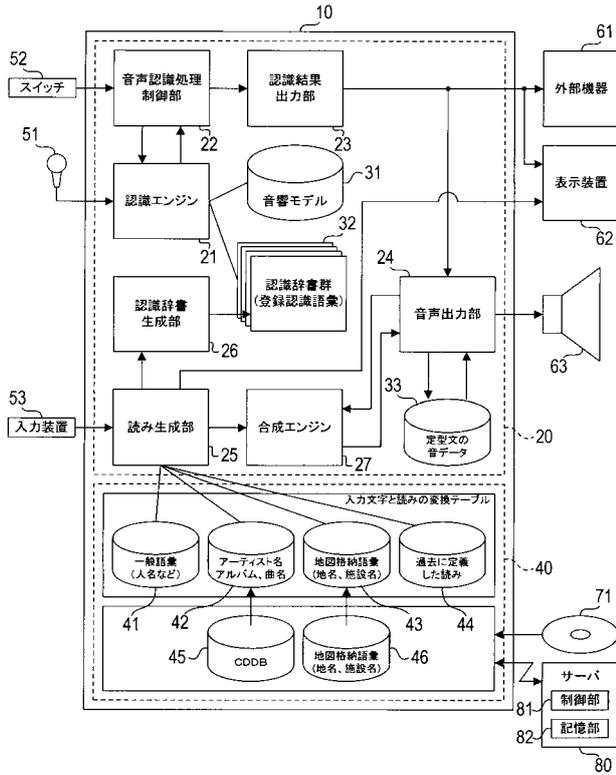
## 【 符号の説明 】

## 【 0 0 9 8 】

10 ... 音声認識装置、 20 ... 制御部、 21 ... 認識エンジン、 22 ... 音声認識処理制御部、 23 ... 認識結果出力部、 24 ... 音声出力部、 25 ... 読み生成部、 26 ... 認識辞書生成部、 27 ... 合成エンジン、 31 ... 音響モデル、 32 ... 認識辞書群、 33 ... 音データ、 40 ... 記憶部、 41 ~ 44 ... 変換テーブル、 45 ... C D D B、 46 ... 地図情報、 51 ... マイクロホン、 52 ... スイッチ、 53 ... 入力装置、 61 ... 外部機器、 62 ... 表示装置、 63 ... スピーカ、 71 ... 可搬型記憶媒体、 72 ... サーバ

30

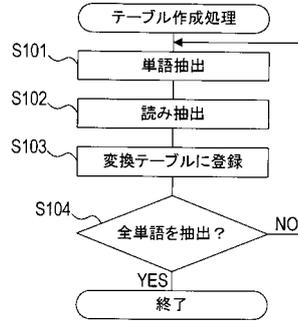
【 図 1 】



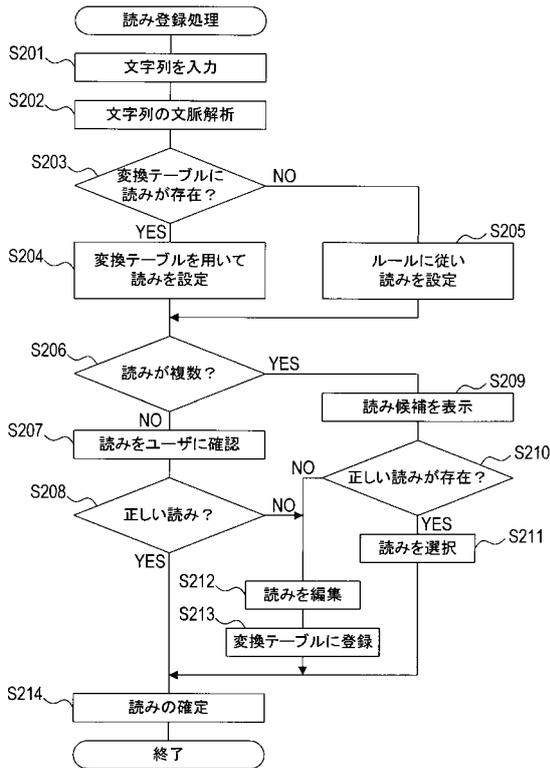
【 図 2 】

単語	読み
DN	デンソー
⋮	⋮
⋮	⋮

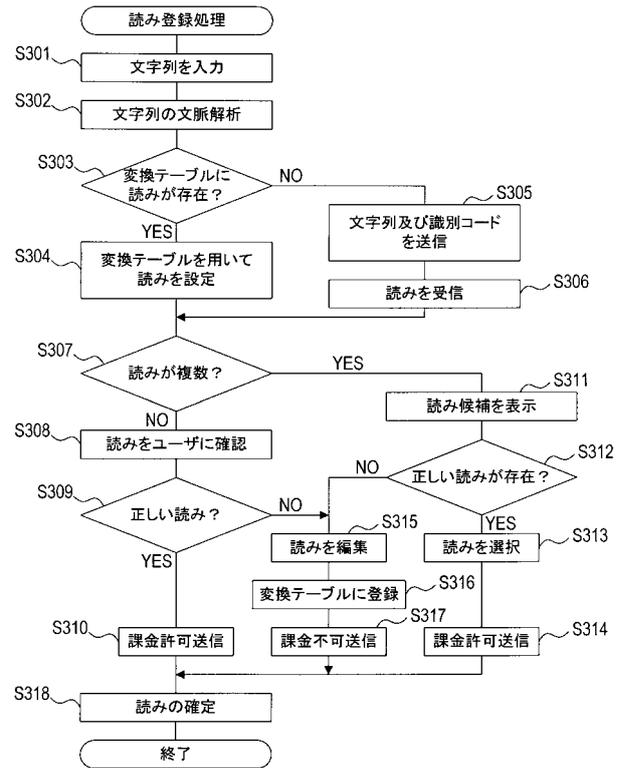
【 図 3 】



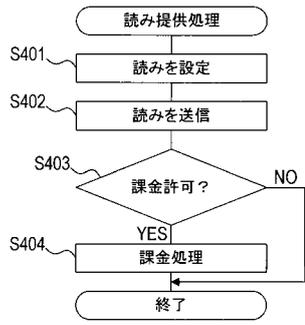
【 図 4 】



【 図 5 】



【 図 6 】



フロントページの続き

(51)Int.Cl.

F I

テーマコード(参考)

G 0 6 F 17/30 2 3 0 Z