(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第3983313号
(P3983313)
(45)発行日　平成19年9月26日(2007.9.26)　(24)登録日　平成19年7月13日(2007.7.13)

(51)Int.Cl.　F I
G10L 13/08 (2006.01)　G10L 13/08　110Z

請求項の数 7　(全 19 頁)

(21)出願番号　特願平8-10399
(22)出願日　平成8年1月24日(1996.1.24)
(65)公開番号　特開平9-204434
(43)公開日　平成9年8月5日(1997.8.5)
審査請求日　平成13年12月27日(2001.12.27)
審判番号　不服2005-499(P2005-499/J1)
審判請求日　平成17年1月6日(2005.1.6)

(73)特許権者　000005223
富士通株式会社
神奈川県川崎市中原区上小田中４丁目１番１号
(74)代理人　100078868
弁理士　河野　登夫
(72)発明者　寺本　良明
神奈川県川崎市中原区上小田中１０１５番地　富士通株式会社内
(72)発明者　片江　伸之
神奈川県川崎市中原区上小田中１０１５番地　富士通株式会社内
(72)発明者　辻内　秀敏
神奈川県川崎市中原区上小田中１０１５番地　富士通株式会社内
最終頁に続く

(54)【発明の名称】音声合成装置及び音声合成方法

(57)【特許請求の範囲】
【請求項１】
文章情報から音声を合成する音声合成装置において、
文章情報の入力手段と、
先に出現する住所を表わす単語によって該単語の次に出現する住所を表わす単語の読みが予め決められている住所を表わす単語群からなる単語列の各単語の表記及び前記読みの情報が、単語を階層化した階層構造の情報とともに格納されている階層辞書と、
入力された文章情報に含まれている文字列の中から住所の単語列に一致する単語列の読みの候補を階層辞書から検索する階層辞書検索手段と、
該読みの候補の中から、文章情報に含まれている前記単語列の読みを前記階層構造の情報に基づいて選択し、前記単語列を該読みに変換する文章解析手段と、
読みの情報から音声波形を生成する音声波形生成手段と、
生成した音声波形に係る音声の出力手段と
を備えたことを特徴とする音声合成装置。
【請求項２】
地域を指定する手段を備え、階層辞書検索手段が、指定された地域に属する階層から検索を開始する手段である請求項１記載の音声合成装置。
【請求項３】
前記階層構造の情報が、その読みを該読みに決めるべき各単語の上位の階層の親単語を特定する情報であり、階層辞書から検索した読みの候補の該情報を参照して文章情報に含

まれる文字列の中から住所の単語列に一致する単語列の読みの候補の接続関係を設定する接続関係設定手段を備えた請求項１又は２記載の音声合成装置。

【請求項４】

前記階層辞書検索手段は、前記単語列のいずれかの階層の単語の表記が省略されている単語列を基に階層辞書を検索する手段を有する請求項１乃至３のいずれかに記載の音声合成装置。

【請求項５】

所定数以上の文字又は単語の表記が階層辞書に格納されている単語列に含まれる該文字又は単語の表記と一致する場合に階層辞書の読みを該文字又は単語の読みと判定する読み判定手段を備えた請求項１乃至４のいずれかに記載の音声合成装置。

【請求項６】

前記単語列に接続される接尾語の表記及び読みの情報が格納されている接尾語辞書と、前記単語列の直後の表記に一致する表記の情報を接尾語辞書から検索し、前記直後の表記に一致する接尾語辞書の表記の読みを前記直後の表記の読みとして選択する接尾語辞書検索手段とを備えた請求項１乃至５のいずれかに記載の音声合成装置。

【請求項７】

先に出現する住所を表わす単語によって該単語の次に出現する住所を表わす単語の読みが予め決められている住所を表わす単語群からなる単語列の各単語の表記及び前記読みの情報が、単語を接続順に従って階層化した階層構造の情報とともに格納されている階層辞書を参照して文章情報から音声を合成する音声合成方法であって、

文章情報を入力し、

入力された文章情報に含まれている文字列の中から住所の単語列に一致する単語列の読みの候補を階層辞書から検索し、

該読みの候補の中から、文章情報に含まれている前記単語列の読みを前記階層構造の情報に基づいて選択し、

前記単語列を該読みに変換し、

読みの情報から音声波形を生成し、

生成した音声波形に係る音声を出力すること

を特徴とする音声合成方法。

【発明の詳細な説明】

【０００１】

【発明の属する技術分野】

本発明は、一つの表記に複数の読みが存在する住所の文章情報を正確な読みで読み上げる音声合成装置及び音声合成方法に関する。

【０００２】

【従来の技術】

従来の音声合成装置では、文章情報が入力されると、単語の表記に対応付けて、単語の読み、音声を合成するためのアクセント等の情報が格納されている辞書を参照して文章情報を読みに変換し、読みの情報及びアクセント等の情報から音声波形を生成して文章情報を合成音声で読み上げる。

【０００３】

【発明が解決しようとする課題】

しかし、一つの表記に複数の読みが存在する住所等が文章情報に含まれている場合、例えば「本町」という表記の文章情報を「ホ＇ンチョー」「ホ＇ンマチ」「モト＇マチ」（「＇」はアクセント記号）といった複数の読みのいずれの読みで読み上げるべきかを正確に決定できずに誤った読みで読み上げてしまう可能性がある。

【０００４】

本発明はこのような問題点を解決するためになされたものであって、一つの表記に対して複数の読みが存在する住所の表記及び読みの情報を階層化した階層構造の辞書を持つことにより、一つの表記に対して複数の読みが存在する住所の文章情報を正確な読みで読み

上げる音声合成装置及び音声合成方法の提供を目的とする。
【０００５】
【課題を解決するための手段】
図１は本発明の音声合成装置の基本ブロック図である。
ＣＤ－ＲＯＭ，光磁気ディスク等の記録媒体から直接的に、又は公衆回線等を介して文章情報を入力する文章入力部１が文章情報を入力すると、階層辞書検索部２は、先に出現する住所を表わす単語によって該単語の次に出現する住所を表わす単語の読みが予め決められている住所を表わす単語群からなる単語列の各単語の表記及び読みの情報が、単語を例えば都道府県・市郡・町村区の順に階層化した階層構造の情報とともに格納されている階層辞書３から、入力された文章情報に含まれている文字列の中から住所の単語列に一致する単語列の読みの候補を検索し、文章解析部４が、読みの候補の中から、文章情報に含まれている、階層辞書３中の単語列の中から住所の単語列に一致する単語列の読みを階層構造の情報に基づいて選択し、この単語列を階層辞書３中の読みに変換する文章解析部４と、読みの情報から音声波形を生成する音声波形生成部６と、生成した音声波形に係る音声を出力するスピーカ７とを備える。
これにより、一つの表記に複数の読みが存在する住所を表わす単語列を正しい読みで読み上げる。
【０００６】
また、本発明の音声合成装置は、地域を特定する情報の入力、文章情報に含まれる地名からの判定等に基づいて地域を指定することにより、指定された地域に属する階層から検索を開始して検索時間を短縮する。
【０００７】
また、本発明の音声合成装置は、階層構造の情報が、その読みを決定する各単語の上位の階層の親単語を特定する情報である階層辞書と、階層辞書から検索した読みの候補のこの情報を参照して文章情報に含まれる文字列の中から住所の単語列に一致する単語列の読みの候補の接続関係を設定する。
【０００８】
また、本発明の音声合成装置は、単語列のいずれかの階層の単語の表記が省略されている単語列を基に階層辞書を検索することにより、文書中の単語列の一部が省略されている場合でも正しい読みで読み上げる。
【０００９】
また、本発明の音声合成装置は、所定数以上の文字又は単語の表記が階層辞書に格納されている単語列に含まれる該文字又は単語の表記と一致する場合に階層辞書の読みを該文字又は単語の読みと判定することにより、複数の読みが存在する住所の単語列以外で、その表記がこの単語列の一部と一致する一般の単語を住所の読みで誤って読み上げることがない。
【００１０】
また、本発明の音声合成装置は、単語列に接続される接尾語の表記及び読みの情報が格納されている接尾語辞書を設け、階層辞書中の単語列と一致する単語列の直後の文章情報の表記と一致する表記の情報を接尾語辞書から検索し、直後の表記に一致する接尾語辞書の表記の読みを、直後の表記の読みとして選択することにより、この単語列に接続されることによって一般の単語の読みと異なる読みになる接尾語を正しい読みで読み上げる。
【００１１】
【発明の実施の形態】
図２は本発明の音声合成装置の一例を示す模式図である。
この例の音声合成装置は、音声出力機能を有する汎用のパーソナルコンピュータのディスクドライブに、以下に述べるような、住所辞書，住所接尾語辞書等の辞書及び音声合成方法のコンピュータプログラムが記録されている光磁気ディスク，ＣＤ－ＲＯＭ等の記録媒体Ｄを装填してコンピュータプログラムをローディングし、文章データベースのディスクから又は公衆回線（図示省略）を介して入力された文章情報から音声を合成し、コンピュ

ータ本体に接続されているスピーカから、又は公衆回線を介して合成音声を出力する構成である。

【００１２】

なお、本発明の音声合成装置は上述のように汎用のパーソナルコンピュータにソフトウェアをローディングする構成以外に、ＦＭ多重放送から受信した交通情報等の文字列情報を合成音声で読み上げるような音声合成専用機であってもよく、その場合、文章情報はアンテナを介して入力される。

【００１３】

〔実施の形態１〕

図３は本発明の音声合成装置の実施の形態１の構成図である。

文章入力部101 はＣＤ－ＲＯＭ，光磁気ディスク等から文章情報を入力し、住所辞書検索部102 は、文章情報に含まれる住所の単語列に一致する単語列を、住所の単語列の表記及び読みの情報が階層化され、この階層構造の情報とともに階層構造の情報が格納されている階層構造の住所辞書103 から検索する。

文章解析部104 は、住所辞書103 から検索された単語の読みを住所辞書103 の読みとし、それ以外の一般の単語に一致する単語を基本辞書105 から検索して文章情報を読みに変換し、音声波形生成部106 は読み及びアクセント情報等から音声波形を生成し、スピーカ107 から文章情報を読み上げる合成音声が出力される。

【００１４】

図４は、住所辞書103 の一例の概念図である。

住所辞書103 には、住所の単語列を構成する各単語の表記及び読みの情報が階層化されて格納されている。階層の先頭は４７都道府県名であり、それ以下の階層として、市区町又は郡の後に町又は村の地名、さらに大字、小字などの地名がその接続順に階層化されている。各々の単語に関してはこのような階層情報の他に、検索キーとなる漢字表記、アクセント句境界情報、アクセント型等の韻律情報を含んだ発音情報、即ち読みの情報を持っている。

【００１５】

図５は実施の形態１における住所辞書103 の基本的な検索アルゴリズムのフローチャートである。

まず、入力された文章情報を格納するテキストバッファの先頭にテキストポインタを設定し（Ｓ101 ）、住所辞書検索ポインタを住所辞書103 の階層構造の先頭に設定する（Ｓ102 ）。テキストポインタを一文字ずつずらしながらそのテキストポインタ位置で始まる単語を住所辞書103 中の検索候補の単語と比較して一致するか否かを判定し（Ｓ103 ，104 ）、住所辞書103 中の単語がテキストバッファ内に存在するか否かを調べていく。住所辞書103 に、テキストポインタで始まる単語が存在している場合は、住所辞書検索ポインタを次の階層に設定するとともに（Ｓ105 ）、テキストポインタを次の単語位置に設定し（Ｓ106 ）、テキストの次の単語位置に次の階層の単語が存在するかどうかを調べていく（Ｓ107 ，Ｓ108 ）。

【００１６】

住所辞書103 中の単語に一致する単語が存在しなくなったら、住所辞書103 中の単語に一致した区間を住所区間とみなし、一致した単語列に設定されている読みの列を、その住所区間の読みとして発音情報を設定し（Ｓ109 ）、文章解析部104 にわたす。また、住所区間以降の文章にも住所が含まれている可能性があるため、住所辞書検索ポインタを階層構造の先頭に設定し（Ｓ110 ）、同様の処理をテキスト情報の最後まで文章全体に対して行う（Ｓ111 ，Ｓ112 ）。

【００１７】

〔実施の形態２〕

図６は実施の形態２の構成図である。なお、上述の実施の形態１と同一部分には同一符号を付してその説明を省略する。実施の形態２では、文章解析部104 で使用する基本辞書105 中には単語が存在しないが、住所辞書103 の検索によって発音情報が既知である場合に

、単語登録しないで、入力文章中に発音情報を発音指定文字列としてテキストに埋め込む住所発音設定部108 と、テキストに埋め込まれた発音指定文字列を識別して発音指定と解析する発音指定解析部109 とが設けられている。

【００１８】

ここで、発音指定文字列のフォーマットを「〈発音：漢字表記：発音情報〉」と定義した場合、「〈：〉」は、文章中で特殊な意味を持たせるための記号であり、「発音」という文字列は発音指定を識別するためのキーワードである。「漢字表記」は単なるコメントの役割で、「発音情報」にはカタカナ及びアクセント記号で表現された発音情報を記述する。

【００１９】

次に、動作について説明する。

文章入力部101 から、例えば「東京都大田区に住んでいます。」という文章が入力されると、住所辞書検索部102 は、実施の形態１と同様に住所辞書103 を検索し、「東京都大田区」の区間が住所単語列であり、その住所の発音が「トーキョ＇ート＿オータ＇ク」（「＇」はアクセント記号、「＿」はアクセント区境界記号）であることを判定する。さらに、住所発音設定部108 では、入力された文章の住所の区間を前述の発音指定文字列に置換し、「〈発音：東京都大田区：トーキョ＇ート＿オータ＇ク〉に住んでいます。」という文字列を発音指定解析部109 に出力する。

【００２０】

発音指定解析部109 は、括弧記号（〈　〉）で区切られた文字を識別して発音指定として解析し、その部分には、発音が「トーキョ＇ート＿オータ＇ク」である名詞が存在しているとする一方、その他の部分の文章はそのまま文章解析部104 にわたす。文章解析部104 は発音指定の情報とその他の文章とを解析して、正しい読み情報を設定する。

【００２１】

実施の形態２では、住所辞書検索部102 に住所発音設定部108 を接続し、住所発音設定部108 と文章解析部104 との間に、住所発音設定部108 からの入力経路の他に、文章入力部101 からの文章情報の入力経路を有する発音指定解析部109 を接続することにより、住所を含まない文章を読み上げる場合に、基本辞書105 を参照して発音情報に変換すべく文章情報を文章解析部104 に直接的に入力できる。即ち、図中、破線で囲んだ住所読み上げ部を独立した装置として構成したり、また文章情報の入力経路を選択的に使用することができる。

【００２２】

従って、第１に、住所読み上げ部を独立の装置として構成した場合には並列処理が可能になる。住所辞書103 の単語数は一般的に十万単語を超え、言語処理部の基本辞書105 の単語数は数万単語から十万単語を超える場合もあるので、検索処理の負荷が大きいが、このような構成にすれば、２個のＣＰＵで住所読み上げ部分の処理とその他の部分の音声合成のための言語処理とを並列処理することが可能になるため、処理時間の増加を防ぐことが可能である。

また、第２に、住所読み上げ部をソフトウェアで構成した場合には、住所読み部と言語処理部とのコマンドとして独立のコマンドを作成することができるので、ソフトウェアの保守作業、システム変更への対応が容易である。

【００２３】

〔実施の形態３〕

図７(a) は実施の形態３の構成図である。なお、上述の実施の形態と同一部分には同一符号を付してその説明を省略する。

この実施の形態では、文章情報の入力の都度、地域名を設定する指定地域入力部110 と、指定された地域を階層構造の住所辞書103 の検索開始地点として保持しておく検索開始位置格納バッファ111 とが設けられており、地域名が指定されている場合は、指定されている地域に属する階層構造の各階層の検索から開始することによって、検索を行うために必要な処理時間を大幅に削減するものである。

【００２４】
住所は４７都道府県名から全部表記する場合もあるが、文脈などから都道府県名が自明であったり、よく知られている地名であるために都道府県名が省略できる地名であったりした場合は都道府県名を省略して表記することが多い。しかし、文章中に階層の途中から始まる住所が含まれている場合、階層化された住所辞書103 は途中の階層からの検索も可能であるが、階層が下がるにつれて検索対象の単語数は増え、検索対象の単語が数十万以上にも及ぶ可能性がある。
そのため、この実施の形態では、階層上のどの地点を起点にして検索を行うかという情報を地域で指定する。
【００２５】
図８は、例えば、指定地域入力部110 より「北海道旭川市」という地域が指定された場合の検索開始位置格納バッファ111 の概念図である。指定地域入力部110 より「北海道旭川市」という地域が指定された場合、指定地域が住所辞書検索部102 により検索され、「階層先頭」及び「北海道」及び「旭川市」という３つのそれぞれの階層構造上の検索ポインタが検索開始位置格納バッファ111 に格納される。住所辞書検索部102 は、検索開始位置格納バッファ111 中の各々の検索ポインタを起点として住所辞書103 を検索する。
なお、検索開始位置の情報は１地域に限らず、複数地域の情報を格納しておく構成であってもよい。
【００２６】
図７(b) は実施の形態３の変形例の構成図である。本変形例が実施の形態３と異なる点は、指定地域入力部110 に替えて、入力された文章情報中から地域の情報を獲得する指定地域獲得部112 が設けられている点である。指定地域獲得部112 は、住所辞書検索部102 で住所辞書103 が検索され、文章中の単語に一致する住所辞書103 中の単語列候補が存在した場合、文章中の地名の地域を獲得して検索開始位置格納バッファ111 に格納し、検索開始位置の情報は上述と同様に利用される。
【００２７】
〔実施の形態４〕
図９は本発明の音声合成装置の実施の形態４の構成図である。なお、上述の実施の形態と同一部分には同一符号を付してその説明を省略する。
実施の形態４では、途中の階層から始まる地名を検索すべく、住所辞書の途中から以下の各階層を全て検索する方法をとった場合、検索に長時間を要するという不具合を解消するために階層構造の住所辞書103 を図10のような構成にするとともに、単語間の接続関係を求めるための接続関係設定部113 が設けられている。
【００２８】
即ち、図10の住所辞書103 では、階層構造の全ての単語が、表記及び読みの情報に、読みがこの読みになる、上位の階層の親単語を特定する親番号のような親情報を付与して親子関係で表現されている（例：親単語＃１，親単語＃２，…）。さらに、表記による検索の簡単のために表記のコード順にソーティングされている。このとき、一つの表記に異なる読みを持つ単語は複数の読みのそれぞれに親情報を持たせ、また異なる表記で同じ読みを持つ単語は各読みに複数の親情報を持たせることで情報量を圧縮することもできる。
【００２９】
図11は接続関係設定部113 のアルゴリズムのフローチャートであって、破線で囲んだステップは接続関係設定部113 での処理を示している。
検索文字位置を文章情報の最初の文字に設定し（Ｓ201 ）、開始位置で始まる単語を検索する（Ｓ202 ）。開始位置で始まる単語が存在するか否かを判定し（Ｓ203 ）、開始位置で始まる単語が存在しない場合は検索文字位置を次の文字に設定し（Ｓ204 ）、テキスト中の最後の文字か否かを判定する（Ｓ205 ）。テキスト中の最後の文字でない場合はステップＳ202 に移行し、開始位置で始まる単語を検索する（Ｓ202 ）。住所辞書検索部102 は文章に含まれる全ての単語に基づいて、住所辞書103 の全階層を検索して文章中の単語に一致する全候補を抽出する。

【００３０】
ステップＳ203 での判定の結果、開始位置で始まる単語が存在する場合、異なる読み、単語を持つ単語を分割する（Ｓ206 ）。
接続関係設定部113 は、全ての単語に関する処理を行ったか否かを判定し（Ｓ207 ）、全ての単語に関して処理を行っていない場合、開始位置で終わる単語が存在するか否かを判定し（Ｓ208 ）、存在しない場合は親単語無しの情報を設定する（Ｓ209 ）。
一方、ステップＳ208 の判定の結果、開始位置で終わる単語が存在する場合は親番号が一致する単語が存在するか否かを判定する（Ｓ210 ）。親番号が一致する単語がない場合は親単語無しの情報を設定する一方（Ｓ209 ）、親番号が一致する単語が存在する場合は親単語へのポインタを設定する（Ｓ211 ）。
【００３１】
ステップＳ207 の判定の結果、全ての単語に関して処理を行った場合はステップＳ204 に移行して検索文字位置を次の文字に設定し、テキスト中の最後の文字になるまでステップＳ204 ～Ｓ211 を繰り返す。
【００３２】
この実施の形態では、親単語が存在する場合には必ずその単語の持つ読みを選択する。住所辞書103 をこのような構成にすることにより、階層構造のどの階層から始まっている住所であっても正しい読みが得られる。
また、住所辞書として、住所の各階層を表す単語にハッシュインデックスを付与し、ハッシュインデックスを介して読みを検索する構成であってもよい。住所辞書をこのような構造にした場合でも、住所辞書の検索により表記に一致する候補を住所辞書から抽出した後で接続関係を求める手順は図10の構成の場合と同様である。
【００３３】
また、文章中の単語列に一致する単語列を住所辞書103 から求めたときに複数通りのマッチングが発生する場合がある。例えば、《東京都－港区－白金》と《東京都－港区－白金台》という地名が含まれている住所辞書103 を用いて『東京都港区白金台は、…。』という文章を音声合成する場合、又は《山形県－南陽市－宮内－新町》と《熊本県－荒尾市－宮内―新町》という地名が含まれている住所辞書103 を用いて『荒尾市宮内新町は、…。』という文章を音声合成する場合に住所辞書103 において両方の地名にマッチングする。このとき、先の例では《東京都－港区－白金》の全文字長は７文字であり、《東京都－港区－白金台》の全文字長は８文字であるので、文字長が長い方の読みを選択する。また後の例で、山形県の地名では《宮内－新町》と４文字しか一致しないが、熊本県の地名では《荒尾市－宮内－新町》と７文字一致するので、文字長が長い方の熊本県の読みを選択する。
【００３４】
以上のアルゴリズムのフローチャートを図12に示す。
全単語から単語候補列を作成し（Ｓ301 ）、一番長い文字数の単語列を選択する（Ｓ302 ）。選択した単語と、この単語と区間が重複する候補列を、単語候補列を格納しているバッファ（図示せず）から削除する（Ｓ303 ）。単語候補列のバッファが空か否かを判定し（Ｓ304 ）、バッファが空になるまでステップＳ302 、Ｓ303 を繰り返す。
このアルゴリズムでは文章中に複数個の住所が含まれている場合も考慮した処理を行う。
【００３５】
〔実施の形態５〕
図13(a) 及び図13(b) は本発明の音声合成装置の実施の形態５の構成図であって、図13(a) は階層構造をそのまま持つ図４の構成の住所辞書103 を使用する場合の構成図、図13(b) は階層情報を単語間の親子関係で表現した図10の構成の住所辞書103 を使用する場合の構成図である。なお、図中、上述の実施の形態と同一部分には同一符号を付してその説明を省略する。この実施の形態では、階層省略情報設定部114 を設けた点が異なる。
【００３６】
即ち、地名を表す場合に省略される部分として、都道府県名から始まる先頭部分だけでは

なく、住所の途中の階層が省略される場合もある。例えば、正確には『山梨県西八代郡上九一色村』であるが、『山梨県上九一色村』と郡の名称が省略されている場合、また正確には『神奈川県横浜市緑区長津田』であるが、『神奈川県横浜市長津田』と区の名称が省略されているような場合がある。このような住所辞書の階層構造の一階層又は何階層かが省略された住所表記が文章中に存在する場合には階層構造を持った住所辞書をそのまま参照して検索しても一致する候補が探し出せず、住所を正しく読み上げることはできない。

【００３７】

従って、実施の形態５では、単語検索時又は接続関係設定時に一階層飛ばした組み合わせも可能であるという規則を設ける。

図13(a) に関しては、階層毎にポインタをずらしながら単語を検索していく方法を基に説明する。この場合、特定の階層又は全ての階層での単語を検索する際に、検索ポインタで示される階層の単語を検索するだけでなく、それらの単語及びそれらの単語の一階層又は何階層か下の全ての単語を検索する。階層省略情報設定部114 は、このような検索の階層省略の情報を設定する手段である。

【００３８】

図13(b) に関しては、文章中に表れる全階層中の全単語を検索し、その後に、辞書引きされた単語間の接続関係を接続関係設定部113 で求める方法を基に説明する。この場合、一階層又は何階層か省略された場合でも単語間の親子関係が存在するという関係を接続関係設定部113 で設定する必要があるが、そのための制御を階層省略情報設定部114 で行う。即ち、親情報のポインタを二回たどることで、一階層省略された場合でも接続関係があるということを簡単に判定することができる。

【００３９】

また、上記問題点を解決する他の方法として、単語検索時又は接続関係設定時に、郡の名称，区の名称等の省略される可能性のある単語に関して、省略される可能性があるという情報を、住所辞書103 の単語の属性として持つ方法が考えられる。

図14(a) 、図14(b) は実施の形態５の変形例の構成図であって、図14(a) は階層構造をそのまま持つ図４の構成の住所辞書103 を使用する場合の構成図、図14(b) は階層情報を単語間の親子関係で表現した図10の構成の住所辞書103 を使用する場合の構成図である。

【００４０】

図14(a) に関しては、階層毎にポインタをずらしながら単語を検索していく方法を基に説明する。この場合、階層省略情報獲得部115 が、住所辞書103 中の単語の属性を調べ、省略される可能性があるという情報が含まれている場合のみ、上述と同様に、それらの単語及びそれらの単語の一階層又は何階層か下の全ての単語を検索する。この方法では、一階層又は何階層か省略された場合に、省略される可能性のある部分だけを検索するので、処理量が増加しないという利点がある。

【００４１】

図14(b) に関しては、文章中に表れる全階層中の全単語を検索し、その後に、辞書引きされた単語間の接続関係を接続関係設定部113 で求める方法を基にして説明する。この場合、階層省略情報獲得部115 が、住所辞書103 中の単語の属性を調べて、省略される可能性があるという情報が含まれている場合のみ、親情報のポインタを二回たどることで、一階層省略されて記載されている場合でも接続関係があることを判定することができる。また、親単語が省略される可能性がある場合には、親単語の親単語を親情報として持つというように、複数個の親情報を持つ構成としてもよい。

【００４２】

〔実施の形態６〕

図15は本発明の音声合成装置の実施の形態６の構成図であって、上述の実施の形態と同一部分には同一符号を付してその説明を省略する。

この実施の形態では、全ての可能な単語列候補を作成した後で、その中から一番確からしい候補を選択する単語列候補選択部116 と、文章中に一致した住所辞書103 中の単語列候補に対して住所の読みを選択するか否かを判定する住所読み判定部117 とが設けられてい

る。
【００４３】
即ち、住所辞書にマッチングする際に必ず住所辞書を参照すると、以下のような、住所以外の部分を住所の読みに置き換えてしまうという不具合がある。例えば、『化石（バケ＇イシ）』、『三角（ミカド、ミ＇スミ）』、『山寺（ヤ＇マジ）』、『小文字（コモンジ）』、『大文字（ダ＇イモンジ）』等のように、普通名詞と異なる読みを持つ地名が存在する場合がある。また、たとえ読みが同じであっても、アクセント型又はアクセント結合属性が違うために、読み上げるアクセントが違ってくる場合も発生する。従って、このような場合には住所の読みで読み上げないようにする必要がある。この不具合を回避するため、実施の形態６では、所定の単語数又は文字数より少ない単語数又は文字数しか一致しない場合は住所の読みを選択しないようにする。
【００４４】
例えば、《岡山県（オカヤマ＇ケン）－上房郡（ジョーボ＇ーグン）－賀陽町（カヨーチョー）－北（キ＇タ）－門（カド）》、《大分県（オーイタ＇ケン）－東国東郡（ヒガシクニサキ＇グン）－国見町（クニミチョー）－中（ナ＇カ）―下（シモ）》という地名が住所辞書103 に登録されている場合、『北門で待つ』『上中下』等の文章が入力された場合でも「北門」の２文字、又は「中下」の２文字しかマッチングしない場合には住所の読みを選択しないようにして基本辞書105 を参照するようにすれば、正しく読み上げることができる。
【００４５】
図16は実施の形態６のアルゴリズムのフローチャートである。
候補単語列の単語数又は文字数を求め（Ｓ401 ）、求めた単語数又は文字数がしきい値より小さいか否かを判定する（Ｓ402 ）。候補単語列の単語数又は文字数がしきい値以上の場合は住所の読みを選択する（Ｓ403 ）。一方、求めた単語数又は文字数がしきい値より小さい場合は住所の読みを選択せずに終了する。
【００４６】
〔実施の形態７〕
図17は本発明の音声合成装置の実施の形態７の構成図であって、上述の実施の形態と同一部分には同一符号を付してその説明を省略する。実施の形態７では、自然な発声の合成音声を得るために韻律境界記号を設定する韻律境界設定部118 が設けられている。
【００４７】
即ち、全体のモーラ数の長い住所を読み上げる場合、呼気段落境界及びフレーズ境界を設定しないと、発声のピッチが低くなりすぎたり、息つぎが無い発声で息苦しく聞こえたりする。例えば、『北海道札幌市南区定山渓定山渓豊羽鉱山くるみ沢』という住所の読みは、「ホッカ＇イドー＿サッポロ＇シ＿ミナミ＇ク＿ジョーザ＇ンケー＿ジョーザンケートヨハコ＇ーザン＿クルミ＇サワ」であるが、一気に読むと非常に不自然に聞こえる。また、区切りすぎでも不自然に聞こえる。そのために、フレーズ境界又は呼気段落境界等の韻律境界記号を適当な位置に設定する必要がある。
【００４８】
単語間境界記号を設定する第１の方法として、住所辞書103 とマッチングした単語のモーラ数を累積していき、累積モーラ数がしきい値を超えないように、又は超える毎にフレーズ境界又は呼気段落境界等の韻律境界記号を設定する方法が考えられる。
【００４９】
図18はこのアルゴリズムのフローチャートである。
累積モーラ数に現単語モーラ数を設定する（Ｓ501 ）。地名単語の読みをバッファ（図示せず）に設定し（Ｓ502 ）、地名単語のポインタを次に進める（Ｓ503 ）。地名単語候補列が終了か否かを判定し（Ｓ504 ）、終了でない場合は地名単語のモーラ数を加算する（Ｓ505 ）。累積モーラ数がしきい値を超えたか否かを判定し（Ｓ506 ）、しきい値を超えるまでステップＳ502 ～Ｓ505 を繰り返す。
累積モーラ数がしきい値を超えると、呼気段落記号をバッファに設定し（Ｓ507 ）、ステ

ップＳ501 に戻って地名単語候補列が終了するまで、ステップＳ501 ～Ｓ507 を繰り返す。

【００５０】

先の『北海道札幌市南区定山渓定山渓豊羽鉱山くるみ沢』の例では、単語毎のモーラ数は｛６，５，４，６，13，５｝であるから、しきい値を13モーラに設定すれば、｛（６，５），（４，６），13，５｝に分割される。ここで、呼気段落記号を「・」で表した場合、読みは「ホッカ＇イドー＿サッポロ＇シ・ミナミ＇ク＿ジョーザ＇ンケー・ジョーザンケートヨハコ＇ーザン・クルミ＇サワ」となり、自然に読み上げることができる。

【００５１】

単語間境界記号を設定する第２の方法として、住所辞書103 中の地名単語の階層データ構造中に境界記号を含めておき、それを参照してフレーズ境界又は呼気段落境界等の韻律境界記号を設定する方法が考えられる。

図19はこのアルゴリズムのフローチャートである。

地名単語の読みをバッファに設定する（Ｓ601 ）。地名単語のポインタを次に進め（Ｓ602 ）、地名単語候補列が終了か否かを判定する（Ｓ603 ）。地名単語候補列が終了でない場合は韻律境界記号があれば獲得して設定し（Ｓ604 ）、ステップＳ601 に戻って、地名単語候補列が終了するまでステップＳ601 ～Ｓ604 を繰り返す。

【００５２】

第２の方法では、モーラ数だけ見るのではなく、予めフレーズ境界及び呼気段落境界記号を区別して入れることができるので、より自然に発声できる。前述の例で示すと、例えば『札幌市』と『南区』とを一緒に発声した方がより自然であるため、それらの情報を住所辞書103 に格納しておく。ここで、呼気段落境界記号を「・」、フレーズ境界記号を「／」で表した場合、前述の例を「ホッカ＇イドー／サッポロ＇シ＿ミナミ＇ク／ジョーザ＇ンケー・ジョーザンケートヨハコ＇ーザン／クルミ＇サワ」と読み上げることが可能になる。

【００５３】

〔実施の形態８〕

図20は本発明の音声合成装置の実施の形態８の構成図であって、図中、上述の実施の形態と同一部分には同一符号を付してその説明を省略する。実施の形態８では、マイナス記号，長音記号等で表記されている番地用記号を、番地用の読みに変換する番地用記号変換部119 が設けられている。

即ち、『東京都千代田区丸の内１－６－１』と表記された住所を含む文章が入力された場合、番地『１－６－１』を「イチ・マイナス・ロク・マイナス・イチ」と読み上げてしまうと、住所の読みとして違和感を与える。そのため、住所に上記のような連続する数値が表記されている場合に、『１の６の１』、即ち、「イチ＇ノ・ロク＇ノ・イチ」と読み上げる。

【００５４】

図21はこのアルゴリズムのフローチャートである。

文章中に一致した住所辞書103 中の単語列から決定される住所を認識した後、その住所区間に続く文字を検索していく。次の文字が数字（０～９、〇～九、十、百、千、万）か否かを判定し（Ｓ701 ）、数字の場合はポインタを一文字分進め（Ｓ702 ）、ステップＳ701 に戻る。次の文字が数字でない場合は、次の単語が助数詞（番地、番、丁目、号）か否かを判定し（Ｓ703 ）、助数詞の場合はポインタを単語の文字分進め（Ｓ704 ）、ステップＳ701 に戻る。

【００５５】

次の単語が助数詞でない場合は次の文字が区切り記号（－、ー）か否かを判定し（Ｓ705 ）、区切り記号の場合は文字を『の』で置き換える（Ｓ706 ）。ここで、区切り記号として長音記号『ー』を含めたのは、マイナス記号『－』を長音記号で誤って表記するケースも多いので、誤って表記されていても読めるようにするためである。

以上のいずれの文字、単語でもない場合は処理を終了する。これらの処理を行うことで、

数値間の区切り記号文字を含む番地の記述を正しく読み上げることができる。

【００５６】

〔実施の形態９〕

図22は本発明の音声合成装置の実施の形態９の構成図である。なお、上述の実施例と同一部分には同一符号を付してその説明を省略する。

例えば、『東京都豊島区立第三中学校』という文章が入力された場合に、文章中に一致した住所辞書103 の単語列から決定される住所区間のみを発音情報に変換してしまうと、「トーキョ＇ート＿トシマ＇ク」＋『立第三中学校』と解釈し、結果的に文章解析部104 で『立』を「タ＇チ」と読んでしまう。

【００５７】

従って、実施の形態９では、『立、内、民、行き、発、着、…』等の語彙からなる住所接尾語及びアクセント結合属性等が格納されている住所接尾語辞書120 と、文章中の単語に一致した住所辞書103 中の単語列から決定される住所区間の直後に続く単語に一致する単語を住所接尾語辞書120 から検索する住所接尾語辞書検索部121 と、住所接尾語辞書検索部121 の検索によって住所の後ろに住所接尾語が存在した場合は、文章中の単語に一致した住所辞書103 中の単語列からなる住所区間に住所接尾語が含まれるように修正し、文章中の単語に一致した住所辞書103 中の単語列の最終単語と住所接尾語とを住所接尾語に設定されているアクセント結合属性に応じてアクセント結合し、読みを設定するアクセント結合処理部122 とが設けられている。

【００５８】

図23は実施の形態９のアルゴリズムのフローチャートである。

住所の単語列の次の単語が住所接尾語辞書120 にあるか否かを判定し（Ｓ801 ）、住所接尾語辞書120 にある場合はアクセント結合処理部122 がアクセント結合する（Ｓ802 ）。このような処理を行うことで、先の例の文章情報を「トーキョ＇ート＿トシマク＇リツ」＋『第三中学校』という情報として文章解析部105 に渡せるため、正しく読み上げることができる。

【００５９】

〔実施の形態10〕

図24は本発明の音声合成装置の実施の形態10の構成図である。なお、上述の実施例と同一部分には同一符号を付してその説明を省略する。

例えば、文章中に住所が出現した後で、その住所に含まれる単語が繰り返し出現する場合、例えば、『神奈川県中原区上小田中に、上小田中公民館はあります。』という文章で、地名の『神奈川県中原区上小田中』は正しく読めるが、次に出てくる『上小田中公民館は…』という文章に関しては、先で読めているにもかかわらず、同じ表記で別の読みが登録されているために間違った読みで読み上げることがある。また、実施の形態６のような構成を採っている場合、文字数（単語数）が所定のしきい値より少ないために住所読みで処理できないことも起こり得る。

【００６０】

従って、実施の形態10では、住所辞書検索部102 で処理された住所の全ての構成単語に対して、漢字表記と発音情報とをペアにして学習用単語バッファ124 に格納する住所単語学習部123 と、最近使用した住所の表記と読みとを格納する学習用単語バッファ124 と、階層構造の住所辞書103 を検索した後の文章に対して学習用単語の検索を行い、学習用単語に一致する単語が存在する場合には、該当する文章の部分に、対応する発音情報を埋め込む学習用単語検索部125 とが設けられている。学習用単語バッファ124 に格納されている単語は、その読みを優先的に使用することによって、文章中に一度でも出現した住所の一部が次に出現した場合に正しく読み上げることができる。

【００６１】

また、学習用単語バッファ124 内の内容を初期化する学習用単語削除部126 を設けてもよい。この場合、学習用単語バッファ124 は通常は初期化せずに使用者の指定に応じて初期化する構成であっても、また、文章情報の入力の都度、初期化する構成であってもよい。

さらに、登録された後に入力された文章の数が所定数を超えた時点で初期化する構成であってもよい。

【００６２】

なお、本発明の音声合成装置において実施される音声合成方法は、音声合成装置のＲＯＭに書き込んでおく以外に、図25に示すように、コンパクトディスク等の記録媒体Ｄに記録しておき、この記録媒体Ｄをパーソナルコンピュータのディスクドライブに装填して音声合成する構成であってもよい。

【００６３】

なお、上述の実施の形態では単語列が住所の場合について説明したが、単語列は住所に限らず、階層構造を有する単語列であれば同様の効果が得られる。

【００６４】

【発明の効果】

以上のように、本発明の音声合成装置及び音声合成方法は、一つの表記に対して複数の読みが存在する住所の表記及び読みの情報を階層化した階層構造の辞書を持つので、一つの表記に対して複数の読みが存在する住所の文章情報を正確な読みで読み上げるという優れた効果を奏する。

【図面の簡単な説明】

【図１】　本発明の音声合成装置の基本ブロック図である。

【図２】本発明の音声合成装置の一例の模式図である。

【図３】本発明の音声合成装置の実施の形態１の構成図である。

【図４】住所辞書の一例の概念図である。

【図５】実施の形態１のアルゴリズムのフローチャートである。

【図６】本発明の音声合成装置の実施の形態２の構成図である。

【図７】本発明の音声合成装置の実施の形態３及びその変形例の構成図である。

【図８】検索開始位置格納バッファの概念図である。

【図９】本発明の音声合成装置の実施の形態４の構成図である。

【図１０】住所辞書の他の例の概念図である。

【図１１】実施の形態４のアルゴリズムのフローチャート（その１）である。

【図１２】実施の形態４のアルゴリズムのフローチャート（その２）である。

【図１３】本発明の音声合成装置の実施の形態５の構成図である。

【図１４】本発明の音声合成装置の実施の形態５の変形例の構成図である。

【図１５】本発明の音声合成装置の実施の形態６の構成図である。

【図１６】実施の形態６のアルゴリズムのフローチャートである。

【図１７】本発明の音声合成装置の実施の形態７の構成図である。

【図１８】実施の形態７のアルゴリズムのフローチャート（その１）である。

【図１９】実施の形態７のアルゴリズムのフローチャート（その２）である。

【図２０】本発明の音声合成装置の実施の形態８の構成図である。

【図２１】実施の形態８のアルゴリズムのフローチャートである。

【図２２】本発明の音声合成装置の実施の形態９の構成図である。

【図２３】実施の形態９のアルゴリズムのフローチャートである。

【図２４】本発明の音声合成装置の実施の形態10の構成図である。

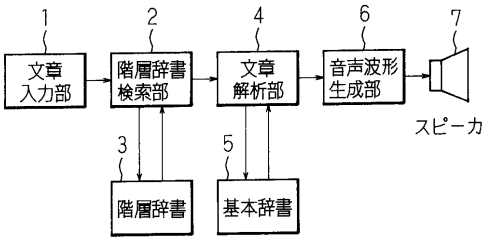
【図２５】　記録媒体の記録状態の概念図である。

【符号の説明】

1　文章入力部
2　階層辞書検索部
3　階層辞書
4　文章解析部
5　基本辞書
6　音声波形生成部
7　スピーカ

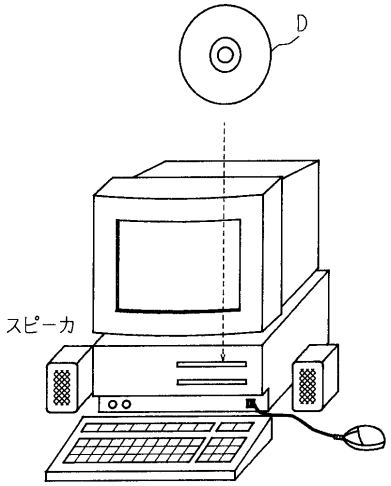
【図１】

本発明の音声合成装置の基本ブロック図



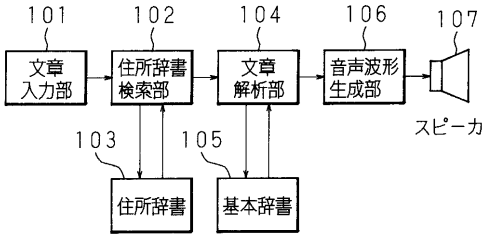
【図２】

本発明の音声合成装置の一例の模式図



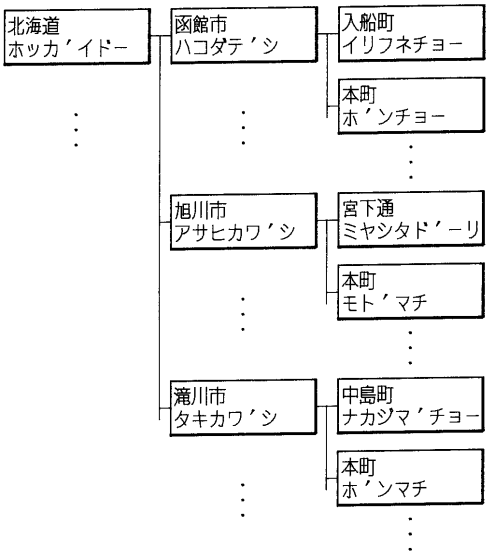
【図３】

本発明の音声合成装置の実施の形態１の構成図



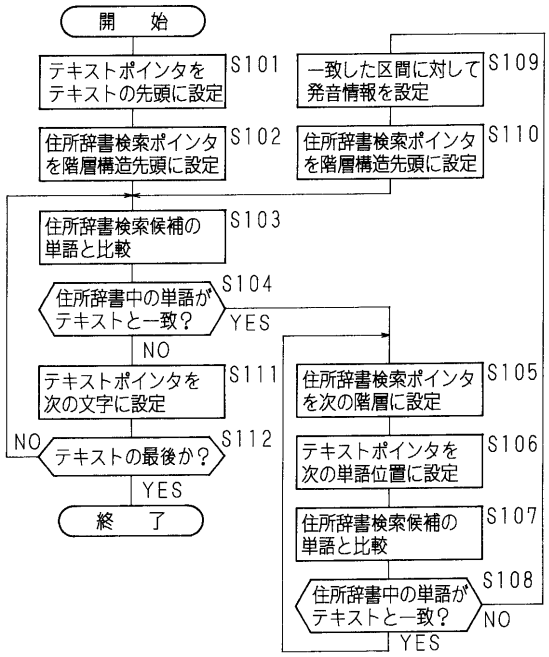
【図４】

住所辞書の一例の概念図



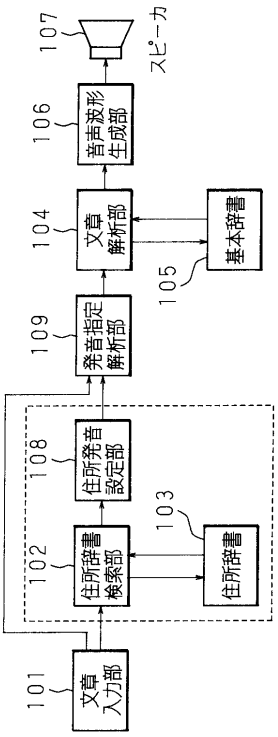
【図５】

実施の形態１のアルゴリズムのフローチャート



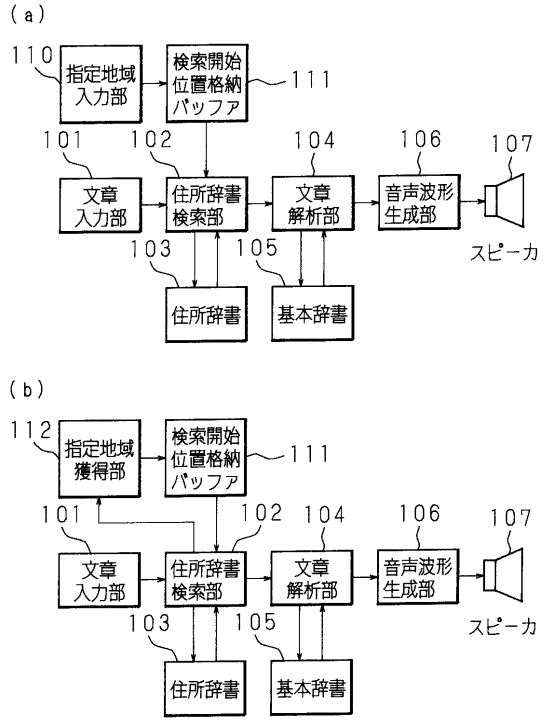
【図６】

本発明の音声合成装置の実施の形態２の構成図



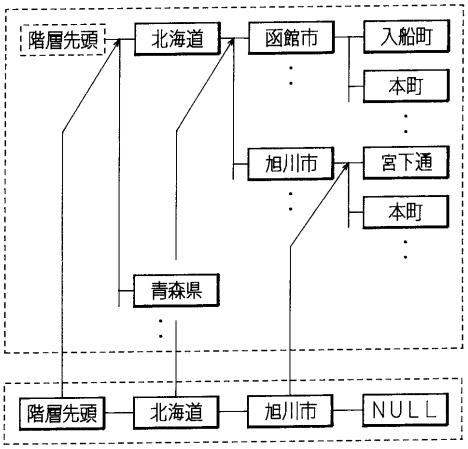
【図７】

本発明の音声合成装置の実施の形態３及びその変形例の構成図



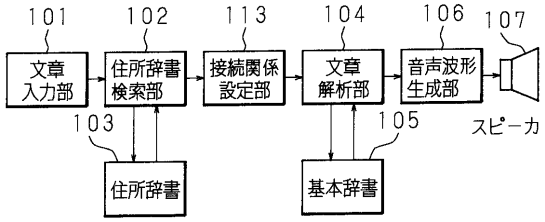
【図８】

検索開始位置格納バッファの概念図



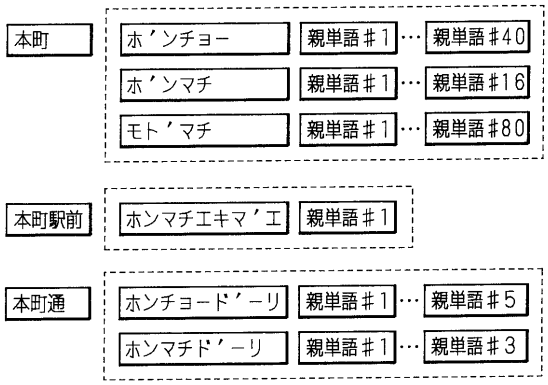
【図９】

本発明の音声合成装置の実施の形態４の構成図



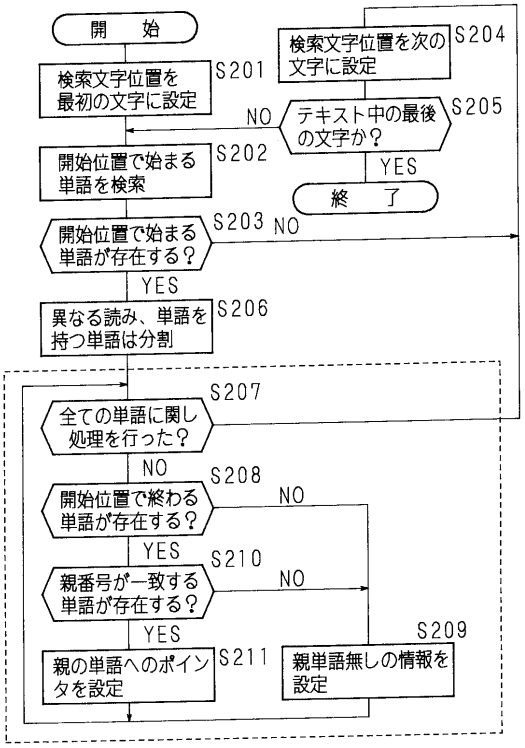
【図１０】

住所辞書の他の例の概念図



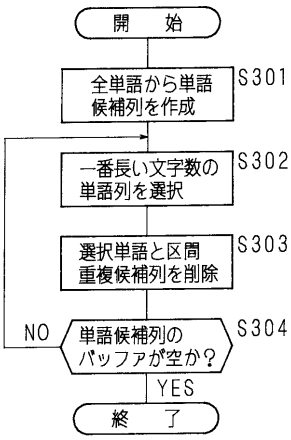
【図１１】

実施の形態４のアルゴリズムのフローチャート（その１）



【図１２】

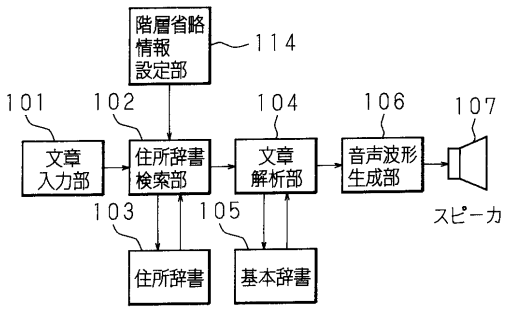
実施の形態４のアルゴリズムのフローチャート（その２）



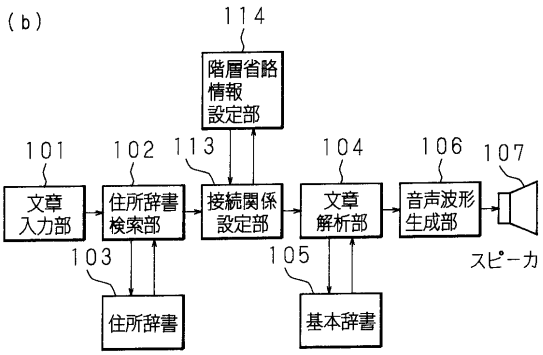
【図１３】

本発明の音声合成装置の実施の形態５の構成図

(a)

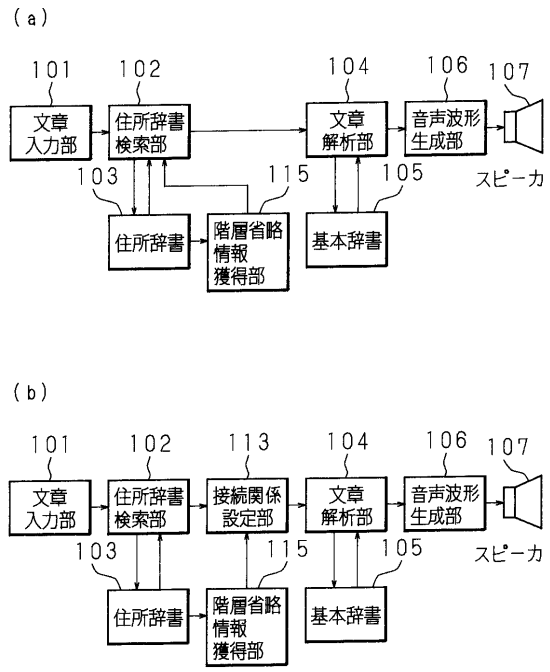


(b)



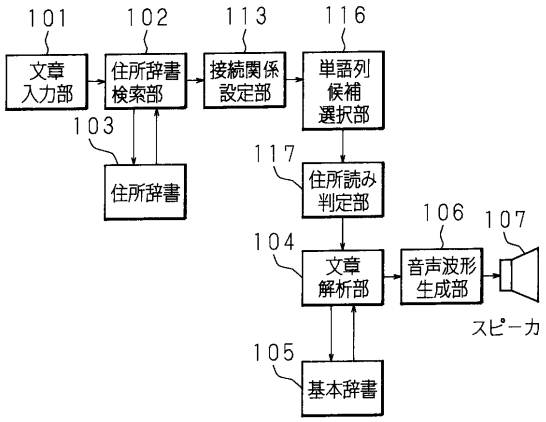
【図１４】

本発明の音声合成装置の実施の形態５の変形例の構成図



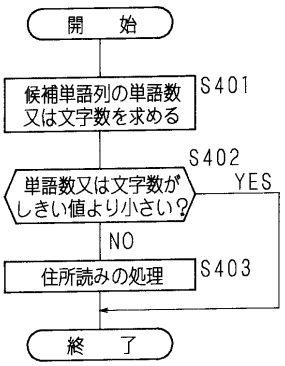
【図１５】

本発明の音声合成装置の実施の形態６の構成図



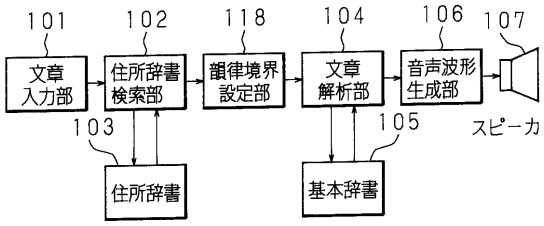
【図１６】

実施の形態６のアルゴリズムのフローチャート



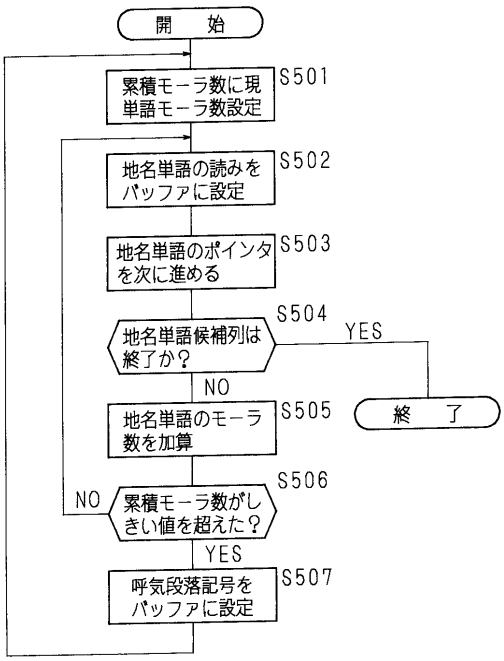
【図１７】

本発明の音声合成装置の実施の形態７の構成図



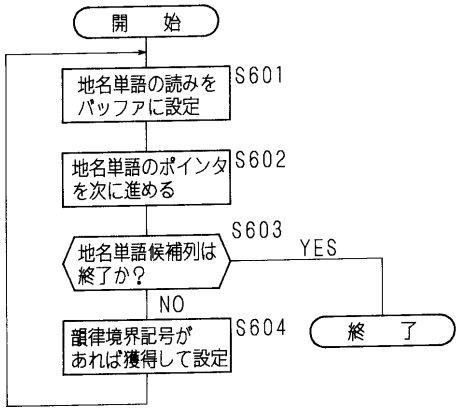
【図１８】

実施の形態７のアルゴリズムのフローチャート（その１）



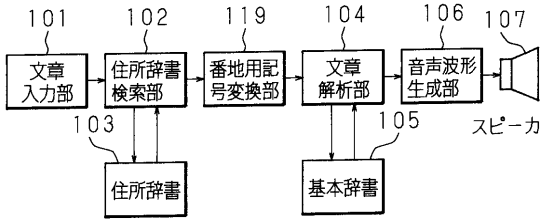
【図１９】

実施の形態７のアルゴリズムのフローチャート（その２）



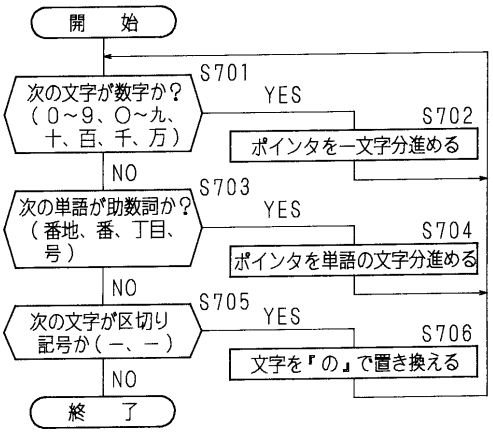
【図２０】

本発明の音声合成装置の実施の形態８の構成図



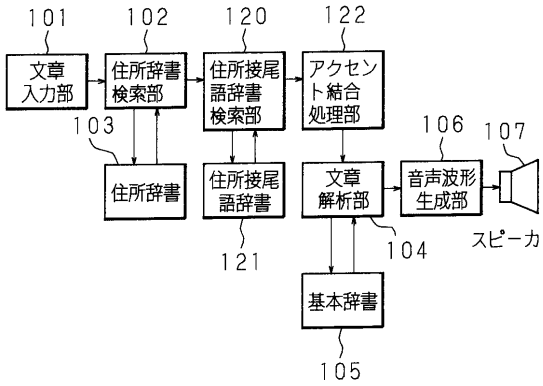
【図２１】

実施の形態８のアルゴリズムのフローチャート



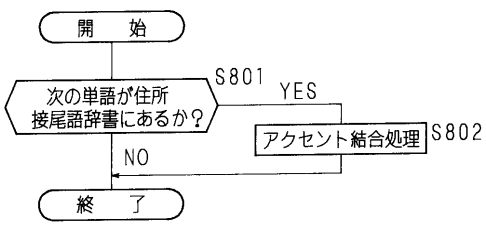
【図２２】

本発明の音声合成装置の実施の形態９の構成図



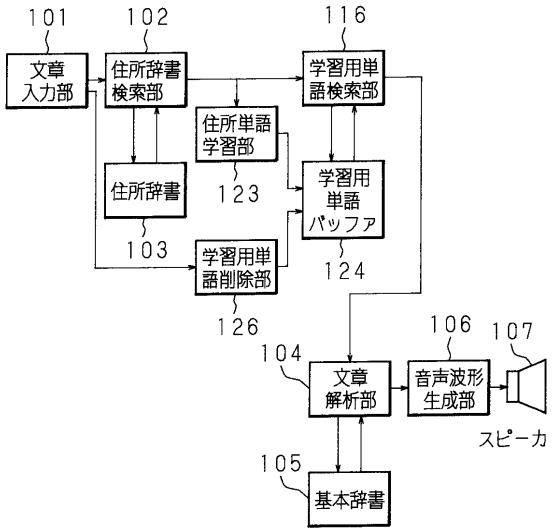
【図２３】

実施の形態９のアルゴリズムのフローチャート



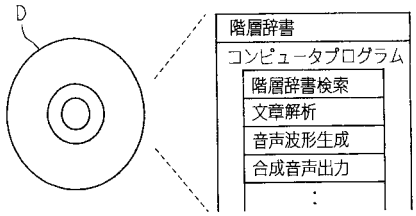
【図２４】

本発明の音声合成装置の実施の形態１０の構成図



【図２５】

記録媒体の記録状態の概念図



フロントページの続き

(72)発明者　木村　晋太
神奈川県川崎市中原区上小田中１０１５番地　富士通株式会社内

合議体
審判長　原　光明
審判官　加藤　恵一
審判官　井上　健一

(56)参考文献　特開平２－２０１６４３（ＪＰ，Ａ）
宮崎、大山：「階層的単語属性を用いた同形語の自動読み分け法」，電子通信学会論文誌　’８５．３　Ｖｏｌ．Ｊ６８－Ｄ　Ｎｏ．３　ｐ．３９２－３９９

(58)調査した分野(Int.Cl.，ＤＢ名)
G10L13/00-13/08