

(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2006-227367

(P2006-227367A)

(43) 公開日 平成18年8月31日(2006.8.31)

(51) Int. Cl.

G 1 0 L 13/08 (2006.01)

F I

G 1 0 L 13/08 1 3 0 E

G 1 0 L 13/08 1 2 7 F

テーマコード (参考)

審査請求 未請求 請求項の数 4 O L (全 11 頁)

(21) 出願番号 特願2005-42130 (P2005-42130)

(22) 出願日 平成17年2月18日 (2005.2.18)

(71) 出願人 000000295

沖電気工業株式会社

東京都港区虎ノ門1丁目7番12号

(74) 代理人 100089093

弁理士 大西 健治

(72) 発明者 矢頭 隆

東京都港区虎ノ門1丁目7番12号 沖電気工業株式会社内

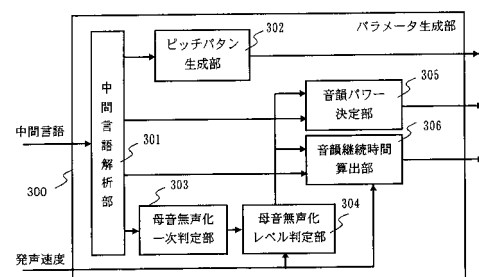
(54) 【発明の名称】 音声合成装置

(57) 【要約】

【課題】 音声の無声化に関する従来の問題点を解決して自然な合成音を得られるような音声合成装置を提供する。

【解決手段】 音声の基本単位となる音声素片が登録された素片辞書と、音韻・韻律記号列に対して少なくとも音声素片、音韻継続時間、基本周波数の合成パラメータを生成するパラメータ生成部と、パラメータ生成部からの合成パラメータを、素片辞書を参照しながら合成波形を生成する波形生成部とを備えた音声合成装置において、パラメータ生成部300は、通常の無声化規則にしたがって無声化判定を行った結果に対して、発声速度と音節の種類に応じて判定結果の補正をおこなう母音無声化レベル判定部304を備えており、この母音無声化レベル判定部304は、少なくとも2以上のレベルによる無声化判定を行うようにしている。

【選択図】 図1



【特許請求の範囲】

【請求項 1】

音声の基本単位となる音声素片が登録された素片辞書と、音韻・韻律記号列に対して少なくとも音声素片、音韻継続時間、基本周波数の合成パラメータを生成するパラメータ生成部と、前記パラメータ生成部からの合成パラメータを前記素片辞書を参照しながら合成波形を生成する波形生成部とを備えた音声合成装置において、

前記パラメータ生成部は母音無声化処理を行うか否かの判定をおこなう母音無声化判定手段を備え、当該母音無声化判定手段は無声化の程度に応じた少なくとも 2 以上の無声化レベルを判定することを特徴とする音声合成装置。

【請求項 2】

前記無声化レベルは、「無声化しない」と「無声化する」の間に少なくとも 1 レベルの中間的レベルを有することを特徴とする請求項 1 項記載の音声合成装置。

【請求項 3】

前記母音無声化判定手段は、文字列の並びとアクセント核の有無に基づいて母音無声化判定を行う一次判定手段を備え、該一次判定手段の結果とユーザから指定された発声速度に応じて、前記無声化レベルを選択する無声化レベル決定基準を有することを特徴とする請求項 1、2 項の何れか 1 項に記載の音声合成装置。

【請求項 4】

前記無声化レベル決定基準は、無声化対象の音節の種類に応じて定められていることを特徴とする請求項 3 項記載の音声合成装置。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、日常読み書きしている漢字・かな混じり文のテキストを入力として、それを音声に変換するテキスト音声合成装置、特に、母音無声化の実現処理に関するものである。

【背景技術】

【0002】

テキスト音声変換技術は、我々が日常読み書きしている漢字かな混じり文を入力し、それを音声に変換して出力するもので、出力語彙の制限がないことから録音・再生型の音声合成に代わる技術として種々の利用分野での応用が期待できる。

【0003】

従来、この種の音声合成装置としては、図 2 に示すような処理形態となっているものが代表的である。図 2 において、101 はテキスト解析部、102 はパラメータ生成部、103 は波形生成部、104 は単語辞書、105 は素片辞書である。

【0004】

テキスト解析部 101 は、漢字かな混じり文を入力し、単語辞書を参照して形態素解析し、読み、アクセント、イントネーションを決定し、韻律記号付き発音記号（中間言語）を出力する。

【0005】

パラメータ生成部 102 は、ピッチ周波数パターンや音韻継続時間等の設定を行い、波形生成部 103 では、音声の合成処理を行う。波形生成部 103 は、中間言語で与えられる発音記号列から、音声合成に使用する音声合成単位を、素片辞書 105 から選択し、パラメータ生成部で決定したパラメータに従って、結合／変形して音声の合成処理を行う。音声合成単位には、音素、音節（CV）、VCV、CVC（C：子音、V：母音）や、音韻連鎖を拡張した単位などが用いられる。

【0006】

以上の構成において、日常読み書きしている漢字・かな混じり文（以下、テキストという）を入力すると、テキスト解析部 101 は、文字情報から音韻・韻律記号列を生成する。音韻・韻律記号列とは、入力文の読み、アクセント、イントネーション等を文字列とし

10

20

30

40

50

て記述したもの（以下、中間言語という）である。単語辞書 104 は、単語の読みやアクセント等が登録された発音辞書で、テキスト解析部 101 はこの単語辞書を参照しながら中間言語を生成する。

【0007】

テキスト解析部 101 で生成された中間言語は、パラメータ生成部 102 で、音声素片（音の種類）、音韻継続時間（音の長さ）、基本周波数（声の高さ、以下ピッチという）等の各パターンからなる合成パラメータを決定し、波形生成部 103 に送る。音声素片とは、接続して合成波形を作るための音声の基本単位で、音の種類等に応じて様々なものがある。

【0008】

パラメータ生成部 102 で生成された各種パラメータは、波形生成部 103 で音声素片等を蓄積する ROM 等から構成された素片辞書 105 を参照しながら、合成波形が生成され、スピーカーを通して合成音声が出力される。以上がテキスト音声変換処理の流れである。

【0009】

次に、パラメータ生成部 102 における処理を、図 3 を参照して詳細に説明する。図 3 は従来の音声合成装置のパラメータ生成部 102 の構成を示すブロック図である。図 3 において、パラメータ生成部 102 は、中間言語解析部 201、ピッチパターン生成部 202、母音無声化判定部 203、音韻パワー決定部 204、音韻継続時間算出部 205、継続時間修正部 206 から構成される。

【0010】

パラメータ生成部 102 に入力される中間言語は、アクセント位置、ポーズ位置などを含んだ音韻文字列であり、これより、ピッチの時間的な変化（以下、ピッチパターンという）、それぞれの音韻の継続時間（以下、音韻継続時間という）、音声パワー等、波形を生成する上で必要なパラメータ（以下、波形生成用パラメータという）を決定する。

【0011】

入力された中間言語は、中間言語解析部 201 で文字列の解析が行われ、中間言語上に記された単語区切り記号から単語境界を判定し、アクセント記号からアクセント核のモーラ位置を得る。

【0012】

アクセント核とは、単語内でピッチが下降する位置のことで、1 モーラ目にアクセント核が存在する単語を 1 型アクセント、n モーラ目にアクセント核が存在する単語を n 型アクセントと称し、総称して起伏型アクセント単語と称する。逆に、アクセント核の存在しない単語（例えば「新聞」や「パソコン」）を 0 型アクセントまたは平板型アクセント単語と称する。

【0013】

ピッチパターン生成部 202 は、中間言語上のフレーズ記号・アクセント記号、単語のモーラ数、フレーズのモーラ数などにより、ピッチパターンを生成する。

【0014】

母音無声化判定部 203 は、中間言語上の音韻記号やアクセント記号などから、母音の無声化判定を行い、その結果を音韻パワー決定部 204 と音韻継続時間算出部 205 に送る。母音の無声化については後述する。

【0015】

音韻継続時間算出部 205 は、音韻文字列からそれぞれの音韻の持続時間を計算し、継続時間修正部 206 に送る。ユーザが発声速度レベルを指定する場合は、指定されたレベルに応じて、205 で算出された音韻継続時間を継続時間修正部 206 で線形伸縮する処理を行う。継続時間修正部 206 で発声速度レベルに応じて伸縮された音韻継続時間は、図示していない波形生成部に送られる。

【0016】

音韻パワー決定部 204 は、波形の振幅値を算出し、波形生成部 103（図 2 参照）へ

10

20

30

40

50

送る。

【0017】

以上説明したこれらの波形生成用パラメータは波形生成部103へ送られ、合成波形が生成される。

【0018】

次に、母音無声化について詳細に説明する。人間が言葉を発する時には、肺から押し出された空気を声帯の開閉運動により音源とし、顎・舌・唇などを動かすことにより声道の共鳴特性を変化させて種々の音韻を表現している。前述したピッチは、声帯の振動周期に対応し、この時間的变化がアクセントやイントネーションの表現となる。

【0019】

このような声帯振動の他に、舌によって声道のある部分に狭い場所を作り、そこを空気流が通り抜けるときに乱流を生じて雑音的な音を生成する摩擦音や、舌や唇で声道を遮断して一時空気流を止めた後に一気に開放してインパルス的な音を生成する破裂音もある。

【0020】

母音や破裂音／b／、／d／、／g／、摩擦音／j／、／z／、鼻子音・流音／m／、／n／、／r／などといった、声帯の振動を伴う音韻を総称して有声音、破裂音／p／、／t／、／k／、摩擦音／s／、／h／、／f／といった、声帯の振動を伴わない音韻を無声音と呼ぶ。特に、子音に注目して、声帯の振動を伴う子音を有声音、伴わない子音を無声音と呼ぶ。有声音では、声帯振動による周期的な波形が生成され、無声音では雑音的な波形が生成される。

【0021】

無声音子音（／p／、／t／、／k／、／s／、／f／、／h／ e t c . ）に母音が／i／、／u／が挟まれると、口構えだけを残して声帯を振動させず、息だけで発音する現象が見られる。これを母音の無声化という。

【0022】

図4は「取材した」の実発音波形を示す図である。図4に示すように、「取材した」の「s h i」の母音は周期的な波形として現れず、無声摩擦音「s h」に融合された様な波形となっている。

【0023】

テキスト音声変換システムにおいても、母音の無声化表現は聴感品質を向上させるために重要であり、この判定を行うのが図3の母音無声化 判定部203である。ここで無声化すると判定された母音は、音韻パワー決定部204と音韻継続時間算出部205において、それに応じた特殊な処理が施される。無声化母音は、通常の母音とは異なり、音韻パワー0、音韻継続時間0として波形生成部103に送られる。

【0024】

無声化判定は、たとえば以下のような規則に従って行われる。

- (1) 無声音子音（無音を含む）に挟まれた母音／i／、／u／は無声化する
- (2) 但し、アクセント核が存在すれば無声化しない
- (3) 但し、前母音がすでに無声化していれば無声化しない
- (4) 但し、疑問文末は無声化しない

【0025】

但し、これらは一般的な傾向から導き出された規則であり、実際の発声では上述の規則通りに無声化 が起こるとは限らない。また、個人によっても差がある。上記規則は、通常話速の発声においては至極、妥当な規則であるが、一般的に発話速度が遅くなった場合、無声化は起き難くなる。図5に通常速度と発話速度が遅い場合の／k i s o k u／の音声波形を示す。／k i s o k u／下線部は無声化の条件に合致し、通常速度では無声化が起こっているのに対し、遅い発声ではいずれも母音が明確に現れている。

【0026】

このような現象に対する対策の一つとして特許文献1において、発声速度に応じて母音無声化 処理を行うか否かの判定基準を変える母音無声化判定手段を備えた音声合成装置

10

20

30

40

50

が提案されている。この文献では、発声速度に閾値を設け、閾値以下となる発声速度においては、たとえ規則による無声化要件を満たしても無声化処理を行わないとしている。

【特許文献1】特開平11-116263号公報

【発明の開示】

【発明が解決しようとする課題】

【0027】

発話速度が遅くなると無声化が起き難くなる傾向にあることは確かであるが、すべての音節がそうであるかと言えば、必ずしもそうではない。「～でした：～d e s h i t a」などの音節では、たとえ発話速度が遅くなっても、無声化して発音される場合が多い。傾向として子音部が持続して発音可能な摩擦音／s／、／f／、／h／などは、低話速でも無声化しやすい。 10

【0028】

発話速度が遅い場合、このような音節も含めて一律に無声化を抑止しては、たどたどしく不自然な合成音となる。また、発話速度が遅くなると無声化が起き難くなる音節についても、閾値を境にして、完全に無声化する／まったく無声化しない、というように発声の性質が両極端に変化するわけではない。もともと無声化という現象は、開口面積が狭い子音に挟まれた母音区間では、狭い開口面積が維持されるために、有声音として声帯を振動させるのに必要な呼気流が確保できないことから起こるものである。人間の発声器官の構造と発声機構によるもので、性質の変化は段階的に変わるものである。しかし、従来の一 20
律的、画一的な無声化判定処理では、発話速度の段階的变化に伴う、無声化の段階的性質の変化は表現できず、不自然な合成音にならざるを得ない。

【0029】

従って、本発明の課題は、音声の無声化に関する従来の問題点を解決して自然な合成音 30
が得られるような音声合成装置を提供する点にある。

【課題を解決するための手段】

【0030】

本発明に係る音声合成装置は、音声の基本単位となる音声素片が登録された素片辞書と、音韻・韻律記号列に対して少なくとも音声素片、音韻継続時間、基本周波数の合成パラメータを生成するパラメータ生成部と、パラメータ生成部からの合成パラメータを、素片辞書を参照しながら合成波形を生成する波形生成部とを備えた音声合成装置において、パ 30
ラメータ生成部は、通常は無声化規則にしたがって無声化判定を行った結果に対して、発声速度と音節の種類に応じて判定結果の補正をおこなう判定補正手段を備えており、この判定補正手段は、無声化の程度に応じて少なくとも2以上のレベルによる無声化判定を行うようにしている。

【発明の効果】

【0031】

本発明に係る音声合成装置では、無声化の判定において、少なくとも2以上の無声化レベルを設け、発声速度レベルに応じて、無声化のレベルを段階的に割り当てるようにしたことにより、発声速度の境界点前後においても無声化の程度が急激に変化することなく、 40
いかなる発声速度においても自然な無声化を実現できる。また、発声速度レベルに応じた無声化レベルの設定は、音節の種別ごとに適した設定則を与える構成にしているため、発声速度の変化に対して子音の特性を反映した自然な無声化が実現可能であり、自然で滑らかな合成音声を得ることができる。

【発明を実施するための最良の形態】

【0032】

以下、図面を参照して本発明の実施の形態について説明する。尚、各図面はこの発明が理解できる程度に概略的に示しているにすぎない。

【0033】

図1は実施形態に係る音声合成装置のパラメータ生成部の構成を示すブロック図である。本発明の特徴部分は、無声化判定手段及びその実現方法にある。本実施形態に於いては 50

、図 2 に示すテキスト解析部 101、単語辞書 104、波形生成部 103、素片辞書 105 は従来の技術を用いることが出来る。

【0034】

図 1 において、パラメータ生成部 300 は、中間言語解析部 301、ピッチパターン生成部 302、母音無声化一次判定部 303、母音無声化レベル判定部 304、音韻パワー決定部 305、音韻継続時間算出部 306 を備えている。

【0035】

パラメータ生成部 300 への入力は、従来と同様に韻律記号の付加された中間言語とユーザから指定される発声速度パラメータである。また、ユーザの好みや利用形態などにより、声の高さやイントネーションの大きさを示す抑揚などの声質パラメータを外部から指定する場合もある。 10

【0036】

合成対象の中間言語は、中間言語解析部 301 に入力され、ユーザ指定の発声速度パラメータは、母音無声化レベル判定部 304 に入力される。中間言語解析部 301 からの出力データのうち、例えば、フレーズ区切り記号、単語区切り記号、アクセント記号のようなパラメータはピッチパターン生成部 302 に入力され、音韻記号列や単語区切り記号、アクセント記号のようなパラメータは音韻パワー決定部 305 と音韻継続時間算出部 306 に入力され、音韻記号列やアクセント記号のようなパラメータは母音無声化一次判定部 303 に入力される。

【0037】

ピッチパターン生成部 302 は、入力されたパラメータから、フレーズ指令の生起時点と大きさ、アクセント指令の開始時点・終了時点と大きさ等のデータを算出し、ピッチパターンを生成する。生成されたピッチパターンは波形生成部 103（図 2 参照）に入力される。ピッチパターン生成過程については、本発明と直接関係がないので説明を省略する。 20

【0038】

母音無声化一次判定部 303 は、字面やアクセントなどの入力テキストのみを基準に母音の無声化判定を行い、その判定結果を母音無声化レベル判定部 304 に出力する。

【0039】

母音無声化レベル判定部 304 は、母音無声化一次判定結果とユーザから指定される発声速度レベルとから最終的な無声化判定を行い、その最終的な判定結果を音韻パワー決定部 305 と音韻継続時間算出部 306 に出力する。 30

【0040】

音韻パワー決定部 305 は、母音無声化判定結果と、中間言語解析部 301 から入力される音韻記号列とから、音韻それぞれの振幅形状を算出し、波形生成部 103 に出力する。

【0041】

音韻継続時間算出部 306 は、母音無声化判定結果と、中間言語解析部 301 から入力される音韻記号列とから、音韻それぞれの継続時間を算出し、その結果を波形生成部 103 に出力する。

【0042】

以下、上述のように構成された音声合成装置の動作を説明する。従来技術と異なる点は、パラメータ生成部 300 内の処理であるので、それ以外の処理については省略する。 40

【0043】

まず、ユーザはあらかじめ発声速度レベルを指定する。発声速度は通常、1 分間に何モールの割合で発声するかといった形式のパラメータとして与えられ、利用便利上、5～10 段階程度に量子化してそのレベル値を与える。このレベルに応じて、音韻継続時間の伸長などの処理が行われる。また、声の高さや抑揚などの声質制御のためのパラメータなども指定することができる。ユーザが特に指定しない場合は、あらかじめ定められた値（デフォルト値）が指定値として設定される。

【0044】

ユーザにより指定された発声速度制御用パラメータは、母音無声化レベル判定部 304 と継続時間算出部 306 に送られる。もう一方の入力の間言言語は、中間言語解析部 301 に送られ、中間言語解析部 301 で入力文字列の解析が行われる。ここでの解析単位は、仮に 1 文章単位とする。

【0045】

1 文章に対応する中間言語から、ピッチパタンの生成に関わるパラメータとして例えば、フレーズ指令の数とそれぞれのフレーズ指令のモーラ数、アクセント指令の数とそれぞれのアクセント指令のモーラ数・アクセント型などの情報がピッチパターン生成部 302 に送られる。

【0046】

ピッチパターン生成部 302 では、入力されたパラメータから、フレーズ指令・アクセント指令それぞれの大きさや立ち上げ・立ち下げ位置などの算出を行い、あらかじめ規定した応答関数を用いてピッチパタンの生成を行う。算出されたピッチパターンは波形生成部 103 (図 2 参照) に送られる。

【0047】

また、アクセント記号・音韻文字列などは、母音無声化一次判定部 303 に送られ、母音の無声化判定が行われる。この一次判定では文字列の並びとアクセント核の有無のみから判定が行われ、暫定的な判定結果が母音無声化レベル判定部 304 に送られる。

【0048】

母音無声化レベル判定部 304 には、ユーザから指定される発声速度レベルも入力されており、前記一次判定結果と併せて最終判定処理が行われる。この判定処理に於いては、発声速度がある規定値を超えたか否かを比較し、比較結果から発声速度が遅いと判定された場合に限り母音無声化処理を行わないようにする。この処理の後、母音無声化の最終判定を行うために、判定結果が、音韻パワー決定部 305 と音韻継続時間算出部 306 に送られる。母音無声化一時判定部 303 と母音無声化レベル判定部 304 の処理については、後に詳述する。

【0049】

音韻パワー決定部 305 では、中間言語解析部 301 から入力された音韻文字列などのパラメータから、音韻あるいは音節それぞれの波形振幅値を算出し、波形生成部 103 に出力する。

【0050】

音韻継続時間算出部 306 では、中間言語解析部 301 から入力された音韻文字列などのパラメータから音韻あるいは音節それぞれの継続時間を算出する。

【0051】

次に、母音無声化一時判定部 303 と母音無声化レベル判定部 304 における母音無声化判定処理について、フローチャートを参照して詳細に説明する。図 6 は母音無声化判定処理の具体例を示すフローチャートであり、図中、S T はフローの各処理ステップを示している。

【0052】

S T 7 までは母音無声化一時判定部 303 の処理に相当し、以降が母音無声化レベル判定部 304 の処理に相当する。S T 7 までの母音無声化一次判定部の処理は、従来技術における無声化判定と同様であり、一般の無声化規則にしたがって、着目母音に対して前後の音韻環境、アクセント核の有無などから一次の無声化判定を行う。

【0053】

まず、ステップ S T 1 では、入力された中間言語を音節単位に検索するための音節ポイント i を 0 に初期化し、ステップ S T 2 で第 i 番目の音節の母音の種類 (a, i, u, e, o) と、当該音節の子音の種類 (無声子音あるいは無音・有声子音)、および後続音節の子音の種類を、それぞれ V 1、C 1、C 2 に設定する。

【0054】

ステップ S T 3、S T 4 の処理は、母音が無声化する前提条件についての判定処理であ

10

20

30

40

50

り、ステップ S T 5 ～ S T 7 の処理は、ステップ S T 3, S T 4 に於ける母音無声化の前提条件を満たしても、無声化が起きない場合の判定処理である。

【 0 0 5 5 】

ステップ S T 3 では、当該音節の母音 V 1 が「 i 」または「 u 」であるかの判定を行う。当該音節の母音 V 1 が「 i 」または「 u 」であるときはステップ S T 4 に進み、そうでなければ、無声化の対象ではないと判断してステップ S T 1 5 に進む。

【 0 0 5 6 】

ステップ S T 4 では、当該音節の子音 C 1 が無声子音で、かつ、後続音節の子音 C 2 が無声子音か、もしくは文末・ポーズなのかを判定する。双方を満足すればステップ S T 5 に進み、該当音節にアクセント核が存在するか否かの判定を行う。日本語はピッチの高低変化でアクセントの位置を表現している。したがって高ピッチから低ピッチへの遷移が存在してアクセントの位置を示すアクセント核のある音節では、ピッチ構造のない無声化が起き難い。

【 0 0 5 7 】

ステップ S T 5 では、当該音節がアクセント核であるか否かを判定し、アクセント核であれば、無声化しないと判断してステップ S T 1 5 に進む。該当音節にアクセント核が存在しなければ（アクセント核でなければ）、ステップ S T 6 で前音節が無声化したか否かを判定する。無声化は、連続しては発生し難い性質があり、前音節が既に無声化していれば、後続する音節では通常、無声化は起こらない。

【 0 0 5 8 】

ステップ S T 7 では、該当音節が疑問文終端であるか否かを判定する。疑問文末はピッチの急激な上昇が起こるため無声化は発生しない。例えば、「～します」と「～します？」では後者の疑問文末の音節には明らかに強調の意図が含まれた発声になるため無声化は起こらない。

【 0 0 5 9 】

以上のステップ S T 3 ～ S T 7 の条件を満たし、S T 8 へと到達したものが無声化の候補となる。S T 8 以降、本発明の特徴である発声速度と音節の種別に応じた無声化レベルの判定をおこなう。

【 0 0 6 0 】

まず、無声化候補となった母音を含む音節の種類の切り分けをおこなう。好適には音節を細かく分類してそれぞれに規則化することも可能であるが、処理構造を理解し易くするため本実施形態では音節を 3 種に分類し、個別に発声速度による判定レベルを設定している。

【 0 0 6 1 】

前述のように、子音部 C 1 が摩擦音である音節では、発話速度が遅くなっても無声化して発音される場合が多い。子音部 C 1 が / p /, / t /, / k / などの破裂音では発声速度が遅くなると無声化しにくくなる。子音に関して、継続時間が伸びても発声時の開口面積が狭い状態が維持される摩擦音 / s /, / h /, / f / では、発声速度が遅くても無声化が起きる可能性が高い。この中では比較的、開口面積が広い / h /, / f / では、/ s / に比べれば無声化が起きにくい。破裂音は、破裂時に開口面積が極端に狭くなるが、その後は広くなる。開口面積を狭い状態で維持して発声することが困難であるため、発声速度が遅くなれば時間経過とともに開口面積が広がり声帯振動が開始する。

【 0 0 6 2 】

このような性質を反映し、本実施形態では、C 1 = / s / (S T 8, S T 1 0)、C 1 = / h /, / f / (S T 9, S T 1 1)、それ以外 (S T 1 2) とで、3 種の音節グループに分類し、それぞれの分類に対して、発声速度に応じた無声化判定レベル T H 1、T H 2 を設定する。ここで、T H 1、T H 2 は、モーラ数 / 分を単位とする発声速度の判定閾値を表し、後続ステップにおいて、T H 1 は、(無声化なし) ←→ (準無声化) の境界判定に、T H 2 は、(準無声化) ←→ (無声化) の境界判定に用いられる。

【 0 0 6 3 】

10

20

30

40

50

上記閾値は、具体的には、第一分類（ $C1 = /s/$ ）の場合、 $TH1 = Na1$ 、 $TH2 = Na2$ に設定し、第二分類（ $C1 = /h/$ 、 $/f/$ ）の場合、 $TH1 = Nb1$ 、 $TH2 = Nb2$ に設定し、第三分類（上記以外）の場合、 $TH1 = Nc1$ 、 $TH2 = Nc2$ に設定する。これらの各閾値は、各音節の性質から明らかなように、 $Na1 \leq Nb1 \leq Nc1$ 、 $Na2 \leq Nb2 \leq Nc2$ 、また $Na1 \leq Na2$ 、 $Nb1 \leq Nb2$ 、 $Nc1 \leq Nc2$ 、の関係に設定する。

【0064】

ST13からのステップでは、設定された発声速度判定閾値 $TH1$ 、 $TH2$ を用いて、無声化レベルを決定する。無声化レベルとは、発声速度に対して、無声化する／まったく無声化しない、の両極端に切り分けるのではなく、無声化の程度に応じて複数のレベルを設けたものである。本実施形態では、「無声化しない」と「完全に無声化する」の間に1レベルの中間的レベル「準無声化」を設け、合計3レベルとして説明する。 10

【0065】

無声化母音は、通常の母音とは異なり、音韻パワー0、音韻継続時間0として実現される。「準無声化」とは、無声化のように母音を無くすのではなく、パワーも継続時間も通常の母音より小さいながら、ある程度、母音を残すようにするものである。本実施形態では、音節ごとに母音の無声化レベルを表わすフラグ変数 $uvflag[i]$ を設け、発声速度判定閾値 $TH1$ 、 $TH2$ と発声速度 $level$ を比較することにより、（無声化なし： $uvflag[i] = 0$ ）、（準無声化： $uvflag[i] = 2$ ）、（無声化： $uvflag[i] = 1$ ）の何れかの母音無声化最終判定を行う。 20

【0066】

ステップST18で音節カウンタ i を1インクリメントし、ステップST19で音節カウンタ i が総モーラ数 sum_mora 以上か（ $i \geq sum_mora$ か）否かを判別し、 $i < sum_mora$ のときは処理が終了していないと判断してステップST2に戻って次音節の処理を同様に繰り返す。

【0067】

上述の処理は、入力テキスト全音節に対して行った後、すなわち、ステップST19で音節カウンタ i が総モーラ数 sum_mora を超えた時点で終了する。

【0068】

以上説明したように、この実施形態によれば、無声化の判定において、無声化のレベルを複数設け、発声速度レベルに応じて、無声化のレベルを段階的に割り当てるようにしたことにより、発声速度の境界点前後においても無声化の程度が急激に変化することなく、いかなる発声速度においても自然な無声化を実現できる。また、発声速度レベルに応じた無声化レベルの設定は、音節の種別ごとに適した設定則を与える構成にしているため、発声速度の変化に対して子音の特性を反映した自然な無声化が実現可能であり、自然で滑らかな合成音声を得ることができる。 30

【図面の簡単な説明】

【0069】

【図1】実施形態に係る音声合成装置のパラメータ生成部の構成を示すブロック図である。 40

【図2】テキスト音声変換処理の機能ブロック図である。

【図3】従来技術に於けるパラメータ生成部の構成を示す機能ブロック図である。

【図4】「取材した」の実発生波形を示す図である。

【図5】「規則」の実発生波形を示す図である。

【図6】実施形態に係る母音無声化判定のフローチャートである。

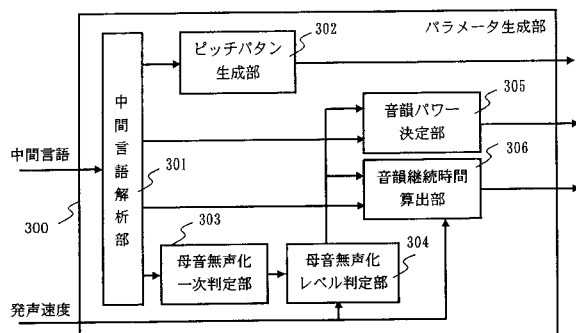
【符号の説明】

【0070】

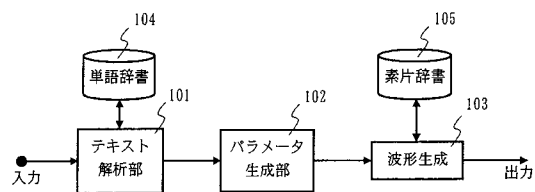
- 300 パラメータ生成部
- 301 中間言語解析部
- 302 ピッチパターン生成部

- 303 母音無声化一次判定部
- 304 母音無声化レベル判定部
- 305 母音パワー決定部
- 306 音韻継続時間算出部

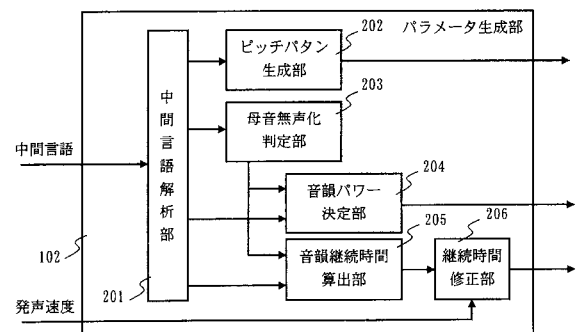
【図 1】



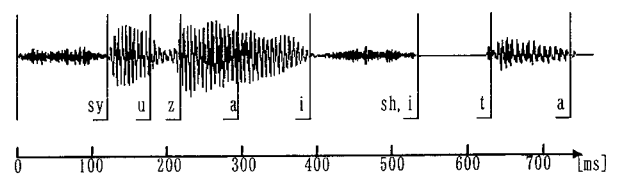
【図 2】



【図 3】

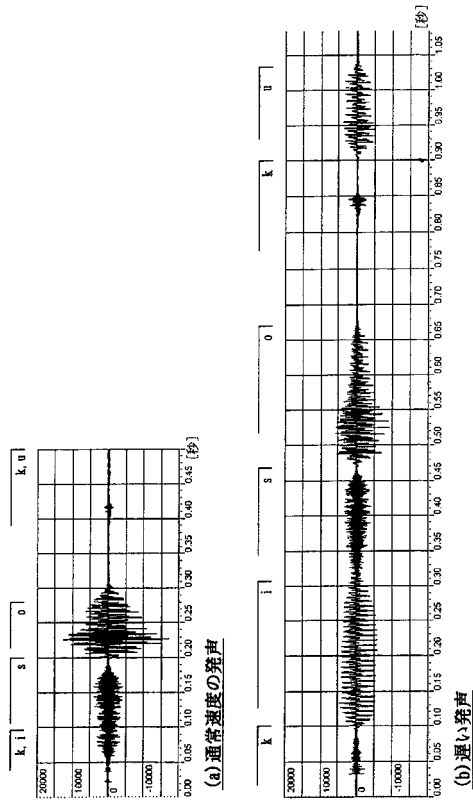


【図 4】



「取材した」の実発声波形

【図 5】



【図 6】

