

(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2009-139677

(P2009-139677A)

(43) 公開日 平成21年6月25日(2009.6.25)

(51) Int.Cl.	F I	テーマコード (参考)
G 1 0 L 13/08 (2006.01)	G 1 0 L 13/08 1 2 2	
G 1 0 L 13/00 (2006.01)	G 1 0 L 13/00 1 0 0 A	
G 1 0 L 13/02 (2006.01)	G 1 0 L 13/02 1 2 3	
	G 1 0 L 13/08 1 2 4	

審査請求 未請求 請求項の数 11 O L (全 16 頁)

(21) 出願番号 特願2007-316637 (P2007-316637)
 (22) 出願日 平成19年12月7日 (2007.12.7)

(71) 出願人 000003078
 株式会社東芝
 東京都港区芝浦一丁目1番1号
 (74) 代理人 100059225
 弁理士 葛田 璋子
 (74) 代理人 100076314
 弁理士 葛田 正人
 (74) 代理人 100112612
 弁理士 中村 哲士
 (74) 代理人 100112623
 弁理士 富田 克幸
 (72) 発明者 籠嶋 岳彦
 東京都港区芝浦一丁目1番1号 株式会社
 東芝内

最終頁に続く

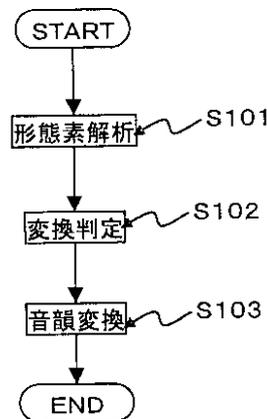
(54) 【発明の名称】 音声処理装置及びそのプログラム

(57) 【要約】

【課題】 文法的、音韻的、韻律的に言語らしさを保存しつつ意味が不明であるような合成音声を生成する。

【解決手段】 単語を表記する文字列と、単語の読みを表す音韻列と、単語の品詞との組から構成される単語辞書を参照して、入力テキストを解析し、入力テキストを1つ以上の部分テキストに分割し、部分テキスト毎に音韻列と品詞列とを生成し、この部分テキストの品詞列と、品詞列のリストとを照合して、部分テキストの音韻の変換を行うかどうかを判定し、音韻の変換を行うと判定された部分テキストの音韻列の各音韻を変換する。

【選択図】 図2



【特許請求の範囲】**【請求項 1】**

テキストを入力する入力部と、

単語を表記する文字列と、前記単語の読みを表す音韻列と、前記単語の品詞との組から構成される辞書と、

前記辞書に基づいて、前記テキストを 1 つ以上の部分テキストに分割し、分割した前記部分テキスト毎に音韻列を含む音声情報を生成する生成部と、

前記部分テキストの音声情報と、予め記憶された音声情報のリストとを照合して、前記部分テキストの前記音韻列に属する音韻の変換を行うかどうかを判定する判定部と、

(1) 前記音韻の変換を行うと判定された前記部分テキストの前記音韻列の前記各音韻を予め記憶した変換規則に従って異なる音韻に変換して出力し、(2) 前記音韻の変換を行わないと判定された前記部分テキストの前記音韻列は、無変換で出力する処理部と、

を備える音声処理装置。

【請求項 2】

テキスト、及び、前記テキストにおける各音韻のそれぞれについて、異なる音韻へ変換を行う部分と変換を行わない部分を表す判別情報を入力する入力部と、

単語を表記する文字列と、前記単語の読みを表す音韻列と、前記単語の品詞との組から構成される辞書と、

前記辞書と前記判別情報とに基づいて、前記テキストを 1 つ以上の部分テキストに分割し、分割した前記部分テキスト毎に、音韻列と前記変換の要否を表す属性とを生成する生成部と、

(1) 前記属性が変換が必要となっている前記部分テキストの前記音韻列の前記各音韻を、予め記憶した変換規則に基づいて、異なる音韻に変換して出力し、(2) 前記属性が変換が不要となっている前記部分テキストの前記音韻列は、無変換で出力する処理部と、

を備える音声処理装置。

【請求項 3】

テキストを入力する入力部と、

音韻の変換を行う単語について、前記単語を表記する文字列と、前記単語の読みを表す音韻の組合せが任意の変換規則に基づいて異なる音韻の組合せに変換された変換音韻列と、前記単語の品詞との組とから構成される第 1 の辞書と、

音韻の変換を行わない単語について、前記単語を表記する文字列と、前記単語の読みをそのまま表す無変換音韻列と、前記単語の品詞との組から構成される第 2 の辞書と、

(1) 前記第 1 の辞書と前記第 2 の辞書とに基づいて、前記テキストを 1 つ以上の部分テキストに分割し、(2) 前記第 1 の辞書に含まれる前記部分テキストは、前記第 1 の辞書に基づいて前記変換音韻列を生成して出力し、(3) 前記第 2 の辞書に含まれる前記部分テキストは、前記第 2 の辞書に基づいて前記無変換音韻列を生成して出力する処理部と、

を備える音声処理装置。

【請求項 4】

前記部分テキスト毎の音韻列に基づいて、前記音韻列の各音韻の継続時間及び声の高さで構成される韻律情報を生成する韻律生成部と、

前記部分テキスト毎の前記音韻列と前記韻律情報とから合成音声を生成する合成部と、をさらに備える請求項 1 から 3 の少なくとも一項に記載の音声処理装置。

【請求項 5】

前記音声情報が、文字列、音韻列、または、品詞列であり、

前記判定部は、

前記部分テキストの文字列が、予め記憶した文字列リスト中の文字列を含むかどうか、

前記部分テキストの音韻列が、予め記憶した音韻列リスト中の音韻列を含むかどうか、

または、前記部分テキストの品詞列が、予め記憶した品詞列リスト中の品詞列を含むかどうかのいずれかに基づいて、

10

20

30

40

50

前記部分テキストの前記音韻の変換を行うかどうかを判定する、
請求項 1 記載の音声処理装置。

【請求項 6】

前記処理部は、

前記変換規則を、変換元の音韻と変換先の音韻との組で表される音韻交換テーブル、または、変換元の音韻列内での音韻の位置と、変換先の音韻列内での音韻の位置との組で表される音韻置換テーブルに記憶している、

請求項 1 または 2 記載の音声処理装置。

【請求項 7】

前記部分テキストは、単語単位、形態素単位、または、アクセント句単位である、

請求項 1 ~ 3 の少なくとも一項に記載の音声処理装置。

【請求項 8】

前記音韻は、音節単位、モーラ単位、または、音素単位である、

請求項 1 ~ 3 の少なくとも一項に記載の音声処理装置。

【請求項 9】

単語を表記する文字列と、前記単語の読みを表す音韻列と、前記単語の品詞との組から構成される辞書を有し、

テキストが入力する入力機能と、

前記辞書に基づいて、前記テキストを 1 つ以上の部分テキストに分割し、分割した前記部分テキスト毎に音韻列を含む音声情報を生成する生成機能と、

前記部分テキストの音声情報と、予め記憶された音声情報のリストとを照合して、前記部分テキストの前記音韻列に属する音韻の変換を行うかどうかを判定する判定機能と、

(1) 前記音韻の変換を行うと判定された前記部分テキストの前記音韻列の前記各音韻を予め記憶した変換規則に従って異なる音韻に変換して出力し、(2) 前記音韻の変換を行わないと判定された前記部分テキストの前記音韻列は、無変換で出力する処理機能と、
をコンピュータによって実現する音声処理プログラム。

【請求項 10】

単語を表記する文字列と、前記単語の読みを表す音韻列と、前記単語の品詞との組から構成される辞書を有し、

テキスト、及び、前記テキストにおける各音韻のそれぞれについて、異なる音韻へ変換を行う部分と変換を行わない部分を表す判別情報が入力する入力機能と、

前記辞書と前記判別情報とに基づいて、前記テキストを 1 つ以上の部分テキストに分割し、分割した前記部分テキスト毎に、音韻列と前記変換の要否を表す属性とを生成する生成機能と、

(1) 前記属性が変換が必要となっている前記部分テキストの前記音韻列の前記各音韻を、予め記憶した変換規則に基づいて、異なる音韻に変換して出力し、(2) 前記属性が変換が不要となっている前記部分テキストの前記音韻列は、無変換で出力する処理機能と

をコンピュータによって実現する音声処理プログラム。

【請求項 11】

音韻の変換を行う単語について、前記単語を表記する文字列と、前記単語の読みを表す音韻の組合せが任意の変換規則に基づいて異なる音韻の組合せに変換された変換音韻列と、前記単語の品詞との組とから構成される第 1 の辞書と、

音韻の変換を行わない単語について、前記単語を表記する文字列と、前記単語の読みをそのまま表す無変換音韻列と、前記単語の品詞との組から構成される第 2 の辞書と、

を有し、

テキストを入力する入力機能と、

(1) 前記第 1 の辞書と前記第 2 の辞書とに基づいて、前記テキストを 1 つ以上の部分テキストに分割し、(2) 前記第 1 の辞書に含まれる前記部分テキストは、前記第 1 の辞書に基づいて前記変換音韻列を生成して出力し、(3) 前記第 2 の辞書に含まれる前記部

10

20

30

40

50

分テキストは、前記第2の辞書に基づいて前記無変換音韻列を生成して出力する処理機能と、

をコンピュータによって実現する音声処理プログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、任意のテキストから音声を合成する音声合成装置に係わり、特に、ビデオゲームなどのエンターテインメント応用のための音声処理装置に関する。

【背景技術】

【0002】

従来から、任意の文章(テキスト)から人工的に音声信号を作り出すテキスト音声合成の技術が提案されている。このようなテキスト音声合成を実現する音声合成装置は、一般に言語処理部、韻律処理部及び音声合成部の3つの要素によって構成される。

【0003】

この音声合成装置の動作は次の通りである。

【0004】

まず、言語処理部において、入力されたテキストの形態素解析や構文解析などが行われ、テキストを形態素、単語、アクセント句などの単位に区切ると共に、各単位の音韻列や品詞列などを生成する。

【0005】

次に、韻律処理部においてアクセントやイントネーションの処理が行われ、基本周波数及び音韻継続時間長などの情報が算出される。

【0006】

最後に、音声合成部において、予め合成音声を生成する際の音声の接続単位である合成単位(例えば、音素や音節など)毎に記憶されている音声素片データと呼ばれる特徴パラメータや音声波形を、韻律処理部で算出された基本周波数や音韻継続時間長などに基づいて接続することで合成音声が生成される。

【0007】

このようなテキスト音声合成技術は、ビデオゲームのキャラクタの音声メッセージ出力にも用いられている(例えば、特許文献1参照)。従来の録音音声の再生による音声メッセージ出力では、予め録音しておいた言葉しか発声することができなかったが、テキスト音声合成を用いることにより、プレイヤーが入力した名前など、事前の録音が不可能な言葉も発声することが可能となった。

【特許文献1】特開2001-34282号公報

【発明の開示】

【発明が解決しようとする課題】

【0008】

上記したように、ビデオゲームのキャラクタ、特に人間や人間型ロボットなどのキャラクタの音声メッセージには、テキスト音声合成を用いることができる。

【0009】

しかしながら、ゲームに登場する様々なキャラクタの中には、人間と同じ言語(例えば日本語)を話すことが適当でない場合がある。例えば「知能の発達したエイリアン」のような設定のキャラクタの場合、言葉を話すことは合理的だが、それが日本語や他の実在する言語では真実味に欠けるといった問題点がある。

【0010】

このときに音声の代わりに、無意味な効果音で代用することも可能であるが、この場合は言語らしくなく真実味に欠けるといった問題点がある。

【0011】

そこで本発明は、意味は不明であるが、言語らしく真実味のある音声合成に用いることができる音韻列を生成する音声処理装置を提供する。

10

20

30

40

50

【課題を解決するための手段】

【0012】

本発明は、テキストを入力する入力部と、単語を表記する文字列と、前記単語の読みを表す音韻列と、前記単語の品詞との組から構成される辞書と、前記辞書に基づいて、前記テキストを1つ以上の部分テキストに分割し、分割した前記部分テキスト毎に音韻列を含む音声情報を生成する生成部と、前記部分テキストの音声情報と、予め記憶された音声情報のリストとを照合して、前記部分テキストの前記音韻列に属する音韻の変換を行うかどうかを判定する判定部と、(1)前記音韻の変換を行うと判定された前記部分テキストの前記音韻列の前記各音韻を予め記憶した変換規則に従って異なる音韻に変換して出力し、(2)前記音韻の変換を行わないと判定された前記部分テキストの前記音韻列は、無変換で出力する処理部と、を備える音声処理装置である。

10

【0013】

また、本発明は、テキスト、及び、前記テキストにおける各音韻のそれぞれについて、異なる音韻へ変換を行う部分と変換を行わない部分を表す判別情報を入力する入力部と、単語を表記する文字列と、前記単語の読みを表す音韻列と、前記単語の品詞との組から構成される辞書と、前記辞書と前記判別情報とに基づいて、前記テキストを1つ以上の部分テキストに分割し、分割した前記部分テキスト毎に、音韻列と前記変換の要否を表す属性とを生成する生成部と、(1)前記属性が変換が必要となっている前記部分テキストの前記音韻列の前記各音韻を、予め記憶した変換規則に基づいて、異なる音韻に変換して出力し、(2)前記属性が変換が不要となっている前記部分テキストの前記音韻列は、無変換で出力する処理部と、を備える音声処理装置である。

20

【0014】

さらに、本発明は、テキストを入力する入力部と、音韻の変換を行う単語について、前記単語を表記する文字列と、前記単語の読みを表す音韻の組合せが任意の変換規則に基づいて異なる音韻の組合せに変換された変換音韻列と、前記単語の品詞との組とから構成される第1の辞書と、音韻の変換を行わない単語について、前記単語を表記する文字列と、前記単語の読みをそのまま表す無変換音韻列と、前記単語の品詞との組から構成される第2の辞書と、(1)前記第1の辞書と前記第2の辞書とに基づいて、前記テキストを1つ以上の部分テキストに分割し、(2)前記第1の辞書に含まれる前記部分テキストは、前記第1の辞書に基づいて前記変換音韻列を生成して出力し、(3)前記第2の辞書に含まれる前記部分テキストは、前記第2の辞書に基づいて前記無変換音韻列を生成して出力する処理部と、を備える音声処理装置である。

30

【発明の効果】

【0015】

本発明によれば、文法的、音韻的、韻律的に言語らしさを保存しつつ意味が不明であるような合成音声を生成できる。

【発明を実施するための最良の形態】

【0016】

以下、本発明の一実施形態の音声合成装置について説明する。

【0017】

(第1の実施形態)

第1の実施形態の音声合成装置について図1～図7に基づいて説明する。

40

【0018】

(1)音声合成装置の構成

本実施形態の音声合成装置の構成について図1に基づいて説明する。図1は、音声合成装置を示すブロック図である。

【0019】

音声合成装置は、テキストを入力するテキスト入力部101と、テキスト入力部101で入力されたテキストから単語毎の音韻列や品詞を生成する音韻列生成部109と、それらの情報から各音韻の声の高さと継続時間長などの韻律情報を生成する韻律処理部103

50

と、音韻列と韻律情報とから合成音声を生成する音声合成部 104 と、音声合成部 104 で生成された合成音声を入力する音声出力部 105 とを備えている。

【0020】

なお、この音声合成装置は、例えば、汎用のコンピュータ装置を基本ハードウェアとして用いることでも実現することが可能である。すなわち、音韻生成部 109、韻律処理部 103、音声合成部 104 は、上記のコンピュータ装置に搭載されたプロセッサにプログラムを実行させることにより実現することができる。このとき、音声合成装置は、上記のプログラムをコンピュータ装置に予めインストールすることで実現してもよいし、CD-ROMなどの記憶媒体に記憶して、あるいはネットワークを介して上記のプログラムを配布して、このプログラムをコンピュータ装置に適宜インストールすることで実現してもよい。また、テキスト入力部 101 は、上記コンピュータ装置に内蔵あるいは外付けされたキーボードなどを適宜利用して実現することができる。また、音声出力部 105 は、上記コンピュータ装置に内蔵あるいは外付けされたスピーカやヘッドホンなどを適宜利用して実現することができる。

10

【0021】

(2) 韻律処理部 103、音声合成部 104

韻律処理部 103 及び音声合成部 104 は、従来からある公知の韻律処理手法及び音声合成手法をそれぞれ用いて実現することができる。

【0022】

例えば、韻律処理における声の高さの生成には、典型的なアクセント句単位の声の高さの変化パターンを選択、接続して1文の声の高さの変化パターンを生成する方法、音韻の継続時間長の生成には、数量化1類による推定モデルを用いる方法などがある。

20

【0023】

音声合成手法には、音素単位や音節単位の音声波形（音声素片）を音韻列にしたがって選択し、韻律情報にしたがって韻律を変形して接続する方法などがある。

【0024】

(3) 音韻列生成部 109 の構成

次に、音韻列生成部 109 について図 1 に基づいて説明する。

【0025】

音韻列生成部 109 は、図 1 に示すように、言語処理部 102、言語辞書記憶部 107、音韻変換部 106、無変換リスト記憶部 108、変換規則記憶部 110 から構成されている。

30

【0026】

言語辞書記憶部 107 は、多数の日本語の単語の情報を記憶しており、各単語の情報は、漢字かな混じりの表記（文字列）、読みを表す音韻列、品詞、活用、アクセント位置などから構成されている。

【0027】

言語処理部 102 は、言語辞書記憶部 107 に記憶されている単語情報を参照して入力テキストを解析し、入力テキストを単語に区切ると共に、各単語の音韻列、品詞、アクセント位置などの音声情報を入力する。

40

【0028】

音韻変換部 106 は、無変換リスト記憶部 108 に記憶されている音声情報のリストを参照して、前記単語の音韻列の変換を行うか否かを判定し、変換を行うと判定された場合には、変換規則記憶部 110 に記憶されている変換規則に従って前記単語の音韻列の変換を行い、変換された音韻列を入力する。

【0029】

(4) 音韻列生成部 109 の動作

次に、音韻列生成部 109 の詳細な動作について図 2 ~ 図 7 に基づいて説明する。図 2 は、音韻列生成部 109 の動作を示すフローチャートである。

【0030】

50

(4-1) 言語処理部 102

言語処理部 102 では、テキスト入力部 101 で入力されたテキストの形態素解析が行なわれる (ステップ S101)。例として「太郎さんお早う」というテキストの解析について説明する。

【0031】

まず、言語辞書記憶部 107 の単語情報を参照して、入力テキストを単語列で表現する。単語列は 1 通りに決定されるとは限らず、例えば図 3 に表されるようなネットワークで表現される。この例では、単語「さん」に接尾と数詞の 2 通りがあるため、2 通りの解析結果がありうることを表している。

【0032】

次に、単語の品詞などを用いた、単語間の接続のし易さについてのルールを参照して、解析結果の候補 (ネットワークのパス) に点数付けを行う。

【0033】

最後に、各候補の点数を比較して、最も確からしいパスを選択し、各単語の文字列、音韻列、品詞を解析結果として出力する。この例では、固有名詞と接尾は接続し易いため、図 4 の結果が出力される。

【0034】

(4-2) 音韻変換部 106

次に、音韻変換部 106 では、形態素解析の結果を参照して、各単語の音韻の変換を行うか否かを判定する (ステップ S102)。

【0035】

判定は、無変換リスト記憶部 108 に記憶されている音声情報リストに基づいて行われる。音声情報リストは、音声情報を要素とするリストである。また、音声情報とは入力テキストを単語に区切ると共に、単語情報を参照して解析した結果として単語毎に得られる情報であり、例えば、音韻列・文字列・品詞・アクセント位置などがある。いずれか 1 種類 (例えば、文字列) のリストとしてもよいし、複数種類が混在したリスト (例えば文字列と品詞) としてもよい。あるいは、「文字列が『千葉』で品詞が『人名』」のように、複数種類の組合せを要素とするリストとしてもよい。音声情報リストが、文字列リストである場合の例を図 5 に示す。

【0036】

入力された単語列の各単語の文字列を、文字列リストと照合し、一致するものがある場合は前記単語の音韻変換は行わず、一致するものが無い場合は音韻変換を行うものと判定する。この例では、単語「太郎」は文字列リストに存在するため変換は行わず、「さん」「お早う」は存在しないため変換を行うものと判定する。

【0037】

次に、変換を行うと判定された単語について、変換規則 110 に記憶されている変換規則に従って音韻の変換を行う (ステップ S103)。

【0038】

音韻の変換とは、少なくとも入力された音韻と変換規則とに基づいて、入力音韻とは異なる音韻を出力する操作である。ここで、変換規則とは少なくとも入力された音韻を、入力された音韻とは異なる音韻に変換する際に用いるもので、ある入力された音韻を異なる音韻に変換する規則を表したものである。

【0039】

本実施形態における音韻の変換は、単語内での音韻の位置を置換することによって実現する。変換規則の例を図 6 に示す。このテーブルは、入力の単語内の音韻の位置と、置換された出力での音韻の位置の関係を表しており、N は単語の音韻の数である。この変換規則を用いて、単語「さん」及び「お早う」の音韻列を変換した出力を図 7 に示す。

【0040】

(5) 効果

本実施形態の音声合成装置では、「太郎さんお早う」というテキスト入力に対して、「

10

20

30

40

50

タローンサハヨーオ」という音声合成される。

【0041】

このように、音韻や抑揚は日本語と同じ特徴を持つことから、意味不明でありながら「言葉らしさ」を備えた音声合成することが可能で、ゲームのキャラクタの音声に利用することができる。

【0042】

また、人名などは、言語が異なっても同じように発音されることから、プレイヤーが入力した名前など、特定の単語は変換しないようにすることで、より現実味が増すという効果がある。

【0043】

また、用いる変換の方法によっては、変換前のテキストを類推することができ、ゲームのキャラクタのセリフの意味を推理するという娯楽性を提供することができる。

【0044】

(6) 変更例

本実施形態の音韻変換部106では、文字列リストを参照して変換するか否かを判定したが、判定方法はこれに限られるものではなく、音韻列リストや品詞リストを参照するようにしてもよい。

【0045】

例えば、音韻列リストに「ヒロシ」という登録があれば、入力テキストの「博」「浩」「寛」などは、全て変換されずにそのままの音韻で合成される。

【0046】

また、品詞リストに「固有名詞」という登録があれば、人名などの固有名詞は全て変換されない。ゲームの入力インターフェースで漢字入力ができず、仮名入力のみの場合、音韻列で照合の方が実装が容易となる。

【0047】

また、品詞で変換の判定を制御することにより、変換部分の割合を容易に制御することが可能で、例えば無変換リストの品詞を増やしていくことで、変換部分をだんだんと少なくし、「キャラクタが日本語を覚えてきた」という演出できる。

【0048】

(第2の実施形態)

次に、本発明の第2の実施形態の音声合成装置について、図8～図12に基づいて説明する。

【0049】

(1) 音声合成装置の構成

図8は、音声合成装置を示すブロック図であり、図1と同様の機能を持つ構成要素には同一符号を付与して説明を省略する。

【0050】

本実施形態の音声合成装置には、テキスト合成部201、変換文記憶部203、無変換文記憶部204が付加されている。

【0051】

変換文記憶部203には、音韻の変換を行うテキストが記憶されており、無変換文記憶部104には、音韻の変換を行わないテキストが記憶されている。例えば、ゲームキャラクタのセリフのうち、既定の部分のテキストは予め変換文記憶部203に記憶されており、プレイヤーが入力した名前などが無変換文記憶部に登録される。

【0052】

(2) 音声合成装置の動作

次に、本実施形態の音声合成装置における音韻生成部209の詳細な動作について図9～図11に基づいて説明する。図11は、音韻生成部209の動作を示すフローチャートである。

【0053】

10

20

30

40

50

(2-1) テキスト合成部 201

テキスト合成部 201 は、変換文記憶部 203 と無変換文記憶部 204 の中の指定されたテキストを組み合わせて入力テキストを生成する (ステップ S201)。

【0054】

さらに、入力テキストの中で、音韻を変換する部分と変換しない部分を表す情報である判別情報を生成する (ステップ S202)。

【0055】

判別情報は、入力テキストにタグとして挿入したり、変換、無変換の境界位置と各区間の変換、無変換の別を表すデータを入力テキストとは別に出力したりするなどの実現方法がある。

【0056】

例えば、図 9 で表されるようなテキストのリストが変換文記憶部 203 に記憶されており、図 10 で表されるようなテキストのリストが無変換文記憶部 104 に記憶されている場合について説明する。

【0057】

図 9 の [可変部分] に、図 10 で指定されたテキストを挿入することにより、入力テキストを生成する。図 9 から「[可変部分]さんお早う」が、図 10 から「太郎」が指定された場合は、これらを組み合わせた結果「<無変換>太郎</無変換>さんお早う」という入力テキストが生成される。ここで、<無変換>及び</無変換>は、入力テキストの中で音韻の変換を行わない区間の始めと終わりをそれぞれ表すタグである。無変換区間ではなく、変換区間を表すタグを用いても良い。

【0058】

また、タグの代わりに、「1文字目から2文字の長さの区間が無変換区間」という情報を変換部分判定情報として出力するようにしても良い。

【0059】

(2-2) 言語処理部 202

次に、言語処理部 202 では、第 1 の実施形態における形態素解析 (ステップ S102) と同様に、入力テキストを単語に分割し、各単語の文字列、音韻列、品詞を生成する。

【0060】

さらに、変換部分判定情報を参照して、各単語に変換、無変換の属性を付与する。言語処理部 202 の出力の例を図 12 に示す。

【0061】

(2-3) 音韻変換部 206

次に、音韻変換部 206 では、言語処理部 202 の出力の変換、無変換の属性を参照して、音韻の変換を行う単語を決定する (ステップ S204)。

【0062】

次に、音韻の変換を行うと決定された単語に対して、変換規則 110 に記憶されている変換規則に従って音韻の変換を行う (ステップ S205)。

【0063】

音韻の変換は、第 1 の実施形態と同様に、単語内での音韻の位置を置換することによって実現する。入力テキストが、「<無変換>太郎</無変換>さんお早う」である場合、生成された音韻列は「タローンサハヨーオ」となる。

【0064】

さらに、この音韻列に基づいて韻律処理部 103 で韻律情報が生成され、音声合成部 104 で「タローンサハヨーオ」という合成音声生成されて、合成音声出力部 105 から出力される。

【0065】

(3) 効果

本実施形態の音声合成装置でも、「太郎さんお早う」というテキストに対して、「タローンサハヨーオ」という音声生成され、第 1 の実施形態と同様の効果がある。

10

20

30

40

50

【0066】

(第3の実施形態)

次に、本発明の第3の実施形態の音声合成装置について、図13～図16に基づいて説明する。

【0067】

(1) 音声合成装置の構成

本実施形態の音声合成装置の構成について図13に基づいて説明する。図13は、音声合成装置を示すブロック図であり、図1及び図8と同様の機能を持つ構成要素には同一符号を付与して説明を省略する。

【0068】

本実施形態の音韻列生成部309は、言語処理部302、変換言語辞書記憶部307、無変換言語辞書記憶部308、音韻変換部306、変換規則記憶部110、言語辞書記憶部107から構成されている。

【0069】

言語処理部302は、変換言語辞書記憶部307と無変換言語辞書記憶部308の2つの言語辞書を参照して動作する。変換言語辞書記憶部307に記憶されている単語の情報は、言語辞書記憶部107と同様であるが、音韻列情報は予め変換規則に基づいて変換されたものとなっている。

【0070】

すなわち、音韻変換部306は、言語辞書記憶部107の全ての単語について、音韻列情報を変換規則記憶部110に記憶されている変換規則に基づいて変換し、変換した音韻列とそのほかの情報(文字列、品詞、活用、アクセント位置など)を変換言語辞書記憶部307に記憶する。

【0071】

(2) 音声合成装置の動作

次に、本実施形態の音声合成装置の動作について説明する。

【0072】

言語辞書記憶部107に記憶されている単語情報の例を図14(a)に示す。また、変換規則記憶部110には、図5で表される音韻入換えテーブルが記憶されている。

【0073】

(2-1) 音韻変換部306

音韻変換部306は、音韻入換えテーブルに基づいて言語辞書記憶部107の音韻列を変換して図14(b)で表される単語情報を生成し、変換言語辞書記憶部307に記憶する。

【0074】

無変換言語辞書記憶部308には、図14(c)で表される単語情報が記憶されているものとする。

【0075】

(2-2) 言語処理部302

言語処理部302は、テキスト入力部101より「太郎さんお早う」というテキストが入力されたとすると、第1の実施形態の言語処理部102と同様に形態素解析処理を行って、各単語の文字列、音韻列、品詞列を解析結果として出力する。但し、本実施形態の言語処理部302は、変換言語辞書記憶部307と、無変換言語辞書記憶部308の2つの言語辞書を参照する。

【0076】

もし、同一文字列の単語が2つの辞書の両方に存在した場合は、無変換言語辞書記憶部308の登録内容を優先して解析に用いるものとする。

【0077】

その結果、図15で表される解析結果が出力される。出力された音韻列は、「タローンサハヨーオ」となる。

10

20

30

40

50

【0078】

(2-3) 韻律処理部103

さらに、韻律処理部103では、この音韻列に基づいて韻律情報が生成され、音声合成部104で「タローンサハヨーオ」という合成音声で生成されて、合成音声出力部105から出力される。

【0079】

(3) 効果

本実施形態の音声合成装置でも、「太郎さんお早う」というテキストに対して、「タローンサハヨーオ」という音声で合成され、第1の実施形態と同様の効果がある。

【0080】

(変更例)

本発明は上記各実施形態に限らず、その主旨を逸脱しない限り種々に変更することができる。

【0081】

(1) 変更例1

上記各実施形態では、音韻の変換は単語内の音韻の位置の置換によって実現するものとして説明したが、その他の変換規則を用いても良い。

【0082】

例えば、図16(a)で表されるような音韻の変換テーブルを用いても良い。これは、入力音韻を出力音韻に変換することを意味しており、音韻の対で構成されている。

【0083】

また、音韻の置換、変換のいずれの場合においても、変換のテーブルは固定である必要は無く、例えば複数のテーブルを切り替えて用いるようにしてもよい。

【0084】

また、これらのテーブルは、入力に対して出力が常に一意に決定される必要は無く、例えば図16(b)のテーブルのように、入力音韻1つに対して複数の出力音韻が対応し、出力が周期的に変化するようにしても良い。この例では、「あ」の入力に対しては、「い」と「お」が交互に出力されることになる。

【0085】

また、必ずしも周期的に変化する必要は無く、図16(c)のテーブルのように、1つの入力音韻に対応する複数の出力音韻に出力確率が付与されており、確率的に出力が決定されるようにしてもよい。この例では、「あ」の入力に対しては、「い」と「お」がそれぞれ50%の確率で出力されることを表している。

【0086】

このように、音韻の変換の方法に応じて、変換された合成音声から、元のテキストを類推できる度合いが変化するため、ゲームのキャラクタの設定や進行状況に適した変換を行うことができるという効果がある。

【0087】

(2) 変更例2

また、上記各実施形態では、言語処理部102における処理の結果、単語の列が出力されるものとして説明したが、これに限られるものではなく、例えば形態素やアクセント句などの単位で出力するようにしても良い。

【0088】

第1の実施形態において、単位をアクセント句とした例を図17に示す。

【0089】

無変換リストの登録は「太郎」であり、アクセント句の文字列「太郎さん」とは完全には一致しないが、この場合は無変換リストの登録単語を含んでいる場合に交換しないものと判定したため、アクセント句「太郎さん」全体を交換していない。

【0090】

また、複数の単語から構成されるアクセント句の場合は、1アクセント句に複数の品詞

10

20

30

40

50

が割り当てられる場合があるため、品詞の無変換リストによって判定する場合は、リストへの登録を品詞列（例えば「固有名詞＋接尾」）としてアクセント句の品詞列と一致するかどうかを判定しても良いし、文字列と同様に、リストへの登録は一つの品詞とし、アクセント句の品詞列に含まれるかどうかによって判定するようにしてもよい。

【0091】

(3) 変更例3

また、上記各実施形態では、音韻は音節であるとして説明したが、これに限定されるものではなく、例えば音韻としてモーラや音素などの単位を用いてもよい。

【0092】

音素を単位とした場合、日本語では連続しない子音が変換によって連続する場合があります。10
、外国語のような雰囲気を出すことができる。

【図面の簡単な説明】

【0093】

【図1】本発明の第1の実施形態の音声合成装置を示すブロック図である。

【図2】音韻生成部の動作を示すフローチャートである。

【図3】単語列を表すネットワークである。

【図4】各単語の文字列、音韻列、品詞の解析結果の例である。

【図5】無変換リスト記憶部に記憶されている文字列リストの例である。

【図6】変換規則の例である。

【図7】音韻列を変換した出力の例である。 20

【図8】第2の実施形態の音声合成装置を示すブロック図である。

【図9】変換文記憶部に記憶されているテキストのリストである。

【図10】無変換文記憶部に記憶されているテキストのリストである。

【図11】音韻生成部の動作を示すフローチャートである。

【図12】言語処理部の出力の例を示す図である。

【図13】第3の実施形態の音声合成装置を示すブロック図である。

【図14】(a)は言語辞書記憶部に記憶されている単語情報の例であり、(b)は音韻変換部が音韻入換えテーブルに基づいて言語辞書記憶部の音韻列を変換した例であり、(c)は無変換言語辞書記憶部に記憶されている単語情報の例である。

【図15】解析結果の出力の例である。 30

【図16】変更例1における変換テーブルである。

【図17】変更例2における単位をアクセント句としたテーブルである。

【符号の説明】

【0094】

101 テキスト入力部

102 言語処理部

103 韻律処理部

104 音声合成部

105 合成音声出力部

107 言語辞書記憶部 40

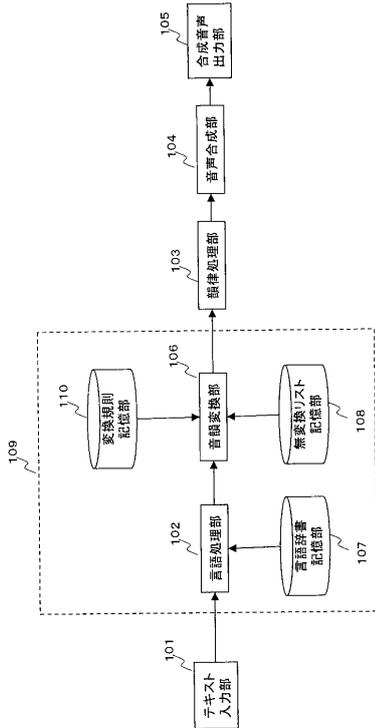
106 音韻変換部

108 無変換リスト記憶部

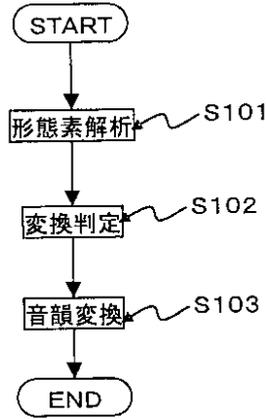
109 音韻列生成部

110 変換規則記憶部

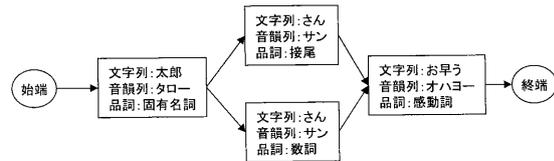
【図1】



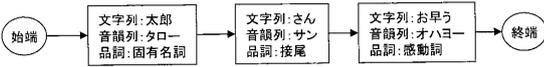
【図2】



【図3】



【図4】



【図6】

音韻入換えテーブル

入力位置	出力位置
1	N
2	1
3	2
4	3
5	4
6	5
.	.
.	.
.	.

【図5】

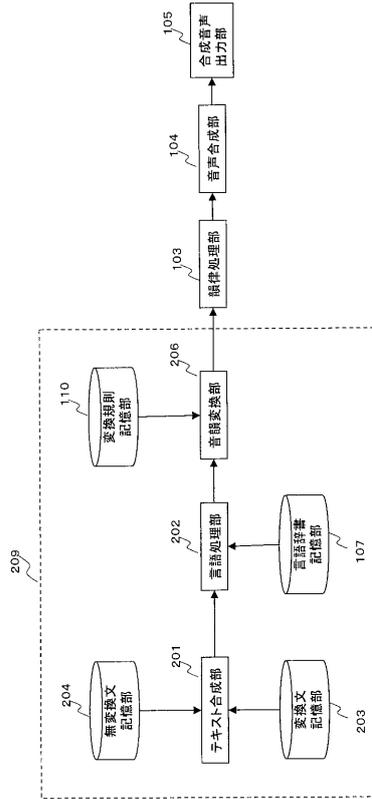
文字列リスト

太郎
次郎
花子

【 図 7 】

テキスト	太郎さんお早う		
文字列	太郎	さん	お早う
音韻列	タロー	サン	オハヨー
品詞	固有名詞	接尾	感動詞
音韻変換	無し	有り	有り
出力音韻列	タロー	ンサ	ハヨーオ

【 図 8 】



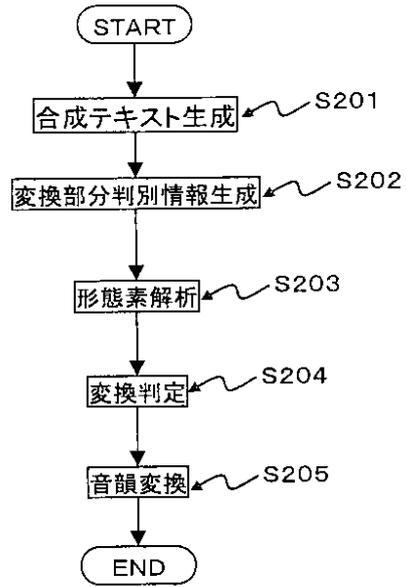
【 図 9 】

[可変部分]さんお早う
今日は[可変部分]に行きましょう
お疲れ様でした
いい天気ですね

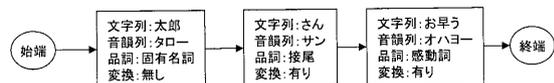
【 図 1 1 】

【 図 1 0 】

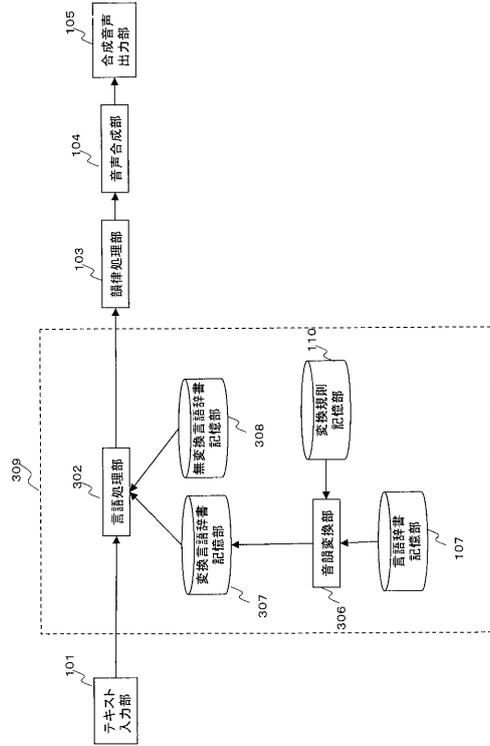
太郎
太郎ランド
花子



【 図 1 2 】



【 図 1 3 】



【 図 1 4 】

文字列	音韻列	品詞
太郎	タロー	固有名詞
さん	サン	固有名詞
花子	ハナコ	固有名詞

(c)

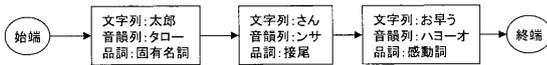
文字列	音韻列	品詞
さん	ンサ	接尾
君	ンク	接尾
様	マサ	接尾
お早う	ハヨ一オ	感動詞
今日は	ンニチワ	感動詞
今晩は	ンバンワ	感動詞

(b)

文字列	音韻列	品詞
さん	サン	接尾
君	ンク	接尾
様	マサ	接尾
お早う	ハヨ一オ	感動詞
今日は	ンニチワ	感動詞
今晩は	ンバンワ	感動詞

(a)

【 図 1 5 】



【 図 1 6 】

音韻変換テーブル

入力音節	出力音節
あ	い
い	お
う	あ
え	あ
お	え
か	し
き	そ
く	さ
け	ず
こ	せ
さ	ち
し	と
ず	た
せ	つ
そ	て
.	.
.	.
.	.
ん	ん

(a)

音韻変換テーブル

入力音節	出力音節
あ	いーお
い	おーえ
う	あーい
え	うーあ
お	えーう
か	し
き	そ
く	さ
け	ず
こ	せ
さ	ち
し	と
ず	た
せ	つ
そ	て
.	.
.	.
.	.
ん	ん

(b)

音韻変換テーブル

入力音節	出力音節
あ	い(0.5)お(0.5)
い	お(0.7)え(0.3)
う	あ(0.9)い(0.1)
え	う(0.6)あ(0.4)
お	え(0.8)う(0.2)
か	し
き	そ
く	さ
け	ず
こ	せ
さ	ち
し	と
ず	た
せ	つ
そ	て
.	.
.	.
.	.
ん	ん

(c)

【 図 1 7 】

テキスト	太郎さんお早う	
文字列	太郎さん	お早う
音韻列	タローサン	オハヨ一
品詞	固有名詞+接尾	感動詞
音韻変換	無し	有り
出力音韻列	タローサン	ハヨ一オ

フロントページの続き

- (72)発明者 山中 紀子
東京都港区芝浦一丁目1番1号 株式会社東芝内
- (72)発明者 矢島 真人
東京都港区芝浦一丁目1番1号 株式会社東芝内