

(19) 日本国特許庁(JP)

(12) 特許公報(B2)

(11) 特許番号

特許第6002598号  
(P6002598)

(45) 発行日 平成28年10月5日(2016.10.5)

(24) 登録日 平成28年9月9日(2016.9.9)

(51) Int.Cl. F I  
G 1 O L 13/10 (2013.01) G 1 O L 13/10 1 1 1 B

請求項の数 9 (全 10 頁)

(21) 出願番号	特願2013-32129 (P2013-32129)	(73) 特許権者	000004226
(22) 出願日	平成25年2月21日 (2013. 2. 21)		日本電信電話株式会社
(65) 公開番号	特開2014-163978 (P2014-163978A)		東京都千代田区大手町一丁目5番1号
(43) 公開日	平成26年9月8日 (2014. 9. 8)	(74) 代理人	100121706
審査請求日	平成27年1月14日 (2015. 1. 14)		弁理士 中尾 直樹
		(74) 代理人	100128705
			弁理士 中村 幸雄
		(74) 代理人	100147773
			弁理士 義村 宗洋
		(72) 発明者	中嶋 秀治
			東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内
		(72) 発明者	水野 秀之
			東京都千代田区大手町二丁目3番1号 日 本電信電話株式会社内

最終頁に続く

(54) 【発明の名称】 強調位置予測装置、その方法、およびプログラム

(57) 【特許請求の範囲】

【請求項1】

テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量と前記テキストの強調位置との関係を表す強調位置予測モデルを格納する強調位置予測モデル格納部と、

入力テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量を得る特徴量構成部と、

前記強調位置予測モデルおよび前記特徴量構成部で得られた前記特徴量を用いて前記入力テキストの強調位置を識別する強調位置予測部と、を有し、

前記特徴量は、主辞の単語の出現形を表す情報および前記主辞の単語の品詞を表す情報を含む、または、前記主辞の単語および前記主辞の単語の前後の所定個数の単語のそれぞれの出現形を表す情報ならびに前記単語のそれぞれの品詞を表す情報を含む、

ことを特徴とする強調位置予測装置。

【請求項2】

請求項1の強調位置予測装置であって、

前記特徴量は、前記主辞の単語から、当該主辞の単語に係り受けで係ってくる前側の単語まで、の距離の最小値を表す情報を含む、

ことを特徴とする強調位置予測装置。

【請求項3】

請求項1または2の強調位置予測装置であって、

10

20

前記特徴量は、単語のカテゴリを表す情報を含む、  
ことを特徴とする強調位置予測装置。

【請求項 4】

請求項 1 から 3 の何れかの強調位置予測装置であって、  
前記強調位置予測モデルは、前記テキストを構成する区間が強調位置であるかと前記特徴量との関係を表すモデルであり、  
前記強調位置予測部は、前記入力テキストを構成する各区間が強調位置であることを識別する、  
ことを特徴とする強調位置予測装置。

【請求項 5】

テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量と前記テキストの強調位置との関係を表す強調位置予測モデルを強調位置予測モデル格納部に格納しておき、  
入力テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量を得る特徴量構成ステップと、  
前記強調位置予測モデルおよび前記特徴量構成ステップで得られた前記特徴量を用いて前記入力テキストの強調位置を識別する強調位置予測ステップと、  
を実行し、  
前記特徴量が、主辞の単語の出現形を表す情報および前記主辞の単語の品詞を表す情報を含む、または、前記主辞の単語および前記主辞の単語の前後の所定個数の単語のそれぞれの出現形を表す情報ならびに前記単語のそれぞれの品詞を表す情報を含む、  
ことを特徴とする強調位置予測方法。

【請求項 6】

請求項 5 の強調位置予測方法であって、  
前記特徴量は、前記主辞の単語から、当該主辞の単語に係り受けで係ってくる前側の単語まで、の距離の最小値を表す情報を含む、  
ことを特徴とする強調位置予測方法。

【請求項 7】

請求項 5 または 6 の強調位置予測方法であって、  
前記特徴量は、単語のカテゴリを表す情報を含む、  
ことを特徴とする強調位置予測方法。

【請求項 8】

請求項 5 から 7 の何れかの強調位置予測方法であって、  
前記強調位置予測モデルは、前記テキストを構成する区間が強調位置であるかと前記特徴量との関係を表すモデルであり、  
前記強調位置予測ステップは、前記入力テキストを構成する各区間が強調位置であることを識別する、  
ことを特徴とする強調位置予測方法。

【請求項 9】

請求項 1 から 4 の何れかの強調位置予測装置としてコンピュータを機能させるためのプログラム。

【発明の詳細な説明】

【技術分野】

【0001】

本発明は、音声合成技術に関し、特に音声の強調位置を予測する技術に関する。

【背景技術】

【0002】

例えば、映画のシーンに応じた台詞を発話する場面、童話の語り聞かせの場面、テレビなどのメディアを通じた商品宣伝の場面、およびコールセンタなどでの電話対応場面などで自然に発せられた「表現豊かな音声」において、強調は頻繁に行われている。音声合成

10

20

30

40

50

によって生成される合成音においても、適切な強調を行うことにより、合成音の自然性が高まる。

【0003】

非特許文献1にあるように、特定の区間が強調されて発話された場合、強調区間の基本周波数が読上げ調で発話された部分に比べて高くなる。従来の音声合成装置で、読み上げとは異なる日常の様々な表現豊かな音声から音声合成用モデルを構築し、そのモデルを用いて音声合成を行っても、このような強調区間での声の高さを十分に再現できない。

【0004】

非特許文献1では、原音声と合成音声との基本周波数の差分が閾値よりも大きい区間を強調区間と推定し、推定した強調区間にマーク（強調マーク）を付与し、それらの強調マ

10

【先行技術文献】

【非特許文献】

【0005】

【非特許文献1】前野，能勢，小林，井島，中嶋，水野，吉岡，「強調音声合成のための局所韻律コンテキスト自動付与の検討」，電子情報通信学会 技術研究報告 音声研究会，112(81)，pp.1-6，2012．

【発明の概要】

【発明が解決しようとする課題】

20

【0006】

非特許文献1の手法によって音声合成を行うためには、合成対象の入力テキストの各区間に強調マークを付与するか否かを決定する必要がある。しかし、非特許文献1の手法では、原音声と合成音声との差分に基づいて強調マークを付与するか否かを決定するため、当該強調マークが付与された区間の性質が、従来の言語処理や音声学の研究で推定されてきたテキストの言語情報に基づく強調区間の性質と同じになるとは限らない。

【0007】

この発明は、このような課題に鑑みてなされたものであり、入力テキストの言語情報に基づいて、入力テキストの強調区間を予測することが可能な技術を提供することを目的と

30

【課題を解決するための手段】

【0008】

本発明では、テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量とテキストの強調位置との関係を表す強調位置予測モデルを格納しておく。音声合成を行う際には、入力テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される前記の特徴量と同種の特徴量を求め、当該特徴量および強調位置予測モデルを用いて入力テキストの強調位置を識別する。

【発明の効果】

【0009】

本発明では、入力テキストの言語情報に基づいて、入力テキストの強調区間を予測することができる。また非特許文献1のような方法では、モデル作成時に学習データに依存して強調区間が変化する。本発明では、そのような変化に対応できる強調区間をテキストから予測できる。

40

【図面の簡単な説明】

【0010】

【図1】強調位置予測装置の機能構成例を示す図。

【図2】強調位置予測装置の動作フローを示す図。

【図3】特徴量構成部で構成される特徴量の組の一例。

【発明を実施するための形態】

【0011】

50

以下、図面を参照して本発明の実施形態を説明する。

図1は、本形態の強調位置予測装置100のブロック図であり、図2はその動作フローを表す。

【0012】

本形態の強調位置予測装置100は、制御部111、特徴量構成部112、強調位置予測部113、および強調位置予測モデル格納部114を有する。強調位置予測装置100は、さらにカテゴリ名辞書を格納したカテゴリ名辞書格納部115を有していてもよい。強調位置予測装置100は、例えば、CPU (central processing unit) やRAM (random-access memory) 等を備える公知のコンピュータに所定のプログラムが読み込まれて構成される装置である。強調位置予測装置100の少なくとも一部が集積回路等のハードウェアで構成されていてもよい。強調位置予測装置100は、制御部111の制御のもとで各処理を実行する。

10

【0013】

強調位置予測装置100の外部の解析結果格納部101には、形態素解析結果および係り受け解析結果が言語情報として付与された入力テキストが格納されている。入力テキストは、音声合成対象となる1個以上の単語からなる系列であり、この系列は強調生起範囲毎に区切られている。

【0014】

形態素解析は、テキストを単語ごとに分割し、それぞれの単語に品詞や読みなどの辞書的情報を付与する技術であり、例えば、参考文献1 (松本裕治, 形態素解析システム『茶筌』, 情報処理, 41(11), pp.1208-1214, 2000) の方法によって実施できる。

20

【0015】

係り受け解析は、形態素解析結果を入力として、複数の単語列からなる文節を構成し、品詞や単語の出現形や単語のアイディの関係に基づいて、文節間の係り受け関係を予測する技術であり、例えば、参考文献2 (工藤拓, 松本裕治, チャンキングの段階適用による日本語係り受け解析, 情報処理学会論文誌, 43(6), pp.1834-1842, 2002) の方法によって実施できる。これら自体は、従来技術であり、それらの詳しい説明は省略する。

【0016】

声の高さのような韻律処理ではアクセント句を単位として処理を行う場合が多い。アクセント句は、1つ以上の文節から構成される系列である。アクセント句の先頭の文節の係り受け情報と最後の文節の係り受け情報をアクセント句間での係り受け情報として用いることができる。アクセント句の境界の予測は、例えば、参考文献3 (Nakajima, H., Miyazaki, N., Yoshida, A., Nakamura, T., Mizuno, H., "Creation and Analysis of a Japanese Speaking Style Parallel Database for Expressive Speech Synthesis", in Proc. Oriental COCOSDA, 2010, [http://desceco.org/0-COCOSDA2010/proceedings/paper\\_30.pdf](http://desceco.org/0-COCOSDA2010/proceedings/paper_30.pdf).) にて実施できることが知られている。

30

【0017】

強調生起範囲は、所定の基準に基づいて定められた単語または単語列からなる区間である。どのような区間を強調生起範囲としてもよいが、例えば、ポーズ位置とポーズ位置で挟まれた区間 (イントネーション句) にある単語または単語列を1つの強調生起範囲とすることができる。ポーズ位置の予測は、例えば、参考文献4 (木暮監修, 山森編著, 「未来ねっと技術シリーズ4 メディア処理技術」, pp.76-77, 電気通信協会) に記載された方法によって実行できる。或いは、アクセント句を強調生起範囲としてもよい。

40

【0018】

特徴量構成部112は、解析結果格納部101から、入力テキストの各強調生起範囲に対応する形態素解析結果および係り受け解析結果の少なくとも一方を読み出し (ステップS11, S12)、読み出した情報に対応する特徴量を生成して出力する。特徴量構成部112は、入力テキストのすべての強調生起範囲についてそれぞれ特徴量を生成する。特徴量は、形態素解析結果および係り受け解析結果の少なくとも一方から抽出可能な情報である。例えば、以下のa) ~ e) の要素の全てまたはそれらの一部の要素の組み合わせか

50

らなる列（例えば、ベクトルや要素の結合値）を特徴量とする。この場合、強調位置予測性能の観点から、a) および b) の両方の要素を含む列を特徴量とすることが望ましい。好ましくは、特徴量が a) および b) の両方の要素を含む列であることは必須である。また、下記の c - 1) ~ c - 4) の一部またはすべての要素を含む列を加えて特徴量としてもよい。

【0019】

a) 強調生起範囲内の着目する単語とその前後のそれぞれN個の単語の出現形（あるいは表層形と呼ぶ）を表す情報を特徴量の要素としてもよい。これは形態素解析結果から取り出すことが可能な情報である。強調生起範囲に対してどの位置の単語を「着目する単語」とするか、各強調生起範囲に対していくつの単語を「着目する単語」とするか、ならびにNをどのような値にするかなどの条件は事前に定められる。例えば、事前に学習データを用いた予備実験が行われ、そこで最高性能を示した条件が採用される。例えば、強調生起範囲に対して特定の関係にある単語のうち主辞の単語のみを「着目する単語」としてもよいし、強調生起範囲に対して特定の関係にあるすべての単語を「着目する単語」としてもよい。「強調生起範囲に対して特定の関係にある単語」は、強調生起範囲に属する単語であってもよいし、強調生起範囲から所定の範囲内にある単語であってもよいし、強調生起範囲から所定距離だけ離れた単語であってもよい。Nは0以上の整数であり、すべての入力データに対して同じ値であってもよいし、入力データの種別等に応じて異なってもよい。これらの条件の設定は、以下の他の要素についても同様な方法で行われる。ただし、a) ~ e) での条件は同じであってもよいし、異なってもよい。

【0020】

b) 上記着目する単語とその前後のそれぞれN個の単語（すなわち、上記a) の $2N + 1$ 個の単語）の品詞を表す情報を特徴量の要素としてもよい。これも形態素解析結果から取り出すことが可能な情報である。

【0021】

c) 強調生起範囲の前後の単語に関して着目する単語と当該着目する単語と係り受けの関係にある他の単語との関係を表す情報を特徴量の要素としてもよい。これは係り受け解析結果から取り出すことが可能な情報である。例えば、以下を特徴量の要素とすることができる。

c - 1) 着目する単語から当該着目する単語に係る文末側の他の単語までの単語数で数えた距離を表す情報。

c - 2) 着目する単語から当該着目する単語に係る文頭側の他の単語までの単語数で数えた最小距離を表す情報。すなわち、着目する単語から当該着目する単語に係る文頭側の最も近い他の単語までの単語数で数えた距離を表す情報。

c - 3) 着目する単語から当該着目する単語に係る文頭側の他の単語までの単語数で数えた最大距離を表す情報。すなわち、着目する単語から当該着目する単語に係る文頭側の最も遠い他の単語までの単語数で数えた距離を表す情報。

c - 4) 着目する単語に係る文頭側の単語の個数を表す情報。

なお、c - 1) では、着目する単語が文末の単語ではなく、かつ、文末方向に係る単語がない場合には、要素の値を0とする。c - 2) と c - 3) では、着目する単語が文頭の単語ではなく、かつ、その単語に係る文頭側の単語がない場合には、要素の値を0とする。また、c - 1) の「文末側」を「文頭側」に置換した情報を要素としてもよいし、c - 2) ~ c - 4) の「文頭側」を「文末側」に置換した情報を要素としてもよい。また、c) の係り受けは通常は文節単位で得られるが、音声合成で頻繁に用いられるアクセント句単位での係り受け情報を得る場合には、アクセント句境界の前後の単語から「着目する単語」への前記の距離や個数を特徴量の要素として用いることができる。なお、アクセント句境界は、特徴量構成部112に入力されたアクセント句境界やポーズの有無を表す情報によって特定されてもよいし、前述の参考文献3等に基づいて特徴量構成部112によって予測されてもよい。「着目する単語」は、強調生起範囲内の個々の単語である。

【0022】

10

20

30

40

50

d) 強調生起範囲内の着目するアクセント句の前後の位置でのポーズの有無を表す情報を要素としてもよい。これは参考文献3および4の結果に基づいて得られる情報である。着目するアクセント句のM1個前側のアクセント句の前の位置のポーズの有無、着目するアクセント句のM2個後ろ側のアクセント句の後ろの位置のポーズの有無を要素として含めても良い。なお、M1とM2は事前に定められる。例えば、事前に学習データを用いた予備実験が行われ、そこで最高性能を示した条件が採用される。「着目するアクセント句」は、強調生起範囲内の着目する単語が含まれているアクセント句である。

#### 【0023】

e) 強調位置予測装置100がカテゴリ名辞書格納部115を備える場合、強調生起範囲に対応する単語をキーとしてカテゴリ名辞書を検索することで得られる「単語のカテゴリ」を表す情報を特徴量の要素としてもよい。例えば、上記a)の $2N+1$ 個の単語のうち的主要部の単語の、品詞以外のカテゴリ名の全てまたはそれらの部分的な組み合わせを表す情報を特徴量の要素としてもよい。例えば、単語が属する組織名や会社名といった細分類カテゴリを表す情報を特徴量の要素としてもよい。このような細分類カテゴリは、例えば、参考文献5(日本語語彙大系, NTTコミュニケーション科学研究所監修, 池原他編集, 1997刊, 岩波書店)を用いて付与することができる。カテゴリ名としては例えば外来語か和語かといった種別がある。また、独自構築したカテゴリ名辞書を参照することにより得られる単語のカテゴリを表す情報を特徴量の要素としてもよい。例えば、「明るい」「楽しい」といったポジティブなカテゴリや、「暗い」「つらい」といったネガティブなカテゴリを設定することもできる。なお、「主要部」とは「主辞」を意味する。すなわち、「主要部」とは係り受けで係ってくる単語(受け側の単語)を意味する。上記の $2N+1$ 個の単語に主要部が複数個存在する場合には、いずれか1つの主要部の単語のみについて特徴量の要素が構成されてもよいし、複数個の主要部の単語について特徴量の要素が構成されてもよい。上記の $2N+1$ 個の単語に主要部が存在しない場合には、主要部が存在しないことを表す情報を特徴量の要素としてもよい。

#### 【0024】

図3に複数の要素からなる列を特徴量とした例を示す。図3の例では、各アクセント句を強調生起範囲とし、アクセント句に対応する主要部の単語の出現形を表す情報(2)の列)、当該主要部の単語の品詞を表す情報(3)の列)、アクセント句の前側にポーズがあるかないかを表す情報(4)の列)、アクセント句の後側にポーズがあるかないかを表す情報(5)の列)、上記主要部の単語から当該単語に係り受けで係ってくる前側の単語までの距離の最小値を表す情報(6)の列)、および上記主要部の単語の品詞以外のカテゴリ名を表す情報(7)の列)を特徴量の要素としている。例えば、アクセント句番号2に対応する特徴量は、a)として図3の2)列の『ソフト』という単語とその前後の『クリーン』や『温風』という単語を表す情報、b)として3)列の『形容詞』と前後の『形容詞』と『名詞』を表す情報、c)として6)列の『1』を表す情報、d)として4)列と5)列の『無』と『無』を表す情報、e)として7)列の『外来語』を表す情報を要素とするベクトルとなる。

#### 【0025】

特徴量構成部112から出力された特徴量は、強調位置予測部113に入力される。強調位置予測部113は、強調位置予測モデル格納部114に格納された強調位置予測モデル、および特徴量構成部112で得られた特徴量を用い、入力テキストの強調位置を識別する(ステップS114)。強調位置予測モデルは、テキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量と当該テキストの強調位置との関係を表すモデルである。すなわち、強調位置予測モデルは、テキストを構成する強調生起範囲(区間)が強調位置であるかと特徴量との関係を表すモデルであり、強調位置予測部113は、特徴量構成部112で得られた特徴量を用い、入力テキストを構成する各強調生起範囲が強調位置であるかを識別する。

#### 【0026】

以下に詳細を例示する。各強調生起範囲 $i = 0, \dots, I - 1$ に対して得られた特徴量を

10

20

30

40

50

$x_i$  とし、各強調生起範囲  $i$  が強調位置であるか否かを表す識別情報（強調マークを付与するか否かを表す識別情報）を  $y_i$  とする。ただし、 $I$  は 1 以上の整数である。例えば、 $I$  は入力テキストに属するすべての強調生起範囲の個数である。特徴量  $x_i$  はベクトル等である。識別情報  $y_i$  の例は、強調生起範囲  $i$  が強調位置である場合に  $y_i = 1$  となり、強調生起範囲  $i$  が強調位置でない場合に  $y_i = 0$  となる二値情報である。特徴量  $x_i$  の系列を  $x = (x_0, \dots, x_{I-1})$  とし、識別情報  $y_i$  の系列を  $y = (y_0, \dots, y_{I-1})$  とする。この場合、強調位置予測モデルは、入力テキストの特徴量の系列  $x$  と識別情報の系列  $y$  とを対応付けるモデルである。例えば、強調位置予測モデルは、特徴量の系列  $x$  を入力変数とし、識別情報の系列  $y$  を出力変数とするモデルである。強調位置予測モデルに限定はないが、強調位置予測モデルの具体例は、系列ラベリング等に用いられる隠れマルコフモデル等の確率モデルである。強調位置予測モデルは、例えば、学習用のテキストの形態素解析結果および係り受け解析結果の少なくとも一方から構成される特徴量と、当該テキストの各強調生起範囲が強調位置であるか否かを表す識別情報との組からなる学習データを用い、一般の機械学習手法を実施することで構築できる。強調位置予測モデルの学習に用いられる特徴量の構成は、特徴量構成部 112 で得られる特徴量の構成と同じである。学習データの識別情報は人手で付与されたものであってもよいし、非特許文献 1 等の従来技術によって自動抽出されたものであってもよい。機械学習手法の詳細は、例えば参考文献 6（高村大也ほか著、「言語処理のための機械学習入門」、コロナ社）等に記載されている。強調位置予測モデルは、例えば、強調生起範囲を処理単位として、その処理単位、および、その前後の単位に対応する特徴量から構成した特徴量の系列を入力とし、その処理単位での強調マークの有無を表す識別情報の系列を予測するためのモデルである。以下に隠れマルコフモデルで構築された強調位置予測モデルを例示する。

10

20

【数 1】

$$P(x, y) = \prod_{i=0}^{I-1} P(y_i | x_i) P(x_i | x_{i-1})$$

ただし、 $P(x, y)$  は  $x = (x_0, \dots, x_{I-1})$ 、 $y = (y_0, \dots, y_{I-1})$  の同時確率、 $P(y_i | x_i)$  は  $x_i$  のもとでの  $y_i$  の条件付き確率、 $P(x_i | x_{i-1})$  は  $x_{i-1}$  のもとでの  $x_i$  の条件付き確率である。 $x_{-1}$  は文の先頭の単語よりも前にあると仮定する文の開始位置という便宜上の強調予測範囲に関して得られる特徴量で定数を要素とするベクトルである。

30

【0027】

強調位置予測部 113 は、強調位置予測モデル格納部 114 から強調位置予測モデルを読み込み、当該強調位置予測モデルの入力変数に特徴量構成部 112 から出力された特徴量の系列  $x$  を設定し、識別情報の系列  $y$  を予測し、当該識別情報の系列  $y$  を強調位置予測結果として出力する。例えば、隠れマルコフモデルで強調位置予測モデルが構成されている場合、入力された特徴量の系列  $x$  に対して  $P(x, y)$  を最大にする識別情報の系列  $y$  が出力される。このような識別情報の系列  $y$  の探索は、例えば公知の Viterbi アルゴリズムを用いて行うことができる。強調生起範囲を処理単位として、その処理単位での強調マークの有無を表す識別情報の系列  $y$  を、その単位、および、その前後の単位に対応する特徴量から構成した特徴量の列を入力した強調位置予測モデルを用いて、文頭から文末までの前記の単位ごとに強調マークを付与する場合と付与しない場合のすべての可能性を列挙して、文頭から文末まで大域的に、強調位置予測モデルが与える確率が最大の系列を Viterbi アルゴリズムで選択すれば良い。或いは、強調生起範囲に代えて、着目する単語、文節、アクセント句を処理単位として識別情報の系列  $y$  を得てもよい。すなわち、特徴量を得る処理単位と識別情報を得る処理単位とは、同一であってもよいし、異なってもよい。あるいは、処理単位の前後数単位間で確率の高い識別情報の系列に探索範囲を絞って探索してもよい。また、特徴量の構成に関わった範囲のみについて強調位置の予測を行うことにし、強調位置の予測自体を局所的に行なうことも可能である。これらの探索法自体は周知であり、例えば、参考文献 6 に記載された方法で実施できる。また、

40

50

強調位置予測モデルを隠れマルコフモデルに限定はしない。例えば、強調位置予測モデルは、スコアや確率値を付与できるモデルであれば、どのようなモデルによって構成されても良い。決定木によって構成されても良いし、対数線形モデルによって構成されても良いし、ニューラルネットワークで構成されても良い。

【0028】

以上のように、本形態では、テキストの言語情報に対応する特徴量とテキストの強調位置との関係を表す強調位置予測モデルを用い、入力テキストの強調位置を識別できる。また、本形態の強調位置予測モデルを用いて予測された強調位置は、非特許文献1のモデルに整合した性質を持つ。その結果、声の高さが精度高く再現された自然な音声を合成できる。

10

【0029】

なお、本発明は上述の実施の形態に限定されるものではない。例えば、強調位置予測装置が、入力テキストの形態素解析や係り受け解析を行う手段を備えていてもよいし、強調位置予測装置が、形態素解析結果や係り受け解析結果を格納する解析結果格納部を備えていてもよい。また上述の各種の処理は、記載に従って時系列に実行されるのみならず、処理を実行する装置の処理能力あるいは必要に応じて並列的あるいは個別に実行されてもよい。その他、本発明の趣旨を逸脱しない範囲で適宜変更が可能であることはいうまでもない。

【0030】

上述の構成をコンピュータによって実現する場合、各装置が有すべき機能の処理内容はプログラムによって記述される。このプログラムをコンピュータで実行することにより、上記処理機能がコンピュータ上で実現される。この処理内容を記述したプログラムは、コンピュータで読み取り可能な記録媒体に記録しておくことができる。コンピュータで読み取り可能な記録媒体の例は、非一時的な(non-transitory)記録媒体である。このような記録媒体の例は、磁気記録装置、光ディスク、光磁気記録媒体、半導体メモリ等である。

20

【0031】

このプログラムの流通は、例えば、そのプログラムを記録したDVD、CD-ROM等の可搬型記録媒体を販売、譲渡、貸与等することによって行う。さらに、このプログラムをサーバコンピュータの記憶装置に格納しておき、ネットワークを介して、サーバコンピュータから他のコンピュータにそのプログラムを転送することにより、このプログラムを流通させる構成としてもよい。

30

【0032】

このようなプログラムを実行するコンピュータは、例えば、まず、可搬型記録媒体に記録されたプログラムもしくはサーバコンピュータから転送されたプログラムを、一旦、自己の記憶装置に格納する。処理の実行時、このコンピュータは、自己の記憶装置に格納されたプログラムを読み取り、読み取ったプログラムに従った処理を実行する。このプログラムの別の実行形態として、コンピュータが可搬型記録媒体から直接プログラムを読み取り、そのプログラムに従った処理を実行することとしてもよく、さらに、このコンピュータにサーバコンピュータからプログラムが転送されるたびに、逐次、受け取ったプログラムに従った処理を実行することとしてもよい。サーバコンピュータから、このコンピュータへのプログラムの転送は行わず、その実行指示と結果取得のみによって処理機能を実現する、いわゆるASP(Application Service Provider)型のサービスによって、上述の処理を実行する構成としてもよい。

40

【0033】

上記実施形態では、コンピュータ上で所定のプログラムを実行させて本装置の処理機能が実現されたが、これらの処理機能の少なくとも一部がハードウェアで実現されてもよい。

【符号の説明】

【0034】

100 強調位置予測装置

50



【図 1】

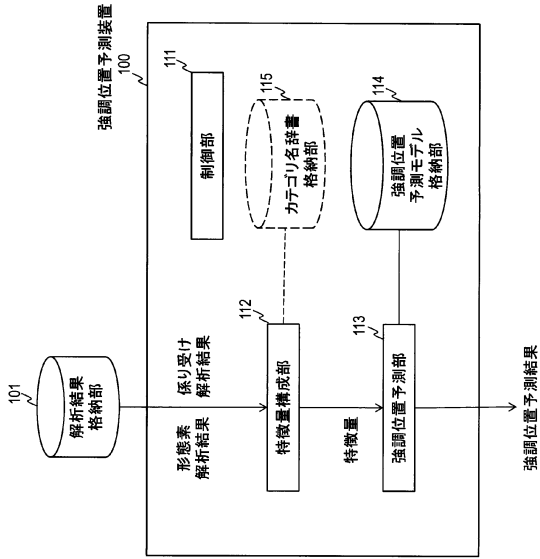


図 1

【図 2】

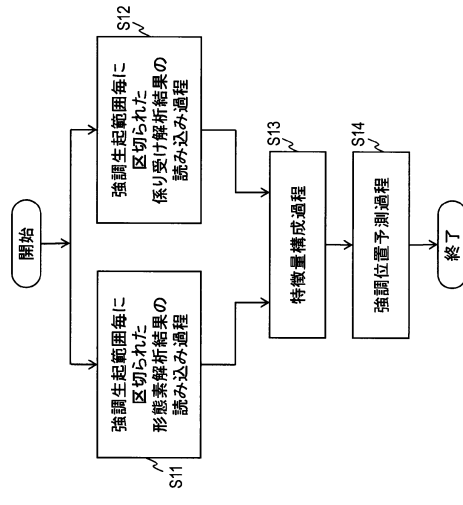


図 2

【図 3】

1) アクセント句番号	2) 単語	3) 品詞	4) アクセント句前ボースの有無	5) アクセント句後ボースの有無	6) 係り受けで係ってくる前側の単語までの距離の最小値	7) カテゴリ名
1	クリーン(で)	形容詞	有	無	0	外来語
2	ソフト(な)	形容詞	無	無	1	外来語
3	温風(で)	名詞	無	有	1	和語
4	しっかり	副詞	有	無	0	和語
5	乾か(ず、から)	動詞	無	有	1	和語
:	:	:	:	:	:	:

図 3

---

フロントページの続き

- (72)発明者 吉岡 理  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内
- (72)発明者 井島 勇祐  
東京都千代田区大手町二丁目3番1号 日本電信電話株式会社内

審査官 大野 弘

- (56)参考文献 特開平07-199981(JP,A)  
特開平11-249678(JP,A)  
特開2004-170983(JP,A)

- (58)調査した分野(Int.Cl., DB名)
- |      |       |
|------|-------|
| G10L | 13/08 |
| G10L | 13/10 |