

(19) 日本国特許庁(JP)

(12) 特 許 公 報(B2)

(11) 特許番号

特許第4512846号
(P4512846)

(45) 発行日 平成22年7月28日(2010.7.28)

(24) 登録日 平成22年5月21日(2010.5.21)

(51) Int.Cl. F I
G 1 0 L 13/06 (2006.01) G 1 0 L 13/06 2 4 0 C

請求項の数 7 (全 19 頁)

<p>(21) 出願番号 特願2004-232297 (P2004-232297)</p> <p>(22) 出願日 平成16年8月9日(2004.8.9)</p> <p>(65) 公開番号 特開2006-47916 (P2006-47916A)</p> <p>(43) 公開日 平成18年2月16日(2006.2.16)</p> <p>審査請求日 平成19年7月19日(2007.7.19)</p> <p>(出願人による申告)平成16年度独立行政法人情報通信研究機構、研究テーマ「大規模コーパスベース音声対話翻訳技術の研究開発」に関する委託研究、産業活力再生特別措置法第30条の適用を受ける特許出願</p>	<p>(73) 特許権者 393031586 株式会社国際電気通信基礎技術研究所 京都府相楽郡精華町光台二丁目2番地2</p> <p>(74) 代理人 100099933 弁理士 清水 敏</p> <p>(72) 発明者 西澤 信行 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p>(72) 発明者 河井 恒 京都府相楽郡精華町光台二丁目2番地2 株式会社国際電気通信基礎技術研究所内</p> <p>審査官 山下 剛史</p> <p style="text-align: right;">最終頁に続く</p>
---	---

(54) 【発明の名称】 音声素片選択装置および音声合成装置

(57) 【特許請求の範囲】

【請求項1】

合成器指令に応答して、音声素片データベースから前記合成器指令によって定まるコスト関数が所定の条件を充足する音声素片系列を選択するための音声素片選択装置であって、

前記コスト関数は、時刻 i におけるコストが、時刻 $i - 1$ までのコストと、時刻 i における音声素片候補の特徴量とによって決定されるように定められ、前記装置は、

合成器指令の入力に응答して、ある時刻における前記音声素片系列を構成する音声素片として選択された音声素片に対応するノードを根ノード、および前記ある時刻より所定の時間だけ後の時刻における音声素片の候補に対応するノードを葉ノードとし、各ノードには、当該ノードに対応する音声素片までの音声素片系列の累積コストの値がそれぞれ関連付けられている、可能な音声素片の系列を表す探索木を決定するための手段と、

前記探索木が決定されたことに응答して、前記ある時刻および前記所定の時間だけ後の時刻に関連して予め定められる深さ制限を課したサーチ手法にしたがって、前記探索木をサーチすることにより、前記コストが所定の条件を充足する音声素片を選択して出力するとともに、前記探索木を更新するための探索手段とを含む、音声素片選択装置。

【請求項2】

前記探索手段は、

前記葉ノードのうち、前記累積コストの値が所定の条件を充足するノードを決定するための葉ノード決定手段と、

前記根ノードの子ノードのうち、前記根ノードから前記所定の条件を充足するノードへ

の経路とは別の経路上に存在する子ノード、および前記別の経路上に存在する子ノードから分岐する全てのノードを前記探索木から削除するための第1のノード削除手段と、

前記根ノードから前記所定の条件を充足するノードへの経路上に存在する、前記根ノードの子ノードに対応する音声素片を、前記音声素片系列の、前記ある時刻に続く時刻における音声素片として選択し出力するための音声素片選択手段と、

前記根ノードから前記所定の条件を充足するノードへの経路上に存在する、前記根ノードの子ノードを新たな根ノードとして選択し、古い前記根ノードを前記探索木から削除するための根ノード選択手段と、

前記新たな根ノードから到達可能な前記葉ノードのうち、葉ノード数に対する所定の限定を充足するものと、前記新たな根ノードと、前記新たな根ノードから前記所定の限定を充足する葉ノードまでの経路上のノードとを残し、それ以外のノードを前記探索木から削除するための第2のノード削除手段と、

前記ある時刻から所定時間後の時刻における根ノードとして、前記新たな根ノードを前記葉ノード決定手段に与えるための手段とを含む、請求項1に記載の音声素片選択装置。

【請求項3】

前記葉ノード決定手段は、前記葉ノードのうち、前記累積コストの値が最小のノードを決定するための手段を含む、請求項2に記載の音声素片選択装置。

【請求項4】

前記第2のノード削除手段は、

前記新たな根ノードから到達可能な前記葉ノードのうち、対応する累積コストが小さいものから順に、かつ対応する音声素片が互いに異なる、予め定める上限個数以下の複数の葉ノードを決定するための手段と、

前記予め定める上限個数以下の複数の葉ノードと、前記新たな根ノードから前記予め定める上限個数以下の葉ノードまでの経路上のノードとを残し、それ以外のノードを前記探索木から削除するための手段を含む、請求項2に記載の音声素片選択装置。

【請求項5】

前記探索手段は、予め定める深さ制限を課した、所定ビーム幅のDP(Dynamic Programming)サーチ手法にしたがって、前記探索木をサーチし更新することにより、逐次的に音声素片を選択して前記音声素片系列を出力するための手段を含む、請求項1に記載の音声素片選択装置。

【請求項6】

コンピュータにより実行されると、当該コンピュータを請求項1～請求項5のいずれかに記載の音声素片選択装置として動作させる、コンピュータプログラム。

【請求項7】

入力される合成目標のテキストを処理し、音声合成のための合成器指令を生成するためのテキスト処理手段と、

複数の音声素片を格納した音声素片データベースと、

前記テキスト処理手段が出力する前記合成器指令を入力として受けるように接続され、当該合成器指令に合致した音声素片を前記音声素片データベースから読み出すことにより音声素片系列を出力する、請求項1～請求項5のいずれかに記載の音声素片選択装置と、

前記音声素片選択装置により出力された音声素片系列にしたがって前記音声素片データベースに記憶されている音声素片を順に接続することにより、前記合成目標のテキストに対応する音声波形を合成するための接続手段とを含む、音声合成装置。

【発明の詳細な説明】

【技術分野】

【0001】

この発明は、波形接続型音声合成において、実時間で所与の条件を満足する波形候補を検索する際等に使用される検索技術に関し、特に、波形候補が多数の場合でも所定の制限時間内にできるだけ精度良く所与の波形候補を検索できる音声素片選択装置およびそれを用いた音声合成装置に関する。

10

20

30

40

50

【背景技術】

【0002】

波形接続型音声合成において、高品質の音声を得るためには多数の音声素片が必要である。そのため、音声素片の数はより多く、音声素片データベースはより大きくなる。このように音声素片データベースが大きくなると、様々な問題が生じる。その第1の問題は、音声の録音がむずかしくなることである。なぜなら、音声素片は、合成音声を高品質にするためには、基本的に一人の話者の発声から収集しなければならないからである。しかしこの問題に対処することは、困難ではあっても決して不可能ではない。それよりも問題となるのは、音声素片データベースが大きくなることにより、音声素片の検索に要する時間が長くなることである。

10

【0003】

機械による自動応答システムなどでは、音声合成を所定時間内に完了しなければならない。一方、音声素片データベース中の音声素片の数が多くなれば適切な音声素片の検索に要する時間が長くなることは直感的に分かる。さらに、波形接続型音声合成では、音声合成テキストから得られる合成器指令に最もよく合致するというだけでなく、前後の音声素片との間の接続についても考慮する必要がある。そのため、素片選択では前後の音声素片との組合せも考慮して素片を検索しなければならない。しかし、音声素片データベース中の音声素片の数が多くなると、それらの組合せの数は文字通り爆発的に増加するため、音声素片の検索を所与の時間内に完了することが困難になる。

【非特許文献1】T. Toda他2名「知覚的評価に基づく、波形接続型音声合成における素片選択のための統合コスト関数の最適化」、EUROSP E E C H 予稿集、ジュネーブ、スイス、pp. 297 - 300、2003年9月(T. Toda, H. Kawai, and M. Tsuzaki, "Optimizing Integrated Cost Function for Segment Selection in Concatenative Speech Synthesis Based on Perceptual Evaluations", Proc. EUROSP E E C H, Geneva, Switzerland, pp. 297-300, Sep. 2003.)

20

【非特許文献2】戸田 智基、河井 恒、津崎 実、「波形接続型音声合成における知覚的評価に基づく素片選択サブコスト関数の最適化」、信学技報、SP2003-81、pp. 43 - 48、2003年8月

【非特許文献3】阿部匡伸、匂坂芳典、梅田哲夫、桑原尚夫：「研究用日本語音声データベース利用解説書（連続音声データ編）」、TR-I-0166、ATR 自動翻訳電話研究所、1990。

30

【発明の開示】

【発明が解決しようとする課題】

【0004】

こうした問題を解決するために、主として二つの対策がある。一方は、選択すべき音声素片のカテゴリ範囲を狭くして、選択対象として考慮する必要のある音声素片の数を少なくする手法である。他方は、音声素片選択の際のアルゴリズムを工夫し、いわゆる枝刈り（プルーニング）という手法により検索空間を動的に小さくする手法である。前者の方法はオフラインで実行され、素片選択の際の計算量を直接減少させる。これに対し後者はオンラインで実行され、素片選択の際の計算量が減少するかどうかはわからず、むしろ計算量は変わらないことが多い。

40

【0005】

しかし、自動応答システムのように実時間性と合成音声の品質との双方が要求されるアプリケーションでは、所定時間内にできるだけ高品質の音声合成を行なうことが必要である。したがって最初から選択対象が限定される前者の方策ではなく、オンラインで動的に選択対象を限定し、かつ時間が許す限りで合成音声の品質ができるだけ高くなるような音声素片の検索を行なう後者的方策を採ることが望ましい。

【0006】

波形接続型音声合成の中に、「コスト」と呼ばれる値を計算することにより、音声素片系列を選択する手法がある（非特許文献1、2を参照されたい。）。この手法では、合成

50

器指令となる一文の先頭から最後までについて、種々の可能性を考慮して複数の音声素片系列を累積コストとともに算出する。そうした音声素片系列の中でコストが最小となるものを選択し、その系列の音声素片を結合することにより、合成器指令の一文の合成音声波形を作成する。

【0007】

この手法では、一文の最後まで到達しなければ最適解が得られず、したがって、それまでは音声を出力できない。もちろん、最適な音声素片系列を探索する場合、その探索空間は広大となるため、探索を効率的にする手法が考えられている。しかしそれでも結局は一文の最後まで最適解が得られないことには変わりがない。したがって、実時間で音声応答をするシステムに利用するには不利である。一文の最後まで到達しなくても、合成音声の品質を下げないような音声素片を検索できるようにすることが望まれる。

10

【0008】

したがって本発明の目的は、合成器指令の途中でも、合成器指令に合致した音声素片系列を合成音声の品質低下を防ぎながら選択できる、音声素片選択装置および当該装置を用いた音声合成装置を提供することである。

【課題を解決するための手段】

【0009】

本発明の第1の局面に係る音声素片選択装置は、合成器指令に応答して、音声素片データベースから合成器指令によって定まるコスト関数が所定の条件を充足する音声素片系列を選択するための音声素片選択装置であって、コスト関数は、時刻 i におけるコストが、時刻 $i - 1$ までのコストと、時刻 i における音声素片候補の特徴量とによって決定されるように定められ、装置は、ある時刻において音声素片系列を構成する音声素片として選択された音声素片に対応するノードを根ノード、およびある時刻より所定の時間だけ後の音声素片の候補に対応するノードを葉ノードとし、各ノードには、当該ノードに対応する音声素片までの音声素片系列の累積コストの値がそれぞれ関連付けられている、可能な音声素片の系列を表す探索木を決定するための手段と、探索木が決定されたことに応答して、ある時刻および所定の時刻に関連して予め定められる深さ制限を課したサーチ手法にしたがって、探索木をサーチすることにより、コストが所定の条件を充足する音声素片を選択して出力するとともに、探索木を更新するための探索手段とを含む。

20

【0010】

コスト関数が、時刻 $i - 1$ までのコストと、時刻 i における音声素片候補の特徴量とによって決定される。音声素片候補の探索は、ある時刻において選択された音声素片に対応するノードを根ノード、それより所定の時間だけ後の時刻における音声素片の候補に対応するノードを葉ノードとする探索木を探索手段が探索することにより、ある時刻の次の時刻における音声素片が決定される。音声素片が逐次的に決定されるので、合成器指令の入力が完了するまえに音声素片系列を順に決定することができる。選択対象の音声素片数が多くなっても、所定時間のうちに音声素片の検索を終えることができるので、利用可能な音声素片数が大きくなったのに伴って合成音声の品質も向上する。

30

【0011】

好ましくは、探索手段は、葉ノードのうち、累積コストの値が所定の条件を充足するノードを決定するための葉ノード決定手段と、根ノードの子ノードのうち、根ノードから所定の条件を充足するノードへの経路とは別の経路上に存在する子ノード、および別の経路上に存在する子ノードから分岐する全てのノードを探索木から削除するための第1のノード削除手段と、根ノードから所定の条件を充足するノードへの経路上に存在する、根ノードの子ノードに対応する音声素片を、音声素片系列の、ある時刻に続く時刻における音声素片として選択し出力するための音声素片選択手段と、根ノードから所定の条件を充足するノードへの経路上に存在する、根ノードの子ノードを新たな根ノードとして選択し、古い根ノードを探索木から削除するための根ノード選択手段と、新たな根ノードから到達可能な葉ノードのうち、葉ノード数に対する所定の限定を充足するものと、新たな根ノードと、新たな根ノードから所定の限定を充足する葉ノードまでの経路上のノードとを残し、

40

50

それ以外のノードを探索木から削除するための第2のノード削除手段と、ある時刻から所定時間後の時刻における根ノードとして、新たな根ノードを葉ノード決定手段に与えて新たな探索木を決定させるための手段とを含む。

【0012】

より好ましくは、葉ノード決定手段は、葉ノードのうち、累積コストの値が最小のノードを決定するための手段を含む。

【0013】

さらに好ましくは、第2のノード削除手段は、新たな根ノードから到達可能な葉ノードのうち、対応する累積コストが小さい順に所定個数の葉ノードを決定するための手段と、所定個数の葉ノードと、新たな根ノードと、新たな根ノードから所定個数の葉ノードまでの経路上のノードとを残し、それ以外のノードを探索木から削除するための手段を含む。

10

【0014】

第2のノード削除手段は、新たな根ノードから到達可能な葉ノードのうち、対応する累積コストが小さいものから順に、かつ対応する音声素片が互いに異なる、予め定める上限個数以下の複数の葉ノードを決定するための手段と、予め定める上限個数以下の複数の葉ノードと、新たな根ノードから予め定める上限個数以下の葉ノードまでの経路上のノードとを残し、それ以外のノードを探索木から削除するための手段を含んでもよい。

【0015】

好ましくは、探索手段は、予め定める深さ制限を課した、所定ビーム幅のビームサーチ手法にしたがって、探索木をサーチし更新することにより、逐次的に音声素片を選択して音声素片系列を出力するための手段を含む。

20

【0016】

より好ましくは、探索手段は、予め定める深さ制限を課した、所定ビーム幅のDPサーチ手法にしたがって、探索木をサーチし更新することにより、逐次的に音声素片を選択して音声素片系列を出力するための手段を含む。

【0017】

本発明の第2の局面に係るコンピュータプログラムは、コンピュータにより実行されると、当該コンピュータを上記したいずれかの音声素片選択装置として動作させるものである。

【0018】

本発明の第3の局面に係る音声合成装置は、入力される合成目標のテキストを処理し、音声合成のための合成器指令を生成するためのテキスト処理手段と、複数の音声素片を格納した音声素片データベースと、テキスト処理手段が出力する合成器指令を入力として受けるように接続され、当該合成器指令に合致した音声素片を音声素片データベースから読み出すことにより音声素片系列を出力する、上記したいずれかの音声素片選択装置と、音声素片選択装置により出力された音声素片系列にしたがって音声素片データベースに記憶されている音声素片を順に接続することにより、合成目標のテキストに対応する音声波形を合成するための接続手段とを含む。

30

【発明を実施するための最良の形態】

【0019】

[第1の実施の形態]

図1に、本実施の形態に係る音声合成装置36の概略ブロック図を示す。図1を参照して、この音声合成装置36は、音声合成の目標となるテキストを含む入力32を受け、それに対する音声波形からなる出力38を得るものであって、入力32を受け、音素、基本周波数、音声の継続時間長、各音声素片のMFCC(Mel Frequency Cepstrum Coefficient)のセントロイドなどを含む合成器指令を生成するためのテキスト処理部64と、多数の音声素片を記憶する音声素片データベース(DB)62と、テキスト処理部64から合成器指令を受け、後述するコスト計算と、深さ制限をしたビームサーチとによって、短遅延で合成器指令に適した音声素片系列を選択するための素片選択部66と、素片選択部66が選択した音声素片系列にしたがって音声素片DB62から音声素片を抽出し、合

40

50

成器指令にしたがって加工および接続し出力 3 8 を生成するための接続部 6 8 とを含む。

【 0 0 2 0 】

図 2 を参照して、音声合成装置 3 6 においては、テキスト処理部 6 4 によるテキスト処理 8 0 と、素片選択部 6 6 による素片選択 8 2 と、接続部 6 8 による波形接続 8 4 とが、所定の遅延時間をもって同時並列的に行なわれる。そして、素片選択 8 2 が開始された後所定の遅延時間が経過した後に、波形接続 8 4 が開始されるようにする。そのために素片選択部 6 6 は、遅延時間の間に少なくとも最初の音声素片を確定しなければならない。この遅延時間は外部から指定可能であり、したがって図 1 に示すように素片選択部 6 6 は外部からこの遅延時間の入力を受ける。

【 0 0 2 1 】

素片選択部 6 6 において音声素片を選択する際の基準となるコストの計算方法について説明する。本実施の形態では、コストは複数のサブコストから構成される。本実施の形態では、コストは 6 種類のサブコスト、すなわち C_{F0} 、 C_{dur} 、 C_{cen} 、 C_{env} 、 C_{spg} および C_{F0C} からなる。 C_{F0} および C_{dur} は $\log F_0$ および音の継続時間長の相違による不自然さに対応する。 C_{cen} は音声素片候補と合成器指令との間のケプストラム平均の相違による不自然さに対応する。本実施の形態では、ターゲットパラメータはすべて HMM による音声合成方式により生成されるので、ケプストラムについても目標特徴量として使用できる。 C_{F0} 、 C_{dur} および C_{cen} はターゲットコスト C_T を構成する。

【 0 0 2 2 】

C_{env} は音声素片候補と目標との間の音素環境の不一致による不自然さに対応する。 C_{env} はまた、素片接続による影響も含む。二つの音声セグメントがコーパス中で連続したものである場合には、仮にそれらセグメントの音声環境が一致していたとしても C_{env} は 0 にならない。 C_{spg} はセグメント境界でのスペクトルの不連続による不自然さに対応する。 C_{spg} は、境界の直前のセグメントと直後のセグメントとの間のメルケプストラム距離の加重和によって算出できる。 C_{F0C} は $\log F_0$ の不連続による不自然さに対応する。 C_{env} 、 C_{spg} および C_{F0C} は接続コスト C_c を構成する。コーパス内で二つの音声セグメントが連続している場合、 C_{env} 、 C_{spg} および C_{F0C} はすべて 0 に設定される。

【 0 0 2 3 】

これらサブコスト関数の設計は、知覚実験の結果により最適化されている。

【 0 0 2 4 】

音声素片の選択は、累積コストを最小化することによって行なわれる。累積コスト関数 C はセグメント系列 $\{u_i\}$ を用いたときに以下の式で計算される、不自然さを表す指標である。

【 0 0 2 5 】

【数 1】

$$C(u_0, u_1, \dots, u_{N-1}) = w_T \cdot \left(\frac{1}{N} \sum_{i=0}^{N-1} C_T(u_i, t_i)^{p_T} \right)^{\frac{1}{p_T}} + w_C \cdot \left(\frac{1}{N-1} \sum_{j=1}^{N-1} C_c(u_{j-1}, u_j)^{p_C} \right)^{\frac{1}{p_C}} \quad (1)$$

ただし C_T 、 C_c および $\{t_i\}$ はそれぞれ、ターゲットコスト、接続コスト、および目標セグメント系列を示す。 w_T 、 p_T 、 w_C および p_C はそれぞれ、ターゲットコストおよび接続コストの重みおよび指数重みをそれぞれ表す。なお、出願人は、従来は p_T および p_C としてそれぞれ 1 と 1.5 の値を用いている。

【 0 0 2 6 】

ターゲットコスト C_T および接続コスト C_c はそれぞれ、以下の式により与えられる。

【 0 0 2 7 】

10

20

30

40

【数2】

$$\begin{aligned}
 C_T(u_i, t_i) &= w_{F_0} \cdot C_{F_0}(u_i, t_i) \\
 &+ w_{dur} \cdot C_{dur}(u_i, t_i) \\
 &+ w_{cen} \cdot C_{cen}(u_i, t_i) \quad (2)
 \end{aligned}$$

$$\begin{aligned}
 C_C(u_{j-1}, u_j) &= w_{env} \cdot C_{env}(u_{j-1}, u_j) \\
 &+ w_{spg} \cdot C_{spg}(u_{j-1}, u_j) \\
 &+ w_{F_0C} \cdot C_{F_0C}(u_{j-1}, u_j) \quad (3)
 \end{aligned}$$

10

ただし w_{F_0} 、 w_{dur} 、 w_{cen} 、 w_{env} 、 w_{spg} および w_{F_0C} はそれぞれ、各サブコストの重みを表す。これらの係数もまた、知覚実験により推定された。

【0028】

上記した基準により音声素片を選択する場合、一つの発話文の目標情報の全てが必要となるので、このままでは一つの文の入力が終了するまで素片選択が実行できない。本実施の形態では、応答を速くするため、文の入力が終了する前に素片選択の開始が行なえるようにする。そのため、以下に述べるように、合計コストの算出方法を改良し、深さ制限をしたビームサーチを採用した音声素片の探索方法を用いる。

【0029】

20

本実施の形態では、素片選択の遅延を少なくするため、上記したコスト計算において p_T および p_C をいずれも 1 に固定する。すると、時刻 i における累積コスト $C(i) = C(u_0, u_1, \dots, u_{i-1})$ は以下の式で与えられる。

【0030】

【数3】

$$\begin{aligned}
 C(0) &= w'_T \cdot C_T(u_0, t_0) \\
 C(i) &= C(i-1) \\
 &+ w'_C \cdot C_C(u_{i-1}, u_i) \\
 &+ w'_T \cdot C_T(u_i, t_i) \quad (i \geq 1) \quad (4)
 \end{aligned}$$

30

ただし w'_C および w'_T はそれぞれ、ターゲットコストと接続コストとの重みである。

【0031】

式(4)は漸化式で、時刻 i のコストが、時刻 $i-1$ までの累積コストと時刻 i における音声素片候補の特徴量とによって決定されるので、累積コストを効率よく算出できる。前述した式で p_T および p_C がそれぞれ 1 と 1.5 であったことを考えると、式(4)を採用したとしてもその影響は少ないと考えられる。

【0032】

遅延時間をより小さくするためには、上記したようにコスト計算の方法を変えた上で、さらに音声素片選択のための探索空間を時間的に制限する必要がある。そのために本実施の形態では、深さ制限をしたビームサーチを採用して音声素片の選択を行なう。本明細書では「深さ」とは探索木の根ノードから葉ノードまでのノード数のことをいう。ただし根ノード自体はこの数には数えない。この場合、深さの最大値は音声素片のための最大遅延に対応する。

40

【0033】

この方法では、すべての枝の深さは互いに等しい。深さが深さの制限値を超えたなら、根ノードの直下の子ノードのうちいずれかが次の根ノードとしてまず保存される。この子ノードは、最小の累積コストを与える枝を逆にたどることで決定される。

【0034】

次に、根ノードの子ノードのうち、保存された子ノード以外のノードとそれ以下のノードとを探索木から削除する。これら子ノードのコストが最適解を与えるものとは考えられ

50

ないためである。この後、根ノードを削除し、保存しておいた子ノードを新たな根ノードとする。この新たな根ノードにおける音声素片が、音声合成のために選択される。こうした処理により、探索対象となる探索木の深さは深さ制限以内に維持される。

【 0 0 3 5 】

図3は、ビーム幅 $w = 4$ 、深さ制限 $L = 2$ の場合の探索木の例を示す。図3において、各ノード内のアルファベットは音素名ではなくノード名を示す。時刻0において、根ノードがノードAとすると、ノードAから深さが3の階層にある葉ノードH, I, J, K, L, M, N, Oのうちで累積コストが最小となるのは、A - C - F - Mという経路である。この結果、ノードCが時刻1において選択されるノード（音声素片）となる。A - B - D - HおよびA - B - D - Iという経路は枝刈りされる（削除される）。図3において、このようにして枝刈りされた枝およびノードは点線で示してある。その後ノードAを削除し、ノードCを新たな根ノードとする。さらに、ノードCの下にある葉ノードJ, K, L, M, N, Oのうち、ビーム幅 $w = 4$ を超える葉ノードが枝刈りされる。図3に示す例では、コストが大きくなる経路C - G - OおよびC - E - Kが枝刈りされ、葉ノードJ, L, M, Nが残される。この処理で枝刈りされる枝およびノードは、図3では破線で示してある。

10

【 0 0 3 6 】

実際には、探索速度を制御するために、ビーム幅を変えることが必要である。累積コスト計算のための計算量には、ビーム幅が影響するためである。連続して音声合成出力を行なうためには、音声素片選択はリアルタイム以上の速度で実行される必要がある。

20

【 0 0 3 7 】

上記した音声合成装置36は、実際にはコンピュータハードウェアおよび当該コンピュータハードウェア上で実行されるソフトウェアとからなるコンピュータシステムにより実現される。図4はこの音声合成装置36を実現するコンピュータシステム130の外観を示し、図5はコンピュータシステム130の内部構成を示す。

【 0 0 3 8 】

図4を参照して、このコンピュータシステム130は、FD（フレキシブルディスク）ドライブ152およびCD-ROM（コンパクトディスク読出専用メモリ）ドライブ150を有するコンピュータ140と、キーボード146と、マウス148と、モニタ142とを含む。

30

【 0 0 3 9 】

図5を参照して、コンピュータ140は、FDドライブ152およびCD-ROMドライブ150に加えて、CPU（中央処理装置）156と、CPU156、FDドライブ152およびCD-ROMドライブ150に接続されたバス166と、ブートアッププログラム等を記憶する読出専用メモリ（ROM）158と、バス166に接続され、プログラム命令、システムプログラム、および作業データ等を記憶するランダムアクセスメモリ（RAM）160とを含む。コンピュータシステム130はさらに、プリンタ144を含んでいる。

【 0 0 4 0 】

ここでは示さないが、コンピュータ140はさらにローカルエリアネットワーク（LAN）への接続を提供するネットワークアダプタボードを含んでもよい。

40

【 0 0 4 1 】

コンピュータシステム130に音声合成装置36としての動作を行なわせるためのコンピュータプログラムは、CD-ROMドライブ150またはFDドライブ152に挿入されるCD-ROM162またはFD164に記憶され、さらにハードディスク154に転送される。または、プログラムは図示しないネットワークを通じてコンピュータ140に送信されハードディスク154に記憶されてもよい。プログラムは実行の際にRAM160にロードされる。CD-ROM162から、FD164から、またはネットワークを介して、直接にRAM160にプログラムをロードしてもよい。

【 0 0 4 2 】

50

このプログラムは、コンピュータ140にこの実施の形態に係る音声合成装置36としての動作を行なわせる複数の命令を含む。この方法を行なわせるのに必要な基本的機能のいくつかはコンピュータ140上で動作するオペレーティングシステム(OS)またはサードパーティのプログラム、もしくはコンピュータ140にインストールされる各種ツールキットのモジュールにより提供される。したがって、このプログラムはこの実施の形態のシステムおよび方法を実現するのに必要な機能全てを必ずしも含まなくてよい。このプログラムは、命令のうち、所望の結果が得られるように制御されたやり方で適切な機能または「ツール」を呼出すことにより、上記した音声合成装置36としての動作を実行する命令のみを含んでいればよい。コンピュータシステム130の動作は周知であるので、ここでは繰返さない。

10

【0043】

図6は、コンピュータシステム130に音声合成装置36としての動作を実行させるためのプログラムのうち、素片選択部66を実現するプログラム部分の制御構造を示すフローチャートである。図6を参照して、まずステップ180で初期化処理が実行される。このステップでは、必要とする記憶領域の確保および初期化、必要な定数のロードなどが実行される。

【0044】

ステップ182では、入力要求があるか否かについての判定が行なわれる。この入力要求は、設計遅延時間での音声合成が可能ないように外部から制御される。入力要求があればステップ184に進み、入力要求がなければステップ188に進む。

20

【0045】

ステップ184では、合成器指令を読み込む。この合成器指令は、図1に示すテキスト処理部64により生成された合成器指令を読み込む処理である。続いてステップ186では、読み込まれた合成器指令に回答し、合成器指令にしたがって、素片候補追加および探索木更新が実行される。ステップ186で実行される処理の詳細については図7を参照して後述する。

【0046】

ステップ188では、出力要求があるか否かについての判定が行なわれる。この出力要求も、コンピュータシステム130に備えられたタイマ等に基づき、設計遅延時間での音声合成が可能ないように外部から制御されるものとする。通常は、探索木が深くなるほど音声素片の選択結果はよくなるので、基本的にはこの出力要求は、設計遅延時間におさまる最大の遅延になるように制御される。出力要求があればステップ190に進み、出力要求がなければステップ194に進む。

30

【0047】

ステップ190では、素片確定・探索木更新処理が実行される。ここで実行されるのが、上記した探索木内での深さ制限によるビームサーチによる探索である。この処理では、図3を参照して説明したとおり、根ノードの直下の子ノードのうち、深さ制限により決まる葉ノードまでの経路の累積コストが最小となるようなものが一旦保存される。さらに、根ノードの子ノードのうち、それ以外のノードが、各ノードの下に存在するノードとともに削除される。根ノードが削除され、保存されていた子ノードが新たな根ノードとなる。さらにこの根ノードのうち、ビーム幅を超えるノードがコストの大きいものから順に削除される。

40

【0048】

ステップ192では、ステップ190で新たな根ノードとなったノードに対応する素片が選択され、選択結果として出力される。この後、ステップ194に進み、ステップ194では全ての入力に対して出力を行なったか否かが判定される。判定結果がYESなら処理は終了し、判定結果がNOであれば制御はステップ182に戻る。

【0049】

図7は、図6のステップ186で実行される素片候補追加・探索木更新処理(ビームサーチ処理)を実現するためのプログラム部分の制御構造を示すフローチャートである。図

50

7を参照して、まずステップ270で、合成器指令入力に対応する素片候補列を配列 $\{u_0, u_1, \dots, u_{k-1}\}$ に代入する。同じく、葉ノード番号のリストを配列 $\{I_0, \dots, I_{w-1}\}$ (w はビーム幅)に代入する。続いてステップ272で変数 i に0を代入する。

【0050】

ステップ274では、変数 i の値が素片候補列の素片候補数 k より小さいか否かを判定する。変数 i の値が k より小さければステップ276に進む。変数 i の値が k 以上であればステップ288に進む。

【0051】

ステップ276では、変数 j に0を代入する。続くステップ278では、変数 j の値がビーム幅 w より小さいか否かを判定する。変数 j がビーム幅 w より小さければステップ280に進み、さもなければステップ286に進む。

【0052】

ステップ280では、葉ノード I_j に登録されているコストおよび素片候補と素片候補 u_i とを用いて累積コスト $C(I_j, u_i)$ を算出し、コストを表す変数 C に代入する。続いてステップ282で、ノード I_j の子ノードを作成し、そのノードに素片候補 u_i とコスト C とを登録する。ステップ284で変数 j に1を加算し、ステップ278に戻る。こうして、全ての葉ノード I_j の下に素片候補 u_i に対応する子ノードが作成され、それらにはそれぞれ素片候補 u_i と累積コストとが登録される。

【0053】

ステップ278での判定の結果、変数 j の値が葉ノードの数と等しくなると、ステップ286で変数 i に1を加算し、ステップ274に戻る。ステップ274からステップ286の処理が、ステップ274で変数 i の値が変数 k の値と等しいと判定されるまで繰返される。すなわち、全ての素片候補について、全ての葉ノードの下に対応の子ノードを作成し、それら全てに素片候補と累積コストとが登録される。

【0054】

ステップ274において、変数 $i < 変数k$ ではないと判定されると、ステップ288においてビーム幅制限による枝刈りが実行される。この枝刈りについては図8を参照して後述する。ステップ288の処理が終了すると、素片候補追加・探索木更新処理は終了する。

【0055】

図8は、図7のステップ288で実行されるビーム幅制限による枝刈りを実現するプログラム部分のフローチャートである。図8を参照して、まずステップ300で探索木の葉ノードの数がビーム幅 w より大きいと判定する。葉ノードの数がビーム幅 w より大きければステップ302に進み、それ以外の場合にはこの処理を終了する。

【0056】

ステップ302では、葉ノードのうちで最大の累積コストを持つものの番号を変数 m に代入する。この後、ステップ304でノード m を削除する。ノードの削除については図10を参照されたい。

【0057】

図9は、図6のステップ190で実行される素片確定・探索木更新処理のフローチャートである。図9を参照して、この処理では、まずステップ210で累積コストが最小となる葉ノードの番号を変数 m に代入する。ステップ212で葉ノード m (変数 m により表されるノードのことをノード m と呼ぶ。以下同じ) から根ノードに向けて木構造をたどり、その経路上での根ノードの直下の子ノードの番号を変数 n に格納する。すなわち、新たな根ノードとなるノードの番号を変数 n に保存する。なお以下のフローチャートの説明では、変数として i 、 j などが使用される。これら変数は、図が異なればたとえ同じ名称 (i 、 j など) であっても互いに異なる変数であることに注意されたい。

【0058】

ステップ214では、以下の処理の繰返し変数 i に0を代入し、ステップ216で変数 i の値が葉ノードの数より大きいと判定する。変数 i の値が葉ノードの数より大き

10

20

30

40

50

ければステップ 2 2 6 に進み、それ以外の場合にはステップ 2 1 8 に進む。

【 0 0 5 9 】

ステップ 2 1 8 では、葉ノード i から根ノードに向けて木構造をたどり、当該経路上における根ノードの直下の子ノード番号を変数 j に格納する。続いてステップ 2 2 0 で、変数 j の値が変数 n の値と等しいか否かについて判定する。両者が等しければステップ 2 2 4 に進み、両者が等しくなければステップ 2 2 2 に進む。

【 0 0 6 0 】

ステップ 2 2 2 では、変数 i により示される葉ノードが、変数 n により示されるノードの下にないということであるから、その葉ノードを削除する。ステップ 2 2 2 の処理については図 1 0 を参照して後述する。ステップ 2 2 2 の後、処理はステップ 2 2 4 に進む。

10

【 0 0 6 1 】

ステップ 2 2 4 では、変数 i に 1 を加算し、ステップ 2 1 6 に戻る。

【 0 0 6 2 】

以上の処理により、葉ノードのうちで、次の根ノードになるノードの下に存在するもの以外の葉ノードが全て削除される。

【 0 0 6 3 】

一方、ステップ 2 1 6 で変数 i の値が葉ノード数より大きいと判定された場合には、全ての葉ノードに対しステップ 2 1 8 ~ 2 2 4 の処理が完了したということである。そこで、ステップ 2 2 6 でそれまでの根ノードを削除し、変数 n により特定されるノードが新たな根ノードとなる。

20

【 0 0 6 4 】

図 1 0 は、図 9 のステップ 2 2 2 で実行されるノード削除のためのプログラムの制御構造を示すフローチャートである。図 1 0 を参照して、まずステップ 2 4 0 で、削除する枝の子ノードの番号を変数 i に代入する。ステップ 2 4 2 で、ノード i が根ノードか否かを判定する。ノード i が根ノードであれば、根ノードの子ノードで削除すべきものは全て削除したということである。したがってこの場合には処理を終了する。ノード i が根ノードでなければステップ 2 4 4 に進む。

【 0 0 6 5 】

ステップ 2 4 4 では、ノード i から子ノードへの枝数が 0 か否かを判定する。枝数が 0 以外であれば、このノードを削除できないので処理を終了する。枝数が 0 であればステップ 2 4 6 に進み、変数 j にノード i の親ノードの番号を代入する。

30

【 0 0 6 6 】

続いてステップ 2 4 8 で、枝 $j i$ (ノード j とノード i とを結ぶ枝) とノード i とを削除する。ステップ 2 5 0 で変数 j の値を変数 i に代入し、ステップ 2 4 2 に戻る。以下、ステップ 2 4 2 で判定が YES またはステップ 2 4 4 で判定が NO となるまで、上記した処理を繰返すことにより、根ノードの直下の子ノードまで、不要なノードが削除される。

【 0 0 6 7 】

上記した素片選択部 6 6 の動作によるノードの選択 (すなわち素片候補の選択) は、図 3 を参照して説明したとおりとなる。したがって、ここではそれらについての詳細な説明は繰返さない。

40

【 0 0 6 8 】

以上のようにこの第 1 の実施の形態に係る音声合成装置 3 6 によれば、コスト計算を漸化式にしたがって行なうようにし、さらに深さ制限を用いたビームサーチによって所定の計算量内でコスト最小の探索木上を探索することにより、素片候補の選択を所定時間ごとに逐次的に行なう。そのため、音声素片 DB 6 2 に含まれる音声素片数が大きくなっても、短い時間で素片を選択できる。そのため合成音声の品質を維持することができる。深さ制限を行なっているため、所定の短遅延時間ごとに素片を選択でき、合成器指令の入力の完了を待つ必要はない。その結果、実時間で高品質の音声合成を行なうことができる。

【 0 0 6 9 】

[第 2 の実施の形態]

50

第1の実施の形態では、探索木の探索に深さ制限を課したビームサーチを採用した。しかし、本発明はそのような実施の形態には限定されない。ビームサーチではなく、同じように深さ制限を課したDP（動的プログラミング）サーチを採用することもできる。第2の実施の形態に係る音声合成装置は深さ制限を課したDPサーチを採用したものである。

【0070】

この第2の実施の形態に係る音声合成装置の構成は、図1に示したものと同様で、ただし素片選択部66で実行される探索方法のみが第1の実施の形態と異なる。したがって、以下では本実施の形態で第1の実施の形態で用いたビームサーチに替えて採用された深さ制限が課されたDPサーチについてのみ説明する。

【0071】

ビームサーチの場合と異なり、DPサーチでは、サーチの途中の各時刻において、探索木のあるノードから分岐するノードに対応する音声素片は、互いに異なっている必要がある。また、ある時刻において可能な音声素片の数がビーム幅未満であれば、その時刻におけるノードの数はビーム幅ではなくその可能な音声素片の数となる。すなわち、各時刻における探索木の葉ノードの数は、ビーム幅以下の値となる。

【0072】

図11は、深さ制限が課されたDPサーチを説明するための探索木を示す図である。ここでは、ビーム幅 $w = 4$ であり、深さ制限 $L = 2$ とする。図11に示す例では、時刻3において可能な音声素片の数は2となっているものとする。したがって、図11において、ノードH, J, L, Nは互いに同じ音声素片に対応し、ノードI, K, M, Oも互いに同じ音声素片に対応することが想定されている。このサーチでは、可能な音声素片の各々に対し、枝との間の関係を調べて最適な枝が一つだけ作成される。したがって葉ノードの数の総和は、時刻3においては2である。

【0073】

ビーム幅およびある時刻における音声素片候補の数をそれぞれ w' および n とすると、累積コストを算出するための計算量は $O(w'n)$ である。これはビームサーチにおける計算量 $O(w'n)$ と同程度である。したがって、もし $w = w'$ であれば、DPサーチの場合にはビームサーチと比較して探索木内により多くの種類の音声素片を持つことができるので、ビームサーチよりもDPサーチの方がより効率的な音声素片の選択ができる。これは、DPサーチの場合には無駄な枝の数がビームサーチの場合より少ないためである。

【0074】

しかし、短遅延の音声合成では探索木の枝を刈り込む必要が生ずる。DPサーチの場合にビームサーチと同様の枝刈りを行なうと、サーチの際に有望な音声素片候補が多数捨てられてしまう可能性がある。図11に示す例では、時刻3において可能な二つの音声素片のうち一方は、次の枝刈りで削除されてしまう。したがって、深さ制限をしたDPサーチにおける音声素片選択の結果は悪化する可能性がある。

【0075】

図11に示す例では、時刻0における根Aから、時刻3における可能なノードの累積コストを見ると、ノードMが最小(1.4)である。したがって、時刻1では、ノードB、Cのうち、ノードM側のノードCが根ノードとして選択される。ノードBと、ノードBから枝分かれするノード(ノードD, H, I)とは全て削除される。削除された枝およびノードは、図11では点線で示してある。

【0076】

さらに、可能なノード数が2であるから、残りのノードのうちノードJ, M以外のノード(ノードK, L, N, O)が削除される。ノードGからの葉ノードが一つもなくなるので、ノードGも削除される。ここで削除されるノードは、図11では破線で示してある。

【0077】

図12は、図11に示すようなDPサーチを実現するためのプログラム部分のフローチャートである。図12に示す例は、DPに基づく方法に相当するが、ここではN-ベスト

10

20

30

40

50

サーチとビームサーチとを組合せた方法としてより一般化したものを示してある。N = 1 のときはDPサーチに相当する。

【0078】

図12に示す例では、各音声素片候補について、コスト計算結果をヒープ領域に格納した後、コストの小さい方からN個の可能性のみを探索木に追加する。Nを小さくすれば、最適ではない特定の素片ばかりが探索木に登録されてしまうことを防止できる。一方、Nを大きくすれば、出力結果を確定させたとき必要となる枝刈り操作において、ある音声素片候補が完全に探索木から失われてしまうことが防止できる。

【0079】

図12を参照して、まずステップ320において、合成器指令入力に対応する素片候補列を配列 $\{u_0 \dots u_{k-1}\}$ に代入し、葉ノード番号のリストを配列 $\{I_0 \dots I_{w-1}\}$ に代入する。ただしkは素片候補の数、wはビーム幅である。

【0080】

ステップ322において、繰返し変数iに0を代入する。続いてステップ324において変数iの値が素片候補の数k未満か否かを判定する。判定結果がYESであればステップ326に進み、それ以外であればステップ352に進む。

【0081】

ステップ326では、サイズNのヒープを初期化する。ヒープとは、ノードに値が保持される木構造で、親ノードの値が子ノードの値よりも常に大きく（または常に小さく）なっているものをいう。この場合、根に最大（最小）の値が格納される。各ノードの値の間に大小の関係を課す事以外には、木の構造に意味はない。したがって、木の構造に強い制約を課すことが可能である。ヒープは、リスト構造ではなく配列を用いて実装するのが普通である。深さが等しい2分木K構造により実現されている場合、ヒープの挿入・削除に要する処理は最悪で $\log N$ となる。ヒープの取扱いに関しては、古典的なアルゴリズムが存在する（ドナルド E. クヌース (Donald E. Knuth)、「The Art of Computer Programming, Volume 3: Sorting and Searching」、1973年）。

【0082】

ステップ328では、変数jに0を代入する。ステップ330では、変数jがビーム幅wより小さいか否かを判定する。変数jがビーム幅wより小さければステップ332に進み、それ以外の場合にはステップ340に進む。

【0083】

ステップ332では、葉ノード I_j と素片候補 u_i との間で累積コスト $C(I_j, u_i)$ を計算し、コストCに代入する。続いてステップ334において、コストCの値が、ヒープ中で最大のコストより小さいか、またはヒープに空きがあるか否かを判定する。判定結果がYESであればステップ336に進み、さもなければステップ338に進む。

【0084】

ステップ336では、コストCをキーにして $\{u_i, I_j, C\}$ をヒープに登録し、ステップ338に進む。

【0085】

ステップ338では、変数jに1を加算し、ステップ330に戻る。

【0086】

ステップ340では、N(N - ベストのN)を変数hに代入する。ステップ342では変数hの値が0より大きいか否かを判定する。変数hの値が0より大きいときにはステップ344に進み、さもなければステップ350に進む。

【0087】

ステップ344では、小さい方からh番目のコストCを含む $\{u, I, C\}$ をヒープから取除く。ステップ346では、Iの子ノードを作成し、そのノードにuとコストCとを登録する。その後ステップ348において変数hから1を減算し、ステップ342に戻る。

。

10

20

30

40

50

【 0 0 8 8 】

ステップ 3 4 2 で変数 h が 0 以下であると判定されると、ステップ 3 5 0 で変数 i に 1 を加算し、ステップ 3 4 2 に戻る。

【 0 0 8 9 】

一方、ステップ 3 2 4 で変数 i の値が素片候補の数 k 未満ではないと判定されたときは、ステップ 3 5 2 でビーム幅を w に制限した枝刈りを行ない、処理を終了する。

【 0 0 9 0 】

以上のプログラムにより、深さ制限を課した DP サーチを行なうことができる。

【 0 0 9 1 】

[深さ制限によるサーチの評価]

深さ制限を課したビームサーチおよび DP サーチを評価するために、ビーム幅および深さ制限を様々に変えて累積コストを算出した。この評価で合成に使用した発話コーパスは、一人の男性話者による 4 5 0 文の発話を収録したものである。この発話は、非特許文献 3 に記載の 5 0 3 文コーパスのセット A からセット I までのものと同じものである。合成目標は、上記した 5 0 3 文コーパスのセット J の 5 3 文である。合成目標に関する情報は、同じ話者が発話したセット J の文から抽出した。データベース内の各音声素片は音素に対応する。音声素片の数は、無音をのぞいて 2 3 , 6 3 2 である。テスト文に含まれる音声素片の平均数は 4 3 . 3 であった。

【 0 0 9 2 】

図 1 3 および図 1 4 はそれぞれ、深さ制限を課したビームサーチと深さ制限を課した DP サーチとによって選択された音声素片の、1 素片辺りの平均累積コストを示す。図 1 3 には、ビーム幅 $w = 1 6$ 、6 4、2 5 6、1 0 2 4、4 0 9 6 として得られた結果が、それぞれ折れ線 3 7 0、3 7 2、3 7 4、3 7 6、3 7 8 として示されている。また図 1 4 には、ビーム幅 $w = 1 6$ 、6 4、2 5 6、1 0 2 4 に対して得られた結果が、それぞれ折れ線 4 0 0、4 0 2、4 0 4、および 4 0 4 ($w = 2 5 6$ の折れ線と $w = 1 0 2 4$ の折れ線とはほとんど重なっている。) で示されている。

【 0 0 9 3 】

この結果から、ビーム幅が等しい場合には、DP サーチの方がビームサーチより累積コストの点で優れていることが分かる。例外は深さ制限の値が小さく、ビーム幅が大きくなる場合である。ビームサーチでは、制限された深さが大きくなると累積コストは悪化する。これにより、ビームサーチでの探索木が非効率であることが分かる。DP サーチでは、ビームサーチのときと異なり、深さ制限が小さいときにはビーム幅は累積コストの低減には貢献しない。これは枝刈りの影響によるものと思われる。

【 0 0 9 4 】

一般的にあって、遅延時間を非常に小さくすることが求められる場合には、深さ制限を課したビームサーチを採用すべきである。他の場合には、DP サーチの方が優れている。もちろん結果は様々な要因により左右される。例えばコスト関数の設計、音声素片のカテゴリ分けの仕方、音声素片データベースの構造、評価で用いるテスト文などがそうした要因となる。しかし、将来の音声合成器のコストも、音声合成の原理に大きな変化が生じない限り、上記したのと同様の傾向を示すであろう。また、どんな場合にも、DP サーチの方がビームサーチよりもよい結果を示すような最小の深さを確認すべきである。

【 0 0 9 5 】

今回開示された実施の形態は単に例示であって、本発明が上記した実施の形態のみに制限されるわけではない。本発明の範囲は、発明の詳細な説明の記載を参酌した上で、特許請求の範囲の各請求項によって示され、そこに記載された文言と均等の意味および範囲内のすべての変更を含む。

【 図面の簡単な説明 】

【 0 0 9 6 】

【 図 1 】 本発明の第 1 の実施の形態に係る音声合成装置のブロック図である。

【 図 2 】 図 1 に示す装置で行なわれる処理の時間的関係を説明するための図である。

10

20

30

40

50

【図3】ビームサーチによる葉ノード選択処理を説明するための図である。

【図4】本発明の一実施の形態に係る音声合成装置36を実現するためのコンピュータシステムの外観を示す図である。

【図5】図4に示すコンピュータシステムのブロック図である。

【図6】ビームサーチを実現するためのプログラムのメイン処理の制御構造を示すフローチャートである。

【図7】第1の実施の形態における素片候補追加・探索木更新処理のフローチャートである。

【図8】ビーム幅制限による枝刈り処理のフローチャートである。

【図9】素片確定・探索木更新処理のフローチャートである。

10

【図10】ノード削除処理のフローチャートである。

【図11】第2の実施の形態における葉ノード選択処理を説明するための図である。

【図12】第2の実施の形態における素片候補追加・探索木更新処理のフローチャートである。

【図13】深さ制限を課したビームサーチの評価結果を示すグラフである。

【図14】深さ制限を課したDPサーチの評価結果を示すグラフである。

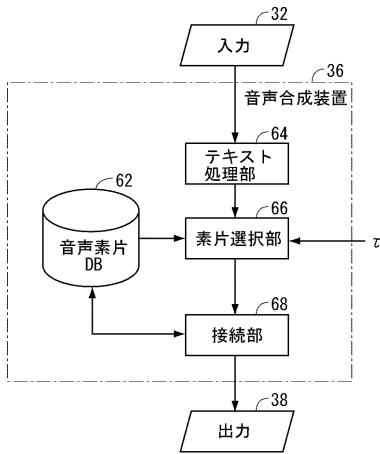
【符号の説明】

【0097】

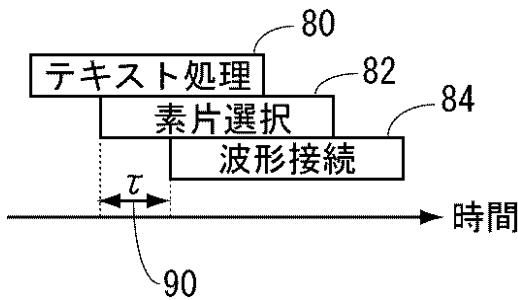
32 入力、36 音声合成装置、38 出力、62 音声素片DB、64 テキスト処理部、66 素片選択部、68 接続部

20

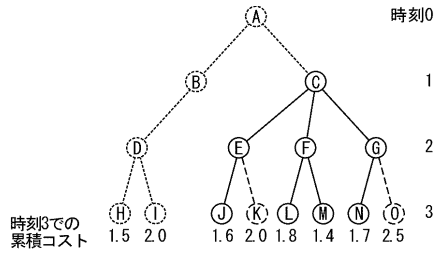
【図1】



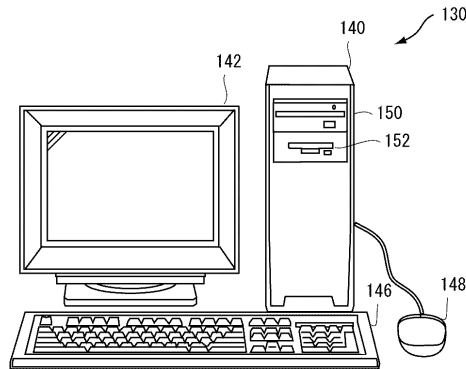
【図2】



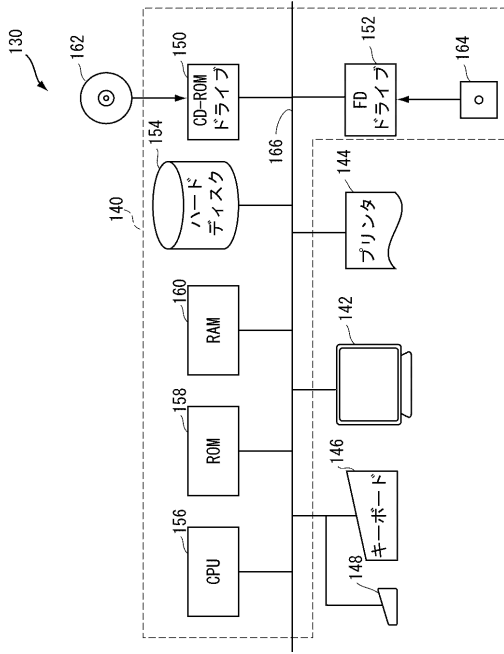
【図3】



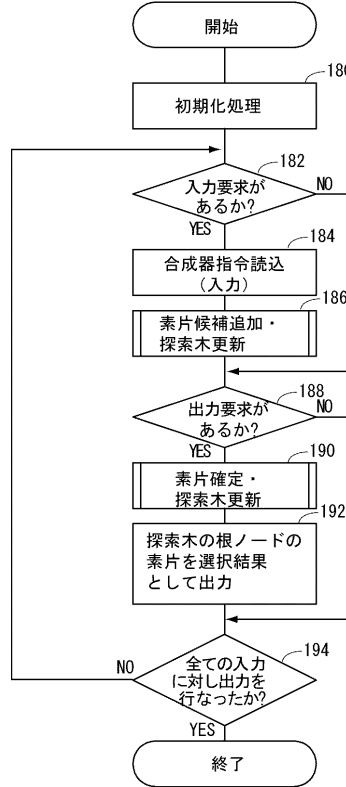
【図4】



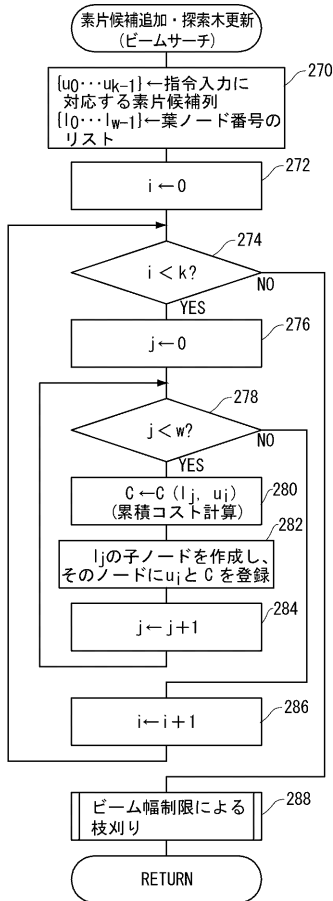
【図5】



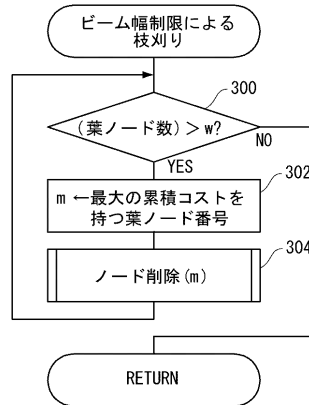
【図6】



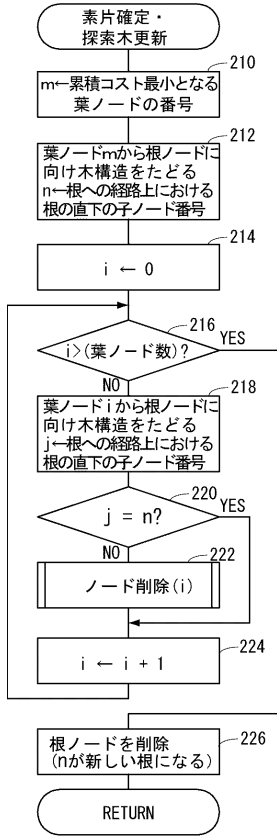
【図7】



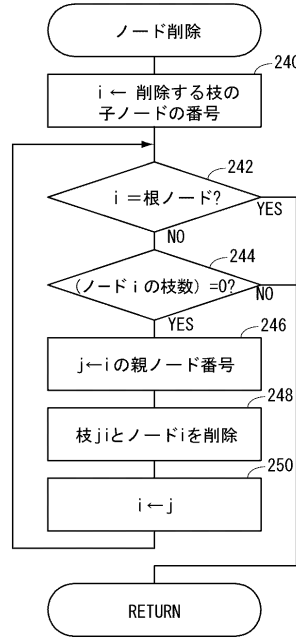
【図8】



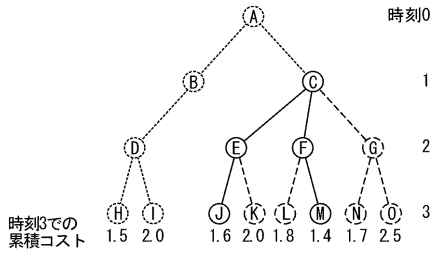
【図9】



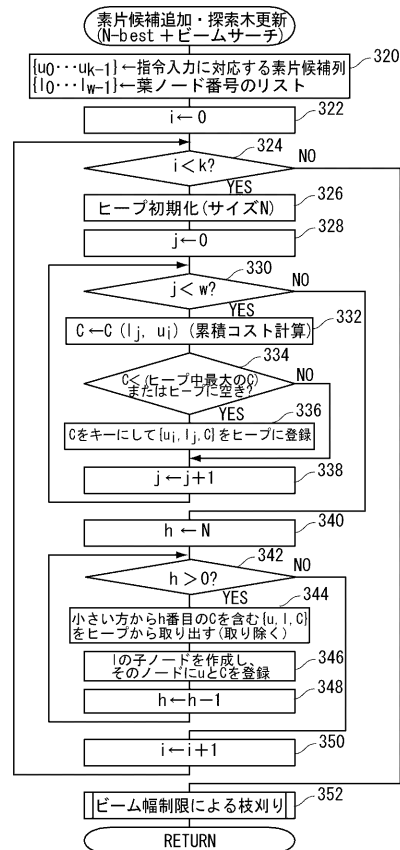
【図10】



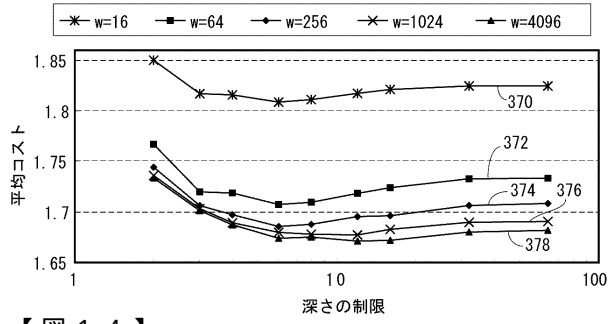
【図11】



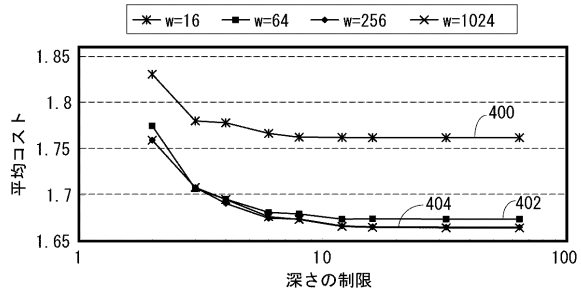
【図12】



【 図 1 3 】



【 図 1 4 】



フロントページの続き

- (56)参考文献 特表2002-530703(JP,A)
特開平10-097290(JP,A)
特表2000-516749(JP,A)
特開2001-242884(JP,A)

- (58)調査した分野(Int.Cl., DB名)
G10L 13/00 - 17/00