(19)日本国特許庁(JP)　(12)特　許　公　報(B2)　(11)特許番号
特許第4424024号
(P4424024)
(45)発行日　平成22年3月3日(2010.3.3)　(24)登録日　平成21年12月18日(2009.12.18)

(51)Int.Cl.　F I
G10L 13/06 (2006.01)　G10L 13/06　220C

請求項の数 14　(全 16 頁)

(21)出願番号　特願2004-75185(P2004-75185)
(22)出願日　平成16年3月16日(2004.3.16)
(65)公開番号　特開2005-266010(P2005-266010A)
(43)公開日　平成17年9月29日(2005.9.29)
審査請求日　平成19年2月26日(2007.2.26)

特許権者において、実施許諾の用意がある。

(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
(74)代理人　100099933
弁理士　清水　敏
(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
(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

審査官　井上　健一

最終頁に続く

(54)【発明の名称】素片接続型音声合成装置及び方法

(57)【特許請求の範囲】
【請求項１】
合成音声の目標と音声素片候補との間で、複数のサブコストを含むコストを算出し、当該コストに基づいて、複数の音声素片候補を含む音声素片データベースから音声素片を選択し接続することにより音声合成を行なう素片接続型音声合成装置であって、各音声素片候補は、各音声素片の音響特徴量データと、各音声素片の波形データとを含み、
前記音声素片データベースを記憶するための第１の記憶装置と、
前記音声素片データベースに記憶された複数の音声素片候補の音響特徴量データと、前記音声素片データベースに記憶された複数の音声素片候補の中で、所定の基準で選択された音声素片候補の波形データとを記憶するための、前記第１の記憶装置より高速アクセス可能な第２の記憶装置とを含み、
前記複数のサブコストは、各音声素片候補の波形データが記憶されている記憶装置へのアクセス速度に関するアクセス速度コストを含み、
前記複数の音声素片候補の各々の音響特徴量データは、当該音声素片候補の波形データが前記第１及び第２の記憶装置のいずれに記憶されているかを示す第１のフラグを含み、
前記音声合成装置はさらに、
前記合成音声の目標との間で、前記複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、前記第２の記憶装置に記憶された音響特徴量に基づいて前記複数の音声素片候補から選択するための選択手段と、
前記選択手段により選択された音声素片候補の音声波形を、当該選択された音声波形に

対応する第１のフラグに基づいて、前記第１の記憶装置又は前記第２の記憶装置のいずれかから読出して前記合成音声の目標に従って接続し、合成音声波形を出力するための接続手段とを含む、素片接続型音声合成装置。

【請求項２】

前記第１の記憶装置に記憶された前記音声素片データベースは、当該音声素片データベースに含まれる複数の音声素片候補のうち、対応する波形データを前記第２の記憶装置に記憶すべき音声素片候補を選択するための基準となる選択基準情報が付されており、

前記素片接続型音声合成装置はさらに、前記第１の記憶装置に記憶された前記音声素片データベースに含まれる前記複数の音声素片候補のうち、前記選択基準情報により選択された音声素片候補の波形データを前記第２の記憶装置にロードするためのロード手段を含む、請求項１に記載の素片接続型音声合成装置。

【請求項３】

テストデータに基づいて、前記音声素片データベースの前記選択基準情報を生成するための選択基準情報生成手段をさらに含む、請求項２に記載の素片接続型音声合成装置。

【請求項４】

前記選択基準情報生成手段は、

前記テストデータに基づき、前記音声素片データベースに含まれる音声素片候補を使用して音声合成をシミュレートするための音声合成シミュレート手段と、

前記音声素片データベースに含まれる前記複数の音声素片候補の各々について、前記音声合成シミュレート手段による音声合成の際に選択された頻度を記録するための頻度記録手段とを含み、

前記選択基準情報は、前記頻度記録手段により記録された、前記音声素片データベースに含まれる前記複数の音声素片候補の各々の頻度情報である、請求項３に記載の素片接続型音声合成装置。

【請求項５】

前記ロード手段は、前記頻度情報に基づき、前記音声素片データベースに含まれる音声素片のうち、前記音声合成シミュレート手段による音声合成で選択された頻度の高いものを上位から所定個数選択し、選択された音声素片の波形データを前記第２の記憶装置にロードするための手段を含む、請求項４に記載の素片接続型音声合成装置。

【請求項６】

前記音声合成シミュレート手段は、前記テストデータから得られる合成音声の目標と、前記音声素片データベースに含まれる前記複数の音声素片候補との間で、前記複数のサブコストのうち、前記アクセス速度コストを除くアクセスコストと、前記頻度記録手段により記録された各音声素片候補が選択された頻度に基づいて算出される選択頻度コストとを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、前記音声素片データベースから選択するための手段を含む、請求項４又は請求項５に記載の素片接続型音声合成装置。

【請求項７】

前記素片接続型音声合成装置は、

キャッシュメモリと、

前記キャッシュメモリに対して設けられたキャッシュメモリ管理機構と、

前記複数の音声素片候補に対応して設けられ、対応の音声素片候補が前記選択手段により選択されたか否かを記録するための第２のフラグを記憶するための手段と、

前記第２のフラグの値を所定の第１の値に初期化するための手段と、

前記複数の音声素片候補のいずれかが前記選択手段により選択されるたびに、選択された音声素片候補に対応する第２のフラグを前記第１の値と異なる所定の第２の値に更新するためのフラグ更新手段とをさらに含み、

前記アクセス速度コストは、前記第１及び第２のフラグに基づいて算出され、

前記選択手段は、前記合成音声の目標との間で、前記アクセス速度コストを含む前記複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を

、前記第２の記憶装置に記憶された音響特徴量に基づいて前記複数の音声素片候補から選択するための手段を含む、請求項１～請求項６のいずれかに記載の素片接続型音声合成装置。

【請求項８】

合成音声の目標と音声素片候補との間で、複数のサブコストを含むコストを算出し、当該コストに基づいて、複数の音声素片候補を含む音声素片データベースから音声素片を選択し接続することにより音声合成を行なう素片接続型音声合成方法であって、各音声素片候補は、各音声素片の音響特徴量データと、各音声素片の波形データとを含み、

前記音声素片データベースを第１の記憶ステップに記憶させるステップと、

前記音声素片データベースに記憶された複数の音声素片候補の音響特徴量データと、前記音声素片データベースに記憶された複数の音声素片候補の中で、所定の基準で選択された音声素片候補の波形データとを、前記第１の記憶装置より高速アクセス可能な第２の記憶装置に記憶させるステップとを含み、

前記複数のサブコストは、各音声素片候補の波形データが記憶されている記憶装置へのアクセス速度に関するアクセス速度コストを含み、

前記複数の音声素片候補の各々の音響特徴量データは、当該音声素片候補の波形データが前記第１及び第２の記憶装置のいずれに記憶されているかを示す第１のフラグを含み、

前記音声合成方法はさらに、

前記合成音声の目標との間で、前記複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、前記第２の記憶装置に記憶された音響特徴量に基づいて前記複数の音声素片候補から選択する選択ステップと、

前記選択ステップにおいて選択された音声素片候補の音声波形を、当該選択された音声波形に対応する第１のフラグに基づいて、前記第１の記憶装置又は前記第２の記憶装置のいずれかから読出して前記合成器指令に従って接続し、合成音声波形を出力する接続ステップとを含む、素片接続型音声合成方法。

【請求項９】

前記第１の記憶装置に記憶された前記音声素片データベースは、当該音声素片データベースに含まれる複数の音声素片候補のうち、対応する波形データを前記第２の記憶装置に記憶すべき音声素片候補を選択する基準となる選択基準情報が付されており、

前記素片接続型音声合成方法はさらに、前記第１の記憶装置に記憶された前記音声素片データベースに含まれる前記複数の音声素片候補のうち、前記選択基準情報により選択された音声素片候補の波形データを前記第２の記憶装置にロードするロードステップを含む、請求項８に記載の素片接続型音声合成方法。

【請求項１０】

テストデータに基づいて、前記音声素片データベースの前記選択基準情報を生成する選択基準情報生成ステップをさらに含む、請求項９に記載の素片接続型音声合成方法。

【請求項１１】

前記選択基準情報生成ステップは、

前記テストデータに基づき、前記音声素片データベースに含まれる音声素片候補を使用して音声合成をシミュレートする音声合成シミュレートステップと、

前記音声素片データベースに含まれる前記複数の音声素片候補の各々について、前記音声合成シミュレートステップによる音声合成の際に選択された頻度を記録する頻度記録ステップとを含み、

前記選択基準情報は、前記頻度記録ステップにおいて記録された、前記音声素片データベースに含まれる前記複数の音声素片候補の各々の頻度情報である、請求項１０に記載の素片接続型音声合成方法。

【請求項１２】

前記ロードステップは、前記頻度情報に基づき、前記音声素片データベースに含まれる音声素片のうち、前記音声合成シミュレートステップによる音声合成で選択された頻度の高いものを上位から所定個数選択し、選択された音声素片の波形データを前記第２の記憶装

置にロードするステップを含む、請求項１１に記載の素片接続型音声合成方法。
【請求項１３】
前記音声合成シミュレートステップは、前記テストデータから得られる合成音声の目標と、前記音声素片データベースに含まれる前記複数の音声素片候補との間で、前記複数のサブコストのうち、前記アクセス速度コストを除くアクセスコストと、前記頻度記録ステップにおいて記録された各音声素片候補が選択された頻度に基づいて算出される選択頻度コストとを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、前記音声素片データベースから選択するステップを含む、請求項１１又は請求項１２に記載の素片接続型音声合成方法。
【請求項１４】
前記素片接続型音声合成方法は、キャッシュメモリと、前記キャッシュメモリに対して設けられたキャッシュメモリ管理機構とを有するコンピュータ上で実行され、さらに、
前記複数の音声素片候補に対応して設けられ、対応の音声素片候補が前記選択ステップにおいて選択されたか否かを記録する第２のフラグを所定の記憶装置に記憶するステップと、
前記第２のフラグの値を所定の第１の値に初期化するステップと、
前記複数の音声素片候補のいずれかが前記選択ステップにおいて選択されるたびに、選択された音声素片候補に対応する第２のフラグを前記第１の値と異なる所定の第２の値に更新するステップとをさらに含み、
前記アクセス速度コストは、前記第１及び第２のフラグに基づいて算出され、
前記選択ステップは、前記合成音声の目標との間で、前記アクセス速度コストを含む前記複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、前記第２の記憶装置に記憶された音響特徴量に基づいて前記複数の音声素片候補から選択するステップを含む、請求項８～請求項１３のいずれかに記載の素片接続型音声合成方法。

【発明の詳細な説明】
【技術分野】
【０００１】
この発明は音声合成装置に関し、特に、所定のコスト関数に基づいて音声素片を選択し接続することにより合成器指令に合致した音声合成を行なう音声合成装置に関する。
【背景技術】
【０００２】
音声認識、音声合成は、人間とコンピュータを用いた諸システムとのインターフェースを実現する技術として重要である。これらと人工知能技術とを併用することにより、利用者は相手がコンピュータシステムであることを意識せずに様々なサービスを利用することができる。
【０００３】
中でも音声合成は、人間に対するシステム出力のためのインターフェースとしてその重要性は大きい。人間は、合成された音声の不自然さを敏感に感じ取る。合成された音声が不自然であると利用者が感じると、発話にも影響を及ぼし、その結果、人間とシステムとの間の対話がうまく行かなくなるおそれもある。
【０００４】
最近の音声合成技術としては、予め人間の発話を多数集めて語・音節・音素等を単位とする音声素片を音素ラベルと関連付けてデータベース化しておき、合成時には、指定された語・音節・音素等に対応する音声素片の中から、最も適切と思われるものを選択して接続するものが知られている。これを素片接続型音声合成と呼ぶ。なお、音素ラベルとは、通常は各音素の音素記号とその開始・終了時刻を記述したものをいう。これに加えて、その区間におけるＭＦＣＣ（Ｍｅｌ－Ｆｒｅｑｕｅｎｃｙ Ｃｅｐｓｔｒｕｍ Ｃｏｅｆｆｉｃｉｅｎｔ）、基本周波数（Ｆ０）等の音響特徴量、さらに前後の素片の音素記号を含

む場合もある。
【０００５】
素片接続型音声合成では、与えられた合成目標を基準として、いかにして適切な音声素片をデータベース中から取出すかが問題となる。
【０００６】
合成目標を構成するデータは、典型的には音素と、Ｆ０、持続時間、ＭＦＣＣ、及びパワー等の音声特徴量とを含む。これらを以下「合成器指令」と呼ぶ。
【０００７】
素片接続型音声合成では、合成器指令と音声素片のＦ０、持続時間、ＭＦＣＣ、パワー等とのずれ、及び接続に伴う自然劣化を表現するための「コスト」と呼ばれる評価関数を定義し、コストを最小とする音声素片を求めることにより、最適な音声素片系列を決定する。
【０００８】
本件出願の出願人は、上記した「コスト」を、それぞれある音声の特徴に対応するような「サブコスト」に分解し、それらを結合したもの（例えば線形和）により定義した素片接続型音声合成を提案している。例えば特許文献１を参照されたい。
【０００９】
サブコストには、物理量から計算されるものと、シンボリックな情報から事前に作成した規則から基づき得られるものとがある。前者は、複数のサンプル値に対する非線形演算であることも多く、その計算量は相対的に大きい。後者は、単純なテーブル参照の形であることが多く、テーブル参照で実現される場合にはサブコスト計算に必要な計算量は非常に少ない。
【００１０】
以上はあくまで一例であるが、この例に限らず、各サブコストの計算量はその種類により大きなばらつきがある場合が多い。
【００１１】
一方、上記とは別に、サブコストは、ターゲットコストに関係するものと接続コストに属するものとの二つに大別することもできる。ターゲットコストは、合成目標と素片候補との間の誤差を表す。接続コストは、合成音声において隣接する素片間の誤差（不連続性）を表す。
【００１２】
このような音声合成をリアルタイムで行なおうとする場合、いかにして素片選択と合成とを高速に行なうかが問題となる。この処理を高速化するためには、素片選択のコスト計算を高速に行なうとともに、選択された音声素片を接続する処理も高速にすることが望ましい。
【００１３】
実際に音声素片を接続する際には音声素片の波形データが必要となるが、個々の波形データのデータ量が比較的大きく、また、合成音声の品質を高くするためには、音声素片データベースに格納される音声素片の数を大きくする必要がある。その結果、音声素片データベース全体の容量は大きくなる。従って従来は、素片データのうち音響特徴量のみをメモリに格納して素片選択のコスト計算を行ない、音声素片データベースは、固定ハードディスク等比較的容量が大きな記憶装置に格納しておき、素片が選択された後、素片の接続時に波形データを読出すようにしている。
【００１４】
【特許文献１】特開２００３－２０８１８８号公報（段落００１４～００４７）
【発明の開示】
【発明が解決しようとする課題】
【００１５】
しかし、容量が大きな記憶装置のアクセス速度は比較的低速である。そのため、素片選択後、音声素片の波形データの取得に要する時間が大きく、音声合成に要する時間も長く

かかるという問題があった。

【００１６】

それゆえに、本発明の目的は、利用可能な音声素片の数を多く保ったまま、高速に音声合成を行なうことができる素片接続型音声合成装置及び方法を提供することである。

【課題を解決するための手段】

【００１７】

本発明の第１の局面に係る素片接続型音声合成装置は、合成音声の目標と音声素片候補との間で、複数のサブコストを含むコストを算出し、当該コストに基づいて、複数の音声素片候補を含む音声素片データベースから音声素片を選択し接続することにより音声合成を行なう素片接続型音声合成装置であって、各音声素片候補は、各音声素片の音響特徴量データと、各音声素片の波形データとを含み、音声素片データベースを記憶するための第１の記憶装置と、音声素片データベースに記憶された複数の音声素片候補の音響特徴量データと、音声素片データベースに記憶された複数の音声素片候補の中で、所定の基準で選択された音声素片候補の波形データとを記憶するための、第１の記憶装置より高速アクセス可能な第２の記憶装置とを含み、複数のサブコストは、各音声素片候補の波形データが記憶されている記憶装置へのアクセス速度に関するアクセス速度コストを含み、複数の音声素片候補の各々の音響特徴量データは、当該音声素片候補の波形データが第１及び第２の記憶装置のいずれに記憶されているかを示す第１のフラグを含み、音声合成装置はさらに、合成音声の目標との間で、複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、第２の記憶装置に記憶された音響特徴量に基づいて複数の音声素片候補から選択するための選択手段と、選択手段により選択された音声素片候補の音声波形を、当該選択された音声波形に対応する第１のフラグに基づいて、第１の記憶装置又は第２の記憶装置のいずれかから読出して合成音声の目標に従って接続し、合成音声波形を出力するための接続手段とを含む。

【００１８】

好ましくは、第１の記憶装置に記憶された音声素片データベースは、当該音声素片データベースに含まれる複数の音声素片候補のうち、対応する波形データを第２の記憶装置に記憶すべき音声素片候補を選択するための基準となる選択基準情報が付されており、素片接続型音声合成装置はさらに、第１の記憶装置に記憶された音声素片データベースに含まれる複数の音声素片候補のうち、選択基準情報により選択された音声素片候補の波形データを第２の記憶装置にロードするためのロード手段を含む。

【００１９】

さらに好ましくは、テストデータに基づいて、音声素片データベースの選択基準情報を生成するための選択基準情報生成手段をさらに含む。

【００２０】

選択基準情報生成手段は、テストデータに基づき、音声素片データベースに含まれる音声素片候補を使用して音声合成をシミュレートするための音声合成シミュレート手段と、音声素片データベースに含まれる複数の音声素片候補の各々について、音声合成シミュレート手段による音声合成の際に選択された頻度を記録するための頻度記録手段とを含んでもよい。選択基準情報は、頻度記録手段により記録された、音声素片データベースに含まれる複数の音声素片候補の各々の頻度情報でもよい。

【００２１】

好ましくは、ロード手段は、頻度情報に基づき、音声素片データベースに含まれる音声素片のうち、音声合成シミュレート手段による音声合成で選択された頻度の高いものを上位から所定個数選択し、選択された音声素片の波形データを第２の記憶装置にロードするための手段を含む。

【００２２】

さらに好ましくは、音声合成シミュレート手段は、テストデータから得られる合成音声の目標と、音声素片データベースに含まれる複数の音声素片候補との間で、複数のサブコストのうち、アクセス速度コストを除くアクセスコストと、頻度記録手段により記録され

た各音声素片候補が選択された頻度に基づいて算出される選択頻度コストとを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、音声素片データベースから選択するための手段を含む。

【００２３】

素片接続型音声合成装置は、キャッシュメモリと、キャッシュメモリに対して設けられたキャッシュメモリ管理機構と、複数の音声素片候補に対応して設けられ、対応の音声素片候補が選択手段により選択されたか否かを記録するための第２のフラグを記憶するための手段と、第２のフラグの値を所定の第１の値に初期化するための手段と、複数の音声素片候補のいずれかが選択手段により選択されるたびに、選択された音声素片候補に対応する第２のフラグを第１の値と異なる所定の第２の値に更新するためのフラグ更新手段とをさらに含んでもよい。アクセス速度コストは、第１及び第２のフラグに基づいて算出され、選択手段は、合成音声の目標との間で、アクセス速度コストを含む複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、第２の記憶装置に記憶された音響特徴量に基づいて複数の音声素片候補から選択するための手段を含んでもよい。

【００２４】

本発明の第２の局面に係る素片接続型音声合成方法は、合成音声の目標と音声素片候補との間で、複数のサブコストを含むコストを算出し、当該コストに基づいて、複数の音声素片候補を含む音声素片データベースから音声素片を選択し接続することにより音声合成を行なう素片接続型音声合成方法であって、各音声素片候補は、各音声素片の音響特徴量データと、各音声素片の波形データとを含み、音声素片データベースを第１の記憶装置に記憶させるステップと、音声素片データベースに記憶された複数の音声素片候補の音響特徴量データと、音声素片データベースに記憶された複数の音声素片候補の中で、所定の基準で選択された音声素片候補の波形データとを、第１の記憶装置より高速アクセス可能な第２の記憶装置に記憶させるステップとを含み、複数のサブコストは、各音声素片候補の波形データが記憶されている記憶装置へのアクセス速度に関するアクセス速度コストを含み、複数の音声素片候補の各々の音響特徴量データは、当該音声素片候補の波形データが第１及び第２の記憶装置のいずれに記憶されているかを示す第１のフラグを含み、音声合成方法はさらに、合成音声の目標との間で、複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、第２の記憶装置に記憶された音響特徴量に基づいて複数の音声素片候補から選択する選択ステップと、選択ステップにおいて選択された音声素片候補の音声波形を、当該選択された音声波形に対応する第１のフラグに基づいて、第１の記憶装置又は第２の記憶装置のいずれかから読出して合成器指令に従って接続し、合成音声波形を出力する接続ステップとを含む。

【００２５】

好ましくは、第１の記憶装置に記憶された音声素片データベースは、当該音声素片データベースに含まれる複数の音声素片候補のうち、対応する波形データを第２の記憶装置に記憶すべき音声素片候補を選択する基準となる選択基準情報を有し、素片接続型音声合成方法はさらに、第１の記憶装置に記憶された音声素片データベースに含まれる複数の音声素片候補のうち、選択基準情報により選択された音声素片候補の波形データを第２の記憶装置にロードするロードステップを含む。

【００２６】

さらに好ましくは、素片接続型音声合成方法はテストデータに基づいて、音声素片データベースの選択基準情報を生成する選択基準情報生成ステップをさらに含む。

【００２７】

選択基準情報生成ステップは、テストデータに基づき、音声素片データベースに含まれる音声素片候補を使用して音声合成をシミュレートする音声合成シミュレートステップと、音声素片データベースに含まれる複数の音声素片候補の各々について、音声合成シミュレートステップによる音声合成の際に選択された頻度を記録する頻度記録ステップとを含んでもよい。選択基準情報は、頻度記録ステップにおいて記録された、音声素片データベ

ースに含まれる複数の音声素片候補の各々の頻度情報でもよい。

【００２８】

好ましくは、ロードステップは、頻度情報に基づき、音声素片データベースに含まれる音声素片のうち、音声合成シミュレートステップによる音声合成で選択された頻度の高いものを上位から所定個数選択し、選択された音声素片の波形データを第２の記憶装置にロードするステップを含む。

【００２９】

さらに好ましくは、音声合成シミュレートステップは、テストデータから得られる合成音声の目標と、音声素片データベースに含まれる複数の音声素片候補との間で、複数のサブコストのうち、アクセス速度コストを除くアクセスコストと、頻度記録ステップにおいて記録された各音声素片候補が選択された頻度に基づいて算出される選択頻度コストとを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、音声素片データベースから選択するステップを含む。

【００３０】

素片接続型音声合成方法は、キャッシュメモリと、キャッシュメモリに対して設けられたキャッシュメモリ管理機構とを有するコンピュータ上で実行される方法でもよい。素片接続型音声合成方法はさらに、複数の音声素片候補に対応して設けられ、対応の音声素片候補が選択ステップにおいて選択されたか否かを記録する第２のフラグを所定の記憶装置に記憶するステップと、第２のフラグの値を所定の第１の値に初期化するステップと、複数の音声素片候補のいずれかが選択ステップにおいて選択されるたびに、選択された音声素片候補に対応する第２のフラグを第１の値と異なる所定の第２の値に更新するステップとをさらに含んでもよい。アクセス速度コストは、第１及び第２のフラグに基づいて算出される。選択ステップは、合成音声の目標との間で、アクセス速度コストを含む複数のサブコストを含んで算出されるコストが所定の条件を充足する一つの音声素片候補を、第２の記憶装置に記憶された音響特徴量に基づいて複数の音声素片候補から選択するステップを含んでもよい。

【発明を実施するための最良の形態】

【００３１】

素片選択型音声合成において、実際に選ばれる音声素片には偏りが生じる。従って、選ばれやすい音声素片は音声素片データベースより高速にアクセス可能な記憶装置（例えばメモリ）に格納しておくことで、音声合成の速度を全体として上げることができる。さらに、素片選択の際に、各音声素片の波形データが記憶されている記憶装置のアクセス速度をコストに加える。記憶装置のアクセス速度に対応するコストを、本明細書では「アクセス速度コスト」と呼ぶ。アクセス速度コストは、記憶装置のアクセス速度が高いほど小さく（０に近く）、低いほど大きくなるように、予め算出式を設計する。高速な記憶装置に波形データが記憶されている音声素片ほど、実際に音声合成で選択される可能性が高くなり、音声合成の速度を全体として高くすることができる。

【００３２】

図１に、本発明の一実施の形態に係る音声合成システム２０のブロック図を示す。図１を参照して、この音声合成システム２０は、従来と同様の音声素片ＤＢ３０と、多数の音声合成のためのテキストからなるテストデータ４２を使用して、音声素片ＤＢ３０を用いた音声合成をシミュレートし、音声素片ＤＢ３０に含まれる音声素片ごとに音声合成で使用された頻度を算出して、頻度情報４４を有する頻度情報付き音声素片ＤＢ３４を生成するための頻度情報生成装置３２とを含む。

【００３３】

音声合成システム２０はさらに、目標となるテキストを分析した結果得られる合成器指令３６を入力として受け、頻度情報付き音声素片ＤＢ３４に含まれる音声素片から適切な音声素片を選択し接続して合成音声波形４０を出力するための音声合成装置３８を含む。音声合成装置３８は、素片選択のための、頻度情報付き音声素片ＤＢ３４内の各音声素片の音響情報と、頻度情報付き音声素片ＤＢ３４内の音声素片のうち、頻度情報４４により

表される出現頻度が高いものを予め記憶しておくためのメモリ４８を含む。

【００３４】

音声合成システム２０はさらに、音声合成装置３８の起動時に、頻度情報付き音声素片ＤＢ３４内の各音声素片の音響情報と、頻度情報４４を参照することにより、頻度情報付き音声素片ＤＢ３４の中から選択した出現頻度の上位の所定個数の音声素片の波形データとをメモリ４８にロードするための音声素片データロード装置４６とを含む。

【００３５】

図２を参照して、頻度情報生成装置３２は、音声素片ＤＢ３０を使用し、テストデータ４２に含まれる各入力に対して実際に音声合成と同様の処理をして、各音声素片の使用頻度を算出する機能を持つ。頻度情報生成装置３２は、テストデータ４２の各テキスト文を受け、合成目標となる合成器指令６２を作成するための合成器指令作成部６０と、この合成器指令６２と音声素片ＤＢ３０中の各音声素片の音響特徴量とのターゲットコストを算出するためのターゲットコスト算出部６８と、合成器指令６２と音声素片ＤＢ３０中の各音声素片の音響特徴量との接続コストを算出するための接続コスト算出部７０と、音声素片ＤＢ３０中の各音声素片が音声合成で選択された頻度を反映する選択頻度コストを算出するための選択頻度コスト算出部７２とを含む。

【００３６】

選択頻度コストとは、頻度情報生成装置３２において選択される音声素片に対し、意図的に偏りを生じさせるために導入されたコストである。選択頻度コストは、音声素片が選択された頻度が高くなるほど小さく、低くなるほど大きくなるような算出式で算出される。本実施の形態では、ｉ番目の音声素片の選択頻度コスト$Ｃｓ_{i}$を次の式により算出する。

【００３７】

【数１】

$$Cs_i = a\left(1-\left(\frac{n_i}{N_0+N}\right)^{\frac{1}{r}}\right) \qquad (1)$$

ただし$ｎ_{i}$はｉ番目の音声素片の選択回数、Ｎは全ての素片の出現回数の和、ａ、$Ｎ_{0}$、及びｒは適当な定数である（ただしｒ＞０、ａ＞０、及びＮ≧０）。この式によれば、ｉ番目の音声素片の選択頻度コスト$Ｃｓ_{i}$の値は、選択頻度$ｎ_{i}$が０であればａであり、選択頻度$ｎ_{i}$が多くなるにつれて０に近づいて行く。選択頻度コスト算出部７２はこのため、頻度情報４４に記憶されている各音声素片の頻度情報を使用する。

【００３８】

頻度情報生成装置３２はさらに、ターゲットコスト算出部６８により算出されたターゲットコストと、接続コスト算出部７０により算出された接続コストと、選択頻度コスト算出部７２により算出された選択頻度コストとに基づいて総コストを算出し、総コストの最も小さな音声素片を選択することにより、実際の音声合成時の素片選択をシミュレートするための素片選択部６４と、素片選択部６４により選択された音声素片に関し、頻度情報４４を更新するための頻度情報更新部６６とを含む。

【００３９】

図３を参照して、図１に示す音声素片データロード装置４６がメモリ４８にロードするデータについて説明する。図３に示すように、メモリ４８は、頻度情報付き音声素片ＤＢ３４の全ての音声素片の音響特徴量１３０,…を格納するための音響特徴量格納領域１２０と、頻度情報付き音声素片ＤＢ３４内の音声素片のうち、頻度情報４４に記録された出現頻度が所定の値以上のものの波形データを記憶するための波形データ格納領域１２２とを含む。これらはいずれも音声素片データロード装置４６により、音声合成装置３８の起動時にメモリ４８にロードされる。

【００４０】

音響特徴量１３０の各々は、前述したとおり、音素ラベル、基本周波数（Ｆ０），ＭＦＣＣ、パワー（図示せず）、持続時間（図示せず）を含むが、これらに加えて、音声素片

の波形データが頻度情報付き音声素片ＤＢ３４に格納されているか、メモリ４８の波形データ格納領域１２２に格納されているかをあらわす第１のフラグ（F_1）１４０と、音声素片の波形データが最近読出されたか否かを示す第２のフラグ（F_2）１４２とを含む。フラグ１４０は、音声素片データロード装置４６が音響特徴量格納領域１２０に音響特徴量をロードした後、高頻度の音声素片の波形データを波形データ格納領域１２２にロードする際に、音声素片データロード装置４６によって設定される。

【００４１】

本実施の形態では、第１のフラグ１４０が０の場合には波形データは頻度情報付き音声素片ＤＢ３４に格納されていることを表し、フラグが１の場合には波形データがメモリ４８の波形データ格納領域１２２に格納されていることを示す。従ってこの第１のフラグ１４０の値を見ることで波形データをどこから読出せばよいかが判定できる。

【００４２】

フラグ１４２は最初に０に初期化され、実際に音声合成を行ないながら、波形データが読出されたときに「１」に更新される。これは、装置をコンピュータで実現する場合、ハードディスク又はメモリから読出されたデータはメモリよりもさらに高速アクセス可能なキャッシュに格納されることがあることを考慮したものである。この第２のフラグ１４２の値はアクセス速度コストに反映される。

【００４３】

図４を参照して、音声合成装置３８は、前述のメモリ４８に加え、それぞれ合成目標を定める合成器指令３６を受け、メモリ４８に記憶されている音声素片であって、かつ合成器指令３６により指定された音素ラベルを持つ音声素片の音響特徴量と合成器指令３６との間のターゲットコストを算出するためのターゲットコスト算出部８２と、同じくメモリ４８内の音声素片と合成器指令３６との間の接続コストを算出するための接続コスト算出部８４と、メモリ４８に記憶されている音声素片に対応する第１及び第２のフラグ１４０、１４２に基づいてアクセス速度コストを算出するためのアクセス速度コスト算出部８６とを含む。

【００４４】

音声合成装置３８はさらに、ターゲットコスト算出部８２により算出されたターゲットコスト、接続コスト算出部８４により算出された接続コスト、及びアクセス速度コスト算出部８６により算出されたアクセス速度コストに基づいて総コストを算出し、総コストの最小の音声素片を選択するための素片選択部８０と、素片選択部８０により選択された音声素片の波形データを接続して合成音声波形４０を出力するための接続部８８とを含む。

【００４５】

接続部８８は、素片選択部８０により指定された音声素片を頻度情報付き音声素片ＤＢ３４又はメモリ４８から読出すため、次のような信号を出力する。すなわち、接続部８８は、素片選択部８０により指定された音声素片の第１のフラグ１４０に対応するレベルをとり、波形データをメモリ４８と頻度情報付き音声素片ＤＢ３４とのいずれから読出すかを指定するための選択信号１００と、波形データを読出すアドレスを指定するアドレス信号１０２とを出力する機能を持つ。選択信号１００は、指定された音声素片のフラグが１のときにはＨレベルをとり、それ以外のときにはＬレベルをとる。

【００４６】

音声合成装置３８は、接続部８８による波形データの読出を行なうための機能ブロックとして、第１及び第２の入力を持ち、選択信号１００のレベルに応じて第１及び第２の入力の信号のいずれかを選択して出力するための選択回路９０と、選択信号１００のレベルを反転して反転選択信号１０４を出力するための反転回路９２と、Ｈレベルの反転選択信号１０４を受けると、アドレス信号１０２により指定されるアドレスの波形データを頻度情報付き音声素片ＤＢ３４から読出して選択回路９０の第１の入力に与えるためのアクセス部９４とを含む。一方、メモリ４８は、Ｈレベルの選択信号１００を受けると、アドレス信号１０２により指定されるアドレスの波形データを選択回路９０の第２の入力に与える。選択回路９０は、選択信号１００がＨレベルのときは第２の入力の信号を、Ｌレベル

のときには第１の入力の信号を選択して出力する。

【００４７】

音声合成装置３８はさらに、接続部８８により読出が指示された波形データについて、メモリ４８のうち、対応する音声素片の第２のフラグ１４２（F_2）を「１」に更新するためのフラグ更新部９６を含む。頻度情報付き音声素片ＤＢ３４又はメモリ４８から読出されたデータは、いずれの場合も、コンピュータのキャッシュメモリに格納されることが通常である。キャッシュメモリはメモリ４８と比較してもさらに高速にアクセス可能である。一度でも読出された波形データはキャッシュメモリに格納されている可能性が高いので、このように第２のフラグを更新し,次のアクセス速度算出部でのコスト計算に反映させ、より選択されやすくする。

【００４８】

なお、キャッシュメモリの容量には限りがあるため、何らかのアルゴリズムによってキャッシュに格納されているデータを選択して削除し、そこに新しいデータを格納する。本来であればキャッシュメモリにどの波形データが格納されているかを把握できればよいが、キャッシュメモリはコンピュータハードウェアにより管理されており、ソフトウェアでキャッシュの内容について知ることはできない。従って本実施の形態では、キャッシュメモリに実際にどのようなデータが格納されているかとは別に、一度でも読出されたことのあるデータについてはキャッシュメモリに格納されているものと想定した設計としている。もちろん、キャッシュメモリに記憶されている波形データがどの音声素片に対応するものであるかを容易に知ることができれば、それをアクセス速度コストの計算に反映させることが好ましい。

【００４９】

本実施の形態で使用されるサブコストは、前述したアクセス速度コスト以外に、基本周波数（Ｆ０）誤差、継続長誤差、ＭＦＣＣ誤差、Ｆ０不連続誤差、ＭＦＣＣ不連続誤差、音素環境誤差にそれぞれ対応する６種類のサブコストを含む。これらのうち、前３者はターゲットコストに属し、後３者は接続コストに属する。

【００５０】

本実施の形態に係る素片選択部６４によるコスト計算では、コストC_0は以下のようにしてサブコストから計算される。

【００５１】

【数２】

$$C_o=\left(\sum_{i1=1}^{N_1}(w_{i1}C_{i1})^{p_1}\right)^{\frac{1}{p_1}}+\left(\sum_{i2=1}^{N_2}(w_{i2}C_{i2})^{p_2}\right)^{\frac{1}{p_2}}+w_s(2-F_1-F_2)\qquad(2)$$

ただし、C_{i1}（ｉ1＝１～３）はターゲットサブコスト、C_{i2}（ｉ2＝１～３）は接続コスト、w_{i1}（ｉ1＝１～３）はターゲットサブコスト間に定義された重み、w_{i2}（ｉ2＝１～３）は接続サブコスト間に定義された重み、p_1及びp_2はそれぞれ、ターゲットコストと接続コスト間に定義された重み、w_sはアクセス速度コストの重み、F_1及びF_2はそれぞれ第１及び第２のフラグの値である。また重みwsの大きさは、他のサブコストの値の範囲を考慮して適当に決定し、品質に影響を及ぼす極端な偏りが生じないようにする。

【００５２】

上記した第１の実施の形態に係る音声合成システム２０は以下のように動作する。大きく分けてこの音声合成システム２０の動作には二つの局面がある。第１の局面は頻度情報付き音声素片ＤＢ３４の作成であり、第２の局面は第１の局面で作成された頻度情報付き音声素片ＤＢ３４を用いた音声合成である。以下、順に説明する。

【００５３】

まず第１の局面では頻度情報生成装置３２が以下のようにして頻度情報付き音声素片ＤＢ３４を作成する。図２を参照して、まず音声素片ＤＢ３０を用意する。音声素片ＤＢ３０には頻度情報は付されていない。また、頻度情報４４を記憶すべき領域をメモリ上に確

保しておく。さらに、頻度情報４４中の、各音声素片の選択頻度回数を全て０に初期化する。

【００５４】

テストデータ４２の第１のテキストを頻度情報生成装置３２に与えると、合成器指令作成部６０がそのテキストに基づいて合成目標の音素ごとに合成器指令６２を作成し、素片選択部６４に与える。素片選択部６４は、この合成器指令６２により指定された音響特徴量をターゲットコスト算出部６８及び接続コスト算出部７０に与える。ターゲットコスト算出部６８及び接続コスト算出部７０は、与えられた音響特徴量に基づき、音声素片ＤＢ３０に含まれる各音声素片との間でターゲットコスト及び接続コストを算出し素片選択部６４に与える。選択頻度コスト算出部７２は、式（１）に従って各音声素片候補の選択頻度コストを算出し素片選択部６４に与える。第１回目の処理では選択頻度はいずれも０であるから、選択頻度コストはいすれも式（１）より「ａ」となる。

【００５５】

素片選択部６４は、ターゲットコスト算出部６８から与えられたターゲットコスト、接続コスト算出部７０から与えられた接続コスト、及び選択頻度コスト算出部７２から与えられた選択頻度コストから式（２）によって総コストを算出する。素片選択部６４はこの総コストが最小の音声素片を選択し、選択された音声素片を示す情報を頻度情報更新部６６に与える。

【００５６】

頻度情報更新部６６は、頻度情報４４の中の頻度情報のうち、素片選択部６４により選択された音素の頻度に１を加算する。以上で第１番目の音素に対する処理を終了する。

【００５７】

同様の処理を、最初のテキストの各音素に対して繰り返す。この繰り返しにより、頻度情報４４は徐々に更新されていく。選択頻度コスト算出部７２による選択頻度コストの算出においては、頻度情報４４の内容が反映される。すなわち、選択された回数が多くなるほど選択頻度コストは小さくなる。従って、選択されたことのある素片候補についてはその後の素片選択で選択される可能性が高くなる。その結果、こうした処理を繰り返すと、互いによく似た音響特徴慮をもつ音声素片同士であって、選択されたことのある素片選択候補はさらに選択されやすく、選択されたことのない候補はさらに選択されにくくなる。

【００５８】

テストデータ４２の全てのテキストについて上記した処理を繰り返すことにより、頻度情報付き音声素片ＤＢ３４が完成する。頻度情報付き音声素片ＤＢ３４が完成すると、音声合成装置３８による音声合成が可能となる。

【００５９】

音声素片の出現頻度を意図的に偏らせることにより、特徴空間において素片の密度が高い領域において、一部の素片のみがよく選択されるようになり、それ以外の素片の出現頻度は下がる。その分、素片の密度が低い部分でそれ以外の素片の頻度の順位が相対的に上がる。これにより、実際の合成時に高速アクセス可能な記憶装置に波形データが格納される音声素片の分布が広がり、高速な記憶装置に格納された波形データに対応する音声素片がより頻繁に選択されるようになる。

【００６０】

実際の音声合成は以下のようにして行なわれる。図１を参照して、最初に音声素片データロード装置４６により、頻度情報付き音声素片ＤＢ３４中の音声素片の音響特徴量がメモリ４８に格納される。音響特徴量に付随する第１及び第２のフラグの値は０で初期化される。次に、頻度情報４４を基準とし、上位の所定個数の音声素片の波形データがメモリ４８に格納される。波形データをメモリ４８にロードした音声素片については、メモリ４８に格納された音響特徴量に付随する第１のフラグ１４０（図３参照）の値を「１」に設定する。

【００６１】

音声素片データロード装置４６によるメモリ４８へのデータのロードが終わると、実際

の音声合成が開始される。図４を参照して、合成器指令３６が与えられると、素片選択部８０はターゲットコスト算出部８２、接続コスト算出部８４及びアクセス速度コスト算出部８６に合成器指令６２により指定された音響特徴量を与える。ターゲットコスト算出部８２及び接続コスト算出部８４は、与えられた音響特徴量を用い、メモリ４８に格納されている音声素片候補のうち、指定された音素ラベルの音声素片候補の音響特徴量との間でターゲットコスト及び接続コストを算出し、素片選択部８０に与える。アクセス速度コスト算出部８６は、各音声素片候補の第１のフラグ１４０及び第２のフラグ１４２の値に基づきアクセス速度コストを算出し素片選択部８０に与える。

【００６２】

素片選択部８０は、ターゲットコスト算出部８２、接続コスト算出部８４、及びアクセス速度コスト算出部８６から与えられたターゲットコスト、接続コスト、及びアクセスコストに基づき、式（２）に従って総コストを算出する。素片選択部８０はさらに、そのようにして算出された総コストが最小の音声素片候補を選択し、その音声素片候補を示す情報と、音響特徴量とを接続部８８に与える。接続部８８は、与えられた音響特徴量の中の、波形データアドレスをアドレス信号１０２に、第１のフラグ１４０を選択信号１００に、それぞれ出力する。

【００６３】

例えば選択された音声素片の波形データがメモリ４８に格納されている場合、その音声素片の第１のフラグの値は１であり、選択信号１００はＨレベルとなる。アドレス信号１０２はメモリ４８の、選択された音声素片の波形データのアドレスとなる。アドレス信号１０２はメモリ４８に与えられる。Ｈレベルの選択信号１００がメモリ４８に与えられるので、メモリ４８はアドレス信号１０２により指定されるアドレスの波形データを読出し、選択回路９０の第２の入力に与える。反転選択信号１０４はＬレベルなのでアクセス部９４は何もしない。

【００６４】

選択信号１００がＨレベルなので、選択回路９０は第２の入力を選択する。すなわち、選択回路９０はメモリ４８からの出力を接続部８８に与える。接続部８８はこの波形データを用いて波形接続を行なう。

【００６５】

また、選択された音声素片の波形データが頻度情報付き音声素片ＤＢ３４に格納されている場合、その音声素片の第１のフラグの値は０であり、選択信号１００はＬレベルとなる。アドレス信号１０２は頻度情報付き音声素片ＤＢ３４の、選択された音声素片の波形データのアドレスとなる。アドレス信号１０２はアクセス部９４に与えられる。Ｈレベルの反転選択信号１０４がアクセス部９４に与えられるので、アクセス部９４はアドレス信号１０２により指定されるアドレスの波形データを頻度情報付き音声素片ＤＢ３４から読出し、選択回路９０の第１の入力に与える。選択信号１００はＬレベルなのでメモリ４８からは何も出力されない。

【００６６】

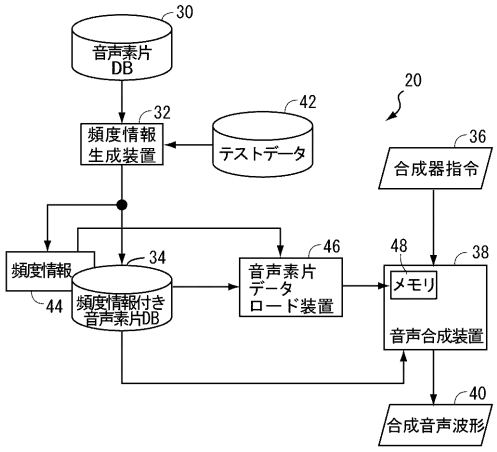
選択信号１００がＬレベルなので選択回路９０は第１の入力の信号を選択して接続部８８に与える。すなわちこの場合、頻度情報付き音声素片ＤＢ３４から読出された波形データが接続部８８に与えられ、波形接続に用いられる。

【００６７】

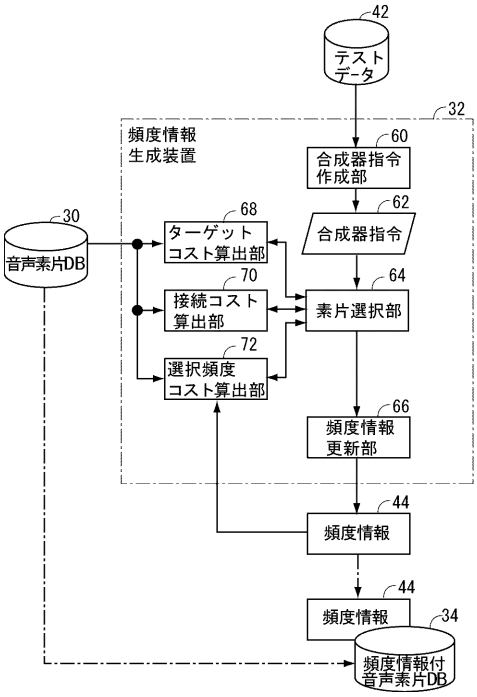
メモリ４８の波形データにせよ、頻度情報付き音声素片ＤＢ３４の波形データにせよ、最近接続部８８により読出されたものは図示しないキャッシュメモリに格納される可能性が高い。従ってフラグ更新部９６は、メモリ４８中の、接続部８８により波形データが読出された音声素片の音響特徴量に付随する第２のフラグ１４２の値を「１」に更新する。この結果、同じ音声素片が次に選択された場合、アクセス速度コストはより小さくなり、同じ音声素片が選択される可能性が高くなる。この音声素片に対応する波形データはキャッシュに格納されていて高速アクセス可能である可能性が高い。従って全体として波形データの読出が高速化される可能性が高い。

【００６８】
以上のように本実施の形態に係る音声合成システム２０では、音声素片ＤＢ３０を使用してテストデータ４２による音声合成実験を行なって、各音声頻度が選択された頻度を調べる。音声合成実験の途中では、音声素片が選択された回数に応じ、選択された素片はさらに選択されやすく、そうでない素片はさらに選択されにくくなるように、選択頻度コストというサブコストを導入し、音声素片の選択に人為的な偏りが生じるようにする。
【００６９】
音声合成時には、この頻度が上位の所定個数の音声素片候補の波形データをメモリに記憶しておく。さらに、サブコストとして、波形データが記憶されている記憶装置のアクセス速度を反映したアクセス速度コストを定義し、高速アクセス可能な記憶装置に波形データが記憶されている音声素片候補が選択されやすくする。その結果、メモリなどの高速アクセス可能な記憶媒体に記憶された音声素片が選択されやすくなり、全体として音声合成処理が高速化される。
【００７０】
さらに、キャッシュメモリなどが使用されることを考慮し、最近選択された音声波形についてはアクセス速度コストが少なく算出されるようにアクセス速度コストを設計する。これにより、コンピュータによるキャッシュ制御を利用した処理速度の向上を図ることができる。
【００７１】
また、上記した実施の形態の装置では、一旦頻度情報４４が作成された後は、頻度情報４４を更新しないことを前提としている。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。例えば音声合成時、頻度情報４４のをメモリ４８に転記し、素片候補が選択されるたびに頻度情報４４を更新し、全体の処理が終了したとき、または所定回数の素片候補の選択が行なわれるたびごとに、もとの頻度情報４４に書き戻すようにしてもよい。こうすることで、実験だけでなく実際のデータに基づく音声合成での選択頻度に基づいて、波形データの格納場所を決定することができる。
【００７２】
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でのすべての変更を含む。
【図面の簡単な説明】
【００７３】
【図１】本発明の一実施の形態にかかる音声合成システム２０のブロック図である。
【図２】図１に示す頻度情報生成装置３２のブロック図である。
【図３】図１に示すメモリ４８の一部の記憶領域の構成を模式的に示す図である。
【図４】図１に示す音声合成装置３８のブロック図である。
【符号の説明】
【００７４】
２０　音声合成システム、３０　音声素片ＤＢ、３２　頻度情報生成装置、３４　頻度情報付き音声素片ＤＢ、３６,６２　合成器指令、３８　音声合成装置、４０　合成音声波形、４２　テストデータ、４４　頻度情報、４６　音声素片データロード装置、４８　メモリ、６０　合成器指令作成部、６４,８０　素片選択部、６６　頻度情報更新部、６８,８２　ターゲットコスト算出部、７０,８４　接続コスト算出部、７２　選択頻度コスト算出部、８６　アクセス速度コスト算出部　８８　接続部、９０　選択回路、１００　選択信号、１０２　アドレス信号、１０４　反転選択信号

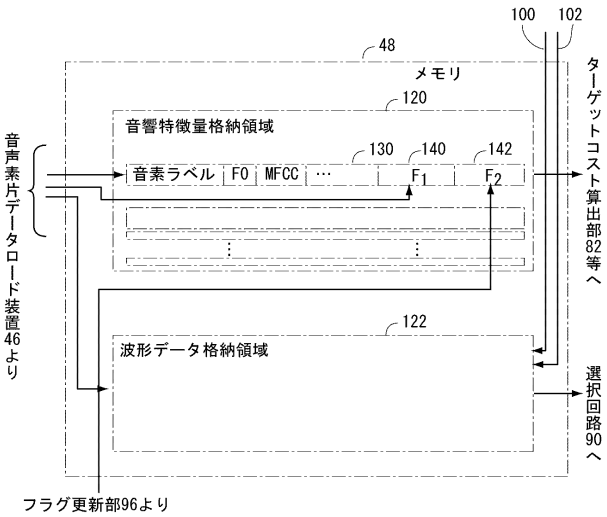
【図１】


30
音声素片DB
32
頻度情報生成装置
42
テストデータ
20
36
合成器指令
34
頻度情報付き音声素片DB
44
頻度情報
46
音声素片データロード装置
48
メモリ
38
音声合成装置
40
合成音声波形

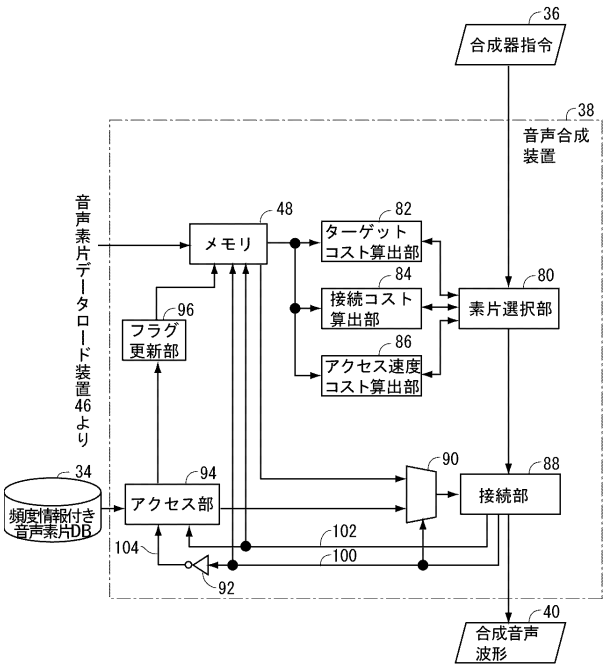
【図２】


42
テストデータ
32
頻度情報生成装置
60
合成器指令作成部
62
合成器指令
30
音声素片DB
68
ターゲットコスト算出部
70
接続コスト算出部
72
選択頻度コスト算出部
64
素片選択部
66
頻度情報更新部
44
頻度情報
34
頻度情報付き音声素片DB

【図３】


100
102
48
メモリ
120
音響特徴量格納領域
130
140
142
音素ラベル
F0
MFCC
…
F1
F2
音声素片データロード装置46より
ターゲットコスト算出部82等へ
122
波形データ格納領域
選択回路90へ
フラグ更新部96より

【図４】


36
合成器指令
38
音声合成装置
音声素片データロード装置46より
48
メモリ
82
ターゲットコスト算出部
84
接続コスト算出部
86
アクセス速度コスト算出部
80
素片選択部
96
フラグ更新部
90
88
接続部
34
頻度情報付き音声素片DB
94
アクセス部
102
100
104
92
40
合成音声波形

フロントページの続き

(56)参考文献　特開昭６３－０５２１９９（ＪＰ，Ａ）
特開２０００－０７５８７７（ＪＰ，Ａ）
特開平０７－２１０１９４（ＪＰ，Ａ）
特開２００１－２８２２７３（ＪＰ，Ａ）

(58)調査した分野(Int.Cl.，ＤＢ名)
Ｇ１０Ｌ　１３／０６