

(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2004-77608

(P2004-77608A)

(43) 公開日 平成16年3月11日(2004.3.11)

(51) Int. Cl. ⁷	F I	テーマコード (参考)
G 1 0 L 13/00	G 1 0 L 3/00	J 5 D 0 4 5
G 1 0 L 13/06	G 1 0 L 3/00	H
G 1 0 L 13/08	G 1 0 L 3/00	R
G 1 0 L 21/04	G 1 0 L 3/02	A
	G 1 0 L 5/04	F
審査請求 未請求 請求項の数 12 O L (全 24 頁)		

(21) 出願番号 特願2002-235039 (P2002-235039)
 (22) 出願日 平成14年8月12日 (2002. 8. 12)

(特許庁注：以下のものは登録商標)
 フロッピー

(71) 出願人 000004075
 ヤマハ株式会社
 静岡県浜松市中沢町10番1号
 (74) 代理人 100098084
 弁理士 川▲崎▼ 研二
 (72) 発明者 劔持 秀紀
 静岡県浜松市中沢町10番1号 ヤマハ株式会社内
 (72) 発明者 ジョルディ ボナダ
 パセイ デ シルコンバルーラシオ、8.
 08003バルセロナ スペイン
 Fターム(参考) 5D045 AA09 AA20 AB01

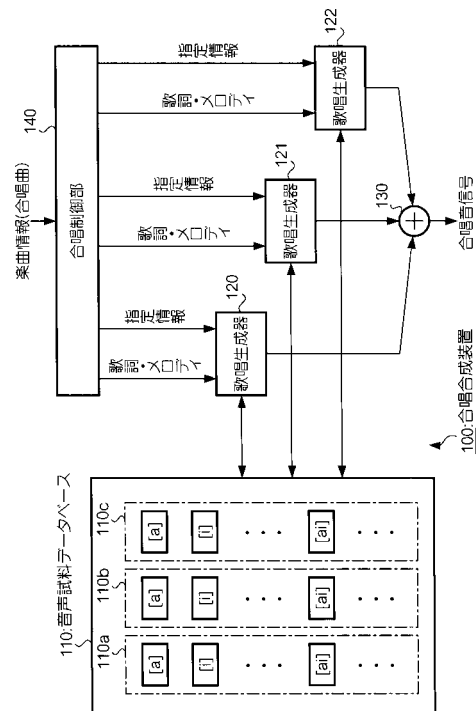
(54) 【発明の名称】 合唱合成装置、合唱合成方法およびプログラム

(57) 【要約】

【課題】より自然な印象を聴取者に与えることが可能な合唱音を合成する。

【解決手段】この合唱合成装置100は、各々異なる音声に基づいて作成した3つの音声試料データ群110a, 110b, 110cを記憶する音声試料データベース110と、3つの歌唱生成器120, 121, 122とを有している。3パートからなる楽曲の合唱音信号を合成する際には、合唱制御部140の制御に下、歌唱生成器120, 121, 122は各々パートに歌詞情報やメロディ情報に応じて歌唱音信号を生成し、各合唱音信号を合成する。この生成の際に、各歌唱生成器120, 121, 122は異なる音声試料データ群110a, 110b, 110cに含まれる音声素片試料データを用いる。

【選択図】 図1



【特許請求の範囲】**【請求項 1】**

楽曲データに基づいて合唱音信号を合成する合唱合成装置であって、同一種類の音声試料データについて、複数の異なる音声に基づいて各々作成された前記音声試料データを記憶するデータベースと、前記楽曲データにしたがって歌唱音信号を生成する手段であって、必要となる前記音声試料データを前記データベースから読み出して当該歌唱音信号の生成に用いる複数の歌唱生成手段と、前記複数の歌唱生成手段で生成された歌唱音信号から合唱音信号を合成する歌唱合成手段とを具備し、前記楽曲データが複数のパートからなり、前記複数の歌唱生成手段の各々が各前記パートに対応する歌唱音信号を生成する際に、少なくとも2つの前記歌唱生成手段の各々は、前記データベースから異なる音声に基づいて作成された前記音声試料データを読み出して前記歌唱音信号の生成に用いることを特徴とする合唱合成装置。

10

【請求項 2】

前記データベースは、音素、あるいは2つ以上の音素のつながりである音素連鎖である音声素片についての音声試料データであって、同一の音素あるいは音素連鎖について複数の異なる音声に基づいて作成した音声素片試料データを記憶しており、前記歌唱生成手段は、前記データベースから前記楽曲データに示される歌詞に対応した音声素片試料データを読み出して接続し、接続した音声素片試料データを前記楽曲データに示されるピッチに応じて調整して歌唱音信号を生成することを特徴とする請求項 1 に記載の合唱合成装置。

20

【請求項 3】

前記データベースは、複数の異なる音声に基づいて各々作成された音声試料データであって、音声のビブラート部分の特徴を示すビブラート音声試料データを記憶しており、前記歌唱生成手段は、ビブラート部分の歌唱音信号を生成する際に、前記データベースに記憶されている前記ビブラート音声試料データを読み出して用いることを特徴とする請求項 1 または 2 に記載の合唱合成装置。

30

【請求項 4】

楽曲データにしたがって合唱音信号を合成する合唱合成装置であって、音声に基づいて作成された所定の時間長を有する音声試料データを記憶するデータベースと、前記楽曲データにしたがって歌唱音信号を生成する手段であって、必要となる前記音声試料データを前記データベースから読み出して当該歌唱音信号の生成に用いる複数の歌唱生成手段と、前記複数の歌唱生成手段で生成された歌唱音信号から合唱音信号を合成する歌唱合成手段とを具備し、前記楽曲データが複数のパートからなり、前記複数の歌唱生成手段の各々が各前記パートに対応する歌唱音信号を生成する際に、少なくとも2つの前記歌唱生成手段の各々は、前記データベースから読み出した前記音声試料データの異なる時間に対応する部分から使用を開始して前記歌唱音信号を生成することを特徴とする合唱合成装置。

40

【請求項 5】

前記データベースは、音素、あるいは2つ以上の音素のつながりである音素連鎖である音声素片についての前記所定の時間長を有する音声素片試料データを記憶しており、前記歌唱生成手段は、前記データベースから前記楽曲データに示される歌詞に対応した音声素片試料データを読み出して接続し、接続した音声素片試料データを前記楽曲データに示されるピッチに応じて調整して歌唱音信号を生成することを特徴とする請求項 4 に記載の合唱合成装置。

50

【請求項 6】

前記データベースは、音声のビブラート部分の特徴を示す前記所定の時間長を有するビブラート音声試料データを記憶しており、
前記歌唱生成手段は、ビブラート部分の歌唱音信号を生成する際に、前記データベースに記憶されている前記ビブラート音声試料データを読み出して用いる
ことを特徴とする請求項 4 または 5 に記載の合唱合成装置。

【請求項 7】

前記楽曲データが複数のパートからなり、前記複数の歌唱生成手段の各々が各前記パートに対応する歌唱音信号を生成する際に、少なくとも 2 つの前記歌唱生成手段の各々によって生成された歌唱音信号の出力タイミングをずらすタイミング調整手段をさらに具備することを特徴とする請求項 1 ないし 6 のいずれかに記載の合唱合成装置。

10

【請求項 8】

前記複数の歌唱生成手段の各々に対応して設けられ、前記データベースから読み出された音声試料データに含まれる音色に関するデータを変更する音色変更手段を有しており、
前記楽曲データが複数のパートからなり、前記複数の歌唱生成手段の各々が各前記パートに対応する歌唱音信号を生成する際に、少なくとも 2 つの前記歌唱生成手段に対応する前記音色変更手段の各々は、互いに異なる方法によって前記音色に関するデータを変更することを特徴とする請求項 1 ないし 7 のいずれかに記載の合唱合成装置。

【請求項 9】

楽曲データに基づいて生成された複数の歌唱音信号から合唱音信号を合成する合唱合成方法であって、
複数のパートからなる前記楽曲データにしたがって前記複数のパートに対応する歌唱音信号を生成する際には、複数の異なる音声に基づいて各々作成された前記音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出し、
少なくとも 2 つの前記パートに対応する歌唱音信号の生成には、前記データベースから読み出された異なる音声に基づいて作成された前記音声試料データを用いる
ことを特徴とする合唱合成方法。

20

【請求項 10】

楽曲データに基づいて生成された複数の歌唱音信号から合唱音信号を合成する合唱合成方法であって、
複数のパートからなる前記楽曲データにしたがって前記複数のパートに対応する歌唱音信号を生成する際には、音声に基づいて作成された所定の時間長を有する音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出し、
少なくとも 2 つの前記パートに対応する歌唱音信号の生成には、前記データベースから読み出した前記音声試料データの異なる時間に対応する部分から使用を開始して前記歌唱音信号を生成する
ことを特徴とする合唱合成方法。

30

【請求項 11】

コンピュータを、
楽曲データにしたがって、複数の異なる音声に基づいて各々作成された音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出して歌唱音信号を生成する手段であって、前記楽曲データが複数のパートからなり、前記複数のパートに対応する歌唱音信号を生成する場合には、少なくとも 2 つの前記パートに対応する歌唱音信号の生成の際に、前記データベースから読み出された異なる音声に基づいて作成された前記音声試料データを用いて前記歌唱音信号を生成する歌唱音生成手段と、
前記生成された歌唱音信号から合唱音信号を合成する歌唱合成手段として機能させるためのプログラム。

40

【請求項 12】

コンピュータを、
楽曲データにしたがって、音声に基づいて作成された所定の時間長を有する音声試料デー

50

タを記憶するデータベースから必要となる前記音声試料データを読み出して歌唱音信号を生成する手段であって、前記楽曲データが複数のパートからなり、前記複数のパートに対応する歌唱音信号を生成する場合には、少なくとも2つの前記パートに対応する歌唱音信号を生成する際に、前記データベースから読み出した前記音声試料データの異なる時間に対応する部分から使用を開始して前記歌唱音信号を生成する歌唱音生成手段と、前記生成された歌唱音信号から合唱音信号を合成する歌唱音合成手段として機能させるためのプログラム。

【発明の詳細な説明】

【0001】

【発明の属する技術分野】

本発明は、合唱音信号を合成する合唱合成装置、合唱合成方法、および合唱音を合成するためのプログラムに関する。

【0002】

【従来技術】

従来より、歌詞情報やメロディ情報に基づいて、歌唱音信号を合成して歌声を発音する合唱合成装置が提案されている。このように歌唱音信号を合成する装置としては、規則音声合成技術を応用した装置等の種々の装置が提案されている。規則合成技術を応用した歌唱音合成装置では、予め発声者が発した音声から、音素や複数の音素を含む音素連鎖を単位とする音声試料データを作成してデータベースに記憶しておく。そして、歌詞情報にしたがって必要となる音素等の音声試料データを読み出して接続することにより歌唱音信号を合成している。

【0003】

ところで、上記のような歌唱音を合成する歌唱音合成装置では、文章読み上げ装置等の音声合成装置と異なり、斉唱や重唱といった合唱時の歌唱音を電子的に出力するといった利用形態も考えられる。したがって、合唱時の歌唱音（合唱音）を合成する機能を備えた合唱音合成装置の開発も行われている。

【0004】

このような合唱時の合唱音信号を合成する機能を備えた合唱音合成装置は、複数のパートの各々に基づいて、音声試料データを読み出して接続することにより合唱音信号を生成する。そして、各々のパートについて生成した歌唱音信号を重ね合わせて出力することにより、合唱音を電子的に出力することができるようになっている。

【0005】

【発明が解決しようとする課題】

しかし、従来技術の合唱音信号を合成する機能を備えた合唱音合成装置では、各パート毎に歌詞情報やメロディ情報にしたがって歌唱音信号を生成する際に、同一の音声試料データを用いているため、各パート毎に生成された歌唱音はメロディが異なっているものの、生成された各パート毎の音声波形の微細な特徴（ピッチのゆらぎ等）は基本的に同一となってしまう。したがって、これらを重ね合わせた合唱音は、聴取者にとって不自然な合唱音に聴こえてしまう。これは、各パート間の相関関係（微細な特徴が一致する）を聴取者が聴き取ってしまい、不自然な印象を与えているものと考えられる。

【0006】

また、斉唱時の合唱音信号を合成する場合には、上記のように各パート毎に単純に歌唱音信号を生成して重ね合わせる手法では、全く同じ歌唱音を重ねられて出力されてしまい、この結果聴取者に不自然な印象を与えてしまうことになる。そこで、従来技術の合唱音合成装置において、斉唱時の合唱音信号を合成する場合には、各パート（内容は同一）毎に生成した歌唱音の発音タイミングを若干ずらしたり、各パート毎に生成した歌唱音のピッチを若干ずらしたりすることにより、全く同一の歌唱音を重ねられて発音されてしまうことを防止していた。しかしながら、発音タイミングやピッチを若干ずらした場合にも、上記のように各パート毎に生成された音声波形の微細な特徴（ゆらぎ等）は基本的に同一となってしまう。したがって、これらを重ね合わせた合唱音は、上記と同様、聴取者にとって不

10

20

30

40

50

自然な合唱音に聴こえてしまう。

【0007】

また、特開平7-146695号公報には、合唱音信号を生成する装置が開示されており、この装置では、各パート毎に歌唱音信号を生成する際に、各パート毎に異なるピッチのゆらぎ成分を付与した歌唱音信号を生成している。このように各パート毎に異なるピッチのゆらぎ成分を付与した歌唱音信号を重ねて出力することにより、各パート間の相関関係を小さくすることができる。しかしながら、この公報に記載された装置において、各パート毎の歌唱音信号に付与されるピッチ成分は、人の音声を基にしたものではなく、人工的に作られたものであるため、各パート間の相関関係は小さくなるものの、合成された合唱音が不自然に聴こえてしまうことがある。

10

【0008】

本発明は、上記の事情を考慮してなされたものであり、より自然な印象を聴取者に与えることが可能な合唱音を合成することができる合唱合成装置、合唱合成方法およびプログラムを提供することを目的とする。

【0009】

【課題を解決するための手段】

上記課題を解決するため、本発明に係る合唱合成装置は、楽曲データに基づいて合唱音信号を合成する合唱合成装置であって、同一種類の音声試料データについて、複数の異なる音声に基づいて各々作成された前記音声試料データを記憶するデータベースと、前記楽曲データにしたがって歌唱音信号を生成する手段であって、必要となる前記音声試料データを前記データベースから読み出して当該歌唱音信号の生成に用いる複数の歌唱生成手段と、前記複数の歌唱生成手段で生成された歌唱音信号から合唱音信号を合成する歌唱合成手段とを具備し、前記楽曲データが複数のパートからなり、前記複数の歌唱生成手段の各々が各前記パートに対応する歌唱音信号を生成する際に、少なくとも2つの前記歌唱生成手段の各々は、前記データベースから異なる音声に基づいて作成された前記音声試料データを読み出して前記歌唱音信号の生成に用いることを特徴としている。

20

【0010】

この構成によれば、各歌唱生成手段が対応するパートの歌唱音信号を生成する際に、少なくとも2つの歌唱生成手段が異なる音声に基づいて作成した音声試料データを用いることになる。ここで、異なる音声に基づいて作成した音声試料データは、微細な特徴等が異なっているため、上記少なくとも2つの歌唱生成手段から出力される歌唱音信号は微細な特徴が異なったものとなる。したがって、各パートに応じた歌唱音として、固有の特徴を有する歌唱音が放音されるので、聴取者に対してより自然な印象を与えることができる。

30

【0011】

また、本発明の別の態様の合唱合成装置は、楽曲データにしたがって合唱音信号を合成する合唱合成装置であって、音声に基づいて作成された所定の時間長を有する音声試料データを記憶するデータベースと、前記楽曲データにしたがって歌唱音信号を生成する手段であって、必要となる前記音声試料データを前記データベースから読み出して当該歌唱音信号の生成に用いる複数の歌唱生成手段と、前記複数の歌唱生成手段で生成された歌唱音信号から合唱音信号を合成する歌唱合成手段とを具備し、前記楽曲データが複数のパートからなり、前記複数の歌唱生成手段の各々が各前記パートに対応する歌唱音信号を生成する際に、少なくとも2つの前記歌唱生成手段の各々は、前記データベースから読み出した前記音声試料データの異なる時間に対応する部分から使用を開始して前記歌唱音信号を生成することを特徴としている。

40

【0012】

この構成によれば、各歌唱生成手段が対応するパートの歌唱音信号を生成する際に、少なくとも2つの歌唱生成手段が音声試料データの異なる時間に対応する部分から使用を開始して生成を行うことになる。ここで、音声に基づいて作成されたある時間長を有する音声試料データは、その時間長の間微細な特徴（音波形のゆらぎ）が一定ではなく、時間によって微細な特徴等が異なっている。このため、上記少なくとも2つの歌唱生成手段から

50

出力される歌唱音信号は微細な特徴が異なったものとなる。したがって、各パートに応じた歌唱音として、固有の特徴を有する歌唱音が放音されるので、聴取者に対してより自然な印象を与えることができる。

【0013】

また、本発明に係る合唱合成方法は、楽曲データに基づいて生成された複数の歌唱音信号から合唱音信号を合成する合唱合成方法であって、複数のパートからなる前記楽曲データにしたがって前記複数のパートに対応する歌唱音信号を生成する際には、複数の異なる音声に基づいて各々作成された前記音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出し、少なくとも2つの前記パートに対応する歌唱音信号の生成には、前記データベースから読み出された異なる音声に基づいて作成された前記音声試料データを用いることを特徴としている。

10

【0014】

また、本発明の別の態様の合唱合成方法は、楽曲データに基づいて生成された複数の歌唱音信号から合唱音信号を合成する合唱合成方法であって、複数のパートからなる前記楽曲データにしたがって前記複数のパートに対応する歌唱音信号を生成する際には、音声に基づいて作成された所定の時間長を有する音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出し、少なくとも2つの前記パートに対応する歌唱音信号の生成には、前記データベースから読み出した前記音声試料データの異なる時間に対応する部分から使用を開始して前記歌唱音信号を生成することを特徴としている。

【0015】

また、本発明に係るプログラムは、コンピュータを、楽曲データにしたがって、複数の異なる音声に基づいて各々作成された音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出して歌唱音信号を生成する手段であって、前記楽曲データが複数のパートからなり、前記複数のパートに対応する歌唱音信号を生成する場合には、少なくとも2つの前記パートに対応する歌唱音信号の生成の際に、前記データベースから読み出された異なる音声に基づいて作成された前記音声試料データを用いて前記歌唱音信号を生成する歌唱音生成手段と、前記生成された歌唱音信号から合唱音信号を合成する合唱合成手段として機能させることを特徴としている。

20

【0016】

また、本発明の別の態様のプログラムは、コンピュータを、楽曲データにしたがって、音声に基づいて作成された所定の時間長を有する音声試料データを記憶するデータベースから必要となる前記音声試料データを読み出して歌唱音信号を生成する手段であって、前記楽曲データが複数のパートからなり、前記複数のパートに対応する歌唱音信号を生成する場合には、少なくとも2つの前記パートに対応する歌唱音信号の生成の際に、前記データベースから読み出した前記音声試料データの異なる時間に対応する部分から使用を開始して前記歌唱音信号を生成する歌唱音生成手段と、前記生成された歌唱音信号から合唱音信号を合成する合唱合成手段として機能させることを特徴としている。

30

【0017】

【発明の実施の形態】

以下、図面を参照して本発明の実施形態について説明する。

40

A. 第1実施形態

A-1. 第1実施形態の基本構成

まず、図1は本発明の第1実施形態に係る合唱合成装置の基本構成を示すブロック図である。同図に示すように、この合唱合成装置100は、音声試料データベース110と、複数(図示の例では3つ)の歌唱生成器120, 121, 122と、合唱制御部140と、歌唱生成器120, 121, 122の各々が出力する歌唱音信号を加算して合成し、出力する加算器130とを備えている。

【0018】

音声試料データベース110には、人が発声した自然の音声に基づいて作成された音声試料データが記憶されている。この音声試料データベース110には、単一の音素または複

50

数の音素で構成される音素連鎖を1つの単位とする音声試料データ(以下、音声素片試料データという)が記憶されている。

【0019】

多数の短時間長の音声試料データをデータベースに蓄積しておいて、歌詞等に応じてこれらの音声試料データを接続して音声合成処理技術では、合成単位として音素が用いられるのが基本である。このため、この合唱合成装置100における音声試料データベース110に、音素(30~50種類程度)単位のみ音声素片試料データを蓄積するようにしてもよいが、音素間の結合規則は複雑であるため、音素単位のみ音声試料データを蓄積した場合には、良好な品質を得ることが難しい。したがって、音声試料データベース110には、音素単位のみ音声素片試料データに加え、音素よりもやや大きい単位(音素連鎖)の音声素片試料データも蓄積しておくことが好ましい。音素よりも大きい単位としては、CV(子音 母音)、VC(母音 子音)、VCV(母音 子音 母音)、CVC(子音 母音 子音)といった単位がある。これらの単位の音声素片試料データを全て蓄積しておくことも考えられるが、合唱音を合成する合唱合成装置100においては、歌唱において使用頻度の高い母音など長く発音する伸ばし音を1単位とした音声素片試料データ、子音から母音(CV)および母音から子音(VC)を1単位とした音声素片試料データ、子音から子音を1単位とした音声素片試料データ、および母音から母音を1単位とした音声素片試料データを蓄積しておくようにすればよい。

10

【0020】

音声試料データベース110には、上述したような音素あるいは音素連鎖を1単位とした音声素片試料データが格納されているが、この音声試料データベース110では、同一種類の音素(例えば「a」)あるいは音素連鎖(例えば、「ai」)について3つの音声素片試料データを記憶している。すなわち、音声試料データベース110には、音素あるいは音素連鎖を1単位とした所定数の単位音声素片試料データからなる3つの音声試料データ群110a, 110b, 110cが記憶されているのである。

20

【0021】

音声試料データベース110に記憶されている3つの音声試料データ群110a, 110b, 110cは、各々異なる音声に基づいて作成されたデータである。ここで、異なる音声とは、発声者が異なることのみを意味するわけではなく、同じ発声者であっても別の機会に発した音声や異なる発声部分を用いたものであってもよい。このように音声試料データ群110a, 110b, 110cは、別の発声者または同じ発声者であっても別の機会に発した音声や別の発声部分に基づいて作成されているのである。このように各音声試料データ群110a, 110b, 110cに含まれる同一の音素(あるいは音素連鎖)についてのデータは、各々のデータを作成するために使用した基となる音声が異なっているため、微細な特徴(ピッチのゆらぎ等)が異なったものとなっている。

30

【0022】

音声試料データベース110には、上述したような3つの音声試料データ群110a, 110b, 110cが記憶されており、各歌唱生成器120, 121, 122は、歌唱音信号を生成する際にこの音声試料データベース110から音声素片試料データを読み出して用いることになる。

40

【0023】

歌唱生成器120, 121, 122の各々は、歌詞情報およびメロディ情報を有する楽曲情報にしたがって、音声試料データベース110から必要となる音声素片試料データを読み出し、読み出した音声素片試料データを用いて歌唱音信号を生成する。

【0024】

より具体的には、歌唱生成器120, 121, 122の各々は、歌詞情報にしたがって音素列を求め、その音素列を構成するために必要な音声素片試料データを決定し、音声試料データベース110から読み出す。そして、読み出した音声素片試料データを時系列に接続し、接続した音声素片試料データをメロディ情報にしたがったピッチに応じて適宜調整し、歌唱音信号を生成するのである。

50

【0025】

本実施形態に係る合唱合成装置100は、歌詞情報およびメロディ情報にしたがって歌唱音信号を生成することができる3つの歌唱生成器120, 121, 122を備えており、これにより3つのパートかならなる合唱曲の楽曲情報(歌詞情報およびメロディ情報)にしたがって、この合唱曲に対応した合唱音信号を合成することができるようになっている。

【0026】

合唱制御部140は、当該合唱合成装置100において、合唱曲の楽曲情報に基づいて合唱音に対応した合唱音信号を合成する際に、楽曲情報を各パート毎に分割して各歌唱生成器120, 121, 122に出力する。これにより、3つのパートからなる楽曲情報にしたがって合唱音信号を合成する場合には、各歌唱生成器120, 121, 122が合唱制御部140から供給される各々のパートの歌詞情報およびメロディ情報にしたがって歌唱音信号を生成し、各歌唱生成器120, 121, 122の各々が生成した歌唱音信号が加算器130に出力される。これにより、加算器130からは3つのパートからなる合唱曲の楽曲情報にしたがって、この合唱曲に対応した合唱音信号を合成することができるのである。

【0027】

また、この合唱合成装置100において、上記のように合唱音信号を合成する際には、合唱制御部140は、各歌唱生成器120, 121, 122の各々が、音声試料データベース110に記憶されている音声試料データ群110a, 110b, 110cのうち、どの音声試料データ群から音声素片試料データを読み出して用いるかを指定する指定情報を歌唱生成器120, 121, 122に出力する。ここで、合唱制御部140は、各歌唱生成器120, 121, 122が互いに異なる音声試料データ群110a, 110b, 110cに含まれる音声素片試料データを用いて歌唱音信号を生成するように、各歌唱生成器120, 121, 122に異なるデータ群を指定する指定情報を出力する。

【0028】

具体的に例示すると、歌唱生成器120に対しては音声試料データ群110aを指定する指定情報を出力し、歌唱生成器121に対しては音声試料データ群110bを指定する指定情報を出力し、歌唱生成器122に対しては音声試料データ群110cを指定する指定情報を出力するといった具合である。このような指定情報が合唱制御部140から供給されると、歌唱生成器120は音声試料データ群110aに含まれる音声素片試料データを読み出して歌唱音信号の生成に用い、歌唱生成器121は音声試料データ群110bに含まれる音声素片試料データを読み出して歌唱音信号の生成に用い、歌唱生成器122は音声試料データ群110cに含まれる音声素片試料データを用いて歌唱音信号を生成することになる。

【0029】

合唱合成装置100において、3つのパートからなる合唱曲の歌唱音信号を合成する場合に、上述したように各歌唱生成器120, 121, 122が互いに異なる音声試料データ群110a, 110b, 110cに含まれる音声素片試料データを用いることにより、より自然な印象を聴取者に与えることが可能な合唱音信号を合成することができる。すなわち、音声試料データベース110に記憶されている音声試料データ群110a, 110b, 110cは、各々異なる音声に基づいて作成されたものであり、同一種類の音素や音素連鎖についてのデータであっても、各音声試料データ群110a, 110b, 110cに含まれるデータに示される音声の微細な特徴(ピッチのゆらぎ等)は異なっている。このように微細な特徴が異なっている音声素片試料データが含まれる音声試料データ群110a, 110b, 110cのうち、各歌唱生成器120, 121, 122が異なる音声試料データ群に含まれる音声素片試料データを用いて歌唱音信号を生成することにより、各歌唱生成器120, 121, 122によって生成される歌唱音信号は、互いに微細な特徴が異なるものとなっている。したがって、これらを重ね合わせた合唱音は、各パート間の相関関係がほとんどない固有の特徴を有するものとなり、聴取者に不自然な印象を与えてし

10

20

30

40

50

まうことを低減することができる。

【0030】

また、音声試料データベース110に記憶されている音声試料データ群110a, 110b, 110cに含まれる音声素片試料データは、各々人が発声した音声に基づいて作成されたデータである。したがって、各音声試料データ群に含まれる音声素片試料データに示される音声の微細な特徴の相違は、予め用意されたピッチのゆらぎを付与するといった人工的に作り出されたものではない。したがって、合成された合唱音が不自然なものとなってしまうことを低減することができる。

【0031】

A-2. 合唱合成装置の具体的な構成

以上説明したのが本実施形態に係る合唱合成装置100の基本的な構成である。この合唱合成装置100においては、歌唱生成器120, 121, 122として、歌詞情報にしたがって音声試料データベース110から音声素片試料データを読み出して接続し、メロディ情報にしたがったピッチに応じて接続した音声素片試料データを調整して歌唱音信号を出力するといった歌唱生成器であれば、規則音声合成技術等を応用した歌唱生成器等の公知の種々の歌唱生成器を用いることができ、音声試料データベース110には採用する歌唱生成器に対応した音声素片試料データを記憶させておけばよい。以下においては、歌唱生成器120, 121, 122として、米国特許第5029509号や特許第2906970号において提案されているスペクトルモデリング合成(SMS: Spectral Modeling Synthesis)技術を利用した歌唱生成器を適用した場合を例

10

20

【0032】

まず、SMS技術を利用した歌唱生成器120, 121, 122を備えた合唱合成装置100における音声試料データベース110の作成手法について説明する。

【0033】

上述したように、この合唱合成装置100における音声試料データベース110には、発声者の発した音声に基づいて作成された音声素片試料データが記憶されている。SMS技術は、オリジナルの音を2つの成分、すなわち調和成分(deterministic component)と、非調和成分(stochastic component)で表すモデルを使用して楽音の分析および合成を行う技術であり、SMS技術を利用した音声合成においては、音素あるいは音素連鎖といった1単位の音声素片試料データとして、上記調和成分および非調和成分からなるデータが音声合成に用いられる。したがって、SMS技術を利用した合唱合成装置100においては、音声試料データベース110に、発声者の発した音声をSMS分析することにより得られた調和成分および非調和成分を示すデータが1つの音声素片試料データとして記憶される。以下、図2を参照しながら、音声試料データベース110の作成手法について説明する。

30

【0034】

同図に示すように、音声試料データベース110の作成のために発声者が発した音声は、SMS分析部200に入力され、SMS分析部200においてSMS分析される。ここで、音声試料データベース110には、異なる音声に基づいて作成した音声試料データ群110a, 110b, 110cを記憶しておく必要があるため、3つの異なる音声がSMS分析部200に入力されることになる。なお、図示では、3つの異なる音声が並列にSMS分析部200に入力されるように表されているが、各音声についてのSMS分析は同時に並列して行う必要はなく、個別に行うようにしてもよい。

40

【0035】

SMS分析部200は、入力される音声に対してSMS分析を行い、各フレーム毎のSMS分析データを出力する。より具体的には、以下の手法により各フレーム毎のSMS分析データを出力する。

【0036】

まず、入力される音声を一連のフレームに分ける。ここで、SMS分析に用いるフレーム

50

周期としては、一定の固定長であってもよいし、入力音声のピッチ等に応じてその周期を変更する可変長の周期であってもよい。

【0037】

次に、フレームに分けた音声に対して高速フーリエ変換（FFT：Fast Fourier Transform）等の周波数分析を行う。この周波数分析によって得られた周波数スペクトル（複素スペクトル）から振幅スペクトルと位相スペクトルを求め、振幅スペクトルのピークに対応する特定の周波数のスペクトルを線スペクトルとして抽出する。このとき、基本周波数およびその整数倍の周波数近傍の周波数を持つスペクトルを線スペクトルとする。このようにして抽出した線スペクトルが上述した調和成分に対応している。

10

【0038】

次に、上記のように入力音声から線スペクトルを抽出するとともに、抽出した線スペクトルをそのフレームの入力音声（FFT後の波形）から減算することにより、残差スペクトルを得る。あるいは、抽出した線スペクトルから合成した調和成分の時間波形データをそのフレームの入力音声波形データから減算して残差成分の時間波形データを取得した後、これに対してFFT等の周波数分析を行うことにより残差スペクトルを得るようにしてもよい。このようにして得られた残差スペクトルが上述した非調和成分に対応している。

【0039】

SMS分析部200は、上記のようにして取得した線スペクトル（調和成分）および残差スペクトル（非調和成分）からなる各フレーム毎のSMS分析データを区間切り出し部201に出力する。

20

【0040】

区間切り出し部201は、SMS分析部200から供給される各フレーム毎のSMS分析データを、音声試料データベース110に記憶すべき音声素片試料データの1単位（音素あるいは音素連鎖）の長さに対応するように切り出す。区間切り出し部201は、各素片の単位長さに対応するようにSMS分析データを切り出し、音声試料データベース110に記憶させる。

【0041】

ここで、音声試料データベース110に記憶される音声素片試料データは、音素あるいは音素連鎖毎に切り出されたSMSデータであり、調和成分については、その音素あるいは音素連鎖に含まれるフレーム全てのスペクトル包絡（線スペクトル（倍音系列）の強度（振幅）および位相のスペクトル）が記憶される。なお、このようなスペクトル包絡そのものを調和成分として記憶させるようにしてもよいが、該スペクトル包絡を何らかの関数で表現したものを記憶させるようにしてもよいし、調和成分を逆FFT等して得た時間波形として記憶させるようにしてもよい。本実施形態では、非調和成分についても調和成分と同様に、強度スペクトルと位相スペクトルとして記憶させることとするが、上記調和成分と同様、関数や時間波形として記憶させるようにしてもよい。

30

【0042】

このような音声に対するSMS分析および区間切り出しが3つの異なる入力音声の各々について行われ、この結果、音声試料データ群110a, 110b, 110cといった3つの異なる音声に基づいて作成された音声素片試料データ（音素あるいは音素連鎖毎のSMS分析データ）の群が音声試料データベース110に記憶される。

40

【0043】

以上が本実施形態に係る合唱合成装置100の音声試料データベース110の作成手法の詳細である。

【0044】

次に、上述したように異なる音声に基づいて作成された3つの音声試料データ群110a, 110b, 110cを記憶する音声試料データベース110を用いて歌唱音信号を生成する各歌唱生成器120, 121, 122について説明する。なお、歌唱生成器120, 121, 122は、各々同様の構成であるため、以下においては歌唱生成器120の構成

50

について図3を参照しながら説明し、他の歌唱生成器121, 122についての説明を割愛する。

【0045】

同図に示すように、この歌唱生成器120は、音声素片選択部301と、ピッチ決定部302と、継続時間長調整部303と、音声素片接続部304と、調和成分生成部305と、加算部306と、逆FFT(高速フーリエ変換)部307と、窓掛け部308と、オーバーラップ部309とを備えている。

【0046】

音声素片選択部301は、合唱制御部140(図1参照)から供給される歌詞情報および指定情報にしたがって、必要となる音声素片試料データを音声試料データベース110から読み出す。より具体的には、供給される歌詞情報を音声記号(音素あるいは音素連鎖)列に変換し、変換した音声記号列にしたがって音声試料データベース110から音声素片試料データを読み出す。例えば、「サイタ」(s a i t a)といった歌詞情報にしたがって歌唱音信号を生成する場合には、該歌詞情報が「#s」、「s」、「sa」、「a」、「ai」、「i」、「it」、「t」、「ta」、「a」、「a#」といった音声記号列に変換され、これらの各音声記号に対応する音声素片試料データが音声試料データベース110から読み出されることになる。

【0047】

音声素片選択部301は、上記のように歌詞情報にしたがって読み出すべき音声素片試料データを決定し、合唱制御部140から供給される指定情報に指定される音声試料データ群の中から決定した音声素片試料データを読み出す。例えば、指定情報が音声試料データ群110aを指定している場合には、音声試料データベース110の音声試料データ群110aに含まれる「#s」、「s」、「sa」、「a」、「ai」、「i」、「it」、「t」、「ta」、「a」、「a#」に対応した音声素片試料データを読み出す。

【0048】

ピッチ決定部302は、合唱制御部140(図1参照)から供給されるメロディ情報に応じて歌唱音のピッチを決定し、決定したピッチを示すピッチ情報を調和成分生成部305に出力する。

【0049】

継続時間長調整部303には、音声素片選択部301によって読み出された音声素片試料データ(調和成分および非調和成分)が供給される。ここで、音声素片選択部301は、読み出した音声素片試料データをそのまま継続時間長調整部303に供給するようにしてもよいが、メロディ情報に示されるピッチ等に応じて適当な補正処理を施してから継続時間長調整部303に供給するようにしてもよい。

【0050】

継続時間長調整部303は、メロディ情報等によって決定される音素あるいは音素連鎖毎の発音時間長に応じて音声素片選択部301から供給された各音声素片試料データの時間長を変更する処理を行う。より具体的には、ある音声素片試料データを、その時間長より短い時間として使用する場合には、該音声素片試料データからフレームを間引く処理を行う。一方、ある音声素片試料データを、その時間長よりも長い時間継続して使用する場合には、その音声素片試料データを使用する時間長の間繰り返して時間を長くするループ処理を行う。このループ処理において、ある音声素片試料データを繰り返す場合には、当該音声素片試料データの最初から最後(0~t)までのデータの後に、当該音声素片試料データjを最初(0)からデータを接続して繰り返すようにしてもよいし、最初から最後(0~t)までのデータの後に、当該音声素片試料データの時間的に最後(t)の部分から最初の部分に向かってデータを接続して繰り返すようにしてもよい。

【0051】

継続時間長調整部303は、上記のように各音声素片の発音時間長に応じて音声素片試料データ(調和成分および非調和成分)の継続時間長を調整した後、時間調整後の音声素片試料データを音声素片接続部304に出力する。

10

20

30

40

50

【0052】

音声素片接続部304は、継続時間長調整部303から供給された音声素片試料データの調和成分のデータを時系列に接続するとともに、非調和成分のデータを時系列に接続する。このような接続に際し、接続する2つの調和成分のスペクトル包絡の形状の差が大きい場合には、スムージング処理等を施すようにすればよい。音声素片接続部304は、接続した調和成分のデータを調和成分生成部305に出力するとともに、接続した非調和成分のデータを加算部306に出力する。

【0053】

調和成分生成部305には、音声素片接続部304から調和成分のデータ(スペクトル包絡情報)が供給されるとともに、ピッチ決定部302からメロディ情報にしたがったピッチ情報が供給される。調和成分生成部305は、音声素片接続部304からのスペクトル包絡情報に示されるスペクトル包絡形状を維持しつつ、ピッチ決定部302からのピッチ情報に対応する倍音成分を生成する。

【0054】

加算部306には、音声素片接続部304からの非調和成分のデータと、調和成分生成部305からの調和成分のデータが供給され、加算部306は両者を合成して逆FFT部307に出力する。逆FFT部307は、加算部306から供給される加算された周波数領域の信号に対し、逆FFTを施すことにより時間領域の波形信号に変換し、変換後の波形信号を窓掛け部308に出力する。窓掛け部308では、時間領域の波形信号に対してフレーム長に対応した窓関数が乗算され、オーバーラップ部309が乗算後の波形信号をオーバーラップさせながら歌唱音信号を生成する。このようにして歌唱生成器120では、合唱制御部140(図1参照)から供給された楽曲情報のあるパートの歌詞情報およびメロディ情報にしたがった歌唱音信号が生成され、生成された歌唱音信号が加算器130(図1参照)に出力される。

【0055】

以上が歌唱生成器120の詳細な構成であり、図1に示す他の歌唱生成器121, 122(歌唱生成器120と同様の構成)からも上記のように合唱制御部140から供給された楽曲情報のあるパートの歌詞情報およびメロディ情報にしたがって生成された歌唱音信号が出力される。ここで、上述したように各歌唱生成器120, 121, 122は、合唱制御部140から振り分けられたパートに対応する歌唱音信号を生成する際に、各々異なる音声試料データ群110a, 110b, 110cから音声素片試料データを読み出して生成に用いているので、各々が生成する歌唱音信号の微細な特徴(ピッチのゆらぎ等)は異なったものとなる。

【0056】

加算器130は、このように合唱曲の楽曲情報の各パートにしたがって歌唱生成器120, 121, 122が生成した歌唱音信号を合成して出力する。加算器130から出力された3つのパートの歌唱音信号が合成された合唱音信号は、図示せぬD/A(Digital to Analog)変換器によってアナログの音声波形信号に変換された後、アンプ等を介してスピーカから放音される。これにより、聴取者は、複数パートからなる合唱曲の楽曲情報にしたがった合唱音を聴くことができる。この合唱合成装置100から放音される合唱音は、各パートの歌唱音の微細な特徴(ピッチのゆらぎ等の相違に起因する声質等)が相違しており、聴取者により自然な印象を与えることが可能な合唱音を発音することができるのである。

【0057】

B. 第2実施形態

次に、本発明の第2実施形態に係る合唱合成装置について、図4を参照しながら説明する。同図に示すように、上記第1実施形態における合唱合成装置100の音声試料データベース110が3つの音声試料データ群110a, 110b, 110cを記憶していたのに対し、第2実施形態に係る合唱合成装置400における音声試料データベース110には、同一の音素または音素連鎖については1種類の音声素片試料データしか記憶されていな

10

20

30

40

50

い点で相違している。第2実施形態に係る合唱合成装置400は、このように1つの音素または音素連鎖について1つの音声素片試料データのみを記憶する音声試料データベース110を用いて、上記第1実施形態と同様により自然な印象を与えることが可能な合唱音信号を合成することができるようになっている。以下、合唱合成装置400の構成について、上記第1実施形態に係る合唱合成装置100との相違点を中心に説明する。

【0058】

合唱合成装置400における歌唱生成器120, 121, 122の各々は、上記第1実施形態と同様であり、歌詞情報およびメロディ情報を有する楽曲情報にしたがって音声試料データベース110から必要となる音声素片試料データを読み出し、読み出した音声素片試料データを用いて歌唱音信号を生成する。第2実施形態においては、音声試料データベース110には1つの音素あるいは音素連鎖については1つの音声素片試料データしか記憶されていないため、合唱曲の楽曲情報にしたがって歌唱音信号を生成する際には、各歌唱生成器120, 121, 122が同一の音声素片試料データを用いることもあり得る。上述したように複数のパートの歌唱音信号を同一の音声素片試料データを用いて生成した場合、微細な特徴(ピッチのゆらぎ等)が基本的に同一になるため、聴取者に不自然な印象を与えてしまう。

10

【0059】

そこで、この合唱合成装置400では、合唱制御部140が合唱曲の楽曲情報の歌詞情報およびメロディ情報を各パート毎に分割して各歌唱生成器120, 121, 122に出力するとともに、音声試料データベース110に記憶されている音声素片試料データをどの時間に対応する部分から使用を開始するかを指定する指定情報を各歌唱生成器120, 121, 122に出力するようになっている。

20

【0060】

上述した第1実施形態で説明したように、音声試料データベース110に記憶される音声素片試料データは、発声者の発した音声に基づいて作成されたものであり、所定の時間長(1フレーム~数フレーム等)の音声波形に基づいて作成されたデータである。すなわち、前記所定の時間内における時間と振幅との関係で表される音声波形に基づいて作成されたデータである。したがって、上記第1実施形態のように音声素片試料データが周波数領域のデータとして記憶されている場合にも、そのデータは時間領域の音声波形にFFT等を施して得られたものである。合唱制御部140は、このように時間に伴って変化する情報である音声素片試料データをどの時間に対応する部分から使用するかを指定する指定情報を各歌唱生成器120, 121, 122に供給するのである。

30

【0061】

ここで、合唱制御部140は、各歌唱生成器120, 121, 122が音声素片試料データを、互いに異なる時間に対応する部分から使用を開始して歌唱音信号を生成するように、各歌唱生成器120, 121, 122に異なる使用開始時間を指定する指定情報を出力する。

【0062】

各歌唱生成器120, 121, 122は、合唱制御部140から供給される各パートの歌詞情報に基づいて必要となる音声素片試料データを音声試料データベース110から読み出すと共に、読み出した音声素片試料データを、合唱制御部140から指定情報に指定される時間に対応する部分から使用を開始して歌唱音信号の生成を行う。

40

【0063】

以下、3つのパートの歌詞情報にしたがって読み出される音声素片試料データが母音の「a」であり、音声素片試料データ「a」がF0~F13といった13のフレーム(時間0~T)からなり、該音声素片試料データを13フレーム分の長さを使用して各歌唱生成器120, 121, 122が歌唱音信号を生成する場合について、図5および図6を参照しながら具体的に例示して説明する。

【0064】

図5に示す例では、歌唱生成器120に対しては最初のフレームF0から使用を開始する

50

ように指定する指定情報が供給されており、歌唱生成器 1 2 1 に対してはフレーム F 3 から使用を開始するように指定する指定情報が供給されており、歌唱生成器 1 2 2 に対してはフレーム F 6 から使用を開始するように指定する指定情報が供給されている。なお、図示では説明の便宜上、音声素片試料データが時間領域の音声波形として示されているが、音声試料データベース 1 1 0 に記憶しておくデータは、上記第 1 実施形態のように周波数領域で表現される調和成分（線スペクトル）および非調和成分（残差スペクトル）といった形態であってもよい。

【 0 0 6 5 】

このような指定情報が供給されている場合には、図 6 に示すように、歌唱生成器 1 2 0 は、フレーム F 0 , F 1 , F 2 F 1 3 といった順序、つまり音声素片試料データ「a」をそのまま使用して歌唱音信号の生成に用いる。また、歌唱生成器 1 2 1 は、フレーム F 3 , F 4 , F 5 F 1 3 , F 0 , F 1 , F 2 , F 3 といった順次で音声素片試料データ「a」を使用して歌唱音信号の生成を行う。さらに、歌唱生成器 1 2 2 は、フレーム F 6 , F 7 F 1 3 , F 0 , F 1 F 5 といった順序の音声素片試料データ「a」を使用して歌唱音信号の生成を行う。

10

【 0 0 6 6 】

このように合唱制御部 1 4 0 が互いに異なる時間に対応する部分から使用を開始して歌唱音信号を生成するように指定情報を出力することにより、同じ音素「a」を同じ時間長（0 ~ T まで）だけ用いて歌唱音信号を生成する際にも、各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 が実際に用いるデータは異なるものとなっている。すなわち、各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 が実際に用いる音声素片試料データに示される微細な特徴（ピッチのゆらぎ等）は異なったものとなり、1つの音素あるいは音素連鎖について1種類の音声素片試料データを用いて、各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 が微細な特徴の異なる歌唱音信号を生成することができるのである。

20

【 0 0 6 7 】

ところで、音素「a」のように単一の音素についての音声素片試料データを用いる場合には、上記のように単純にデータ中の使用開始時間をずらすといった手法により、各パートについて生成される歌唱音の微細な特徴を変えてより自然な合唱歌唱音を合成することができるが、複数の音素が連なる音素連鎖についての音声素片試料データの場合には、単純にデータ中の使用開始時間をずらすだけでは不都合が生じることもある。例えば、「a i」といった音素連鎖についての音声試料データの場合、時間領域における前半部分は「a」の音素をより強く反映したデータであり、後半部分は「i」の音素をより強く反映したデータである。したがって、音素連鎖「a i」の歌唱音信号を生成するために、音素「i」の影響の強い後半部分から使用を開始した場合には、音素連鎖「i a」に類似した傾向を持つデータを用いることになってしまう虞があり、この場合、生成すべき音素連鎖「a i」についての信号が正確に生成できなくなってしまう。

30

【 0 0 6 8 】

そこで、本実施形態では、複数の音素連鎖に対応する音声素片試料データを用いる場合には、合唱制御部 1 4 0 は、図 7 に示すような指定情報を各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 に出力するようにしている。同図に示す例では、歌唱生成器 1 2 0 に対しては最初のフレーム F 0 から使用を開始するように指定する指定情報が供給されており、歌唱生成器 1 2 1 に対してはフレーム F 2 から使用を開始するように指定する指定情報が供給されており、歌唱生成器 1 2 2 に対してはフレーム F 4 から使用を開始するように指定する指定情報が供給されている。すなわち、上記単一の音素についての指定情報と比較すると、各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 に対して指定する使用開始時間が前半部分（「a」の影響の強い）に集中している。このように各パートの使用開始時間をデータの前半部分に集中させることで、上記のように実際に使用するデータが音素連鎖「i a」に類似してしまうことを抑制している。

40

【 0 0 6 9 】

また、上記のように指定情報が供給されている場合に、歌唱生成器 1 2 1 がフレーム F 2

50

、F 3、F 4 …… F 1 3、F 0、F 1、F 2といった順序、すなわち一方向に順番に音声素片試料データを使用して歌唱音信号の生成を行うと、音素「a」の影響の強いフレームF 0～F 2が本来「i」の影響を強くすべき後半部分のデータとして用いられてしまうことになる。そこで、本実施形態では、複数の音素からなる音素連鎖についての音声素片試料データを用いる場合には、最後のフレーム（F 1 3）の後にフレームF 1に戻るのではなく、フレームF 1 2、F 1 1 ……といったように逆方向に戻る順序でフレームを用いるようにしている。したがって、図7に示すように使用開始フレームが指定されている場合には、図8に示すように、歌唱生成器1 2 0は、フレームF 0、F 1、F 2 …… F 1 3といった順序、つまり音声試料データベース1 1 0に記憶されている音声素片試料データをそのまま使用して歌唱音信号の生成に用いる。また、歌唱生成器1 2 1は、フレームF 2、F 3、F 4 …… F 1 3、F 1 2、F 1 1といった順次で音声素片試料データを使用し

て歌唱音信号の生成を行う。さらに、歌唱生成器1 2 2は、フレームF 4、F 5、F 6 …… F 1 3、F 1 2、F 1 1、F 1 0、F 9といった順序の音声素片試料データを使用して歌唱音信号の生成を行う。なお、フレームF 1 3からフレームF 1 2といったように逆方向に戻る順序でフレームを使用する際には両者の接続部分に雑音等が生じる虞があるため、各フレームの接続部分において振幅調整処理やクロスフェード処理等を施すようにすればよい。

10

20

30

40

50

【0070】

複数の音素からなる音素連鎖の音声素片試料データを各歌唱生成器1 2 0、1 2 1、1 2 2で用いる場合には、以上のようにすることでより正確に音素連鎖を生成することができ、また各歌唱生成器1 2 0、1 2 1、1 2 2から出力される歌唱音信号の微細な特徴等が異なるものとなる。

【0071】

以上説明したように、第2実施形態に係る合唱合成装置4 0 0では、1つの音素あるいは音素連鎖について1つの音声素片試料データしか記憶されていなくても、1つの音声素片試料データを用いて、上記第1実施形態と同様により自然な印象を与えることが可能な合唱音信号を合成することができる。すなわち、音声試料データベース1 1 0に記憶させておくデータ量を抑制しつつ、より自然な印象を与えることが可能な合唱音信号を合成することができるのである。

【0072】

C. 変形例

なお、本発明は、上述した第1および第2実施形態に限定されるものではなく、以下に例示するような種々の変形が可能である。

【0073】

（変形例1）

上述した各実施形態においては、音素あるいは音素連鎖といった単位の音声素片試料データを接続して歌唱音信号を生成するようにしているが、ビブラートといわれる歌唱表現法があり、上記各実施形態における合唱合成装置にこのビブラートによる歌唱表現を加える機能を付加するようにしてもよい。

【0074】

従来より、ビブラートによる歌唱音を電子的に発音するための歌唱音信号を生成する手法としては、上記各実施形態のように音素あるいは音素連鎖単位の音声素片試料データを接続するとともに、該接続した音声素片試料データによって表現される波形に約6 Hz程度の周波数変調を付与する方法が知られている。このような方法を実施するための構成を上記各実施形態における合唱合成装置に加えるようにしてもよいが、聴取者により自然な印象を与えることが可能なビブラート歌唱音信号の生成方法として、発声者がビブラート歌唱法で歌唱した時の音声に基づいて作成したビブラート音声試料データを用いる方法があり、この方法を実施するための構成を上記各実施形態に係る合唱合成装置に付加することが好ましい。

【0075】

以下、図9を参照しながら、発声者のビブラート歌唱音声に基づいて作成したビブラート音声試料データを用いて歌唱音信号を生成する機能を上記第1実施形態における合唱合成装置に付加した場合を例に挙げて説明する。

【0076】

同図に示すように、この合唱合成装置100'における音声試料データベース110には、上記音声試料データ群110a, 110b, 110cといった音素あるいは音素連鎖を単位とした音声素片試料データに加え、ビブラート歌唱時の歌唱音声に基づいて作成されたビブラート音声試料データが記憶されている。ここで、音声試料データベース110には、各々異なる音声に基づいて作成された3つのビブラート音声試料データBDa, BD b, BD cが記憶されている。

10

【0077】

この構成の下、合唱制御部140は、上述した第1実施形態と同様、各歌唱生成器120, 121, 122に各パートの歌詞情報およびメロディ情報と、使用する音声試料データ群を指定する指定情報とに加え、3つのビブラート音声試料データBDa, BD b, BD cのいずれを使用するかを指定する第2の指定情報を供給するようになっている。ここで、各歌唱生成器120, 121, 122に供給される第2の指定情報は、異なるビブラート音声試料データを使用するように指定する情報である。このような第2の指定情報を各歌唱生成器120, 121, 122に供給することによって、各歌唱生成器120, 121, 122はビブラート歌唱音信号を生成する際に各々異なるビブラート音声試料データを読み出し、上記実施形態と同様に接続した音声素片試料データによって表現される音声

20

【0078】

このようにビブラート歌唱音信号を生成する際に、各歌唱生成器120, 121, 122が異なる音声に基づいて作成された3つのビブラート音声試料データBDa, BD b, BD cを各パート毎に使い分けることにより、生成されるビブラート歌唱音信号の微細な特徴(ビブラート時の周波数の変動具合等)も各パート毎に異なったものとなる。このようにビブラート歌唱音の各パート毎の相関関係がほとんどなく、各々のパートが固有の特徴を持つことになるため、当該合唱合成装置100'によって合成された合唱音信号に基づいた歌唱音のビブラート部分を聴いた聴取者に対して、より自然な印象を与えることが可能となる。

30

【0079】

ところで、合唱音において各パートのビブラート部分の特徴が基本的に同一であることは、聴取者にとって他の部分の特徴が同一である場合よりも不自然な印象を与えるものである。したがって、ビブラート部分だけでも各パート毎に固有の特徴を付与した装置が要望されることもあり得る。このような場合には、上記各実施形態のように音素あるいは音素連鎖についての音声試料データは、各パートで同一のものをそのまま使用して歌唱音信号を生成し、生成した歌唱音信号に各パート毎に異なるビブラート音声試料データによって表現される波形を加算してビブラート効果を付与するようにしてもよい。

【0080】

(変形例2)

また、図9に示すように、各歌唱生成器120, 121, 122の数に対応して3つのビブラート音声試料データを用いるようにしてもよいが、図10に示す合唱合成装置400'のように、歌唱生成器120, 121, 122が同一のビブラート音声試料データを用いてビブラート部分の歌唱音信号を生成するようにしてもよい。

40

【0081】

上述した実施形態で説明したように歌唱生成器120, 121, 122は、音声試料データ群110a, 110b, 110cを使い分けることにより各々異なる固有の特徴を有する歌唱音信号を生成することができるので、このように生成した歌唱音信号に同一のビブラート音声試料データによって表現される波形を加算しても、各々の歌唱生成器120,

50

1 2 1 , 1 2 2 から出力されるビブラート部分の歌唱音信号は固有の特徴を有したものとなる。したがって、単純に1つのビブラート音声試料データを各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 が用いるようにしてもよいが、ビブラート音声試料データについても、上記第2実施形態において各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 による音声素片試料データの使用方法として説明したように、各々の歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 が同一のビブラート音声試料データの異なる時間に対応する部分から使用を開始するようにしてもよい。この場合、合唱制御部 1 4 0 がどの時間に対応する部分から使用を開始するかを指定する指定情報を各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 に供給するようにすればよい。このようにすることで、各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 がビブラート付与のために用いる実際の音声試料データは異なる特徴を有するものとなる。したがって、各々のパートのビブラート部分の歌唱音信号が固有の特徴を有するものとなり、当該合唱合成装置 4 0 0 ' によって合成された合唱音信号に基づいた歌唱音のビブラート部分を聴いた聴取者に対して、より自然な印象を与えることが可能となる。

10

【0082】**(変形例3)**

また、上記変形例においては、生成する歌唱音信号にビブラート効果を付与するために音声試料データベース 1 1 0 にビブラート音声試料データを記憶させておくようにしていたが、ビブラート以外のトレモロ、ポルタメント等の種々の歌唱法による歌唱音を電子的に放音するために、発声者によるトレモロ部分の歌唱音声や、ポルタメント部分の歌唱音声に基づいて作成した音声試料データを音声試料データベース 1 1 0 に記憶させておくようにしてもよい。この場合にも、上述した変形例におけるビブラート音声試料データと同様、各パート毎に音声試料データを用意しておいたり、同じ音声試料データであっても異なる時間に対応した部分から使用を開始したりすることにより、各パートのトレモロやポルタメント部分の歌唱音信号に固有の特徴を持たせることができる。

20

【0083】**(変形例4)**

また、上述した第1実施形態では、3つの音声試料データ群 1 1 0 a , 1 1 0 b , 1 1 0 c を音声試料データベース 1 1 0 に記憶させるようにしていたが、高音、中音、低音といったように異なるピッチの音声に基づいて、各音声試料データ群 1 1 0 a , 1 1 0 b , 1 1 0 c に含まれる音声素片試料データを作成するようにしてもよい。例えば、高音の音声に基づいて作成した音声素片試料データは音声試料データ群 1 1 0 a に含ませるようにし、中音の音声に基づいて作成した音声素片試料データは音声試料データ群 1 1 0 b に含ませるようにし、低音の音声に基づいて作成した音声素片試料データを音声試料データ群 1 1 0 c に含ませるようにしてもよい。

30

【0084】

このように各音域毎に作成された音声試料データ群 1 1 0 a , 1 1 0 b , 1 1 0 c を記憶した音声試料データベース 1 1 0 を用いる場合、合唱制御部 1 4 0 は、楽曲情報に含まれる複数のパートのうち、高音域のメロディからなるパートの歌唱音信号の生成を担当する歌唱生成器に対し、高音の音声に基づいて作成した音声試料データ群 1 1 0 a を用いるように指定する指定情報を出力する。また、中音域のメロディからなるパートの歌唱音信号の生成を担当する歌唱生成器に対し、中音の音声に基づいて作成した音声試料データ群 1 1 0 b を用いるように指定する指定情報を出力し、さらに低音域のメロディからなるパートの歌唱音信号の生成を担当する歌唱生成器に対し、低音の音声に基づいて作成した音声試料データ群 1 1 0 c を用いるように指定する指定情報を出力する。これにより各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 は、各々が担当するパートの歌唱音信号の生成により好適な音声素片試料データを用いることができ、より高品位の歌唱音信号を生成することができる。

40

【0085】

なお、上記のようにある楽曲に対応する歌唱音信号の生成時には各歌唱生成器 1 2 0 , 1 2 1 , 1 2 2 が使用する音声試料データ群 1 1 0 a , 1 1 0 b , 1 1 0 c を固定するよう

50

にしてもよいが、各パート毎にメロディ情報によって決定される各パート毎のピッチの高低が時間毎に変化することも考えられる。この場合には、ある1つの楽曲の歌唱音信号を生成する際に、各パート毎のメロディ情報によって決定される各パート毎のピッチに高低に応じて合唱制御部140が各歌唱生成器120, 121, 122に対して指定する音声試料データ群を楽曲の途中で逐次変更するような指定情報を出力するようにしてもよい。

【0086】**(変形例5)**

また、上述した変形例では、異なるピッチ毎の音声に基づいて作成した音声試料データ群110a, 110b, 110cを音声試料データベース110に記憶させるようにしていたが、歌唱時には同じ音韻を発声している間にピッチが大きく変動することもある。したがって、音声試料データベース110に、同じ音韻、例えば「a」を発声している間にピッチ(音高)を変動させて発した音声に基づいて音声素片試料データを作成し、該音声素片試料データを音声試料データベース110に記憶させるようにしてもよい。このように音声試料データベース110には、上述した各実施形態において説明した同一ピッチの音素あるいは音素連鎖だけではなく、歌唱時に起こりうる様々なピッチ変動等を考慮して音声素片試料データを作成しておくようにしてもよい。

10

【0087】**(変形例6)**

また、上述した第1実施形態では、音声試料データベース110に記憶されている音声試料データ群110a, 110b, 110cを各歌唱生成器120, 121, 122が使い分けて用い、第2実施形態では、同一の音声素片試料データを異なる時間に対応する部分から使用を開始することにより、より自然な印象を与えることが可能な合唱音信号を合成していた。このような第1および第2実施形態に係る合唱合成装置に、音声試料データベース110から読み出した音声素片試料データに示される何らかの値(すなわち音を決定付けるパラメータ)を各歌唱生成器120, 121, 122毎に変更してから供給するパラメータ変更手段を設けるようにしてもよい。このようなパラメータ変更手段を上記第1実施形態に係る合唱合成装置に付加した場合の構成を図11に示す。

20

【0088】

同図に示すように、この合唱合成装置100”は、上記第1実施形態における合唱合成装置100の構成に加え、各歌唱生成器120, 121, 122に対応して設けられるパラメータ変更部220, 221, 222を備えている。この構成の下、合唱制御部140は、上述した第1実施形態と同様、歌唱生成器120, 121, 122に各パートの歌詞情報およびメロディ情報と、どの音声試料データ群を使用するかを指定する指定情報を出力するとともに、パラメータ変更部220, 221, 222の各々に対してパラメータの変更内容を示す変更情報を出力する。ここで、合唱制御部140は、音声試料データベース110から読み出した音声素片試料データに対して各々異なる内容の変更が施されるような変更情報を各パラメータ変更部220, 221, 222に出力する。

30

【0089】

パラメータ変更部220, 221, 222は、対応する歌唱生成器120, 121, 122が必要とする音声素片試料データを指定情報に示される音声試料データ群の中から読み出し、合唱制御部140から供給される変更情報にしたがって読み出した音声素片試料データを変更する。そして、変更後の音声素片試料データが対応する歌唱生成器120, 121, 122に供給する。そして、各歌唱生成器120, 121, 122が変更後の音声素片試料データを用いて歌唱音信号を生成する。

40

【0090】

ここで、パラメータ変更部220, 221, 222が音声試料データベース110から読み出した音声素片試料データに対して行う変更処理の内容としては、音韻性を損なわない程度に音色等を変更する処理であれば種々の変更処理を適用することができる。例えば、音声試料データベース110から読み出したある音声素片試料データによって表現される音声のフォルマント構造をモデル化し、フォルマントのバンド幅を数%変更したり、バン

50

ドの中心周波数を10Hz程度シフトする等によって音色を微妙に変更する方法がある。この場合、変更するフォルマントのバンド幅の割合や、バンドの中心周波数のシフトする量を各パラメータ変更部220, 221, 222毎に異なる値とすることにより、各パラメータ変更部220, 221, 222によって読み出された音声素片試料データに示される音声の音色が微妙に異なるものとなる。

【0091】

(変形例7)

また、上述した各実施形態において、各歌唱生成器120, 121, 122によって生成された合唱音がより自然な印象を聴取者に与えるために、各パート毎に生成された歌唱音信号による歌唱音の発音タイミングをずらすようにしてもよい。この場合、合唱制御部140が各パートに対して発音タイミングをどの程度ずらすかを指定するタイミング指定情報を供給する。この際、合唱制御部140は、各歌唱生成器120, 121, 122での発音タイミングが微妙にずれるようなタイミング指定情報を各歌唱生成器120, 121, 122に供給する。例えば、歌唱生成器120に対しては、合唱制御部140から供給される歌詞情報およびメロディ情報にしたがって生成した歌唱音信号を遅延させることなく加算器130に出力させ、歌唱生成器121に対しては、10ms遅延させて歌唱音信号を加算器130に出力させ、歌唱生成器122に対しては20ms遅延させて歌唱音信号を加算器130に出力させるようにすれば、各パートの歌唱音が微妙にずれて発音され、聴取者に対してより自然な印象を与えることができる。

10

【0092】

なお、上記のようにある1つの楽曲の歌唱音信号を生成している際に、各歌唱生成器120, 121, 122の発音タイミングの相関関係を固定するようにしてもよいが、ある楽曲の途中であっても歌唱生成器120, 121, 122の発音タイミングの相関関係を変動させるようにしてもよい。例えば楽曲の前半部分では、上記例のように歌唱生成器120, 121, 122といった順序で発音するようにし、楽曲の後半部分では歌唱生成器122, 121, 120といった順序で発音するようにしてもよい。

20

【0093】

(変形例8)

また、上述した第1実施形態では、歌唱生成器120, 121, 122の数(3つ)に応じた種類の音声試料データ群を音声試料データベース110に記憶させるようにしていたが、歌唱生成器の数よりも多い種類の音声試料データ群を記憶させるようにしてもよい。

30

【0094】

また、歌唱生成器120, 121, 122といった3つの歌唱生成器を備えている場合に、音声試料データベース110に2つの音声試料データ群110a, 110bしか記憶されていない場合には、少なくとも2つの歌唱生成器が異なる音声試料データ群110a, 110bを用いて歌唱音信号を生成するようにすればよい。この場合には、歌唱生成器120が音声試料データ群110aを用い、歌唱生成器121が音声試料データ群110bを用い、歌唱生成器122が音声試料データ群110a, 110bのいずれかを歌唱生成器120, 121と異なる時間に対応する部分から使用を開始すれば、3つの歌唱生成器120, 121, 122が実際には異なる音声素片試料データを用いて歌唱音信号を生成

40

【0095】

(変形例9)

上述した各実施形態および変形例における合唱合成装置は、専用のハードウェア回路で構成するようにしてもよいが、図12に示すようなコンピュータシステムによるソフトウェアによって構成するようにしてもよい。同図に示すように、このコンピュータシステムは、装置全体を制御するCPU(Central Processing Unit)320、各種制御データやプログラム群を記憶するROM(Read Only Memory)321、ワークエリアとして使用されるRAM(Random Access Mem

50

ory) 322、楽曲情報やプログラム群を記憶するハードディスクやCD-ROM (Compact Disc Read Only Memory) ドライブ等の外部記憶装置 323、キーボードやマウス等の操作部 324、各種情報をユーザに表示する表示部 325、D/A変換器 326、アンプ 327、スピーカ 328を備えている。

【0096】

CPU 320は、ROM 321もしくはハードディスク等の外部記憶装置 323に記憶されているプログラム群にしたがって、音声試料データベース 110をRAM 322もしくは外部記憶装置 323に構築し、音声試料データベース 110を用いて上記各実施形態や変形例と同様に各パート毎の歌唱音信号合成処理を行う。そして、CPU 320は、生成した各パート毎の歌唱音信号を加算した後、加算後の合唱音信号をD/A変換器 326に出力する。D/A変換器 326では合唱音信号がアナログ信号に変換され、該合唱音のアナログ信号アンプ 327によって増幅された後、スピーカ 328から放音される。

10

【0097】

このように上記各実施形態および変形例における合唱合成装置は、コンピュータシステムによるソフトウェアによって構成することが可能であり、上記各実施形態等と同様の合唱音合成処理をコンピュータシステムに実行させるためのプログラムの形態でユーザに提供するようにしてもよい。このようなプログラムの提供方法としては、CD-ROMやフロッピーディスク等の各種記録媒体に記憶して提供する方法や、インターネット等の通信回線を介して提供する方法等がある。

【0098】

20

【発明の効果】

以上説明したように、本発明によれば、より自然な印象を聴取者に与えることが可能な合唱音を合成することができる。

【図面の簡単な説明】

【図1】本発明の第1実施形態に係る合唱合成装置の基本構成を示すブロック図である。

【図2】前記合唱合成装置の構成要素である音声試料データベースの作成手法を説明するための図である。

【図3】前記合唱合成装置の構成要素である歌唱生成器の機能構成を示すブロック図である。

【図4】本発明の第2実施形態に係る合唱合成装置の基本構成を示すブロック図である。

30

【図5】第2実施形態に係る前記合唱合成装置による歌唱音信号生成方法を説明するための図である。

【図6】第2実施形態に係る前記合唱合成装置による歌唱音信号生成方法を説明するための図である。

【図7】第2実施形態に係る前記合唱合成装置による歌唱音信号生成方法を説明するための図である。

【図8】第2実施形態に係る前記合唱合成装置による歌唱音信号生成方法を説明するための図である。

【図9】第1実施形態に係る前記合唱合成装置の変形例の基本構成を示すブロック図である。

40

【図10】第2実施形態に係る前記合唱合成装置の変形例の基本構成を示すブロック図である。

【図11】第1実施形態に係る前記合唱合成装置の他の変形例の基本構成を示すブロック図である。

【図12】前記合唱合成装置による機能をソフトウェアによって実現するためのコンピュータシステムの構成を示すブロック図である。

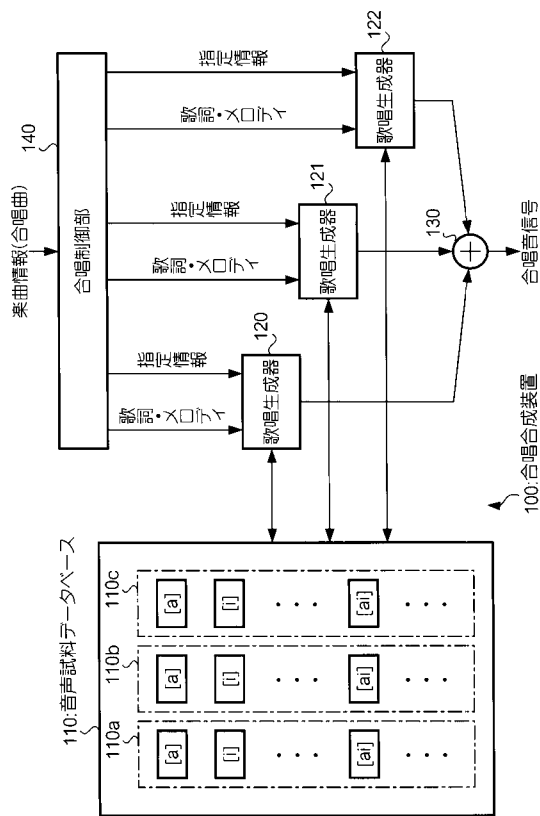
【符号の説明】

100、100'、100" ... 合唱合成装置、110 ... 音声試料データベース、110a, 110b, 110c ... 音声試料データ群、120 ... 歌唱生成器、121 ... 歌唱生成器、122 ... 歌唱生成器、130 ... 加算器、140 ... 合唱制御部、200 ...

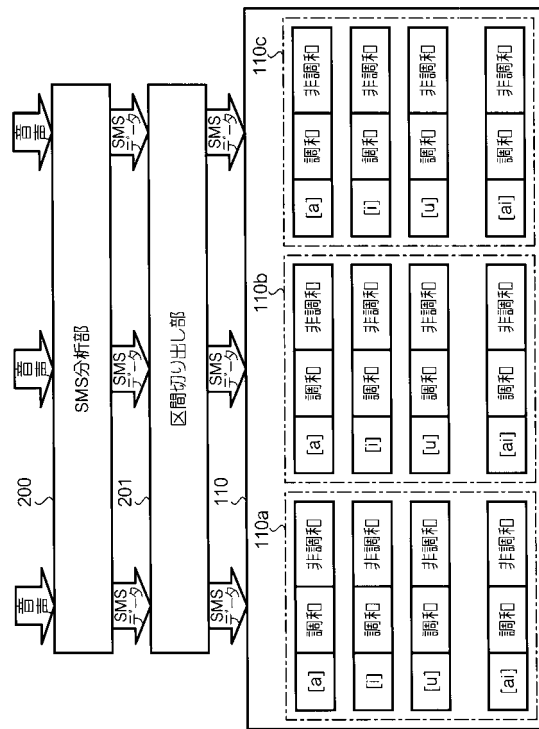
50

... SMS分析部、201... 区間切り出し部、220、221、222... パラメータ変更部、301... 音声素片選択部、302... ピッチ決定部、303... 継続時間長調整部、304... 音声素片接続部、305... 調和成分生成部、306... 加算部、307... 逆FFT部、308... 窓掛け部、309... オーバラップ部、400、400'... 合唱合成装置。

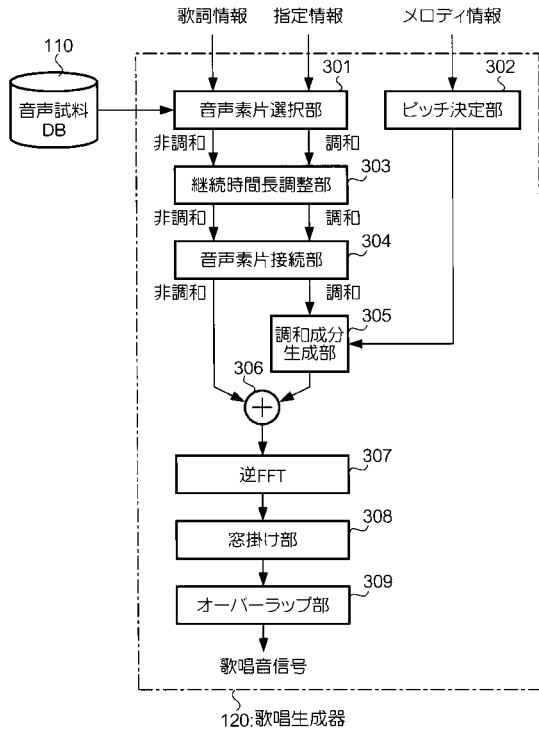
【 図 1 】



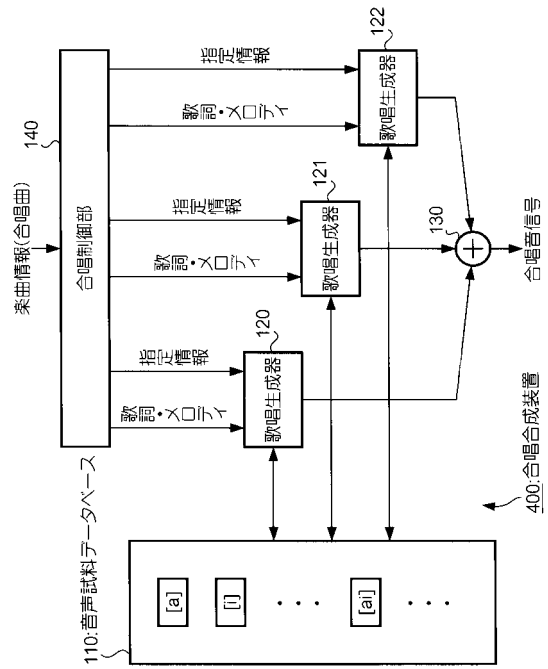
【 図 2 】



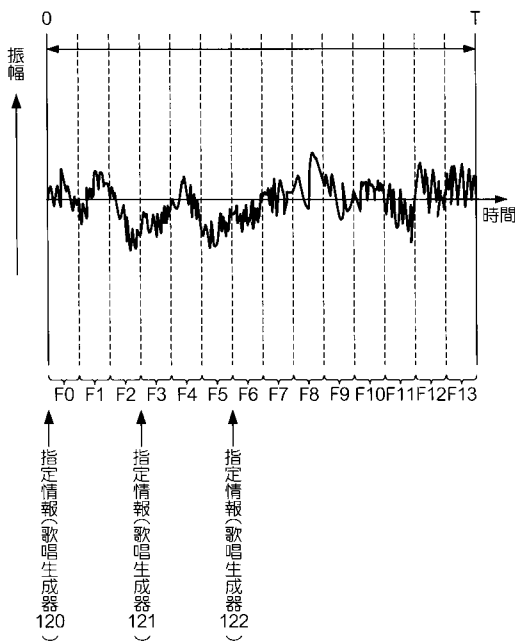
【 図 3 】



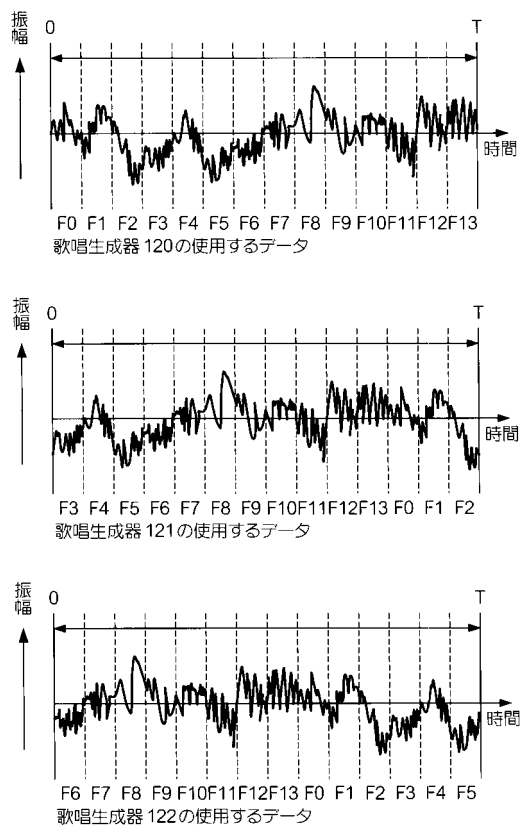
【 図 4 】



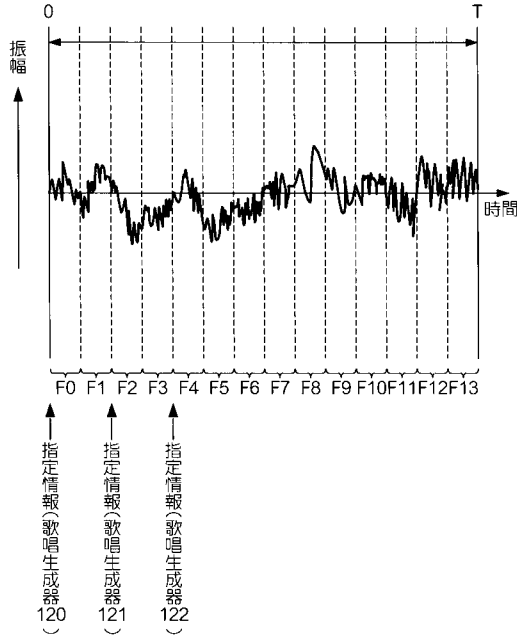
【 図 5 】



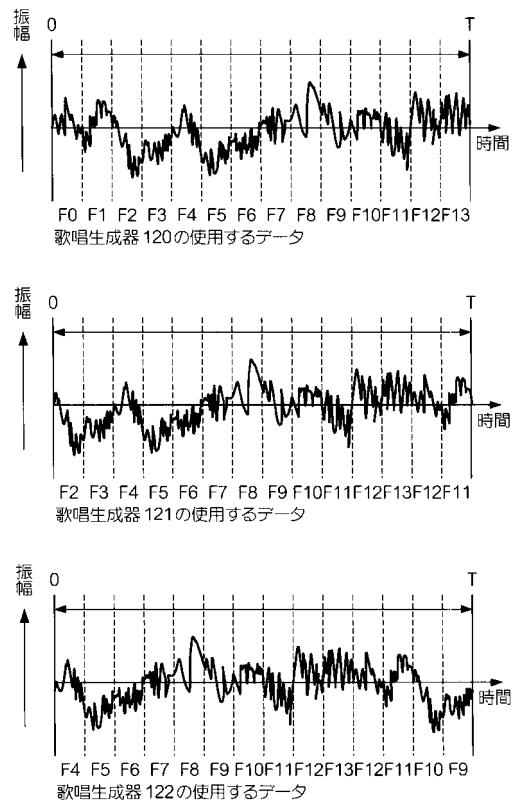
【 図 6 】



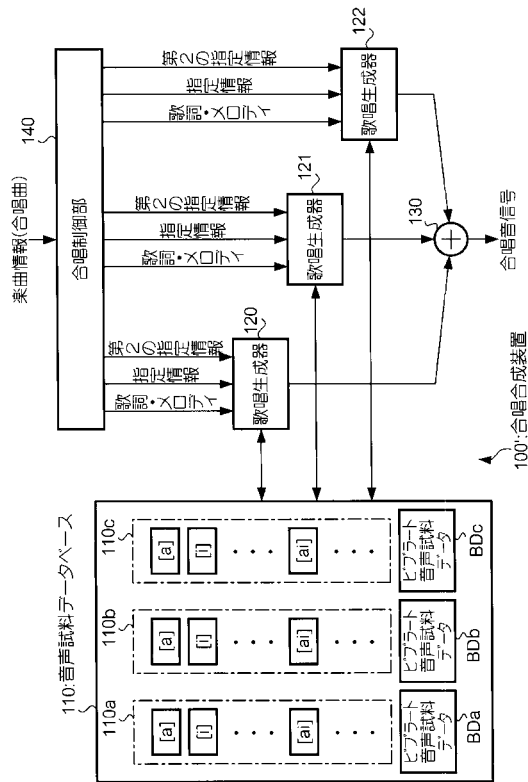
【 図 7 】



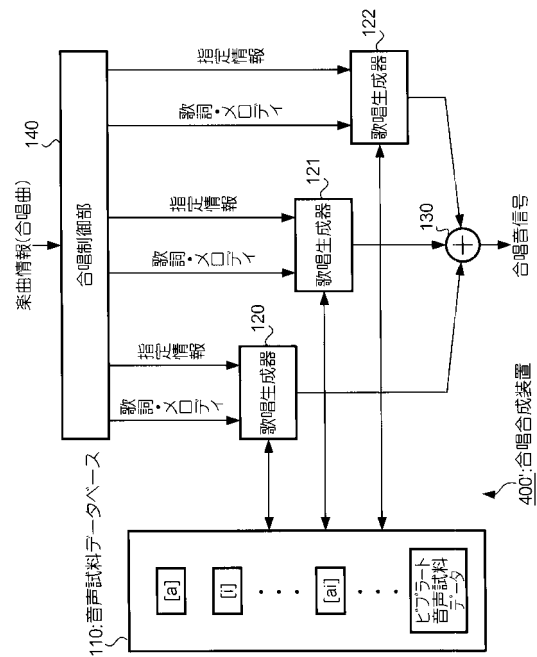
【 図 8 】



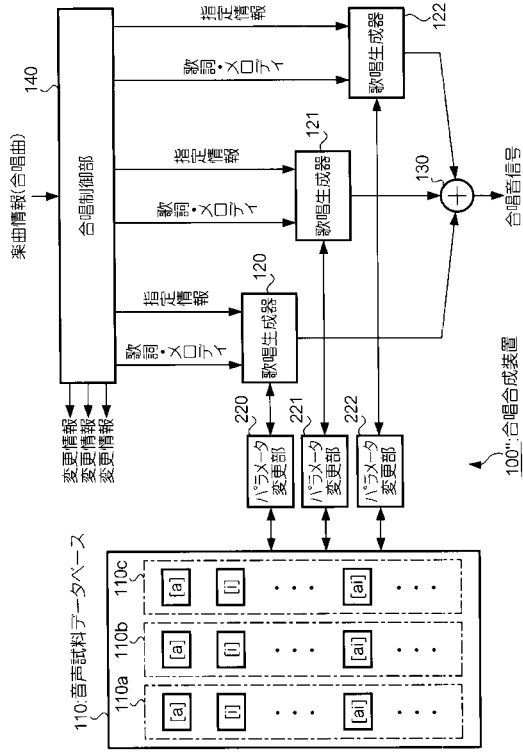
【 図 9 】



【 図 10 】



【 図 1 1 】



【 図 1 2 】

