(19) **日本国特許庁(JP)**

(12) 特 許 公 報(B2)

(11)特許番号

特許第3901475号 (P3901475)

(45) 発行日 平成19年4月4日(2007.4.4)

(24) 登録日 平成19年1月12日 (2007.1.12)

(51) Int.C1. F 1

G1OL 13/04 (2006.01) G1OL 5/02 F **G1OL** 19/00 (2006.01) G1OL 9/00 H

請求項の数 6 (全 13 頁)

(21) 出願番号 特願2001-201408 (P2001-201408) (22) 出願日 平成13年7月2日 (2001.7.2) (65) 公開番号 特開2003-15681 (P2003-15681A) 平成15年1月17日 (2003.1.17) 審查請求日 平成15年8月11日 (2003.8.11) 審判番号 不服2004-20919 (P2004-20919/J1) 審判請求日 平成16年10月7日 (2004.10.7)

(73)特許権者 000003595

株式会社ケンウッド

東京都八王子市石川町2967番地3

(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 100095407

弁理士 木村 満

(72)発明者 佐藤 寧

東京都渋谷区道玄坂1丁目14番6号 株

式会社ケンウッド内

|(72)発明者 ダビン パトリック

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

最終頁に続く

(54) 【発明の名称】信号結合装置、信号結合方法及びプログラム

(57)【特許請求の範囲】

【請求項1】

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれフーリエ変換し、 当該フーリエ変換の結果に基づいて、当該2個の入力信号のスペクトルの上限の周波数を 特定し、当該上限の周波数に基づいて、当該2個の入力信号の結合部分をフィルタリング する帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、を備え、

前記フィルタリング手段がフィルタリングする前記 2 個の入力信号の結合部分について、当該結合部分の先頭から当該 2 個の入力信号の境界までの時間長は、当該 2 個の入力信号のうち先行するものの時間長の 1 0 分の 1 以下であり、当該境界から当該結合部分の末尾までの時間長は、当該 2 個の入力信号のうち後続するものの時間長の 1 0 分の 1 以下である、

ことを特徴とする信号結合装置。

【請求項2】

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれハイパスフィルタ

によりフィルタリングした結果の振幅に基づいて、当該 2 個の入力信号の結合部分をフィルタリングする帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、を備える、

ことを特徴とする信号結合装置。

【請求項3】

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成し

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれフーリエ変換し、 当該フーリエ変換の結果に基づいて、当該2個の入力信号のスペクトルの上限の周波数を 特定し、当該上限の周波数に基づいて、当該2個の入力信号の結合部分をフィルタリング する帯域幅を決定し、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、決定した帯域幅でフィルタリングする、

<u>ことを特徴とする信号結合方法であって、</u>

フィルタリングする前記 2 個の入力信号の結合部分について、当該結合部分の先頭から 当該 2 個の入力信号の境界までの時間長は、当該 2 個の入力信号のうち先行するものの時間長の 1 0 分の 1 以下であり、当該境界から当該結合部分の末尾までの時間長は、当該 2 個の入力信号のうち後続するものの時間長の 1 0 分の 1 以下である、

ことを特徴とする信号結合方法。

【請求項4】

波形を表す複数の入力信号を入力し、各該入力信号を互いに結合して出力信号を生成し

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれハイパスフィルタによりフィルタリングした結果の振幅に基づいて、当該2個の入力信号の結合部分をフィルタリングする帯域幅を決定し、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、決定した帯域幅でフィルタリングする、

ことを特徴とする信号結合方法。

【請求項5】

コンピュータを、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれフーリエ変換し、 当該フーリエ変換の結果に基づいて、当該2個の入力信号のスペクトルの上限の周波数を 特定し、当該上限の周波数に基づいて、当該2個の入力信号の結合部分をフィルタリング する帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、

して機能させるためのプログラム<u>であって、</u>

前記フィルタリング手段がフィルタリングする前記 2 個の入力信号の結合部分について、当該結合部分の先頭から当該 2 個の入力信号の境界までの時間長は、当該 2 個の入力信号のうち先行するものの時間長の 1 0 分の 1 以下であり、当該境界から当該結合部分の末尾までの時間長は、当該 2 個の入力信号のうち後続するものの時間長の 1 0 分の 1 以下である、

ことを特徴とするプログラム。

【請求項6】

コンピュータを、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

20

40

30

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれハイパスフィルタによりフィルタリングした結果の振幅に基づいて、当該2個の入力信号の結合部分をフィルタリングする帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、

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】

この発明は、信号結合装置及び信号結合方法に関し、特に、波形を表す信号を結合するための信号結合装置及び信号結合方法に関する。

[0002]

【従来の技術】

近年、音声合成の技術により合成された音声が広く利用されている。具体的には、たとえば、テキスト読み上げソフトウェアや、電話番号案内や、株式案内、旅行案内、店舗案内、交通情報など、多くの場面で利用されている。

[0003]

音声合成の手法には、大別して、波形編集方式と、規則合成方式とがある。

規則合成方式は、音声を合成する対象のテキストについて形態素解析を行い、解析の結果に基づき、テキストに音韻論的処理を施すことにより音声を生成する手法である。規則合成方式では、音声合成に用いるテキストの内容についての制約が少なく、多様な内容のテキストを音声合成に用いることができる。しかし、規則合成方式では、波形編集方式に比べ、出力される音声の品質が劣っている。

[0004]

一方、波形編集方式は、人間が実際に発話した音声を録音して、録音した音声を分割して 得られる構成部分をつなぎ合わせることにより、目的とする音声を得る手法である。波形 編集方式は、音声の品質の点で規則合成方式より有利である。

しかし、波形編集方式では、録音された音声から取り出すことのできない部分を含む音声は合成できない。このため、録音された音声を分割する単位が大きいほど、合成する音声についての制約が多くなる。

このため、波形編集方式では、録音された音声を個々の母音や子音のレベルにまで細分化することにより、多様な音声を合成できるようにする手法も提案されている。

[0005]

【発明が解決しようとする課題】

しかし、録音した音声の構成部分をつなぎ合わせる接続部分の波形は、たとえば図 6 (a) に示すように不連続となり、これがノイズの発生源になる。そして、録音された音声を細分化する単位が小さい場合、接続部分が不連続であることにより生じるこのノイズが目立つようになり、合成音声の品質の低下を招く。

[0006]

このノイズを軽減する手法としては、たとえば、図6(b)に示すように、不連続な接続 部分を直線で接続し改善する手法が考えられる。しかし、接続された部分は高調波成分を 発生することとなり、この高調波成分もノイズとなる。

[0007]

また、接続部分が不連続であることにより生じるノイズを軽減する手法としてはMDS(Minimum Distance Search)方式がある。MDS方式は、図6(c)に示すように、2個の波形をつなぎ合わせる際、前に来る波形のうちなるべく後端に近い部分と、後ろに来る波形のうちなるべく前端に近い部分とから、瞬時値と接線の勾配とが互いにほぼ一致する点を1個ずつ探し出して、これらの点同士を接続するという手法である。

しかし、MDS方式では、波形相互の接続部分が、つなぎ合わされる各波形の端ではない ことが通常である。このため、つなぎ合わされる波形の一部は通常切り捨てられることに 30

20

50

なり、この結果、合成音声が不自然なものとなる。

[0008]

この発明は上記実状に鑑みてなされたものであり、ノイズが少なく、あるいは自然な合成 音声を生成できる信号結合装置及び信号結合方法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】

上記目的を達成すべく、この発明の第1の観点にかかる信号結合装置は、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれフーリエ変換し、 当該フーリエ変換の結果に基づいて、当該2個の入力信号のスペクトルの上限の周波数を 特定し、当該上限の周波数に基づいて、当該2個の入力信号の結合部分をフィルタリング する帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、を備え、

前記フィルタリング手段がフィルタリングする前記 2 個の入力信号の結合部分について、当該結合部分の先頭から当該 2 個の入力信号の境界までの時間長は、当該 2 個の入力信号のうち先行するものの時間長の 1 0 分の 1 以下であり、当該境界から当該結合部分の末尾までの時間長は、当該 2 個の入力信号のうち後続するものの時間長の 1 0 分の 1 以下である、

ことを特徴とする。

[0010]

このような信号結合装置により結合される2個の入力信号の結合部分は、これらの入力信号のスペクトルの上限の周波数で決まる帯域幅でフィルタリングされるので、出力信号は、高調波成分に起因するノイズが少ないものとなる。また、このような信号結合装置は、入力信号の端を切り捨てることがないので、入力信号が音声の波形を表す場合、自然な合成音声を生成する。

[0013]

また、この発明の第2の観点にかかる信号結合装置は、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれハイパスフィルタによりフィルタリングした結果の振幅に基づいて、当該2個の入力信号の結合部分をフィルタリングする帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記2個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、を備える、

ことを特徴とする。

[0014]

このような信号結合装置により結合される2個の入力信号の結合部分は、これらの入力信号の高域成分の内容に基づいて決まる帯域幅でフィルタリングされるので、出力信号から、高調波成分に起因するノイズを除去することが可能となる。また、このような信号結合装置は、入力信号の端を切り捨てることがないので、入力信号が音声の波形を表す場合、自然な合成音声を生成する。

[0015]

また、この発明の第3の観点にかかる信号結合方法は、

波形を表す複数の入力信号を入力し、各該入力信号を互いに結合して出力信号を生成し

<u>入力した</u>前記出力信号内で互いに隣接する2個の入力信号をそれぞれフーリエ変換し、 当該フーリエ変換の結果に基づいて、当該2個の入力信号のスペクトルの上限の周波数を 特定し、当該上限の周波数に基づいて、当該2個の入力信号の結合部分をフィルタリング

20

30

40

する帯域幅を決定し、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、決定した帯域幅でフィルタリングする、

ことを特徴とする。

[0016]

このような信号結合方法により結合される2個の入力信号の結合部分は、これらの入力信号のスペクトルの上限の周波数で決まる帯域幅でフィルタリングされるので、出力信号は、高調波成分に起因するノイズが少ないものとなる。また、このような信号結合方法では、入力信号の端が切り捨てられることがないので、入力信号が音声の波形を表す場合、自然な合成音声が生成される。

[0017]

また、この発明の第4の観点にかかる信号結合方法は、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成し

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれハイパスフィルタによりフィルタリングした結果の振幅に基づいて、当該2個の入力信号の結合部分をフィルタリングする帯域幅を決定し、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、決定した帯域幅でフィルタリングする、

ことを特徴とする信号結合方法であって、

フィルタリングする前記 2 個の入力信号の結合部分について、当該結合部分の先頭から 当該 2 個の入力信号の境界までの時間長は、当該 2 個の入力信号のうち先行するものの時間長の 1 0 分の 1 以下であり、当該境界から当該結合部分の末尾までの時間長は、当該 2 個の入力信号のうち後続するものの時間長の 1 0 分の 1 以下である、

ことを特徴とする。

[0018]

このような信号結合方法により結合される2個の入力信号の結合部分は、これらの入力信号の高域成分の内容に基づいて決まる帯域幅でフィルタリングされるので、出力信号から、高調波成分に起因するノイズを除去することが可能となる。また、このような信号結合方法では、入力信号の端が切り捨てられることがないので、入力信号が音声の波形を表す場合、自然な合成音声が生成される。

[0019]

また、この発明の第5の観点にかかるプログラムは、

コンピュータを、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれフーリエ変換し、 当該フーリエ変換の結果に基づいて、当該2個の入力信号のスペクトルの上限の周波数を 特定し、当該上限の周波数に基づいて、当該2個の入力信号の結合部分をフィルタリング する帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、

して機能させるための<u>プログラムであって、</u>

前記フィルタリング手段がフィルタリングする前記 2 個の入力信号の結合部分について 、当該結合部分の先頭から当該 2 個の入力信号の境界までの時間長は、当該 2 個の入力信 号のうち先行するものの時間長の 1 0 分の 1 以下であり、当該境界から当該結合部分の末 尾までの時間長は、当該 2 個の入力信号のうち後続するものの時間長の 1 0 分の 1 以下で ある、

ことを特徴とする。

[0020]

10

20

30

このようなプログラムを実行するコンピュータにより結合される2個の入力信号の結合部分は、これらの入力信号のスペクトルの上限の周波数で決まる帯域幅でフィルタリングされるので、出力信号は、高調波成分に起因するノイズが少ないものとなる。また、このようなプログラムを実行するコンピュータは、入力信号の端を切り捨てることがないので、入力信号が音声の波形を表す場合、自然な合成音声を生成する。

[0021]

また、この発明の第6の観点にかかるプログラムは、

コンピュータを、

波形を表す複数の入力信号を<u>入力し、各該入力信号を</u>互いに結合して出力信号を生成する結合手段と、

入力した前記出力信号内で互いに隣接する2個の入力信号をそれぞれハイパスフィルタによりフィルタリングした結果の振幅に基づいて、当該2個の入力信号の結合部分をフィルタリングする帯域幅を決定する帯域幅決定手段と、

前記出力信号のうち、前記 2 個の入力信号の結合部分を、前記帯域幅決定手段が決定した帯域幅でフィルタリングするフィルタリング手段と、

して機能させるためのものであることを特徴とする。

[0022]

このようなプログラムを実行するコンピュータにより結合される2個の入力信号の結合部分は、これらの入力信号の高域成分の内容に基づいて決まる帯域幅でフィルタリングされるので、このコンピュータによれば、出力信号から高調波成分に起因するノイズを除去することが可能となる。また、このようなプログラムを実行するコンピュータは、入力信号の端を切り捨てることがないので、入力信号が音声の波形を表す場合、自然な合成音声を生成する。

[0023]

【発明の実施の形態】

以下に、図面を参照して、この発明の実施の形態を、音声合成装置を例として説明する。この発明の実施の形態に係る音声合成装置10は図1に示すように、予め録音した音声を個々の母音や子音のレベルに細分化することによって得られる波形信号が入力端IN-A及びIN-Bから供給され、その供給された波形信号を合成した合成音声信号が出力端OUTから出力される基本構成から成っている。

[0024]

この音声合成装置10は、具体的な内部構成を図2に示すように、入力端IN-Aに連なった遅延部1A及びフーリエ変換部2Aと、入力端IN-Bに連なった遅延部1B及びフーリエ変換部及び2Bと、加算部3と、フィルタ特性決定部4と、LPF(ローパスフィルタ)5と、入力端IN-A及びIN-Bと、出力端OUTとから構成されている。

[0025]

遅延部1A及び1Bは、互いに実質的に同一の構成を有しており、それぞれ、たとえばシフトレジスタ等の遅延回路や、DSP(Digital Signal Processor)やCPU(Central Processing Unit)等から構成されている。遅延部1Aは入力端IN-Aに接続されており、遅延部1Bは入力端IN-Bに接続されている。

[0026]

遅延部1Aは、入力端IN-Aより信号を供給されると、この信号を一定時間遅延させて加算部3に供給する。遅延部1Bは、入力端IN-Bより信号を供給されると、この信号を一定時間遅延させて加算部3に供給する。

[0027]

なお、遅延部1A及び1Bが各自に供給された信号を遅延させる時間長は、実質的に同一である。この時間長は、フィルタ特性決定部4がLPF5へと後述の制御信号を供給するタイミングが後述する通りとなるよう選ばれている。

[0028]

フーリエ変換部2A及び2Bは、互いに実質的に同一の構成を有しており、それぞれ、D

10

20

30

40

20

30

40

50

SPやCPU等から構成されている。フーリエ変換部2Aは入力端IN-Aに接続されており、フーリエ変換部2Bは入力端IN-Bに接続されている。従って、フーリエ変換部2A及び遅延部1Aには、入力端IN-Aから同一の信号が実質的に同時に供給される。また、フーリエ変換部2B及び遅延部1Bには、入力端IN-Bから同一の信号が実質的に同時に供給される。

[0029]

フーリエ変換部 2 A は、入力端 I N - A より波形を表す信号を供給されると、高速フーリエ変換の手法(あるいは、信号をフーリエ変換した結果を表すデータを生成する他の任意の手法)により、この信号が表す波形のスペクトルを表すスペクトルデータを生成し、フィルタ特性決定部 4 へと供給する。フーリエ変換部 2 B は、入力端 I N - B より波形を表す信号を供給されると、フーリエ変換部 2 A と実質的に同一の動作を行い、この信号が表す波形のスペクトルを表すスペクトルデータを生成して、フィルタ特性決定部 4 へと供給する。

[0030]

加算部3は、加算回路等より構成されている。加算部3は、遅延部1Aより供給される信号の値と遅延部1Bより供給される信号の値の和を表す信号を生成して、LPF5へと供給する。

[0031]

フィルタ特性決定部4は、DSPやCPUより構成されている。フィルタ特性決定部4は、フーリエ変換部2A及び2Bよりそれぞれスペクトルデータを供給されると、これらのスペクトルデータに基づいて、LPF5のカットオフ周波数(具体的には、たとえばLPF5の利得が高周波側でピークより3デシベル低下する周波数)を決定し、決定したカットオフ周波数を示す制御信号を生成してLPF5に供給する。

[0032]

具体的には、フィルタ特性決定部4は、たとえば図3(a)に示すように、フーリエ変換部2Aより供給されたスペクトルデータが示すスペクトルSaの強度が高周波側でピークより20デシベル減衰する周波数を、このスペクトルSaの上限faと特定する。また、フィルタ特性決定部4は、たとえば図3(b)に示すように、フーリエ変換部2Bから供給されたスペクトルデータが示すスペクトルSbの強度が高周波側でピークより20デシベル減衰する周波数を、このスペクトルSbの上限fbとして特定する。そして、特定した2個の周波数fa及びfbのうち高い方の値を、LPF5のカットオフ周波数と決定する。なお、図3(c)は、fa<fbである場合のLPF5の周波数特性(ただし、制御信号がLPF5に供給されている間の周波数特性)を示すグラフである。

[0033]

LPF5は、たとえば、FIR (Finite Inpulse Response)型のディジタルフィルタ等より構成されている。LPF5は、加算部3より供給された信号をフィルタリングして出力する。LPF5は、フィルタリングの特性を、フィルタ特性決定部4からの制御信号の供給の有無、及び、この制御信号が示す周波数に基づいて決定する。

[0034]

具体的には、LPF5は、たとえば、フィルタ特性決定部4から制御信号が供給されている間は、加算部3より供給された信号が表す波形のうち、この制御信号が示す周波数をカットオフ周波数とするような512次のローパスフィルタを通過する成分を表す信号を生成し、生成した信号を、フィルタリングの結果を表す信号として出力端OUTより出力する。

一方、LPF5は、制御信号を供給されていない間は、加算部3より供給された信号を実質的にフィルタリングすることなくそのまま出力端OUTより出力する。

[0035]

この音声合成装置に音声の合成を行わせるには、入力端IN-A及びIN-Bに、波形信号を交互に供給する。すなわち、たとえば図4(a)及び(b)に示すように、n番目(nは任意の正の奇数)の波形信号s(n)を入力端IN-Aに供給したとすると、n番目

の波形信号が終端に達するのと実質的に同時に、入力端 I N - B に、(n + 1)番目の波 形信号 s (n + 1)の供給を開始する、というようにして、波形信号を順次供給する。

[0036]

入力端IN-Aにn番目の波形信号が供給され、入力端IN-Bに(n+1)番目の波形信号が供給されると、n番目の波形信号は遅延部1Aにより遅延を受け、また、(n+1)番目の波形信号は遅延部1Bにより遅延を受けた上で、加算部3に供給される。遅延部1A及び1Bが信号を遅延させる時間長(図4(b)及び(c)で「t0」として示す時間長)は互いに実質的に等しいので、加算部3からは、図4(c)に示すように、n番目の波形信号と(n+1)番目の波形信号とが実質的に隙間なく連続してLPF5に供給される。

[0037]

一方、 n 番目の波形信号はフーリエ変換部 2 A にも供給され、(n + 1)番目の波形信号はフーリエ変換部 2 B にも供給される。すると、フーリエ変換部 2 A は、n 番目の波形信号が表す波形のスペクトルを表すスペクトルデータを生成し、フィルタ特性決定部 4 へと供給する。また、フーリエ変換部 2 B は、(n + 1)番目の波形信号が表す波形のスペクトルを表すスペクトルデータを生成し、フィルタ特性決定部 4 へと供給する。

[0038]

フィルタ特性決定部4は、n番目及び(n+1)番目の波形信号のスペクトルを表す2個のスペクトルデータを供給されると、これらのスペクトルデータが示す各々のスペクトルの強度が高周波側で平均値より20デシベル減衰する周波数を特定する。そして、特定した2個の周波数のうち高い方の値を、LPF5のカットオフ周波数と決定し、決定したカットオフ周波数を示す制御信号をLPF5に供給する。

[0039]

n番目及び(n + 1)番目の波形信号に基づいて決定されたカットオフ周波数を示す制御信号は、図4(d)にタイミングを示すように、加算部3が出力する信号がn番目の波形信号から(n + 1)番目の波形信号へと切り替わる時点(図4(d)で「T(n)」として示す時点)を含む期間、フィルタ特性決定部4からLPF5へと供給される。(なお、理解を容易にするため、本明細書及び図面においては、LPF5自身による信号伝搬の遅延時間は無視できる程度に短いものとする。)

[0040]

なお、この音声合成装置が出力する音声信号が表す音声の劣化を防ぐために、制御信号の供給開始から波形信号が切り替わる時点までの時間長は、n番目の波形信号の時間長(図4(a)で「L(n)」として示す時間長)の10分の1以下とすることが望ましい。また、波形信号が切り替わる時点から制御信号の供給終了までの時間長は、(n+1)目の波形信号の時間長(図4(b)で「L(n+1)」として示す時間長)の10分の1以下とすることが望ましい。

[0041]

そして、LPF5は、

(A) (n-1)番目及びn番目の波形信号に基づいて決定された周波数を示す制御信号の供給が終了してから、n番目及び(n+1)番目の波形信号に基づいて決定された周波数を示す制御信号が供給されるまでの期間(図4(d)で「t1」として示す期間)は、n番目の波形信号を、実質的にフィルタリングすることなく出力端OUTから出力し、(B) n番目及び(n+1)番目の波形信号に基づいて決定された周波数を示す制御信号が供給されている期間(図4(d)で「t2」として示す期間)は、この周波数をカットオフ周波数とする512次のローパスフィルタを通過する成分を表す信号を生成して出力端OUTより出力し、

(C) n番目及び(n+1)番目の波形信号に基づいて決定された周波数を示す制御信号の供給が終了してから、(n+1)番目及び(n+2)番目の波形信号に基づいて決定された周波数を示す制御信号が供給されるまでの期間(図4(d)で「t3」として示す期間)は、(n+1)番目の波形信号を、実質的にフィルタリングすることなく出力端O

10

20

30

40

UTから出力する。

[0042]

LPF5が上述した通りにフィルタリングを行う結果、 n番目及び(n+1)番目の波形信号が、不要な高調波成分を生じることなく、また、各波形に元来含まれていた周波数成分を実質的に損なうことなく、互いに結合される。従って、結合された波形信号が表す音声は、ノイズが少なく、また、自然に聞こえるものとなる。

[0043]

なお、この音声合成装置の構成は上述のものに限られない。

たとえば、LPF5のフィルタ段数は任意であり、また、フーリエ変換部2A及び2Bが供給するスペクトルデータが示すスペクトルの上限の周波数の定義の仕方や、LPF5のカットオフ周波数の定義の仕方も、上述の定義に限らず任意である。

また、遅延部 1 A、遅延部 1 B、フーリエ変換部 2 A、フーリエ変換部 2 B、加算部 3、フィルタ特性決定部 4 及び L P F 5 の機能の全部又は一部を、単一の D S P や C P U が行ってもよい。

[0044]

また、この音声合成装置は、入力端IN-A及びIN-Bに代えて、波形信号が記録された記録媒体(たとえば、フレキシブルディスクや、MO(Magneto-Optical Disk)など)から波形信号を読み出して遅延部1A及び1Bやフーリエ変換部2A及び2Bに供給する記録媒体ドライブ装置(たとえば、フレキシブルディスクドライブや、MOドライブなど)を備えていてもよい。

また、この音声合成装置は、出力端OUTに代えて、LPF5が生成した信号を記録媒体に書き込む記録媒体ドライブ装置を備えていてもよい。

なお、同一の記録媒体ドライブ装置が、記録媒体からの波形信号を読み出す機能とLPF 5 が生成した信号を記録媒体に書き込む機能とを両方行うようにしてもよい。

[0045]

なお、入力端IN-A又はIN-Bに供給される波形信号は、無音状態を表すものであっても差し支えない。有音状態を表す波形信号と無音状態を表す波形信号とが結合されることにより、有音状態を表す信号の端を含む部分(具体的には、たとえば、音声の始まりや終わり、あるいは息継ぎ部分など)がノイズを発生することが避けられ、またこの部分が自然に聞こえるものとなる。

[0046]

また、この発明の音声合成装置においては必ずしもフーリエ変換部 2 A 及び 2 B を必要とせず、その代わりに、例えば入力端 I N - A 及び I N - B に供給される波形信号の候補を識別する識別データとこの候補のスペクトルの上限の周波数を示す周波数データとを互いに対応付けて格納記憶するテーブルを備える手法が考えられる。

この手法の場合、入力端 I N - A 及び I N - B に供給された波形信号を識別する識別データを別途外部より取得し、その取得した識別データに対応付けられた周波数データをテーブルから読み出してフィルタ特性決定部 4 に供給し、フィルタ特性決定部 4 は、その周波数データが示す周波数のうち高い方の値を、LPF5のカットオフ周波数と決定する。

[0047]

また、図 5 に示すように、この音声合成装置は、フーリエ変換部 2 A 及び 2 B に代えて H P F (ハイパスフィルタ) 6 A 及び 6 B を備えていてもよい。

HPF6A及び6Bは、互いに実質的に同一の構成を有しており、それぞれ、たとえばIIR (Infinite Inpulse Response)型のディジタルフィルタ等より構成されている。

[0048]

HPF6Aは入力端IN-Aに接続されており、HPF6Bは入力端IN-Bに接続されていて、HPF6A及び遅延部1Aには、入力端IN-Aから同一の信号が実質的に同時に供給され、HPF6B及び遅延部1Bには、入力端IN-Bから同一の信号が実質的に同時に供給される。

[0049]

50

20

10

30

20

30

40

50

HPF6Aは、入力端IN-Aより波形を表す信号を供給されると、この信号のうち所定のカットオフ周波数以下の成分を実質的に遮断し、他の成分フィルタ特性決定部4へと供給する。HPF6Bは、入力端IN-Bより供給された信号のうち所定のカットオフ周波数以下の成分を実質的に遮断し、他の成分フィルタ特性決定部4へと供給する。なお、HPF6A及び6Bのカットオフ周波数は、互いに実質的に等しいものとする。

[0050]

この音声合成装置がフーリエ変換部2A及び2Bに代えてHPF6A及び6Bを備えている場合、フィルタ特性決定部4は、HPF6A及び6Bよりそれぞれ供給された波形信号の成分に基づいて(具体的には、たとえばHPF6Aが供給した成分の振幅及びHPF6Bが供給した成分の振幅のうち大きい方の値に基づいて)、LPF5のカットオフ周波数を決定するものとする。

[0051]

この音声合成装置がフーリエ変換部 2 A 及び 2 B に代えて H P F 6 A 及び 6 B を備えていれば、比較的複雑なフーリエ変換の処理が省略されるので、この音声合成装置の処理をより高速にすることが可能になる。

[0052]

以上、この発明の実施の形態を説明したが、この発明にかかる信号結合装置は、専用のシステムによらず、通常のコンピュータシステムを用いて実現可能である。

例えば、パーソナルコンピュータに上述の遅延部1A(又はHPF6A)、遅延部1B(又は)HPF6B、フーリエ変換部2A、フーリエ変換部2B、加算部3、フィルタ特性決定部4及びLPF5の動作を実行させるためのプログラムを格納した媒体(CD-ROM、MO、フレキシブルディスク等)から該プログラムをインストールすることにより、上述の処理を実行する音声合成装置を構成することができる。

[0053]

また、例えば、通信回線の掲示板(BBS)に当該プログラムを掲示し、これを通信回線を介して配信してもよく、また、当該プログラムを表す信号により搬送波を変調し、得られた変調波を伝送し、この変調波を受信した装置が変調波を復調して当該プログラムを復元するようにしてもよい。

そして、当該プログラムを起動し、OSの制御下に、他のアプリケーションプログラムと同様に実行することにより、上述の処理を実行することができる。

[0054]

なお、 O S が処理の一部を分担する場合、あるいは、 O S が本願発明の 1 つの構成要素の一部を構成するような場合には、記録媒体には、その部分を除いたプログラムを格納してもよい。この場合も、この発明では、その記録媒体には、コンピュータが実行する各機能又はステップを実行するためのプログラムが格納されているものとする。

[0055]

【発明の効果】

以上説明したように、この発明によれば、ノイズが少なく、あるいは自然な合成音声を生成できる信号結合装置及び信号結合方法が実現される。

【図面の簡単な説明】

【図1】この発明の実施の形態に係る音声合成装置を示す図である。

【図2】この発明の実施の形態に係る音声合成装置の内部構成を示すブロック図である。

【図3】(a)は、入力端IN-Aに供給された信号のスペクトルを表すグラフであり、 (b)は、入力端IN-Bに供給された信号のスペクトルを表すグラフであり、(c)は 、LPFの周波数特性を表すグラフである。

【図4】(a)は、入力端IN-Aに供給された波形信号を表すグラフであり、(b)は、入力端IN-Bに供給された波形信号を表すグラフであり、(c)は、加算部が出力する信号を表すグラフであり、(d)は、LPFが出力する信号を表すグラフである。

【図5】図2の音声合成装置の変形例の内部構成を示すブロック図である。

【図6】(a)は、つなぎ合わされる信号が不連続になる様子を表す図であり、(b)は

10

、不連続部分を直線で接続する従来の手法を表す図であり、(c)は、MDS方式により 信号をつなぎ合わせた様子を表す図である。

【符号の説明】

1 A 、 1 B 遅延部

1 A 、 1 B遅延部2 A 、 2 Bフーリエ変換部

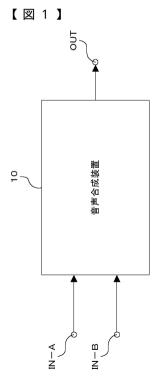
加算部

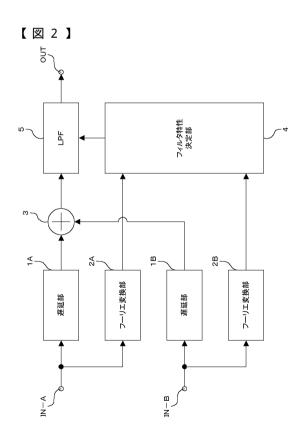
フィルタ特性決定部

5 LPF

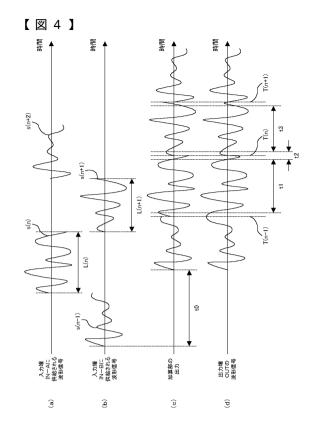
6A、6B HPF

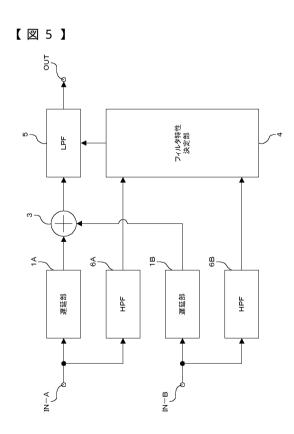
10 音声合成装置

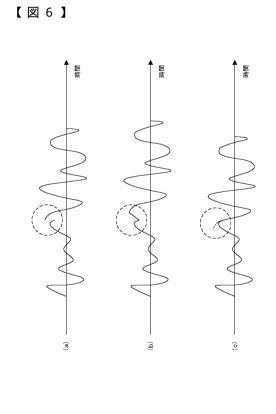




(図3)(図3)(図3)(図3)(図3)(図3)(図3)(図3)(図3)(図3)(図3)(図4)(図5)(図6)(図7)(図7)(図8)</l







フロントページの続き

合議体

審判長 原 光明

審判官 南 義明

審判官 益戸 宏

(56)参考文献 特開2000-310994(JP,A)